Institute for Economic Studies, Keio University

Keio-IES Discussion Paper Series

The Three Puzzles in the Gravity Model of International Trade

笹原 彰

2026 年 2 月 20 日
DP2026-002
https://ies.keio.ac.jp/publications/27245/

Keio University

1858
CALAMVS GLADIO FORTIOR

Institute for Economic Studies, Keio University
2-15-45 Mita, Minato-ku, Tokyo 108-8345, Japan
ies-office-group@keio.jp
20 February, 2026

The Three Puzzles in the Gravity Model of International Trade
笹原 彰
IES Keio DP2026-002
2026 年 2 月 20 日
JEL Classification: F14, F43, O47
キーワード: The gravity model of trade, the Glick-Rose puzzle, the border puzzle, the missing globalization puzzle

【要旨】

This article provides an overview of three puzzles related to the gravity model of trade---the Glick–Rose puzzle, the border puzzle, and the missing globalization puzzle. It summarizes recent developments in estimation methods, including the shift from cross-sectional to panel settings, the inclusion of appropriate fixed effects, and the shift from log-linear specifications to nonlinear approaches.

笹原 彰
慶應義塾大学
sasahara@keio.jp

謝辞：This paper is an English version of the Japanese article of the same title and author, published in the Mita Journal of Economics (Mita Gakkai Zasshi), Vol. 117, No. 2, October 2024. Permission to translate the original article and make it publicly available is granted by the Keio Economic Society. The author would like to thank Kozo Kiyota, Toshiyuki Matsuura, and an anonymous reviewer for their constructive feedback. The author would like to thank Alexandra Sinulingga, Enrique Magnaye, and the Keio Economic Society for their assistance in preparing this English version. The author is grateful for the financial support from JSPS Grants-in-Aid KAKENHI under grant numbers 21K1329, 21H00713, and 22H00063. All errors are the author’s responsibility.

The Three Puzzles in the Gravity Model of International Trade*

Akira Sasahara†
Keio University

February 2026

Abstract

This article provides an overview of three puzzles related to the gravity model of trade—the Glick–Rose puzzle, the border puzzle, and the missing globalization puzzle. It summarizes recent developments in estimation methods, including the shift from cross-sectional to panel settings, the inclusion of appropriate fixed effects, and the shift from log-linear specifications to nonlinear approaches.

Keywords: The gravity model of trade, the Glick-Rose puzzle, the border puzzle, the missing globalization puzzle
JEL codes: F14, F43, O47

*This paper is an English version of the Japanese article of the same title and author, published in the *Mita Journal of Economics* (*Mita Gakkai Zasshi*), Vol. 117, No. 2, October 2024. Permission to translate the original article and make it publicly available is granted by the Keio Economic Society. The author would like to thank Kozo Kiyota, Toshiyuki Matsuura, and an anonymous reviewer for their constructive feedback. The author would like to thank Alexandra Sinulingga, Enrique Magnaye, and the Keio Economic Society for their assistance in preparing this English version. The author is grateful for the financial support from JSPS Grants-in-Aid KAKENHI under grant numbers 21K1329, 21H00713, and 22H00063. All errors are the author's responsibility.

†Faculty of Economics, Keio University. E-mail: `sasahara@keio.jp`

1 Introduction

The gravity model is commonly used in empirical studies in international trade for several reasons. First, they represent the simple idea of explaining the trade between two countries as a function of their GDP and the distance between them. Second, they perform well to explain actual trade flows. Third, they are easily applicable to examine on various topics such as the effects of financial crises or free trade agreements (FTAs) on international trade (Head and Mayer, 2014). Hence, trade economists community refers to them as the "workhorse model" (Head and Mayer, 2014). Specifically, the value of exports from country i to country j, X_{ij} can be written as the following equation:

$$X_{ij} = A \frac{GDP_i \times GDP_j}{dist_{ij}},$$

where A is a constant term, GDP_i denotes the GDP of country i, and $dist_{ij}$ denotes the distance between the two countries. When estimating this equation, it is log-linearized by taking the log. Replacing A with β_0, adding βs as the coefficients for the variables, and adding an error term u_{ij} leads to the following expression:

$$\ln(X_{ij}) = \beta_0 + \beta_1 \ln(GDP_i) + \beta_2 \ln(GDP_j) + \beta_3 \ln(dist_{ij}) + u_{ij},$$

where $\beta_1 = \beta_2 = 1$ and $\beta_3 = -1$ theoretically.

This paper discusses three puzzles that trade economists have faced while working with the gravity model. We also summarize the gravity model's development. The discussion focuses on how the analysis evolved from cross-sectional to panel approaches, how the inclusion of appropriate fixed effects became prevalent, and how Poisson Pseudo Maximum Likelihood (PPML) estimation replaced log-linear modeling as the mainstream method.[1]

The remainder of this paper is organized as follows. Section 2 discusses the Glick-Rose puzzle, currency unions' surprisingly large impact on international trade. Section 3 discusses the border puzzle, national borders' surprisingly large negative impact on international trade. Section 4 discusses the missing globalization puzzle, the unexpectedly stable distance coefficient between 1975 and 2000 even though the importance of distance should have decreased due to globalization. Section 5 offers a concluding remark. Appendix A presents a supplemental explanation on the interpretation of the coefficient of a dummy variable, and Appendix B presents Stata commands.

2 The Glick-Rose puzzle

2.1 Rose (2000)

What is the effect of employing a common currency between countries on international trade across the countries? We might expect that it increases international trade by reducing exchange rate volatility. Nevertheless, the magnitude of the trade facilitation effect remains a relevant research question.

