HIAS-E-119

Transportation Infrastructure and Trade

Han Zheng [(a)], Li Hongtao [(b)]

(a) Hitotsubashi Institute for Advanced Study, Hitotsubashi University
(b) The University of Tokyo

March, 2022

Hitotsubashi Institute for Advanced Study, Hitotsubashi University
2-1, Naka, Kunitachi, Tokyo 186-8601, Japan
tel:+81 42 580 8668 http://hias.hit-u.ac.jp/

Transportation Infrastructure and Trade

Han Zheng*

Li Hongtao†

March 21, 2022

Abstract

This paper offers a variant of Ricardian model able to structurally interpret the estimate of country-specific variable—transportation infrastructure in a commonly used fixed effect gravity estimation. Guided by this new theoretical framework, this paper shows that transportation infrastructure enhances international trade more than internal trade and this result is robust to various estimation methods and different versions of transportation infrastructure measures. Moreover, it shows that the transportation infrastructure has a non-negative effect on internal trade. Further quantitative analysis suggests 10% increase in transportation infrastructure induce 3.9% increase in real income and more than 95% of the gains concentrate on the infrastructure improving country. All the above results suggest that better infrastructure leads to sizable gains providing additional empirical support to policies aiming to improve transportation infrastructure. This paper also suggests, contrary to what ACR formula claims, domestic goods expenditure share change is no longer sufficient to predict how real income changes.

Keywords: Gravity model, Transportation infrastructure, Internal trade cost

JEL Codes: F10, F14

*Hitotsubashi Institute for Advanced Study, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo 186-8601, Japan. Email address: hanzheng211@gmail.com.

†Graduate School of Public Policy, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. Email address: li-hongtao@g.ecc.u-tokyo.ac.jp.

1 Introduction

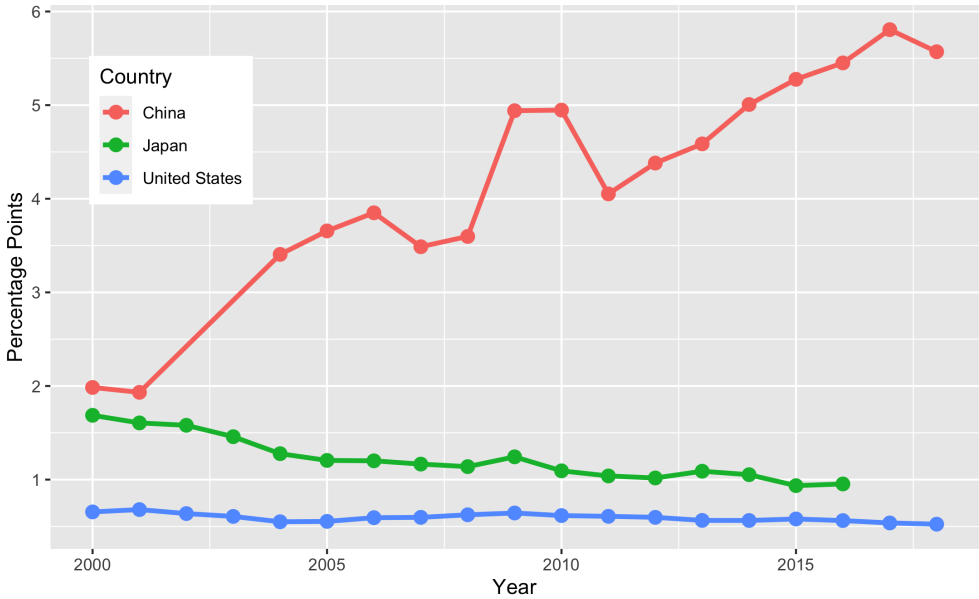
Job creations and investment payoff are two common pros for infrastructure investment discussed in economics literature while relatively few literature analyze how infrastructure affects trade flows. The scarcity of researches is partly because of the nondiscriminatory feature of transportation infrastructure 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. 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.

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 and on the contrary, intuitions 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 Atkin and Donaldson (2017), Ramondo et al. (2016), Kerem Coşar and Fajgelbaum (2016). 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 moti-

Figure 1: Total Inland Transport Infrastructure Expenditure over GDP



Note: The data comes from the OECD Database. The y axis is the total inland transport infrastructure expenditure percentage points out of GDP.

vated by the real-time infrastructure investment difference shown in Table 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 are following. A new index of infrastructure is created in Donaubauer et al. (2016b), and it is used to study various issues in Donaubauer et al. (2016a), Donaubauer et al. (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 trade deficit. They use fully modified least squares (FMOLS) and dynamic ordinary least square (DOLS) estimator to deal with endogeneity issue by adding the leads and lags, while we use 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 above, 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. Recently, using a newly assembled sub-country regional bilateral trade data set, 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 summarizes the empirical regularities about the gravity estimated trade cost. Section 2 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 to 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 robustness check. Section 8 provides policy implication for the policymaker. The final section concludes.

[1]Many other noteworthy studies mainly analyze the relations between domestic infrastructure and but not confined to real income. For example, Li and Ma (2021) finds better domestic infrastructure contributes to the rapid spread of Covid-19 in the early stage of the global pandemic.

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 intermediates goods. Each country sources each variety of intermediates 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 derives utility on 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:

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

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 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:

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

where τ_{ij} is the standard iceberg trade cost and $t_{\cdot j}$ is the trade cost stemming from the importing country's infrastructure condition, which applies unanimously to the goods from every exporting country. Similarly $t_{i\cdot}$ 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 a reminiscent of what Han (2021) has convincingly shown that the internal trade cost at country level is a func-

[2] Cavallo et al. (2021) arrive at a similar formula as in equation (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.

tion 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 hightlighted in Han (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]

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

with price index P_i given as:

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

where $A = \left[\Gamma\left(\frac{\theta+1-\sigma}{\theta}\right)\right]^{1/(1-\sigma)}$, a constant of parameters. Equation (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. This variant incorporates two new components into the country's competitiveness measure $T_i\left(c_i t_{i\cdot} \tau_{ij} t_{\cdot j}\right)^{-\theta}$: $t_{i\cdot}$ and $t_{\cdot j}$, which function the same way as τ_{ij}—higher cost lower competitiveness, except both are country-specific

[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 (2020). However, several other versions of gravity specifications not using fixed effect to control multilateral resistance terms permit analysing the effect of country-specific variables such as infrastructure, 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 tend 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.

rather than pair specific. Some straight forward substitutions yield the 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 parameterzied 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 equation (8) can be alternatively writtern as:

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

The hat algebra version of equation (8) is given as:

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

Equation (10) looks very similar to the famous ACR formula in Arkolakis et al. (2012) where the ACR formula only contains the first parenthesis in equation (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 both trade cost to 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 an 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 equation (10) will adjust the gains from international trade by the internal trade conditions.[6] More importantly, equations (8) and (10) are useful to structurally estimate the country-specific components of the internal

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

[6] 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.

trade cost, where the standard gravity estimation falls short. Additionally equation (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 equation (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 equation (10).[7]

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 country-specific infrastructure proxy is used in estimation.

