All Roads Lead to Rome: Global Air Connectivity and Bilateral Trade*

Zheng Wang† Feicheng Wang‡ Zhuo Zhou§

This version: August 19, 2023
First version: July 11, 2021

Abstract

Using a unique itinerary-level global air traffic database, this paper proposes a novel instrumental variable to estimate the effect of in-person interactions on bilateral trade. Our identification strategy exploits variations in third-country connecting capacities driven by external travel demand in the global flight network to leverage exogenous variations in the air connectivity between two countries. Several findings stand out. First, air connectivity boosts trade between countries, especially for industries with a high reliance on relationship-specific investments. Second, trade in new products, relative to existing products, responds more positively to improved air travel links. Third, tentative evidence suggests that air connectivity is only to some extent substituted by online social media in facilitating trade. These findings confirm the role of in-person communications in reducing transaction costs associated with tacit knowledge. Our estimation approach also provides a useful tool for evaluating the social and economic impacts of a complex transport network.

Keywords: air connectivity, face-to-face communications, connecting flights, trade

JEL classifications: F1, F2, R4

*We are grateful to the editor Treb Allen, two anonymous reviewers, Kazunobu Hayakawa, Lianming Zhu, Mo Tian, Richard Kneller, and the participants of seminars and conferences at De Montfort University, University of Nottingham, Zhejiang University, University of Göttingen, the CITP Inaugural Conference, the 23rd Göttinger Workshop on "Internationale Wirtschaftsbeziehungen", and the 2022 ETSG conference for their valuable comments. Zheng Wang acknowledges the British Academy and the Sino-British Fellowship Trust (grant no. SRG1920\100546) for financial support on this research and thanks De Montfort University for the provision of high performance computing facility for the data analysis. The usual disclaimer applies. Please address all correspondence to: Zheng Wang. Postal address: Hugh Aston Building, De Montfort University, Leicester, LE1 9BH, United Kingdom. Email: zheng.wang@dmu.ac.uk.

†De Montfort University and GEP, University of Nottingham

‡University of Groningen and IZA, Bonn

§Shanghai Academy of Social Sciences

1 Introduction

Digital communication technology has become increasingly accessible and versatile, while long-distance travel remains expensive and time-consuming. Therefore, one might expect that face-to-face meetings may have become less important in business relations. This notion, however, has been challenged by real-world data which reveals a remarkable resilience in the demand for business travel amidst this pervasive technological revolution.[1] This contrast leads to an inquiry: To what extent could in-person communications be replaced, or rather not replaced, by the virtual communication technology? In this paper, we address this question in the context of international trade: How do face-to-face interactions, measured by air connectivity between countries, facilitate exports?

While recent evidence has indicated the enduring importance of face-to-face contact in workplace and urban living environments,[2] the significance of in-person communications holds distinct relevance in the context of international trade for several reasons. First, the share of knowledge- and technology-intensive goods in global trade has been rising as a result of general technological progress (National Science Board, 2018). Second, the prevalence of globally organized production, as exemplified by the growing prominence of intermediate products (UNCTAD, 2020) and value-added creation (Wang et al., 2017; Li et al., 2019) in global trade, means a buyer-seller relationship is frequently embedded within an intricate network of international production and trade transactions.

Collectively, these forces contribute to the increasing complexity of international trade and foster a growing reliance on the exchange of intricate information between transaction parties, fuelling the demand for international business travel for face-to-face meetings (Cristea, 2011). The importance of in-person contact in sustaining international business networks is also evident in a cross-country survey which reveals that a majority of surveyed companies deem air services vital or of high importance in catering to their domestic and foreign customers (Air Transport Action Group, 2005). Nevertheless, the prevailing belief that face-to-face interactions stimulate international trade has so far relied almost entirely on anecdotal or descriptive evidence. The lack of a rigorous causal evaluation, which is crucial for informing policy decisions, leaves a

[1] According to the data from the World Travel and Tourism Council, the average year-on-year growth rate of business travel spending for the period 2000-2019 was 4.1%, similar to the world annual GDP growth rate (WTTC, 2022).

[2] It is shown that in-person interactions remain a powerful means of stimulating creativity (Xiao et al., 2021), improving productivity (Battiston et al., 2021), generating knowledge spillover (Frakes and Wasserman, 2021), and delivering technological know-how that cannot be easily codified (Storper and Venables, 2004). Büchel and v. Ehrlich (2020) finds that mobile phone usage increases significantly with the population density of a city area even after the effect of geographical distance is accounted for, which suggests that telecommunications and face-to-face interactions complement each other in daily communications.

significant gap for empirical quests.

Our paper fills this void of knowledge by using a novel identification strategy to quantify the causal impact of in-person communications on international trade on a global scale. Following the literature (Andersen and Dalgaard, 2011; Cristea, 2011; Hovhannisyan and Keller, 2015), we use international air passenger traffic to measure in-person communications across countries. Figure 1 shows evidence that countries that attract a higher number of international passengers from a specific partner country also export more to this partner. This pattern appears to be slightly stronger for exports of high-contract-intensity industries than for exports of low-contract-intensity industries, where the former is believe to rely more on in-person communications (see Section 3.2 for the economic rationale behind the contract intensity measure). However, correlation at this face value by no means implies causality because of the obvious endogeneity of passenger flows.

To isolate the causal effect from spurious correlations, we use an instrumental variable approach to leverage the exogenous variations in air connectivity between two countries. We do so by exploiting changes in the connecting capacity of transit airports located in third countries. This identification strategy depends crucially on the geographical coverage and granularity of data: it requires detailed information on air passenger flows of each flight leg of the entire global flight network.[3] This demanding task is made possible by a uniquely rich air traffic dataset we use. The data, disaggregated at the itinerary level, encompasses virtually the universe of scheduled passenger flights in the entire world, allowing us to track passenger flows and link associated flight legs and transit points (if any). Thus, it serves as an ideal source of data for studying connecting options in a global flight network.

Specifically, this estimation strategy is based on the idea that for two countries j and i, the presence of a third country m with a higher connecting capacity for passengers traveling from the rest of the world (excluding j) to i enhances the air connectivity between countries j and i. Importantly, the connecting capacity of the third country m defined here excludes connections between j and i. Thus, this capacity is primarily driven by the demand for connecting travel from the rest of the world to destination i, which is independent of the demand for travel between the pair j and i.

Conditional on a rich set of fixed effects (including exporter-industry and importer-industry specific time trends to capture demand or supply shocks), along with diligent efforts to mitigate the possible influence of confounding channels (e.g., increased investment in connecting capac-

[3] As actual seat capacity data is not available at this granular level, we use actual annual passenger traffic as a proxy for capacity. In Section 3.1, we explain why this is a reasonable approach.

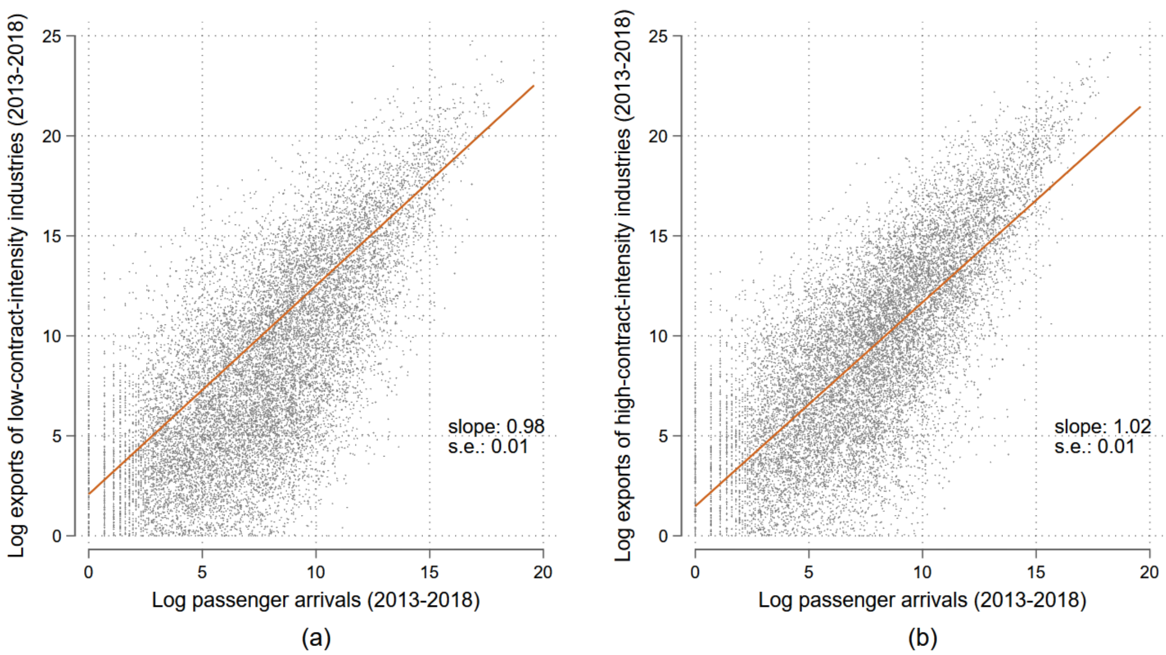

Figure 1. Passenger arrivals and exports by country-pair: 2013-2018. In panel (a) the vertical axis is the log exports of low-contract-intensity industries, and in panel (b) it is the log exports of high-contract-intensity industries. Each gray dot is an exporter-importer observation. Contract intensity measure is from Nunn (2007) which is in turn based on Rauch (1999) (liberal definition). The distinction between low and high contract intensities is based on the value relative to the sample mean. The straight lines are linear fits. Sources of data: OAG for passenger flows and CEPII-BACI for export value.

ity in response to trade growth, effect of distance between countries, and trade competition with third countries), this external connecting capacity provides a source of plausibly exogenous variation for air connectivity at the country-pair level. To illustrate the point, consider the non-stop flight connections between the United Arab Emirates (UAE) and South Africa. Our strategy is based on the fact that an expanded capacity of non-stop fight services between the UAE and South Africa, driven by India's travel demand for South Africa, facilitates the ease of booking a ticket for a Chinese visitor to travel to South Africa via the UAE. In this example, the increase in non-stop service capacity between the UAE and South Africa, resulting from the heightened travel demand from India, can be considered independent of the trade relationship between China and South Africa, conditional on baseline controls.

Our findings establish a trade-creation effect of air passenger connectivity. Notably, this effect is more pronounced in industries characterized by a higher degree of contract intensity, an index which measures the extent to which an industry relies on relationship-intensive inputs (Nunn, 2007). The amplified impact observed in these industries can be attributed to the fact that, due to greater contractual frictions, transactions within contract-intensive sectors are anticipated to rely more on face-to-face interactions (Cristea, 2011). Specifically, it is estimated that a 10% increase in passenger flows leads to a 0.31% expansion in trade for an industry with an average contract intensity. For industries exceeding the average contract intensity by one standard deviation, the trade growth is estimated to be 0.42%. Given the rapid growth of global air traffic, which more than doubled between 2013 and 2018, these effects translate into an economically sizable impact on trade volumes. In comparing the trade-promoting effect of doubling air connectivity to the impact of import tariffs, we find that it is equivalent to a 15.3% reduction in the tariff rate for industries with an average contract intensity. For industries whose contract intensity that is one standard deviation above average, achieving the same level of trade growth as doubling air connectivity would require a tariff cut of 20.8 percent.

We undertake extensive efforts to mitigate the influences of confounding channels on our estimation. For example, to address the possibility that changes in external connecting capacity could be the result of trade competition between countries, we use an import similarity index to alleviate the influence of import competitions in our external connecting capacity measure. This is supplemented by tests which rule out the correlation between current external connecting capacity and past trade growth. Our results are also robust to the exclusion of the largest economies from our sample as the variations in connecting capacity and trade flows are driven more by these countries than by others.

Our estimate could be affected by the distance between countries. To the extent that longer-distance countries rely more on connecting flights for travels between each other, this influence of distance on our estimate has been absorbed by our country-pair fixed effects. However, international trips can also be facilitated through road or railway connections, particularly for short-distance countries or smaller countries that share land borders. To address this concern, we conduct a series of robustness checks, including examining the impact of excluding country pairs that primarily rely on such alternative modes of transportation and allowing for the trade effect of air connectivity to vary across country pairs of different distances. Another concern is that our estimate could be picking up the trade impact of improved air cargo shipping. This suspicion comes from the fact that air cargo shipping often shares the same routes as passenger flights and passenger flights could sometimes carry freight. To minimize the influence of this channel, we use external data sources and two different methods to control for the probability that goods in an industry are shipped by air. Our main findings withstand these checks.

Our extended analysis uncovers additional insights regarding the trade-creation effect of air connectivity. We find this effect is enhanced by the strength of contract enforcement in the exporting country, and is larger for new products than for existing products. These results highlight the positive role of institutional environment in reaping the trade benefits of air connections and the importance of face-to-face interactions in breaking market barriers that involves significant sunk costs. Furthermore, our research reveals an intriguing relationship between the strength of online social connections and the effect of air connections on trade. We observe that as countries possess stronger online social connections, the positive effect of air connections on trade diminishes. This diminishing effect is particularly noticeable in contract-intensive industries. This finding offers tentative evidence that online communications may, to some extent, substitute in-person interactions in facilitating complex trade transactions.

This research contributes to the existing literature in several ways. First, with a novel identification strategy, our paper adds new credible evidence on the impact of transport networks or transport costs on economic outcomes. While numerous studies have explored the effects of transport network or technology development on economic growth (North, 1958; Fogel, 1964; Blonigen and Cristea, 2015; Donaldson and Hornbeck, 2016; Donaldson, 2018; Campante and Yanagizawa-Drott, 2018; Cristea, 2020), international trade (Baier and Bergstrand, 2001; Feyrer, 2009; Hummels, 2007; Bernhofen et al., 2016; Alderighi and Gaggero, 2017), foreign investment (Fageda, 2017; Campante and Yanagizawa-Drott, 2018), and firm performances (Bernard et al., 2019), a key empirical challenge remains in finding a suitable strategy that effectively isolates exogenous variations in connections at the country-pair level. In a recent study, Söderlund (2022) uses the liberalization of the Soviet airspace in 1985 as a natural experiment to identify the effects of reduced air travel time on trade between countries. This paper contributes a novel instrumental variable for the estimation of the causal effect. The global itinerary-level air traffic data that we use enables us to exploit all possible connecting routes between countries and utilize the connecting capacity of third countries as a leverage to isolate exogenous variations in air connectivity. The application of this instrument can be extended to many other settings, enabling the evaluation of the impact of transport networks on a broad array of social and economic outcomes.

Second, our finding that air passenger connectivity fosters trade and more so in contract-intensive industries aligns with the widely acknowledged notion about the importance of in-person communications in trade transactions. Face-to-face meetings are believed to help overcome informational barriers and facilitate knowledge transfer (Andersen and Dalgaard, 2011; Hovhannisyan and Keller, 2015), alleviate transaction frictions, thereby increasing the likelihood of successful deals. This mechanism suggests that the trade-promoting effect of in-person communications is more pronounced in transactions involving greater levels of relation-specific investments or complex knowledge that are challenging to specify without physical visits or meetings. Cristea (2011) finds that indeed demand for business travel is higher in industries that are more relation intensive. Our research supplements her study in that we estimate the causal impact in the opposite direction: better air connectivity leads to more trade opportunities in industries where building relationships and navigating intricate knowledge are integral to successful outcomes. This paper is also closely related to the work of Startz (2021), which uses Nigerian trade survey data to demonstrates that engaging in business travel for face-to-face interactions can effectively lower search and contracting costs in international trade, particularly for differentiated products. Rather than focusing on one specific country, our study offers comprehensive evidence using global trade and air travel data in a cross-country context.

Third, this study also enriches the literature on the role of social networks in international trade. Previous evidence suggests that countries tend to trade more with each other when they are more closely connected by ethnic and cultural links (Rauch, 1999, 2001; Rauch and Trindade, 2002), or by friendship networks formed on social media (Bailey et al., 2021). Our findings imply that economic relationships between countries can also be strengthened by international transport links, particularly industries that involve complicated contracts. By combining the online social connectedness data from Bailey et al. (2021) with ours, our study reveals an interesting insight: the trade-promoting effect of air connectivity in contract-intensive industries is diminished for countries that already have strong online connections. This implies that virtual communications can, to a certain degree, serve as a substitute for in-person meetings in facilitating international trade.