To answer this question, Rose (2000) estimates the following regression equation:[2]

$$\ln(T_{ij}) = \beta_0 + \beta_1 \ln(GDP_i \times GDP_j) + \beta_2 \ln(dist_{ij}) + \beta_3 D_{ij}^{CU} + \mathbf{X}_{ij}' \boldsymbol{\beta}_4 + u_{ij}, \quad (1)$$

where $T_{ij} = X_{ij} + X_{ji}$, and X_{ij} denotes the value of exports from country i to country j. The variable

[1] For comprehensive literature reviews of studies using the gravity model, see Anderson (2011), Head and Mayer (2014), Yotov (2022), and Larch and Yotov (2024). For comprehensive guidebooks on estimating gravity models, see Shepherd (2016), Shepherd (2019), and Yotov et al. (2016).

[2] The parameters should be written with different letters since the regression equations discussed in this paper differ from one another. For the sake of simplicity, however, all the regression equations discussed use the same parameter and variable notations.

D_{ij}^{CU} is a dummy variable that takes unity if country i and country j use a common currency and zero otherwise, and $\mathbf{X}_{ij}$ is a vector containing control variables such as adjacency and common language dummies. Of particular interest is the coefficient β_3 that captures the impact of using a common currency on international trade. Equation (1) is estimated for each year with data from 186 countries. The estimates are at 5-year intervals for the period 1970–1990.

Rose (2000)'s estimates of the coefficient β_3 are summarized in Table 1. It shows that the estimated coefficient is 0.87 in 1970. The coefficient increases over time and reaches 1.51 in 1990. The estimate using data from all periods, shown in the rightmost column, is 1.21. Given that the dependent variable is the log of trade value, and the variable of interest D_{ij}^{CU} is a dummy variable, the exact impact of a common currency on the value of trade is $e^{1.21} = 3.35$, which can be interpreted as follows: Trade between countries sharing a common currency is 3.35 times greater than trade between countries that do not. Or, countries using a common currency have 235% greater trade than countries not using a common currency.[3]

This is a much larger number than we would expect. Rose (2000) explains possible reasons for the appearance of such extreme values, such as "Perhaps hedging exchange rate risk is much more difficult than commonly believed." However, he admits he does not have a definitive answer in his blunt statement, "I don't know." The puzzle has prompted many follow-up studies because the estimated impact of a common currency on international trade has been much larger than researchers expected.

Table 1: Rose (2000)'s estimates of the effects of currency unions on trade

	1970	1975	1980	1985	1990	All
Common currency dummy coefficient	0.87	1.28	1.09	1.40	1.51	1.21
	(0.43)	(0.41)	(0.26)	(0.27)	(0.27)	(0.14)
Sample size	4,052	4,474	5,092	5,091	4,239	22,948

Note: This table comes from Table 2 in Rose (2000). Standard errors are in parentheses.

Are there any econometric problems with Rose (2000)'s regression equation? The first problem is that the five regressions for 1970, 1975, 1980, 1985, and 1990 are based on cross-sectional data only and thus do not control for the effects of unobserved time-invariant variables. Although Rose (2000) includes several time-invariant variables, such as the adjacency and common language dummies, there are many other potential time-invariant determinants, such as labor market institutions and geographical factors. The regression equation in the rightmost column of Table 1 pools data for all time periods, but it does not include country-pair fixed effects although it could have done so.

2.2 Estimation with country-pair fixed effects

Rose (2000) seemed to have recognized that the coefficients may have been overestimated due to the failure to account for fixed effects. Therefore, his own follow-up study, Glick and Rose (2002) report their estimated coefficients with fixed effects. Specifically, they estimate the following regression equation:

$$\ln\left(T_{ijt}\right) = \beta_0 + \beta_1 \ln(GDP_{it} \times GDP_{jt}) + \beta_2 D_{ijt}^{CU} + \mathbf{X}'_{ijt}\boldsymbol{\beta}_3 + \phi_{ij} + u_{ijt}, \tag{2}$$

where ϕ_{ij} denotes country-pair fixed effect.[4] Since the variable $\ln(dist_{ij})$ is absorbed by the country-pair fixed effects, we cannot obtain an estimate of the distance coefficient. However, this coefficient is our focus because β_2, the coefficient of the common currency dummy D_{ijt}^{CU}, is of particular interest. Additionally, the subscript t, representing time (years), appears in the regression equation due to the addition of time

[3]It could be calculated as $e^{1.21} - 1 = 2.35$ and expressed as "The value of trade between countries using a common currency is 235% higher than the value of trade between countries not using a common currency." The reason the exponential function must be used here to calculate the effect is explained in Appendix A.

[4]"Country-pair fixed effects" are also sometimes expressed as "dyadic fixed effects".

series information. D_{ijt}^{CU} also has a t subscript, meaning that the status of whether country i and country j use a common currency can change over time. Equation (2) is estimated using data from 217 countries for the period 1948–1997.

Note that the interpretation of the coefficient changes depending on the inclusion of country-pair fixed effects. When country-pair fixed effects are not included, coefficient identification relies on cross-sectional variation in the data, measuring how much larger (or smaller) the trade value is for a pair of countries that use a common currency compared with a pair that does not. When country-pair fixed effects are included, coefficient identification relies on time-series variation in the data, measuring how much trade increases (or decreases) if a pair of countries began using a common currency, or how much it decreases (or increases) if they left a currency union.[5]

Table 2 summarizes Glick and Rose (2002)'s estimates. First, column (1) shows that estimation without country-pair fixed effects leads to a coefficient of 1.30, confirming the reproducibility of Rose (2000)'s results with Glick and Rose (2002)'s updated data. Column (2) shows that the inclusion of country-pair fixed effects reduces the value of the coefficient to 0.65. Since $e^{0.65} - 1 = 0.92$, it suggests that joining a currency union increases trade by 92%. This is almost half the size of Rose (2000)'s estimate. These results suggest that Rose (2000)'s estimates are biased since it does not include country-pair fixed effects.

Table 2: Glick and Rose (2002)'s estimates of the effects of currency unions on trade

	(1)	(2)
Common currency dummy coefficient	1.30	0.65
	(0.13)	(0.05)
Country-pair fixed effects		Yes
Sample size	219,558	
Number of country-pairs	11,178	

Note. This table comes from Tables 2 and 4 in Glick and Rose (2002). Standard errors are in parentheses.