Inspection on equation (5) reveals that the standard fixed effect gravity model is plagued by the econometric difficulty of the perfect collinearity. The country-specific trade cost $t_{i.}$ and $t_{.j}$ 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 equation (7):

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

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

[7] 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 equation (5) reveals that the technology parameter has partial trade elasticity one, whereas 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.

When apply the fixed effect specification to estimate above equations, the first part of the above equations is the exporter fixed effect exp_i and the importer fixed effect imp_j captures the second and third terms. 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 equation (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 shouldn't appear in the internal trade flow observation as in equation (12). To balance this effect, the second term in equation (12) has to contain $t_{j\cdot}^{-\theta}$ which in turn will appear in the importer fixed effect in equation (11) because of the very same importer fixed effect in both international trade flow and internal trade flow observations and, as a result, the forth term in equation (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 proxy:

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

Then the trade flow equations become:

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

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, what the coefficient in front of INF_j in the following specification will capture is $\theta(b + c - a)$, namely the infrastructure's effect on the international flow relative to the internal flow as follows:

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

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 leads to biased estimates.

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

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

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

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

$\theta a \ln INF_i$ 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 statistics of several variables of interest.

4.1 Data Description

Trade Flows — The trade flow data is from the ITPD-E Database.[8] Its coverage spans 17 years, from 2000 to 2016, including 243 countries and 170 industries. Because the analysis in this paper is country level, the industrywise trade flows are aggregated to obtain country level trade flows. 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 gravity method, therefore, the database is suitable for the gravity estimation. The details of trade data used for quantitative analysis will be separately presented in the Appendix C.

Bilateral Trade Proxies — The proxies come from the CEPII website, which covers bilateral distance over 225 countries. The merits of the proxies constructed there are

[8]Its link is given as `https://www.usitc.gov/data/gravity/itpde.htm.`

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.[9]

Regional Trade Agreement — The regional trade agreement dummy comes from the Mario Larch's Regional Trade Agreements Database.[10] It covers 516 RTA notified to WTO from year 1950 to 2019. Dummies separating the type of trade agreements, such as currency union, free trade agreement, economic integration agreement, are also included.

Infrastructure Proxy — As the main variable of interest, several types of infrastructure variables are collected from various sources to ensure 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.[11] It measures how extensive one country's distribution network is. Considering the innate heterogeneity of 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. (2016b), where an unobserved component model is employed in order to structurally clump together multiple highly correlated infrastructure measures. Donaubauer et al. (2016b) 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.

Institution Control — Numerous studies have analyzed the impact of institutional quality on bilateral trade flows and trade patterns.[12] 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 differential impact on international trade flow and internal trade flow. In light of those literature, we decide

[9]See Yotov (2012), Agnosteva et al. (2014), Ramondo et al. (2016), Beverelli et al. (2018) and Heid et al. (2021).

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

[11]Though the railroad and road transportation are substitutes for goods delivering, the sum of both transportation modes more accurately represents 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.

[12]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, and the institutional index coming from Economic Freedom of the World is another alternative. Beverelli et al. (2018) also choose WGI as one of the options.

to include institutional variables as control to obviate the omitted variable bias. Following Beverelli et al. (2018), we use the institutional quality indices from the Wonld Bank's World Governance Indicators which contains 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. To remedy this ad hoc approach of taking average, we use the summary rating from the Economic Freedom of the World: 2021 Annual Report issued by Fraser Institute. Their data covers even longer time span than World Governance Indicators database and it offers a summary index of 165 countries over plenty of minor sub-categories.

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 World Bank's World Development Indicator database as the innovation intensity proxy. This data set documents annual number of researchers engaged in certain activities, such as developing concepts and theories, 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, 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 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 abnormal large number exactly testifies the validity of this data set, because this data set contains internal trade flows and large countries generally absorb majority of productions domestically. Moreover, the internal trade flows are computed using value of gross production rather than GDP as reference sum, which further enlarges the size of internal trade flows. The observation attains this highest trade flows in record is the internal trade of US in 2008. The infrastructure proxy in length has 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 infrastructure index and WGI institution index have mean around 0 and standard deviation 1, because both are an index composite clumping several measures together using unobserved components model method. This method generally assumes the underlying true variables across countries follow standard normal distribution. The negative values of those proxies prevents a theoretical consistent interpretation of their coefficients, however the signs of their coefficients have qualitative meanings. As for innovation intensity, there is huge disparity among countries with lowest record in this sample from Lesotho and highest record from Denmark.

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
Institution EFW	2236	6.820568	.9877112	2.7	9.09
Innovation Intensity	1340	2014.32	1957	5.91	8065.89

Note: This table only includes variables of interest. The second row corresponds to infrastructure measure constructed by summing railroad length and paved road length. The third row is the index constructed in Donaubauer et al. (2016b). The forth row is the institutional index available from World Bank's World Governance Indicator database. The fifth row is the institutional index comes from Economic Freedom of the World: 2021 Annual Report. The sixth row is the number of researchers engaged in certain activities per million people from World Bank's World Development Indicator database.

5 Empirical Findings

This section first presents the empirical results in augmented gravity estimation motivated by equation (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 decreases according to ACR formula, the real income increases because equation (10) shows that effect from absorption rate change has to be adjusted by the change of internal trade cost.

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 insignificant. Either way could happen even if the infrastructure has positive impact on trade flow is well established. The exact empirical model is given as follows:

$$X_{ij,t} = \exp(imp_{i,t} + exp_{j,t} + border + INF_{j,t} \times border + INS_{j,t} \times border + BTP_{ij} \times \eta) + \epsilon_{ij,t} \, \forall i, j, \tag{18}$$

where $INS_{j,t}$ is the institutional control added to obviate omitted variable bias, as numerous studies have found significant impact of institution quality on trade flows.

The benchmark results of the above empirical model using the total length of railroad and paved road as infrastructure proxy is 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 forth column estimate equation (18) using bilateral trade proxies with different method 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 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 0.5. Contrary to the assumption Cavallo et al. (2021) made that the domestic trade costs impacts 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 able to capture all the observed or even unobserved components of symmetric trade cost, and 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. Another interesting observation is that institutional quality remains significant at 1% level but its signs alternate across specifications. This indicates studies using institution alone might suffer from omitted variable biases and its effect is sensitive to the regression methods picked, which is confirmed in the robustness analysis.

