Supplementary Information for: Knowledge Diffusion in the Network of International Business Travels

Michele Coscia[1,2,*], Frank Neffke[1], and Ricardo Hausmann[1]

[1]Center for International Development - Harvard University, 79 Jfk St, Cambridge 02138 US
[2]ITU Copenhagen, Rued Langgaards Vej 7, 2300 Copenhagen, Denmark
[*]To whom correspondence should be addressed; E-mail:
michele_coscia@hks.harvard.edu.

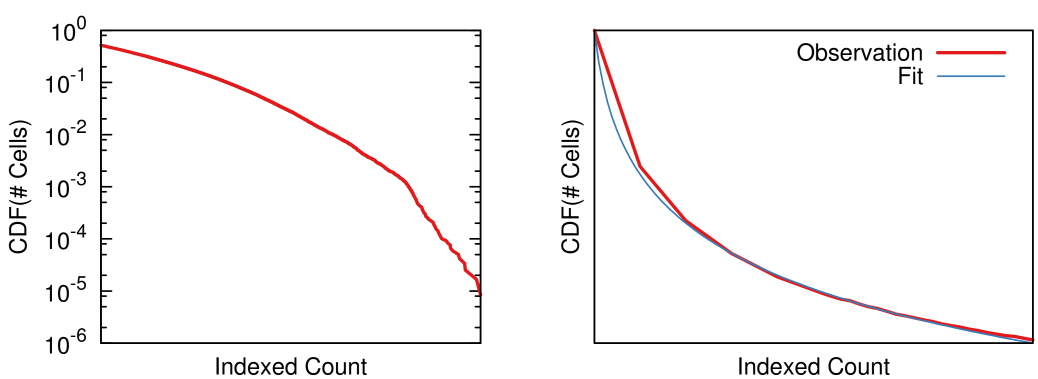
Supplementary Results

This document contains supplementary results for the paper "Knowledge Diffusion in the Network of International Business Travel." All p-values reported in this document and in the main text for estimated regression parameters and correlations refer to two-sided t-tests. The associated confidence intervals are set at an α of 5% (i.e., we report 95% confidence intervals throughout both documents).

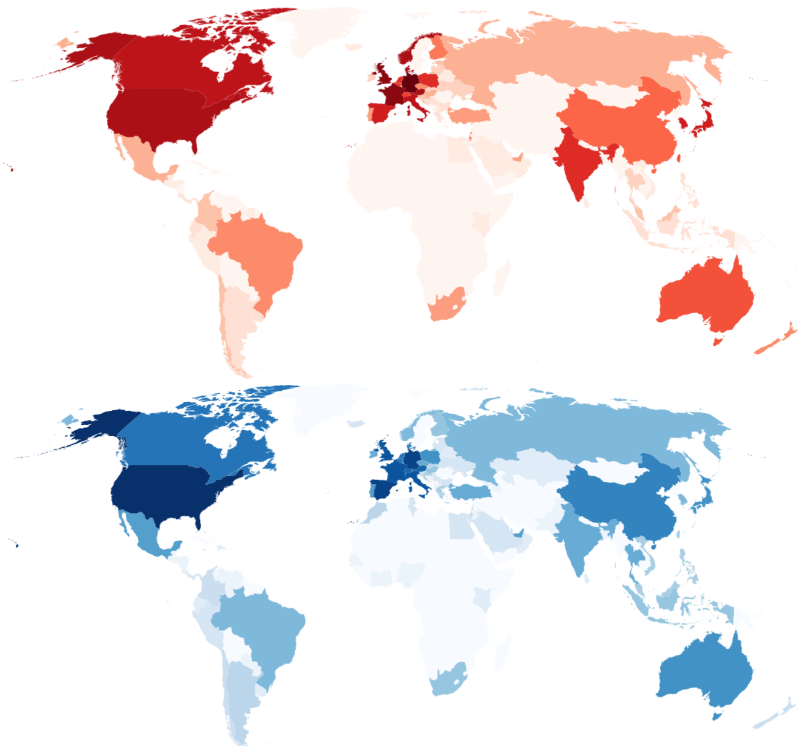
S1 Supplementary Tables and Figures for the Methods Section

Albania, *Andorra*, Angola, *Antigua and Barbuda*, Argentina, Armenia, *Aruba*, Australia, Austria, Azerbaijan, Bahrain, Bangladesh, *Barbados*, Belarus, Belgium, Belize, *Bermuda*, Bolivia, Bosnia and Herzegovina, Brazil, Brunei Darussalam, Bulgaria, Côte d'Ivoire, Canada, *Cayman Islands*, Chile, China, Colombia, Costa Rica, Croatia, Cyprus, Czech Republic, Denmark, Dominican Republic, Ecuador, Egypt, El Salvador, Estonia, Finland, France, Georgia, Germany, Ghana, Greece, *Guam*, Guatemala, Guyana, Haiti, Honduras, Hong Kong SAR - China, Hungary, Iceland, India, Indonesia, Iraq, Ireland, Israel, Italy, Jamaica, Japan, Jordan, Kazakhstan, Kenya, Korea, Kuwait, Latvia, Lebanon, Lithuania, Luxembourg, Macao SAR - China, Macedonia, Malaysia, Malta, Mauritius, Mexico, Moldova, Morocco, Netherlands, *Netherlands Antilles*, New Zealand, Nicaragua, Nigeria, Norway, Oman, Pakistan, Panama, Papua New Guinea, Paraguay, Peru, Philippines, Poland, Portugal, Qatar, Russia, Rwanda, *San Marino*, Saudi Arabia, Senegal, *Serbia*, *Seychelles*, Singapore, Slovak Republic, Slovenia, South Africa, Spain, Sri Lanka, *St. Kitts and Nevis*, Suriname, Swaziland, Switzerland, *Taiwan*, Tajikistan, Thailand, The Bahamas, Trinidad and Tobago, Tunisia, Turkey, *Turks and Caicos Islands*, Ukraine, United Arab Emirates, United Kingdom, United States, Uruguay, Uzbekistan, Venezuela, Vietnam, Zimbabwe.

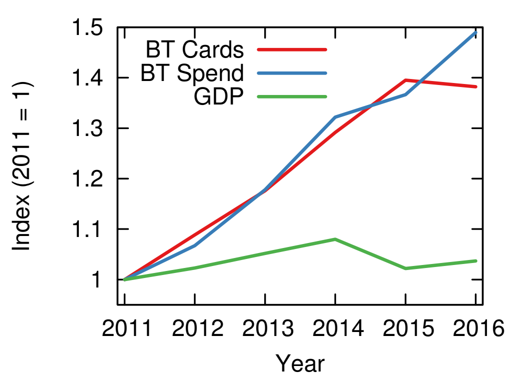
Supplementary Table 1: The list of 127 countries included in this study. Countries with populations below 2.5 million underlined; countries for which information is missing on some of the alternative bilateral links in italics.



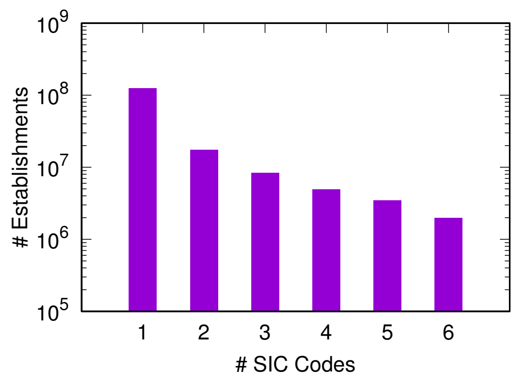
Supplementary Figure 1: CDF of the number of cells (y-axis) containing a given number of cards (x-axis). (Left) The full distribution of the observed (non-censored) values. Both axes are plotted on log-scales. Labeling on the horizontal axis is omitted as a privacy and data protection control. (Right) The head of the distribution, with the added CDF points extracted from the interval frequencies. The vertical axis is scaled in levels (not logs); labeling is omitted as a privacy and data protection control. In blue we add our best log-linear fit.



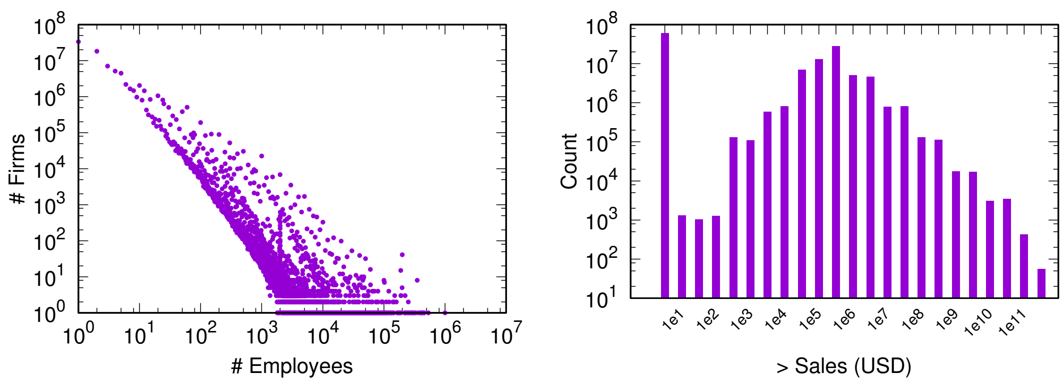
Supplementary Figure 2: Number of business travelers by origin (top) and destination (bottom).



Supplementary Figure 3: Growth of business travel indexed expenditures compared to global GDP growth. Both figures in current (nominal) USD and indexed against 2011.



Supplementary Figure 4: Distribution of number of SIC codes per establishment in the D&B dataset.



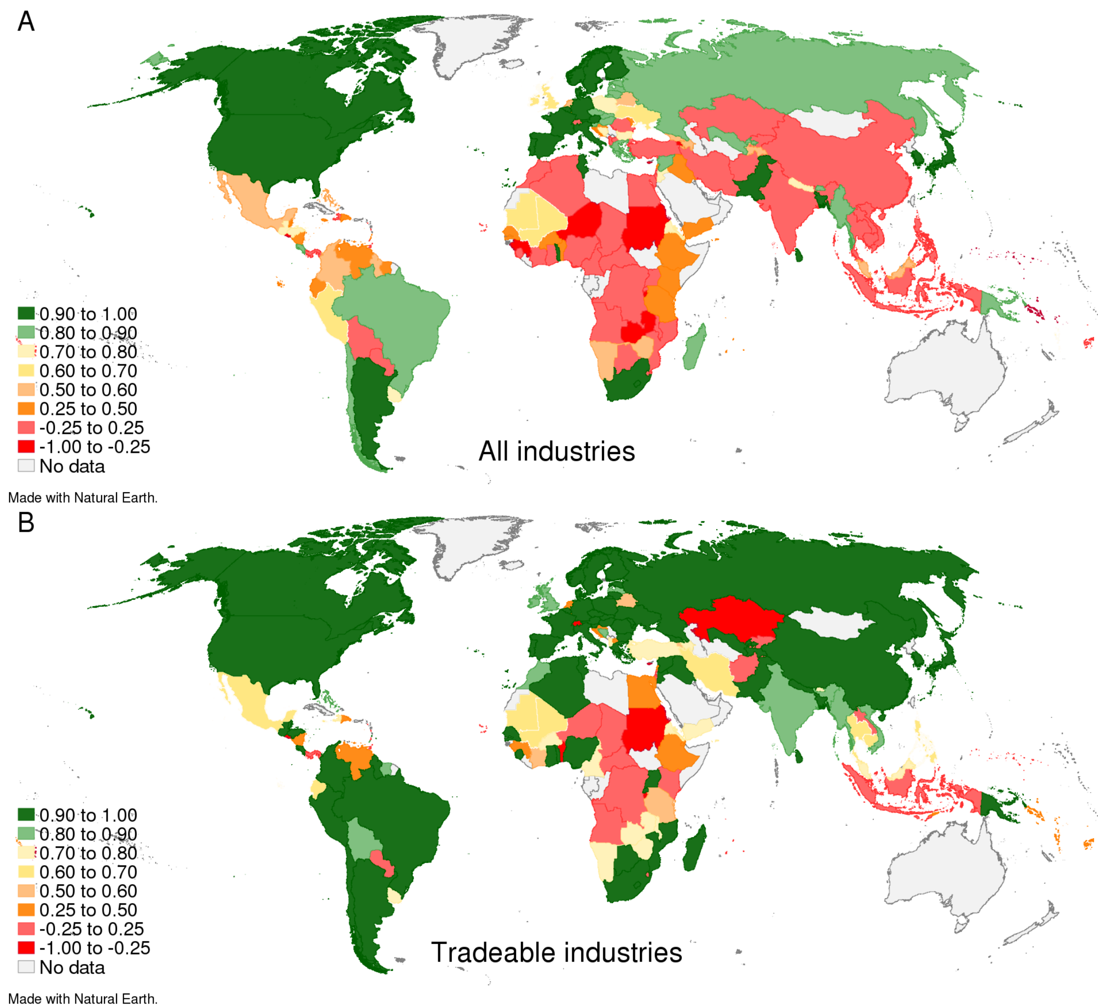
Supplementary Figure 5: Descriptive statistics of the D&B dataset. (Left) Number of firms with a given number of employees. (Right) Number of firms with a given sale volume.

S2 Representativeness D&B data

The D&B data contain information on establishments across the entire world. However, their coverage differs between countries and sectors. In particular, because the D&B data can be seen as a registry of establishments, economic activities in the informal sector of a country will not be listed. To explore how representative D&B data are across countries and sectors, we collected official industry employment statistics from IPUMS (*1, 2*), the Statistics Bureau of Japan (*3*), Eurostat (*4*), the OECD (*5*) and the World Bank (*6*). We match industry codes in these data sets to the SIC87 classification at the level of ten broad sectors. Next, we represent each country as a vector of industries that records the number of people employed in each sector. Finally, we calculate the Pearson correlation between these vectors based on official data to an analogous vector based on the D&B data.

Supplementary Figure 6 (top) shows the correlation results when using all industries. For some countries, like Australia and New Zealand, we were unable to find census data at the required level of granularity (colored grey on the maps). However, given the overall number of establishments listed in D&B, as well as the typical coverage of industrialized countries, we expect the D&B coverage to be good in these countries. The figure shows that for a substantial number of countries, the correlation between D&B data and official statistics is poor. Crucially, these include populous countries with large economies such as India and China.