2.3 "Exporter × year fixed effects" and "importer × year fixed effects"

A follow-up study, Glick and Rose (2016), points out three potential problems with Glick and Rose (2002). First, Glick and Rose (2002) includes 16 country-pairs that joined currency unions and 130 country-pairs that left currency unions during the sample period. Nevertheless, the approach implicitly assumes that joining and leaving a common currency increases or decreases trade by the same magnitude. It also implicitly assumes that all currency unions have the identical trade-creating effect. Second, because the sample period ends in 1997, the analysis does not include the Euro (the Economic and Monetary Union in Europe, EMU), which was introduced in 2002. Therefore, Glick and Rose (2016)'s analysis focuses specifically on the Euro's trade-creating effects. Third, Glick and Rose (2002)'s regression model includes neither "exporter × year fixed effects" nor "importer × year fixed effects."[6]

Regarding the first point, Glick and Rose (2016) show that the assumption that joining and leaving a currency union result in comparable changes in trade is largely correct. Therefore, the following discussion of the estimates focuses on the second and third points. Table 3 summarizes Glick and Rose (2016)'s estimates. Columns (1) and (2) use the same dataset as Glick and Rose (2002), confirming the results in their prior study. Columns (3) and (4) show results from running Glick and Rose (2002)'s original regression model using the dataset that is extended to 2013. The extended dataset leads to slightly

[5]When the dataset is a panel, estimations that do not include country-pair fixed effects identify coefficients using both cross-sectional and time-series variation. However, because cross-sectional variation in the data tends to be larger than time-series variation, the estimates essentially capture a cross-sectional relationship.

[6]If exporter (importer) × year fixed effects are not introduced, the estimation does account for the exporter's (importer's) multilateral resistance. This point is discussed in Section 3.

smaller coefficients, decreasing from 1.30 (column (1)) to 0.92 (column (3)), and decreasing from 0.65 (column (2)) to 0.63 (column (4)).

Table 3: Glick and Rose's (2016)'s estimates of the effects of currency unions on trade

	Glick-Rose (2002) data		Glick-Rose (2016) data			
Dependent variable	$\ln(T_{ij})$		$\ln(T_{ij})$		$\ln(X_{ij})$	
	(1)	(2)	(3)	(4)	(5)	(6)
Common currency dummy coefficient	1.30	0.65	0.92	0.63	0.51	0.34
	(0.13)	(0.05)	(0.09)	(0.07)	(0.02)	(0.02)
Country-pair fixed effects		Yes		Yes		Yes
Exporter × year fixed effects					Yes	Yes
Importer × year fixed effects					Yes	Yes
Sample size	219,558		426,935		879,794	
Number of country-pairs	11,178		14,801		33,886	
Number of country-year observations					22,438	
Sample period	1948–1997		1948–2013			

Note: This table comes from Tables 1, 2, and 5 in Glick and Rose (2016). Standard errors are in parentheses.

Columns (5) and (6) differ from Rose (2000) and Glick and Rose (2002) in terms of the data structure. In Rose (2000) and Glick and Rose (2002), each country-pair had one trade flow $X_{ijt} + X_{jit}$ per year. However, in Glick and Rose (2016), two-way trade, X_{ijt} and X_{jit}, are included in the dataset separately. Therefore, the sample size in Columns (5) and (6) is about twice as large as the sample in Columns (3) and (4). This data structure allows us include "exporter × year fixed effects" and "importer × year fixed effects," making it possible to control for all observable and unobservable macroeconomic conditions of exporters and importers.[7]

The regression equation is as follows:

$$\ln\left(X_{ijt}\right) = \beta_0 + \beta_1 D_{ijt}^{CU} + \mathbf{X}_{ijt}'\boldsymbol{\beta}_3 + \phi_{ij} + \phi_{it} + \phi_{jt} + u_{ijt}, \quad (3)$$

where $\mathbf{X}_{ijt}$ denotes a vector of time-variant controls, for example, FTA dummies and colonial relation dummies. The variable ϕ_{ij} is the country-pair fixed effect, ϕ_{it} denotes the exporter × year fixed effect, and ϕ_{jt} denotes the importer × year fixed effect.

Column (5) shows the results from estimating a model without country-pair fixed effects but with "exporter × year fixed effects" and "importer × year fixed effects." The coefficient is 0.51, indicating a slight decrease relative to the estimate in Column (4). Column (6) includes "exporter × year fixed effects" and "importer × year fixed effects" in addition to country-pair fixed effects. The coefficient is 0.34, showing a further decrease from the coefficient in Column (5).

Column (6)'s suggested trade creation effect is $e^{0.34} = 1.40$, implying that joining a currency union increases trade by 40%. This estimate seems plausible compared to Rose (2000)'s 235% trade creation effect. These results suggest that earlier studies' estimates of the effects of currency unions on trade may have been overstated due to the omission of appropriate fixed effects.

[7] The exporter dummies (and the importer dummies) are perfectly multicollinear with country-pair fixed effects. Therefore, the model cannot include these fixed effects jointly. However, by interacting the exporter dummies (and the importer dummies) with year dummies, these interacted dummies can be jointly included with country-pair fixed effects.

2.4 Heterogeneous trade creation effects of currency unions

Table 4 estimates the effects of the EMU and other currency unions separately to allow for heterogeneity in currency unions' trade creation effects.[8] The results presented in Columns (1) and (2) show that the estimated effect of other currency unions exceeds that of the EMU. The estimated coefficients for the non-EMU currency union dummies are 1.12 and 0.75, respectively, which can be interpreted as meaning that trade between non-EMU currency union countries is $e^{1.12} = 3.06$ times larger than trade between countries that do not use a common currency, and joining a non-EMU currency union increases trade by $e^{0.75} = 2.11$ times.

