Japan & The World Economy 64 (2022) 101162

Contents lists available at ScienceDirect

Japan & The World Economy

journal homepage: www.elsevier.com/locate/jwe

Transportation infrastructure and trade☆

Han Zheng[a,b], Li Hongtao[c,*]

[a] Bay Area International Business School, Beijing Normal University, No.18 Jinfeng Road, Tangjiawan, Zhuhai City, Guangdong Province 51-9087, China
[b] Hitotsubashi Institute for Advanced Study, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo 186-8601, Japan
[c] Graduate School of Public Policy, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

ARTICLE INFO

JEL classification:
F10
F14

Keywords:
Gravity model
Transportation infrastructure
Internal trade cost

ABSTRACT

This paper offers a variant of the Ricardian model able to structurally interpret the estimate of country specific variable—transportation infrastructure. Guided by this new theoretical framework, this paper shows that transportation infrastructure enhances international trade more than internal trade. Further quantitative analysis suggests 10% increase in transportation infrastructure induces 3.9% increase in real income and more than 95% of the gains concentrate on the infrastructure improving country. This paper also suggests that transportation infrastructure improvement increases real income mostly through internal trade cost reduction. All the above results suggest that better infrastructure leads to sizable gains providing additional empirical support to policies aiming to improve transportation infrastructure.

1. Introduction

Job creations and investment payoff are two common pros for infrastructure investment discussed in economics literature while relatively few literature analyze how infrastructure affects trade flows. The scarcity of research is partly because of the nondiscriminatory feature of transportation infrastructure as a country specific variable, which precludes the identification of transportation infrastructure out of country specific fixed effect in the popular fixed effect gravity specification. However, this paper capitalizes on the methodology devised by Beverelli et al. (2018) and finds that the infrastructure boosts international trade more than internal trade. Further quantitative counterfactual exercise shows that 10% transportation infrastructure increase induces sizable 3.9% increase in real income. This provides solid support to President Biden's Bipartisan Infrastructure Law by arguing this Law can lead to higher trade openness.

It is commonly believed that a better infrastructure can lead to both lower international and internal trade costs, where the international trade refers to the trade flows crossing the country border while the internal trade is defined conversely, regardless of whether those trade flows are inter-industry or not. Further academic evidence confirms this intuition—Jaworski et al. (2022) find a significant impact of the US highways on both internal and international trade flows. However, whether infrastructure primarily facilitates international trade or internal trade is a question of significance, because this heterogeneous impact affects expenditure share on domestically produced goods which is one of the sufficient statistics on real income. To provide analysis for this question, this paper proposes a variant of the Eaton and Kortum (2002) framework, capable of identifying the country specific variable separately from the country specific fixed effect, to estimate how much infrastructure can affect the trade cost and the gains from trade. Contrary to an intuitive statement that infrastructure benefits internal trade more than international trade since domestically produced goods seem to have unfettered access to internal transportation infrastructure, the empirical results show that better infrastructure boosts international trade more than internal trade, rendering infrastructure a favorable policy instrument to effectively elevate the trade openness. Further quantitative analysis reinforces the aforementioned claim and finds infrastructure improvements can generate sizable gains and the infrastructure improving country reaps more than 95% of its gains.

Another noteworthy merit of this framework is that it can be readily applicable to assess the impact of other more convoluted country specific factors, e.g., institution quality, political party turnover, on both international trade and internal trade flows.

The internal trade cost is the same across countries is a convenient assumption imposed in international trade literature, whereas little evidence is shown to warrant this assumption. On the contrary, intuitions

☆ We wish to thank Kenichi Ueda, Taiji Furusawa, Yoichi Sugita, Hirokazu Ishise, Daisuke Fujii, and Konstantin Kucheryavyy for providing useful comments at the various stages of this paper and two anonymous referees for their constructive comments. All remaining errors and omissions are our sole responsibility. Han Zheng is supported by JSPS Grant-in-Aid for Early-Career Scientists (KAKENHI), Project Number JP22K13380. Li Hongtao is supported by Ministry of Education Culture Sports Science and Technology (MEXT), WINGS-GSDM Designing Future Society (DFS) Fellowship, Project Number 150201SP01.

* Corresponding author.
E-mail addresses: hanzheng211@gmail.com (Z. Han), li-hongtao@g.ecc.u-tokyo.ac.jp (H. Li).

https://doi.org/10.1016/j.japwor.2022.101162
Received 1 July 2022; Received in revised form 19 October 2022; Accepted 26 October 2022
Available online 31 October 2022

and both anecdotal and academic evidence hint at the opposite—transportation infrastructure and investment exhibits considerable variations across countries. Several recent studies also analyze the presence and implication of the internal trade cost, such as Allen and Arkolakis (2014), Ramondo et al. (2016), Kerem CoŞar and Fajgelbaum (2016), Atkin and Donaldson (2017) and Allen and Arkolakis (2022). Unlike the aforementioned studies, this paper attempts to estimate the impact of internal infrastructure on both international and internal trade flows using a popular fixed effect gravity specification. In addition, building on the empirical specification proposed by Beverelli et al. (2018), this paper provides a framework capable of structurally interpreting the estimate on country specific variables, such as infrastructure which affects both internal trade and international trade.

Suspecting that the infrastructure should affect trade cost across time and be motivated by the real-time infrastructure investment difference shown in Fig. 1, this paper builds a theoretical model, a variant of Eaton and Kortum (2002), to accommodate the fact that infrastructure affects both the international trade cost and the internal trade cost. The theoretical model yields a novel welfare formula predicting that the change in the internal trade cost works as a multiplier, amplifying the traditional gains from trade. This modification is not only quantitative important but also qualitative crucial, because the quantitative analysis section shows that there are sizable gains in real income even when domestic goods expenditure share increases. More importantly, this formula provides a theoretical foundation to structurally estimate a country specific internal trade cost component. Based on the new framework, this paper offers a structural interpretation of the coefficient in front of the infrastructure measure. Using an estimation method developed by Beverelli et al. (2018), this paper finds that a better infrastructure boosts international trade flow more than internal trade flow, a result, though odd at first sight, which suggests that there might be an economy of scale in the domestic distribution, that is exports and imports have to go through several major ports or nodes for clearance, whereas internal goods flow has no such requirements. Another possible explanation is that countries provide more favorable shipping policies to exports and imports to foster international trade. Additionally, this result also implies that, other than the two channels through which people commonly perceive infrastructure investments to yield benefits, creating jobs and investment returns, infrastructure investment has the potential to levitate trade openness, which in turn yields higher gains from trade.

The most related literature which study the effect of infrastructure on international trade using the gravity model are the following. Feenstra and Ma (2014) examine how the port efficiency, part of transportation infrastructure, affects the extensive margin of trade and confirm a significant positive impact. A new index of infrastructure is created in Donaubauer et al. (2016a), and it is used to study various issues in Donaubauer et al. (2016b, 2015), and Rehman et al. (2020). Using the new index of infrastructure, Donaubauer et al. (2015) find a country's endowment with overall infrastructure to be positively associated with more intensive trade relations and the effects of infrastructure prove to be non-linear. Similar to his work, We further investigate the effects of infrastructure on international and internal trade and results show that better infrastructure boosts international trade more than internal trade. Rehman et al. (2020) show that infrastructure positively improves exports while negatively affecting the trade deficit. They use fully modified least squares (FMOLS) and dynamic ordinary least square (DOLS) estimator to deal with endogeneity issues by adding the leads and lags, while we use the Poisson Pseudo Maximum Likelihood Estimator for Structural Gravity Models, whose solutions define the exporter and importer specific multilateral resistance terms.

Admittedly, this paper is by no means the first paper to discuss the internal trade cost and the domestic infrastructure. Other than the studies mentioned before, Donaldson and Hornbeck (2016) and Donaldson (2018) analyze, respectively, the impact of the US railroad networks and the railroad construction project during the Indian colonial period, both of which find a significant, positive impact on welfare measures. Banerjee et al. (2020) find access to domestic transportation infrastructure has a positive causal effect on per capita GDP levels in China. Using detailed Chinese custom data, Fan et al. (2021) develop a route choice model and embed it into an input–output multi-sector trade model to evaluate the massive expressway construction program in China, finding positive welfare gain. Recently, using a newly assembled sub-country regional bilateral trade dataset, under a causal inference framework, Santamaria et al. (2020) unveil a significant Border effect, unexplained by usual geographical proxies, such as distance. As an intermediate measure to find how the theoretical cost of living measures altered by the variety of goods, Cavallo et al. (2021) discover that the internal trade cost constructed by the margins, captured by the transportation sector and the various taxes, is positively correlated with trade openness which frequently and positively correlates with the per capita income.[1]

The following section presents a theoretical model capable of dealing with factors affecting the internal trade cost and the international trade cost, simultaneously. Section 3 builds a roadmap from the theory to the empirical counterparts and offers a structural interpretation of the coefficient of the infrastructure proxy. Section 4 describes the data sources and briefly discusses their merits. Section 5 presents the estimation results and a discussion ensues. Section 6 considers several counterfactual situations and assesses both the qualitative and quantitative importance of the new welfare formula. Section 7 reports the robustness check. The final section concludes.

2. Theoretical framework

The section presents a theoretical framework as a variant of Eaton and Kortum (2002) (thereafter the EK model) with roundabout production. However, the qualitatively same results shown here can also be derived under the Anderson–Wincoop or Melitz–Chaney settings.[2]

As in Alvarez and Lucas (2007), the tradable sector produces intermediate goods. Each country sources each variety of intermediate goods from different countries and combines them into a CES composite used to produce both intermediate goods and final goods which is non-traded. The consumers derive utility from the consumption of competitively sold final goods. The production technology of each variety of intermediate goods is given as

$$q_i(\omega) = z_i(\omega) l_i(\omega)^{\beta} m_i(\omega)^{1-\beta}, \tag{1}$$

where $l_i(\omega)$ is the labor and $m_i(\omega)$ is the intermediate aggregate. $z_i(\omega)$ is drawn independently from Type-2 Gumbel distribution $F(z) = P(z_i(\omega) \leq z) = \exp(-T_i z^{-\theta})$. Because of perfect competition and a constant return to scale, the factory price is given as:

$$p_i(\omega) = \frac{w_i^{\beta} P_i^{1-\beta}}{z_i(\omega)\beta^{\beta}(1-\beta)^{1-\beta}} = \frac{c_i}{z_i(\omega)}, \tag{2}$$

[1] Many other noteworthy studies mainly analyze the relations between domestic infrastructure and the economic activities not confined to real income. For example, using the rich road network data for Peru, Volpe Martincus et al. (2017) confirm the positive effect from the expansion of road network on firms' export and job growth; Ma and Tang (2020) study the migration within China through the transportation infrastructure and discover the welfare impacts of migrations crucially hinge on the extensive margin of trade—the entry and exit of firms; Li and Ma (2021) find better domestic infrastructure contributes to the rapid spread of Covid-19 in the early stage of the global pandemic; Felbermayr and Tarasov (2022) emphasize the uneven distribution of transportation infrastructure inside countries arising from the endogenous infrastructure investment decisions could account for a significant portion of the border effect.

[2] Cavallo et al. (2021) arrive at a similar formula as in Eq. (8) with the Melitz–Chaney heterogeneous firm framework. The only difference between this model and theirs is that the goods variety holds to be constant.

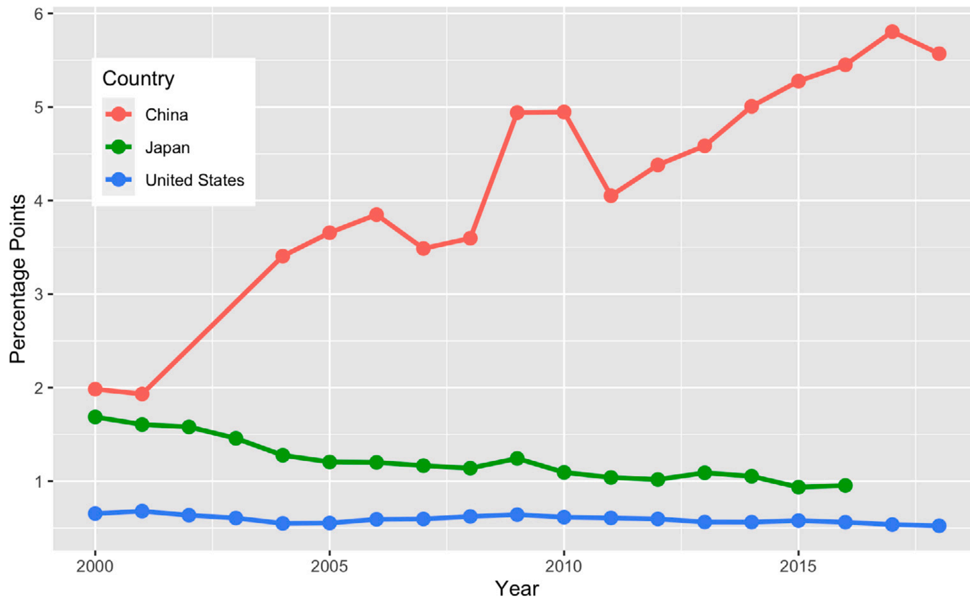

Fig. 1. Total inland transport infrastructure expenditure over GDP
Note: The data come from the OECD Database. The y-axis is the total inland transport infrastructure expenditure percentage points out of GDP.

where $c_i = \frac{w_i^\beta P_i^{1-\beta}}{\beta^\beta (1-\beta)^{1-\beta}} = \frac{w_i^\beta P_i^{1-\beta}}{\gamma}$. P_i is the price index of intermediate aggregate. Similar to the intermediate goods sector, the production technology of the final goods sector is deterministic Cobb–Douglas type meaning:

$$p_F = \frac{w_i^\beta P_i^{1-\beta}}{\beta^\beta (1-\beta)^{1-\beta}}. \tag{3}$$