To a large extent, however, these poor correlations are the result of large informal sectors. To some extent, this is due to the poor coverage of the informal sector. Whereas D&B intends to record the economic activities by registered companies, activities in the informal sector of a country will typically not be registered. This is a particularly important factor in developing countries. Activities in the informal sector will be poorly covered in our D&B data as well as in official government statistics. Moreover, the main hypothesis in our paper is that, for an industry



Supplementary Figure 6: Correlation between official industry employment statistics and the statistics as derived from the D&B dataset. A: considering all industries. B: considering only tradeable industries. Country is coded as missing if the country is either not covered by D&B or if no suitable official data sources were available.

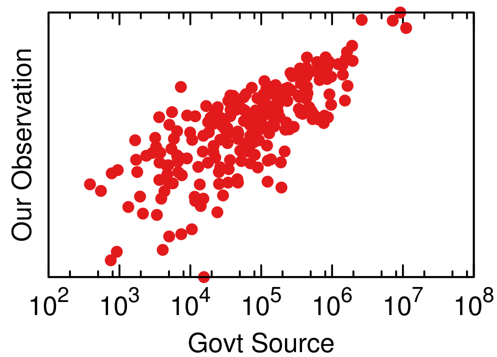
to emerge or grow in a country, the country must have access to the industry-specific know-how used in the industry's production processes. However, the location of some industries depends less on know-how and more on geo-physical characteristics or the demand for their products. For instance, the location of agricultural and extractive industries will not be pinned down by access to know-how. Moreover, commodity prices are highly volatile leading to boom-bust cycles that are unpredictable from information on access to knowledge. Similarly, the location of retail activities and restaurants will depend more on the amount of wealth in a local market than on access to industry-specific know-how. For these reasons, we restrict our analysis to a subset of industries that are neither part of the extractive or nontraded sectors of the economy, nor of sectors that are prone to high informality rates.

In particular, we restrict our data to the SIC divisions D (Manufacturing), E (Transportation, Communications, Electric, Gas, And Sanitary Services), and H (Finance, Insurance, And Real Estate), as well as industry groups 73, 81, and 87 within division I (Business, Legal, and Engineering, Accounting, Research, Management Services). Supplementary Figure 6 (bottom) shows the correlation results when focusing on these economic activities. Using only this selection of industries almost uniformly leads to a much close alignment of D&B and official employment data. Most critically, the coverage of the D&B data for large emerging markets is much better in these industries.

S3 Representativeness BT data

It is difficult to obtain reliable estimates of bilateral international business travel. However, many countries publish aggregated statistics about their foreign visitors, sometimes distinguishing them by their reason of visit. We collect data about business travels from:

- Eurostat (*7*) for European countries: Austria, Belgium, Bulgaria, Croatia, Cyprus, CzechRepublic, Denmark, Estonia, Finland, Germany, Greece, Hungary, Ireland, Latvia, Lithua-



Supplementary Figure 7: Comparison to governmental business travel data sources.

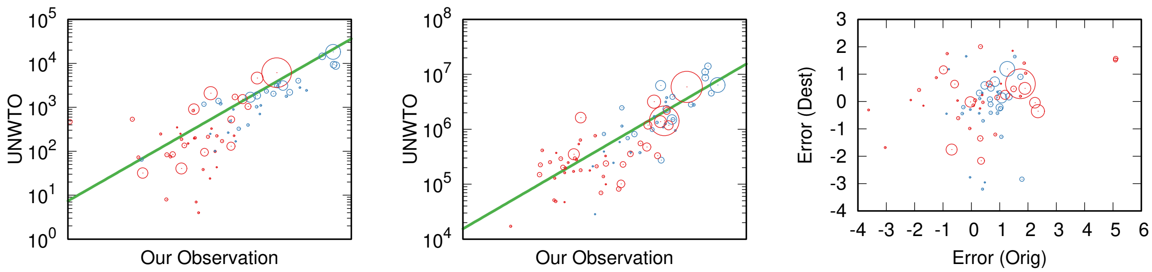
nia, Luxembourg, Malta, Poland, Portugal, Slovakia, Slovenia, Spain, Sweden and Switzerland.

- Government sources for individual countries: France (*8*), Israel (*9*), Italy (*10*), Japan (*11*), The Netherlands (*12*), South Africa (*13*), UK (*14*), USA (*15*)

This yields data on 212 country-country pairs, plus 36 aggregate statistics (where either the origin or the destination is the entire world).

The data cover the period 2014-2016 (endpoints included), which is completely included in our observation period (2011-2016). Supplementary Figure 7 depicts the relationship between the two. Overall the data line up well, with a Pearson correlation of 0.84 ± 0.04 ($p < 0.001$, $N = 237$).

The country producing such statistics are overwhelmingly developed countries, thus we still cannot exclude a systematic bias against developing countries. Therefore, we perform a second representativity check. We obtained data on international business travel from the United Nations World Tourism Organization (UNWTO) (*16*). UNWTO covers most countries in the world, including major emerging markets such as China, India, Indonesia and Brazil. As a consequence, UNWTO provides information on either the total incoming or the total outgoing business travel volume or both for 90 of the 96 countries we study in our main analysis. The



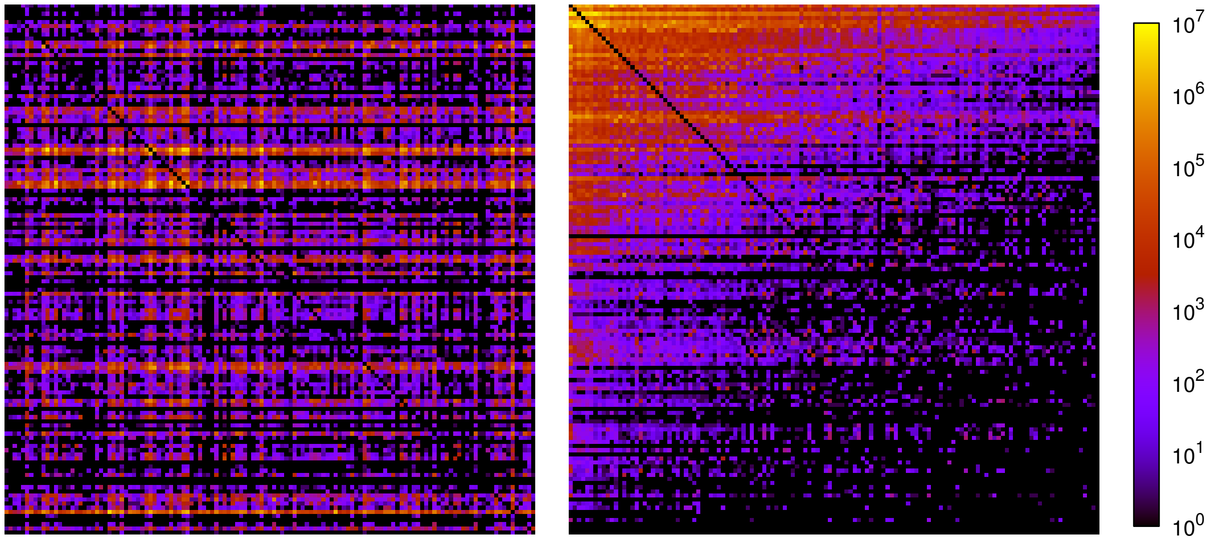
Supplementary Figure 8: In all plots, point color encodes the country type (red = developing, blue = developed). Point size is proportional to the square root of the population (in hundreds of millions). The green line shows the best fit, population weighted. (Left) Payment card spend (x axis) against UNWTO business travel expenditure (y axis) per country of origin. (Middle) Number of cards observed in the destination (x axis) against number of business trips in UNWTO (y axis). (Right) Residuals from the previous regressions.

remaining six countries are: Iraq, Malaysia, Portugal, Slovakia, Switzerland and the United Arab Emirates.

There are two possible sources of systematic bias when considering developing countries: (i) travelers from these countries do not use payment cards when they travel, and (ii) developing countries do not accept credit cards at home – because of lacking financial infrastructure or because they are predominantly cash-based economies.

For the analysis that follows, we divide countries into two bins. A developing country is a country whose GDP per capita PPP is lower than $20k.

Supplementary Figure 8 shows the relationship between our estimates and UNWTO data. In Supplementary Figure 8(left) we see that both data sources are in strong agreement when it comes to total spend by country of origin. Although the relation between our estimates and the UNWTO data is less tight for developing countries, there is no clear systematic under- or overcounting of either group in our data. In fact, the noisier relation for developing countries seems mostly driven by some small countries. In sum, travelers from developing countries seem to use payment cards as much as, if not more than, travelers from developed countries. This answers the first concern.

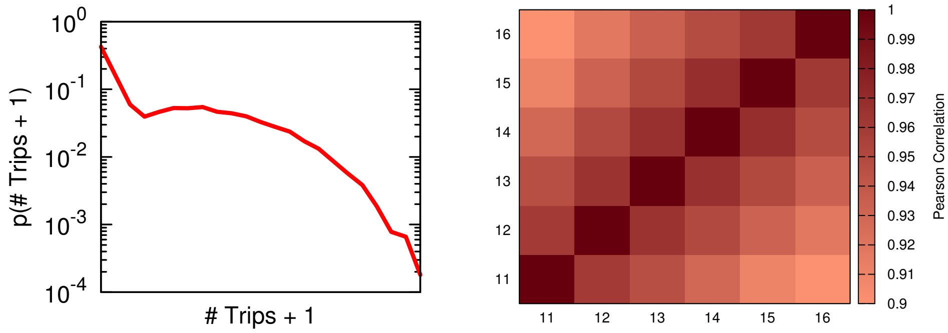


Supplementary Figure 9: Matrix representation of the business travel dataset. Each row is a country of destination and each column a country of origin. (Left) The matrix with rows and columns sorted alphabetically by ISO code. (Right) The same matrix, with rows and columns sorted by a country's number of incoming travelers (i.e., by the row sum).

In Supplementary Figure 8(middle) we again see agreement of our estimates with the UNWTO data. That is, also the total number of trips by destination is not systematically biased towards developed countries. In fact, plotting the error terms of both regressions against each other shows no clear biases in either direction (Supplementary Figure 8(right)). There is no bias towards developing countries because, if we create a dummy variable for whether or not a country is a developing country, this variable is not correlated with the error terms (t-statistic (1,69): 1.588, $p = 0.117$).

S4 Business Travel Weight Distribution

The time-averaged business travel matrix, $\bar{B}$, is sparse. As reported in the paper almost 43% of the possible directed connections between countries have zero travelers and fewer than 9% of non-zero entries exceed five travelers, due to the data censoring, the lowest possible non-zero entry (see Methods section in the main paper). Fig. 9 (left) illustrates this with a graphical



Supplementary Figure 10: (Left) Log-binned distribution of $B_{cc'}$ values. The x-axis records the number of business travelers ($+1$, so that we can include zero entries in the log-log scale). The y-axis records the probability of observing a cell in $B_{cc'}$ for the given amount of trips. (Right) The Pearson correlation of the (log-transformed) matrix values for each pair of years. Note that the color scale starts at 0.9.

representation of the business travel matrix. The black cells are cells for which we observe fewer than one visitor (note that the color scale is in logs), which effectively means zero.

Fig. 9 (left) lacks a clear pattern. In Fig. 9 (right), we reorder rows and columns by a country's total incoming business travel (i.e., the matrix' row sum). Now, an upper-triangular shape emerges, which is typical for nested systems – as originally discovered in ecology (*17,18*) and more recently in economics at the macro (*19*) and micro levels (*20*). The nestedness implies that countries that are visited by every other country also visit most of them. On the other hand, the countries that visit only one or two countries end up visiting the countries that are visited by everybody else. As a consequence, the business travel network exhibits well-connected hubs. We leave a more in-depth analysis of further structural characteristics of this network for future research.

The sparseness of the matrix is evident once we plot its value distribution (Fig. 10, left). The mode of the distribution is zero, which is observed in 43% of all cells. Furthermore, the distribution of non-zero values is highly skewed. While $\sim 4\%$ of the non-zero cells are equal to the minimum possible weight (after censoring), only three cells are larger than 5,000,000. In fact, these three cells contain more than 9% of all observed business travels. Even though they do not observe a power-law, these values still span several orders of magnitude.

Finally, the business travel structure is highly stable. Supplementary Figure 10 (right) shows this by depicting the Pearson correlation of the stacked vector of log-transformed matrix elements between different years. This correlation never falls below 0.90 ± 0.004 ($p < 0.001$, $N = 9,487$) and exceeds 0.96 ± 0.002 ($p < 0.001$, $N = 9,487$) when comparing consecutive years.

S5 Explaining Business Travel

To understand how business travel correlates with other phenomena that may affect a country's development path, we explore how much of the variance in bilateral business travel patterns we can explain by the following factors:

- Economy size. Two large economies are likely to connect to each other, because rich countries will be able to send and attract many travelers. We measure an economy's size using its GDP in 2011 (in current USD) as provided by the United Nations (*21*).
- Geographical distance. The costs of travel in monetary terms and in terms of time increase with distance. We use the weighted average distance between countries' major cities from CEPII (*22*).
- Migration links. Two countries may also be connected by a large stock of migrants. These migrants may maintain strong links with their countries of origin, some of which may result in business ties and business travel. We use migration data published by the United Nations, using 2010 stocks (*23*).
- Greenfield foreign direct investments (FDI). If companies from one country set up new economic establishments in another country, we expect a certain amount of business

travels to emerge between the two countries. These travels may help assess the profitability of the potential investment, or implement and monitor the investment itself. Our FDI investments come from the Financial Times' FDI Markets Library (`http://www.fdimarkets.com/`) for the period 2003-2015. The data were originally compiled from press releases on new cross-border investments. We record the FDI intensity between two countries in millions of dollars invested.

- Equity links (ownership). Greenfield and other foreign investments result in a presence of establishments in one country whose headquarters are located in another. We expect that the headquarter country will send business travelers to exert control over its subsidiaries. The Dun & Bradstreet dataset (`http://www.dnb.com/`) contains an entry for each establishment in the world, with a link to its global headquarter establishment. We measure equity connections, both in terms of numbers of establishments owned and number of employees employed by these global head quarters.

- Trade. Trade relations are likely to involve travel as well. Customers may need to visit suppliers and vice versa. We use cleaned version of the CEPII trade dataset (UN Comtrade data, (*24*)) that can be downloaded here: `http://atlas.cid.harvard.edu/engage#data-download`. We record bilateral trade links in USD.

- Shared history and language. Countries may also be strongly connected because they share a language or parts of their history. We rely again on the CEPII distance dataset (*22*). Language links are proxied by having shared languages and evaluated by a dummy variable that equals one if the countries share a language that is spoken by a majority or a minority of their population. For historical links we use variables on a country's colonial past. This information is again coded in a dummy variable, which equals one if two countries have or had a colonial link or if they share a common colonizer.