However, these estimates do not control for "exporter × year fixed effects" and "importer × year fixed effects." Column (3) shows the results after controlling for those fixed effects, and Column (4) shows the results after applying controls for country-pair fixed effects in addition to "exporter × year fixed effects" and "importer × year fixed effects." The results in Column (4), which may be the most reliable, show that joining the EMU increases trade by a factor of $e^{0.43} = 1.53$, and joining a non-EMU currency union increases trade by a factor of $e^{0.30} = 1.34$.

Rose (2017) surveys and meta-analyzes 45 relevant papers to examine why scholars' estimates of EMU trade creation effects vary widely, and it finds that excluding countries with smaller economies and lower income levels from the sample tends to lead to a downward bias in estimates of the coefficient of the currency union dummy. According to Rose (2017)'s estimates, trade increases by 54% by joining a currency union. This estimate is close to the estimate suggested in Rose and Stanley (2005)'s meta-analysis, 47%.

Table 4: Glick and Rose (2016)'s estimates of the effects of currency unions on trade

Dependent variable	$\ln(T_{ij})$		$\ln(X_{ij})$	
	(1)	(2)	(3)	(4)
EMU dummy coefficient	0.02	0.41	-0.65	0.43
	(0.08)	(0.05)	(0.03)	(0.02)
Non-EMU currency union dummy coefficient	1.12	0.75	0.76	0.30
	(0.11)	(0.10)	(0.02)	(0.03)
Country-pair fixed effects		Yes		Yes
Exporter × year fixed effects			Yes	Yes
Importer × year fixed effects			Yes	Yes
Sample size	426,935		879,794	
Number of country-pairs	14,801		33,886	
Number of country-year observations			22,438	
Sample period	1948–2013			

Note: This table comes from Tables 1, 2, and 5 in Glick and Rose (2016). Standard errors are in parentheses.

2.5 Adding country-pair dummies × trend

Campbell (2013) lists several problems in Glick and Rose (2002). First, the analysis does not control for the decline in economic ties between former colonial powers and their colonies since 1948. Second, there are currency allies that are not properly included in the analysis over all periods due to a lack of data for some variables. Third, countries in conflict or war are included in the analysis without controlling for those events' occurrence.

Estimating the Glick and Rose (2002) regression equation yields a coefficient of 0.65 for the currency union dummy, but this coefficient drops to 0.46 when the interaction term with the former British colony

[8] Glick and Rose (2016) also report separate estimates of eight non-EMU currency unions' effects, but the current paper discusses a simplified version.

dummy and the year trend are included. It further declines to 0.21 when countries at war or in conflict are excluded from the sample. It further declines to 0.11 when currency unions for which data are not available over the entire sample period are excluded. It further declines to −0.05 when the "country-pair dummies × year trend" is included. Campbell (2013) states that policymakers in countries facing the policy decision to join a currency union should not expect large trade creation effects.

2.6 PPML estimation

Larch, Wanner, et al. (2019) report their findings on revisiting currency unions' impact on trade volume using PPML instead of a log-linear model.[9] They justify PPML estimation by pointing out that estimating a log-linear model with ordinary least squares (OLS) does not ensure a lack of bias nor does the method ensure the consistency of the coefficients when heteroscedasticity is present in the data. Another advantage of using PPML is that the log-linear model cannot include observations with zero trade flows in the analysis, whereas PPML can account for zero trade flows. The regression equation can be written as follows:

$$X_{ijt} = \exp\left(\beta_0 + \beta_1 D_{ijt}^{CU} + \mathbf{X}_{ijt}\boldsymbol{\beta}_3 + \phi_{ij} + \phi_{it} + \phi_{jt}\right) \times u_{ijt}. \quad (4)$$

Larch, Wanner, et al. (2019) also use Cameron et al. (2011)'s multi-way clustering standard errors, which report standard errors that account for the correlation of error terms within the multiple clusters of exporting country, importing country, and year.

Table 5 summarizes Larch, Wanner, et al. (2019)'s estimation results. Columns (1) and (2) show the results of OLS estimation of the linear model, which replicate Glick and Rose (2016)'s results. Column (1) corresponds to column (6) in Table 3, and Column (2) corresponds to Column (4) in Table 4. Columns (3) and (4) show the results of estimating equation (4) using PPML. In Column (3), the coefficient is 0.13, which is $e^{0.13} - 1 = 0.138$, suggesting that currency union membership increases trade by about 14%. However, this coefficient is not statistically significant. Column (4) shows the results of estimating the effect of monetary unions separately for the EMU versus other monetary unions. The EMU coefficient is 0.030, which means that joining the EMU would increase trade by only 3%. However, this coefficient is also not statistically significant. For non-EMU currency unions, the coefficient is 0.70, so $e^{0.70} = 2.01$ and it is statistically significant, suggesting that currency unions roughly double trade.

Table 5: Larch et al. (2019)'s estimates of the effects of currency unions on trade

	Linear model		PPML	
	(1)	(2)	(3)	(4)
Common currency dummy's coefficient	0.343		0.130	
	(0.089)		(0.081)	
EMU dummy's coefficient		0.429		0.030
		(0.149)		(0.092)
Non-EMU currency union dummy's coefficient		0.298		0.700
		(0.097)		(0.172)
Country-pair fixed effects	Yes	Yes	Yes	Yes
Exporter × year fixed effects	Yes	Yes	Yes	Yes
Importer × year fixed effects	Yes	Yes	Yes	Yes
Sample size	877,736			
Number of countries	32,000			
Number of exporters / Number of importers	212 / 212			
Sample period	1948–2013			

Note: This table comes from Table 1 in Larch, Wanner, et al. (2019). Robust standard errors, clustered by exporting country, importing country, and year, are in parentheses.

[9] See Silva and Tenreyro (2006) for a discussion of the importance of using PPML in gravity model estimation in the international trade context.