The c.i.f price for any price is given as:

$$p_{ij}(\omega) = \begin{cases} p_i(\omega) t_{i.} \tau_{ij} t_{.j} & \forall i \neq j, \\ p_i(\omega) t_{ii} & i = j, \end{cases} \tag{4}$$

where τ_{ij} is the standard iceberg trade cost and $t_{.j}$ is the trade cost stemming from the importing country's infrastructure condition, which applies unanimously to the goods from every exporting country. Similarly $t_{i.}$ is the trade cost stemming from the exporting country's infrastructure condition. t_{ii} is the internal trade cost from this country's infrastructure. This country specific internal trade cost is reminiscent of what Han (2021) has convincingly shown that the internal trade cost at the country level is a function of country size and the trade cost between the basic geographical units defined in that paper. The basic geographical unit is a swathe of area within a country satisfying these three requirements: equal labor endowment, equal technology level, and equal price index. The heterogeneity in t_{ii} across countries reflects the innate differences in the trade costs between basic geographical units. Notice that the above formulation of prices across country pairs lies at the very core of this new extension of the EK model, because it allows a country specific infrastructure to affect both internal trade and international trade heterogeneously. This heterogeneous impact on internal trade and international trade might reflect the economy of scale in shipping or the nonlinear pricing phenomenon in shipping as highlighted in Han and Fujii (2020). Normally, the effect of country specific variables is precluded from an estimation even when only using the international trade flows in a fixed effect estimation approach, as the country specific variables will be perfectly collinear with the country specific fixed effects.[3] The same procedure as in EK yields the trade share of country j's expenditure on country i's goods as:[4]

$$\pi_{ij} = \begin{cases} \dfrac{T_i (c_i t_{i.} \tau_{ij} t_{.j})^{-\theta}}{\sum_{s \neq j} \left[T_s (c_s t_{s.} \tau_{sj} t_{.j})^{-\theta} \right] + T_j (c_j t_{jj})^{-\theta}} & \forall i \neq j, \\ \dfrac{T_i (c_i t_{ii})^{-\theta}}{\sum_{s \neq j} \left[T_s (c_s t_{s.} \tau_{sj} t_{.j})^{-\theta} \right] + T_j (c_j t_{jj})^{-\theta}} & i = j, \end{cases} \tag{5}$$

with price index P_i given as:

$$P_i = A \left\{ \sum_{i \neq j} \left[T_i \left(c_i t_{i.} \tau_{ij} t_{.j} \right)^{-\theta} \right] + T_j (c_j t_{jj})^{-\theta} \right\}^{-1/\theta}, \tag{6}$$

where $A = \left[\Gamma \left(\frac{\theta + 1 - \sigma}{\theta} \right) \right]^{1/(1-\sigma)}$, a constant of parameters. Eq. (5) immediately gives the trade flow from country i to country j as multiplication of trade share and country j's total expenditure:

$$X_{ij} = \pi_{ij} X_j \quad \forall i, j, \tag{7}$$

where X_j is country j's total expenditure. When $i = j$, similar to the definition in Ramondo et al. (2016), X_{jj} are the domestic absorption and also the internal trade flows including both inter-industry trade and intra-industry trade within one country's boundary. This variant incorporates two new components into the country's competitiveness measure $T_i \left(c_i t_{i.} \tau_{ij} t_{.j} \right)^{-\theta}$: $t_{i.}$ and $t_{.j}$, which function the same way as τ_{ij}—higher cost lower competitiveness, except both are country specific rather than pair specific. Some straightforward substitutions yield the

[3] Fixed effect estimation approach of the gravity model is preferred here because it is consistent with a broader class of trade models, including Melitz and Ottaviano (2008) and Han and Fujii (2020). However, several other versions of gravity specifications not using the fixed effect to control multilateral resistance terms permit analyzing the effect of country specific variables such as infrastructure, and institution quality. One notable example is Donaubauer et al. (2018) use a weighted average of all bilateral trade costs to control for multilateral resistance terms and embed it into a traditional gravity model with GDP to represent economy sizes. Their results show that higher infrastructure quality in both pair countries tends to increase trade flows between them. However, their approach might suffer from omitted variable biases because it could miss some country specific forces captured by country specific fixed effects but not able to be exhausted in their enumeration of variables included.

[4] Curious readers can refer to Eaton and Kortum (2002) or Dekle et al. (2008) for details in derivation.

following:[5]

$$\frac{w_j}{p_F} = B\left(\frac{T_j}{\pi_{jj} t_{jj}^{\theta}}\right)^{\frac{1-\beta}{\theta\beta}}, \tag{8}$$

where π_{jj} is also dubbed absorption rate in trade literature and B is a collection of constants including A, γ. Following Ramondo et al. (2016) and Han (2021), the technology parameter T_j can be sensibly parameterized as $T_j = \phi_j L_j$ where L_j is country j's labor endowment and ϕ_j is called innovation intensity measuring how creative on average each personnel can be in country j. Then Eq. (8) can be alternatively written as:

$$\frac{w_j}{p_F} = B\left(\frac{\phi_j L_j}{\pi_{jj} t_{jj}^{\theta}}\right)^{\frac{1-\beta}{\theta\beta}}. \tag{9}$$

Following Dekle et al. (2008), the hat algebra version of Eq. (8) is given as:[6]

$$\frac{\hat{w}_j}{\hat{p}_F} = \underbrace{\left(\frac{\hat{T}_j}{\hat{\pi}_{jj}}\right)^{\frac{1-\beta}{\theta\beta}}}_{\text{ACR formula}} \left(\hat{t}_{jj}\right)^{-\frac{1-\beta}{\beta}}. \tag{10}$$

Eq. (10) looks very similar to the famous ACR formula in Arkolakis et al. (2012) where the ACR formula only contains the first parenthesis in Eq. (10). However, it extends what Arkolakis et al. (2012) claim: the welfare measure not only depends on two sufficient statistics, trade elasticity and the absorption rate change, but also the internal trade cost change. A simple thought experiment would illustrate the intuition behind this equation. A negative shock devastates the infrastructure lifting up the trade cost to both internal trade and international trade rendering the absorption rate somewhat unchanged. The formula endorsed by Arkolakis et al. (2012) would predict a small welfare change or even higher welfare in case internal trade shrinks more than international trade proportionately, completely silencing the adverse impact of the negative shock. The internal trade cost change component in Eq. (10) will adjust the gains from international trade by the internal trade conditions.[7] More importantly, Eqs. (8) and (10) are useful to structurally estimate the country specific components of the internal trade cost, where the standard gravity estimation falls short. Additionally Eq. (10) is reminiscent of the theoretical framework in Ramondo et al. (2016) and Han (2021) where the regional level EK setting is aggregated into a country level EK setup. Their welfare formula is similar to Eq. (10), under the assumption that the regions are all symmetric. Meanwhile, Ramondo et al. (2016)'s results show that the symmetric assumption is not restrictive at all, reassuring the validity of Eq. (10).[8]

3. From theory to empirics

This section offers a road map from the theoretical framework to our empirical specifications. More importantly, it provides insights on how to structurally interpret the coefficients estimated and explains why the same coefficient appears regardless either side of the country specific infrastructure proxy is used in estimation.

Inspection on Eq. (5) reveals that the standard fixed effect gravity model is plagued by the econometric difficulty of the perfect collinearity. The country specific trade cost $t_{i\cdot}$ and $t_{\cdot j}$ would be absorbed into the country specific fixed effect. Fortunately, Beverelli et al. (2018) provide hints at how the country specific trade cost can be estimated indirectly using the internal trade flows. Built upon their results, the following derivation demonstrates how the effects can be estimated and their theory-consistent interpretation. The trade flow equations can be written as follows which comes from Eq. (7):

$$X_{ij} = \exp\left(\ln\left(T_i c_i^{-\theta} t_{i\cdot}^{-\theta}\right) - \ln\left(\sum_s \left[T_s \left(c_s \tau_{sj} t_{\cdot j}\right)^{-\theta}\right] t_{j\cdot}^{-\theta}\right) + \ln X_j + \ln\left(t_{j\cdot}\tau_{ij} t_{\cdot j}\right)^{-\theta}\right), \tag{11}$$

$$X_{jj} = \exp\left(\underbrace{\ln\left(T_j c_j^{-\theta} t_{j\cdot}^{-\theta}\right)}_{exp_j} \underbrace{- \ln\left(\sum_s \left[T_s \left(c_s \tau_{sj} t_{\cdot j}\right)^{-\theta}\right] t_{j\cdot}^{-\theta}\right) + \ln X_j}_{imp_j} + \ln\left(\tau_{jj} t_{jj}\right)^{-\theta}\right). \tag{12}$$

Notice that the first term of the above equation system is exporting country specific and the second and third terms are importing country specific, which respectively will be controlled for using exporter country fixed effect exp_i and importer country fixed effect imp_j in the cross-section empirical setting.[9] τ_{ij} is the standard iceberg trade cost able to be captured by bilateral trade proxies or country pair fixed effect and τ_{jj} is one as usual. Notice that $t_{i\cdot}^{-\theta}$ is an exporting country specific term which will be absorbed into i's exporter fixed effect as in Eq. (11). And because of the existence of exporter fixed effect in internal trade flow observations, the same i's exporter fixed effect appearing in its own internal trade observation contains an effect from $t_{j\cdot}^{-\theta}$ which should not appear in the internal trade flow observation as in Eq. (12). To balance this effect, the second term in Eq. (12) has to contain $t_{j\cdot}^{-\theta}$ which in turn will appear in the importer fixed effect in Eq. (11) because of the very same importer fixed effect in both international trade flow and internal trade flow observations and, as a result, the fourth term in Eq. (11) will have to contain both effects stemming from the importer's infrastructure $t_{j\cdot}^{-\theta}$ and $t_{\cdot j}^{-\theta}$.

To proceed, parametric assumptions are imposed on the country specific trade cost—they are power functions on the infrastructure

[5] The derivation procedure strictly follows Han (2021)'s cookbook steps.

[6] The detailed derivation is provided in Appendix D.

[7] Another way to show the comparative statics of $\hat{t}_{jj}$ on the real income change is to combine the direct effect from $\hat{t}_{jj}$ and indirect effect from $\hat{\pi}_{jj}$, as $\hat{t}_{jj}$ affects $\hat{\pi}_{jj}$ as well. It is straightforward to show that the net effect of $\hat{t}_{jj}$ on real income change is negative, meaning higher $\hat{t}_{jj}$ leads to a lower real income change. However, this simplistic discussion of the impact from $\hat{t}_{jj}$ on the real income change does not provide much valuable insight, because this paper allows infrastructure to affect not only the internal trade cost but the international trade cost as well, both of which affects the absorption rate. Therefore, a discussion of $\hat{t}_{jj}$ conditional on the absorption rate change serves the purpose of this paper better.

[8] One may argue that the ACR formula has already included the effects of internal trade cost change because its effects can be attributed to the change of technology as a residual. This kind of argument has a logic like national accounting in growth literature. However, this claim is not very well grounded for the following two reasons. First, careful inspection of Eq. (5) reveals that the technology parameter has partial trade elasticity one, whereas the internal trade cost parameter has partial trade elasticity θ meaning the variation on the internal trade cost parameter could be used to estimated trade elasticity. Second, though sounds like the logic used in national accounting, they are essentially different. National accounting reduces all the unknown elements affecting production to a technology parameter after netting out all the known production factors. However, the immediate attribution of change in the internal trade cost to the change in technology does not help in understanding what constitutes the technology parameter. As a matter of fact, Han (2021) shows that in a parsimonious trade model with roundabout production, after filtering out the effect of internal trade cost, the technology parameter can be well approximated by R&D data.

[9] This paper is using panel data, therefore the exp_i and imp_j fixed effects will be transformed into the exporter-time fixed effect $exp_{i,t}$ and the importer-time fixed effect $imp_{j,t}$ in this paper's empirical specifications.

proxy for expository simplicity:[10]

$$t_{jj} = INF_j^a, \quad t_{\cdot j} = INF_j^b, \quad t_{j\cdot} = INF_j^c. \tag{13}$$

Then the trade flow equations become:

$$\begin{aligned} X_{ij} &= \exp\left(exp_i + imp_j - \theta(b+c)\ln\left(INF_j\right) + BTP_{ij}\times\eta\right), \\ X_{jj} &= \exp\left(exp_i + imp_j - \theta a \ln(INF_j)\right), \end{aligned} \tag{14}$$

where BTP_{ij} is used to capture τ_{ij}. Notice that $\theta a \ln(INF_j)$ is importer specific and will be absorbed into imp_j. Therefore, in the specification in Eq. (15), what the coefficient in front of $\ln(INF_j)\times Border$ will capture is $\theta(b+c-a)$, namely the infrastructure's effect on the international flow relative to the internal flow:

$$X_{ij} = \exp\left(imp_i + exp_j + Border + \ln(INF_j)\times Border + BTP_{ij}\times\eta\right) \quad \forall i,j, \tag{15}$$

where $Border$ is a dummy variable taking values of 1 for the international trade flows and 0 for the internal trade flows. The inclusion of a $Border$ dummy follows Beverelli et al. (2018) and also Han (2021) shows compelling evidence that the border effect is an indispensable part of internal trade cost and failing to include it might lead to biased estimates.