Finally, to the best of our knowledge, our constructed measure of air connectivity is the first of its kind at the country-pair level and on a global scale. In particular, this measure supplements and improves the country-specific air connectivity score published by International Air Transport Association (2020). Moreover, our measure can be easily integrated with other country-pair-level data, offering a fine-grid and valuable resource for research across various disciplines.[4]

The rest of our paper proceeds as follows. Section 2 introduces the empirical strategy, elaborating on our instrumental variable approach utilizing the flight network structure. Section 3 describes the data and reports summary statistics. Section 4 presents the baseline results, and Sections 5 conducts a number of robustness checks to address various threats to the credibility of our estimates. Section 6 delves deeper into several aspects by exploring the role of contractual environment, the extensive and intensive margins of trade, and the interactions between online connections and air links. Section 7 concludes the paper.

[4]This constructed measure will be provided as an open-access resource to the public in the format of a panel dataset deposited in the public domain.

2 Empirical Methods

Our main empirical strategy employs an instrumental variable to establish the causality of interest through the utilization of two-stage least squares (2SLS) estimations. In this section, we first motivate and discuss our second-stage estimation specification, and then elaborate our first-stage strategy which exploits the third-country connecting capacities in the global flight network to leverage the exogenous variation in air connectivity.

2.1 Estimation specification

Our baseline empirical specification is grounded in a gravity framework, which aligns with a substantial body of literature exploring the impact of information barriers on international trade (e.g., Rauch, 1999; Rauch and Trindade, 2002; Freund and Weinhold, 2004; Fink et al., 2005; Tang, 2006; Choi, 2010; Juhász and Steinwender, 2018; Bailey et al., 2021). A key hypothesis of these studies is that a reduction in information costs in trade, facilitated by reduced communication costs or improved access to social or business networks, leads to more trade between countries. This relationship is often tested using a gravity equation in which a measure of communication costs or network access is included as an explanatory variable alongside standard trade cost measures (e.g., distance, tariffs, and the use of a common language). This reduced-form empirical approach is easily justified when information costs affect only the variable costs of trade in a standard, theoretically derived gravity model (Anderson and van Wincoop, 2004). In cases where information costs affect the fixed costs of trade, it is also possible to derive an empirical equation that exhibits a gravity-like structure by aggregating firm responses to the country-pair level (Freund and Weinhold, 2004; Krautheim, 2012).

Our research builds on this body of work by focusing on the relationship between air connectivity and trade. In particular, we examine the interaction between air connectivity and the industry contract intensity to evaluate the impact of in-person communications on trade outcomes. Improved air connectivity has the potential to reduce trade frictions and foster the establishment or expansion of business networks, thereby affecting both the variable and fixed costs associated with trade. As a result, the trade response to air connectivity at the aggregate level can be formulated with a gravity framework.[5]

Specifically, we regress exports of industry k from country i to j on our air connectivity measure to estimate how the ease of traveling from a buyer's country to the seller's country has an impact on the trade volume. To account for the fact that industries which are more susceptible

[5] In Section 6.2, we investigate the impact of air connectivity on the extensive and intensive margins of trade, allowing us to differentiate between its effects on variable and fixed costs.

to information frictions and incomplete contracts rely more on face-to-face communications (Cristea, 2011; Startz, 2021), we interact an industry-specific contract intensity index with the air connectivity measure. This interaction term plays a central role in this study as it establishes a connection between the trade impact of air connectivity and an industry's demand for in-person interactions, thus shedding light on the underlying mechanisms at play. Our baseline empirical specification takes the following form:

$$\begin{aligned} Exports_{ijkt} = \beta Connectivity_{ijt} + \theta ContractIntensity_{k} \times Connectivity_{ijt} \\ + \gamma Tariff_{ijkt} + \psi RTA_{ijt} + \delta_{ikt} + \omega_{jkt} + \eta_{ij} + \epsilon_{ijkt}, \end{aligned} \tag{1}$$

where the outcome variable $Exports_{ijkt}$ is the log value of exports of industry k from country i to country j in year t. $Connectivity_{ijt}$ is the air connectivity measure constructed as the log number of air passengers from j to i in year t. This measure differs from the country-level connectivity measure of International Air Transport Association (2020) in that ours is defined at the country-pair level and thus more granular. $ContractIntensity_{k}$ measures the intensity of contract requirement for each six-digit industry k (an input-output table sector in our data), which is derived from Nunn (2007) but standardized so that it has a mean of zero and a standard deviation of one for the ease of interpretation. The interaction between industry contract intensity and air connectivity allows us to discern how the trade effect of air connectivity changes as exports become more contract intensive. The parameter β captures the effect of air connectivity on exports of industries with an average contract intensity, while θ estimates how this effect varies for each one-standard-deviation increase in contract intensity relative to the sample mean. A positive θ indicates an enhanced effect of air connectivity on trade as industries become more contract intensive.

$Tariff_{ijkt}$ is the average tariff rate imposed by country j on products within industry k originating from country i in year t.[6] RTA_{ijt} is a dummy variable which takes on the value of one when a country pair ij is part of the same regional trade agreement in year t, and zero otherwise. Taking advantage of the multidimensional nature of our data, we include a set of stringent fixed effects in our estimations. Here, δ_{ikt} and ω_{jkt} denote the exporter-industry-year and importer-industry-year fixed effects, respectively, capturing all time-varying factors at the respective country-industry level that that could influence both trade and air connectivity between countries simultaneously. Specifically, δ_{ikt} captures all supply shocks to the country-industry dyad ik that affect its export supply, including factors such as productivity improvement and

[6] Due to the large number of zero tariff rates, they are added the value of one before being converted to logarithms.

technological progress, which could also stimulate travel demand. Analogously, ω_{jkt} accounts for demand shocks to importer-industry dyads jk, encompassing factors such as income growth that elevate both the demand for foreign goods and the demand for travel. For the rest of the parameters, η_{ij} is the country-pair fixed effects, controlling for all time-invariant country-pair specific factors such as distance, adjacency, common language, common currency, and colonial links; and ϵ_{ijkt} is the error term. We cluster standard errors at the country-pair level to allow for possible correlations of estimation residuals within country pairs.

2.2 Identification strategy

2.2.1 Construction of the instrument

The identification of the effects of air connectivity on exports in Equation (1) derives from inter-temporal variations in air passengers from country j to i conditional on all baseline controls and fixed effects, while allowing for heterogeneity across industries with different levels of contract intensity. Potential threats to the identifying assumption would come from unobserved time-variant factors at the country-pair level that are correlated with both travel and trade. One possible example is bilateral economic linkages, through foreign direct investment for instance, which can contribute to increased demand for air travels (Cristea, 2011), particularly between geographically distant countries, and may also promote bilateral trade (Fuchs et al., 2020). Another source of bias arises from reverse causality, where trade induces travel demand as documented in Cristea (2011).

To address the endogeneity of our air connectivity measure, we employ an instrumental variable approach by exploiting improvements in external connecting capacity that increase connectivity between an origin country and a destination country through connecting flights routed via third countries. Intuitively, our identifying assumption is that for a given pair of origin and destination countries, an improvement in a third country's air connectivity to the destination country facilitates travel from the origin country to the same destination but is uncorrelated with trade between them.

The idea behind this instrument is illustrated in Figure 2. In this illustrated example, to travel from airport a in country j to airport b in country i, one can either travel by non-stop flight services (if there are any), or by connecting flight services routed via an airport c in a third country m. In practice, connecting services often become more relevant for travel between distant countries due to the limited availability of non-stop flight services, as shown in Figure 3. The logic behind our instrumentation strategy is that when external connecting capacity increases, there are overall more seats available for booking a trip from a to b. From a logistics point of view, with such an expansion it becomes easier to book a ticket for a planned journey as the traveler is now faced with more travel options. An external capacity expansion of the connections between airports c and b due to, for instance, an increase in flight frequencies or an upgraded fleet or plane size, could accommodate more passengers traveling from a to b and thereby raise the overall connectivity between a and b.[7] The connecting capacity, which is driven by travel demand *external* to the trading pair in question, is proxied by the total number of air passengers from the rest of the world to b through airport c, including all those who depart or transit in airport c but, importantly, excluding those from country j to i.[8]

The instrumental variable for the connectivity between a and b is represented by the thick solid lines. In our actual computation, the above air connectivity and external connecting capacity measures are first constructed at the airport-to-airport level, and then aggregated to the country-pair level so that they can be mapped to the trade data to facilitate our analysis.[9] Note that we exclude passenger flows originating from country j from our instrument. This exclusion is to address the concern that connectivity improvement between a third country m and i could be influenced by the demand for travel from country j to i, particularly when c serves as a major transit hub for travelers between these two countries.

Another concern regarding the instrumental variable arises from the possibility that the improved air connectivity between c and i may be driven by an increase in travel demand due to shocks in country i, which, in turn, affect its exports. This scenario could occur if country i experiences a positive productivity shock in industry k, leading to an influx of international visitors and a corresponding boost in i's exports of k (including exports to country j). In this case, improved air connectivity between c and i could exhibit a positive correlation with i's exports to country j through a channel other than air connectivity between j and i, thereby violating the exclusion restriction assumption. This concern is addressed by the inclusion of exporter-industry-year fixed effects (δ_{ikt}), which absorb all industry-specific shocks on the exporter's side that could potentially correlate with both air connectivity and trade between country j and i. Similarly, ω_{jkt} controls for all industry-specific shocks on the importer's side

[7] Ticket price is a consideration for travellers, but it is not the only constraint they are faced with when planning for a journey, especially if the trip is a business one which is then less sensitive to prices. Moreover, ticket prices could go up or down depending on a number of factors, including not only the demand of seats relative to the supply (capacity), but also operation cost and the pricing strategies of airlines and ticket agents. Conceptually and more importantly, our air connectivity is about the *physical* connections between countries. So, at the country-pair level, when more seats become available for booking, we consider it an improvement in air connectivity in a *physical* sense even though the average price does not necessarily get cheaper.

[8] In the data some flight routes appear to be exceptionally long and irregular. We test the sensitivity of our results by reducing the weight of connecting routes which are particularly long and thus unlikely to be regular travel options. See section 5.1 for more details.

[9] The OAG data reports flights with up to two transit points (or three legs of flights). In online Appendix A, we provide more details on how we calculate external connecting capacities for the cases of flight journeys with one or two transit points, as well as how the constructed measures are aggregated to the country-pair level.

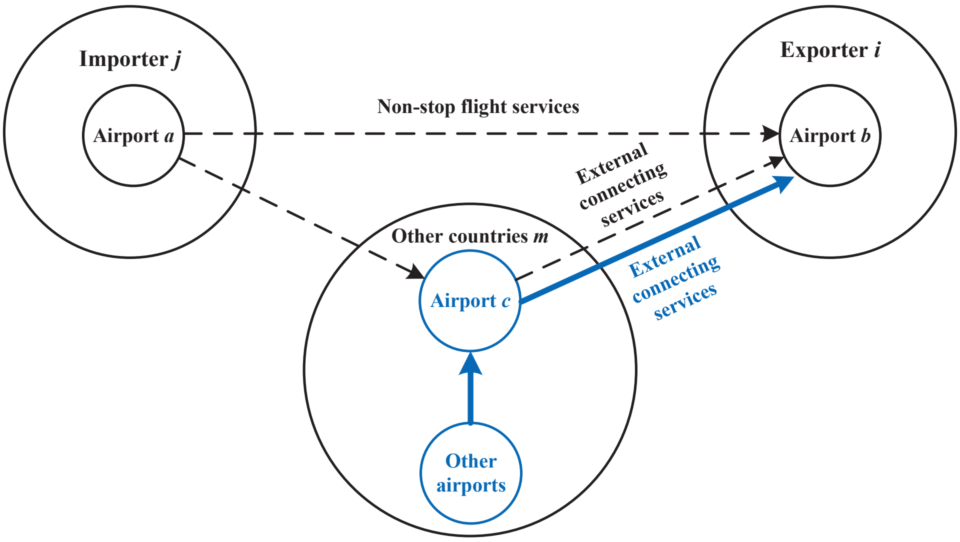

Figure 2. Illustration of the instrumentation strategy. Dashed lines represent flight services that carry passengers for at least part of their journey from airport a in origin country (importing country) j to airport b in destination country (exporting country) i. The dashed line connecting a to b represents non-stop flight services from a to b. The dashed line connecting a to c represents all (non-stop and connecting) flight services from origin airport a to airport c that is located in a third country m which is outside countries j and i. The dashed line connecting c to b represents all (non-stop and connecting) flight services from third-country airport c to destination airport b. The solid line connecting c to b also represents all (non-stop and connecting) flight services from third-country airport c to destination airport b, but their passengers originally depart from countries other than j and i. The solid line connecting "other airports" to airport c represents all (non-stop and connecting) flight services from airports outside countries j and i to airport c, and their passengers originally depart from countries other than j and i.

that could drive the correlation between passenger flows and trade. In Section 5, we perform various robustness checks to address additional concerns about the construction and validity of our instrumental variable, such as measurement errors due to irrelevant connecting routes and correlation of trade patterns between importers.

A second endogenous variable in our main estimation is the interaction term between air connectivity and industry contract intensity. As the endogenous variation of this interaction term comes exclusively from air connectivity, we interact external connecting capacity with industry contract intensity as an instrument for this endogenous variable. Following Wooldridge (2010), our first-stage estimations for each endogenous variable incorporate both instrumental variables and a full set of control variables. The first-stage estimations are formulated as follows:

$$Connectivity_{ijkt} = \rho ExternalCapacity_{ijt} + \lambda ContractIntensity_k \times ExternalCapacity_{ijt} + \gamma Tariff_{ijkt} + \psi RTA_{ijt} + \delta_{ikt} + \omega_{jkt} + \eta_{ij} + \nu_{ijkt}, \quad (2)$$

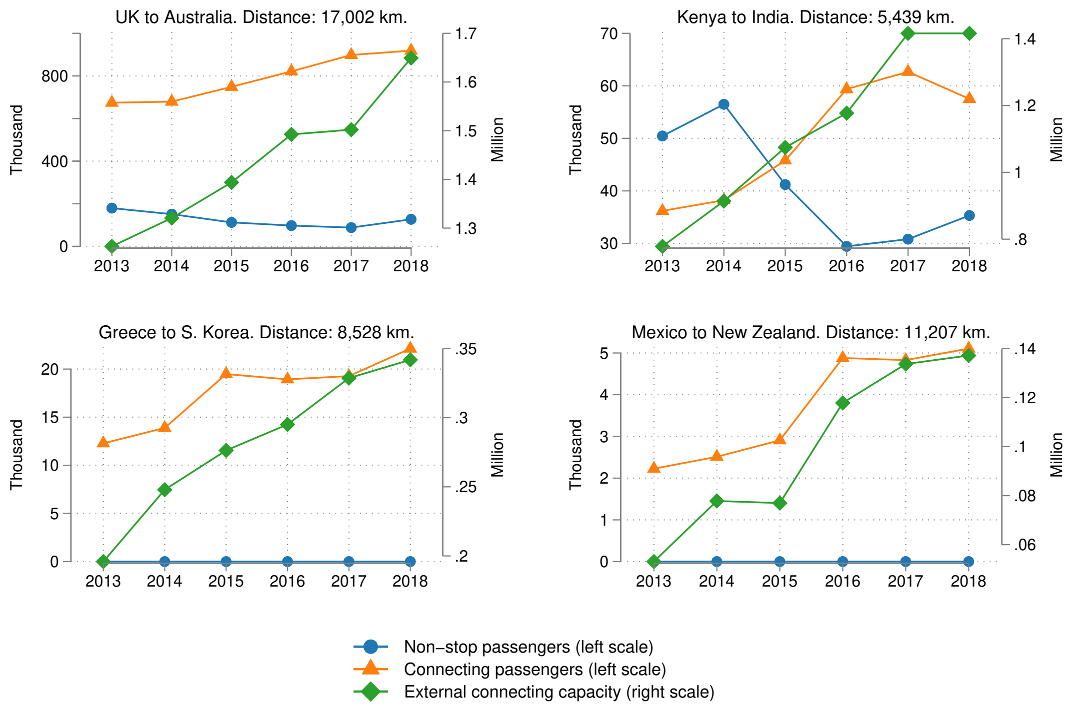

Figure 3. Air connectivity and external connecting capacity for selected country pairs. Sources of data: OAG for air connectivity and CEPII Gravity for population-weighted distance between countries.

$$\begin{aligned} ContractIntensity_k \times Connectivity_{ijt} = {} & \phi ExternalCapacity_{ijt} \\ & + \kappa ContractIntensity_k \times ExternalCapacity_{ijt} \\ & + \gamma Tariff_{ijkt} + \psi RTA_{ijt} + \delta_{ikt} + \omega_{jkt} + \eta_{ij} + \mu_{ijkt}. \end{aligned} \quad (3)$$

Our instrumentation strategy with the above specification means that the identifying variation comes from within-country-pair air travel induced by third-country connecting capacity shocks which are uncorrelated with industry-specific shocks in both the importing and exporting countries.[10] Conceptually, we ask the question: Do two countries trade more with each other during periods when the external connecting capacity for travel between the pair is increased? Furthermore, by incorporating industry heterogeneity in contract intensity, we extend our inquiry to examine whether the aforementioned effect is more pronounced in industries characterized by higher contract intensity.