In summary, by using a PPML methodology that considers heteroscedasticity, further controlling for country-pair fixed effects, "exporter × year fixed effects," and "importer × year fixed effects," and using robust standard errors that account for the correlation of error terms within three clusters of exporting country, importing country, and year, Larch, Wanner, et al. (2019) conclude that a currency union exerts a trade creation effect of only 14% and that joining the EMU yields almost no trade creation effect. Since estimation methods for the gravity model improved after 2000 and addressed the shortcomings of earlier approaches, the puzzle of the currency union's large trade-creating effect has gradually been resolved.

3 The border puzzle

3.1 McCallum (1995)

The border puzzle, first proposed by McCallum (1995), refers to the unexpectedly large negative effects of national borders on trade flows. McCallum (1995) documents that, amid the substantial trade liberalization of the 1980s and 1990s, there was a prevalent (mis)understanding that borders no longer mattered for global economic activity at the time. To examine whether this understanding was correct, he estimated a gravity model of trade to assess the effects of national borders on trade flows.

Using data on domestic and international trade between Canadian provinces and US states (hereafter, simply regions), McCallum (1995) estimated the following regression equation:

$$\ln (X_{ij}) = \beta_0 + \beta_1 \ln(GDP_i) + \beta_2 \ln(GDP_j) + \beta_3 \ln(dist_{ij}) + \beta_4 D_{ij}^{Canada} + u_{ij}, \tag{5}$$

where X_{ij} denotes the value of exports from region i to region j, GDP_i denotes GDP of region i, $dist_{ij}$ denotes the distance between region i and j, and D_{ij}^{Canada} is a dummy variable that takes unity if regions i and j are both located in Canada and zero otherwise. McCallum (1995) estimated equation (5) using 1988 data, and the coefficient of D_{ij}^{Canada} is estimated to be 3.09.[10] Since $e^{3.09} = 22$, it suggests that after controlling for GDP and the distances, Canada's intra-country trade is about 22 times larger than international trade between Canada and the US.

It is important to understand that the estimate is obtained by controlling for GDP and distance. Let us consider "trade between British Columbia and Alberta" and "trade between British Columbia and the state of Washington." Suppose that the three economies, British Columbia, Alberta, and Washington, have the same GDP. Suppose that the "distance between British Columbia and Alberta" and the "distance between British Columbia and Washington" are identical. In this case, the gravity model predicts that "trade between British Columbia and Alberta" and "trade between British Columbia and the state of Washington" would be the same. Nevertheless, the regression result suggested that the border between British Columbia and Washington reduced trade between the two regions, resulting in a substantial difference between intra-country trade within Canada and international trade between Canada and the US. I hope this illustrative example helps understand that the estimate $e^{3.09} = 21.98$ seems too large. McCallum (1995) does not clarify the reasons for this huge negative effect of national borders on trade.

3.2 Remoteness and multilateral resistance

Anderson and van Wincoop (2003) propose one solution to McCallum (1995)'s border puzzle. They focus on the size difference between the US and Canadian economies. As of 1990, the US had about 10 times greater GDP than Canada, making Canada a small country compared to the US. In such an environment, Canadian provinces have less domestic trading partners than US states do, making domestic

[10] See Column (2) of Table 1 in McCallum (1995).

trade with other Canadian provinces important. Anderson and van Wincoop (2003) describe the following thought experiment on pp. 177–178 in their article.

> Consider the following example without any reference to gravity equations and multilateral resistance variables. A small economy with two regions and a large economy with 100 regions engage in international trade. All regions have the same GDP. What matters here is not the number of regions, but the relative size of the two economies as measured by total GDP. We only introduce regions in this example because it is illustrative in the context of the U.S. states and Canadian provinces that are the focus of the empirical analysis. Under borderless trade, all regions sell one unit of one good to all 102 regions (including themselves). Now impose a barrier between the small and the large country, reducing trade between the two countries by 20 percent. Region 1 in the small country then reduces its exports to the large country by 20. It sells ten more goods to itself and ten more goods to region 2 in the small country. Trade between the two regions in the small country rises by a factor 11, while trade between two regions in the large country rises by a factor of only 1.004. This shows that even a small drop in international trade can lead to a very large increase in trade within a small country. Trade between the two regions in the small country is now 13.75 times trade between regions in both countries, while trade between two regions in the large country is only 1.255 times trade between regions in the two countries.

To theoretically demonstrate the implication of different country sizes on domestic trade flows, Anderson and van Wincoop (2003) show that a variable called "multilateral resistance" appears on the right-hand side of the gravity model:

$$X_{ij} = \frac{GDP_i \times GDP_j}{GDP^W} \left(\frac{t_{ij}}{P_i \times P_j} \right)^{1-\sigma},$$

where GDP^W denotes the world GDP, and t_{ij} denotes trade costs of exporting from region i to region j. The variable P_i denotes the multilateral resistance of region i—the constant elasticity of substitution (CES) price index derived from the maximization of a CES utility, which can be written as follows:

$$P_i = \left[\sum_k (\beta_k p_k t_{ik})^{1-\sigma} \right]^{\frac{1}{1-\sigma}}, \tag{6}$$

where β_k denotes the utility function's demand shift parameter, p_k denotes the price of variety produced in region k, and σ denotes the elasticity of substitution.

The following are three ways to estimate the gravity model by controlling for multilateral resistance:

1. Introducing the "remoteness," a proxy variable for multilateral resistance, into the regression equation.
2. Introducing the "multilateral resistance" in a manner that is consistent with theory, equation (6), into the regression equation.
3. Introducing appropriate fixed effects to absorb the effects of multilateral resistance.
   - 3.1 Introducing "exporter fixed effects" and "importer fixed effects" into the regression with cross-sectional data.
   - 3.2 Introducing "exporter × year fixed effects" and "importer × year fixed effects" into the regression with panel data.

We discuss these three approaches in the following sections.