An alternative way of reformulating the trade flow equation is given as follows:

$$\begin{aligned} X_{ij} &= \exp\left(\ln\left(T_i c_i^{-\theta} t_{\cdot i}^{\theta}\right) - \ln\left(\sum_s\left[T_s\left(c_s\tau_{sj}t_{\cdot j}\right)^{-\theta}\right]t_{\cdot j}^{\theta}\right)\right. \\ &\quad \left. + \ln X_j + \ln\left(t_{i\cdot}\tau_{ij}t_{\cdot i}\right)^{-\theta}\right), \\ X_{jj} &= \exp\left(\ln\left(T_j c_j^{-\theta} t_{\cdot j}^{\theta}\right) - \ln\left(\sum_s\left[T_s\left(c_s\tau_{sj}t_{\cdot j}\right)^{-\theta}\right]t_{\cdot j}^{\theta}\right)\right. \\ &\quad \left. + \ln X_j + \ln\left(\tau_{jj}t_{jj}\right)^{-\theta}\right). \end{aligned} \tag{16}$$

Similar to the above analysis, the fourth term in the above equation contains both effects from the exporter side $t_{i\cdot}$ and $t_{\cdot i}$ and its empirical counterpart is given as:

$$\begin{aligned} X_{ij} &= \exp\left(exp_i + imp_j - \theta(b+c)\ln\left(INF_i\right) + BTP_{ij}\times\eta\right), \\ X_{jj} &= \exp\left(exp_i + imp_j - \theta a \ln(INF_i)\right). \end{aligned} \tag{17}$$

$\theta a \ln(INF_i)$ will be absorbed into exp_i, leaving the coefficient of INF_i capturing the relative effect $\theta(b+c-a)$. The above analysis predestines that regardless of which side's measure is used, the estimate of the infrastructure will be the same and its interpretation will be the exporting and importing effects combined, relative to the internal effect, which is confirmed in the robust analysis section. This analysis confirms what Beverelli et al. (2018) discovered and their interpretation.

4. Data description and summary statistics

This section discusses the data sources and their merits, then followed by a summary of statistics of several variables of interest.

4.1. Data description

Trade Flows—The trade flow data is from the ITPD-E Database.[11] The ITPD-E contains consistent data on international and domestic trade for 243 countries, 170 industries, and 17 years. The time period covered by the dataset includes the years from 2000 until 2016, the most recent year for which data across all sectors are available. The dataset is constructed using reported administrative data and intentionally excludes information estimated using the statistical technique to preserve the authenticity of the data. The data are constructed at the industry level covering the broad sectors of agriculture, mining and energy, manufacturing, and services which means the ITPD-E describes nearly completely the traded sectors of each economy.

About the construction of the international trade data, the dataset uses importer-reported values which are considered more reliable than data on exports, for goods trade. Also, the authors of this dataset use exports reported by partner countries to fill the values whenever they are missing from the importer side. For service trade, it uses services trade data provided by the exporter side, since services export data are collected by mandatory surveys and are more accurate than from the importer side. Each observation in the original ITPD-E Database documents the bilateral trade value of one industry in one particular year. Therefore, this database has no information about how the imported goods are used meaning the imported materials should be the aggregate value used by all the industries. Therefore, the trade value should contain both inter-industry trade and intra-industry trade. Because the analysis in this paper is at the country level and to map the industry level trade value in the data to the country level model in the paper, the industrywise trade flows are aggregated to obtain country level trade flows through summing the trade values of all the industries for each bilateral country pair. Relative to the other trade flow database, the most outstanding advantage is that it contains internal trade flows whenever possible and its computation uses gross production rather than GDP as the reference sum, which is more consistent with the roundabout structure of production in reality. The database's construction does not rely on an imputation using the gravity method, therefore, the database is suitable for gravity estimation. The details of trade data used for quantitative analysis will be separately presented in Appendix F.

We compared ITPD-E with other trade and production datasets to highlight its merits. Compared with World Bank's TPP, ITPD-E covers data post 2004 and includes significantly more countries. Also, ITPD-E covers the complete economy while TPP only covers data from manufacturing industries. Compared with CEPII's TradeProd which covers manufacturing industries for over 150 countries from 1980 to 2006, ITPD-E covers a larger set of countries and more recent years. GTAP dataset covers the entire economies of 121 countries like ITPD-E. However, its data only covers 2004, 2007, 2011, and 2014 specific years. GTAP's reliance on statistical models to fill missing values renders it more appropriate for simulation than estimation. ITPD-E is not balanced and constructed from administrative data to preserve the authenticity of the raw data so that it is suitable for estimation. Compared with WIOD, which covers 56 industries from 2000 to 2014, ITPD-E covers more countries and wider industries. Similarly to the GTAP, WIOD heavily relies on economic models to fill the missing values. Therefore, compared to other frequently used trade datasets, ITPD-E generally covers wider countries and industries with a longer time span including more recent years. Its construction accesses the accurate administrative data making it suitable for estimation.

Bilateral Trade Proxies—The proxies come from the CEPII website, which covers bilateral distance, common language, border continuity

[10] The true components of t_{jj}, $t_{\cdot j}$ and $t_{j\cdot}$ could potentially include confounding differential trade cost terms other than infrastructure proxy. The existence of other differential trade cost terms introduces additional country specific border fixed effect in its empirical specification. However, when estimating using symmetric pair fixed effect to capture bilateral trade cost the method endorsed by Baier and Bergstrand (2007), the flexible specification of bilateral trade cost subsumes the country specific Border effect situation. That is one of the reasons symmetric pair fixed effect specification is favored when picking the preferred estimates in the later section.

[11] Its link is given as https://www.usitc.gov/data/gravity/itpde.htm. Its detailed procedure of construction and comparison with other frequently used trade databases are well documented in Borchert et al. (2021).

etc. over 225 countries. The merits of the distance proxies constructed there are twofold: first, it computes intra-national distances in the same manner as international distances; second, it offers several versions of bilateral distances, including population weighted distances taking into account all the cities in a typical country. Therefore, the distance component of the internal trade cost can be accounted for and the internal distance is also used to approximate the internal trade cost in several studies.[12]

Regional Trade Agreement—The regional trade agreement dummy comes from the Mario Larch's Regional Trade Agreements Database.[13] It covers 516 Regional Trade Agreement (RTA) notified to WTO from year 1950 to 2019. Dummies separating the type of trade agreements, such as currency union, free trade agreement, and economic integration agreement, are also included. Because the primary focus of this study is not the impact of various RTAs, we include only the RTA dummy to control for bilateral trade policies' induced impact.

Infrastructure Proxy—As the main variable of interest, several types of infrastructure variables are collected from various sources to ensure the robustness of our main result. First, we pick the sum of the total length of the railroad and the paved road as the infrastructure proxy.[14] It measures how extensive one country's distribution network is. Considering the innate heterogeneity of the country's territorial size, a transportation density measure is also constructed to check the robustness. The railroad length is from the United Nations' World Development Indicator Database and the road length is from the OECD Transport Database. Second, an alternative measure of infrastructure proxy is the index composite constructed in Donaubauer et al. (2016a), where an unobserved component model is employed in order to structurally clump together multiple highly correlated infrastructure measures. Donaubauer et al. (2016a) categorize those infrastructure measures into four board categories—transportation, ICT, energy, and financial infrastructure and assemble indices for each of these categories and a total index. In Appendix A, we experiment with each of these sub-indices and find the robustness of our empirical results.

Institution Control—Numerous studies have analyzed the impact of institutional quality on bilateral trade flows and trade patterns.[15] Despite their vastness, their main message is simple and clear—better institution in the pair countries generally leads to higher trade flows. More importantly, Beverelli et al. (2018), the most similar paper to ours in terms of methodology concludes that institution has a differential impact on international trade flow and internal trade flow. In light of those literature, in the robustness analysis, we decide to include institutional variables as a control to obviate the omitted variable bias. Following Beverelli et al. (2018), we use the institutional quality indices from the World Bank's World Governance Indicators which contain the following categories: Voice and Accountability, Political Stability and Absence of Violence, Government Effectiveness, Regulatory Quality, Rule of Law and Control of Corruption. However, a structurally composite indicator over those categories is not available. As a compromise, we use the simple arithmetic average as the aggregate index.

Innovation Intensity—Han (2021) has convincingly shown in a model with heterogeneous internal trade cost, the R&D data can approximate innovation intensity quite well. Drawing upon this insight, we use the number of researchers per million population from the World Bank's World Development Indicator database as the innovation intensity proxy. This dataset documents the annual number of researchers engaged in certain activities, such as developing concepts and theories, and creating software of operational methods. Though this database has another innovation intensity measure—percentage point of expenditure on R&D activities out of GDP, the number of researchers is adopted as it is more consistent with the parameterization of $T_i = \phi_i L_i$.

Other Variables—All the other variables are from Penn World Table 9.1 such as real income, GDP, GDP per capita, population, and employment. Its coverage spans from year 1950 to 2017 and its construction and features are well documented in Feenstra et al. (2015). The preferred real income measure is real GDP estimated from the expenditure approach suitable for comparison across countries and years.

4.2. Summary statistics

The following Table 1 provides descriptive statistics of interesting variables used in our regression study. Given the trade flows are measured in million US dollars, the max value of trade flow seems overly large. However, this abnormally large number exactly testifies to the validity of this dataset, because this dataset contains internal trade flows and large countries generally absorb the majority of productions domestically. Moreover, the internal trade flows are computed using the value of gross production rather than GDP as reference sum, which further enlarges the size of internal trade flows. The observation that attains this highest trade flow on record is the internal trade of the US in 2008. The infrastructure proxy in length has a relatively low number of observations because the OECD Transport Database only provides paved road data for OECD countries and partner countries. However, we prefer this proxy when we quantitatively interpret the results, because the other infrastructure index does not permit a structural interpretation consistent with the theory provided. Both the infrastructure index and the WGI institution index have its mean around 0 and the standard deviation 1, because both are an index composite clumping several measures together using the unobserved components model method. This method generally assumes the underlying true variables across countries follow the standard normal distribution. The negative values of those proxies prevent a theoretically consistent interpretation of their coefficients, however, the signs of their coefficients have qualitative meanings. As for innovation intensity, there is a huge disparity among countries with the lowest record in this sample from Lesotho and the highest record from Denmark.

5. Empirical findings

This section first presents the empirical results in augmented gravity estimation motivated by Eq. (15) and establishes the result that infrastructure boosts international trade more than internal trade. Then comes the analysis that estimates the elasticity of infrastructure on internal trade, which shows better infrastructure is conducive to higher internal trade flows. A discussion of better infrastructure's impact on real income shows that though the traditional gains from trade decrease according to the ACR formula, the real income increases because Eq. (10) shows that the effect from absorption rate change has to be adjusted by the change of internal trade cost.

[12] See Yotov (2012), Agnosteva et al. (2019), Ramondo et al. (2016), Beverelli et al. (2018) and Heid et al. (2021).

[13] It can be downloadable from the following website https://www.ewf.uni-bayreuth.de/en/research/rta-data/index.html.

[14] Though the railroad and road transportation are substitutes for goods delivery, the sum of both transportation modes more accurately represents the country's total transportation investment intensity than either component, under the assumption that a country would invest in both modes until their marginal benefits equalize.

[15] See Anderson and Marcouiller (2002), Levchenko (2007), Tang (2012), Manova (2013), Francois and Manchin (2013), Nunn and Trefler (2014), Beverelli et al. (2018) and Demir and Hu (2020). Samadi and Alipourian (2021) has suggested several indicators in the literature to measure the quality of institutions from different dimensions. We select World Bank's Worldwide Governance Indicators, WGI for short, as a candidate to measure the quality of institutions. Beverelli et al. (2018) also choose WGI as one of the options.

Table 1
Descriptive statistics.

	Count	Mean	St. Dev.	Min	Max
Trade flows (Million US$)	566570	1997.801	104128.3	0	18844254
Infrastructure length (km)	486	457003	1209203	3320.798	7155777
Infrastructure index	1477	.0090378	1.005249	−1.760078	3.216129
Institution WGI	3011	−.0338481	.9144237	−2.449376	1.969566
Innovation intensity	1340	2014.32	1957	5.91	8065.89

Note: This table only includes variables of interest. The second row corresponds to the infrastructure measure constructed by summing railroad length and paved road length. The third row is the index constructed in Donaubauer et al. (2016a). The fourth row is the institutional index available from World Bank's World Governance Indicator database. The fifth row is the number of researchers engaged in certain activities per million people from the World Bank's World Development Indicator database.

5.1. Augmented gravity estimation

Notice that though the previous analysis establishes that the coefficient in front of the infrastructure proxy captures the relative effect of the export side and import side combined over the internal trade cost, it does not preordain that the effect should be positive or negative, or even significant. Either way could happen even if the infrastructure has a positive impact on trade flow is well established. The exact empirical model is given as follows:

$$\log X_{ij,t} = \alpha \times Border + \rho \times INF_{j,t} \times Border + \eta \times BTP_{ij} + imp_{i,t} + exp_{j,t} + \epsilon_{ij,t} \quad \forall i, j, \tag{18}$$

where $X_{ij,t}$ is the bilateral trade flow at time t; $Border$ is the border dummy taking value 1 for international trade flows, to signify any bilateral trade barrier either tangible or intangible because of the traversing of the country borders; $INF_{j,t} \times Border$, as elaborated in Section 3, measures how much the infrastructure improvement enhances the international trade relative to the internal trade with positive ρ meaning the infrastructure improvement facilitates the international trade more than the internal trade; BTP_{ij} stands for the standard bilateral trade cost proxies used in gravity literature such as distance, common language *etc.*; $imp_{i,t}$ is the importer-year fixed effect to control all the importer specific factors; $exp_{j,t}$ is the exporter-year fixed effect to control all the exporter specific factors; $\epsilon_{ij,t}$ is the error term.