VARIABLES	(1) distance	(2) FDI	(3) ownership	(4) trade	(5) migration	(6) language	(7) colonizer	(8) all
log(GDP origin)	1.145*	0.600*	0.730*	0.769*	0.987*	1.151*	1.159*	0.442*
	(0.087)	(0.074)	(0.071)	(0.097)	(0.092)	(0.091)	(0.092)	(0.078)
	0.97 - 1.32	0.45 - 0.75	0.59 - 0.87	0.58 - 0.96	0.80 - 1.17	0.97 - 1.33	0.98 - 1.34	0.29 - 0.60
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log(GDP destination)	0.714*	0.432*	0.384*	0.388*	0.471*	0.720*	0.727*	0.176*
	(0.044)	(0.047)	(0.044)	(0.049)	(0.050)	(0.048)	(0.048)	(0.050)
	0.63 - 0.80	0.34 - 0.52	0.30 - 0.47	0.29 - 0.48	0.37 - 0.57	0.62 - 0.82	0.63 - 0.82	0.08 - 0.27
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
log(km. dist.)	-1.372*							-0.547*
	(0.094)							(0.099)
	-1.56 - -1.19							-0.74 - -0.35
	(0.000)							(0.000)
log(FDI)		0.623*						0.272*
		(0.036)						(0.031)
		0.55 - 0.70						0.21 - 0.33
		(0.000)						(0.000)
log(ownership)			1.105*					0.615*
			(0.082)					(0.077)
			0.94 - 1.27					0.46 - 0.77
			(0.000)					(0.000)
log(trade)				0.259*				0.129*
				(0.029)				(0.020)
				0.20 - 0.32				0.09 - 0.17
				(0.000)				(0.000)
log(migrants)					0.286*			0.072
					(0.027)			(0.022)
					0.23 - 0.34			0.03 - 0.12
					(0.000)			(0.001)
common language						1.213*		0.120
						(0.225)		(0.168)
						0.77 - 1.66		-0.21 - 0.45
						(0.000)		(0.478)
common col. hist.							0.947*	-0.234
							(0.238)	(0.163)
							0.47 - 1.42	-0.56 - 0.09
							(0.000)	(0.153)
origin FE	no	no	no	no	no	no	no	no
destination FE	no	no	no	no	no	no	no	no
R^2	0.607	0.641	0.654	0.559	0.560	0.510	0.501	0.739
# obs.	12,656	12,656	12,656	12,656	12,656	12,656	12,656	12,656
F-stat	189.8	349.2	344.7	115.2	238.6	143.8	127.1	304.9
d.o.f. F-stat	(3, 112)	(3, 112)	(3, 112)	(3, 112)	(3, 112)	(3, 112)	(3, 112)	(9, 112)
Prob. > F	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Supplementary Table 2: Explaining business travels with bilateral linkages between countries: OLS.

Supplementary Table 2 reports the results of a series of gravity models. The dependent variable is the number of business travelers from an origin to a destination country. All variables, dependent and independent, are log-transformed, except for the language and colonial dummies. We restrict the sample in each model to the 113 countries for which data on all bilateral links is available, leaving us with $113 * 112 = 12,656$ directed edges without self loops. We also control for the size of both economies in each model. In models 1 to 7 we add one bilateral relation at a time. In the final model, we estimate the effect of all variables simultaneously. To avoid missing values when bilateral links are absent, we add 1 to each variable we log-transform.

The R^2 of the gravity models never falls below 50%, the fit for the model when only using the colonial variable. Among the models that feature only one bilateral relation at a time, the most predictive is the model that uses the number of establishments in the destination country that are owned by firms head quartered in the origin country ($R^2 = 65.4\%$, F-statistic (3,12652): 7969, $p < 0.001$), followed by the model that uses greenfield FDI as an explanatory variable ($R^2 = 64.1\%$, F-statistic (3,12652): 7540, $p < 0.001$).

When all variables enter the estimation model simultaneously, most bilateral relations retain a statistically significant effect. The exceptions are the variables that capture whether countries share a common language or colonial past.

In Fig. 1C of the main paper, we provide an estimate of the relative explanatory force each factor has in predicting the flow of business travelers. Tab. 3 reports the full set of contributions. The contribution to the R^2 were derived following the method by Feldman (*25*) as implemented in (*26*). The variable that explains most of the variation in business travel patterns is FDI, followed by the number of foreign-owned establishments, and the economy of origin's size. Language and colonial links offer little to no explanatory power.

Variable	R^2 contribution
Unexplained	0.2605
Equity ownership	0.2540
FDI	0.2100
Trade	0.1717
GDP origin	0.0613
Distance (km)	0.0170
Migration	0.0163
GDP destination	0.0088
Colonial	0.0003
Language	0.0001

Supplementary Table 3: The contributions to the R^2 of each variable in model (8) from Supplementary Table 2.

S6 Business Travel and economic transformation

In the main paper, we show that TQI predicts the intensive margin of economic growth: how fast do existing industries in a country grow? Here, we expand this analysis by also studying the extensive margin of growth: which new industries emerge in a country? and which old industries disappear?

The observations in these regression analyses are industry-country pairs, i.e., national industries, such as manufacturing *Pharmaceutical Preparations* in *Switzerland*. Tab. 4 reports outcomes of OLS regressions that look at different aspects of growth. The first model regresses the growth of the logarithm of an existing national industry's number of establishments between 2011 and 2016. The second model focuses on existing industries' growth in terms of number of employees instead. The third and fourth column analyze the extensive margin of growth: they study the likelihood that a nonexisting industry enters a national economy or that an existing industry exits from it. The final two columns combine the intensive and extensive margins of growth by pooling data for all existing and nonexisting industries in a country. In these regressions, we augment the establishment and employment counts by one unit to avoid observations

VARIABLES	(1) $\Delta\log(\#\text{est.})$	(2) $\Delta\log(\text{emp})$	(3) entry	(4) exit	(5) $\Delta\log(\#\text{est.}+1)$	(6) $\Delta\log(\text{emp}+1)$
log(TQI)	0.060	0.077*	0.022	-0.007	0.039	0.113*
	(0.021)	(0.022)	(0.007)	(0.005)	(0.018)	(0.027)
	0.02 - 0.10	0.03 - 0.12	0.01 - 0.04	-0.02 - 0.00	0.00 - 0.07	0.06 - 0.17
	(0.005)	(0.001)	(0.003)	(0.149)	(0.032)	(0.000)
log(#est. in 2011)	-0.039			-0.037*		
	(0.020)			(0.006)		
	-0.08 - 0.00			-0.05 - -0.02		
	(0.055)			(0.000)		
log(emp. in 2011)		-0.126*				
		(0.016)				
		-0.16 - -0.09				
		(0.000)				
log(#est.+1 in 2011)					-0.092*	
					(0.017)	
					-0.13 - -0.06	
					(0.000)	
log(emp.+1 in 2011)						-0.178*
						(0.015)
						-0.21 - -0.15
						(0.000)
country FE	yes	yes	yes	yes	yes	yes
industry FE	yes	yes	yes	yes	yes	yes
R^2	0.341	0.251	0.198	0.319	0.317	0.282
# obs.	37,596	35,149	20,076	39,917	60,000	60,000
F-stat	4.6	32.4	9.4	21.2	15.2	72.9
d.o.f. F-stat	(2, 95)	(2, 95)	(1, 91)	(2, 95)	(2, 95)	(2, 95)
Prob. $>$ F	0.012	0.000	0.003	0.000	0.000	0.000
base prob.			0.1215	0.0581		

Supplementary Table 4: Industry growth along different margins. Dependent variables: Intensive margin: Col. 1: growth in number of establishments, Col. 2: growth in employees, Extensive margin: Col. 3: entry dummy variable, Col. 4: exit dummy variable, Col. 5: growth in number of establishments + 1, Col. 6: Growth in employees + 1. For columns The independent variables include a mean-reversion term, as well as dummy variables for the country of origin and for the country of destination.

that would result in $\log(0)$. All models control for country of origin and country of destination fixed effects, as well as the base-year's (2011) industry size. Standard errors are clustered at the level of the country and at the level of the industry to adjust confidence intervals for correlated errors within an industry or a country.

The variable of interest is the logarithm of TQI. Tab. 4 shows that a 10% increase in business travel flows is associated with a $0.60\% \pm 0.42\%$ higher growth in number of establishments of existing industries, a $0.77\% \pm 0.44\%$ higher growth in employment in these industries, and a

VARIABLES	(1) Δ log(#est.)	(2) Δ log(emp)	(3) entry	(4) exit	(5) Δ log(#est. + 1)	(6) Δ log(emp + 1)
log(TQI)	0.070*	0.079*	0.023*	-0.009	0.038	0.100*
	(0.018)	(0.019)	(0.005)	(0.004)	(0.012)	(0.019)
	0.04 - 0.10	0.04 - 0.12	0.01 - 0.03	-0.02 - 0.00	0.01 - 0.06	0.06 - 0.14
	(0.000)	(0.000)	(0.000)	(0.051)	(0.003)	(0.000)
log(#est. in 2011)	-0.031			-0.036*		
	(0.018)			(0.006)		
	-0.07 - 0.00			-0.05 - -0.02		
	(0.088)			(0.000)		
log(emp. in 2011)		-0.126*				
		(0.014)				
		-0.15 - -0.10				
		(0.000)				
log(#est.+1 in 2011)					-0.084*	
					(0.015)	
					-0.11 - -0.05	
					(0.000)	
log(emp.+1 in 2011)						-0.169*
						(0.013)
						-0.19 - -0.14
						(0.000)
country FE	yes	yes	yes	yes	yes	yes
industry FE	yes	yes	yes	yes	yes	yes
R^2	0.337	0.239	0.173	0.296	0.302	0.260
# obs.	43,688	39,852	33,056	46,319	79,375	79,375
F-stat	8.0	40.7	21.7	25.1	16.3	87.8
d.o.f. F-stat	(2, 126)	(2, 126)	(1, 122)	(2, 126)	(2, 126)	(2, 126)
Prob. > F	0.001	0.000	0.000	0.000	0.000	0.000
base prob.			0.1011	0.0568		

Supplementary Table 5: Industry growth along different margins, using all 127 countries. Dependent variables: Intensive margin: Col. 1: growth in number of establishments, Col. 2: growth in employees, Extensive margin: Col. 3: entry dummy variable, Col. 4: exit dummy variable, Col. 5: growth in number of establishments + 1, Col. 6: Growth in employees + 1. For columns The independent variables include a mean-reversion term, as well as dummy variables for the country of origin and for the country of destination.

0.22 ± 0.14 percentage points increase (over a 12% base) in the probability that a new industry will enter the country's economy. The estimated coefficient for industry-exit, in contrast, is negative, yet not statistically significant.

Tab. 5 repeats these analyses for the entire sample of 127 countries, including the ones with fewer than 2.5 million inhabitants. Results are similar to the ones in the restricted sample. If anything, the estimated effect of business travel is stronger in this larger sample: the negative effect on exit probabilities has turned significant at the 1% level and the same holds for the positive effect on the growth in number of establishments in model (5), which pools observations for the intensive and extensive margins.

S6.1 Alternative bilateral links

Tab. 2 shows that business travel flows correlate with many other bilateral relations between countries: their distance, how much they trade with one another, invest in each other's economies etc.. In fact, prior research has already shown that countries are more likely to diversify into products in which their neighbors (*27*) and the host countries of their diaspora (*28*) specialize. An important question, therefore, is whether the business travel network tells us anything beyond other bilateral linkages between countries.

To answer this question, we create matrices with elements X_{di} that mimic the TQI, but for which we replace B_{od} (the elements of the origin-destination matrix of business travel) with a different bilateral relation R_{od}:

$$X = AR^{rel},$$

where each R^{rel} represents one of the variables presented in Section S5.

Next, we use X^{rel}_{ci} as a new control variable. Because the elements of some of the X^{rel} matrix may be zero, we add 1 before we log-transform to these variable. Tab. 6 reports results.

When the alternative indexes enter in isolation (models 1-8), each displays its expected correlation: negative for geographical distance and positive in all other cases. However, the variables based on a shared language or colonial history do not exhibit statistically significant associations with national industry growth rates. Note, moreover, that the statistically strongest effect is associated with the TQI.

Model 9 of Tab. 6 shows the results when all indexes enter the regression equation simultaneously. In this specification, the only variable that retains a somewhat statistically significant

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
VARIABLES	TQI	dstQI	lanQI	colQI	migQI	fdiQI	ownQI	trdQI	all	all
log(GDP/cap)										-0.061
										(0.064)
										-0.19 - 0.07
										(0.343)
log(pop)										0.057
										(0.046)
										-0.03 - 0.15
										(0.219)
log(#est. in ind.)										0.005
										(0.044)
										-0.08 - 0.09
										(0.914)
log(#est. in 2011)	-0.041	-0.045	-0.039	-0.039	-0.042	-0.041	-0.041	-0.041	-0.046	-0.048
	(0.020)	(0.021)	(0.020)	(0.020)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
	-0.08 - -0.00	-0.09 - -0.00	-0.08 - 0.00	-0.08 - 0.00	-0.08 - -0.00	-0.08 - -0.00	-0.08 - -0.00	-0.08 - -0.00	-0.09 - -0.00	-0.09 - -0.01
	(0.047)	(0.035)	(0.056)	(0.056)	(0.045)	(0.049)	(0.049)	(0.047)	(0.033)	(0.021)
log(TQI)	0.065								0.025	0.077
	(0.020)								(0.027)	(0.034)
	0.02 - 0.11								-0.03 - 0.08	0.01 - 0.14
	(0.002)								(0.355)	(0.025)
log(dstQI)		-0.187*							-0.136	0.031
		(0.050)							(0.054)	(0.041)
		-0.29 - -0.09							-0.24 - -0.03	-0.05 - 0.11
		(0.000)							(0.014)	(0.457)
log(lanQI+1)			0.079						0.014	-0.014
			(0.098)						(0.112)	(0.149)
			-0.11 - 0.27						-0.21 - 0.24	-0.31 - 0.28
			(0.418)						(0.902)	(0.925)
log(colQI+1)				0.120					0.113	0.239
				(0.119)					(0.136)	(0.235)
				-0.12 - 0.36					-0.16 - 0.38	-0.23 - 0.70
				(0.316)					(0.405)	(0.311)
log(migQI+1)					0.025				0.009	-0.018
					(0.008)				(0.008)	(0.024)
					0.01 - 0.04				-0.01 - 0.02	-0.07 - 0.03
					(0.002)				(0.272)	(0.450)
log(fdiQI)						0.056			0.031	0.054
						(0.023)			(0.026)	(0.042)
						0.01 - 0.10			-0.02 - 0.08	-0.03 - 0.14
						(0.016)			(0.248)	(0.194)
log(ownQI)							0.047		-0.007	-0.069
							(0.022)		(0.027)	(0.037)
							0.00 - 0.09		-0.06 - 0.05	-0.14 - 0.00
							(0.034)		(0.784)	(0.066)
log(trdQI)								0.064	-0.002	0.054
								(0.021)	(0.022)	(0.057)
								0.02 - 0.11	-0.05 - 0.04	-0.06 - 0.17
								(0.003)	(0.924)	(0.339)
country FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	no
industry FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	no
R^2	0.346	0.346	0.345	0.345	0.346	0.346	0.346	0.346	0.347	0.073
# obs.	36,689	36,689	36,689	36,689	36,689	36,689	36,689	36,689	36,689	36,689
F-stat	5.6	7.0	2.0	2.3	5.3	3.3	3.1	4.8	2.2	12.1
d.o.f. F-stat	(2, 93)	(2, 93)	(2, 93)	(2, 93)	(2, 93)	(2, 93)	(2, 93)	(2, 93)	(9, 93)	(12, 93)
Prob. $>$ F	0.005	0.002	0.139	0.104	0.007	0.042	0.048	0.010	0.026	0.000