3.2.1 Including remoteness

When the paper Anderson and van Wincoop (2003) was written, a variable called remoteness

$$Rem_i = \sum_{k \neq i} \frac{dist_{ik}}{GDP_k},$$

was commonly used. Therefore, Anderson and van Wincoop (2003) also present results from running a regression introducing the remoteness variable into McCallum (1995)'s original regression equation. The regression equation is as follows:

$$\begin{aligned} \ln(X_{ij}) &= \beta_0 + \beta_1 \ln(GDP_i) + \beta_2 \ln(GDP_j) + \beta_3 \ln(dist_{ij}) \\ &+ \beta_4 \ln(Rem_i) + \beta_5 \ln(Rem_j) + \beta_6 D_{ij}^{Canada} + \beta_7 D_{ij}^{USA} + u_{ij}. \end{aligned} \tag{7}$$

There are three major differences between McCallum (1995)'s and Anderson and van Wincoop (2003)'s models. First, McCallum (1995) does not introduce remoteness into the regression equation, whereas Anderson and van Wincoop (2003) do. Second, McCallum (1995) does not use US domestic trade data, whereas Anderson and van Wincoop (2003) include US domestic trade data in their dataset and add a US domestic trade dummy D_{ij}^{USA} to the regression equation. Third, McCallum (1995) uses 1988 data, whereas Anderson and van Wincoop (2003) use 1993 data.

Table 6 summarizes Anderson and van Wincoop (2003)'s results from running equation (7). Column (1) shows the results of excluding remoteness from the model to verify if the updated data produces different results from McCallum (1995). It shows that Canada's domestic trade is about 16 times international trade between Canada and the US, which is slightly smaller than McCallum (1995)'s number of 22 times. It also shows that US domestic trade is only 1.49 times that of international trade, clearly demonstrating national borders' asymmetric impacts on domestic trade in small and large economies. Column (2) shows the results from introducing the remoteness variable into the model. The border effect is slightly smaller than column (1). To be consistent with theory, columns (3) and (4) show results under the unity constraints for GDP coefficients. However, it leads to no significant change in the border effect.[11] These results suggest that the remoteness variable performs poorly to account for the multilateral resistance of the regions.

Table 6: Anderson and van Wincoop (2003)'s estimates with the remoteness variable

	(1)	(2)	(3)	(4)
exp(Coef. of the dummy for the Canadian domestic trade)	15.70	15.60	14.20	15.00
	(1.90)	(1.90)	(1.60)	(1.80)
exp(Coef. of the dummy for the US domestic trade)	1.49	1.38	1.62	1.42
	(0.85)	(0.07)	(0.09)	(0.08)
Remoteness		✓		✓
Unity restriction			✓	✓

Note: This table comes from Table 1 of Anderson and van Wincoop (2003). Standard errors are in parentheses.

[11] Practically, the unity restrictions can be introduced as follows. First, we set $\beta_1 = \beta_2 = 1$, and rewrite the gravity model as follows:

$$\ln X_{ij} = \beta_0 + \ln(GDP_i) + \ln(GDP_j) + \ln(dist_{ij}) + u_{ij},$$

where the explanatory variables other than GDP and distance are omitted. Moving GDP variables to the left-hand side and using logic formation to combine the X_{ij} and GDP variables into a single variable leads to the following equation:

$$\ln\left(\frac{X_{ij}}{GDP_i \times GDP_j}\right) = \beta_0 + \ln(dist_{ij}) + u_{ij}.$$

By using $\ln\left(\frac{X_{ij}}{GDP_i \times GDP_j}\right)$ as the dependent variable and removing the GDP variables from the right-hand side of the equation, the model introduces unity restrictions on the GDP coefficients.

Based on these results, Anderson and van Wincoop (2003) argue that although the remoteness variable was commonly used in the literature, it is "entirely disconnected from the theoretical model." They instead propose the multilateral resistance term, which is consistent with the theoretical framework. Head and Mayer (2014) also state that remoteness is only a proxy variable for multilateral resistance because distance alone cannot measure trade costs, and "proxy variables do not take the theory seriously enough."

3.2.2 Including theory-consistent multilateral resistance

Anderson and van Wincoop (2003) also perform an analysis with the "multilateral resistance consistent with theory." In general equilibrium, P_i cannot be written as a closed-form solution because it depends on multilateral resistance in other regions. Anderson and van Wincoop (2003) estimate the model's parameters from the data and solve P_i recursively to account for multilateral resistance in a manner consistent with theory. As summarized in Table 7, when multilateral resistance consistent with theory is considered, the Canadian border effect drops to about 11 in both the two-country and the multi-country models, whereas the US border effect rises slightly, compared to the estimates in Table 7, to about 2.2–2.6. Although we continue to observe larger border effects in Canada than in the US, we can correct for at least some of the bias in McCallum (1995)'s estimates.

Table 7: Anderson and van Wincoop (2003)'s border effects consistent with theory

	Canada	The US
Two-country model	10.50	2.56
	(1.16)	(0.13)
Multi-country model	10.70	2.24
	(1.06)	(0.12)

Note: This table comes from Table 5 of Anderson and van Wincoop (2003). Standard errors are in parentheses.

3.2.3 Including exporter fixed effects and importer fixed effects

Outside the border effects context, Head and Mayer (2014) recommend including "exporter fixed effects" and "importer fixed effects" when estimating a gravity model with cross-sectional data to control for multilateral resistance.[12] With a panel dataset, including "exporter × year fixed effects" and "importer × year fixed effects" would control for (potentially) time-varying multilateral resistance terms.

Baldwin and Taglioni (2007) describe estimation without controlling for multilateral resistance as the "gold medal of classic gravity model mistakes." Cipollina and Salvatici (2010) avoid this mistake by including exporter fixed effects and importer fixed effects in their estimation with cross-sectional data. They include country-pair fixed effects in their estimation with panel data. They state that, by doing so, the omitted variable bias arising from omitting the multilateral resistance terms can be partially mitigated.

The use of country-pair, "exporter × year," and "importer × year" fixed effects in panel data analyses has become standard practice, so long as their inclusion does not preclude coefficient identification.[13]

[12] All factors specific to the exporting or importing country, including its multilateral resistance, are absorbed in the exporting or importing country fixed effects, respectively.