The benchmark results of the above empirical model using the total length of the railroad and paved road as infrastructure proxy are tabulated in Table 2 and all specifications are estimated with both importer-time and exporter-time fixed effects implemented. The first column offers the results from the standard log-linear gravity estimation. Almost all the estimates on the bilateral trade proxies lie within the one standard error range from the estimates in the meta analysis of Head and Mayer (2014). Column (2) confirms the validity and reliability of the data set and the PPML estimates located within a reasonable range from the ones in Santos Silva and Tenreyro (2006). The third and fourth column estimate Eq. (18) using bilateral trade proxies with different methods and both of them confirm a positive and significant effect of the infrastructure measure comparable to the size of the RTA across different specifications. The border dummy capturing the average unexplained internal trade cost exerts a considerable impact on the trade flows. Column (5) shows the result with a symmetric country pair fixed effect using the PPML estimate. The effect of the infrastructure is still positive under a specification so rich in fixed effects and its impact is stable across all the specifications, ranging from 0.25 to 0.65. Contrary to the assumption Cavallo et al. (2021) made that the domestic trade costs impact unanimously on domestically produced goods and imported goods, Column 3 to 5 suggest the domestic trade costs, at least for the portion proxied by the combination of road length and railroad length, are very likely to affect goods discriminatively. Because the symmetric country pair fixed effects specification is flexible enough to subsume country specific border effect specification as mentioned in Footnote 10 and able to capture all the observed or even unobserved components of symmetric trade cost, an approach advocated by Baier and Bergstrand (2007). In addition, estimates of PPML automatically satisfy the add-up constraints: $\sum_i \hat{X}_{ij} = \sum_i X_{ij} = E_j$, $\sum_j \hat{X}_{ij} = \sum_j X_{ij} = Y_i$, according to Fally (2015). The estimate in column (5) is picked as our preferred estimate. Having said that, however, viewing Eq. (18) as a reduced form estimation might suggest an upward bias of the estimate on the infrastructure interaction term because of the reverse causality. For example, higher trade generally leads to higher welfare and rich countries already have much better domestic market access meaning the infrastructure built is for the purpose of promoting international market access. Therefore, failing to control for the relative effect of the size and welfare status of the country might bias the estimates upward, even though the country-time fixed effect has controlled the unanimous effect. So we conduct the exercises taking care of the size of countries in Appendix B and find robust estimates using infrastructure indices. Also reassuringly even smaller estimates do not affect the main implications in the quantification analysis as shown in the working paper version of this paper.

Because of the relatively low observations of total length infrastructure proxy, we corroborate our primary results using the infrastructure index constructed in Donaubauer et al. (2016a), which is tabulated in Table 3. Generally, the results are preserved when using the infrastructure index as the proxy instead. The estimated coefficients are positive across most specifications, though in the PPML estimation with symmetric pair fixed effect it is not precisely estimated. Another noteworthy point is that the coefficients of the infrastructure index under the OLS estimation method in column 1 are similar to the ones in Donaubauer et al. (2018) using the same infrastructure index in a traditional gravity model specification. This similarity cross-validates the validity of the results shown here. Further robustness check using sub-indices of the infrastructure index is offered in Appendix A.

Therefore, we can safely conclude that better infrastructure tend to affect trade flows heterogeneously—it boosts international trade flows more than internal trade flows, which answers the question posited by Coughlin and Novy (2021) whether infrastructure primarily facilitates within-border trade or cross-border trade.[16] Several economic mechanisms could potentially be behind this observation. First, generally more infrastructure stocks are dedicated to facilitating internal trade than international trade reflecting the general assumption that internal trade cost is one with international trade cost bigger than one. If the marginal infrastructure improvement exhibits diminishing return, given the same degree of infrastructure improvement, internal trade

[16] On the contrary, Fan et al. (2021) generate opposite result—infrastructure facilitate internal trade flows more than international flows. The methodology used and the entity of focus could be the cause of different results. Fan et al. (2021) exploit the cross-time variation of the expressway expansion in one single country China, whereas this study utilizes the cross-country variations to estimate the transport infrastructure to trade cost elasticity. Therefore, their results are not expected to be applicable to other countries, because of the idiosyncrasies of China. Despite those differences, reassuringly, both studies yield similar quantitative results, which will be discussed in Section 6.

Table 2
Augmented gravity estimations with infrastructure.

	(1) Inter_OLS	(2) Inter_PPML	(3) INF_OLS	(4) INF_PPML	(5) INF_Pair_PPML
INF × Border			0.556***	0.276***	0.650***
			(0.124)	(0.039)	(0.146)
Distance	−1.570***	−0.742***	−1.277***	−0.571***	
	(0.016)	(0.030)	(0.053)	(0.042)	
Contiguity	0.717***	0.300***	0.218	0.455***	
	(0.088)	(0.059)	(0.150)	(0.080)	
Language	0.731***	0.161**	0.393***	0.486***	
	(0.033)	(0.065)	(0.090)	(0.087)	
Colony	0.902***	0.161**	0.949***	−0.034	
	(0.085)	(0.082)	(0.119)	(0.100)	
RTA	0.515***	0.304***	0.418***	0.309***	0.251***
	(0.027)	(0.055)	(0.069)	(0.084)	(0.040)
Border			−9.744***	−7.399***	
			(1.559)	(0.573)	
Constants	13.751***	15.223***	15.738***	18.783***	12.741***
	(0.146)	(0.272)	(0.399)	(0.269)	(0.364)
Observations	456066	629336	95319	106838	106397
R^2	0.734		0.818		
Pseudo R^2		0.946		0.992	0.998
Country pair FE	No	No	No	No	Yes
Importer-Time FE	Yes	Yes	Yes	Yes	Yes
Exporter-Time FE	Yes	Yes	Yes	Yes	Yes

Note: Standard errors clustered at country pair level are reported in parenthesis.
The dependent variable is the bilateral trade flows in PPML estimation and log trade flows in OLS estimations. The total length of the railroad and the paved road is selected as the infrastructure proxy. INF is short for infrastructure proxy and FE is short for fixed effect.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3
Augmented gravity estimations with infrastructure index.

	(1) INF_OLS	(2) INF_PPML	(3) INF_Pair_OLS	(4) INF_Pair_PPML
INF Index × Border	1.577***	0.455***	0.493***	−0.033
	(0.199)	(0.043)	(0.187)	(0.076)
Distance	−1.444***	−0.575***		
	(0.020)	(0.049)		
Contiguity	0.843***	0.426***		
	(0.093)	(0.076)		
Language	0.754***	0.480***		
	(0.038)	(0.064)		
Colony	0.865***	−0.094		
	(0.092)	(0.097)		
RTA	0.482***	0.128*	0.116***	−0.097
	(0.033)	(0.073)	(0.036)	(0.068)
Border	−4.134***	−4.337***		
	(0.189)	(0.184)		
Constants	16.768***	18.436***	0.610***	14.001***
	(0.219)	(0.309)	(0.033)	(0.027)
Observations	237560	313243	236472	305287
R^2	0.749		0.864	
Pseudo R^2		0.989		0.998
Country pair FE	NO	NO	Yes	Yes
Importer-Time FE	Yes	Yes	Yes	Yes
Exporter-Time FE	Yes	Yes	Yes	Yes

Note: Standard errors clustered at country pair level are reported in parenthesis.
The dependent variable is the bilateral trade flows in PPML estimation and log trade flows in OLS estimations. The infrastructure index is selected as the infrastructure proxy. INF is short for infrastructure and FE is short for fixed effect.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

flows will respond less than international trade flows. This could cause the phenomenon that marginal infrastructure improvement facilitates more international trade flows than internal trade flows. Second, it is well known that there exists the economy of scale at the port level, see *e.g.,* Clark et al. (2004), whose the economy of scale story could similarly hold in shipping because exports and imports need to be transported to several major ports and nodes for declaration or clearance. This would tend to create more international trade flows than internal trade flows given the same level of infrastructure improvement and reinforce the heterogeneous effects discussed above. The exact cause and economic mechanism are, unfortunately, out of the realm of this paper and can be a very promising venue for future research.

An additional exercise of replacing the importer side infrastructure measure with the exporter side infrastructure measure is also performed as the robustness analysis, confirming the same coefficient obtained as claimed in Section 3. This duality presents researcher options to maximize the information used in the trade flow data, that is the importer side infrastructure can be used when the countries with infrastructure data have more accurate and more observations.

5.2. Elasticity on internal trade

It turns out to be quite challenging to obtain infrastructure's elasticity directly from the trade flows, because the other elements that constitute internal trade cost are unobserved. As a theory consistent roundabout solution, we capitalize on Eq. (9) and draw on the result in Han (2021) that after controlling for internal trade cost term, the innovation technology term ϕ_i can be well approximated by R&D data. The logic of recovering process is similar to the indirect least squares, but different from the indirect least squares, the structural empirical equation directly comes from the theory. The log transformation of Eq. (9) is given as follows:

$$\log\left(\frac{w_j}{p_F}\right) + \frac{1-\beta}{\theta\beta}\log(\pi_{jj}) = C + \frac{1-\beta}{\theta\beta}\log(\phi_j) + \frac{1-\beta}{\theta\beta}\log(L_j) - \frac{1-\beta}{\beta}\log(t_{jj}), \tag{19}$$

where C is a constant composite. As indicated in Footnote 10, it is very likely other factors other than the infrastructure condition affect the internal trade cost. Therefore, this equation motivates the following empirical specification:

$$\log(RealGDP_{j,t}) + \frac{1-\beta}{\theta\beta}\log(\pi_{jj,t}) = m \times R\&D_{j,t} + n \times Employ_{j,t} + r \times INF_{j,t} + \tilde{t}_{jj} + \epsilon_{j,t}, \tag{20}$$

where $RealGDP_{j,t}$ is the real GDP data from the Penn World Table 9.1; $\pi_{jj,t}$ is the domestic goods expenditure share also called absorption rate calculated from the ITPD-E data; $R\&D_{j,t}$ is the number of researchers engaged in certain activities per million people from World Bank's World Development Indicator database and $Employ_{j,t}$ is the number of employment from the Penn World Table 9.1; $INF_{j,t}$ is our variable of interest and the infrastructure proxy used in this paper; $\tilde{t}_{jj}$ represents other factors of the internal trade cost which will be controlled by country fixed effect. Infrastructure as an element of internal trade cost does not preclude other unobserved factors from influencing the internal trade cost, which suggests this specification should be estimated with country fixed effect. Failing to account for the internal trade cost fixed effect would render the estimates endogenous and biased downward, because countries are supposed to reduce their internal trade cost across various methods up to the equalized marginal value. The estimation of the above equation requires us to take a stand on the values of β and θ. Following Waugh (2010) we pick $\beta = 0.33$ and Simonovska and Waugh (2014) estimate the value of trade elasticity from 4 to 8. We pick $\theta = 8$ which together with $\beta = 0.33$, according to Eq. (19) implies the coefficient a should roughly equal to 0.25.[17] Moreover, Eq. (19) presents another testable prediction that the coefficients on innovation intensity and employment are roughly the same, namely $m = n$ which will be tested later on. The estimation results are tabulated in Table 4 which also reports the P-value of testing the equality of coefficients on R&D and employment $m = n$.

The column 1 estimates the specification using heteroscedastic robust OLS. As anticipated, because of the presence of other unobserved internal trade cost components, the erratic behavior of infrastructure proxy suggests that failing to control other unobserved internal trade cost components downward bias the estimate. Column 2 reports a specification adding a country specific fixed effect to control unobserved internal trade cost terms and any country specific terms. The coefficient on innovation intensity is similar to the coefficient on employment. The statistical test does not reject the hypothesis that those two coefficients are the same. Moreover, the estimated coefficient is close to 0.25 confirming the validity of our specification. Moving to column 3, the positiveness and statistical significance of infrastructure persist when the infrastructure index is used. Even in situations where different values are assigned to θ to which plenty of results in trade literature are sensitive, most of the results are preserved. Column 4 displays the results under $\theta = 6$. The estimate on infrastructure proxy remains positive and significant. Its value is close to the ones in column 2. The coefficients on innovation intensity and employment are very close to each other and pass the equality test.

Another possibility of endogeneity comes from reverse causality. The government would improve infrastructure conditions to facilitate trade when it anticipates an economic boom. As Francois and Manchin (2013) point out, it is particularly hard to find a satisfactory instrument to obviate the reverse causality concern. In light of Arellano and Bond (1991), we use the two period lag of infrastructure proxy as an instrument for infrastructure. This instrument would be valid when the contemporary shocks on real income would not be correlated with the infrastructure condition one period ahead. Then we apply the two stage least squares and report the results in the last column of Table 4. The 2SLS results are similar to the ones using plain fixed estimation in column 2. Both results are positively significant at 1% level, except that the point estimate on infrastructure length is slightly higher, which does not qualitatively change our result. The statistic tests on the equality of the coefficients on innovation intensity and employment are not rejected. To further ensure the exclusion restriction, we also conduct the 2SLS using 10 period lag as the instrument and report

Table 4
Internal trade elasticity from real income.