Supplementary Table 6: Alternative bilateral relations. Dependent variable: change in logarithm of the number of establishments in existing national industries. Main independent variable: logarithm of Travel Quantity Index + 1 ($\log(TQI)$). Alternative Quantity Indexes based on: geographical distance ($\log(dstQI)$), shared language ($\log(lanQI+1)$), shared colonial history ($\log(colQI+1)$), migration ($\log(migQI+1)$), FDI ($\log(fdiQI)$), ownership ($\log(ownQI)$), trade ($\log(trdQI)$) information.

effect is the one based on the distance between countries. The collinearity among the variables is essentially too high. When instead of our preferred specification, which includes industry and country fixed effects, we use the specification of column (3) in Tab. 1 of the main text, the effect of TQI becomes more statistically significant again ($p = 0.027$, column 10, Tab. 6).

How should we interpret these results? It is hard to interpret the exact meaning of the models of columns (9) and (10). Some of the alternative bilateral relations, most notably the FDI, ownership and trade links, represent channels through which we would expect a potential effect of business travel to materialize. For instance, business travel may be used to find a trading partner or to monitor an investment. Controlling for such channels is therefore not necessarily appropriate. Similarly, given that geographical distance affects the costs – and therewith the intensity – of business travel, it is hard to interpret the meaning of the effect of business travel on the diffusion of industries keeping distance constant. For the present paper, however, it is reassuring that the explanatory force of business travel does not disappear when we control for a number of plausible diffusion channels.

S6.2 Changes in industry aggregation

So far, all regressions have been conducted using the 4-digit level of the SIC87 industry classification. The classification distinguishes among over 1,000 different industries, of which we use 625 that are tradable, but not based on natural resources. While this level of aggregation provides a large number of observations, the downside is that it ignores the fact that know-how may spill over between such narrowly defined industries. For instance, know-how related to "Knit Outerwear Mills"(SIC 2253) will presumably also be useful to gain a competitive advantage in "Knit Underwear and Nightwear Mills" (SIC 2254).

To explore this issue further, we aggregate the data to the 3-, and 2-digit level as well as to the level of sectors (so-called "divisions"). This reduces the number of (tradable, non-natural-resource-based) industries to 220, 40 and 4.

Tab. 7 reports the results when we re-estimate our preferred model (column (3) of Tab. 1 of the main text), using four, three or two digit industries, as well as using sectors. The effect of TQI is statistically significant at the four- ($p = 0.005$) and three-digit ($p < 0.002$) levels. As we move to more aggregated industries, more and more of the spillovers of business travel in one industry to narrowly defined industries in the same sector should be captured. In line with this, we see that the point estimate of the elasticity of industry growth with respect to business travel increases as we successively aggregate industries. However, due to the reduction in the number of observations, the statistical power of the models weaken and the effect of business travel becomes insignificant in the final model.

	(1)	(2)	(3)	(4)
VARIABLES	4-digit	3-digit	2-digit	sector
log(#est. in 2011)	-0.039	-0.078*	-0.121*	-0.153*
	(0.020)	(0.022)	(0.021)	(0.041)
	-0.08 - 0.00	-0.12 - -0.03	-0.16 - -0.08	-0.23 - -0.07
	(0.055)	(0.001)	(0.000)	(0.000)
log(TQI)	0.060	0.095	0.116	0.274
	(0.021)	(0.030)	(0.061)	(0.176)
	0.02 - 0.10	0.04 - 0.15	-0.01 - 0.24	-0.08 - 0.62
	(0.005)	(0.002)	(0.061)	(0.124)
country FE	yes	yes	yes	yes
industry FE	yes	yes	yes	yes
R^2	0.341	0.455	0.631	0.837
# obs.	37,596	16,731	3,667	384
F-stat	4.6	8.0	17.5	7.3
d.o.f. F-stat	(2, 95)	(2, 95)	(2, 95)	(2, 95)
Prob. > F	0.012	0.001	0.000	0.001

Supplementary Table 7: Growth of industries: aggregation. Idem model (3) of Tab. 1 in the main text, but aggregating industry classification at the level of 4-(model 1), 3- (model 2) and 2-(model 3) digits, as well as the level of sectors ("divisions"). Because of the small number of industries in the sector and 2-digit classification, we cluster standard errors at the country level only in these models.

2714, 2742, 2789, 2815, 2816, 2860, 2871, 2872, 2873, 2874, 2875, 2876, 2877, 2879, 2890, 3221, 3222, 3223, 3224, 3232, 3330, 3343, 3352, 3353, 3354, 3413, 3414, 3415, 6672, 6811, 6812, 6821, 6831, 6841, 6851, 6861, 6871, 6891, 6899, 7283, 9710.

Supplementary Table 8: Harmonized System 4-digit commodity codes classified as natural resources by (*29*)

S6.3 Growth of exports

So far, results suggest that TQI helps predict the industrial transformation of a country. As a final robustness check, we test our framework in a second dataset. The dataset we choose are the international trade data collected by United Nations Comtrade and downloaded from `http://atlas.cid.harvard.edu/engage#data-download`. The data contain information on countries' exports by product category. Once again, we drop countries with fewer than 2.5 million inhabitants and drop natural-resource exports, using the definition applied in (*29*) (see Tab. 8).

Once again, we distinguish between growth at the intensive and at the extensive margin. Growth at the intensive margin is measured by the change in the logarithm of the total value of a country's exports of products that the country had already exported in 2011. At the extensive margin, we distinguish between the appearance of new products and the disappearance of existing products in a country's export basket. TQI is defined as before, with the understanding that the activity matrix A is now based on countries' export profiles.

Tab. 9 repeats the analyses of Tab. 1 of the main text. Tab. 10 shows results for growth at the intensive and extensive margins as in Tab. 4. The results in both tables point in the same direction as what we found so far: a larger inflow of business travelers is associated with a stronger growth in exports ($p < 0.001$). However, we now find this effect only for existing exports. At the extensive margin in columns (2) and (3) of Tab. 10, the estimated coefficients of TQI are not statistically significant.

	(1)	(2)	(3)
VARIABLES	$\Delta\log(\text{exprts})$	$\Delta\log(\text{exprts})$	$\Delta\log(\text{exprts})$
log(TQI)	0.160*	0.150*	0.321*
	(0.016)	(0.029)	(0.047)
	0.13 - 0.19	0.09 - 0.21	0.23 - 0.41
	(0.000)	(0.000)	(0.000)
log(exprts in 2011)	-0.233*	-0.239*	-0.305*
	(0.015)	(0.014)	(0.015)
	-0.26 - -0.20	-0.27 - -0.21	-0.33 - -0.27
	(0.000)	(0.000)	(0.000)
log(GDP/cap)		-0.002	
		(0.041)	
		-0.08 - 0.08	
		(0.962)	
log(pop)		0.042	
		(0.028)	
		-0.01 - 0.10	
		(0.141)	
log(global exprts in prod.)		0.036	
		(0.030)	
		-0.02 - 0.10	
		(0.225)	
country FE	no	no	yes
product FE	no	no	yes
R^2	0.101	0.103	0.204
# obs.	59,777	59,777	59,776
F-stat	125.8	57.8	202.6
d.o.f. F-stat	(2, 95)	(5, 95)	(2, 95)
Prob. $>$ F	0.000	0.000	0.000

Supplementary Table 9: Growth of exports.

VARIABLES	(1) Δ log(exprts)	(2) entry	(3) exit	(4) Δ log(exprts + 1)
log(TQI)	0.321*	0.013	0.000	0.343*
	(0.047)	(0.009)	(0.005)	(0.059)
	0.23 - 0.41	-0.00 - 0.03	-0.01 - 0.01	0.23 - 0.46
	(0.000)	(0.134)	(0.936)	(0.000)
log(exprts in 2011)	-0.305*		-0.008*	
	(0.015)		(0.001)	
	-0.33 - -0.27		-0.01 - -0.01	
	(0.000)		(0.000)	
log(exprts+1 in 2011)				-0.472*
				(0.014)
				-0.50 - -0.44
				(0.000)
country FE	yes	yes	yes	yes
product FE	yes	yes	yes	yes
R^2	0.204	0.313	0.194	0.289
# obs.	59,776	7,840	62,421	70,333
F-stat	202.6	2.3	28.5	575.0
d.o.f. F-stat	(2, 95)	(1, 95)	(2, 95)	(2, 95)
Prob. $>$ F	0.000	0.134	0.000	0.000
base prob.		0.3052	0.0424	

Supplementary Table 10: Evolution of countries' export baskets: intensive and extensive margin.

S7 Causality

The results presented thus far do not necessarily show that business travel *causes* growth in economic activities. On the one hand, business travel may help forge trade, FDI and equity links between countries by facilitating the exchange of know-how and know-who. On the other hand, growth in economic activities may make it attractive for business travelers to visit a country to seek trade, FDI and equity links. This suggests that the observed correlation between business travel and economic growth may equally mean that business travel *causes* economic growth or that it *results* from it.

However, note that either explanation for the correlation between business travel and growth builds on the idea that business travel enables countries to establish economic linkages, i.e., that it *accompanies* such linkages. In other words, business travel creates a temporary physical proximity between economic agents from two different countries. This proximity, in turn, may enable transactions that otherwise wouldn't have taken place. For instance, business travelers may be dispatched to identify relevant trading partners or help implement new work flows or technologies in local subsidiaries. The physical proximity afforded by business travel in these examples enables the transmission of know-who or know-how, regardless of whether growth leads to business travel or vice versa. As such, business travel is at best a proximate, not a distal cause: it accompanies economic transactions, but is not their ultimate driver. These transactions, in turn, may either cause or be caused by economic growth. However, absent such transactions, the mere presence of representatives of foreign firms in a country will unlikely have causal effects on the country's economic growth beyond the Keynesian effect of the travel-related expenditures by these representatives.

Business travel's nature as an enabling device for knowledge transfer complicates the estimation of causal effects. In particular, the models in Sec 6 of SI – where we estimated the effect of TQI on the growth of local industries, while partialling out the effects of trade, FDI and equity links – effectively close down the channels through which business travel would affect growth. As such, they are inadequate tests of the causal effect of business travel on economic growth. In this section, we therefore take a different approach and estimate this causal effect by means of instrumental variables (IV) estimation.

IV estimation requires that we find a variable that predicts business travel from an industry to a country of destination, but that – conditional on control variables – is unrelated to the growth of that activity, except for through its effect on business travel. The variable we select for this task is based on bilateral visa regimes between countries. Such a variable is attractive, because visa regimes solely affect the movement of people. As such, the only way in which they should affect economic ties between two countries is through the extent to which they facilitate or hinder the flow of travelers.

The data were collected from `https://github.com/ilyankou/passport-index-dataset`. The visa regimes refer to the current year, as the data source does not provide historical data. While this might introduce noise, since the regimes do not refer to the period of observation, note that visa regimes are remarkably static. Data from the DEMIG policy database shows only 13 bilateral or multilateral visa regimes changes in the 2009-2014 period (*30, 31*).

The visa regime data code the visa regime for travel from country o to country d into a categorical variable that assumes one of four values: (0) no visa required, (1) electronic travel authorization required, (2) visa can be obtained on arrival and (3) visa required. We use this information in three different ways to construct instrumental variables that predict the number of business travelers associated with industry i that arrive in country d, BT_{di}.

For the first instrument, we regress the bilateral business travel flows between countries, BT_{od}, on a dummy group that encodes the visa-regime categories:

$$\log BT_{od} = \alpha + \gamma_1 V_{od}^{(1)} + \gamma_2 V_{od}^{(2)} + \gamma_3 V_{od}^{(3)} + \epsilon^{od}$$

where $V_{od}^{(iv)}$ is a dummy for visa regime iv governing the travel from country o to country d. We use the estimated coefficients to predict business travel flows from o to d:

$$\hat{BT}_{od} = e^{\hat{\alpha}+\hat{\gamma}_1 V_{od}^{(1)}+\hat{\gamma}_2 V_{od}^{(2)}+\hat{\gamma}_3 V_{od}^{(3)}},$$

Note that this prediction only uses information on visa regimes, not on countries' economies. Next, we distribute these predicted flows across industries in the same way in which TQI distributes actual business travel flows, i.e., by using the shares of each industry in a country of origin o and summing across origins:

$$\hat{BT}_{di}^{(I)} = \sum_o \hat{BT}_{od} \frac{A_{io}}{A_o}$$

where A_{io} represents the number of establishments in industry i in country o and omitted subscripts denote summations over the corresponding dimension. These predicted business travel inflows form our first instrument.

For the second instrument, instead of first predicting business travel flows, we distribute the visa-category dummies directly by the industry shares of the country of origin. This yields four distinct instruments:

$$V_{di}^{(II),iv} = \sum_o V_{od}^{(iv)} \frac{A_{io}}{A_o}$$

where $iv \in 0, 1, 2, 3$. Note that, because industry shares will add up to one within each origin, one of these instruments is redundant and will be omitted. The resulting group of variables forms our second set of instruments. Essentially, these instruments sum an industry's local establishment shares across origins that can reach the destination through a given visa regime.