[10] In the robustness check section (Section 5), we introduce different weighting schemes to address possible threats to the validity of our instrumental variable. For example, the inclusion of irregular connecting routes may introduce measurement errors. We address this concern by weighting passenger flows on each route by the reciprocal of the squared relative distance to minimize the influence of such routes. Another example is that for a given exporter i, third countries may share similar travel and import patterns to those of importer j, potentially invalidating our instrument and resulting in biased estimates. We then refine the instrument based on the similarity between j's imports and a third country's imports from i such that a third country with a more similar import pattern to j weighs less in the instrument for ji travel. Our results are robust to these refined instruments.

2.2.2 *Relevance of the instrument*

To assess the relevance of our instrument, we provide concrete examples by presenting the time trends of connecting passengers, non-stop passengers, and external connecting capacity for four specific long-distance country pairs in Figure 3: the UK to Australia, Kenya to India, Greece to South Korea, and Mexico to New Zealand. We observe that connecting flights dominate non-stop flights in all four cases. Crucially, the external connecting capacity and the number of connecting passengers exhibits a remarkably similar growth pattern, suggesting a significant role played by external connecting capacity in facilitating air connectivity between geographically distant countries.

Figure 4 depicts the general relationship between air connectivity and external connecting capacity using the sample of all country pairs. The graphical representation highlights a positive correlation between the number of air passengers (both total and connecting) and external connecting capacity, offering visual evidence that supports the relevance of our instrument.

3 Data and Summary Statistics

In this section we explain the sources of key data used in this research and how they are linked together to build the dataset for analysis. We also describe and discuss the key features of the data.

3.1 Itinerary-level air traffic data

Our air passenger traffic data comes from the Traffic Analyzer database of OAG, a leading commercial provider of aviation data and analysis. The database covers 97% of scheduled flights worldwide, including both domestic and international routes. It is updated on a monthly basis, covering the period from January 2013 to December 2018. In addition to the nearly universal coverage, another notable feature of this database is its exceptional granularity. Traffic flows are reported at the itinerary level, i.e., based on unique combinations of departure and arrival airports, including any connecting airports involved, as well as the airlines operating each flight leg. Other key information provided includes the number of passengers carried, the corresponding airline for each flight leg, the total great-circle distance traveled,[11] and the city and country names and codes of all airports associated with an itinerary. Compared to other widely used air traffic databases, such as the data from International Civil Aviation Organization (ICAO), the highly disaggregated air travel information in this OAG data allows us to trace the

[11] In cases of connecting flights, the reported great-circle distance is the sum of great-circle distances of each flight leg.

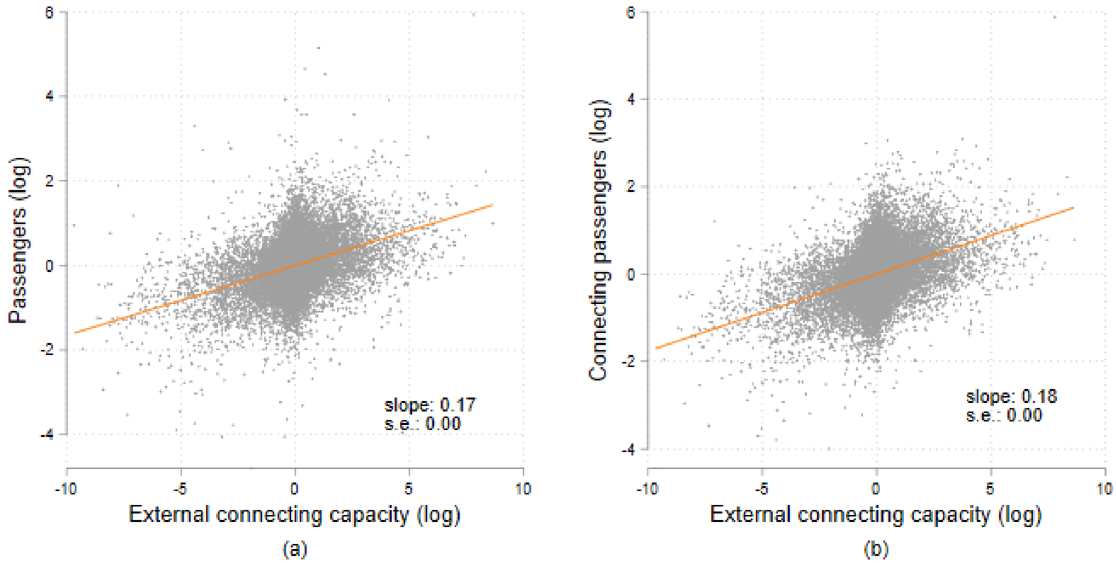

Figure 4. Relevance of the instrumental variable. Each grey dot is an exporter-importer-year observation. The straight lines are linear fits. The data is removed of exporter-year, importer-year, and exporter-importer fixed effects, i.e., each variable y_{ijt} here is taken as the estimation residual $\hat{\varepsilon}_{ijt}$ from the regression $y_{ijt} = \delta_{it} + \omega_{jt} + \eta_{ij} + \varepsilon_{ijt}$, where δ_{it}, ω_{jt}, and η_{ij} are the above fixed effects respectively. Source of data: OAG.

complete air journeys of travelers, thereby enabling the differentiation of passenger flows based on different routes and flights for any given pair of departure and destination airports. Crucial to our research, this detailed information on global flight route network enables us to establish connections between passenger flows and the connecting capacity offered by third countries acting as transit points.

As the itinerary-level seat capacity data does not exist, we use actual annual air traffic of third-country routes as a proxy for annual external connecting capacity, assuming that actual annual air traffic is positively correlated with actual annual capacity. We believe this is a reasonable approach because, according to the experience of the aviation industry, the break-even load factor (i.e., the minimum percentage of available seating capacity that is occupied by passengers to ensure profitability) for most airlines is around 70%, and airlines always try to maximize their load factor by closely monitoring the market and by actively using marketing and route optimization (including flight cancellation).[12] This experience accords well with the statistics released in the annual reports of the ICAO (see, e.g., ICAO, 2018), which show that the actual load factor of scheduled passenger flights is consistently between 70% and 85% in all regions for the period 2013-2018.

The international connections between countries are extracted from the international segments of all itineraries. This makes sure that both the first and last legs of the flight journeys we

[12]See, for example, this article: https://executiveflyers.com/what-is-load-factor-in-aviation/.

look at are served by international flights. The Schengen Area is treated as a "country" (with an artificial code "SCH") in most parts of this paper due to its no-internal-border policy and the establishment of a single market, which enable unrestricted movement of goods and people within the area akin to that of a single nation.[13]

It is worth noting that the vast majority of the international air traffic observed in our data is carried by existing flight services. During the sample period of 2013-2018, the occurrence of new international flight services was relatively rare, and the introduction of new international routes and airports was even rarer. In total, passengers who travelled by new airports, new routes (via existing airports), and new flight services (via existing routes) account for only 0.4% of the variation in the external connecting capacity.[14]

Online Appendix Table C1 contains the aggregated statistics of international air traffic from the OAG data. For the period 2013-2018, on average, over one billion passengers were transported annually across more than 26 thousand country pairs. Of these international trips, approximately 85% were carried out through non-stop flights, while the remaining 15% involved connecting flights, where passengers transited at least once before reaching their final destination airport.[15]

[13]Here an exception is that Iceland and Greece, who are both Schengen and single market members, are treated as individual countries separate from the bloc. This is because these two countries do not share any land borders with other Schengen members, which makes travel between these two countries and the rest of the Schengen Area, as well as *vice versa*, more restricted for non-EU travelers. Our results are not much affected if all Schengen members are treated as individual countries. See a robustness check in Table 6 in Section 5.3.

[14]For the sample period, we identify only two newly opened international airports that provide international connecting services, which are (with their IATA airport codes and years of opening for scheduled flights in parentheses): Blaise Diagne International Airport (DSS, 2017) in Senegal and Zhukovsky International Airport (ZIA, 2016) in Russia. The numbers of international connecting passengers these airports serve in each year after opening are: 1,306 (2017) and 29,797 (2018) for Blaise Diagne International Airport, and 6 (2016), 538 (2017) and 3,010 (2018) for Zhukovsky International Airport. For either airport, these numbers represent less than 0.3% of the total passenger flows for the country-pair connections they serve, and are a neglectable proportion of the global connecting passengers. As for new routes (via existing airports) and new flight services (via existing routes), they carry 0.75% and 4.62% of the passengers for the country-pair connections they serve, and 0.34% and 3.72% of the global connecting passengers, respectively. See online Appendix Table C2 for details. Overall, the shares of these margins are too small to be a reliable source of identification on a global scale. Another caveat about these margins is that as we move from the margin of new airports to the margin of new flights, the concern about the endogeneity of the margin also increases because of the substantial decrease in the cost of adjustment at the margin in response to trade. For example, for logistical, financial, technical, and policy reasons, the addition of a new flight service to an existing route is expected to be much less costly than building a new airport or launching a new route. Therefore, flight services, relative to airports and flight routes, could much more easily and frequently respond to observed changes in trade between countries.

[15]Our estimation strategy leverages the connecting capacities of third countries, but travel between countries with short distances is typically served by non-stop flights or ground transportation, unless a traveler deliberately chooses to transit via a third country. We make several attempts to ensure our estimates are not driven by the influence of distance. First, the country-pair fixed effects included in baseline estimations have absorbed the influence of country-pair-specific factors - including distance - that could be correlated with both air connectivity and trade. Second, in our robustness checks reported in Section 5.3, we rerun our estimations by: (1) excluding country pairs with shared land borders, (2) excluding country pairs where more than half of air passengers travel via non-stop flights, (3) excluding country pairs with a distance shorter than 6,000 miles, and (4) employing a specification which allows air connectivity to affect country pairs with different distances differently. Results from these alternative samples and specifications are only marginally different from our baseline estimates.

Table 1 reports more detailed summary statistics of the international air traffic data, summarized at the country-pair-year level. While an average country pair experiences over 43 thousand air passengers traveling between them in a year, approximately 15% of them are served by connecting flights through a third country. The data also reveals a significant discrepancy in passenger flows across countries, as indicated by the substantial standard deviations. However, these patterns should be interpreted with caution, considering that they are derived from raw data. Factors such as country size and distance can heavily influence the distribution.

Table 1. Summary statistics of the international air traffic data (2013-2018)

	N	Mean	SD	Min	Max
Number of all passengers	160,115	43,444	590,777	1	61,786,240
Number of non-stop passengers	160,115	36,963	575,238	0	61,617,648
Number of connecting passengers	160,115	6,481	46,740	0	3,376,696
External connecting capacity	159,769	764,435	2,685,379	0	85,929,021

Notes. Data is reported for country-pair-year observations. Source of data: authors' calculation based on OAG.

3.2 Contract intensity, trade, and other data

The contract intensities for industries are from Nunn (2007), where the contract intensity of an industry is measured by the value share of relationship-specific inputs (defined as those which are neither sold on an organized exchange nor reference priced) in all inputs.[16] Industries with a higher contract intensity are expected to place greater significance on relationship-specific investments. This expectation arises due to the increased utilization of differentiated inputs within these industries, resulting in more negotiations and a higher likelihood of encountering hold-up problems (Nunn, 2007). Consequently, it is also expected that trade in these industries are more likely to derive benefits from face-to-face communications. Contract intensities from Nunn (2007) are available by US Input-Output (IO) Table (1997) sector.

The global trade data used in this paper is from CEPII-BACI trade database (Gaulier and Zignago, 2010), where trade records are available by country pair and by six-digit Harmonized System (HS) product. Our tariff data is sourced from the UNCTAD Trade Analysis Information System (TRAINS) database (World Bank, 2020b), specified for each exporter-importer pair and six-digit HS product. Data on shared borders, distance between countries, and regional

[16] In Nunn's (2007) baseline contract intensity measure, relationship-specific inputs are defined as those inputs which are neither sold on an organized exchange nor reference priced. However, he shows that this measure is highly correlated with an alternative measure where reference priced products are also regarded as relationship specific. His lists of products sold on an organized exchange and reference priced are, in turn, from Rauch's (1999) liberal version of classification, but according to Nunn (2007) his results are unaffected by the use of the conservative version of Rauch's (1999) classification. In online Appendix Table C6, we show that when we turn to use Nunn's (2007) alternative measure of contract intensity (where reference priced products are added to the relationship-specific product category), the estimated trade effect of air connectivity becomes a bit smaller. This reduced effect is expected as reference priced products are less relations-specific than those sold on organized exchanges. Our results are, however, virtually unaffected by the use of Rauch's (1999) conservative classification.

trade agreements is obtained from CEPII Gravity database (Head and Mayer, 2014).[17] We use the index of enforcing contracts from the World Bank Doing Business database to measure a country's business regulatory quality (World Bank, 2020a).

3.3 Merged data

We map the import tariff data to the trade data using three-digit country ISO codes and six-digit HS product codes. These product codes are then converted to the US IO (1997) sectors using the concordance provided by the US Bureau of Economic Analysis,[18] and further merged with Nunn's (2007) contract intensities which are also specified at the same IO sector level.

Described in Table 2, the merged dataset covers nearly 18 thousand country pairs and over 300 six-digit industries (US IO sectors) in a span of six years, giving us over one million exporter-importer-industry triads and nearly three million exporter-importer-industry-year observations. The two variables that vary at the most granular (observation) level are exports and import tariff rates.

Table 2. Summary statistics of the analytical sample

	N	Mean	SD	Min	Max
Number of exporters	190				
Number of importers	152				
Number of exporter-importer pairs	17,759				
Number of industries (six-digit)	339				
Number of exporter-importer-industry triads	1,013,355				
Number of years	6				
Log export value (exporter-importer-industry-year level)	2,877,929	6.07	3.64	0.00	22.62
Log passengers (exporter-importer-year level)	57,992	6.16	3.69	0.00	17.94
Log non-stop passengers (exporter-importer-year level)	57,992	2.05	4.27	0.00	17.94
Log connecting passengers (exporter-importer-year level)	57,992	5.77	3.27	0.00	15.03
Log external connecting capacity (exporter-importer-year level)	57,992	10.53	4.71	0.00	17.95
Tariff rate (%) (exporter-importer-industry-year level)	2,877,929	4.78	15.18	0.00	3,000.00
Dummy for regional trade agreement (exporter-importer-year level)	57,992	0.17	0.38	0.00	1.00
Industry contract intensity (unstandardized, industry level)	335	0.51	0.22	0.02	0.98
Index of enforcing contracts (unstandardized, exporter-year level)	920	54.11	12.85	20.82	89.16
Index of enforcing contracts (unstandardized, importer-year level)	537	56.38	12.60	20.82	89.16

Notes. In each row, summary statistics is reported at the most disaggregated level where the variable in question varies. Source of data: merged from OAG, CEPII-BACI, and others; see text in Section 3 for exact details.

All passengers-related variables exhibit significant variations. Comparing the means and standard deviations of these variables, it becomes evident that the instrument (external connecting capacity) is much larger and more dispersed than the connectivity measure (number of passengers). Converted to unlogged terms, the average external connecting capacity is 78 $((\exp(10.53) - \exp(6.16)) / \exp(6.16))$ times that of the passenger flows for a country pair at the sample mean. A closer examination of the data shows that for 95% of the country pairs, their bilateral traffic is less than 26% of the size of their external connecting capacity. We are therefore

[17] The distance variable we extracted from CEPII Gravity is the population weighted distance between countries, which is different from the great-circle distance in the OAG data.