	(1) OLS	(2) Length	(3) Index	(4) Theta	(5) 2SLS
INF Length	0.014	0.326***		0.351**	0.401***
	(0.027)	(0.117)		(0.139)	(0.120)
INF Index			0.089***		
			(0.033)		
R&D Intensity	0.647***	0.364***	0.230***	0.324***	0.346***
	(0.036)	(0.048)	(0.027)	(0.057)	(0.060)
Employment	0.042*	0.407**	1.063***	0.297	0.405**
	(0.022)	(0.161)	(0.084)	(0.192)	(0.191)
Constant	4.776***	2.711**	5.882***	2.876*	
	(0.392)	(1.354)	(0.181)	(1.618)	
Observations	469	469	619	469	420
R^2	0.599	0.926	0.987	0.899	
$m = n$ P-value	0.000	0.819	0.000	0.904	0.783
Country fixed effect	No	Yes	Yes	Yes	Yes

Note: Standard errors in parentheses.
The dependent variable is the log of real GDP per capita adjusted by the absorption rate. R&D Intensity means researcher count engaged in certain activities in per million people. INF Length is the infrastructure proxy constructed using summation of length of railroad and paved road. INF Index is the infrastructure index constructed by Donaubauer et al. (2016a). R&D, Employment, INF Length are in their log values. Other indices are in their level values.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

both first stage and second stage results in Appendix C, confirming the results using the 2 period lag in column 5. However, the 2SLS using the 10 period lag drops most of the observations and the estimates might suffer from the weak instrument problem. So we decide to pick the estimates under 2SLS using the 2 year lag as our preferred estimates. Additionally, the working paper version of this paper shows that the estimates are robust to the inclusion of institutional control. To summarize, we can safely conclude that better infrastructure enhance the real income and internal trade, echoing the results in Allen and Arkolakis (2014), Donaldson and Hornbeck (2016), Donaldson (2018) and Fan et al. (2021).

From the estimates of our preferred specification column 5, the internal trade elasticity of infrastructure can be recovered as 1.604, which corresponds to θa in Section 3. Together with the preferred coefficient 0.650 in Table 2, the elasticity of infrastructure on importing trade flows and exporting trade flows combined namely $\theta(b + c)$ equals to 2.254. Because there is no systematic evidence showing the cross country variation of infrastructure affects importing flows and exporting flows heterogeneously, following Fan et al. (2021), we impose the assumption that to the same extent infrastructure enhances both trade flows, which results in $\theta b = 1.127$. The qualitative repercussions from this assumption would not be altered if the discriminatory impacts of infrastructure on both flows are not overly unbalanced. The relative magnitude of θa and θb decides the direction of absorption rate change with respect to the unit increase of infrastructure. Here, the estimation result implies domestic goods expenditure share also known as the absorption rate will increase from better infrastructure, meaning the traditional gains from trade will be dampened according to the ACR formula due to better infrastructure. Though seems odd at the first sight, Eq. (10) shows that the traditional gains from trade have to be adjusted by the change of internal trade cost. The better infrastructure reduces internal trade cost meaning domestically produced goods are more competitive and higher expenditure share on the low cost goods naturally leads to higher real income, although the absorption rate appears to increase.

Moreover, the results of the above two subsections combined have important ramifications. It implies that a better infrastructure tends to increase trade openness, increasing both export and import. This opens the door to additional benefits associated with higher trade openness, such as increasing mobility of the international capital flow and higher

[17] Several notable studies, *e.g.,* Eaton and Kortum (2002) and Allen and Arkolakis (2014), pick $\theta = 8$ as their preferred value. In fact, the following analysis shows that the estimation result is insensitive to the value of θ.

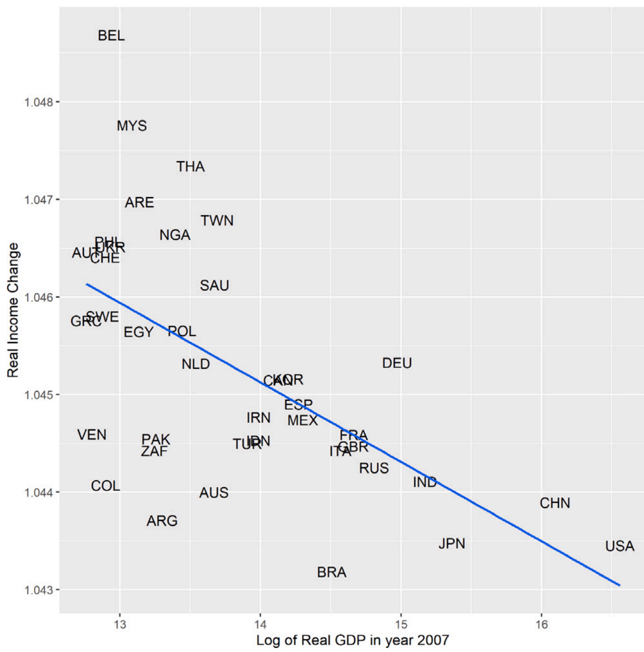

Fig. 2. Cross country real income change.

effectiveness of the financial market, which has been shown extensively by growth and finance literature.

6. Quantitative analysis

Given the elasticity estimated in the previous sections, this section provides several quantitative exercises, highlighting the importance of internal trade cost in assessing the change of real income. The first counterfactual scenario is to increase the transport infrastructure worldwide by 10%, within which two situations are considered. The first situation is to assess the resulting effects using all the estimates we preferred in the previous sections. The second situation is assuming the infrastructure is not affecting the internal trade cost, whose purpose is twofold. On one hand, it provides a conservative result to the elasticity estimation on internal trade because of the lack of ideal instruments, on the other hand, it provides square contrast between the traditional ACR formula and the welfare formula advocated in this paper. The second counterfactual exercise is to compare the counterfactual changes if the 10% increase in transport infrastructure happened in either China or the United States, in light of the recently passed the US infrastructure bill. The details of the data used in this section are delegated into Appendix F.

6.1. 10% worldwide infrastructure increase

Table 5 displays the absorption rate change and real income change under two different scenarios. The columns on the left correspond to the counterfactual situation where all the elasticities are specified as we preferred in the previous sections. Meanwhile, the specification of the two columns on the right is largely the same except that the elasticity of transportation infrastructure on the internal trade cost is taken as 0.

Though it is a worldwide uniform increase of transport infrastructure, its impact is quite diverse across countries in both scenarios with small countries generally having higher real income change than large countries. Fig. 2 plots the real income change against country size for the situation of $a = 0.201$. The fitted blue line shows a clear downward sloping trend confirming the conventional wisdom that small countries generally benefit more than large countries in the uniform trade liberalization with the highest gain for Belgium almost 13% higher than the lowest for Brazil.

Table 5
Cross countries comparison of absorption rate and real income change.

	$a = 0.201$		$a = 0$	
ISOCode	Absorption	Real income	Absorption	Real income
ARE	0.981	1.047	0.933	1.018
ARG	0.993	1.044	0.976	1.006
AUS	0.992	1.044	0.972	1.007
AUT	0.983	1.046	0.940	1.016
BEL	0.975	1.049	0.912	1.024
BRA	0.995	1.043	0.983	1.004
CAN	0.988	1.045	0.957	1.011
CHE	0.983	1.046	0.941	1.016
CHN	0.993	1.044	0.974	1.007
COL	0.992	1.044	0.971	1.007
DEU	0.987	1.045	0.955	1.012
EGY	0.986	1.046	0.950	1.013
ESP	0.989	1.045	0.960	1.010
FRA	0.990	1.045	0.964	1.009
GBR	0.991	1.044	0.966	1.009
GRC	0.986	1.046	0.949	1.013
IDN	0.990	1.045	0.965	1.009
IND	0.992	1.044	0.971	1.008
IRN	0.990	1.045	0.962	1.010
ITA	0.991	1.044	0.967	1.009
JPN	0.994	1.043	0.979	1.005
KOR	0.988	1.045	0.957	1.011
MEX	0.990	1.045	0.962	1.010
MYS	0.978	1.048	0.923	1.020
NGA	0.983	1.047	0.938	1.016
NLD	0.987	1.045	0.955	1.012
PAK	0.990	1.045	0.965	1.009
PHL	0.983	1.047	0.939	1.016
POL	0.986	1.046	0.950	1.013
RUS	0.991	1.044	0.969	1.008
SAU	0.984	1.046	0.944	1.015
SWE	0.986	1.046	0.948	1.014
THA	0.980	1.047	0.929	1.019
TUR	0.991	1.044	0.966	1.009
TWN	0.982	1.047	0.936	1.017
UKR	0.983	1.047	0.939	1.016
USA	0.994	1.043	0.980	1.005
VEN	0.990	1.045	0.964	1.009
XTW	0.986	1.046	0.950	1.013
ZAF	0.991	1.044	0.967	1.009

The wage level in country NGA is normalized to be the same as factual level as country NGA has the lowest export value to China in year 2007.

When compared across two scenarios, several noteworthy points emerge. First, both the absorption rate change and real income change exhibit lower variance across countries when the transport infrastructure's internal trade cost elasticity is in place. Almost all regions have real income increases around 4.5%. The likely reason behind this phenomenon is that the incremental on the internal trade flow because of the reduction of the internal trade cost is the deciding component of the overall impact. Although transport infrastructure has roughly similar elasticity on both importing trade flows and internal trade flows according to our estimates, just as Waugh (2010) points out, the factual absorption rate is roughly the same regardless of the country size and generally higher than 0.5 meaning more than half of the expenditures goes to domestic goods producers. Given the sheer size of the factual absorption rate and its rough homogeneity across countries, the incremental of the internal trade flow is likely to inherit the dominant role in shaping the real income gain. The second noteworthy point can be obtained from comparing the absorption rate across the two scenarios. It is clear that both situations imply the gains from openness and the reduction of absorption rate is uniformly smaller in the situation where the internal trade cost reduces. According to the ACR formula, the situation where $a = 0.201$ should have lower real income gains whereas the model in this paper implies higher real income gains in all the countries and regions, a crucial difference between this model and the ACR formula. As Eq. (10) implies, the gains from the change of absorption rate have to be adjusted by the internal trade cost change to arrive at the real income change. It is possible that one country could have the absorption rate change higher than 1 and the real income gain at the same time. Intuitively, though this country seems to be “more closed from the world" and relies more on its own production, it can actually be better off from consuming more of the more competitive domestically produced goods. In addition to the qualitative difference between the model in this paper and the classical ACR formula, the third point highlights the quantitative difference. The scenario where $a = 0.201$ has much higher real income gains compared with the scenario where $a = 0$, with 5 times higher on average. The lowest is Belgium being 2.06 times higher and the highest is Brazil with 10.10 times higher.

Table 5 also provides conservative quantitative results where the causality effect of infrastructure on internal trade flows is not properly identified due to the lack of an appropriate instrument. The mean of the real income gain for the situation where $a = 0$ is 1.16%, with Belgium being the highest having 2.4%.

6.2. 10% infrastructure increase in China vs. USA

Tables 6 and 7 present the key changes for the counterfactual increase of transport infrastructure in either China or the US. Those two tables share many things in common. The most striking and crucial message comes in the absorption rate change. The unilateral trade cost reduction resulting from the increase in infrastructure decreases all the countries’ absorption rate by a similar magnitude in both cases. Yet, the gains in real income are the highest for the country itself. This result echos the key difference between this model and the ACR formula described in the theory section and the previous counterfactual exercise—the gains from openness has to be adjusted by the internal trade cost change to arrive at the real income change. The second noteworthy point is that the gains from better infrastructure densely concentrate on the country itself, though better infrastructure facilitates both internal and international trade flow. The gains reaped by the infrastructure improving country are around 4.2% under both situations, which are more than one hundred times higher than the benefits spilling over to its trading partners. However, the gains for the infrastructure improving country are fairly sizable given they come from merely a 10% increase in the infrastructure. This echoes the result in Fan et al. (2021) where shutting down the international trade channel will only marginally diminish the gains from the domestic infrastructure improvement. Interestingly, despite the differences in methodology, this study implies similar size welfare gain as Fan et al. (2021). 50,000 km expressway is roughly 13%, close to this exercise, increase of the infrastructure measure used in this paper and yields 5.1% welfare gains which are also tantamount in size to the gains of this exercise. The third insight appears when comparing this counterfactual analysis with the previous one where worldwide infrastructure is increased by 10%. The improvement on the country’s own infrastructure contributes more than 95% of its total gains, which echos the last point—the infrastructure improving country reaps most of its gains.

Despite the above similarity, there are differences when zooming in on particular countries. For example, the US is the major trading partner for Canada and Mexico. They benefit one order higher than other countries from 10% improvement in the US’s infrastructure. On the contrary, 10% improvement in China’s infrastructure benefits other countries to the roughly same extent except for Chinese Taipei who also has China as its major trading partner.

To summarize, several salient points stand out in the entire quantitative analysis. First, transport infrastructure improvement can generate sizable real income gains. Second, the infrastructure improving country reaps almost the entirety of the gains from trade. Third, the gains from absorption rate change are much smaller than the gains directly from the internal trade cost change. Failing to account for the change of the internal trade cost will result in underestimating the real income change by several orders. Slightly lower estimates of a, b, and c—the trade elasticities will not qualitatively neither quantitatively change the above results, as has been shown in the working paper version of this paper.

Table 6
10% increase of transport infrastructure in China.

ISOcode	Wage	Price index	Absorption	Real income
ARE	1.00047	0.99993	0.99857	1.00036
ARG	1.00037	1.00015	0.99941	1.00015
AUS	1.00068	1.00031	0.99902	1.00025
AUT	1.00022	1.00013	0.99976	1.00006
BEL	1.00027	1.00000	0.99931	1.00018
BRA	1.00032	1.00015	0.99957	1.00011
CAN	1.00008	0.99979	0.99925	1.00019
CHE	1.00039	1.00017	0.99943	1.00014
CHN	1.04998	0.98717	0.99920	1.04219
COL	1.00006	0.99990	0.99956	1.00011
DEU	1.00028	1.00006	0.99941	1.00015
EGY	1.00010	0.99980	0.99921	1.00020
ESP	1.00014	1.00002	0.99968	1.00008
FRA	1.00019	1.00008	0.99969	1.00008
GBR	1.00013	1.00000	0.99966	1.00009
GRC	1.00013	1.00000	0.99964	1.00009
IDN	1.00048	1.00008	0.99894	1.00027
IND	1.00035	1.00008	0.99929	1.00018
IRN	1.00080	1.00014	0.99826	1.00044
ITA	1.00014	1.00003	0.99969	1.00008
JPN	1.00090	1.00048	0.99890	1.00028
KOR	1.00151	1.00055	0.99746	1.00064
MEX	0.99996	0.99981	0.99961	1.00010
MYS	1.00119	1.00015	0.99727	1.00069
NGA	1.00000	0.99971	0.99922	1.00020
NLD	1.00009	0.99989	0.99949	1.00013
PAK	1.00034	0.99996	0.99901	1.00025
PHL	1.00144	1.00047	0.99747	1.00064
POL	1.00009	0.99998	0.99971	1.00007
RUS	1.00023	0.99994	0.99922	1.00020
SAU	1.00065	1.00011	0.99857	1.00036
SWE	1.00025	1.00008	0.99956	1.00011
THA	1.00114	1.00033	0.99787	1.00054
TUR	1.00009	0.99995	0.99962	1.00010
TWN	1.00252	1.00092	0.99579	1.00107
UKR	1.00009	0.99981	0.99926	1.00019
USA	1.00005	0.99983	0.99942	1.00015
VEN	1.00018	0.99996	0.99942	1.00015
XTW	1.00049	1.00004	0.99880	1.00031
ZAF	1.00032	1.00004	0.99925	1.00019

The wage level in country NGA is normalized to be the same as factual level.