For the third instrument, we interact the visa-category dummies, not with the share, but with the number of an industry's establishments in the country of origin:

$$V_{di}^{(III),iv} = \sum_o V_{od}^{(iv)} A_{io}$$

Note that, in this case, none of the instruments is redundant. The instruments now reflect how large the component of the global industry is that can access the destination country through each of the four visa regimes.

We use these instruments in Two-Stage Least Squares (2SLS) models. In a first stage, the instruments are used to predict the number of business travelers associated with each industry in a country of destination. The instruments enter these first-stage regressions log-transformed. To avoid losing observations to $\log(0)$ issues, we add a small number (the midpoint between zero and the next smallest observed value in the data) to each variable before taking logarithms.

Tables 11 - 13 show the results of a 2SLS estimator, each using a different (set of) instrumental variable(s). The columns correspond to the different margins of growth: the growth of the number of establishments in existing industries, the creation of new industries, the disappearance of existing industries and a pooled regression of the growth in the logarithm of the number of establishments plus 1. The estimation controls for country fixed effects and a mean reversion term, but not for industry fixed effects.

In none of the models does the Hansen J statistic (χ^2(2) for models using $V_{di}^{(II)}$ and χ^2(3) for models using $V_{di}^{(III)}$ as instruments) reject the null hypothesis that the instruments are exogenous at any conventional level, except for the entry-regression in Supplementary Table 13. However, note that the Hansen J statistic is not necessarily appropriate here: the instruments reflect a categorical variable that was coded into a dummy group. They should thus not be regarded as separate variables. A more convincing sign of the instruments' validity is that we find remarkably similar causal effect estimates when we use markedly different ways to construct these instruments.

The causal effects of business travel in the 2SLS models are somewhat greater than the correlational effects from a simple OLS regression (see Supplementary Table 14, which reports the results of this OLS regression with the same control variables as in the 2SLS regressions).

	(1)	(2)	(3)	(4)
VARIABLES	$\Delta\log(\#\text{est.})$	entry	exit	$\Delta\log(\#\text{est.}+1)$
log(TQI)	0.138*	0.055*	0.005	0.104*
	(0.019)	(0.005)	(0.004)	(0.016)
	0.10 - 0.18	0.04 - 0.07	-0.00 - 0.01	0.07 - 0.13
	(0.000)	(0.000)	(0.191)	(0.000)
log(#est. in 2011)	-0.038		-0.038*	
	(0.020)		(0.006)	
	-0.08 - 0.00		-0.05 - -0.02	
	(0.063)		(0.000)	
log(#est.+1 in 2011)				-0.085*
				(0.016)
				-0.12 - -0.05
				(0.000)
country FE	yes	yes	yes	yes
industry FE	no	no	no	no
Kleibergen-Paap	58.8739	40.3545	60.8726	63.9723
p-value	0.0000	0.0000	0.0000	0.0000
Hansen J	n.a.	n.a.	n.a.	n.a.
p-value	n.a.	n.a.	n.a.	n.a.
# obs.	37,596	20,083	39,917	60,000

Supplementary Table 11: Growth across different margins: 2SLS estimates - instrument: $V_{di}^{(I),iv}$ (country fixed effects)

VARIABLES	(1) $\Delta\log(\#\text{est.})$	(2) entry	(3) exit	(4) $\Delta\log(\#\text{est.}+1)$
log(TQI)	0.150*	0.066*	0.008	0.114*
	(0.020)	(0.007)	(0.004)	(0.017)
	0.11 - 0.19	0.05 - 0.08	-0.00 - 0.02	0.08 - 0.15
	(0.000)	(0.000)	(0.065)	(0.000)
log(#est. in 2011)	-0.044		-0.039*	
	(0.021)		(0.007)	
	-0.09 - -0.00		-0.05 - -0.03	
	(0.033)		(0.000)	
log(#est.+1 in 2011)				-0.091*
				(0.017)
				-0.13 - -0.06
				(0.000)
country FE	yes	yes	yes	yes
industry FE	no	no	no	no
Kleibergen-Paap	59.6528	38.0487	61.6796	63.4145
p-value	0.0000	0.0000	0.0000	0.0000
Hansen J	2.3533	2.4689	5.2639	0.5569
p-value	0.3083	0.2910	0.0719	0.7569
# obs.	37,596	20,083	39,917	60,000

Supplementary Table 12: Growth across different margins: 2SLS estimates - instruments: $V_{di}^{(II),iv}$, $iv \in \{1,2,3\}$ (country fixed effects)

VARIABLES	(1) $\Delta\log(\#\text{est.})$	(2) entry	(3) exit	(4) $\Delta\log(\#\text{est.}+1)$
log(TQI)	0.130*	0.052*	0.007	0.101*
	(0.019)	(0.005)	(0.004)	(0.015)
	0.09 - 0.17	0.04 - 0.06	-0.00 - 0.01	0.07 - 0.13
	(0.000)	(0.000)	(0.052)	(0.000)
log(#est. in 2011)	-0.033		-0.039*	
	(0.020)		(0.006)	
	-0.07 - 0.01		-0.05 - -0.03	
	(0.102)		(0.000)	
log(#est.+1 in 2011)				-0.084*
				(0.016)
				-0.12 - -0.05
				(0.000)
country FE	yes	yes	yes	yes
industry FE	no	no	no	no
Kleibergen-Paap	60.0721	40.4654	62.1349	65.4926
p-value	0.0000	0.0000	0.0000	0.0000
Hansen J	3.9973	7.8852	4.9895	1.9221
p-value	0.2618	0.0484	0.1726	0.5887
# obs.	37,596	20,083	39,917	60,000

Supplementary Table 13: Growth across different margins: 2SLS estimates - instruments: $V_{di}^{(III),iv}$, $iv \in \{0, 1, 2, 3\}$. (country fixed effects)

The most likely explanation is that our estimates on incoming business travel by industry may be somewhat noisy. This is particularly plausible given that we *impute* these business travel flows. Such mismeasurement would bias OLS estimates towards zero (*32*). Moreover, it would explain why adding industry fixed effects, as in Supplementary Table 4, weakens the OLS effects further: these fixed effects would absorb part of the signal in our business travel variable, leading to a deterioration of the effective signal-to-noise ratio (*33*).

VARIABLES	(1) $\Delta \log(\#\text{est.})$	(2) entry	(3) exit	(4) $\Delta \log(\#\text{est.} + 1)$
log(TQI)	0.112*	0.039*	0.004	0.086*
	(0.016)	(0.004)	(0.003)	(0.013)
	0.08 - 0.14	0.03 - 0.05	-0.00 - 0.01	0.06 - 0.11
	(0.000)	(0.000)	(0.236)	(0.000)
log(#est. in 2011)	-0.024		-0.037*	
	(0.019)		(0.006)	
	-0.06 - 0.01		-0.05 - -0.02	
	(0.212)		(0.000)	
log(#est.+1 in 2011)				-0.076*
				(0.015)
				-0.11 - -0.05
				(0.000)
country FE	yes	yes	yes	yes
industry FE	no	no	no	no
R^2	0.305	0.109	0.301	0.287
# obs.	37,596	20,083	39,917	60,000
F-stat	59.6	103.6	20.7	22.3
d.o.f. F-stat	(2, 95)	(1, 91)	(2, 95)	(2, 95)
Prob. > F	0.000	0.000	0.000	0.000

Supplementary Table 14: OLS estimates (country fixed effects)

Including such industry fixed effects in the 2SLS regressions (Tables 15 - 17) yields causal effect estimates that are even stronger than the ones reported in Tables 11 - 13. However, note that these fixed effects severely weaken the instruments (as evidenced by the reduction in the Kleibergen-Paap statistic, which has a χ^2 distribution with degrees of freedom equal to the number of instruments, i.e., 1 for $V_{di}^{(I)}$, 3 for $V_{di}^{(II)}$ and 4 for $V_{di}^{(III)}$), substantially widening confidence intervals. This increases the risk of running into issues related to weak instruments, which typically bias point estimates upward (*34*). In contrast, dropping country fixed effects (Tables 18 - 20) yields results that are fairly similar to the ones in Tables 15 - 17, albeit somewhat less stable across specifications and with the Hansen J test occasionally rejecting exogeneity of the instruments at any conventional level of statistical significance.

VARIABLES	(1) $\Delta \log(\#\text{est.})$	(2) entry	(3) exit	(4) $\Delta \log(\#\text{est.} + 1)$
log(TQI)	0.257	0.117	0.021	0.164
	(0.175)	(0.046)	(0.048)	(0.097)
	-0.09 - 0.61	0.03 - 0.21	-0.07 - 0.12	-0.03 - 0.36
	(0.147)	(0.012)	(0.656)	(0.096)
log(#est. in 2011)	-0.045		-0.038*	
	(0.023)		(0.006)	
	-0.09 - 0.00		-0.05 - -0.03	
	(0.054)		(0.000)	
log(#est.+1 in 2011)				-0.096*
				(0.018)
				-0.13 - -0.06
				(0.000)
country FE	yes	yes	yes	yes
industry FE	yes	yes	yes	yes
Kleibergen-Paap	15.8551	20.2596	15.5550	21.5368
p-value	0.0001	0.0000	0.0001	0.0000
Hansen J	n.a.	n.a.	n.a.	n.a.
p-value	n.a.	n.a.	n.a.	n.a.
# obs.	37,596	20,076	39,917	60,000

Supplementary Table 15: Growth across different margins: 2SLS estimates - instrument: $V_{di}^{(I),iv}$ (country and industry fixed effects)

A potential concern with our identification strategy is that the visa regimes a country of destination can negotiate could depend on its growth patterns. Such a dependence of visa regimes on future growth prospects would violate the instrument's exclusion restriction. In principle, we believe that that such dependencies are unlikely to be industry-specific. If they are not, they are accounted for in the models that condition the analysis on country fixed effects. However, (*35*) show that, in the US, lobby groups that represent industries with an interest in hiring

VARIABLES	(1) $\Delta\log(\#\text{est.})$	(2) entry	(3) exit	(4) $\Delta\log(\#\text{est.}+1)$
log(TQI)	0.456	0.183	-0.077	0.180
	(0.198)	(0.102)	(0.069)	(0.231)
	0.06 - 0.85	-0.02 - 0.39	-0.21 - 0.06	-0.28 - 0.64
	(0.024)	(0.077)	(0.264)	(0.438)
log(#est. in 2011)	-0.052		-0.035*	
	(0.023)		(0.006)	
	-0.10 - -0.01		-0.05 - -0.02	
	(0.027)		(0.000)	
log(#est.+1 in 2011)				-0.097*
				(0.017)
				-0.13 - -0.06
				(0.000)
country FE	yes	yes	yes	yes
industry FE	yes	yes	yes	yes
Kleibergen-Paap	9.9926	14.6511	10.3527	14.5717
p-value	0.0186	0.0021	0.0158	0.0022
Hansen J	2.4807	2.6963	6.5125	1.5583
p-value	0.2893	0.2597	0.0385	0.4588
# obs.	37,596	20,076	39,917	60,000

Supplementary Table 16: Growth across different margins: 2SLS estimates - instruments: $V_{di}^{(II),iv}$, $iv \in \{1,2,3\}$ (country and industry fixed effects)

VARIABLES	(1) $\Delta \log(\#\text{est.})$	(2) entry	(3) exit	(4) $\Delta \log(\#\text{est.} + 1)$
log(TQI)	0.348	0.043	0.013	0.176
	(0.172)	(0.056)	(0.048)	(0.131)
	0.01 - 0.69	-0.07 - 0.15	-0.08 - 0.11	-0.09 - 0.44
	(0.045)	(0.449)	(0.790)	(0.184)
log(#est. in 2011)	-0.048		-0.037*	
	(0.023)		(0.006)	
	-0.09 - -0.00		-0.05 - -0.02	
	(0.038)		(0.000)	
log(#est.+1 in 2011)				-0.097*
				(0.017)
				-0.13 - -0.06
				(0.000)
country FE	yes	yes	yes	yes
industry FE	yes	yes	yes	yes
Kleibergen-Paap	14.7300	14.4968	15.2205	17.9117
p-value	0.0053	0.0059	0.0043	0.0013
Hansen J	2.9715	2.9624	6.0538	2.6410
p-value	0.3960	0.3975	0.1090	0.4503
# obs.	37,596	20,076	39,917	60,000

Supplementary Table 17: Growth across different margins: 2SLS estimates - instruments: $V_{di}^{(III),iv}$, $iv \in \{0, 1, 2, 3\}$ (country and industry fixed effects)

VARIABLES	(1) $\Delta\log(\#\text{est.})$	(2) entry	(3) exit	(4) $\Delta\log(\#\text{est.}+1)$
log(TQI)	0.179*	0.045*	-0.016	0.125*
	(0.038)	(0.007)	(0.008)	(0.025)
	0.10 - 0.25	0.03 - 0.06	-0.03 - 0.00	0.08 - 0.17
	(0.000)	(0.000)	(0.051)	(0.000)
log(#est. in 2011)	-0.118*		-0.010	
	(0.034)		(0.006)	
	-0.18 - -0.05		-0.02 - 0.00	
	(0.001)		(0.092)	
log(#est.+1 in 2011)				-0.137*
				(0.024)
				-0.19 - -0.09
				(0.000)
country FE	no	no	no	no
industry FE	no	no	no	no
Kleibergen-Paap	38.2298	32.1679	39.4260	49.9051
p-value	0.0000	0.0000	0.0000	0.0000
Hansen J	n.a.	n.a.	n.a.	n.a.
p-value	n.a.	n.a.	n.a.	n.a.
# obs.	37,596	20,083	39,917	60,000

Supplementary Table 18: Growth across different margins: 2SLS estimates - instrument: $V_{di}^{(I),iv}$ (no fixed effects)

VARIABLES	(1) $\Delta\log(\#\text{est.})$	(2) entry	(3) exit	(4) $\Delta\log(\#\text{est.}+1)$
log(TQI)	0.122	0.012	-0.013	0.038
	(0.057)	(0.010)	(0.013)	(0.037)
	0.01 - 0.23	-0.01 - 0.03	-0.04 - 0.01	-0.04 - 0.11
	(0.036)	(0.226)	(0.293)	(0.318)
log(#est. in 2011)	-0.078		-0.011	
	(0.047)		(0.009)	
	-0.17 - 0.02		-0.03 - 0.01	
	(0.104)		(0.209)	
log(#est.+1 in 2011)				-0.067
				(0.033)
				-0.13 - -0.00
				(0.044)
country FE	no	no	no	no
industry FE	no	no	no	no
Kleibergen-Paap	22.6347	20.3075	24.0044	29.1297
p-value	0.0000	0.0001	0.0000	0.0000
Hansen J	3.9023	21.3419	9.0375	11.8039
p-value	0.1421	0.0000	0.0109	0.0027
# obs.	37,596	20,083	39,917	60,000

Supplementary Table 19: Growth across different margins: 2SLS estimates - instruments: $V_{di}^{(II),iv}$, $iv \in \{1,2,3\}$ (no fixed effects)

VARIABLES	(1) $\Delta\log(\#\text{est.})$	(2) entry	(3) exit	(4) $\Delta\log(\#\text{est.}+1)$
log(TQI)	0.107	0.019	0.001	0.049
	(0.043)	(0.007)	(0.006)	(0.027)
	0.02 - 0.19	0.00 - 0.03	-0.01 - 0.01	-0.01 - 0.10
	(0.015)	(0.013)	(0.919)	(0.077)
log(#est. in 2011)	-0.067		-0.021*	
	(0.038)		(0.006)	
	-0.14 - 0.01		-0.03 - -0.01	
	(0.080)		(0.001)	
log(#est.+1 in 2011)				-0.075
				(0.025)
				-0.13 - -0.03
				(0.004)
country FE	no	no	no	no
industry FE	no	no	no	no
Kleibergen-Paap	38.3826	24.9453	40.3951	45.4125
p-value	0.0000	0.0001	0.0000	0.0000
Hansen J	3.3325	13.3073	10.0306	8.7837
p-value	0.3431	0.0040	0.0183	0.0323
# obs.	37,596	20,083	39,917	60,000

Supplementary Table 20: Growth across different margins: 2SLS estimates - instruments: $V_{di}^{(III),iv}$, $iv \in \{0,1,2,3\}$ (no fixed effects) migrant workers have managed to loosen US policy on work permits. If lobbyists also try to relax requirements for visitor visas, the exclusion restriction of our instruments might be violated. However, visa regimes rarely change. As mentioned before, there were only 13 changes recorded Between 2009 and 2014, the five-year period for which the DEMIG database reports such changes that is closest to the one in which we measure growth. Therefore, if an industry's future growth prospects in a country of destination were to affect visa regimes in that country, then we would expect that the industry's past growth should have affected current visa regimes as well.