[18] The concordance can be found at: `https://www.bea.gov/industry/historical-benchmark-input-output-tables`

confident that, in the majority of cases, the size of the passenger flows between two countries is very small compared to the external connecting capacity these countries are exposed to.

To better understand what drives the variation in our instrument, we run regressions of the instrument on three different sets of fixed effects: (a) time fixed effects only, (b) country-pair fixed effects only, and (c) time plus country-pair fixed effects. The resulting R^2 are: 0.002, 0.795, and 0.796, respectively. The high R^2 values for regressions (b) and (c) indicate that the discrepancies in the instrument observed across country pairs are substantial and persistent, hence the predominant source of the observed variation in our instrument. In our estimations, however, these cross-country-pair differences are absorbed by the country-pair fixed effects, so the source of identification comes from within-country-pair time variation in external connecting capacity, which accounts for about 20% ($(1 - 0.795) \times 100$) of the observed variation in the instrument.

To find out which specific countries drive the variation in external connectivity capacity between countries, we plot in Figure 5 the top 300 country pairs that drive the variation in external connecting capacity in the data. It is seen that country pairs involving two large economies, including the US, China, the Schengen Area, the UK, Japan, and India, contribute most to the variations in external connecting capacity. They are followed by country pairs comprising one of these influential economic powerhouses and a smaller partner. Meanwhile, these top country pairs have some of the highest bilateral trade values globally (Panel (a)), and they are predominantly high- and medium-income countries (Panel (b)).[19] Together, these patterns show the dominance of large and affluent countries in global trade, as well as their superior air connectivity with the rest of the world.

4 Main Results

This section presents and analyzes the basic empirical findings. We begin with ordinary least squares (OLS) and Poisson pseudo-maximum likelihood (PPML) estimations, and then discuss the results from our preferred 2SLS estimations, where external connecting capacity is used as an instrument for air connectivity between countries. We also estimate the dynamic effects to rule out reverse causality and to examine the lagged impacts of air connectivity on trade.

[19] Figure C1 in online Appendix C shows more exactly how the incomes of both the origin and destination countries jointly drive the variation in external connecting capacity within the data. It is found that country pairs in which both the origin and destination countries are high-income countries contribute most to the variation.

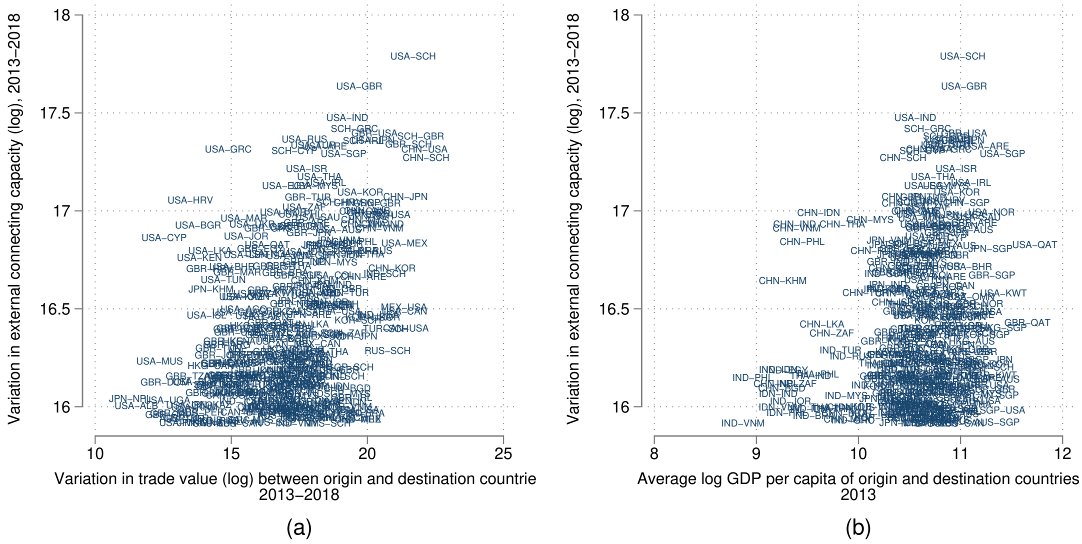

Figure 5. Top 300 country pairs which drive the variation in external connecting capacity. In both panels, the vertical axis measures the variation in external connecting capacity (in logs) that is contributed by country pairs. Contributions of country pairs to the variation in external connecting capacity are extracted from the coefficients of country pair dummies in the regressions where the dependent variable is external connecting capacity and the explanatory variables are these dummies. In Panel (a), the horizontal axis measures the variation in trade value (in logs) between countries. In Panel (b), the horizontal axis measures the average income (log GDP per capita) of the origin and destination countries in 2013. The code "SCH" denotes the Schengen Area. Sources of data: authors' calculation based on data complied from various sources (see Section 3.3).

4.1 Baseline estimates

4.1.1 OLS estimates

We start with the OLS method to assess the effect of improved air connectivity on trade. In all OLS specifications, we use the log value of exports as the dependent variable and the log number of air passengers as the measure of air connectivity. Before estimating Equation (1) at the most granular level (i.e., exporter-importer-industry-year level), we collapse the data to the exporter-importer-year level to see how passenger flows are associated with trade flows between countries. The results are presented in Table 3. Columns (1) and (2) report estimates for exports of low (defined as lower than sample mean) and high (defined as higher than sample mean) contract intensity products, respectively. In both estimations, we include tariff and an RTA dummy as control variables as well as the exporter-year, importer-year, and exporter-importer fixed effects to eliminate the influence of country-specific shocks and country-pair-specific influencing factors that may affect passenger and trade flows simultaneously. It is found that a positive correlation between passenger and trade flows only exists for high contract intensity trade, which is significant at the 10% level. For low contract intensity trade, the estimated relationship is not significantly different from zero.

To more precisely estimate how the trade effect of air connectivity varies across products

Table 3. Air connectivity and exports: OLS and PPML estimates

	Dep var: log exports					
	Exporter-importer-year				Exporter-importer-industry-year	
	Low contract intensity trade. OLS (1)	High contract intensity trade. OLS (2)	Low contract intensity trade. PPML (3)	High contract intensity trade. PPML (4)	OLS (5)	PPML (6)
Passenger flows	-0.011	0.054*	0.121**	0.119***	0.061***	0.265***
	(0.030)	(0.032)	(0.048)	(0.045)	(0.009)	(0.050)
Contract intensity×Passenger flows					0.015***	-0.009
					(0.002)	(0.014)
Tariff	-0.058*	-0.339***	-0.039**	-0.070***	-0.203***	0.113*
	(0.032)	(0.034)	(0.016)	(0.017)	(0.007)	(0.060)
RTA dummy	Yes	Yes	Yes	Yes	Yes	Yes
Exporter-year fixed effects	Yes	Yes	Yes	Yes	No	No
Importer-year fixed effects	Yes	Yes	Yes	Yes	No	No
Exporter-importer fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Exporter-product-year fixed effects	No	No	No	No	Yes	Yes
Importer-product-year fixed effects	No	No	No	No	Yes	Yes
Observations	54,194	54,194	52,834	52,788	2,839,161	2,839,161
Adjusted R^2	0.866	0.866			0.410	
Pseudo R^2			0.999	0.999		0.656

Notes. OLS and PPML estimates are reported. Dependent variable: log exports. Time period: 2013-2018. Unit of observation is exporter-importer-year in columns (1)-(4), and exporter-importer-industry-year in columns (5)-(6). "Passenger flows" is the log number of air passengers from the importing to the exporting country, "Contract intensity" is an industry-level standardised measure of reliance on relation-specific investments, "Tariff" is the log of one plus the weighted average tariff rate for an exporter-importer-year observation (columns (1)-(4)) or exporter-importer-industry-year observation (columns (5)-(6)), and "RTA dummy" is an indicator for both countries being in a same regional trade agreement. See text for sources of data. Standard errors reported in parentheses and clustered at the importer-exporter level. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels respectively.

of different levels of contract intensity, we add industry variations to the data and estimate Equation (1) with an interaction term between contract intensity and passenger flows. For ease of interpretation, the contract intensity measure is standardized to have a mean of zero and a standard deviation of one, which then indicates an industry's degree of reliance on relationship-specific investments relative to an average industry in the sample. Column (5) reports this result, where the estimation is conducted at the country-pair-industry level. Compared to columns (1) and (2), column (5) uses more stringent fixed effects, i.e., exporter-industry-year and importer-industry-year fixed effects, to filter out the influence of more disaggregated, country-industry-level shocks. The result shows that the trade effect of passenger flows is 0.061 for industries with an average contract intensity. For every one-standard-deviation increase in the contract intensity, the trade effect increases by 0.015. Both coefficients are statistically significant at the 1% level. The coefficients on tariff rates are negative and statistically significant, indicating that lower import tariff rates are associated with less bilateral trade.

4.1.2 PPML estimates

We now switch to the use of the PPML method for the above estimations. In the context of this research, a possible advantage of PPML over OLS is that it is robust to heteroscedasticity, which is a common problem in log-linearized models (Silva and Tenreyro, 2006).[20] The results for the country-pair-year data are reported in columns (3) and (4), and that for the country-pair-industry-year data in column (6). In comparison to the OLS estimates, we observe a stronger trade effect in both low- and high-contract-intensity industries, although it appears that the effect does not exhibit significant variation as the contract intensity increases.

4.1.3 2SLS estimates

The above OLS and PPML results should be interpreted with great caution for casual inferences. Even with a set of stringent fixed effects and other control variables, the above results can be biased if there are unobserved factors that are simultaneously correlated with passenger flows and trade. To address this threat from the endogeneity of passenger flows, we now resort to our preferred estimation strategy for this study, which involves the use of an instrumental variable and follows a 2SLS procedure.[21] In the main first-stage equation, Equation (2), passenger flows is instrumented by external connecting capacity. In an auxiliary first-stage equation, Equation (3), the interaction term between contract intensity and passenger flows is instrumented by contract intensity interacted with external connecting capacity.

The 2SLS estimation results are reported in Table 4 in which columns (1)-(3) correspond to the first-stage results, and columns (4) and (5) to the second-stage results. In line with Table 3, we report results at both the country-pair-year level and the country-pair-industry-year level. The first-stage result in columns (1) show that external connecting capacity is positively correlated with the number of passengers when we only explore country-pair-level variations. This is consistent with the descriptive evidence in Figure 4. The positive correlation becomes even stronger when we turn to explore country-pair-industry-year variations in column (2). In the auxiliary first-stage estimation in column (3), the main instrument of interest is now the

[20]Another advantage of using PPML instead of OLS is that it can include observations with zero-value outcomes, which would be dropped in a log-linear estimation due to the nonexistence of the natural logarithm of zero. In practice, however, adding all zeros back to the data is not always practical because it could lead to (1) a significant expansion of the data, making it computationally challenging for non-linear estimations, and (2) a biased result in a standard count data model when there is a large number of zeros (known as "zero inflation") (e.g., Mullahy, 1997; Hinde and Demétrio, 1998). This is especially the case as the data becomes more disaggregated. In our case, according to Table 2, a full rectangularization of the data would give us 190 (exporters)×152 (importers)×339 (industries)×6 (years) = 58,741,920 observations, of which over 95% (58,741,920 - 2,877,929 = 55,863,991 observations) contain zero trade flows.

[21]Instrumental-variable estimators of count data models (including PPML) with fixed effects suffer from the incidental parameters problem, which leads to inconsistent estimates (Jochmans, 2022). To the best of our knowledge, at the time of writing, there are no known practical solutions for estimating a count data model with an instrumental variable and high-dimensional fixed effects.

variable of contract intensity interacted with external connecting capacity. By construction, this variable appears to explain much of the variation in the endogenous interaction term (contract intensity $\times$ passenger flows) in a positive way. Overall, the first-stage results suggest that external connecting capacity is a strong predictor of our air connectivity measure.

Columns (4) and (5) of Table 4 presents the second-stage estimates of Equation (1). The large values of first-stage F-statistics suggest that it is unlikely that the instrument is weak. At the country-pair-year level, we find that a positive (statistically significant at the 10%) trade effect of air connectivity exists for high-contract-intensity industries (column (5)), but not for low-contract-intensity industries (column (4)). This suggests that air connectivity is more important for facilitating trade in industries with a greater reliance on relationship-specific investments. To obtain a smoother estimation of the effect across products with different contract intensities, we turn to the country-pair-industry-level data presented in column (6). The positive coefficient on passenger flows alone suggests that improved air connectivity leads to a higher value of trade in industries of average contract intensity. Moreover, the positive sign of the coefficient of the interaction term suggests that as an industry becomes more contract intensive, the trade promoting effect of air connectivity also increases. Compared to the OLS result (column (5) of Table 3), the trade effect estimated by the instrumental variable approach is larger, implying an upward bias in the OLS estimates.

Our preferred 2SLS estimates in column (6) suggest that a 10% increase in passenger flows leads to a 0.31% growth in trade for an industry with an average contract intensity. This effect is comparable to the findings of Brugnoli et al. (2018), who find that an expansion in aviation capacity in the Lombardy region of Italy led to an elasticity of trade in the region of just above zero to 0.13.[22] This effect increases by 0.11 percentage points for every one-standard-deviation increase in contract intensity. To provide a reference point, consider that the average pair of countries experienced a growth of approximately 120% in international air traffic (measured by the number of international passengers) over the six-year period from 2013 to 2018. Based on the sample mean for the given period, this improvement in air connectivity translates to a 3.72% ($0.031 \times 120 = 3.72$) increase in export value for industries with an average contract intensity. An industry with a contract intensity that is one standard deviation above average would see a 5.04% ($(0.031 + 0.011) \times 120 = 5.04$) increase in export value.

[22]Numerous studies conducted in different specific contexts consistently demonstrate a positive correlation between a decrease in transportation costs and trade. For example, in a recent study Söderlund (2022) finds that a 10% reduction in flight time, induced by the liberalization of the Soviet airspace, was associated with an over 100% increase in trade.

Table 4. Air connectivity and exports: Baseline 2SLS estimates

	1st stage			2nd stage Dep var: log exports		
	Exporter-importer level. Dep var: passenger flows.	Exporter-importer-industry level. Dep var: passenger flows.	Exporter-importer-industry level. Dep var: contract intensity ×passenger flows.	Exporter-importer-year level. Low contract intensity	Exporter-importer-year level. High contract intensity	Exporter-importer-industry-year level.
	(1)	(2)	(3)	(4)	(5)	(6)
External connecting capacity	0.165***	0.257***	-0.071***			
	(0.002)	(0.023)	(0.009)			
Contract intensity×External connecting capacity		-0.000***	0.725***			
		(0.000)	(0.002)			
Passenger flows				-0.015	0.114*	0.031**
				(0.064)	(0.065)	(0.014)
Contract intensity×Passenger flows						0.011***
						(0.000)
Tariff	-0.005	-0.004***	0.056***	-0.058***	-0.339***	-0.202***
	(0.004)	(0.001)	(0.005)	(0.022)	(0.022)	(0.002)
RTA dummy	Yes	Yes	Yes	Yes	Yes	Yes
Exporter-year fixed effects	Yes	No	No	Yes	Yes	No
Importer-year fixed effects	Yes	No	No	Yes	Yes	No
Exporter-importer fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Exporter-industry-year fixed effects	No	Yes	Yes	No	No	Yes
Importer-industry-year fixed effects	No	Yes	Yes	No	No	Yes
Observations	54,194	2,839,161	2,839,161	54,194	54,194	2,839,161
Adjusted R^2	0.985	0.992	0.966			
Kleibergen-Paap F				5,885.41	5,885.41	232,069.20

Notes. First and second-stage results of 2SLS estimations are reported. Time period: 2013-2018. Unit of observation is exporter-importer-year in columns (1), (4), and (5), and exporter-importer-industry-year in columns (2), (3) and (6). "External connecting capacity" is the capacity of connecting flight services departing from a third country, "Contract intensity" is an industry-level standardized measure of reliance on relation-specific investments, "Passenger flows" is the log number of air passengers from the importing to the exporting country, "Tariff" is the log of one plus the weighted average tariff rate for an exporter-importer-year observation (columns (1), (4) and (5)) or an exporter-importer-industry-year observation (columns (2), (3), and (6)), and "RTA dummy" is an indicator for both countries being in a same regional trade agreement. "Passenger flows" is instrumented by external connecting capacity in columns (4)-(6). See text for sources of data. Standard errors reported in parentheses and clustered at the importer-exporter level. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels respectively.