7. Robustness analysis

The section provides several types of robustness checks. First, we show the prediction in Section 3 is valid meaning regardless of either side of the infrastructure proxy is employed, the estimates always capture the export and import side combined relative to internal trade. Second, we show our main results in Table 2 are robust to the inclusion of institutional variables whose impact numerous studies have found to be significant on trade flows. The third robustness check is to show our main results are robust to infrastructure proxy accounting for the country’s geographical size. It further shows that the results remain valid even when infrastructure is proxied by the railroad or paved road separately.

7.1. Estimations with infrastructure of exporter side

Table 8 displays the estimation results with bilateral trade proxies using OLS and PPML as showcasing and validates the analysis in

Table 7
10% Increase of transport infrastructure in the US.

ISOcode	Wage	Price index	Absorption	Real income
ARE	0.99835	0.99798	0.99903	1.00025
ARG	0.99820	0.99799	0.99945	1.00014
AUS	0.99823	0.99799	0.99938	1.00016
AUT	0.99849	0.99825	0.99937	1.00016
BEL	0.99852	0.99808	0.99884	1.00029
BRA	0.99827	0.99806	0.99946	1.00014
CAN	1.00031	0.99818	0.99437	1.00143
CHE	0.99855	0.99813	0.99889	1.00028
CHN	0.99887	0.99857	0.99921	1.00020
COL	0.99889	0.99823	0.99828	1.00044
DEU	0.99844	0.99818	0.99930	1.00018
EGY	0.99862	0.99810	0.99863	1.00035
ESP	0.99838	0.99824	0.99961	1.00010
FRA	0.99840	0.99822	0.99952	1.00012
GBR	0.99850	0.99820	0.99921	1.00020
GRC	0.99835	0.99811	0.99935	1.00016
IDN	0.99850	0.99832	0.99951	1.00013
IND	0.99856	0.99834	0.99940	1.00015
IRN	0.99837	0.99824	0.99964	1.00009
ITA	0.99842	0.99828	0.99962	1.00010
JPN	0.99849	0.99827	0.99941	1.00015
KOR	0.99851	0.99809	0.99887	1.00029
MEX	1.00031	0.99829	0.99467	1.00136
MYS	0.99897	0.99821	0.99798	1.00051
NGA	1.00000	0.99869	0.99655	1.00088
NLD	0.99838	0.99813	0.99935	1.00017
PAK	0.99892	0.99855	0.99903	1.00025
PHL	0.99875	0.99807	0.99821	1.00045
POL	0.99830	0.99820	0.99973	1.00007
RUS	0.99839	0.99825	0.99965	1.00009
SAU	0.99881	0.99814	0.99822	1.00045
SWE	0.99849	0.99823	0.99932	1.00017
THA	0.99877	0.99825	0.99863	1.00035
TUR	0.99834	0.99820	0.99962	1.00010
TWN	0.99881	0.99812	0.99817	1.00047
UKR	0.99839	0.99823	0.99957	1.00011
USA	1.04795	0.98533	0.99936	1.04215
VEN	1.00050	0.99923	0.99666	1.00085
XTW	0.99870	0.99820	0.99870	1.00033
ZAF	0.99845	0.99824	0.99943	1.00014

The wage level in country NGA is normalized to be the same as factual level.

Section 3. The 1 and 3 columns show the results for OLS and PPML estimations for the exporter side infrastructure proxy and the 2 and 4 columns display the results using the importer side infrastructure proxy. The coefficients of infrastructure proxies confirm that they capture the relative effect regardless of which side's measure is used. The estimates of the infrastructure are the same and their interpretation is that the exporting and importing effect combined, relative to the internal effect. The same results are obtained when estimating using other alternative methods, such as PPML with pair fixed effect.

7.2. Robustness to institutional control

After including the interaction term of the WGI index and border dummy as institutional control, Table 9 displays that most of the coefficients of the infrastructure proxy are significantly positive and slightly smaller than the ones in Table 2 dissipating the concern that infrastructure can easily succumb to omitted variable biases. Those slightly smaller estimates will not quantitatively change our quantitative results which have been shown in the working paper version. As a side note, Table 9 also shows that the signs of coefficients of the institution are unstable in the estimations we conduct. The method in Beverelli et al. (2018) having the institution as the only variable of interests without controls can lead to omitted variable bias, while the estimation with infrastructure as the only dependent variable has higher immunity.

Table 8
Augmented gravity estimations with infrastructure of exporter side.

	(1) EXP_OLS	(2) IMP_OLS	(3) EXP_PPML	(4) IMP_PPML
EXP INF × Border	0.409***		0.196***	
	(0.128)		(0.045)	
IMP INF × Border		0.409***		0.196***
		(0.128)		(0.045)
Distance	−1.251***	−1.251***	−0.627***	−0.627***
	(0.069)	(0.069)	(0.051)	(0.051)
Contiguity	0.420***	0.420***	0.394***	0.394***
	(0.104)	(0.104)	(0.089)	(0.089)
Language	0.333***	0.333***	0.563***	0.563***
	(0.114)	(0.114)	(0.102)	(0.102)
Colony	0.038	0.038	−0.101	−0.101
	(0.125)	(0.125)	(0.110)	(0.110)
RTA	0.487***	0.487***	0.107	0.107
	(0.089)	(0.089)	(0.097)	(0.097)
Border	−8.105***	−8.105***	−6.119***	−6.119***
	(1.574)	(1.574)	(0.663)	(0.663)
Constants	18.715***	18.693***	19.215***	19.213***
	(0.449)	(0.448)	(0.324)	(0.324)
Observations	17406	17406	17409	17409
R^2	0.912	0.912		
Pseudo R^2			0.994	0.994
Country pair FE	No	No	No	No
Importer-Time FE	Yes	Yes	Yes	Yes
Exporter-Time FE	Yes	Yes	Yes	Yes

Note: Standard errors clustered at country pair level are reported in parenthesis. The 1st row shows the results for the infrastructure proxy of the importer side and the 2nd row shows the results for the infrastructure proxy of the exporter side. These results are not the same as the results in Table 2, because we drop some observations with missing values of either the infrastructure proxy of the exporter or importer side and estimate only with observations that have both of them. The main goal of this exercise is to verify our analysis at the end of Section 3. FE is short for fixed effect.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 9
Augmented gravity estimations with institutions.

	(1) OLS	(2) PPML	(3) Pair_OLS	(4) Pair_PPML
INF × Border	0.490***	0.260***	0.039	0.369**
	(0.101)	(0.038)	(1.205)	(0.184)
WGI × Border	0.808***	0.326***	0.602	−0.607***
	(0.259)	(0.071)	(0.608)	(0.143)
Observations	89437	99867	89261	99396
R^2	0.819		0.925	
Pseudo R^2		0.992		0.998
Country pair fixed effect	No	No	Yes	Yes
Importer-Time fixed effect	Yes	Yes	Yes	Yes
Exporter-Time fixed effect	Yes	Yes	Yes	Yes

Note: Standard errors clustered at country pair level are reported in parenthesis. Total length of the railroad and the paved road is selected as infrastructure proxy. WGI index is the institution quality index from World Bank's World Governance Indicator database. The standard gravity variables such as distance are not reported here.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

7.3. Alternative proxies for infrastructure

To account for the intrinsic heterogeneity in the country size, in Table 10, we present the results in which we use the paved road, railroad, and infrastructure density, as the proxy of infrastructure, respectively. Table 10 implies that results similar to our main result can be obtained regardless of what out of those three proxies we pick. However, it reports a negative coefficient of the length of the railroad from PPML estimation with symmetric pair fixed effect, although reassuringly it is insignificant. This might come from the lack of variations in the railroad length variable. A sensible explanation is that the geomorphic feature of some countries is not suitable to have more railroads built, whereas paved road is more adaptable to various geomorphic features.

Table 10
Augmented gravity estimations with alternative proxies for infrastructure.

	Paved Road			Railroad			Total length density		
	(1) OLS	(2) PPML	(3) Pair_PPML	(4) OLS	(5) PPML	(6) Pair_PPML	(7) OLS	(8) PPML	(9) Pair_PPML
INF_Paved × Border	0.455***	0.256***	0.896***						
	(0.130)	(0.037)	(0.154)						
INF_Rail × Border				0.477***	0.211***	−0.124			
				(0.138)	(0.050)	(0.131)			
INF_Density × Border							0.703***	0.232***	0.650***
							(0.244)	(0.052)	(0.146)
Observations	108486	121892	121210	95319	106838	106397	95319	106838	106397
R^2	0.814			0.818			0.818		
Pseudo R^2		0.992	0.998		0.992	0.998		0.992	0.998
Country pair FE	No	No	Yes	No	No	Yes	No	No	Yes
Importer-Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Exporter-Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Standard errors clustered at country pair level are reported in parenthesis.
The independent variables are the log of the length of the paved road, the length of the railroad, and the total length of the paved road and the railroad per area, respectively. The standard gravity variables such as distance are not reported here. FE is short for fixed effect.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

8. Conclusion

Motivated by cross-country differences in the domestic infrastructure investment, this paper attempts to estimate how much infrastructure can affect the trade cost and the gains from trade and it proposes a variant of the EK framework, capable of identifying the country specific variable separately from the country specific fixed effect. This framework can be widely applicable to assess the impact of other country specific factors, *e.g.*, institution quality and political party turnover, on both international trade and internal trade flows. The proposed theoretical model enables the structural interpretation of coefficients on infrastructure and explains why regardless of whichever side's infrastructure is used, the estimate would be the same. Moreover, this theoretical model offers a new welfare formula that better aligns with intuition when the variable of interest affects both the international trade cost and the internal trade cost. Proven by the quantitative evidence, failing to account for the impact of infrastructure improvement on internal trade cost in the welfare formula will not only quantitative but also qualitative misinterpret the gains from trade.

Guided by the theoretical model, the empirical results show that better infrastructure boosts international trade more than internal trade and it has a non-negative impact on internal trade. Further quantitative analysis shows that 10% transportation infrastructure increase induces a sizable 4.2% increase in real income and more than 95% of total gains concentrate on the infrastructure improving country itself. All the above results make infrastructure improvement a favorable policy instrument.

An interesting venue to extend this paper is to embed this framework into a broader setting and assess the relative importance of several interesting factors in shaping the world trade flows and real income changes in a comparative way.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgment

We are thankful to Dr. Julian Donaubauer for providing us with the complete dataset of Global Infrastructure Index.

Table 11
PPML estimations with sub-indices of infrastructure.

	(1) PPML	(2) PPML	(3) PPML	(4) PPML
INF_Tran × Border	0.389***			
	(0.038)			
INF_Com × Border		0.274***		
		(0.044)		
INF_En × Border			0.143***	
			(0.046)	
INF_Fin × Border				0.589***
				(0.047)
Observations	321889	330789	272238	304552
Pseudo R^2	0.989	0.988	0.988	0.989
Country pair fixed effect	No	No	No	No
Importer-Time fixed effect	Yes	Yes	Yes	Yes
Exporter-Time fixed effect	Yes	Yes	Yes	Yes

Note: Standard errors clustered at country pair level are reported in parenthesis. Independent variables from the 1st to the 4th row are sub-indices of the infrastructure index constructed by Donaubauer et al. (2016a) interacted with border dummy. Their categories are transportation, communication, energy, and financial infrastructure, respectively. The standard gravity variables such as distance are not reported here.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix A. Augmented gravity estimations with sub-indices of infrastructure

Tables 11 and 12 showcase the results when we estimate with sub-indices in Donaubauer et al. (2016a) using PPML and OLS with symmetric pair fixed effect. The coefficients of those sub-indices are positive and most of them remain significant. These results are consistent with the results in Table 3 showing our results are not driven by the aggregation of sub-indices.

Appendix B. Augmented gravity estimations with real GDP and real GDP per capita

As Frankel and Romer (1999) and Feyrer (2019) point out there exist reverse causality between trade and income which complicates the identification of the positive effect of trade on income, in the same vein, it is reasonable to have a concern that missing GDP in the estimation could cause omitted variable bias. To alleviate this concern, we run tests that include real GDP and real GDP per capita. However, the simple inclusion of real GDP and real GDP per capita introduces another complexity—multicollinearity between real GDP and the length of transportation infrastructure. Table 13 shows the result of the estimation including GDP instead of the proxy of infrastructure.

Table 12
Pair fixed effect OLS estimations with sub-indices of infrastructure.