Because of the relatively low observations of total length infrastructure proxy, we corroborate our primary results using the infrastructure index constructed in Donaubauer

Table 2: Augmented Gravity Estimations

	(1) OLS	(2) Inter_PPML	(3) INF_OLS	(4) INF_PPML	(5) Pair_PPML
Distance	-1.578***	-0.743***	-1.287***	-0.597***	
	(0.017)	(0.031)	(0.054)	(0.046)	
Contiguity	0.709***	0.297***	0.199	0.441***	
	(0.089)	(0.060)	(0.151)	(0.078)	
Language	0.723***	0.160**	0.390***	0.458***	
	(0.034)	(0.065)	(0.091)	(0.088)	
Colony	0.865***	0.163**	0.956***	-0.036	
	(0.087)	(0.082)	(0.120)	(0.099)	
RTA	0.520***	0.300***	0.411***	0.251***	0.262***
	(0.028)	(0.055)	(0.070)	(0.092)	(0.040)
Border			-9.662***	-7.471***	
			(1.212)	(0.571)	
INF×Border			0.490***	0.260***	0.369**
			(0.101)	(0.038)	(0.184)
INS×Border			0.808***	0.326***	-0.607***
			(0.259)	(0.071)	(0.143)
_cons	13.921***	15.257***	15.776***	18.945***	13.552***
	(0.151)	(0.274)	(0.393)	(0.293)	(0.483)
N	416070	564109	89437	99867	99396
R^2	0.737		0.819		
pseudo R^2		0.945		0.992	0.998
Pair Fixed Effect	NO	NO	NO	NO	Yes

Note: Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
The independent variable is the bilateral trade flows in PPML estimation and log trade flows in OLS estimations. Total length of railroad and paved road is selected as infrastructure proxy and institution quality index uses World Bank's World Governance Indicator database.

et al. (2016b), which is tabulated in Table 3. Generally the results using total length of railroad and paved road are preserved using infrastructure index as proxy. The estimated coefficients are positive across all specifications, though in the PPML estimation with symmetric pair fixed effect it is not precisely estimated. Another noteworthy point is that the coefficients of infrastructure index under OLS estimation method 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. As anticipated, both the sign and significance of institution quality proxy fluctuate and its impact on international trade flows relative to internal trade flows might suffer from severe omitted variable bias. Further robustness check using sub-indices of the infrastructure index and alternative institution quality index is offered in robustness analysis section.

Table 3: Augmented Gravity Estimations with Infrastructure Index

	(1) INF_OLS	(2) INF_PPML	(3) INF_Pair_OLS	(4) INF_Pair_PPML
Distance	-1.447***	-0.502***		
	(0.020)	(0.041)		
Contiguity	0.837***	0.425***		
	(0.094)	(0.072)		
Language	0.761***	0.421***		
	(0.038)	(0.064)		
Colony	0.859***	0.018		
	(0.092)	(0.100)		
RTA	0.489***	0.414***	0.107***	-0.070
	(0.033)	(0.064)	(0.037)	(0.061)
Border	-4.123***	-4.779***		
	(0.187)	(0.149)		
INF×Border	1.432***	1.082***	0.497***	0.056
	(0.316)	(0.083)	(0.186)	(0.075)
WGI index	0.225	-0.826***	0.343	-0.411***
	(0.314)	(0.099)	(0.305)	(0.101)
_cons	16.824***	17.980***	0.621***	14.044***
	(0.221)	(0.258)	(0.035)	(0.032)
N	217151	284609	216034	276544
R^2	0.750		0.866	
pseudo R^2		0.989		0.998
Pair Fixed Effect	NO	NO	Yes	Yes

Note: Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
The independent variable is the bilateral trade flows in PPML estimation and log trade flows in OLS estimations. The infrastructure index is selected as infrastructure proxy and institution quality index uses World Bank's World Governance Indicator database.

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 facilitate within-border trade or cross-border trade. The economic mechanism behind the above observation is, unfortunately, out of the realm of this paper. The observation possibly implies greater benefits of economy of scale in shipping for exports and imports than domestically sold goods, because exports and imports need to be transported to several major ports and nodes for declaration or clearance. Or it is the discriminatory shipping policies favoring exports and imports, a country typically issues to foster international trade that propel the above observation. The exact cause 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 is unobserved. As a theory consistent roundabout solution, we capitalize on equation (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 equation (9) is given as follows:

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

where C is a constant composite. This equation motivates the following empirical specification:

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

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. The estimation of 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 equation (19) implies the coefficient a should roughly equal to 0.25. Moreover, equation (19) presents another testable prediction that the coefficients on innovation intensity and employment are roughly the same, namely $m = n$. Both will be tested later on. The estimation results are tabulated in Table 4.

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 which is our preferred specification adds a country-specific fixed effect to control unobserved internal trade cost terms and any country-specific terms. The coefficient on innovation intensity equals to 0.24 close to 0.25 and it is similar to the coefficient on employment. 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 preferred specification. Moving to column 3 where alternative institution quality control is used, the coefficients on innovation intensity and employment remain to be close to each other, though their values are slightly higher than 0.25. The estimate of infrastructure proxy is positively significant and close to the value in column 2. The positiveness and statistical significance of infrastructure persists when infrastructure index is used in column 4. 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 5 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 and 3. The coefficients on innovation intensity and employment are very close to each other.

As Francois and Manchin (2013) point out, it is particularly hard to find a satisfactory instrument to obviate the reverse causality concern. Following Francois and Manchin (2013), we use the one period lag of infrastructure proxy as an instrument for infrastructure, assuming 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 to find the results are roughly similar except 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, neither are the tests on those coefficients equaling to 0.25. 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) and Donaldson (2018).

From the estimates of our preferred specification column 2, the internal trade elasticity of infrastructure can be recovered as 1.512, which corresponds to θa in section 3. Together with the preferred coefficient 0.369 in Table 2, the elasticity of infrastructure on importing

Table 4: Internal Trade Elasticity from Real Income

	(1) OLS	(2) Fixed	(3) Fixed_INS	(4) Fixed_INF	(5) Theta
R&D	0.531***	0.239***	0.384***	0.236***	0.196***
	(0.040)	(0.048)	(0.050)	(0.028)	(0.059)
Employment	0.082***	0.302**	0.309*	1.056***	0.195
	(0.022)	(0.153)	(0.162)	(0.090)	(0.187)
Inf length	-0.031	0.378***	0.311**		0.406***
	(0.028)	(0.109)	(0.120)		(0.134)
Ins index	0.238***	0.291***			0.295***
	(0.041)	(0.040)			(0.049)
WGI index			0.398***	0.111*	
			(0.095)	(0.060)	
Inf index				0.065*	
				(0.035)	
_cons	4.347***	1.052	2.541*	5.810***	1.173
	(0.365)	(1.294)	(1.406)	(0.204)	(1.584)
N	466	466	444	568	466
R^2	0.621	0.933	0.930	0.988	0.904
Fixe Effect	No	Yes	Yes	Yes	Yes

Note: Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
The dependent variable is the log of real GDP per capita adjusted by the absorption rate. R&D data 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. Ins index is the summary rating of one country's institution from Economic Freedom of the World: 2021 Annual Report. WGI index is the institution quality index from World Bank's World Governance Indicator database. Inf index is the infrastructure index constructed by Donaubauer et al. (2016b). R&D, Employment, Inf length are in their log values. Other indices are in their level values.

trade flows and exporting trade flows combined namely $\theta(b+c)$ equals to 1.881. Because there is no systematic evidence showing cross country variation of infrastructure affects importing flows and exporting flows heterogeneously, we impose the assumption that to the same extent infrastructure enhances both trade flows, which results in $\theta b = 0.94$. 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 absorption rate will increase from better infrastructure, meaning the traditional gains from trade will be dampened according to ACR formula due to better infrastructure. Though seems odd at the first sight, equation (10) shows that the traditional gains from trade has to be adjusted by the change of internal trade cost. Better infrastructure reduces internal trade cost meaning domestically produced goods are more competitive and higher expenditure share on low cost goods naturally leads to higher real income, although absorption rate appears to increase.