Based on the real-world data in 2018,[23] doubling air connectivity would increase global trade in industries with an average contract intensity by $24.5 \times 0.031 = 0.76$ trillion US dollars, which is equivalent to the gross output of Turkey. In industries with a contract intensity one standard deviation above average, a one-fold increase in air connectivity would promote trade by an estimated one ($38 \times 0.042 = 1.0$) trillion US dollars, which is similar to the economic size of Indonesia and larger than the economy of the Netherlands.

By using the estimated coefficient of tariffs in column (6), we can determine the *ad valorem* tariff equivalent of doubling air connectivity. For industries with an average contract intensity, this would result in a reduction of tariff rates by approximately 15.3% ($0.031/0.202 \times 100 = 15.3$). Industries with a contract intensity one standard deviation above average would experience a larger reduction of around 20.8% ($0.042/0.202 \times 100 = 20.8$). Considering that the global average tariff rate was 4.3% in 2018, these reductions translate to a substantial trade promotion effect. Specifically, the estimated impact of a one-fold growth in air connectivity is equivalent to a reduction of 0.66 ($4.3 \times 0.153 = 0.66$) percentage points in the global average import duties for industries with an average contract intensity. For industries with a contract intensity one standard deviation above average, the reduction would be 0.89 ($4.3 \times 0.208 = 0.89$) percentage points.

In summary, our 2SLS estimates show that an increase in air connectivity between two countries leads to an increase in bilateral trade. The estimated effect is nontrivial, and it is larger for trade in more contract-intensive industries.

4.2 Dynamic effects

The above instrumentation strategy does not account for the possibility of a correlation between our instrument and trade in different periods. Using the example of Figure 2, one may speculate that reverse causality could be at play, whereby the trade relationship between exporter i and importer j could be correlated with a transit country's connecting capacity in subsequent periods. Specifically, it is possible that a transit country improves its connecting capacity to exporter i after observing growing trade flows between i and j. This improvement could be realized through an investment in connecting facilities, but such an investment often takes time to materialize due to the size of its financial commitment and logistical challenges. If this is the case, then the correlation between trade and the instrument in the same period could be spurious.

[23] The value of world merchandise trade is 24.5 trillion US dollars in 2018. In the same year, the gross domestic products (GDP) of Turkey, Indonesia, and the Netherlands are 0.78, 1.04, and 0.91 trillion US dollars, respectively. The world trade data is from the World Trade Organization and GDP data is from the World Bank's World Development Indicators.

One way to alleviate this concern is to conduct a reduced-form estimation in an event-study manner, examining how external connecting capacity in subsequent periods, as well as previous periods, influences trade in the current period. However, our data only spans six years, which is too short to include all lags and leads of the treatment variable at the same time in a standard event study specification.[24] We therefore perform this test by conducting two separate estimations. In the first estimation, we replace the concurrent instrument with a set of lags (Equation (4)), and in the second, it is replaced by a set of leads (Equation (5)):

$$Exports_{ijkt} = \sum_{\tau=-3}^{0} \rho_\tau ExternalCapacity_{ij,t+\tau} + \sum_{\tau=-3}^{0} \lambda_\tau ContractIntensity_k \times ExternalCapacity_{ij,t+\tau} + Others_{ijkt}, \quad (4)$$

$$Exports_{ijkt} = \sum_{\tau=1}^{3} \rho_\tau ExternalCapacity_{ij,t+\tau} + \sum_{\tau=1}^{3} \lambda_\tau ContractIntensity_k \times ExternalCapacity_{ij,t+\tau} + Others_{ijkt}, \quad (5)$$

where τ is the number of years away from the current year t, with a negative number indicating a past year and a positive number indicating a subsequent year; *Others* includes all control variables and fixed effects as used in our baseline specifications. Here, Equation (4) estimates the lagged effects of external connecting capacity on trade, and Equation (5) allows us to conduct a falsification test, assessing whether there is reverse causality that runs from trade to external connecting capacity.

The estimated dynamic effects are plotted in Figure 6, where the solid triangles represent the estimates for industries with an average contract intensity (ρ_τ in Equations (4) and (5)) and solid circles denote the estimates for industries whose contract intensity is one-standard above average ($\rho_\tau + \lambda_\tau$ in Equations (4) and (5)). The estimates from the lags ($t-1$, $t-2$, and $t-3$) are all consistently above zero and statistically significant at the 5% level. The sizes of the effects are very close between industries with an average and high (i.e., one standard deviation above average) contract intensity. This indicates that exogenous changes in air connectivity have a lasting impact on trade that goes on for at least three years.

The estimates from the leads ($t+1$, $t+2$, and $t+3$) are statistically insignificant and largely stable, suggesting that there is no strong correlation between external connecting capacity of

[24]With the use of the high-dimensional fixed effects, our identification comes from variations within very fine cells including exporter-importer dyads. As the lags and leads of the instrument only vary over time (not across industries) within exporter-importer dyads, the inclusion of these variables at the same time would lead to their colinearity with time fixed effects within country pairs.

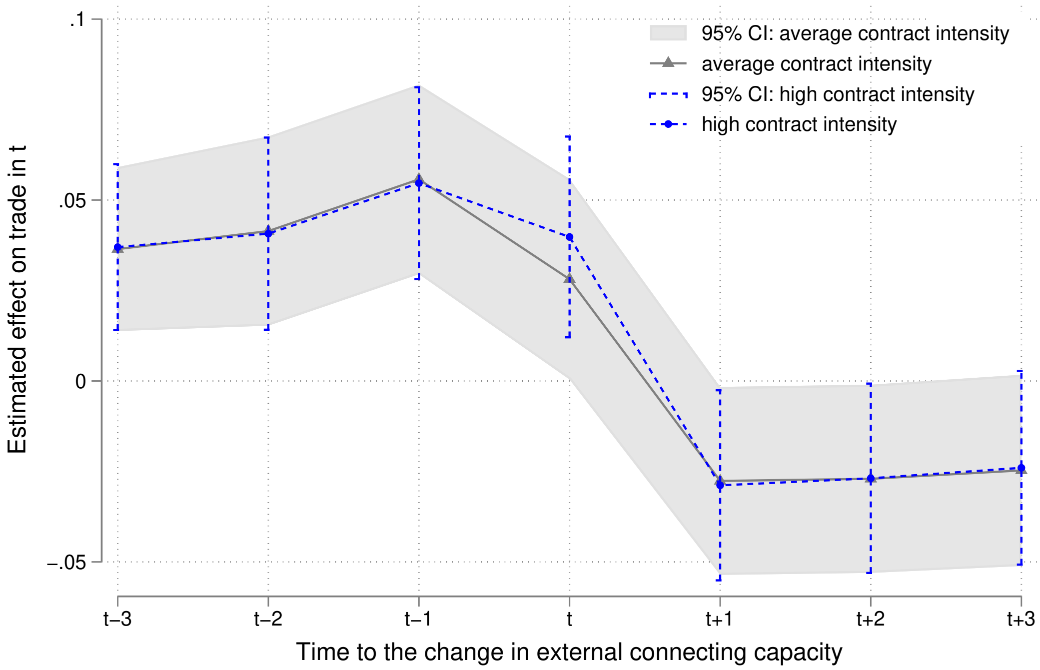

Figure 6. Dynamic effects of air connectivity. The horizontal axis measures the time relative to a change in external connecting capacity, with t being the year of change, $t-1$ one year prior to the change, $t+1$ one year post the change, etc.. The vertical axis gives the reduced-form estimates of the trade effects of external connecting capacity in different relative times, using the specifications in Equations (4) and (5). Gray triangles are estimates for industries of average contract intensity, extracted from ρ_τ in Equations (4) and (5)). Blue solid circles are for industries whose contract intensity is one standard deviation above average, extracted from $\rho_\tau + \lambda_\tau$ in Equations (4) and (5). Grey shaded areas and blue dotted lines are the 95% confidence intervals for the respective estimates. Source of data: authors' calculation based on data compiled from various sources (see Section 3.3).

the current period and trade of past periods (up to three years back). This provides reassurance that the estimated effect of external connecting capacity on trade is unlikely to be driven by reverse causality.[25]

5 Robustness Checks

In this section, we present a number of robustness checks to address specific concerns detailed below.

5.1 Irrelevant connecting routes

Our global flight data contains some irregular transit routes that are rarely used by travelers. These routes are unlikely to be relevant to our construction of external connecting capacity between countries that have access to much shorter routes. For example, the connecting route

[25] Investment in connecting capacity may be more responsive to expected changes in trade flows between larger trade partners. To check for this possibility, we conducted a robustness check in Section 5.3 where we exclude the world's largest trading economies (i.e., the US, China, and the Schengen Area) from our estimation. The results reported in Table 6 show that the size of the key estimates drops only slightly, suggesting that our main findings are not driven by the trade flows between larger trade partners.

from Beijing to London via Singapore is unlikely to affect most travelers' travel decisions, as it is much longer than other more regularly used routes.[26]

To minimize measurement error due to irrelevant routes, we weight passenger flows on each route by the reciprocal of the squared relative distance of the route. The relative distance is defined as the travel distance divided by the shortest air travel distance between the two airports. This weighting scheme ensures that longer routes are weighted less heavily than shorter routes, which helps to reduce the influence of irrelevant routes on our external connecting capacity measure.[27] With this adjustment to the instrument, the estimation result in column (1) of Table 5 shows that the key estimates remain largely unaffected. This is not surprising, considering that only a small proportion of travelers are observed using these unusual connecting routes for their destinations.

5.2 Correlation of trade pattern between importers

It is possible that visits between other parts of the world and exporter i could be correlated with trade between j and i in the same period. This correlation can be attributed to three potential channels. First, importing countries may share a similar trade pattern and the same travel routes for trade with exporter i. If there are common unobserved shocks that affect air connectivity and trade with i simultaneously, our estimated trade effect could partly reflect the trade effects on third countries. Second, enhanced connectivity via an external connecting route may facilitate trade between other importers and i while crowding out j's imports from i due to increased import competition. Third, when the products of exporter i are in demand, more imports by country j could induce a positive travel response from other importers. These importers may seek to secure their trade relationship with i through more frequent in-person interactions. The first two channels represent omitted pathways that could introduce a bias in the estimate, in an upward direction in the first channel and a downward direction in the second. The third channel poses a concern of reverse causality, where trade influences the instrument, potentially resulting in an upward bias in our key estimate. These channels become more pertinent when importers share the same transit hub(s) for the same destination i and when they source similar products from this particular destination.

[26]The non-stop flight distance between Beijing and London is 8,150 kilometers. In contrast, the travel distance between Beijing and London via Singapore is 15,400 kilometers, which is nearly twice the distance of a non-stop service and significantly longer than any connecting route with a transit point in a European or Middle Eastern country.

[27]We could exclude certain exceptionally long detour routes, but we lack an objective rule for choosing a distance threshold to define an irrelevant route. However, we tried excluding connecting routes whose total air travel distance is more than 1.5 times the shortest distance between the two airports. For example, this procedure would exclude connecting routes which go from Beijing to London via a transit point in a Southeast Asian country (e.g., Singapore or Thailand). Reassuringly, such routes account for less than 5% of the global international air traffic, and the results with this restriction criterion do not deviate much from our baseline findings.

Table 5. Robustness checks: Alternative external connecting capacity measures

	Dep var: log exports		
	Routes weighted by $(1/dist\_ratio)^2$ (1)	Competing contries weighted by (1-ISI) (2)	Routes weighted by $(1/dist\_ratio)^2$ & competing contries weighted by (1-ISI) (3)
Passenger flows	0.037**	-0.009	0.000
	(0.015)	(0.019)	(0.019)
Contract intensity×Passenger flows	0.011***	0.012***	0.012***
	(0.001)	(0.001)	(0.001)
RTA dummy & tariff	Yes	Yes	Yes
Exporter-importer fixed effects	Yes	Yes	Yes
Exporter-industry-year fixed effects	Yes	Yes	Yes
Importer-industry-year fixed effects	Yes	Yes	Yes
Observations	2,839,161	2,839,161	2,839,161
Kleibergen-Paap F	61.19	211.07	199.29

Notes. Second-stage results of 2SLS estimations are reported. Dependent variable: log exports. Time period: 2013-2018. Unit of observation is exporter-importer-industry-year. "Passenger flows" is the log number of air passengers from the importing to the exporting country, "Contract intensity" is an industry-level standardised measure of reliance on relation-specific investments, "Tariff" is the log of one plus the weighted average tariff rate, and "RTA dummy" is an indicator for both countries being in a same regional trade agreement. "Passenger flows" is instrumented by external connecting capacity. In the construction of the instrument, each route is weighted by $(1/dist\_ratio)^2$ in column (1), each country is weighted by (1-ISI) in column (2), and both weights are used in column (3). Here "dist_ratio" is the relative distance, defined as the travel distance of a route divided by the shortest travel distance between the two airports. "ISI" is the country's import similarity index with the given importer for the given exporter (Equation (6)). See text for sources of data. Standard errors reported in parentheses and clustered at the importer-exporter level. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels respectively.

To mitigate the influence of the above confounding channels, we reduce the weight of those countries who have a similar import pattern and compete more directly with importer j for products produced by exporter i in our instrument. The weight is based on a measure of the degree of competition between importer j and an external importer m for a given exporter i. This measure is inspired by the widely used export similarity index which is designed to assess the similarity of the export structure of two countries (Finger and Kreinin, 1979). In our analysis, we adapt this index to focus on imports and define the import structure based on imports from a specific exporter instead of all exporters. Specifically, our import similarity index between importers (j and m) for exporter i is computed as:

$$ISI^i_{jm} = \sum_{k \in K} \min \left(\frac{Imports_{jik}}{\sum_{k \in K} Imports_{jik}}, \frac{Imports_{mik}}{\sum_{k \in K} Imports_{mik}} \right), \tag{6}$$

where K is the set of products ever imported by j or m from i in 2013. The import similarity index (ISI) is calculated based on the overlap of products imported by j and m from i, ranging from zero to one. A value of zero indicates no common imported products, while a value of one represents an identical import structure from the given exporter. The computation is based on the initial-year (2013) data so that it is free of potential reverse causality coming from the response of import structure to competition from other countries sharing the same air travel routes for a trade partner. To refine our instrument, we incorporate a weighting scheme based on the import similarity index. Competing importers, including transit countries, are assigned a weight of $(1 - ISI)$. This weight ensures that importers who directly compete with the importer of interest (j) have a reduced influence on our estimations.[28]

The result using this refined instrument is reported in column (2) of Table 5. For industries with an average contract intensity, the estimated effect of connectivity on trade becomes statistically insignificant. However, the positive differential effect on industries with high contract intensity remains in line with our previous findings. We observe a similar outcome in column (3) where the weighting schemes for irrelevant routes and competing countries are both applied to the instrument.

5.3 Additional checks

Our main findings remain robust when subjected to several additional checks that address concerns regarding the reliability of our results. Below we provide motivations for these checks, with the results reported in Table 6.

5.3.1 Schengen members as individual countries

So far we have treated the Schengen Area as a "country" in the data, assuming its absence of restrictions on the movement of goods and people makes it equivalent to a "country". However, many studies have provided evidence suggesting the existence of substantial national border effects in trade within the European Union (most of which overlaps the single market and the Schengen Area) (Nitsch, 2000; Chen, 2004). Specifically, while certain Schengen countries such as Germany and France rank among the world's top trading nations and popular travel destinations, many others are economically smaller and less frequented by international travelers. By treating the entire Schengen Area as one data point, we might overlook valuable variations that could enhance our estimations. A robust check is to include Schengen members as individual countries (column (1)).

5.3.2 Country size

It is possible that third-country connecting capacity could respond to changes in trade between important trade partners. This is because the economically most active countries and the largest traders are also among the most connected regions. For example, the US, China, and the Schengen Area are all major economic powers and also have extensive air connectivity (see

[28] We also tried excluding competing importers with an ISI greater than 0.1 (the median and mean values of the import similarity index are 0.88 and 0.12, respectively), and our main results still hold.