	(1) Pair_OLS	(2) Pair_OLS	(3) Pair_OLS	(4) Pair_OLS
INF_Tran × Border	0.001			
	(0.106)			
INF_Com × Border		0.184		
		(0.136)		
INF_En × Border			0.766***	
			(0.212)	
INF_Fin × Border				0.199*
				(0.109)
Observations	240427	246360	211831	231638
R^2	0.862	0.860	0.874	0.866
Country pair fixed effect	Yes	Yes	Yes	Yes
Importer-Time fixed effect	Yes	Yes	Yes	Yes
Exporter-Time fixed effect	Yes	Yes	Yes	Yes

Note: Standard errors clustered at country pair level are reported in parenthesis. Independent variables from the 1st to the 4th row are sub-indices of the infrastructure index constructed by Donaubauer et al. (2016a) interacted with border dummy. Their categories are transportation, communication, energy, and financial infrastructure, respectively. The standard gravity variables such as distance are not reported here.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

As anticipated, the highly significant coefficient on real GDP indicates the high correlation between real GDP and the length of transportation infrastructure causes spurious positive causal interpretation. Moreover, when both real GDP and the length of transportation infrastructure are included as shown in the first two columns of Table 14, the coefficients of the proxy of infrastructure become negative and not significant, which is exactly the typical symptom of multicollinearity issue. The same difficulty persists when the additional variable real GDP per capita is included as shown in the first two columns of Table 15.

The underlying cause of the strong correlation between real GDP and the length of transportation infrastructure is unknown and hard to identify. One possible reason is that the relatively small observations of the length of transportation infrastructure as documented in the data description section compromises its efficiency as a proxy for transportation infrastructure. As a potential remedy, we use another proxy which has much larger observations—the infrastructure index to redo all the exercises as before whose results are shown in the last two columns of Tables 14 and 15. The positive coefficient on the infrastructure index remains significant and persistent across different specifications confirming the main result that the infrastructure facilitates international trade more than internal trade.

As a side note, all the above exercises are of reduced form. The specification in the main content of this paper is derived from the structural gravity model. The structural gravity model implies we should not find a significant relative effect from real GDP between international trade and internal trade because all the effects from the size of the economy are subsumed in the country specific fixed effects. In the preferred specification of pair fixed effect as in the last column of Tables 14 and 15, where the unobserved differential internal trade costs are controlled for as mentioned in footnote 10, the coefficient on real GDP remains insignificant, echoing the implication of structural gravity model. To summarize, despite the difficulty of the multicollinearity issue, we have enough reduced form support for our results in the main content of this paper, and those results are also backed up by our preferred pair fixed effect estimation from structural gravity theory.

Appendix C. IV estimations of internal trade elasticity

Here we confirm the results in Table 4 using a longer lag of the total infrastructure length as the instrument to further ensure exclusion restriction. The IV estimation results using 2 different lags as instruments are tabulated in Table 16. The first stage results hint that both instruments are relevant to the infrastructure proxy, although the 10 year lag is a bit less significantly correlated with the infrastructure proxy. This is intuitive because as time elapses, the correlation tends to fade away and be diluted by other contemporaneous shocks. The second stage results show both instruments indicate a positive internal trade elasticity estimate with the 10 year lag identifying a slightly higher effect. However, we tend to interpret the results using the 10 year lag with caution. First, the 10 year lag reduces sample size significantly. It reduces the initial 469 observations to 169 observations with around two-thirds dropped, which would significantly impair the validity of the estimates. Second, the Kleibergen–Paap rk F statistic is smaller than the conventionally accepted value 10 suggesting a weak instrument. Third, the erratic behavior of the coefficients on R&D and employment might imply the instrument is weak or under identified. Therefore, our final take is to appoint estimates from the IV estimation with the 2 year lag instrument as our preferred estimates and use them in the quantitative analysis ensued.

Table 13
Augmented gravity estimations with GDP.

	(1) OLS	(2) PPML	(3) Pair_OLS	(4) Pair_PPML
GDP × Border	0.787***	0.384***	0.782***	1.300***
	(0.101)	(0.036)	(0.205)	(0.125)
Observations	397819	519468	396947	519468
R^2	0.748		0.853	
Pseudo R^2		0.985		0.996
Country pair fixed effect	No	No	Yes	Yes
Importer-Time fixed effect	Yes	Yes	Yes	Yes
Exporter-Time fixed effect	Yes	Yes	Yes	Yes

Note: Standard errors clustered at country pair level are reported in parenthesis. Real GDP data are from Penn World Table 9.1. The first column shows the result from the OLS estimation. The second column shows the result from the PPML estimation. The third column shows the result with a symmetric country pair fixed effect using the OLS. The fourth column shows the result with a symmetric country pair fixed effect using the PPML. The standard gravity variables such as distance are not reported here.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 14
Augmented gravity estimations with GDP and infrastructure proxy.

	(1) PPML	(2) Pair_OLS	(3) PPML_Index	(4) Pair_OLS_Index
INF × Border	−0.052	−1.263		
	(0.065)	(0.771)		
INF Index × Border			0.338***	0.394**
			(0.041)	(0.185)
GDP × Border	0.400***	2.157***	0.264***	0.182
	(0.068)	(0.392)	(0.032)	(0.187)
Observations	106838	95163	302736	230486
R^2		0.923		0.866
Pseudo R^2	0.993		0.989	
Country pair fixed effect	No	Yes	No	Yes
Importer-Time fixed effect	Yes	Yes	Yes	Yes
Exporter-Time fixed effect	Yes	Yes	Yes	Yes

Note: Standard errors clustered at country pair level are reported in parenthesis. This table contrasts the results of including the infrastructure index to the ones using infrastructure length. The real GDP data are from Penn World Table 9.1. Column 1 is the PPML estimation including the total length. Column 2 is a symmetric country pair fixed effect using the OLS estimation including the total length. Column 3 is the PPML estimation including the infrastructure index. Column 4 is a symmetric country pair fixed effect using the OLS estimation including the infrastructure index. The standard gravity variables such as distance are not reported here.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix D. Hat algebra of general equilibrium

Following Dekle et al. (2008), we define $\hat{x} = x'/x$ as the change of any equilibrium variables x in its counterfactual x' relative to its factual x. The labor market clear condition in the equilibrium is given as:

$$w_i L_i = \beta w_i L_i + \beta \sum_j \pi_{ij} X_j, \tag{21}$$

Table 15
Estimations with GDP per capita and infrastructure proxy.

	(1) PPML	(2) Pair_OLS	(3) PPML_Index	(4) Pair_OLS_Index
INF × Border	−0.045	−1.277*		
	(0.068)	(0.760)		
INF Index × Border			0.522***	0.346*
			(0.070)	(0.193)
GDP × Border	0.358***	2.469***	0.228***	−0.215
	(0.072)	(0.742)	(0.029)	(0.360)
GDP per Capita × Border	0.807***	−0.361	−0.245**	0.772
	(0.122)	(0.723)	(0.103)	(0.505)
Observations	106838	95163	291575	223095
R^2		0.923		0.869
Pseudo R^2	0.993		0.990	
Country pair fixed effect	No	Yes	No	Yes
Importer-Time fixed effect	Yes	Yes	Yes	Yes
Exporter-Time fixed effect	Yes	Yes	Yes	Yes

Note: Standard errors clustered at country pair level are reported in parenthesis. This table displays the results of including two types of infrastructure proxies, real GDP and real GDP per capita. GDP data are from Penn World Table 9.1. Column 1 is the PPML estimation including the total length. Column 2 is a symmetric country pair fixed effect using the OLS estimation including the total length. Column 3 is the PPML estimation including the infrastructure index. Column 4 is a symmetric country pair fixed effect using the OLS estimation including the infrastructure index. The standard gravity variables such as distance are not reported here.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 16
IV estimations using different lags.

	Second stage	
	2 Year Lag	10 Year Lag
INF Length	0.401***	0.636**
	(0.120)	(0.309)
R&D Intensity	0.346***	−0.240
	(0.060)	(0.180)
Employment	0.405**	0.862**
	(0.191)	(0.362)
	First Stage	
2 Lag Length	0.830***	
	(0.075)	
10 Lag Length		0.302*
		(0.158)
R&D Intensity	0.007	0.056
	(0.015)	(0.062)
Employment	−0.050	−0.577*
	(0.083)	(0.362)
Observations	420	169
KP F-Statistic	121.630	3.630
Country fixed effect	Yes	Yes

Notes: Standard errors in parentheses.
The dependent variable is the log of real GDP per capita adjusted by the absorption rate. R&D Intensity means researcher count engaged in certain activities in per million people. INF Length is the infrastructure proxy constructed using summation of length of railroad and paved road. R&D, Employment, INF Length are in their log values. Other indices are in their level values.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

where the first term on the right denotes the demand for labor in the final goods sector and the second term captures the labor demand in the intermediate goods sector. The total expenditure on intermediate X_j consists of spending for intermediate and final goods usage which are:

$$X_j = (1-\beta)w_jL_j + \frac{(1-\beta)^2}{\beta}w_jL_j = \frac{1-\beta}{\beta}w_jL_j. \tag{22}$$

The above two equations combined imply the following

$$w_iL_i = \sum_j \pi_{ij}w_jL_j. \tag{23}$$

Several other relevant equations are inherited from the main text, which includes the unit cost formula:

$$c_i = \frac{w_i^{\beta}P_i^{1-\beta}}{\gamma}, \tag{24}$$

the price index formula:

$$P_j = A\left\{\sum_{i\neq j}\left[T_i\left(c_it_{i.}\tau_{ij}t_{.j}\right)^{-\theta}\right] + T_j(c_jt_{jj})^{-\theta}\right\}^{-1/\theta}, \tag{25}$$

the trade share formula:

$$\pi_{ij} = \begin{cases} \dfrac{T_i\left(c_it_{i.}\tau_{ij}t_{.j}\right)^{-\theta}}{\sum_{s\neq j}\left[T_s\left(c_st_{s.}\tau_{sj}t_{.j}\right)^{-\theta}\right]+T_j(c_jt_{jj})^{-\theta}} & \forall i\neq j,\\ \dfrac{T_i(c_it_{ii})^{-\theta}}{\sum_{s\neq j}\left[T_s\left(c_st_{s.}\tau_{sj}t_{.j}\right)^{-\theta}\right]+T_j(c_jt_{jj})^{-\theta}} & i=j. \end{cases} \tag{26}$$

Given the variable set $\{T_i, L_i, t'_{i.}, t'_{.j}\}$, Eqs. (21) to (26) provide the panorama of the new counterfactual general equilibrium. However, knowing T_i and L_i would require much more demanding data availability and a more precise model layout. Provided that the rate of change is the focus of interest, a roundabout way is to solve the general equilibrium in the rate of change a.k.a hat algebra. The hat algebra version of the general equilibrium is given as follows:

$$\hat{c}_i = \hat{w}_i^{\beta}\hat{P}_i^{1-\beta}, \tag{27}$$

$$\hat{P}_j = \left\{\sum_i \pi_{ij}(\hat{c}_i\hat{t}_{ij})^{-\theta}\right\}^{-1/\theta}, \tag{28}$$

$$\hat{\pi}_{ij} = \left\{\frac{\hat{c}_i\hat{t}_{ij}}{\hat{P}_j}\right\}^{-\theta}, \tag{29}$$

$$w'_iL_i = \sum_j \pi'_{ij}w'_jL_j, \tag{30}$$

where $\delta_{ij} := \frac{\pi_{ij}X_j}{\sum_j \pi_{ij}X_j}$ and $t_{ij} := t_{i.}\tau_{ij}t_{.j}$. After solving the general equilibrium in changes, the welfare change is given as in formula (10).

Appendix E. Computation algorithm

1. Using Eq. (30) and factual values of π_{ij} to obtain factual values of w_iL_i,
2. Given initial guess of $\hat{w}_i$, Eqs. (27) and (28) together can solve both $\hat{c}_i$ and $\hat{P}_i$.
3. After knowing both $\hat{c}_i$ and $\hat{P}_i$, Eq. (29) generates value of $\hat{\pi}_{ij}$.
4. Using $\hat{\pi}_{ij}$ and factual values of π_{ij} to construct counterfactual π'_{ij} and plugging back the values of π'_{ij} into Eq. (30) to obtain counterfactual values of w'_iL_i.[18]
5. Dividing the counterfactual values w'_iL_i by the factual values w_iL_i obtained in step 1, to get updated $\hat{w}_i$.
6. Repeat the step 2 to 5 with the updated $\hat{w}_i$ until the values of $\hat{w}_i$ converge.

One caveat is that the above equation system can only identify $\hat{w}_i$ up to a scale, *i.e.*, only the relative value of $\hat{w}_i$ but not its absolute value is meaningful. Therefore, in the quantification exercise, additional restrictions are imposed to pinpoint the solution.

Appendix F. Data for quantification

The data used for quantification analysis is from GTAP 8 Data Base, which is publicly available online. The main motive for using GTAP

[18] We choose to work on π'_{ij} rather than its hat algebra version $\hat{\pi}_{ij}$, because π'_{ij} is a stochastic matrix which has guaranteed eigenvectors associated with eigenvalue 1. This would accelerate the iteration process and minimize the approximation error during iteration.

Table 17
Regions concordance.

ISOCode	Region name	ISOCode	Region name
ARE	United Arab Emirates	DEU	Germany
ARG	Argentina	EGY	Egypt
AUS	Australia	ESP	Spain
AUT	Austria	FRA	France
BEL	Belgium	GBR	United Kingdom
BRA	Brazil	GRC	Greece
CAN	Canada	IDN	Indonesia
CHE	Switzerland	IND	India
CHN	China	IRN	Iran (Islamic Republic of)
COL	Colombia	ITA	Italy
JPN	Japan	SAU	Saudi Arabia
KOR	Republic of Korea	SWE	Sweden
MEX	Mexico	THA	Thailand
MYS	Malaysia	TUR	Turkey
NGA	Nigeria	TWN	Chinese Taipei
NLD	Netherlands	UKR	Ukraine
PAK	Pakistan	USA	United States
PHL	Philippines	VEN	Venezuela (Bolivarian Republic of)
POL	Poland	XTW	Rest of the World
RUS	Russian Federation	ZAF	South Africa

data is that it has well-regularized data values such that the adding-up constraints stipulated in Fally (2015) are satisfied which complies better with the theoretical framework in this paper.