Moreover, the results of above two subsections combined has 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 international capital and higher effectiveness of financial market, which has been shown extensively by growth and finance literature.

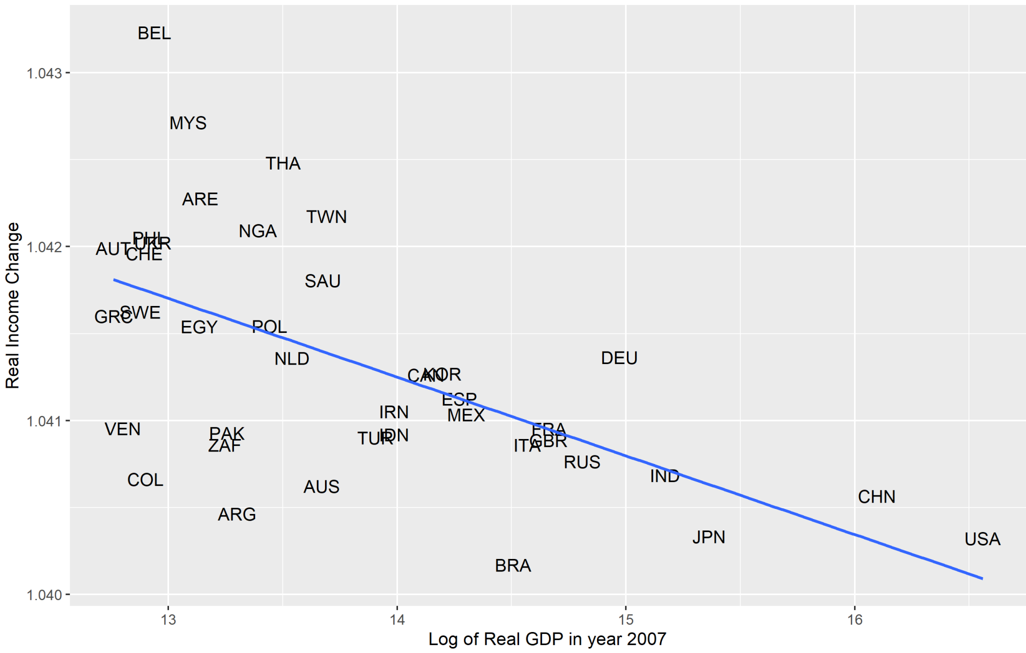
6 Quantitative Analysis

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. First situation is to assess the resulting effects using all the estimates we preferred in the previous sections. 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 lack of ideal instruments, on the other hand it provides square contrast between the traditional ACR formula and welfare formula advocated in this paper. The second counterfactual exercise is to compare the counterfactual changes if the 10% increase of transport infrastructure happened in either China or the United States, in light of the recently passed the US infrastructure bill. The details of the date used in this section is delegated into Appendix C.

6.1 10% Worldwide Infrastructure Increase

The Table 5 displays the absorption rate change and real income change under two different scenarios. The columns on the left corresponds 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.

Figure 2: Cross Country Real Income Change



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. Figure 2 plots the real income change against country size for the situation of $a = 0.189$. The fitted blue line shows clear downward sloping trend confirming the conventional wisdom that small countries generally benefit more than large countries in the uniform trade liberalization with highest gain for Belgium almost 8% higher than the lowest Brazil.

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 increase close around 4.1%. 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 disregard the country size and generally

Table 5: Cross Countries Comparison of Absorption Rate and Real Income Change

	$a = 0.189$		$a = 0$	
ISOCode	Absorption	Real Income	Absorption	Real Income
ARE	0.990	1.042	0.945	1.014
ARG	0.996	1.040	0.980	1.005
AUS	0.996	1.041	0.977	1.006
AUT	0.991	1.042	0.951	1.013
BEL	0.986	1.043	0.927	1.019
BRA	0.997	1.040	0.986	1.003
CAN	0.993	1.041	0.965	1.009
CHE	0.991	1.042	0.951	1.013
CHN	0.996	1.041	0.978	1.006
COL	0.996	1.041	0.976	1.006
DEU	0.993	1.041	0.963	1.010
EGY	0.992	1.042	0.959	1.011
ESP	0.994	1.041	0.967	1.008
FRA	0.995	1.041	0.971	1.008
GBR	0.995	1.041	0.972	1.007
GRC	0.992	1.042	0.958	1.011
IDN	0.995	1.041	0.971	1.007
IND	0.996	1.041	0.976	1.006
IRN	0.994	1.041	0.969	1.008
ITA	0.995	1.041	0.973	1.007
JPN	0.997	1.040	0.983	1.004
KOR	0.993	1.041	0.964	1.009
MEX	0.994	1.041	0.969	1.008
MYS	0.988	1.043	0.937	1.017
NGA	0.990	1.042	0.949	1.013
NLD	0.993	1.041	0.963	1.010
PAK	0.995	1.041	0.971	1.007
PHL	0.990	1.042	0.949	1.013
POL	0.992	1.042	0.959	1.011
RUS	0.995	1.041	0.974	1.007
SAU	0.991	1.042	0.954	1.012
SWE	0.992	1.042	0.958	1.011
THA	0.989	1.042	0.941	1.016
TUR	0.995	1.041	0.972	1.007
TWN	0.990	1.042	0.947	1.014
UKR	0.990	1.042	0.950	1.013
USA	0.997	1.040	0.983	1.004
VEN	0.995	1.041	0.971	1.008
XTW	0.992	1.042	0.959	1.011
ZAF	0.995	1.041	0.973	1.007

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.

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 internal trade flows is likely 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 ACR formula the situation where $a = 0.189$ should has 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 ACR formula. As equation (10) implies, the gains from change of absorption rate has to be adjusted by the internal trade cost change to arrive at the real income change. It is possible that, which happens in the next counterfactual exercise, one country could have 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 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 of the model in this paper from classical ACR formula, the third point highlights the quantitative difference. The scenario where $a = 0.189$ has much higher real income gains compared with scenario where $a = 0$, with 5 times higher on average. The lowest is Belgium being 2.23 times higher and the highest is Brazil with 11.49 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 appropriate instrument. The mean of the real income gain for the situation where $a = 0$ is 0.95%, with Belgium being the highest having 1.9%.

6.2 10% Infrastructure Increase in China vs. USA

Table 6 and Table 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 except the country itself in both cases. Yet, the gains in real income are the highest for the country itself. This result echos the key difference of this model and ACR formula described in the theory section and the previous counterfactual exercise—the direction of the absorption rate change is no longer sufficient to determine the direction real income change. The second noteworthy point is that the gains from better infrastructure densely concentrates on the country itself, though better infrastructure facilitates both internal and international trade flow. The gains reaped by the infrastructure improving country are around 3.9% under both situations, which are more than 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 10% increase in the infrastructure. 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 to particular countries. For example, Japan is one of the major trading partners for both China and the US. But Japan gains around twice as much for 10% improvement in China's infrastructure than 10% improvement in the US's infrastructure.

To summarize, several salient points stand out in the entire quantitative analysis. First, transport infrastructure improvement can generate sizable real income gains. Second, infrastructure improving country reaps almost entirety of the gains from trade. Third, the direction of absorption rate change does not necessarily determine the direction of real income change the magnitude of which is adjusted by the internal trade cost change meanwhile. Failing to account for the change of internal trade cost will result in underestimating the real income change by several times.