Table 6. Additional checks: Borders, distance, country size, and air cargo shipping

	Dep var: log exports							
	Schengen members as individual countries	Excl. US, China, Schengen Area	Excl. country pairs with shared land borders	Country pairs where share of connecting services is consistently >50%	Excl. country pairs with distance shorter than 6,000 miles	Control. for the effect of distance	Control. for the effect of air dependence	Control. for the effect of value-to-weight ratio
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Passenger flows	0.028***	0.020	0.022	0.011	-0.003	0.030**	0.143***	0.027*
	(0.010)	(0.016)	(0.015)	(0.015)	(0.024)	(0.015)	(0.016)	(0.015)
Contract intensity × Passenger flows	0.017***	0.010***	0.013***	0.018***	0.024***	0.011***	0.011***	0.013***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
RTA dummy & tariff	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Exporter-importer fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Exporter-industry-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Importer-industry-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,775,761	2,298,143	2,623,660	1,494,491	796,176	2,821,576	2,758,498	2,775,611
Kleibergen-Paap F	479,238.82	50.14	57.86	390.24	135.35	131.77	40.69	40.64

Notes. Second-stage results of 2SLS estimations are reported. Dependent variable: log exports. Time period: 2013-2018. Unit of observation is exporter-importer-industry-year. "Passenger flows" is the log number of air passengers from the importing to the exporting country, "Contract intensity" is an industry-level standardised measure of reliance on relation-specific investments, "Exporter enforcing contracts" and "Importer enforcing contracts" are standardized measures of the strength of contract enforcement in the exporting and importing country respectively, "Tariff" is the log of one plus the weighted average tariff rates, and "RTA dummy" is an indicator for both countries being in a same regional trade agreement. "Passenger flows" is instrumented by external connecting capacity. See text for sources of data. Standard errors reported in parentheses and clustered at the importer-exporter level. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels respectively.

Section 3.3). When there is an improvement (or deterioration) in a country's trade relationship with a prominent trading partner, it may create an expectation of increased (or decreased) travel demand. Consequently, this expectation could prompt a third country to invest (or divest) in enhancing the connecting capacity between the two trading partners. To address this concern, we reestimate our baseline model by excluding observations where the US, China, or the Schengen Area is involved as a trade partner (column (2)).

5.3.3 *Distance between countries*

Distance could affect the relevance of connecting flights and air connectivity in general. We address several specific concerns here. First, we drop country pairs with shared land borders (column (3)). This is because for these countries, non-air modes of transport (mainly road and railroad transport) could reduce the relevance of air connections for cross-country travel. Second, we exclude country pairs of whom less than half of the passengers travel to the other country via a connecting route (column (4)). We do this because our instrument becomes less relevant for country pairs who rely primarily on non-stop flights for international travel. Third, we remove country pairs with distances less than 6,000 miles (approximately 9,656 kilometers) (column (5)). This distance threshold corresponds to the range of most non-stop flights, as used in Campante and Yanagizawa-Drott (2018). This gives us a sample in which most face-to-face communications are facilitated exclusively by connecting flights. Fourth, we control for the impact of distance by including in our model an interaction term between distance and the air connectivity measure (column (6)). This eliminates from our estimation the influence of air connectivity that is confounded with distance.

5.3.4 *Cargo shipping along with passenger flights*

Some large passenger aircraft, such as the Boeing 747 combi, can carry both passengers and cargo. More generally, connecting capacities for passengers and cargo can be improved at the same time on the same route. In both cases, our baseline estimate would be picking up the effect of cargo shipping capacity on trade. However, this threat only exists for goods that are likely to be shipped by air. To mitigate this confounding influence, we control for the likelihood that a product is shipped by air. This likelihood is captured by either of two product-specific measures. The first is an air dependence index that we estimate from a customs database (column (7)). This index measures the share of a product that is shipped by air. The second is the value-to-weight ratio (column (8)). We use this ratio based on the rationale that products with high value-to-weight ratios are more likely to be shipped by air due to high costs of air shipping. Online Appendix B contains further details on the construction of these two measures.

6 Further Analysis

6.1 The role of contract enforcement

So far, we have found that the trade-creation effect of air connections is stronger within industries that involve substantial relation-specific investments. We can also ask whether and how this effect varies across countries. One important country characteristic is contract enforcement. While there is consistent evidence that stronger contract enforcement itself encourages trade (e.g., Anderson and Marcouiller, 2002; Berkowitz et al., 2006; Nunn, 2007; Levchenko, 2007; Álvarez et al., 2018; Bailey et al., 2021), it is less clear how contract enforcement *interact* with in-person exchanges in affecting trade. On the one hand, in countries where formal institutions are lacking, robust in-person relationships between trading partners can serve as a substitute for reliable contract enforcement (Yu et al., 2015). On the other, stronger local institutions can enhance the effectiveness of business interactions by mitigating the risk of potential non-compliance with agreements (Berkowitz et al., 2006). Here we empirically test whether the net effect is dominated by a substitute or complementary relationship.

We measure the quality of contract enforcement by using the enforcing contracts score from the World Bank Doing Business database (World Bank, 2020a). The score evaluates the overall efficiency and quality of contract enforcement by evaluating practices of revolving commercial disputes in the local court system, with the values ranging from 0 (worst performance) to 100 (best performance). To address the issue of a change in the measurement method in 2015, we standardize the scores annually, ensuring they fall within the range of zero to one. A higher standardized score indicates comparatively easier enforcement of contracts relative to other countries in the same year. Recognizing that contract enforcement in both importing and exporting countries is relevant to trade outcomes (Berkowitz et al., 2006), we introduce two interaction terms in our analysis. They involve the interaction of our air connectivity measure with the contract enforcement scores of both importing and exporting countries separately. The two interaction terms are instrumented by the interaction of external connecting capacity with the respective contract enforcement. By incorporating these interaction terms, we capture how the trade effect of air connectivity changes when a country exhibits a higher contract enforcement score. Note that the stand-alone effects of contract enforcement are absorbed by the fixed effects.

The results are reported in Table 7. Column (1) shows the effects of contract enforcement on trade when the baseline interaction term between contract intensity and passenger flows is excluded. Column (2) adds back the interaction term to see how our baseline effect is affected by the inclusion of the influencing channels of contract enforcement. The positive coefficient of the interaction term between air connectivity and the exporter's contract enforcement score in both specifications suggests a predominantly complementary relationship between the two: air connectivity creates even more trade when the exporting country has stronger enforcement of contracts. The importing country's contract enforcement, however, turns out to be statistically and economically insignificant.

One possible explanation for this observed asymmetry could be attributed to the uneven distribution of risks inherent in international trade transactions. Sellers, relative to buyers, have more responsibilities for the production process, product quality, and the transfer of goods to the international carrier. Therefore, a good reputation and effectiveness of contract enforcement in the exporting country, relative to that in the importing country, could make in-person communications more effective in facilitating trade. The effect of contract intensity remains unaffected by the inclusion of the controls for contract enforcement.

Table 7. The role of contract enforcement

	Dep var: log exports	
	(1)	(2)
Passenger flows	0.020	0.029*
	(0.016)	(0.016)
Contract intensity × Passenger flows		0.011***
		(0.001)
Exporter enforcing contracts × Passenger flows	0.005***	0.006***
	(0.001)	(0.001)
Importer enforcing contracts × Passenger flows	0.000	0.001
	(0.001)	(0.001)
RTA dummy & tariff	Yes	Yes
Exporter-importer fixed effects	Yes	Yes
Exporter-industry-year fixed effects	Yes	Yes
Importer-industry-year fixed effects	Yes	Yes
Observations	2,649,376	2,616,157
Kleibergen-Paap F	58.91	43.66

Notes. Second-stage results of 2SLS estimations are reported. Dependent variable: log exports. Time period: 2013-2018. Unit of observation is exporter-importer-industry-year. "Passenger flows" is the log number of air passengers from the importing to the exporting country, "Contract intensity" is an industry-level standardised measure of reliance on relation-specific investments, "Exporter enforcing contracts" and "Importer enforcing contracts" are standardized measures of the strength of contract enforcement in the exporting and importing country respectively, "Tariff" is the log of one plus the weighted average tariff rates, and "RTA dummy" is an indicator for both countries being in a same regional trade agreement. "Passenger flows" is instrumented by external connecting capacity. See text for sources of data. Standard errors reported in parentheses and clustered at the importer-exporter level. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels respectively.

6.2 Effects on different margins of trade

Standard trade theories predict that fixed export costs can prevent productive firms from exporting (Melitz, 2003; Chaney, 2008). One important source of these costs is market research, which involves finding potential buyers and learning about foreign regulatory environments.

Improved air connectivity between countries can reduce these fixed costs by lowering informational barriers and matching frictions between buyers and sellers (Startz, 2021). This can lead to more firms exporting, including firms that are less productive. Here we estimate the effects of air connectivity on different margins of trade.

We update our merged data to include all products at the six-digit HS level. For a given country pair in a given year, we classify products as new or old based on whether they had been exported by the exporter to the importer in previous years. We then further decompose both old and new products into homogeneous and differentiated products, using the classification of Rauch (1999) or our own criterion based on the price dispersion of products. The differentiation of products, similar to the classification based on industry contract intensity, reflects the fact that both differentiated products and contract-intensive industries often involve exchanges of intricate information throughout their transaction processes.[29] We then aggregate the trade data into four product categories for each country pair in a given year: old homogeneous, old differentiated, new homogeneous, and new differentiated. In order to examine the variation of our estimated effect across these product categories, we conduct regressions at the exporter-importer-category-year level. The coefficients of the interaction terms between product category dummies (with old homogeneous as the reference group) and our air connectivity measure are the key parameters to be assessed.

Table 8 shows the results of our study. In column (1), we use the classification of homogeneous and differentiated products developed by Rauch (1999). In column (2), we use our own calculation of price dispersion, based on the deviation from the median of the price dispersion. The results for the difference between homogeneous and differentiated products are either mixed or minimal. However, the effect for new products is found to be consistently and significantly larger than that for old products. This finding is consistent with the existence of fixed costs in trading new products that have not been previously sold in the market. It also highlights the positive role of in-person business connections in breaking market entry barriers.

[29] Here we use the liberal version of Rauch's (1999) classification, but our results are similar when we use his conservative version. We regard products traded on an organized exchange or with reference prices as homogeneous products, and others as differentiated products. The original classification is at the four-digit SITC level. To map it to our data, we convert the classification to the six-digit HS categories using the correspondence table provided by the UN Trade Statistics at https://unstats.un.org/unsd/trade/classifications/correspondence-tables.asp. The price dispersion for each six-digit product is measured by the coefficient of variation of product prices, which is the ratio of the standard deviation of price to the mean of price for each product. Here the product-specific time trends are removed from prices before the computation to ensure that our price dispersion measure is not driven by product-specific factors, i.e., prices are taken as the residuals $\hat{v}_{ijky}$ from the regression $Price_{ijkt} = \delta_{kt} + v_{ijkt}$, where δ_{kt} is the product-year dummies capturing product-specific time effects.

Table 8. Effects of air connectivity on extensive and intensive margins of trade

	Dep var: log exports Product differentiation based on:	
	Rauch's measure (1)	Price dispersion (2)
Reference products: Old homogeneous		
Passenger flows×Old differentiated	0.048***	-0.011***
	(0.003)	(0.001)
Passenger flows×New homogeneous	0.247***	0.236***
	(0.002)	(0.002)
Passenger flows×New differentiated	0.241***	0.223***
	(0.003)	(0.002)
RTA dummy & tariff	Yes	Yes
Exporter-importer fixed effects	Yes	Yes
Exporter-year fixed effects	Yes	Yes
Importer-year fixed effects	Yes	Yes
Observations	130,619	138,410
Kleibergen-Paap F	183.99	184.81

Notes. Second-stage results of 2SLS estimations are reported. Dependent variable: log exports. Time period: 2013-2018. Unit of observation is exporter-importer-category-year, where category is one of four product categories defined as old homogeneous products (omitted group), old differentiated products, new homogeneous products, and new differentiated products. The differentiation of homogeneous and differentiated products in column (1) follows Rauch's (1999) classification (liberal version), and that in column (2) is based on the deviation from the median of the price dispersion from our own calculation using trade data. The price dispersion for each six-digit product is measured by the coefficient of variation of product prices, i.e. the ratio of the standard deviation of price to the mean of price for each product, and the product-specific time trends are removed from the prices before the computation to ensure that our price dispersion measure is not driven by product-specific macro factors. "Passenger flows" is the log number of air passengers from the importing to the exporting country, "Tariff" is the log of one plus the weighted average tariff rates, and "RTA dummy" is an indicator for both countries being in a same regional trade agreement. "Passenger flows" is instrumented by external connecting capacity. Standard errors reported in parentheses and clustered at the importer-exporter level. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels respectively.

6.3 Online connections and in-person meetings: Substitutes or complements?

The rapid advancement in virtual communication technology has considerably decreased the costs of communication between trade partners. In a recent study by Bailey et al. (2021), it is found that social connectedness, fostered through Facebook friendship networks, alleviates information barriers and promote international trade between nations. This finding challenges the conventional belief that face-to-face meetings are indispensable in international business and raises the possibility of their eventual replacement by virtual communication. To address this inquiry, we incorporate the social connectedness index from Bailey et al. (2021) into our data and empirically investigate the extent to which online and in-person communications act as substitutes or complements.

Specifically, we divide country pairs into four quartiles according to their degree of social connectedness and then estimate Equation (1) for each quartile separately. If virtual connectivity is a substitute for air connectivity, we would observe a weaker trade effect of air connectivity for countries which are better connected online. The key estimated coefficients are plotted in Figure 7. Panel (a) shows the estimated effects for industries with an average contract intensity. The effect decreases as countries become more connected online, although the drop is not statistically significant. Panel (b) illustrates the differential effects on industries with a contract intensity that is one standard deviation above the average, relative to industries with an average contract intensity (Panel (a)). As countries become more connected through online social networks, a noticeable decline in the differential effect becomes apparent, suggesting a reduced sensitivity of trade to air connectivity between countries with well-established online networks. These findings indicate that while air connectivity contributes to trade, its impact is less pronounced in countries with stronger online connections, especially for contract-intensive industries that rely heavily on face-to-face contact. Collectively, these patterns offer suggestive evidence that online communications may, to a certain extent, serve as a viable substitute for in-person interactions in facilitating international trade.[30]

7 Conclusions

Despite the rise of digital communication technology, face-to-face interactions are still widely believed to be essential for business transactions. This paper uses international air passenger flows to assess the importance of in-person communications between buyers and sellers in international trade. We propose a novel instrumental variable that takes advantage of variations in connecting flight capacities in the global air traffic network. For a specific country pair, our identification comes from time variations in the capacity of connecting flights that use third countries as transit points.

Our study finds that international air connectivity has a positive effect on trade, especially for industries that rely heavily on relationship-specific investments and for new products. These results suggest that face-to-face interactions hold particular value in domains characterized by higher risks and the exchange of intricate knowledge that cannot be easily codified. However, we also find that online communications can substitute for in-person meetings to a certain degree.

Our research is also relevant to the COVID-19 pandemic, which restricted international travel

[30]Note that in this research we only look at goods trade. With some caution, however, this analysis could be extended to other contexts. For example, trade in some service sectors are believed to benefit more from the adoption of digital technology. This is because, compared to manufacturing sectors, face-to-face contact in these sectors (e.g., retail and wholesale, and finance and insurance) could be more easily replaced by digital services (Baldwin, 2019).