[13] When estimating the impact of a time-variant country-level variable (e.g., the level of industrial robot penetration) on trade, including exporter × year fixed effects and importer × year fixed effects results in their absorption of the "level of industrial robot penetration," making estimation of the robot penetration coefficient impossible.

4 Missing globalization puzzle

From the 1970s through the 1990s, many countries lowered tariffs. How did globalization change the coefficient of distance in the gravity model? If globalization leads to greater international trade between distant countries than between adjacent countries, we should expect a smaller absolute value of the distance coefficient from recent periods compared with earlier periods. However, using a log-linear gravity model for their estimates, Coe et al. (2007) report no decrease in the absolute value of the coefficient of distance between 1975 and 2000. They call this discrepancy between our intuition and the empirical result the "missing globalization puzzle."

Table 8 summarizes Coe et al. (2007)'s estimates for the period 1975–2000. The absolute value of the distance coefficient in the log-linear model in Panel A is 1.02 for 1975. Its absolute value decreases to 0.92 in 1990, but it rises again to 1.08 in 2000, indicating that the distance coefficient exceeds the 1975 level. Panel B shows estimated coefficients from the non-linear model. The absolute value varies from 0.53 to 0.32 between 1975 and 2000, showing a declining trend. Coe et al. (2007) attribute the counterintuitive distance coefficients in the log-linear model to potential bias stemming from the exclusion of zero trade flows.

Table 8: Coe et al. (2007)'s estimates of the distance coefficient

	1975	1980	1985	1990	1995	2000
Panel A: Log-linear model						
Coefficient on $\ln(distance)$	-1.02	-1.01	-1.04	-0.92	-1.00	-1.08
	(0.07)	(0.08)	(0.07)	(0.07)	(0.06)	(0.06)
Panel B: Non-linear model						
Coefficient on $\ln(distance)$	-0.53	-0.40	-0.41	-0.32	-0.29	-0.32
	(0.21)	(0.14)	(0.18)	(0.15)	(0.11)	(0.12)

Note: This table comes from Table 2 in Coe et al. (2007). Standard errors are in parentheses.

Yotov (2012) emphasizes the importance of including domestic trade as a solution to the missing globalization puzzle. Yotov (2012) points out that prior empirical studies did not include domestic trade. As a result, since the distance coefficient measures the effect of distance for a country-pair, compared to the effect of distance for another country-pair, it could be natural to obtain similar distance coefficients across time. If globalization facilitated economic integration among countries, then among regions within a country—in other words, if international trade increases more than domestic trade—the distance coefficient should decrease over time when domestic trade is also included in the sample.

Based on this idea, Yotov (2012) first estimates the following regression equation without domestic trade:

$$X_{ij} = \exp[\beta_0 + \beta_1 D_{ij}^{Cntg} + \beta_2 D_{ij}^{Lang} + \beta_3 D_{ij}^{Clny} + \beta_4 \ln(dist_{ij}) + \phi_i + \phi_j] \times u_{ij},$$

where X_{ij} denotes the value of exports from country i to country j, D_{ij}^{Cntg} denotes an adjacency dummy, D_{ij}^{Lang} denotes a common language dummy, D_{ij}^{Clny} denotes a colonial relation dummy, $\ln(dist_{ij})$ denotes the log of distance, ϕ_i denotes exporter fixed effects, and ϕ_j denotes importer fixed effects.

Furthermore, Yotov (2012) estimates the following model with domestic trade:

$$X_{ij} = \exp[\alpha_0 + \alpha_1 D_{ij}^{Cntg} + \alpha_2 D_{ij}^{Lang} + \alpha_3 D_{ij}^{Clny} + \alpha_4 \ln(dist_{ij}) + \alpha_5 \ln(dist_{ii}) + \phi_i + \phi_j] \times u_{ij},$$

where $dist_{ij}$ denotes the distance between countries, and it takes zero when $i = j$. The variable $dist_{ii}$ denotes the average distance weighted by population of the distances between domestic regions, and it takes zero when $i \neq j$. Domestic exports (equivalently imports) are defined as GDP minus exports to foreign countries. If globalization decreases the importance of international distance relative to the importance of domestic distance, then the absolute value of $\widehat{\alpha_4} - \widehat{\alpha_5}$ should decrease.

Table 9: Yotov's (2012) estimates of the distance coefficient

	1965	1985	2005
$\widehat{\beta}_4$	-0.81	-0.85	-0.87
	(0.05)	(0.04)	(0.03)
$\widehat{\alpha}_5 - \widehat{\alpha}_4$	-0.65	-0.53	-0.41
	(0.03)	(0.03)	(0.02)

Note: This table comes from the PPML results in Column (1) of Table 1 in Yotov (2012). Standard errors are in parentheses.

Table 9 summarizes the estimates of β_4 and $\alpha_5 - \alpha_4$ from Yotov (2012)'s PPML. As expected, $\widehat{\beta}_4$ remains mostly stable from 1965 to 2005, whereas the absolute value of $\widehat{\alpha}_5 - \widehat{\alpha}_4$ shows a decline from 1965 to 2005. These results suggest that including domestic trade mitigates bias.

5 Conclusion

This paper summarizes the evolution of gravity model estimation methods with reference to three puzzles: the Glick-Rose puzzle, the border puzzle, and the missing globalization puzzle. Along the way, we have realized that the estimation methods have changed substantially over the past 20 years. For example, an appropriate inclusion of fixed effects has become a standard practice. Moreover, the approach has shifted from log-linear models towards non-linear models to account for zero trade flows. It is important to select an appropriate estimation model and understand why the choice matters. The author would like to pay close attention to ongoing and future developments in the gravity model.