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 infrastructure proxy is employed, the estimates always capture the export and import side combined relative to internal trade. Second, we experiment with sub-indices from Donaubauer et al. (2016b) and show our main results are robust to sub-indices. Third, we show that comparing to institution a well established influencing factor, infrastructure is more immune to omitted variable biases. The forth robustness check is to show our main results are robust to infrastructure proxy accounting for country's geographical size. It further shows that the results remain valid even when infrastructure is proxied by railroad or paved road separately.

7.1 Estimations with Infrastructure of Exporter Side

The Table 8 displays the estimation results with bilateral trade proxies using OLS and PPML as showcasing and validates the analysis in Section 3. The coefficients of infrastructure proxies confirms 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

Table 6: 10% Increase of Transport Infrastructure in China

ISOcode	Wage	Price Index	Absorption	Real Income
ARE	1.00034	0.99995	0.99896	1.00026
ARG	1.00027	1.00011	0.99957	1.00011
AUS	1.00049	1.00022	0.99929	1.00018
AUT	1.00016	1.00010	0.99982	1.00004
BEL	1.00019	1.00000	0.99949	1.00013
BRA	1.00023	1.00011	0.99969	1.00008
CAN	1.00006	0.99985	0.99945	1.00014
CHE	1.00028	1.00012	0.99959	1.00011
CHN	1.04653	0.98796	1.00060	1.03934
COL	1.00005	0.99993	0.99968	1.00008
DEU	1.00020	1.00004	0.99957	1.00011
EGY	1.00007	0.99985	0.99943	1.00015
ESP	1.00010	1.00001	0.99977	1.00006
FRA	1.00014	1.00005	0.99977	1.00006
GBR	1.00009	1.00000	0.99975	1.00006
GRC	1.00010	1.00000	0.99974	1.00007
IDN	1.00035	1.00006	0.99923	1.00019
IND	1.00025	1.00006	0.99949	1.00013
IRN	1.00058	1.00010	0.99873	1.00032
ITA	1.00010	1.00002	0.99977	1.00006
JPN	1.00065	1.00035	0.99920	1.00020
KOR	1.00110	1.00040	0.99816	1.00047
MEX	0.99997	0.99986	0.99972	1.00007
MYS	1.00086	1.00011	0.99802	1.00050
NGA	1.00000	0.99979	0.99944	1.00014
NLD	1.00006	0.99992	0.99963	1.00009
PAK	1.00025	0.99997	0.99928	1.00018
PHL	1.00104	1.00035	0.99816	1.00047
POL	1.00006	0.99998	0.99979	1.00005
RUS	1.00017	0.99995	0.99943	1.00014
SAU	1.00048	1.00008	0.99896	1.00026
SWE	1.00018	1.00006	0.99968	1.00008
THA	1.00083	1.00024	0.99845	1.00039
TUR	1.00007	0.99996	0.99973	1.00007
TWN	1.00184	1.00067	0.99694	1.00078
UKR	1.00007	0.99986	0.99946	1.00014
USA	1.00003	0.99987	0.99958	1.00011
VEN	1.00013	0.99997	0.99958	1.00011
XTW	1.00036	1.00003	0.99912	1.00022
ZAF	1.00023	1.00003	0.99945	1.00014

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

Table 7: 10% Increase of Transport Infrastructure in the US

ISOcode	Wage	Price Index	Absorption	Real Income
ARE	0.99880	0.99854	0.99929	1.00018
ARG	0.99869	0.99854	0.99960	1.00010
AUS	0.99871	0.99854	0.99955	1.00011
AUT	0.99890	0.99873	0.99954	1.00012
BEL	0.99893	0.99861	0.99916	1.00021
BRA	0.99874	0.99859	0.99961	1.00010
CAN	1.00023	0.99868	0.99591	1.00104
CHE	0.99895	0.99864	0.99920	1.00020
CHN	0.99918	0.99896	0.99942	1.00015
COL	0.99919	0.99872	0.99875	1.00032
DEU	0.99887	0.99868	0.99949	1.00013
EGY	0.99899	0.99862	0.99900	1.00025
ESP	0.99883	0.99872	0.99972	1.00007
FRA	0.99884	0.99871	0.99965	1.00009
GBR	0.99891	0.99869	0.99942	1.00015
GRC	0.99880	0.99863	0.99953	1.00012
IDN	0.99891	0.99878	0.99964	1.00009
IND	0.99896	0.99879	0.99956	1.00011
IRN	0.99882	0.99872	0.99974	1.00007
ITA	0.99885	0.99875	0.99973	1.00007
JPN	0.99890	0.99874	0.99957	1.00011
KOR	0.99892	0.99861	0.99918	1.00021
MEX	1.00023	0.99876	0.99613	1.00098
MYS	0.99925	0.99870	0.99853	1.00037
NGA	1.00000	0.99905	0.99750	1.00064
NLD	0.99882	0.99864	0.99953	1.00012
PAK	0.99921	0.99894	0.99929	1.00018
PHL	0.99909	0.99860	0.99870	1.00033
POL	0.99877	0.99869	0.99981	1.00005
RUS	0.99883	0.99873	0.99974	1.00006
SAU	0.99914	0.99865	0.99871	1.00033
SWE	0.99890	0.99871	0.99951	1.00012
THA	0.99911	0.99873	0.99900	1.00025
TUR	0.99879	0.99869	0.99972	1.00007
TWN	0.99914	0.99863	0.99867	1.00034
UKR	0.99883	0.99871	0.99969	1.00008
USA	1.04515	0.98660	1.00044	1.03938
VEN	1.00037	0.99945	0.99758	1.00062
XTW	0.99905	0.99870	0.99906	1.00024
ZAF	0.99887	0.99872	0.99959	1.00010

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

PPML with pair fixed effect.

Table 8: Augmented Gravity Estimations with Infrastructure of Exporter side

	(1) EXP_OLS	(2) IMP_OLS	(3) EXP_PPML	(4) IMP_PPML
Distance	-1.333***	-1.333***	-0.641***	-0.641***
	(0.067)	(0.067)	(0.063)	(0.063)
Contiguity	0.318***	0.318***	0.370***	0.370***
	(0.105)	(0.105)	(0.082)	(0.082)
language	0.337***	0.337***	0.470***	0.470***
	(0.122)	(0.122)	(0.106)	(0.106)
Colony	0.053	0.053	-0.129	-0.129
	(0.131)	(0.131)	(0.111)	(0.111)
RTA	0.428***	0.428***	0.154	0.154
	(0.102)	(0.102)	(0.136)	(0.136)
Border	-7.355***	-7.355***	-7.271***	-7.271***
	(1.307)	(1.307)	(0.737)	(0.737)
IMP INF×Border		0.284**		0.218***
		(0.112)		(0.047)
EXP INF×Border	0.284**		0.218***	
	(0.112)		(0.047)	
WGI index	0.928***	0.928***	0.711***	0.711***
	(0.217)	(0.217)	(0.080)	(0.080)
_cons	19.330***	19.328***	19.399***	19.399***
	(0.413)	(0.413)	(0.400)	(0.400)
N	14536	14536	14539	14539
R^2	0.914	0.914		
pseudo R^2			0.995	0.995
Pair Fixed Effect	No	No	No	No

Note: Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
The dependent variable is whether infrastructure proxy of importer side or exporter side. The 7th row shows results for infrastructure proxy of importer side and the 8th row shows results for infrastructure proxy of exporter side. WGI index is the institution quality index from World Bank's World Governance Indicator database. These results are not the same as the results in Table2, because we drop some observations with missing value of either infrastructure proxy of exporter or importer side and estimate with observations which have both of them. The main goal of this exercise is to verify our analysis in the end of 3.