To test this, we can add the pretrend of our dependent variable, i.e., the logarithm of the industry's growth rate in the destination country between 2006 and 2011, as an explanatory variable. Unfortunately, the D&B data were only available for the period 2011-2016. Consequently, we cannot construct these pretrends in our analysis of industry growth. Instead, we turn to the trade data, for which a longer timeseries is available. Results are shown in Tables 21 and 22. To avoid that the sample changes due to observations for which past growth rates are undefined being dropped, we set logarithms to zero whenever the past growth rate is infinite or zero. The former repeats the 2SLS analysis of Supplementary Table 11 with trade data instead of industry data, whereas the latter reports the results when including the pretrend variable.

The analysis first of all shows that the causal effects of business travel on growth also replicate in trade data. Second, the addition of a pretrend does not lead to any major changes in point estimates. In fact, in the first three columns, changes are observable only in the point estimate's third decimal, if at all. In the final column, point estimates are reduced, but only by 8%, which is well within the margin of error.

Taken together, the instrumental variables estimations in this section suggest that business travel has a robust causal effect on the growth of economic activities in the destination country. However, we believe that this effect should not be regarded as separate from the effect of the

VARIABLES	(1) $\Delta\log(\text{exprts})$	(2) entry	(3) exit	(4) $\Delta\log(\#\text{exprts}+1)$
log(TQI)	0.261*	0.052*	-0.009	0.475*
	(0.020)	(0.006)	(0.003)	(0.034)
	0.22 - 0.30	0.04 - 0.06	-0.02 - -0.00	0.41 - 0.54
	(0.000)	(0.000)	(0.002)	(0.000)
log(exprts in 2011)	-0.289*		-0.011*	
	(0.016)		(0.001)	
	-0.32 - -0.26		-0.01 - -0.01	
	(0.000)		(0.000)	
log(exprts+1 in 2011)				-0.388*
				(0.014)
				-0.41 - -0.36
				(0.000)
country FE	yes	yes	yes	yes
product FE	no	no	no	no
Kleibergen-Paap	70.7207	19.1442	71.2718	58.9522
p-value	0.0000	0.0000	0.0000	0.0000
Hansen J	n.a.	n.a.	n.a.	n.a.
p-value	n.a.	n.a.	n.a.	n.a.
# obs.	59,777	7,911	62,422	70,333

Statistics: Point estimate, (standard errors), 95% two-sided confidence interval, (p-value)
Clustered (cntry, prod) standard errors: * p<0.001

Supplementary Table 21: Growth across different margins: 2SLS estimates - instrument: $V_{di}^{(I),iv}$ (country fixed effects, trade data)

VARIABLES	(1) $\Delta\log(\text{exprts})$	(2) entry	(3) exit	(4) $\Delta\log(\#\text{exprts}+1)$
log(TQI)	0.261*	0.051*	-0.008	0.438*
	(0.019)	(0.006)	(0.003)	(0.032)
	0.22 - 0.30	0.04 - 0.06	-0.01 - -0.00	0.38 - 0.50
	(0.000)	(0.000)	(0.003)	(0.000)
log(exprts in 2011)	-0.362*		0.022*	
	(0.020)		(0.002)	
	-0.40 - -0.32		0.02 - 0.03	
	(0.000)		(0.000)	
log(exprts+1 in 2011)				-0.667*
				(0.015)
				-0.70 - -0.64
				(0.000)
pretrend (06-11)	0.076*	-0.076*	-0.034*	0.339*
	(0.011)	(0.008)	(0.002)	(0.018)
	0.05 - 0.10	-0.09 - -0.06	-0.04 - -0.03	0.30 - 0.37
	(0.000)	(0.000)	(0.000)	(0.000)
country FE	yes	yes	yes	yes
product FE	no	no	no	no
Kleibergen-Paap	70.6874	19.2723	71.1957	58.6914
p-value	0.0000	0.0000	0.0000	0.0000
Hansen J	n.a.	n.a.	n.a.	n.a.
p-value	n.a.	n.a.	n.a.	n.a.
# obs.	59,777	7,911	62,422	70,333

Statistics: Point estimate, (standard errors), 95% two-sided confidence interval, (p-value)
Clustered (cntry, prod) standard errors: * $p<0.001$

Supplementary Table 22: Growth across different margins: 2SLS estimates - instrument: $V_{di}^{(I),iv}$ (country fixed effects, trade data with pretrend) trade, FDI or equity linkages that business travelers help establish and maintain. On the contrary, we submit that the causal effect of business travel operates *through* these channels. To test the validity of this claim, Supplementary Table 23 shows how business travel affects FDI, equity and trade links, using the visa-regime dummies as instruments. Note that these regressions are run at the bilateral level, i.e., at the level of origin-destination, not destination-industry combinations. As controls, we use country-of-origin and country-of-destination fixed effects.

	(1)	(2)	(3)
VARIABLES	log(FDI+1)	log(eqty+1)	log(trade+1)
log(BT+1)	0.899*	0.446*	1.109*
	(0.113)	(0.076)	(0.295)
	0.67 - 1.12	0.30 - 0.60	0.52 - 1.69
	(0.000)	(0.000)	(0.000)
country o FE	yes	yes	yes
country d FE	yes	yes	yes
Kleibergen-Paap	24.4404	24.4404	24.4404
p-value	0.0000	0.0000	0.0000
Hansen J	0.6612	6.4536	5.1276
p-value	0.7185	0.0397	0.0770
# obs.	9,120	9,120	9,120

Supplementary Table 23: Causal effect of business travel on economic linkages: 2SLS estimates

The 2SLS estimates suggest that business travel has a strong causal effect on all three types of economic linkages between countries. The effects are again somewhat larger than the OLS effects (Supplementary Table 24). Moreover, they suggest an about one-for-one relation between business travel and FDI and equity links: doubling business travels leads to more or less doubling the intensity of these links. In fact, the confidence intervals contain a value of unity, suggesting that business travelers are not so much a cause of, but rather an *input* to produce such links, in the same way that capital and labor are inputs in the production of regular output. This supports our claim that business travel should not be seen as a regular cause of economic growth, but rather as an enabling cause. For trade, the relation with business travel is weaker,

VARIABLES	(1) log(FDI+1)	(2) log(eqty+1)	(3) log(trade+1)
log(BT+1)	0.694*	0.415*	0.216
	(0.032)	(0.028)	(0.077)
	0.63 - 0.76	0.36 - 0.47	0.06 - 0.37
	(0.000)	(0.000)	(0.006)
country o FE	yes	yes	yes
country d FE	yes	yes	yes
R^2	0.729	0.781	0.682
# obs.	9,120	9,120	9,120
F-stat	63.1	34.8	14.2
d.o.f. F-stat	(1, 95)	(1, 95)	(1, 95)
Prob. > F	0.000	0.000	0.000

Supplementary Table 24: Correlational effect of business travel on economic linkages: OLS estimates

suggesting that doubling the number of business travelers leads to an about 50% increase in trade between countries.

S8 Effects of outgoing business travel: transposing the flow matrix

In the main text, we relate economic growth to inward as opposed to outward business travel. The reason is that we believe that, possibly with the exception of trade, the channels through which business travel operates (see the previous section) are more likely to be accompanied by knowledge flows with than against the direction of travel. For instance, if business travelers visit a country to establish new branch plants (as in FDI) or to visit existing ones (as in equity links), it is more likely that knowledge flows from the headquarters to these subsidiaries than in the reverse direction. However, business travel may still take place from subsidiaries to headquarters and some knowledge may flow against the direction of travel. In this section, we therefore replicate the paper's core findings, using the transpose of the bilateral business travel matrix.

We start by replicating Table 1 of the main text, which reports how business travel affects the growth of industries in a country of destination. This table is reproduced in Supplementary Table 25. Supplementary Table 26 repeats the analysis, but now focuses on the growth of industries in the country of origin. Note that, because some countries that receive business travelers do not send any business travelers abroad themselves, the number of observations is somewhat smaller in the bottom panel. The first two models fail to find a statistically significant effect at any conventional level of outgoing business travel. However, in the preferred specification, where we add country-of-origin and country-of-destination fixed effects, results in both panels are very similar, well within each other's confidence intervals.

Tables 27 and 28 repeat this exercise using trade data. Once again, results are fairly similar, albeit somewhat weaker when focusing on the growth of exports in the country of origin instead of the country of destination.

Finally, we study whether the relation between business travel flows and other bilateral linkages depends on the direction of travel. To do so, we replicate Supplementary Table 2 of SI, once using the actual business flows, and once reversing these flows. Results are reported in Tables 29 and 30. Note that, because some of the bilateral linkage types are symmetric, models (1), (6) and (7) yield identical results in both panels, with origin and destination GDP variables reversed. As before, the effects of business travel on economic linkages (FDI, trade and equity links) are similar, but slightly more pronounced when measured in the direction of the flow rather than in the opposite direction.

To summarize, the results in this section suggest that business travel flows, regardless of their direction, are associated with the growth of economic activities within a country and with important economic linkages between countries. That business travel in either direction predicts growth and economic linkages is in part due to the fact that there is some degree of symmetry

VARIABLES	(1) Δ log(#est.)	(2) Δ log(#est.)	(3) Δ log(#est.)
log(TQI)	0.059*	0.066	0.060
	(0.016)	(0.028)	(0.021)
	0.03 - 0.09	0.01 - 0.12	0.02 - 0.10
	(0.001)	(0.023)	(0.005)
log(#est. in 2011)	-0.033	-0.051	-0.039
	(0.018)	(0.021)	(0.020)
	-0.07 - 0.00	-0.09 - -0.01	-0.08 - 0.00
	(0.071)	(0.017)	(0.055)
log(GDP/cap)		-0.090	
		(0.050)	
		-0.19 - 0.01	
		(0.078)	
log(pop)		0.080	
		(0.030)	
		0.02 - 0.14	
		(0.010)	
log(#est. in ind.)		0.058	
		(0.031)	
		-0.00 - 0.12	
		(0.060)	
country FE	no	no	yes
industry FE	no	no	yes
R^2	0.014	0.062	0.341
# obs.	37,596	37,596	37,596
F-stat	8.9	16.0	4.6
d.o.f. F-stat	(2, 95)	(5, 95)	(2, 95)
Prob. $>$ F	0.000	0.000	0.012

Supplementary Table 25: Growth regressions. Replication of Table 1 in the main text.

VARIABLES	(1) Δ log(#est.)	(2) Δ log(#est.)	(3) Δ log(#est.)
log(TQI)	-0.003	0.005	0.071*
	(0.013)	(0.015)	(0.019)
	-0.03 - 0.02	-0.02 - 0.03	0.03 - 0.11
	(0.826)	(0.748)	(0.000)
log(#est. in 2011)	0.010	-0.042	-0.041
	(0.016)	(0.021)	(0.021)
	-0.02 - 0.04	-0.08 - -0.00	-0.08 - 0.00
	(0.533)	(0.049)	(0.052)
log(GDP/cap)		-0.036	
		(0.041)	
		-0.12 - 0.05	
		(0.384)	
log(pop)		0.116	
		(0.034)	
		0.05 - 0.18	
		(0.001)	
log(#est. in ind.)		0.114*	
		(0.022)	
		0.07 - 0.16	
		(0.000)	
country FE	no	no	yes
industry FE	no	no	yes
R^2	0.000	0.055	0.344
# obs.	36,145	36,145	36,145
F-stat	0.3	13.2	7.0
d.o.f. F-stat	(2, 87)	(5, 87)	(2, 87)
Prob. > F	0.776	0.000	0.002

Statistics: Point estimate, (standard errors), 95% two-sided confidence interval, (p-value)
Standard errors two-way clustered by country and industry: * p<0.001

Supplementary Table 26: Growth regressions. Replication of Table 1, using transposed business travel matrix to calculate TQI.

VARIABLES	(1) Δ log(exprts)	(2) Δ log(exprts)	(3) Δ log(exprts)
log(TQI)	0.160*	0.150*	0.321*
	(0.016)	(0.029)	(0.047)
	0.13 - 0.19	0.09 - 0.21	0.23 - 0.41
	(0.000)	(0.000)	(0.000)
log(exprts in 2011)	-0.233*	-0.239*	-0.305*
	(0.015)	(0.014)	(0.015)
	-0.26 - -0.20	-0.27 - -0.21	-0.33 - -0.27
	(0.000)	(0.000)	(0.000)
log(GDP/cap)		-0.002	
		(0.041)	
		-0.08 - 0.08	
		(0.962)	
log(pop)		0.042	
		(0.028)	
		-0.01 - 0.10	
		(0.141)	
log(global exprts in prod.)		0.036	
		(0.030)	
		-0.02 - 0.10	
		(0.225)	
country FE	no	no	yes
product FE	no	no	yes
R^2	0.101	0.103	0.204
# obs.	59,777	59,777	59,776
F-stat	125.8	57.8	202.6
d.o.f. F-stat	(2, 95)	(5, 95)	(2, 95)
Prob. $>$ F	0.000	0.000	0.000

Supplementary Table 27: Export growth regressions. Replication of Supplementary Table 9 in SI.