The original GTAP database consists of 129 regions and 57 sectors for year 2007. To circumvent the computation limits and ensure bilateral trade data has nonzero entries to guarantee the irreducibility and aperiodicity of the trade share matrix, we keep the highest 39 regions in terms of real GDP and aggregate the rest of the regions into the rest of the world(XTW). The 40 regions' ISO codes and country names concordance is given in Table 17. The complete data for quantitative exercises utilize the bilateral trade flow data and domestic input–output data out of the GTAP database. The original bilateral trade flow data contains bilateral sector level trade flows which we aggregate to obtain country level bilateral trade flows. As for the construction of country level internal trade flows a.k.a absorption, we aggregate the entries of the domestic input-output flows across each pair sector. The resulting internal trade flows resemble more closely the gross expenditure used in the theoretical framework, rather than a value-added term.

References

Agnosteva, Delina E., Anderson, James E., Yotov, Yoto V., 2019. Intra-national trade costs: Assaying regional frictions. Eur. Econ. Rev. 112, 32–50. http://dx.doi.org/10.1016/J.EUROECOREV.2018.11.008.

Allen, Treb, Arkolakis, Costas, 2014. Trade and the topography of the spatial economy. Q. J. Econ. 129 (3), 1085–1139. http://dx.doi.org/10.1093/qje/qju016.

Allen, Treb, Arkolakis, Costas, 2022. The Welfare Effects of Transportation Infrastructure Improvements. Rev. Econom. Stud. http://dx.doi.org/10.1093/restud/rdac001, forthcoming.

Alvarez, Fernando, Lucas, Robert E., 2007. General equilibrium analysis of the Eaton-Kortum model of international trade. J. Monetary Econ. 54 (6), 1726–1768. http://dx.doi.org/10.1016/j.jmoneco.2006.07.006.

Anderson, James E., Marcouiller, Douglas, 2002. Insecurity and the pattern of trade: An empirical investigation. Rev. Econ. Stat. 84 (2), 342–352. http://dx.doi.org/10.1162/003465302317411587.

Arellano, Manuel, Bond, Stephen, 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. Rev. Econom. Stud. 58 (2), 277–297. http://dx.doi.org/10.2307/2297968.

Arkolakis, Costas, Costinot, Arnaud, Rodríguez-Clare, Andrés, 2012. New trade models, same old gains? Amer. Econ. Rev. 102 (1), 94–130. http://dx.doi.org/10.1257/aer.102.1.94.

Atkin, David, Donaldson, Dave, 2017. Who's getting globalized? the size and implications of intra-national trade costs. SSRN Electron. J. http://dx.doi.org/10.2139/ssrn.2818941.

Baier, Scott L., Bergstrand, Jeffrey H., 2007. Do free trade agreements actually increase members' international trade? J. Int. Econ. 71 (1), 72–95. http://dx.doi.org/10.1016/j.jinteco.2006.02.005.

Banerjee, Abhijit, Duflo, Esther, Qian, Nancy, 2020. On the road: Access to transportation infrastructure and economic growth in China. J. Dev. Econ. 145, 102442. http://dx.doi.org/10.1016/j.jdeveco.2020.102442.

Beverelli, Cosimo, Keck, Alexander, Larch, Mario, Yotov, Yoto V., 2018. Institutions, Trade and Development: A Quantitative Analysis. CESifo Working Paper, https://ideas.repec.org/p/ris/drxlwp/2018_003.html.

Borchert, Ingo, Larch, Mario, Shikher, Serge, Yotov, Yoto V., 2021. The international trade and production database for estimation (ITPD-E). Int. Econ. 166, 140–166. http://dx.doi.org/10.1016/j.inteco.2020.08.001.

Cavallo, Alberto, Feenstra, Robert C., Inklaar, Robert, 2021. Product variety, the cost of living and welfare across countries. Amer. Econ. J.: Macroecon. http://dx.doi.org/10.1257/MAC.20210137, forthcoming.

Clark, Ximena, Dollar, David, Micco, Alejandro, 2004. Port efficiency, maritime transport costs, and bilateral trade. J. Dev. Econ. 75 (2 SPEC. ISS.), 417–450. http://dx.doi.org/10.1016/j.jdeveco.2004.06.005.

Coughlin, Cletus C., Novy, Dennis, 2021. Estimating border effects: The impact of spatial aggregation. Internat. Econom. Rev. 62 (4), 1453–1487. http://dx.doi.org/10.1111/IERE.12520.

Dekle, Robert, Eaton, Jonathan, Kortum, Samuel, 2008. Global rebalancing with gravity: Measuring the burden of adjustment. IMF Staff Pap. 55 (3), 511–540. http://dx.doi.org/10.1057/imfsp.2008.17.

Demir, Firat, Hu, Chenghao, 2020. Destination institutions, firm heterogeneity and exporter dynamics: empirical evidence from China. Rev. World Econ. 156 (1), 183–217. http://dx.doi.org/10.1007/s10290-019-00358-x.

Donaldson, Dave, 2018. Railroads of the Raj: Estimating the impact of transportation infrastructure. Amer. Econ. Rev. 108 (4–5), 899–934. http://dx.doi.org/10.1257/aer.20101199.

Donaldson, Dave, Hornbeck, Richard, 2016. Railroads and American economic growth: A "market access" approach. Q. J. Econ. 131 (2), 799–858. http://dx.doi.org/10.1093/qje/qjw002.

Donaubauer, Julian, Glas, Alexander, Meyer, Birgit E., Nunnenkamp, Peter, 2018. Disentangling the impact of infrastructure on trade using a new index of infrastructure. Rev. World Econ. 154 (4), 745–784. http://dx.doi.org/10.1007/s10290-018-0322-8.

Donaubauer, Julian, Glas, Alexander, Nunnenkamp, Peter, 2015. Infrastructure and Trade: A Gravity Analysis for Major Trade Categories Using a New Index of Infrastructure. Kiel Working Paper 2016, (2016), Kiel Institute for the World Economy (IfW), Kiel, http://hdl.handle.net/10419/125932.

Donaubauer, Julian, Meyer, Birgit E., Nunnenkamp, Peter, 2016a. A new global index of infrastructure: construction, rankings and applications. World Econ. 39 (2), 236–259. http://dx.doi.org/10.1111/twec.12290.

Donaubauer, Julian, Meyer, Birgit, Nunnenkamp, Peter, 2016b. Aid, infrastructure, and FDI: assessing the transmission channel with a new index of infrastructure. World Dev. 78, 230–245. http://dx.doi.org/10.1016/j.worlddev.2015.10.015.

Eaton, Jonathan, Kortum, Samuel, 2002. Technology, geography, and trade. Econometrica 70 (5), 1741–1779. http://dx.doi.org/10.1111/1468-0262.00352.

Fally, Thibault, 2015. Structural gravity and fixed effects. J. Int. Econ. 97 (1), 76–85. http://dx.doi.org/10.1016/j.jinteco.2015.05.005.

Fan, Jingting, Lu, Yi, Luo, Wenlan, 2021. Valuing domestic transport infrastructure: A view from the route choice of exporters. Rev. Econ. Stat. 1–46. http://dx.doi.org/10.1162/rest_a_01084.

Feenstra, Robert C., Inklaar, Robert, Timmer, Marcel P., 2015. The next generation of the penn world table. Amer. Econ. Rev. 105 (10), 3150–3182. http://dx.doi.org/10.1257/aer.20130954.

Feenstra, Robert C., Ma, Hong, 2014. Trade facilitation and the extensive margin of exports. Jpn. Econ. Rev. 65 (2), 158–177. http://dx.doi.org/10.1111/jere.12031.

Felbermayr, Gabriel J., Tarasov, Alexander, 2022. Trade and the spatial distribution of transport infrastructure. J. Urban Econ. 130, 103473. http://dx.doi.org/10.1016/J.JUE.2022.103473.

Feyrer, James, 2019. Trade and income—exploiting time series in geography. Amer. Econ. J.: Appl. Econ. 11 (4), 1–35. http://dx.doi.org/10.1257/app.20170616.

Francois, Joseph, Manchin, Miriam, 2013. Institutions, infrastructure, and trade. World Dev. 46, 165–175. http://dx.doi.org/10.1016/j.worlddev.2013.02.009.

Frankel, Jeffrey A., Romer, David, 1999. Does trade cause growth? Amer. Econ. Rev. 89 (3), 142.

Han, Zheng, 2021. Heterogeneous internal trade cost and its implications in trade. SSRN Electron. J. 1–47.

Han, Zheng, Fujii, Daisuke, 2020. Nonlinear pricing in transport industry and the gains from trade. SSRN Electron. J. 1–47.

Head, Keith, Mayer, Thierry, 2014. Gravity equations: workhorse,toolkit, and cookbook. In: Handbook of International Economics. Vol. 4, Elsevier B.V., pp. 131–195. http://dx.doi.org/10.1016/B978-0-444-54314-1.00003-3.

Heid, Benedikt, Larch, Mario, Yotov, Yoto V., 2021. Estimating the effects of non-discriminatory trade policies within structural gravity models. Can. J. Econ. 54 (1), 376–409. http://dx.doi.org/10.1111/caje.12493.

Jaworski, Taylor, Kitchens, Carl, Nigai, Sergey, 2022. Highways and Globalization. NBER Working Paper, National Bureau of Economic Research, Cambridge, MA, pp. 1–54. http://dx.doi.org/10.3386/W27938.

Kerem CoŞar, A., Fajgelbaum, Pablo D., 2016. Internal geography, international trade, and regional specialization. Am. Econ. J. Microecon. 8 (1), 24–56. http://dx.doi.org/10.1257/mic.20140145.

Levchenko, Andrei A., 2007. Institutional quality and international trade. Rev. Econom. Stud. 74 (3), 791–819. http://dx.doi.org/10.1111/j.1467-937X.2007.00435.x.

Li, Bingjing, Ma, Lin, 2021. JUE insight: Migration, transportation infrastructure, and the spatial transmission of COVID-19 in China. J. Urban Econ. 127, 103351. http://dx.doi.org/10.1016/j.jue.2021.103351.

Ma, Lin, Tang, Yang, 2020. Geography, trade, and internal migration in China. J. Urban Econ. 115, 103181. http://dx.doi.org/10.1016/j.jue.2019.06.004.

Manova, Kalina, 2013. Credit constraints, heterogeneous firms, and international trade. Rev. Econom. Stud. 80 (2), 711–744. http://dx.doi.org/10.1093/restud/rds036.

Melitz, Marc J., Ottaviano, Gianmarco I.P., 2008. Market size, trade, and productivity. Rev. Econom. Stud. 75 (1), 295–316. http://dx.doi.org/10.1111/j.1467-937X.2007.00463.x.

Nunn, Nathan, Trefler, Daniel, 2014. Domestic institutions as a source of comparative advantage. In: Handbook of International Economics. Vol. 4, Elsevier B.V., pp. 263–315. http://dx.doi.org/10.1016/B978-0-444-54314-1.00005-7.

Ramondo, Natalia, Rodríguez-Clare, Andrés, Saborío-Rodríguez, Milagro, 2016. Trade, domestic frictions, and scale effects. Amer. Econ. Rev. 106 (10), 3159–3184. http://dx.doi.org/10.1257/aer.20141449.

Rehman, Faheem Ur, Noman, Abul Ala, Ding, Yibing, 2020. Does infrastructure increase exports and reduce trade deficit? Evidence from selected South Asian countries using a new Global Infrastructure Index. J. Econ. Struct. 9 (1), 10. http://dx.doi.org/10.1186/s40008-020-0183-x.

Samadi, Ali Hussein, Alipourian, Masoumeh, 2021. Measuring institutional quality: a review. In: Dynamics of Institutional Change in Emerging Market Economies. Springer International Publishing, pp. 143–171. http://dx.doi.org/10.1007/978-3-030-61342-6_6.

Santamaria, Marta A., Ventura, Jaume, Yesilbayraktar, U., 2020. Borders Within Europe. NBER Working Paper, http://dx.doi.org/10.3386/w28301.

Santos Silva, J.M.C., Tenreyro, Silvana, 2006. The log of gravity. Rev. Econ. Stat. 88 (4), 641–658. http://dx.doi.org/10.1162/rest.88.4.641.

Simonovska, Ina, Waugh, Michael E., 2014. The elasticity of trade: Estimates and evidence. J. Int. Econ. 92 (1), 34–50. http://dx.doi.org/10.1016/j.jinteco.2013.10.001.

Tang, Heiwai, 2012. Labor market institutions, firm-specific skills, and trade patterns. J. Int. Econ. 87 (2), 337–351. http://dx.doi.org/10.1016/j.jinteco.2012.01.001.

Volpe Martincus, Christian, Carballo, Jerónimo, Cusolito, Ana, 2017. Roads, exports and employment: Evidence from a developing country. J. Dev. Econ. 125, 21–39. http://dx.doi.org/10.1016/j.jdeveco.2016.10.002.

Waugh, Michael E., 2010. International trade and income differences. Amer. Econ. Rev. 100 (5), 2093–2124. http://dx.doi.org/10.1257/aer.100.5.2093.

Yotov, Yoto V., 2012. A simple solution to the distance puzzle in international trade. Econom. Lett. 117 (3), 794–798. http://dx.doi.org/10.1016/j.econlet.2012.08.032.