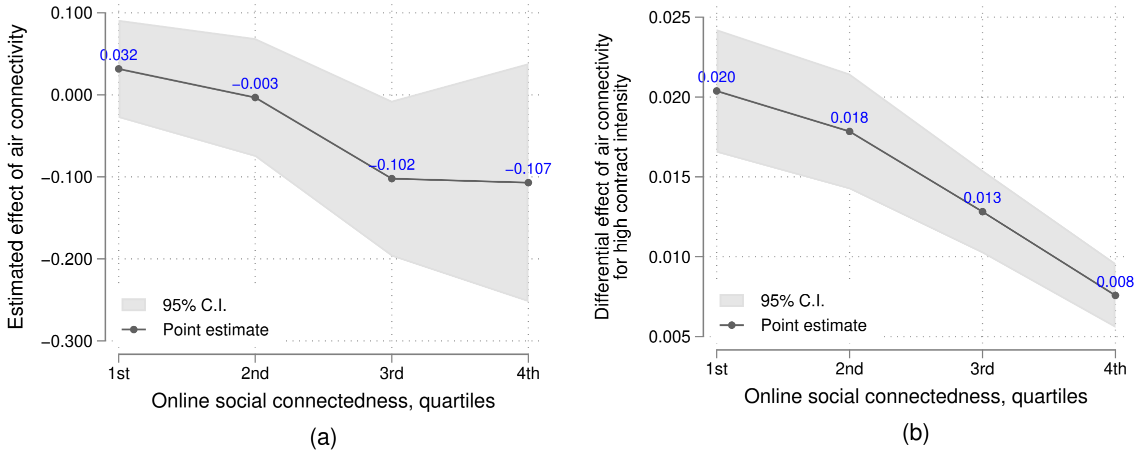

Figure 7. Estimated effects of air connectivity for contract-intensive industries by quartile of social connectedness. This figure shows the 2SLS point estimates and the 95% confidence intervals of the coefficient on the interaction term between air connectivity and contract intensity (θ) in Equation (1) by quartile of social connectedness. Sources of data: authors' calculation based on data compiled from various sources (see Section 3.3); social connectedness index is from Bailey et al. (2021).

and led to a significant drop in trade.[31] The pandemic has forced businesses to rely more on online communications, but it is unclear to what extent this has been able to substitute for face-to-face interactions. As more recent trade and travel data becomes available, it will be interesting to see how the pandemic has affected the role of in-person communications in business. For example, we may see that trade between firms that had established business relationships prior to the outbreak has been more resilient to the pandemic than trade between newly-formed partnerships. We may also see that the proportion of trade that arises from newly-formed partnerships has decreased, as businesses have been reluctant to form new relationships without the opportunity for face-to-face interaction. A full and careful investigation into the trade response to travel restrictions will shed further light on the role of in-person communications in contemporary business activities.

Finally, the instrumental variable we have developed in this research has the potential to be applied in a wider context beyond international trade. It provides a useful tool for researchers to evaluate the causal impact of connections in complex networks on a range of socio-economic outcomes.

References

Air Transport Action Group (2005) 'The economic and social benefits of air transport.' Available at https://www.icao.int/meetings/wrdss2011/documents/jointworkshop2005/atag_socialbenefitsairtransport.pdf

[31] According to the World Trade Organization, the global trade in merchandise decreased by 7.5% in 2020 relative to 2019; see https://data.wto.org/.

Alderighi, Marco, and Alberto A. Gaggero (2017) 'Fly and trade: Evidence from the Italian manufacturing industry.' *Economics of Transportation* 9, 51–60

Álvarez, Inmaculada C., Javier Barbero, Andrés Rodríguez-Pose, and José L. Zofío (2018) 'Does institutional quality matter for trade? Institutional conditions in a sectoral trade framework.' *World Development* 103, 72–87

Andersen, Thomas, and Carl-Johan Dalgaard (2011) 'Flows of people, flows of ideas, and the inequality of nations.' *Journal of Economic Growth* 16(1), 1–32

Anderson, James E., and Douglas Marcouiller (2002) 'Insecurity and the pattern of trade: An empirical investigation.' *The Review of Economics and Statistics* 84(2), 342–352

Anderson, James E., and Eric van Wincoop (2004) 'Trade costs.' *Journal of Economic Literature* 42(3), 691–751

Baier, Scott L., and Jeffrey H. Bergstrand (2001) 'The growth of world trade: Tariffs, transport costs, and income similarity.' *Journal of International Economics* 53(1), 1–27

Bailey, Michael, Abhinav Gupta, Sebastian Hillenbrand, Theresa Kuchler, Robert Richmond, and Johannes Stroebel (2021) 'International trade and social connectedness.' *Journal of International Economics* 129, 103418

Baldwin, Richard (2019) 'Digital technology and telemigration.' In 'World Trade Report 2019' (World Trade Organization)

Battiston, Diego, Jordi Blanes i Vidal, and Tom Kirchmaier (2021) 'Face-to-face communication in organizations.' *Review of Economic Studies* 88(2), 574–609

Berkowitz, Daniel, Johannes Moenius, and Katharina Pistor (2006) 'Trade, law, and product complexity.' *Review of Economics and Statistics* 88(2), 363–373

Bernard, Andrew, Andreas Moxnes, and Yukiko Saito (2019) 'Production networks, geography, and firm performance.' *Journal of Political Economy* 127(2), 639 – 688

Bernhofen, Daniel M., Zouheir El-Sahli, and Richard Kneller (2016) 'Estimating the effects of the container revolution on world trade.' *Journal of International Economics* 98, 36–50

Blonigen, Bruce A., and Anca D. Cristea (2015) 'Air service and urban growth: Evidence from a quasi-natural policy experiment.' *Journal of Urban Economics* 86, 128–146

Brugnoli, Alberto, Antonio Dal Bianco, Gianmaria Martini, and Davide Scotti (2018) 'The impact of air transportation on trade flows: A natural experiment on causality applied to italy.' *Transportation Research Part A: Policy and Practice* 112, 95–107. 2016 ATRS Conference

Büchel, Konstantin, and Maximilian v. Ehrlich (2020) 'Cities and the structure of social interactions: Evidence from mobile phone data.' *Journal of Urban Economics* 119, 103276

Campante, Filipe, and David Yanagizawa-Drott (2018) 'Long-range growth: Economic development in the global network of air links.' *The Quarterly Journal of Economics* 133(3), 1395–1458

Chaney, Thomas (2008) 'Distorted gravity: The intensive and extensive margins of international trade.' *American Economic Review* 98(4), 1707–1721

Chen, Natalie (2004) 'Intra-national versus international trade in the european union: Why do national borders matter?' *Journal of International Economics* 63(1), 93–118

Choi, Changkyu (2010) 'The effect of the internet on service trade.' *Economics Letters* 109(2), 102–104

Cristea, Anca D. (2011) 'Buyer-seller relationships in international trade: Evidence from U.S. States' exports and business-class travel.' *Journal of International Economics* 84(2), 207–220

_ (2020) 'The role of aviation networks for urban development.' Working Paper, University of Oregon

Donaldson, Dave (2018) 'Railroads of the Raj: Estimating the impact of transportation infrastructure.' *American Economic Review* 108(4-5), 899–934

Donaldson, Dave, and Richard Hornbeck (2016) 'Railroads and American economic growth: A "Market Access" approach.' *Quarterly Journal of Economics* 131(2), 799–858

Fageda, Xavier (2017) 'International air travel and FDI flows: Evidence from Barcelona.' *Journal of Regional Science* 57(5), 858–883

Feyrer, James (2009) 'Distance, trade, and income - the 1967 to 1975 closing of the Suez Canal as a natural experiment.' NBER Working Paper No. 15557

_ (2019) 'Trade and income—exploiting time series in geography.' *American Economic Journal: Applied Economics* 11(4), 1–35

Finger, Joseph, and Mordechai Kreinin (1979) 'A measure of 'export similarity' and its possible uses.' *The Economic Journal* 89(356), 905–912

Fink, Carsten, Aaditya Mattoo, and Ileana Cristina Neagu (2005) 'Assessing the impact of communication costs on international trade.' *Journal of International Economics* 67(2), 428–445

Fogel, Robert W. (1964) *Railroads and American Economic Growth: Essays in Econometric History* (Baltimore: Johns Hopkins Press)

Frakes, Michael D., and Melissa F. Wasserman (2021) 'Knowledge spillovers, peer effects, and telecommuting: Evidence from the U.S. Patent Office.' *Journal of Public Economics* 198, 104425

Freund, Caroline L, and Diana Weinhold (2004) 'The effect of the internet on international trade.' *Journal of International Economics* 62(1), 171–189

Fuchs, Andreas, Lennart Kaplan, Krisztina Kis-Katos, Sebastian S. Schmidt, Felix Turbanisch, and Feicheng Wang (2020) 'Mask wars: China's exports of medical goods in times of COVID-19.' Kiel Working Papers 2161

Gaulier, Guillaume, and Soledad Zignago (2010) 'Baci: International trade database at the product-level. the 1994-2007 version.' CEPII Working Paper No. 2010-23

Head, Keith, and Thierry Mayer (2014) 'Gravity equations: Workhorse, toolkit, and cookbook.' vol. 4 (Elsevier) chapter 3, pp. 131–195

Hinde, John, and Clarice G.B. Demétrio (1998) 'Overdispersion: Models and estimation.' *Computational Statistics & Data Analysis* 27(2), 151–170

Hovhannisyan, Nune, and Wolfgang Keller (2015) 'International business travel: An engine of innovation?' *Journal of Economic Growth* 20(1), 75–104

Hummels, David (2007) 'Transportation costs and international trade in the second era of globalization.' *Journal of Economic Perspectives* 21(3), 131–154

ICAO (2018) 'ICAO Annual Report 2018.' Available at https://www.icao.int/annual-report-2018

International Air Transport Association (2020) 'Air connectivity: Measuring the connections that drive economic growth.' Available at https://www.iata.org/en/iata-repository/publications/economic-reports/air-connectivity-measuring-the-connections-that-drive-economic-growth/

Jochmans, Koen (2022) 'Bias in instrumental-variable estimators of fixed-effect models for count data.' *Economics Letters* 212, 110318

Juhász, Réka, and Claudia Steinwender (2018) 'Spinning the web: The impact of ICT on trade in intermediates and technology diffusion.' NBER Working Paper No. 24590

Krautheim, Sebastian (2012) 'Heterogeneous firms, exporter networks and the effect of distance on international trade.' *Journal of International Economics* 87(1), 27–35. Symposium on the Global Dimensions of the Financial Crisis

Levchenko, Andrei (2007) 'Institutional quality and international trade.' *Review of Economic Studies* 74(3), 791–819

Li, Xin, Bo Meng, and Zhi Wang (2019) 'Recent patterns of global productionand GVC participation.' In 'Global Value Chain Development Report 2019: Technical Innovation, Supply Chain Trade, and Workers in a Globalized World' (World Trade Organization) chapter 1

Melitz, Marc J. (2003) 'The impact of trade on intra-industry reallocations and aggregate industry productivity.' *Econometrica* 71(6), 1695–1725

Mullahy, John (1997) 'Heterogeneity, excess zeros, and the structure of count data models.' *Journal of Applied Econometrics* 12(3), 337–350

National Science Board (2018) 'Science and engineering indicators 2018.' Technical Report, National Science Foundation. Available at https://www.nsf.gov/statistics/2018/nsb20181/report

Nitsch, Volker (2000) 'National borders and international trade: Evidence from the european union.' *Canadian Journal of Economics* 33(4), 1091–1105

North, Douglass (1958) 'Ocean freight rates and economic development 1750-1913.' *Journal of Economic History* 18(4), 537–555

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

Rauch, James E. (1999) 'Networks versus markets in international trade.' *Journal of International Economics* 48(1), 7–35

_ (2001) 'Business and social networks in international trade.' *Journal of Economic Literature* 39(4), 1177–1203

Rauch, James E., and Vitor Trindade (2002) 'Ethnic Chinese networks in international trade.' *The Review of Economics and Statistics* 84(1), 116–130

Silva, J. M. C. Santos, and Silvana Tenreyro (2006) 'The log of gravity.' *The Review of Economics and Statistics* 88(4), 641–658

Startz, Meredith (2021) 'The value of face-to-face: Search and contracting problems in Nigerian trade.' Working Paper, Dartmouth College

Storper, Michael, and Anthony J. Venables (2004) 'Buzz: Face-to-face contact and the urban economy.' *Journal of Economic Geography* 4(4), 351–370

Söderlund, Bengt (2022) 'The importance of business travel for trade: Evidence from the liberalization of the soviet airspace.' Working Paper, Lund University

Tang, Linghui (2006) 'Communication costs and trade of differentiated goods.' *Review of International Economics* 14(1), 54–68

UNCTAD (2020) 'Key statistics and trends in international trade 2020.' Technical Report, The United Nations Conference on Trade and Development

Wang, Zhi, Shang-Jin Wei, Xinding Yu, and Kunfu Zhu (2017) 'Characterizing global value chains: Production length and upstreamness.' NBRR Working Paper No. 23261

Wooldridge, Jeffrey M (2010) *Econometric Analysis of Cross Section and Panel Data* (MIT press)

World Bank (2020a) 'Doing Business.' Washington, DC: The World Bank. Available at https://www.doingbusiness.org/en/data

_ (2020b) 'The United Nations Conference on Trade and Development (UNCTAD) Trade Analysis Information System (TRAINS).' Washington, DC: The World Bank. Available at https://wits.worldbank.org/tariff/trains/country-byhs6product.aspx?lang=en

WTTC (2022) 'Travel & Tourism Economic Impact 2022.' World Travel & Tourism Council. Available at https://wttc.org/Portals/0/Documents/Reports/2022/EIR2022-Global%20Trends.pdf

Xiao, Hongyu, Andy Wu, and Jaeho Kim (2021) 'Commuting and innovation: Are closer inventors more productive?' *Journal of Urban Economics* 121, 103300

Yu, Shu, Sjoerd Beugelsdijk, and Jakob de Haan (2015) 'Trade, trust and the rule of law.' *European Journal of Political Economy* 37, 102–115

Online Appendix to

"All Roads Lead to Rome: Global Air Connectivity and Bilateral Trade"

A Construction of the instrument in a flight network

A.1 Flight journeys with two external transit points

We start with the case of a flight journey with two external transit points between an origin country and a destination airport. This is illustrated in Equation (A1), where j is the origin country, i is the destination country, b is the first arrival airport (entry point) in i, airports c and d are the two external transit points located outside countries j and i, and services s and u are flight services that connect c to d and d to b, respectively. The external connecting route is the entire route that starts from airport c all the way to airport b, including the two transit points (c and d) and the two connecting flight services (s and u).

$$\text{Origin country } j \longrightarrow \overbrace{\underbrace{\text{Airport } c \xrightarrow{\text{Service } s} \text{Airport } d}_{\text{Outside countries } i \text{ and } j} \xrightarrow{\text{Service } u} \underbrace{\text{Airport } b}_{\text{Country } i}}^{\textbf{External connecting route}} \tag{A1}$$

We measure the capacity of the above external connecting route as the annual total number of passengers carried through the route by the above flight services, including passengers from all origin countries. However, this measure could still be partly driven by the travel demand between country j and airport d. This is because operating airlines could respond to the connecting demand of passengers from country j by increasing their capacity through more frequent schedules, using larger aircraft, or even adding seats in an existing airplane.

To address this concern, we refine the measure by excluding passengers from the origin country in question (country j) from the total route capacity. This is illustrated in Equation (A2), where we construct the capacity measure based on the same route but only include passengers from countries other than j and i.

$$\underbrace{\text{Origin country } m}_{m \neq i,j} \longrightarrow \overbrace{\underbrace{\text{Airport } c \xrightarrow{\text{Service } s} \text{Airport } d}_{\text{Outside countries } i \text{ and } j} \xrightarrow{\text{Service } u} \underbrace{\text{Airport } b}_{\text{Country } i}}^{\textbf{External connecting route}} \tag{A2}$$

Our measure thus isolates the exogenous part of the capacity that is driven by third countries' travel demand. Intuitively, when there is more demand for flights from third countries to airport b, airlines are more likely to increase their capacity on the route, which in turn improves connectivity for passengers traveling from country j to airport b.

A.2 Flight journeys with one external transit point

Now we consider the case of a journey with only one external transit point en route from country j to airport b. As illustrated in Equation (A3), airport c is now the only transit point, and service v is the flight that links airports d and b.

$$\text{Origin country } j \longrightarrow \overbrace{\underbrace{\text{Airport } c}_{\substack{\text{Outside countries} \\ i \text{ and } j}} \xrightarrow{\text{Service } v} \underbrace{\text{Airport } b}_{\text{Country } i}}^{\text{External connecting route}} \tag{A3}$$

The construction of the exogenous external connecting route capacity is illustrated in Equation (A4). It follows the same steps as in the above case of a journey with two external transit points, except that there is only one transit point in this case.

$$\underbrace{\text{Origin country } m}_{m \neq i,j} \longrightarrow \overbrace{\underbrace{\text{Airport } c}_{\substack{\text{Outside countries} \\ i \text{ and } j}} \xrightarrow{\text{Service } v} \underbrace{\text{Airport } b}_{\text{Country } i}}^{\text{External connecting route}} \tag{A4}$$

A.3 Passengers who start their journey from an external transit point

If connecting flight services pick up local passengers from any of the connecting airports, these passengers should be included in the external connecting capacity. In our example, passengers who travel from airport c as their starting point with services s and u in the route of Equation (A2) are already captured in Equation (A4), where airport c is located in one of the third countries m. For the same reason, we calculate the number of passengers who fly from d as their starting point to b with non-stop service v (in Equation (A2)), and add it back to Equation (A4) as part of the external connecting capacity there.