VARIABLES	(1) Δ log(exprts)	(2) Δ log(exprts)	(3) Δ log(exprts)
log(TQI)	0.085*	0.078*	0.196*
	(0.012)	(0.020)	(0.032)
	0.06 - 0.11	0.04 - 0.12	0.13 - 0.26
	(0.000)	(0.000)	(0.000)
log(exprts in 2011)	-0.191*	-0.221*	-0.293*
	(0.016)	(0.014)	(0.015)
	-0.22 - -0.16	-0.25 - -0.19	-0.32 - -0.26
	(0.000)	(0.000)	(0.000)
log(GDP/cap)		0.007	
		(0.045)	
		-0.08 - 0.10	
		(0.875)	
log(pop)		0.085	
		(0.027)	
		0.03 - 0.14	
		(0.003)	
log(global exprts in prod.)		0.101*	
		(0.025)	
		0.05 - 0.15	
		(0.000)	
country FE	no	no	yes
product FE	no	no	yes
R^2	0.084	0.096	0.197
# obs.	56,437	56,437	56,436
F-stat	78.8	53.2	194.0
d.o.f. F-stat	(2, 87)	(5, 87)	(2, 87)
Prob. > F	0.000	0.000	0.000

Statistics: Point estimate, (standard errors), 95% two-sided confidence interval, (p-value)
Standard errors two-way clustered by country and product: * p<0.001

Supplementary Table 28: Export growth regressions. Replication of Supplementary Table 9, using transposed business travel matrix to calculate TQI.

VARIABLES	(1) distance	(2) FDI	(3) ownership	(4) trade	(5) migration	(6) language	(7) colonizer	(8) all
log(GDP origin)	1.145* (0.087) 0.97 - 1.32 (0.000)	0.600* (0.074) 0.45 - 0.75 (0.000)	0.730* (0.071) 0.59 - 0.87 (0.000)	0.769* (0.097) 0.58 - 0.96 (0.000)	0.987* (0.092) 0.80 - 1.17 (0.000)	1.151* (0.091) 0.97 - 1.33 (0.000)	1.159* (0.092) 0.98 - 1.34 (0.000)	0.442* (0.078) 0.29 - 0.60 (0.000)
log(GDP destination)	0.714* (0.044) 0.63 - 0.80 (0.000)	0.432* (0.047) 0.34 - 0.52 (0.000)	0.384* (0.044) 0.30 - 0.47 (0.000)	0.388* (0.049) 0.29 - 0.48 (0.000)	0.471* (0.050) 0.37 - 0.57 (0.000)	0.720* (0.048) 0.62 - 0.82 (0.000)	0.727* (0.048) 0.63 - 0.82 (0.000)	0.176* (0.050) 0.08 - 0.27 (0.001)
log(km. dist.)	-1.372* (0.094) -1.56 - -1.19 (0.000)							-0.547* (0.099) -0.74 - -0.35 (0.000)
log(FDI)		0.623* (0.036) 0.55 - 0.70 (0.000)						0.272* (0.031) 0.21 - 0.33 (0.000)
log(ownership)			1.105* (0.082) 0.94 - 1.27 (0.000)					0.615* (0.077) 0.46 - 0.77 (0.000)
log(trade)				0.259* (0.029) 0.20 - 0.32 (0.000)				0.129* (0.020) 0.09 - 0.17 (0.000)
log(migrants)					0.286* (0.027) 0.23 - 0.34 (0.000)			0.072 (0.022) 0.03 - 0.12 (0.001)
common language						1.213* (0.225) 0.77 - 1.66 (0.000)		0.120 (0.168) -0.21 - 0.45 (0.478)
common col. hist.							0.947* (0.238) 0.47 - 1.42 (0.000)	-0.234 (0.163) -0.56 - 0.09 (0.153)
origin FE	no	no	no	no	no	no	no	no
destination FE	no	no	no	no	no	no	no	no
R^2	0.607	0.641	0.654	0.559	0.560	0.510	0.501	0.739
# obs.	12,656	12,656	12,656	12,656	12,656	12,656	12,656	12,656
F-stat	189.8	349.2	344.7	115.2	238.6	143.8	127.1	304.9
d.o.f. F-stat	(3, 112)	(3, 112)	(3, 112)	(3, 112)	(3, 112)	(3, 112)	(3, 112)	(9, 112)
Prob. $>$ F	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Supplementary Table 29: Explaining business travel. Replication of Supplementary Table 2 in SI.

VARIABLES	(1) distance	(2) FDI	(3) ownership	(4) trade	(5) migration	(6) language	(7) colonizer	(8) all
log(GDP origin)	0.714*	0.352*	0.362*	0.354*	0.498*	0.720*	0.727*	0.182*
	(0.044)	(0.045)	(0.046)	(0.049)	(0.047)	(0.048)	(0.048)	(0.050)
	0.63 - 0.80	0.26 - 0.44	0.27 - 0.45	0.26 - 0.45	0.40 - 0.59	0.62 - 0.82	0.63 - 0.82	0.08 - 0.28
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log(GDP destination)	1.145*	0.958*	0.866*	0.833*	0.815*	1.151*	1.159*	0.629*
	(0.087)	(0.094)	(0.086)	(0.093)	(0.084)	(0.091)	(0.092)	(0.085)
	0.97 - 1.32	0.77 - 1.14	0.70 - 1.04	0.65 - 1.02	0.65 - 0.98	0.97 - 1.33	0.98 - 1.34	0.46 - 0.80
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log(km. dist.)	-1.372*							-0.587*
	(0.094)							(0.102)
	-1.56 - -1.19							-0.79 - -0.39
	(0.000)							(0.000)
log(FDI)		0.415*						0.028
		(0.032)						(0.033)
		0.35 - 0.48						-0.04 - 0.09
		(0.000)						(0.402)
log(ownership)			0.938*					0.581*
			(0.066)					(0.066)
			0.81 - 1.07					0.45 - 0.71
			(0.000)					(0.000)
log(trade)				0.248*				0.113*
				(0.028)				(0.019)
				0.19 - 0.30				0.08 - 0.15
				(0.000)				(0.000)
log(migrants)					0.389*			0.215*
					(0.033)			(0.035)
					0.32 - 0.45			0.14 - 0.29
					(0.000)			(0.000)
common language						1.213*		0.115
						(0.225)		(0.182)
						0.77 - 1.66		-0.25 - 0.47
						(0.000)		(0.529)
common col. hist.							0.947*	-0.115
							(0.238)	(0.190)
							0.47 - 1.42	-0.49 - 0.26
							(0.000)	(0.547)
origin FE	no	no	no	no	no	no	no	no
destination FE	no	no	no	no	no	no	no	no
R^2	0.607	0.561	0.610	0.553	0.614	0.510	0.501	0.711
# obs.	12,656	12,656	12,656	12,656	12,656	12,656	12,656	12,656
F-stat	189.8	262.7	283.7	115.9	256.7	143.8	127.1	295.7
d.o.f. F-stat	(3, 112)	(3, 112)	(3, 112)	(3, 112)	(3, 112)	(3, 112)	(3, 112)	(9, 112)
Prob. $>$ F	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Statistics: Point estimate, (standard errors), 95% two-sided confidence interval, (p-value)
Standard errors two-way clustered by country of origin and country of destination: * $p<0.001$

Supplementary Table 30: Explaining business travel. Replication of Supplementary Table 2, using transposed business travel matrix to calculate TQI.

in the business-travel-flow matrix: the Spearman rank correlation between BT_{od} and BT_{do} is 0.726 ± 0.009 ($p < 0.001$, $N = 12,656$). Moreover, business travelers may also help know-how and know-who to diffuse from the destination to the origin of their travels. Overall however, the evidence in the growth and economic linkages regressions is somewhat stronger for the direction with compared to against the flow of travel.

S9 Travel Intensity Rankings

In the main text, we report rankings of contributors and recipients of global know-how flows. In this section, we explain how these rankings were constructed. It is important to note that the rankings are based on simplifying assumptions that allow us to attach a value to the know-how we have hypothesized to be embedded in business travel flows. In essence, we predict by how much larger the employment in an industry in a country is due to business travel inflows and then weight this predicted growth with an estimate of a national industry's productivity per worker.

To do so, we repeat the regression of model (3) in Tab. 1 of the main text. However, this time we do not measure an industry's size in terms of its number of establishments, but rather in terms of D&B's estimate of its number of employees. Moreover, we predict, not the growth of the industry, but rather its size in 2016:

$$\log E_{ic,2016} = \alpha + \beta \log TQI_{ic} + \gamma \log E_{ic,2011} + \eta_i + \delta_c,$$

where $\log E_{ic,t}$ is the logarithm of the employment of industry i in country c in year t and η_i and δ_c are industry and country fixed effects. Note that this model non-parametrically accounts for country and industry specific factors in the prediction. In principle, we could use the causal estimates in Sec. S7 of SI. However, these causal estimates become quite imprecise in the specification that includes country and industry fixed effects. We therefore prefer to use the more precisely estimated, and also more conservative, OLS estimates of the main text in these predictions.

Using the estimated parameters, we can predict how many employees an industry will have at different levels of business travel inflows. In particular, we will be interested in three different predictions:

- Actual business travel inflows. The first prediction is based on the actual level of TQI in an industry and a country:

$$\hat{E}^{actual}_{di,2016} = e^{\hat{\alpha}+\hat{\beta}\log TQI_{id}+\hat{\gamma}\log E_{id,2011}+\hat{\eta}_i+\hat{\delta}_c}$$

- Benchmark business travel inflows. The second prediction is based on the level of business travel had business travel been randomly distributed given the global sizes of industries and population of countries. We define this benchmark business travel level as:

$$\widehat{TQI}_{id} = \frac{\sum\limits_{c\in C} E_{ci,2011}}{\sum\limits_{i'\in I,c'\in C} E_{c'i',2011}},$$

  where the first term represents the share of industry i in global employment and the second term the share of country d in global population. In other words, we are assuming that the likelihood that a business traveler will arrive in an industry in a country equals the multiplication of two independent probabilities: the probability that the traveler comes from a given industry (equal to the industry's share in global employment) and the probability that the traveler arrives in the country of destination (equal to the share of the global population that lives in the country). The predicted employment by industry and country under this scenario is:

$$\hat{E}^{bench}_{di,2016} = e^{\hat{\alpha}+\hat{\beta}\log \widehat{TQI}_{id}+\hat{\gamma}\log E_{id,2011}+\hat{\eta}_i+\hat{\delta}_c}$$

- Leave-one-out business travel inflows. The third prediction is based on the level of business travel had one country of origin refrained from sending business travelers. We define the leave-out prediction for country of origin o as:

$$\hat{E}^{ex\, o}_{di,2016} = e^{\hat{\alpha}+\hat{\beta}\log(TQI_{id}-TQI_{oid})+\hat{\gamma}\log E_{id,2011}+\hat{\eta}_i+\hat{\delta}_c},$$

where TQI_{oid} is the business travelers attributed to industry i that were send from country o to country d.

Because not all industries are equally productive, we need to weight the employment of industries by some valuation that can be summed across industries. To do so, we use information on the estimated value of establishments' total sales. Our version of the D&B data provide this information for the year 2016. However, outside the US, reported sales volumes often represent estimates and are not consistently provided for all establishments. Therefore, to calculate the average productivity per worker in an industry, we restrict ourselves to establishments in the US and assume that the estimated dollar value produced by the average employee in the industry in the US is representative – if not in absolute, at least in relative terms – of the industry's productivity in other countries. That is, we assume that the productivity *ratio* between two industries is roughly similar in different countries. To account for productivity differences across countries, we multiply this productivity by the ratio of a country's GDP per capita to US GDP per capita. That is, countries may be more or less productive than the US, but they will be uniformly so across all industries:

$$P_{di} = \frac{O_{i,US}}{E_{i,US}} \frac{\text{GDP}_d/\text{pop}_d}{\text{GDP}_{US}/\text{pop}_{US}},$$

where $O_{US,i}$ represents industry i's output in the US in 2016, measured as its dollar-value sales volume and $\text{GDP}_c/\text{pop}_c$ a country's GDP per capita. We use this productivity estimate to adjust our earlier employment predictions and sum the result across all industries to arrive at a predicted size of a country's economy, $\widehat{\$S}$, under different scenario's:

$$\begin{aligned}
\widehat{\$S}_d^{actual} &= \sum_{i \in I} \hat{E}_{di,2016}^{actual} P_{di} \\
\widehat{\$S}_d^{bench} &= \sum_{i \in I} \hat{E}_{di,2016}^{bench} P_{di} \\
\widehat{\$S}_d^{ex\,o} &= \sum_{i \in I} \hat{E}_{di,2016}^{ex\,o} P_{di}
\end{aligned}$$

The rankings in the main text are based on these quantities. The Incoming Know-how Index expresses how much larger a national economy d is predicted to be when using actual TQIs compared to the benchmark scenario:

$$IKI_d = \frac{\widehat{\$S}_d^{actual}}{\widehat{\$S}_d^{bench}} - 1$$

Similarly The Outgoing Know-how Index expresses how much smaller the world economy would have been, had country of origin o not send any business travelers abroad:

$$OKI_o = \frac{\sum_{d \neq o \in C} \left(\widehat{\$S}_d^{actual} - \widehat{\$S}_d^{ex\,o} \right)}{\sum_{d' \neq o \in C} \widehat{\$S}_{d'}^{actual}} - 1$$

Tab. 31 and 32 report full rankings of countries by their IKI_d and OKI_o indices. In the main text, these are depicted in Fig. 3. The top senders of know-how in the world are composed of industrialized countries like Germany, Canada, the US, the UK and the Republic of Korea. The top recipients are more mixed. They mostly consist of smaller economies close to the big know-how emitters, such as Austria, Ireland, Switzerland and Denmark. The first large know-how emitter in the list of recipients is the UK (rank 13), followed by France (rank 19) and Canada (rank 20).