7.2 Augmented Gravity Estimations with Sub-indices of Infrastructure

Table 9 and 10 showcase the results when we estimate with sub-indices in Donaubauer et al. (2016b) 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 2 and 3 showing our main results are not driven by the aggregation of sub-indices.

Table 9: PPML Estimations with Sub-indices of Infrastructure

	(1) PPML	(2) PPML	(3) PPML	(4) PPML
Distance	-0.551***	-0.618***	-0.623***	-0.567***
	(0.043)	(0.048)	(0.049)	(0.041)
Contiguity	0.362***	0.381***	0.421***	0.524***
	(0.075)	(0.070)	(0.075)	(0.077)
Language	0.505***	0.584***	0.541***	0.409***
	(0.063)	(0.062)	(0.066)	(0.069)
Colony	-0.025	-0.061	-0.069	-0.077
	(0.091)	(0.092)	(0.094)	(0.108)
RTA	0.285***	0.053	0.036	0.200***
	(0.066)	(0.076)	(0.077)	(0.068)
Border	-4.340***	-3.939***	-3.734***	-4.411***
	(0.156)	(0.179)	(0.173)	(0.141)
INF_TRAN	0.490***			
	(0.053)			
INF_COM		0.454***		
		(0.055)		
INF_EN			0.027	
			(0.058)	
INF_FIN				0.661***
				(0.061)
WGI index	-0.197***	-0.248***	0.190**	-0.153**
	(0.075)	(0.079)	(0.079)	(0.065)
_cons	18.285***	18.702***	18.745***	18.392***
	(0.273)	(0.301)	(0.306)	(0.256)
N	291760	300660	247450	277577
R^2				
pseudo R^2	0.989	0.988	0.988	0.989
Pair Fixed Effect	No	No	No	No

Note: Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Dependent variables from the 7th to the 10th row are sub-indices of the infrastructure index constructed by Donaubauer et al. (2016b). They are transportation, communication, energy, and financial infrastructure, respectly.

7.3 The Effect of Institution

Table 11 displays that the coefficients of the proxy of infrastructure are significantly positive and similar to the ones in Table 2 dissipating the concern that infrastructure can easily succumb to omitted variable biases. On the contrary, Table 12 reports the

Table 10: Pair Fixed Effect OLS Estimations with Sub-indices of Infrastructure

	(1) Pair_OLS	(2) Pair_OLS	(3) Pair_OLS	(4) Pair_OLS
RTA	0.110***	0.121***	0.117***	0.107***
	(0.038)	(0.037)	(0.039)	(0.038)
Border				
INF_TRAN	0.002			
	(0.104)			
INF_COM		0.183		
		(0.138)		
INF_EN			0.717***	
			(0.208)	
INF_FIN				0.207*
				(0.110)
WGI index	0.445	0.439	0.465	0.275
	(0.291)	(0.277)	(0.333)	(0.299)
_cons	0.679***	0.606***	0.800***	0.666***
	(0.023)	(0.030)	(0.046)	(0.033)
N	219258	225188	193564	212073
R^2	0.864	0.862	0.876	0.868
Pair Fixed Effect	Yes	Yes	Yes	Yes

Note: Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Dependent variables from the 3rd to the 6th row are sub-indices of the infrastructure index constructed by Donaubauer et al. (2016b). They are transportation, communication, energy, and financial infrastructure, respectly.

estimates when only institution is included and their signs are unanimously positive which conflict with some of the estimates when institution is jointly estimated with infrastructure such as Table 3. These results show that the coefficient of WGI index is not stable in estimations we conduct. The method in Beverelli et al. (2018) can lead to an omitted variable bias, while the estimation with infrastructure as the only dependent variable has higher immunity.

Table 11: Estimations with Instituition

	(1) OLS	(2) PPML	(3) Pair_PPML
Distance	-1.282***	-0.572***	
	(0.053)	(0.043)	
Contiguity	0.202	0.452***	
	(0.151)	(0.080)	
Language	0.394***	0.490***	
	(0.091)	(0.087)	
Colony	0.955***	-0.025	
	(0.120)	(0.099)	
RTA	0.419***	0.306***	0.253***
	(0.070)	(0.084)	(0.041)
Border	-9.535***	-7.341***	
	(1.592)	(0.573)	
INF×Border	0.542***	0.272***	0.534***
	(0.126)	(0.039)	(0.155)
_cons	15.769***	18.789***	13.002***
	(0.401)	(0.270)	(0.391)
N	89437	99867	99396
R^2	0.818		
pseudo R^2		0.992	0.998
Pair Fixed Effect	No	No	Yes

Note: Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
The dependent variable is the log of total length of the paved road and the railroad.

7.4 Alternative Proxies for Infrastructure

To account for the intrinsic heterogeneity in the country size, in Table 13, 14, and 15, we present the results which we use the paved road, railroad, and infrastructure density, as the proxy of infrastructure, respectively. Table 13 and 15 imply that results similar to our main result can be obtained when we choose the length of paved road and the infrastructure density. However, Table 14 reports a negative coefficient of the length of railroad from PPML estimation with symmetric pair fixed effect. This might come from lack of variations in the railroad length variable. A sensible explanation is that the

Table 12: Augmented Gravity Estimations with Institution

	(1) OLS	(2) PPML	(3) Pair_PPML
Distance	-1.558***	-0.638***	
	(0.017)	(0.050)	
Contiguity	0.732***	0.416***	
	(0.089)	(0.081)	
language	0.734***	0.496***	
	(0.034)	(0.070)	
Colony	0.848***	-0.049	
	(0.087)	(0.093)	
RTA	0.533***	0.027	0.242***
	(0.028)	(0.075)	(0.057)
Border	-3.730***	-3.503***	
	(0.196)	(0.190)	
WGI index	1.686***	0.116*	0.193
	(0.213)	(0.063)	(0.166)
_cons	17.353***	18.746***	13.761***
	(0.218)	(0.312)	(0.033)
N	413895	555720	554782
R^2	0.742		
pseudo R^2		0.983	0.996
Pair Fixed Effect	No	No	Yes

Note: Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
The dependent variable is the WGI index which is the institution quality index from World Bank's World Governance Indicator database.

geomorphic feature of some countries is not suitable to have more railroads built, whereas paved road is more adaptable to various geomorphic features.