A.4 Aggregation to the country-pair level

To align our air connectivity measure with the trade data, we further aggregate the above route-level capacity measure to the country-pair level. To do this, for a given year t and origin country j, we aggregate the capacities of country j's external connecting routes by each destination country i (which the destination airport is in). This then allows to have a capacity measure that is specified for each pair of origin and destination countries in each year, i.e., at the ijt level.

In the robustness checks (Sections 5.1 and 5.2) of the paper, each external connecting route is weighted by the reciprocal of its squared relative travel distance (i.e., $1/(\text{dist\_ratio})^2$) and/or by one minus the import similarity index (i.e., 1-ISI). After assigning the weights to these routes, the weighted capacities of these routes are then collapsed to (origin country)-(destination country)-year level (i.e., ijt level) using the same aggregation procedure as above.

B Industry-specific air dependence index and value-weight ratio

We estimate the propensity of air shipping for each HS6 product using two alternative data sources. We first use the Chinese customs data from the General Administration of Customs of China, which differentiates the value of trade goods by mode of transport (air, sea, road, etc.) and by partner country. We use this data to compute an index measuring the intensity, or propensity, of air cargo shipping for each product, conditional on country-pair-specific factors such as distance and geographical conditions. Specifically, we use the 2012 Chinese customs data to estimate the air cargo shipping intensity based on the following regression:

$$AirShare_{kc} = \boldsymbol{\alpha}^k \mathbf{D}_k + \omega_c + \mu_{kc}, \tag{B1}$$

where $AirShare_{kc}$ is the observed percentage of value of trade between China and country c that was shipped by air, $\mathbf{D}_k$ is a vector of dummies for all six-digit industries (converted to six-digit US IO sector codes), $\boldsymbol{\alpha}^k$ is a vector of coefficients associated with the dummies in $\mathbf{D}_k$, ω_c is a country fixed effect that captures all country-specific factors such as distance and geography that could affect the share of air shipping in trade between China and country c, and μ_{kc} is the estimation residual.

Our estimated air shipping intensities are extracted from the estimated values of $\boldsymbol{\alpha}^k$. These intensities can be interpreted as the expected likelihood of air shipping. The underlying assumption is that, after filtering out the influence of distance and geography as well as noise on the choice of transportation method, the air shipping intensities should be similar for all countries. This is because the intensities constructed in this way should mainly reflect the physical and intrinsic features (e.g., unit value and product life) of the product.

Our second measure is based on the product-level bilateral trade records from UN Comtrade in 2012. We map the six-digit HS products in the data to six-digit US IO codes to match our other data. While this dataset does not report trade by transportation method, it includes information on trade value and net weight. This allows us to calculate the average per-kilo value of products in each industry as a proxy for air cargo shipping intensity, based on the assumption that products with a higher value-to-weight ratio are more likely to be transported through air (Feyrer, 2019). In a similar manner to Equation (B1), we run the following regression:

$$VW_{kij} = \boldsymbol{\alpha}^k \mathbf{D}_k + \omega_{ij} + \mu_{kij}, \tag{B2}$$

where the outcome variable VW_{kij} is the per-kilo value of trade for products in industry k between country i and j. Again, $\mathbf{D}_k$ is a vector containing a set of industry (six-digit US IO sector) dummies, ω_{ij} contains country-pair fixed effects that capture the average per-kilo value of trade between country pairs, and μ_{kij} is the error term. The vector $\boldsymbol{\alpha}^k$ contain the estimated coefficients associated with the industry dummies in $\mathbf{D}_k$. These coefficients capture the expected likelihood of air shipping for each industry, net of the influences of country-pair-specific factors.

C Additional tables and figures

Table C1. Description of OAG international air traffic data - number of country pairs and passengers

	Country pairs	All passengers	Non-stop passengers	Connecting passengers
2013	26,276	986,631,620	838,663,903	147,967,717
2014	26,336	1,043,134,160	885,301,996	157,832,164
2015	26,678	1,106,346,345	936,644,490	169,701,855
2016	26,786	1,189,779,881	1,010,109,571	179,670,310
2017	26,996	1,279,156,257	1,091,092,079	188,064,178
2018	27,043	1,350,996,795	1,156,567,488	194,429,307
Total	160,115	6,956,045,058	5,918,379,527	1,037,665,531

Table C2. New airports, new routes, and new flight services in international connecting services

	New airports			New routes			New flight services		
	Passengers (persons) (1)	Share in served country-pair connections (%) (2)	Share in global connecting capacity (%) (3)	Passengers (persons) (4)	Share in served country-pair connections (%) (5)	Share in global connecting capacity (%) (6)	Passengers (persons) (7)	Share in served country-pair connections (%) (8)	Share in global connecting capacity (%) (9)
2014							2,823,116	2.07	1.79
2015				194,143	0.26	0.11	5,646,429	3.57	3.33
2016	6	1.44	0.00	595,446	0.52	0.33	7,919,598	4.62	4.41
2017	1,844	0.11	0.00	1,047,636	0.79	0.56	10,042,242	5.53	5.34
2018	32,807	0.24	0.00	1,679,661	1.15	0.86	12,160,611	6.46	6.26
Total	34,657	0.23	0.00	3,516,886	0.75	0.34	38,591,996	4.62	3.72

Notes. New airports are airports which started their operation of scheduled passenger flights during the period 2013-2018. New routes are flight routes which started their operation of scheduled passenger flights between existing airports during the period 2013-2018. New flight services are scheduled passenger flight services which started their operation on existing flight routes during the period 2013-2018.

Table C3. Most connected country pairs

Ranked by all passengers (two-way)					Ranked by connecting passengers (two-way)				
Rank	Country pairs	All passengers (mn people)	Share of connecting passengers	Trade value (tn USD)	Rank	Country pairs	Connecting passengers (mn people)	Share of connecting passengers	Trade value (tn USD)
1	GBR - SCH	642.87	0.01	3,156.08	1	SCH - USA	33.22	0.16	3,407.08
2	SCH - USA	202.61	0.16	3,407.08	2	IND - USA	22.48	0.78	361.96
3	MEX - USA	153.59	0.01	3,090.91	3	SCH - THA	14.41	0.47	227.72
4	CAN - USA	151.84	0.01	3,129.91	4	AUS - SCH	13.16	1.00	259.44
5	SCH - TUR	121.93	0.01	798.29	5	IND - SCH	12.35	0.50	510.52
6	CHN - KOR	91.63	0.00	1,404.78	6	IND - SAU	10.70	0.35	195.64
7	GRC - SCH	89.53	0.02	213.53	7	AUS - GBR	9.63	0.85	49.94
8	NOR - SCH	86.11	0.01	745.81	8	CHN - SCH	8.68	0.19	3,190.80
9	GBR - USA	85.45	0.09	614.89	9	PHL - USA	8.42	0.64	115.04
10	RUS - SCH	79.30	0.03	1,649.70	10	IDN - SCH	7.92	0.81	178.35
11	JPN - KOR	79.14	0.00	470.06	11	GBR - USA	7.59	0.09	614.89
12	CHN - JPN	74.86	0.05	1,607.47	12	USA - VNM	7.34	0.98	261.47
13	CHN - THA	73.49	0.05	396.51	13	CHN - USA	7.27	0.17	3,291.05
14	IRL - SCH	70.45	0.05	567.85	14	JPN - SCH	7.14	0.24	802.01
15	MAR - SCH	63.86	0.01	233.22	15	PAK - SAU	6.90	0.28	21.39
16	CHN - HKG	62.83	0.00	1,756.45	16	AUS - IND	6.79	0.90	100.84
17	ARE - IND	59.47	0.04	253.67	17	GBR - IND	6.66	0.39	92.29
18	GBR - IRL	55.09	0.00	260.85	18	THA - USA	6.13	0.96	224.65
19	JPN - USA	47.43	0.06	1,153.52	19	PHL - SCH	5.41	0.91	109.37
20	CHN - SCH	46.51	0.19	3,190.80	20	ISR - USA	5.05	0.38	142.88

Notes. The code "SCH" denotes the Schengen Area.

Table C4. Non-stop vs. connecting service passengers

		2SLS	
	Dep var: non-stop passenger flows (1)	1st stage. Dep var: connecting passenger flows (2)	2nd stage. Dep var: log exports (3)
External connecting capacity	-0.016***	0.285***	
	(0.003)	(0.015)	
Connecting passenger flows			0.025*
			(0.013)
Contract intensity×Connecting passenger flows			0.013***
			(0.001)
Non-stop passenger flows			0.006***
			(0.002)
Tariff	-0.002	-0.003**	-0.202***
	(0.009)	(0.001)	(0.007)
RTA dummy	Yes	Yes	Yes
Exporter-year fixed effects	Yes	Yes	No
Importer-year fixed effects	Yes	Yes	No
Exporter-importer fixed effects	Yes	Yes	Yes
Exporter-industry-year fixed effects	No	Yes	Yes
Importer-industry-year fixed effects	No	Yes	Yes
Observations	54,194	2,839,161	2,839,161
Kleibergen-Paap F			188.24

Notes. Time period: 2013-2018. Unit of observation is exporter-importer-year in columns (1)-(2), and exporter-importer-industry-year in column (3). "External connecting capacity" is the capacity of connecting flight services departing from a third country, "Contract intensity" is an industry-level standardized measure of reliance on relation-specific investments, "Connecting passenger flows" is the log number of air passengers from the importing to the exporting country via connecting services, "Non-stop passenger flows" is the log number of air passengers from the importing to the exporting country via non-stop services, "Tariff" is the log of one plus the weighted average tariff rate for an exporter-importer-year observation (columns (1)-(2)) or an exporter-importer-industry-year observation (column (3)), and "RTA dummy" is an indicator for both countries being in a same regional trade agreement. "Connecting passenger flows" is instrumented by external connecting capacity in column (3). See text for sources of data. Standard errors reported in parentheses and clustered at the importer-exporter level. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels respectively.

Table C5. External connecting capacity and exports: Reduced-form estimates

	Dep var: log exports					
	Exporter-importer-year				Exporter-importer-industry-year	
	Low contract intensity trade. OLS (1)	High contract intensity trade. OLS (2)	Low contract intensity trade. PPML (3)	High contract intensity trade. PPML (4)	OLS (5)	PPML (6)
External connecting capacity	-0.002	0.019	0.070**	0.009	0.007*	0.245***
	(0.015)	(0.014)	(0.031)	(0.034)	(0.004)	(0.067)
Contract intensity $\times$ External connecting capacity					0.008***	0.017***
					(0.000)	(0.004)
Tariff	-0.058*	-0.339***	-0.044***	-0.070***	-0.202***	0.112*
	(0.032)	(0.034)	(0.016)	(0.018)	(0.007)	(0.059)
RTA dummy	Yes	Yes	Yes	Yes	Yes	Yes
Exporter-year fixed effects	Yes	Yes	Yes	Yes	No	No
Importer-year fixed effects	Yes	Yes	Yes	Yes	No	No
Exporter-importer fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Exporter-product-year fixed effects	No	No	No	No	Yes	Yes
Importer-product-year fixed effects	No	No	No	No	Yes	Yes
Observations	54,194	54,194	52,834	52,788	2,839,161	2,839,161
Adjusted R^2	0.866	0.866			0.410	
Pseudo R^2			0.999	0.999		0.655

Notes. Reduced-form estimates are reported. Dependent variable: log exports. Time period: 2013-2018. Unit of observation is exporter-importer-year in columns (1)-(4), and exporter-importer-industry-year in columns (5)-(6). "External connecting capacity" is the capacity of connecting flight services departing from a third country, "Contract intensity" is an industry-level standardized measure of reliance on relation-specific investments, "Tariff" is the log of one plus the weighted average tariff rate for an exporter-importer-year observation (columns (1)-(4)) or an exporter-importer-industry-year observation (columns (5)-(6)), and "RTA dummy" is an indicator for both countries being in a same regional trade agreement. See text for sources of data. Standard errors reported in parentheses and clustered at the importer-exporter level. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels respectively.

Table C6. Robustness checks: Alternative definitions of differentiated products

	Dep var: log exports			
	Relation-specific products: Neither sold on organized exchange nor reference priced		Relation-specific products: Not sold on organized exchange	
	Rauch's classification used for organized exchange and reference priced products		Rauch's classification used for organized exchange and reference priced products	
	Conservative (1)	Liberal (2)	Conservative (3)	Liberal (4)
Passenger flows	0.031**	0.031**	0.028*	0.028*
	(0.014)	(0.014)	(0.014)	(0.014)
Contract intensity × Passenger flows	0.011***	0.014***	0.001***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
RTA dummy & tariff	Yes	Yes	Yes	Yes
Exporter-importer fixed effects	Yes	Yes	Yes	Yes
Exporter-industry-year fixed effects	Yes	Yes	Yes	Yes
Importer-industry-year fixed effects	Yes	Yes	Yes	Yes
Observations	2,839,161	2,839,161	2,839,161	2,839,161
Kleibergen-Paap F	232,069.20	232,065.58	232,090.94	232,088.82

Notes. Second-stage results of 2SLS estimations are reported. Dependent variable: log exports. Time period: 2013-2018. Unit of observation is exporter-importer-industry-year. "Passenger flows" is the log number of air passengers from the importing to the exporting country, "Contract intensity" is an industry-level standardised measure of reliance on relation-specific investments, "Tariff" is the log of one plus the weighted average tariff rate, and "RTA dummy" is an indicator for both countries being in a same regional trade agreement. See text for sources of data. Standard errors reported in parentheses and clustered at the importer-exporter level. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels respectively.

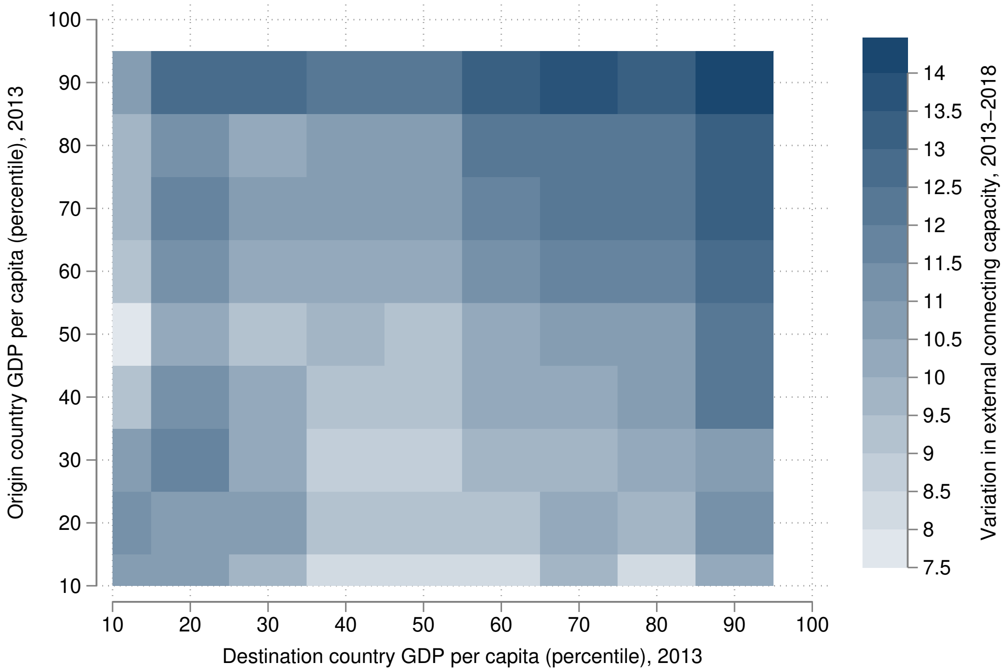

Figure C1. Country-pair incomes and variation in external connecting capacity. The horizontal and vertical axes measure the percentiles of GDP per capita of origin and destination countries in 2013, respectively. The darkness of the colored squares indicates the variation in external connecting capacity for the period 2013-2018 that is contributed by the origin-destination countries in a percentile square. Sources of data: authors' calculation based on data complied from various sources (see Section 3.3).