Finally, the OKI_o can also be calculated for each country of destination separately:

$$OKI_{od} = \frac{\widehat{\$S}_d^{actual} - \widehat{\$S}_d^{ex\,o}}{\widehat{\$S}_d^{actual}} - 1$$

OKI_{od} can be thought of as an estimate of by what percentage the economy of country d would shrink if country o were to stop sending business travelers. In the main text, we depict this quantity for six different countries of origin in Fig. 4.

S9.1 Caveats

It is important to note that the OKI and IKI estimates should not be taken at face value. On the contrary, they crucially depend on the causal nature of the effects of business travel on industrial growth and on how accurately we have been able to estimate a national industry's productivity. For instance, in Sec. S7, we noted that business travel is likely to be an enabling cause, not an ultimate cause. This aligns well with the thought experiment behind the OKI, where we imagine all business travel emanating from a country to cease, but less so with the thought experiment behind the IKI, where we increase or decrease the business travel into a country to a benchmark level. However, absent accompanying motives for business travel, it is unlikely that increasing business travel will cause economic growth (although decreasing it may still hinder growth). Furthermore, productivity ratios of industries may differ from the ones measured in the U.S.. In that case, the change in GDP we attribute to business travel to countries that are similar to the U.S. will be more accurate than to countries that are not.

With these caveats in mind, the maps showing the geographical variation in OKI and IKI still help understand how to interpret what the estimated effects of business travel mean for different national economies. They illustrate that existing business travel leads to a wide range of effects across countries. For instance, some regions of the world are poorly integrated through business travel in the global economy and depend mostly on knowledge inflows from within the region.

c0	c1	c2	c3	c4	c5	c6	c7	c8
01:	Austria	44.3%	33:	Bulgaria	-3.4%	65:	Ecuador	-26.3%
02:	Ireland	38.6%	34:	Jamaica	-5.1%	66:	Philippines	-26.3%
03:	Switzerland	37.0%	35:	Uruguay	-7.4%	67:	Azerbaijan	-27.1%
04:	Denmark	33.3%	36:	Lebanon	-7.9%	68:	Viet Nam	-27.1%
05:	Belgium	28.7%	37:	Serbia	-8.2%	69:	Republic of Moldova	-27.3%
06:	Hong Kong	28.4%	38:	Taiwan	-9.6%	70:	Kenya	-27.8%
07:	Singapore	27.8%	39:	Malaysia	-9.8%	71:	Sri Lanka	-27.9%
08:	Netherlands	27.3%	40:	South Africa	-10.1%	72:	Egypt	-28.6%
09:	Norway	25.6%	41:	Republic of Korea	-10.3%	73:	Papua New Guinea	-31.7%
10:	United Arab Emirates	22.1%	42:	Chile	-12.2%	74:	Senegal	-32.3%
11:	Spain	21.4%	43:	Turkey	-12.4%	75:	China	-32.7%
12:	Finland	20.4%	44:	Argentina	-12.7%	76:	Indonesia	-33.0%
13:	United Kingdom	17.4%	45:	Georgia	-12.9%	77:	Bolivia	-33.9%
14:	Czechia	16.8%	46:	Thailand	-13.0%	78:	Venezuela	-34.0%
15:	Portugal	15.9%	47:	Japan	-13.5%	79:	Nicaragua	-34.0%
16:	Croatia	15.7%	48:	Bosnia and Herzegovina	-13.5%	80:	Dominican Republic	-35.4%
17:	Italy	13.7%	49:	Mexico	-18.1%	81:	India	-36.3%
18:	New Zealand	13.6%	50:	Russia	-18.3%	82:	El Salvador	-39.1%
19:	France	12.7%	51:	Kazakhstan	-18.9%	83:	Haiti	-39.2%
20:	Canada	9.9%	52:	Saudi Arabia	-19.3%	84:	Ghana	-39.9%
21:	Hungary	9.6%	53:	Peru	-19.5%	85:	Honduras	-41.8%
22:	Australia	9.6%	54:	Ukraine	-19.8%	86:	Angola	-42.5%
23:	Germany	9.4%	55:	Colombia	-20.0%	87:	Uzbekistan	-46.2%
24:	Israel	6.7%	56:	Paraguay	-20.1%	88:	Nigeria	-46.3%
25:	Slovakia	6.5%	57:	Brazil	-21.1%	89:	Côte d'Ivoire	-48.1%
26:	Greece	6.1%	58:	Jordan	-21.5%	90:	Rwanda	-49.0%
27:	Lithuania	3.9%	59:	Morocco	-21.8%	91:	Tajikistan	-49.0%
28:	Panama	2.2%	60:	Costa Rica	-22.1%	92:	Guatemala	-49.1%
29:	United States of America	2.0%	61:	Albania	-23.0%	93:	Zimbabwe	-50.1%
30:	Oman	1.7%	62:	Tunisia	-23.7%	94:	Pakistan	-51.3%
31:	Kuwait	-0.7%	63:	Belarus	-24.4%	95:	Bangladesh	-51.4%
32:	Poland	-1.6%	64:	Armenia	-25.1%	96:	Iraq	-57.2%

Supplementary Table 31: Incoming know-how index

On the outgoing side of the equation, our estimates show that some global hubs dispatch their business travelers to all corners of the world, whereas other advanced economies help countries with strong linguistic ties connect to the global body of knowledge. More generally, the maps and indices constructed in this section show that the combination of a skewed distribution of business travel and a skewed distribution of industrial know-how favors well-connected and often already wealthy economies. That is, rich countries are not only better connected in general, they tend to be especially well-connected to other *rich* countries, which typically offer access to the kind of knowledge that is needed to sustain the global economy's most productive industries.

c0	c1	c2	c3	c4	c5	c6	c7	c8
01:	Germany	4.824%	33:	Chile	0.034%	65:	Kuwait	0.002%
02:	Canada	1.229%	34:	Russia	0.019%	66:	Belarus	0.002%
03:	United States of America	1.068%	35:	Uruguay	0.019%	67:	Azerbaijan	0.001%
04:	United Kingdom	0.983%	36:	Czechia	0.019%	68:	Sri Lanka	0.001%
05:	Republic of Korea	0.945%	37:	Ireland	0.018%	69:	Jordan	0.001%
06:	France	0.622%	38:	Philippines	0.016%	70:	Bosnia and Herzegov-ina	0.001%
07:	Japan	0.489%	39:	Ecuador	0.013%	71:	Haiti	0.001%
08:	Denmark	0.464%	40:	Thailand	0.013%	72:	Papua New Guinea	0.001%
09:	New Zealand	0.357%	41:	Costa Rica	0.012%	73:	Nigeria	0.001%
10:	Taiwan	0.323%	42:	Jamaica	0.011%	74:	Republic of Moldova	0.001%
11:	Norway	0.322%	43:	Portugal	0.011%	75:	Pakistan	0.001%
12:	India	0.226%	44:	Greece	0.009%	76:	Tunisia	0.000%
13:	Italy	0.178%	45:	Indonesia	0.009%	77:	Armenia	0.000%
14:	Spain	0.130%	46:	Guatemala	0.009%	78:	Bolivia	0.000%
15:	Brazil	0.122%	47:	Honduras	0.009%	79:	Kenya	0.000%
16:	Singapore	0.121%	48:	Panama	0.009%	80:	Albania	0.000%
17:	China	0.116%	49:	Ukraine	0.008%	81:	Uzbekistan	0.000%
18:	Australia	0.087%	50:	Argentina	0.008%	82:	Senegal	0.000%
19:	Switzerland	0.081%	51:	Hungary	0.007%	83:	Kazakhstan	0.000%
20:	Israel	0.068%	52:	Lebanon	0.006%	84:	Dominican Republic	0.000%
21:	Mexico	0.067%	53:	El Salvador	0.006%	85:	Georgia	0.000%
22:	Turkey	0.062%	54:	Saudi Arabia	0.005%	86:	Venezuela	0.000%
23:	United Arab Emirates	0.061%	55:	Croatia	0.005%	86:	Morocco	0.000%
24:	Belgium	0.060%	56:	Nicaragua	0.004%	88:	Angola	0.000%
25:	Netherlands	0.057%	57:	Egypt	0.004%	88:	Ghana	0.000%
26:	South Africa	0.049%	58:	Lithuania	0.004%	88:	Zimbabwe	0.000%
27:	Hong Kong	0.047%	59:	Oman	0.004%	88:	Paraguay	0.000%
28:	Poland	0.041%	60:	Peru	0.004%	88:	Tajikistan	0.000%
29:	Colombia	0.039%	61:	Serbia	0.003%	88:	Iraq	0.000%
30:	Malaysia	0.039%	62:	Bulgaria	0.003%	88:	Côte d'Ivoire	0.000%
31:	Austria	0.038%	63:	Slovakia	0.003%	88:	Rwanda	0.000%
32:	Finland	0.037%	64:	Viet Nam	0.003%	88:	Bangladesh	0.000%

Supplementary Table 32: Outgoing know-how index

Supplementary References

1. IPUMS. U.S. census data for social, economic, and health research (https://usa.ipums.org/usa/) (2018). Accessed: 2018-03-01.

2. IPUMS. Harmonized international census data for social science and health research (https://international.ipums.org/international/) (2018). Accessed: 2018-03-01.

3. Statistics Bureau of Japan. 2015 population census (https://www.e-stat.go.jp/en/stat-search/files?page=1&toukei=00200521&tstat=000001080615) (2017). Accessed: 2018-03-01.

4. Eurostat. Labor force survey (http://ec.europa.eu/eurostat/web/lfs/data/database) (2018). Accessed: 2018-03-01.

5. OECD. Population and employment by main activity (https://stats.oecd.org/Index.aspx?DataSetCode=SNA_TABLE3#) (2017). Accessed: 2018-03-01.

6. World Bank. Enterprise surveys data (http://www.enterprisesurveys.org/data) (2017). Accessed: 2018-03-01.

7. Eurostat. Annual data on trips of eu residents (http://ec.europa.eu/eurostat/web/tourism/data/database) (2018). Accessed: 2018-03-01.

8. DGE, Banque de France. Enquete aupres de visiteurs venant de l'etranger (http://www.entreprises.gouv.fr/files/files/directions_services/etudes-et-statistiques/4p-DGE/2015-4p47-donnees-EVE.xlsx) (2015). Accessed: 2018-03-01.

9. Research, Statistics and Information Management. Tourism to israel (https://motwebmediastg01.blob.core.windows.net/nop-attachment/7566_Tourism_to_Israel_Statistical_Report-2015-eng.pdf) (2015). Accessed: 2018-03-01.

10. ISTAT. Tourist flow in italy (https://www.istat.it/en/archive/193010) (2016). Accessed: 2018-03-01.

11. JTB Tourism Research & Consulting Co. Japanese tourists statistics (https://www.tourism.jp/en/tourism-database/stats/) (2018). Accessed: 2018-03-01.

12. NBTC Holland Marketing. Inbound tourism (https://www.nbtc.nl/en/home/article/inbound-tourism-en.htm) (2016). Accessed: 2018-03-01.

13. Statistics South Africa. Tourism 2015 (http://www.statssa.gov.za/publications/Report-03-51-02/Report-03-51-022015.pdf) (2016). Accessed: 2018-03-01.

14. Office for National Statistics. Overseas travel and tourism, quarterly (https://www.ons.gov.uk/peoplepopulationandcommunity/leisureandtourism/datasets/overseastravelandtourism) (2018). Accessed: 2018-03-01.

15. National Travel and Tourism Office. 2014 U.S. resident travel: Business and convention travel (http://tinet.ita.doc.gov/outreachpages/download_data_table/2014-US-Business.pdf) (2015). Accessed: 2018-03-01.

16. UNWTO. All countries: Inbound tourism: Arrivals by main purpose 1995 - 2017 (01.2019). (2019).

17. Wright, D. H. & Reeves, J. H. On the meaning and measurement of nestedness of species assemblages. *Oecologia* **92**, 416–428 (1992).

18. Atmar, W. & Patterson, B. D. The measure of order and disorder in the distribution of species in fragmented habitat. *Oecologia* **96**, 373–382 (1993).

19. Bustos, S., Gomez, C., Hausmann, R. & Hidalgo, C. A. The dynamics of nestedness predicts the evolution of industrial ecosystems. *PloS one* **7**, e49393 (2012).

20. Pennacchioli, D., Coscia, M., Rinzivillo, S., Giannotti, F. & Pedreschi, D. The retail market as a complex system. *EPJ Data Science* **3**, 1 (2014).

21. United Nations. *World Population Prospects: The 2017 Revision, custom data acquired via website.* (Department of Economic and Social Affairs, Population Division, 2017).

22. Mayer, T. & Zignago, S. Notes on cepii's distances measures: The geodist database (2011).

23. DESA, U. Trends in international migrant stock: Migrants by destination and origin. *United Nations database, POP/DB/MIG/Stock/Rev* **2013** (2013).

24. Gaulier, G. & Zignago, S. *Baci: international trade database at the product-level (the 1994-2007 version)* (CEPII Working Paper 2010-23, 2010).

25. Feldman, B. E. Relative importance and value. *Available at SSRN 2255827* (2005).

26. Groemping, U. Relative importance for linear regression in r: The package relaimpo". *Journal of Statistical Software* **17**, 1–27 (2006).

27. Bahar, D., Hausmann, R. & Hidalgo, C. A. Neighbors and the evolution of the comparative advantage of nations: Evidence of international knowledge diffusion? *Journal of International Economics* **92**, 111–123 (2014).

28. Bahar, D. & Rapoport, H. Migration, knowledge diffusion and the comparative advantage of nations. *The Economic Journal* **0**, 1–33 (2018).

29. Hausmann, R. *et al.* *The atlas of economic complexity: Mapping paths to prosperity* (Mit Press, 2014).

30. De Haas, H., Natter, K. & Vezzoli, S. Compiling and coding migration policies: Insights from the demig policy database. *International Migration Institute, DEMIG project paper* **16**, 43 (2014).

31. De Haas, H., Natter, K. & Vezzoli, S. Conceptualizing and measuring migration policy change. *Comparative Migration Studies* **3**, 15 (2015).

32. Durbin, J. Errors in variables. *Revue de l'institut International de Statistique* 23–32 (1954).

33. Griliches, Z. & Hausman, J. A. Errors in variables in panel data. *Journal of econometrics* **31**, 93–118 (1986).

34. Stock, J. H. & Yogo, M. Testing for weak instruments in linear iv regression (2002).

35. Facchini, G., Mayda, A. M. & Mishra, P. Do interest groups affect us immigration policy? *Journal of International Economics* **85**, 114–128 (2011).