Table 13: Augmented Gravity Estimations with Paved Road

	(1) OLS	(2) PPML	(3) Pair_OLS	(4) Pair_PPML
Distance	-1.269***	-0.586***		
	(0.048)	(0.044)		
Contiguity	0.152	0.412***		
	(0.152)	(0.075)		
Language	0.409***	0.411***		
	(0.081)	(0.084)		
Colony	1.047***	0.013		
	(0.121)	(0.095)		
RTA	0.414***	0.285***	0.071	0.272***
	(0.066)	(0.084)	(0.048)	(0.048)
Border	-8.599***	-7.225***		
	(1.331)	(0.551)		
INF_paved	0.401***	0.244***	0.800	0.716***
	(0.112)	(0.037)	(1.106)	(0.186)
WGI index	0.843***	0.280***	0.726	-0.429***
	(0.262)	(0.070)	(0.606)	(0.126)
_cons	15.551***	18.794***	-8.193	12.496***
	(0.355)	(0.275)	(13.190)	(0.496)
N	102177	114340	101865	113610
R^2	0.814		0.922	
pseudo R^2		0.992		0.998
Pair Fixed Effect	No	No	Yes	Yes

Note: Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
The dependent variable is log of the length of the paved road. WGI index is the institution quality index from World Bank's World Governance Indicator database.

Table 14: Augmented Gravity Estimations with Railroad

	(1) OLS	(2) PPML	(3) Pair_OLS	(4) Pair_PPML
Distance	-1.286***	-0.620***		
	(0.054)	(0.051)		
Contiguity	0.199	0.412***		
	(0.151)	(0.075)		
Language	0.393***	0.502***		
	(0.091)	(0.088)		
Colony	0.956***	-0.042		
	(0.120)	(0.099)		
RTA	0.409***	0.150	0.079	0.269***
	(0.070)	(0.102)	(0.052)	(0.042)
Border	-8.257***	-6.178***		
	(1.116)	(0.570)		
INF_rail	0.484***	0.216***	0.232	-0.163
	(0.115)	(0.047)	(1.103)	(0.141)
WGI index	1.002***	0.415***	0.601	-0.693***
	(0.275)	(0.074)	(0.597)	(0.124)
_cons	15.772***	19.097***	-0.673	14.808***
	(0.397)	(0.320)	(9.550)	(0.275)
N	89437	99867	89261	99396
R^2	0.819		0.925	
pseudo R^2		0.992		0.998
Fixe Effect	No	No	Yes	Yes

Note: Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
The dependent variable is log of the length of the railroad. WGI index is the institution quality index from World Bank's World Governance Indicator database.

Table 15: Augmented Gravity Estimations with Total Length Density

	(1) OLS	(2) PPML	(3) Pair_OLS	(4) Pair_PPML
Distance	-1.287***	-0.597***		
	(0.054)	(0.046)		
Contiguity	0.199	0.441***		
	(0.151)	(0.078)		
Language_ethno	0.390***	0.458***		
	(0.091)	(0.088)		
Colony	0.956***	-0.036		
	(0.120)	(0.099)		
RTA	0.411***	0.251***	0.079	0.262***
	(0.070)	(0.092)	(0.052)	(0.040)
Border	-9.662***	-7.471***		
	(1.212)	(0.571)		
Total Length Density	0.490***	0.260***	0.039	0.369**
	(0.101)	(0.038)	(1.205)	(0.184)
WGI index	0.808***	0.326***	0.602	-0.607***
	(0.259)	(0.071)	(0.608)	(0.143)
_cons	15.776***	18.945***	0.874	13.552***
	(0.393)	(0.293)	(14.468)	(0.483)
N	89437	99867	89261	99396
R^2	0.819		0.925	
pseudo R^2		0.992		0.998
Pair Fixed Effect	No	No	Yes	Yes

Note: Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
The dependent variable is log of total length of the paved road and the railroad per area. WGI index is the institution quality index from World Bank's World Governance Indicator database.

8 Policy Implication

The debate and discussion about infrastructure investment have been raging, since China proposed the "One Belt One Road Initiative". This world, where multilateralism has been on the retreat, has been under the crossfire of numerous rhetorics pushing the infrastructure investment to the center stage. Though whether the true geopolitical intentions are being camouflaged behind those rhetorics is still an open question, their economic arguments actually contain noteworthy academic and research values. Some raise the awareness of the possible insolvency of financing those infrastructure projects; some champion the infrastructure investment emphasizing the potential investment payoff in the future and its facility of spillover knowledge and know how; some endorse it because of the jobs it creates.

Several years have passed. It seems that this world is gradually taking a more positive attitude toward infrastructure investment and its benefits are more or less summarized into the following two categories—job creations and investment payoff, which are epitomized by the following two quotes from the speech of Joe Biden, the 46th President of the United States of America:

> "It creates jobs to upgrade our transportation infrastructure. Jobs modernizing our roads, bridges, highways. Jobs building ports and airports, rail corridors, transit lines."

> "Independent experts estimate the American Jobs Plan will add millions of jobs and trillions of dollars to economic growth in the years to come."

Though the above two pros for infrastructure are well established in economics literature, few literature discusses infrastructure under the international trade setting. It is understandable because of an econometric difficulty mentioned in previous sections—perfect multicollinearity. The infrastructure affects both internal trade and international trade. Fortunately, equipped with the empirical methodology devised by Beverelli et al. (2018) and the theoretical model built in this paper, this paper is the first paper to arrive at the conclusion that the infrastructure boosts international trade more than internal trade. Facilitated with the results from growth literature, this paper can draw the conclusion that infrastructure boosts trade openness, increasing both export and import. Further quantitative counterfactual exercise shows that 10% transportation infrastructure increase induces sizable 3.9% increase in real income. Although the unilateral infrastructure improvement in one country generally benefits all its trading partners, the benefits concentrate on the country itself and it reaps more than 95% of the total gains. All those results suggest infrastructure development should be in the toolkit of the policymakers ready for appropriate use and the policymakers should be aware of and cautious about the impact on trade openness when deciding on infrastructure spending. Trade openness has preponderant ramifications. Trade openness is very closely related to the absorption rate, one of the sufficient statistics of gains from trade. Besides, international finance literature indicates that trade openness could potentially affect cross-border capital flow which in turn affects the economic growth of a particular country.

9 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 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 non-negative impact on internal trade. Further quantitative analysis shows that 10% transportation infrastructure increase induces sizable 3.9% 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 boarder setting and assess the relative importance of several interesting factors in shaping the world trade flows and real income changes in a comparative way.

Acknowledgements

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

References

Agnosteva, D. E., Anderson, J. E., and Yotov, Y. V. (2014). Intra-national Trade Costs: Measurement and Aggregation. *NBER Working Paper*.

Allen, T. and Arkolakis, C. (2014). Trade and the topography of the spatial economy. *Quarterly Journal of Economics*, 129(3):1085–1139.

Alvarez, F. and Lucas, R. E. (2007). General equilibrium analysis of the Eaton-Kortum model of international trade. *Journal of Monetary Economics*, 54(6):1726–1768.

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

Arkolakis, C., Costinot, A., and Rodríguez-Clare, A. (2012). New trade models, same old gains? *American Economic Review*, 102(1):94–130.

Atkin, D. and Donaldson, D. (2017). Who's Getting Globalized? The Size and Implications of Intra-National Trade Costs. *SSRN Electronic Journal*.

Beverelli, C., Keck, A., Larch, M., and Yotov, Y. V. (2018). Institutions, Trade and Development: A Quantitative Analysis.

Cavallo, A., Feenstra, R. C., and Inklaar, R. (2021). Product Variety, the Cost of Living and Welfare Across Countries.

Coughlin, C. C. and Novy, D. (2021). ESTIMATING BORDER EFFECTS: THE IMPACT OF SPATIAL AGGREGATION. *International Economic Review*.

Dekle, R., Eaton, J., and Kortum, S. (2008). Global rebalancing with gravity: Measuring the burden of adjustment. *IMF Staff Papers*, 55(3):511–540.

Demir, F. and Hu, C. (2020). Destination institutions, firm heterogeneity and exporter dynamics: empirical evidence from China. *Review of World Economics*, 156(1):183–217.

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

Donaldson, D. and Hornbeck, R. (2016). Railroads and American economic growth: A "market access" approach. *Quarterly Journal of Economics*, 131(2):799–858.

Donaubauer, J., Glas, A., Meyer, B. E., and Nunnenkamp, P. (2018). Disentangling the impact of infrastructure on trade using a new index of infrastructure. *Review of World Economics*, 154(4):745–784.

Donaubauer, J., Glas, A., and Nunnenkamp, P. (2015). Infrastructure and trade: A gravity analysis for major trade categories using a new index of infrastructure. Kiel Working Paper 2016, Kiel.

Donaubauer, J., Meyer, B., and Nunnenkamp, P. (2016a). Aid, Infrastructure, and FDI: Assessing the Transmission Channel with a New Index of Infrastructure. *World Development*, 78:230–245.

Donaubauer, J., Meyer, B. E., and Nunnenkamp, P. (2016b). A New Global Index of Infrastructure: Construction, Rankings and Applications. *World Economy*, 39(2):236–259.

Eaton, J. and Kortum, S. (2002). Technology, geography, and trade. *Econometrica*, 70(5):1741–1779.

Fally, T. (2015). Structural gravity and fixed effects. *Journal of International Economics*, 97(1):76–85.

Feenstra, R. C., Inklaar, R., and Timmer, M. P. (2015). The next generation of the penn world table. *American Economic Review*, 105(10):3150–3182.

Francois, J. and Manchin, M. (2013). Institutions, Infrastructure, and Trade. *World Development*, 46:165–175.

Han, Z. (2020). Nonlinear Pricing in Transport Industry and the Gains from Trade.

Han, Z. (2021). Heterogeneous Internal Trade Cost and Its Implications in Trade.

Head, K. and Mayer, T. (2014). Gravity Equations: Workhorse,Toolkit, and Cookbook. In *Handbook of International Economics*, volume 4, pages 131–195. Elsevier B.V.

Heid, B., Larch, M., and Yotov, Y. V. (2021). Estimating the effects of non-discriminatory trade policies within structural gravity models. *Canadian Journal of Economics*, 54(1):376–409.

Kerem Coşar, A. and Fajgelbaum, P. D. (2016). Internal geography, international trade, and regional specialization. *American Economic Journal: Microeconomics*, 8(1):24–56.

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

Li, B. and Ma, L. (2021). JUE insight: Migration, transportation infrastructure, and the spatial transmission of COVID-19 in China. *Journal of Urban Economics*.

Manova, K. (2013). Credit constraints, heterogeneous firms, and international trade. *Review of Economic Studies*, 80(2):711–744.

Melitz, M. J. and Ottaviano, G. I. (2008). Market size, trade, and productivity. *Review of Economic Studies*, 75(1):295–316.

Nunn, N. and Trefler, D. (2014). Domestic institutions as a source of comparative advantage. In *Handbook of International Economics*, volume 4, pages 263–315. Elsevier B.V.

Ramondo, N., Rodríguez-Clare, A., and Saborío-Rodríguez, M. (2016). Trade, domestic frictions, and scale effects. *American Economic Review*, 106(10):3159–3184.

Rehman, F. U., Noman, A. A., and Ding, Y. (2020). Does infrastructure increase exports and reduce trade deficit? Evidence from selected South Asian countries using a new Global Infrastructure Index. *Journal of Economic Structures*, 9(1):10.

Samadi, A. and Alipourian, M. (2021). *Measuring Institutional Quality: A Review*, pages 143–171.

Santamaria, M. A., Ventura, J., and Yesilbayraktar, U. (2020). Borders Within Europe.

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

Simonovska, I. and Waugh, M. E. (2014). The elasticity of trade: Estimates and evidence. *Journal of International Economics*, 92(1):34–50.

Tang, H. (2012). Labor market institutions, firm-specific skills, and trade patterns. *Journal of International Economics*, 87(2):337–351.

Waugh, M. E. (2010). International trade and income differences. *American Economic Review*, 100(5):2093–2124.

Yotov, Y. V. (2012). A simple solution to the distance puzzle in international trade. *Economics Letters*, 117(3):794–798.

Appendices

A 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:

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

where the first term on the right denotes the demand of 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:

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

The above two equations combined imply the following

$$w_i L_i = \sum_j \pi_{ij} w_j L_j. \tag{23}$$

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:

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

the trade share formula:

$$\pi_{ij} = \begin{cases} \frac{T_i (c_i t_{i \cdot} \tau_{ij} t_{\cdot j})^{-\theta}}{\sum_{s \neq j} \left[T_s (c_s t_{s \cdot} \tau_{sj} t_{\cdot j})^{-\theta} \right] + T_j (c_j t_{jj})^{-\theta}} & \forall i \neq j \\ \frac{T_i (c_i t_{ii})^{-\theta}}{\sum_{s \neq j} \left[T_s (c_s t_{s \cdot} \tau_{sj} t_{\cdot j})^{-\theta} \right] + T_j (c_j t_{jj})^{-\theta}} & i = j \end{cases}. \tag{26}$$

Given the variable set $\{T_i, L_i, t'_{i \cdot}, t'_{\cdot j}\}$, equations (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 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:

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

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

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

$$w_i' L_i = \sum_j \pi_{ij}' w_j' L_j, \tag{30}$$

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

B Computation Algorithm

1. Using equation (30) and factual values of π_{ij} to obtain factual values of $w_i L_i$,
2. Given initial guess of $\hat{w}_i$, equation (27) and (28) together can solve both $\hat{c}_i$ and $\hat{P}_i$.
3. After knowing both $\hat{c}_i$ and $\hat{P}_i$, equation (29) generates value of $\hat{\pi}_{ij}$.
4. Using $\hat{\pi}_{ij}$ and factual values of π_{ij} to construct counterfactual π_{ij}' and pluging back the values of π_{ij}' into equation (30) to obtain counterfactual values of $w_i' L_i$.[13]
5. Dividing the counterfactual values $w_i' L_i$ by the factual values $w_i L_i$ 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.

C 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 data is that it has well-regularized

[13] 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.

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

The original GTAP data base consists of 129 regions and 57 sectors for year 2007. To circumvent the computation limits and ensure bilateral trade data has nonzero entries, 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 16. 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 sectors. The resulting internal trade flows resemble more closely to the gross expenditure used in the theoretical framework, rather than a value-added term.

Table 16: Regions Concordance

ISOCode	Country Name	ISOCode	Country 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