Appendix A. Interpreting the coefficients of dummy variables when the dependent variable is with natural log

When the dependent variable is in natural log form and the explanatory variable is a dummy variable with a large coefficient, interpreting the coefficient in terms of an exponential function is necessary to restore accuracy. In this Appendix, we explain the accompanying theoretical background. For example, suppose we use cross-sectional data at a single time point to estimate the following regression equation in which the dependent variable is the log value of trade and the explanatory variable is a currency union dummy:

$$\ln(T) = \beta_0 + \beta_1 D^{CU} + u.$$

In this case, the average trade value between countries using a common currency $\ln(T_1)$ and that between countries not using a common currency $\ln(T_2)$ are

$$\ln(T_1) = \widehat{\beta}_0 + \widehat{\beta}_1 \quad (\text{because } D^{CU} = 1)$$
$$\ln(T_2) = \widehat{\beta}_0 \quad (\text{because } D^{CU} = 0)$$

respectively, where the error term is ignored. The difference between these two is $\ln(T_2) - \ln(T_1) = \widehat{\beta}_1$. When the difference between T_2 and T_1 is large, the log difference is not an accurate approximation of the rate of change. Since our interest is the rate of change $(T_2 - T_1)/T_1$, we express it as follows:

$$\begin{aligned}\frac{T_2 - T_1}{T_1} &= \frac{T_2}{T_1} - 1 \\ &= \exp\left[\ln\left(\frac{T_2}{T_1}\right)\right] - 1 \\ &= \exp\left[\widehat{\beta}_1\right] - 1.\end{aligned}$$

The third line of the equation follows from the fact that $\ln(T_2) - \ln(T_1) = \widehat{\beta}_1$. Thus, the rate of change is the exponential of the coefficient estimate $\widehat{\beta}_1$ minus 1. If we add 1 to the both sides, we obtain $T_2/T_1 = \exp[\widehat{\beta}_1]$. The exponential of the coefficient estimate $\widehat{\beta}_1$ is the ratio of trade between countries using a common currency divided by trade between countries not using a common currency.

Appendix B. Stata commands

This Appendix provides Stata commands used to estimate the gravity model. As an example, we estimate the effects of free trade agreements' (FTAs) on international trade using Conte et al. (2022)'s dataset.[14]

B.1. Preparation

First, read the dataset in `.dta` format using the following command:

```
use Gravity_V202211.dta, clear
```

Type the following:

```
ssc install reghdfe
ssc install ftools
ssc install ppmlhdfe
```

to install tools to use the high-dimensional fixed effects command, which can substantially reduce computation time. Type the following commands to generate 'data identification codes' to be used for fixed effects estimation:

```
egen pair_id = concat(country_id_o country_id_d)
egen o_year = concat(country_id_o year)
egen d_year = concat(country_id_d year)
```

The first line of the command above generates the exporting country × importing country variable, the second line generates the exporting country × year variable, and the third generates the importing country × year variable.

Use the following command to make Stata recognize the data as panel data for country pair × year:

```
encode pair_id, gen(pair_idn)
xtset pair_idn year
```

[14] The dataset is available for download at https://www.cepii.fr/CEPII/en/bdd_modele/bdd_modele_item.asp?id=8

The first line is a command to convert `pair_id`, a non-numeric string variable, into a numeric variable. The second line instructs Stata to treat the data as panel data, using `pair_idn` to identify the cross-section and `year` to indicate the time series.

The dependent variable is exports by "Reporter (Origin country)" obtained from UN Comtrade. Assuming that there is no trade for observations with missing export values, replace the missing values with zero using the following command:

```
replace tradeflow_comtrade_o = 0 if tradeflow_comtrade_o == .
```

To estimate a log-linear model, we find the log of trade flows, GDP, and distance using the following commands:

```
gen lntrade = ln(tradeflow_comtrade_o)
gen lngdp_o = ln(gdp_o)
gen lngdp_d = ln(gdp_d)
gen lndist = ln(distcap)
```

B.2. Running regressions

Since the log of zero is negative infinity, the log of zero trade flows would again be missing, but these are included in PPML estimations. We first estimate the log-linear pooled OLS, with no fixed effects, using the following command:

```
reg lntrade fta_wto lngdp_o lngdp_d lndist, vce(cluster pair_id)
```

Note that the standard errors are robust, clustered by country pairs. The regression controlling for country-pair fixed effects is estimated using the following command:

```
xtreg lntrade fta_wto lngdp_o lngdp_d, fe vce(cluster pair_id)
```

Note that this command assumes that the command "`tsset pair_idn year`" has been typed previously to utilize the panel structure of the data.[15]

B.3. Running high-dimensional fixed effect regressions

We run regressions controlling for "exporter × year fixed effects," "importer × year fixed effects," and "country-pair fixed effects," using the following command:

```
reghdfe lntrade fta_wto, absorb(pair_id o_year d_year) vce(cluster pair_id)
```

The following command conducts the same estimation as the previous model, but with robust standard errors clustered by exporters, importers, and year:

```
reghdfe lntrade fta_wto, absorb(pair_id o_year d_year) cluster(country_id_o country_id_d year)
```

Finally, we estimate PPML with "exporters × year fixed effects," "importers × year fixed effects," and "country-pair fixed effects" using the following command:

```
ppmlhdfe tradeflow_comtrade_o fta_wto, absorb(pair_id o_year d_year) vce(cluster pair_id)
```

[15] A similar estimation can be performed with the following command using `reghdfe`:
`reghdfe lntrade fta_wto lngdp_o lngdp_d, absorb(pair_id) vce(cluster pair_id)`.
In this command, `absorb(pair_id)` specifies that the cross-section is recognized by the `pair_id` variable, eliminating the need to run `tsset pair_idn year` beforehand.

In the case of PPML, the dependent variable is actual trade values (rather than the log of trade value). The following command conducts the same estimation as the previous model, but with robust standard errors clustered by exporters, importers, and year:

```
ppmlhdfe tradeflow_comtrade_o fta_wto, absorb(pair_id o_year d_year) cluster(country_id_o
    country_id_d year)
```

B.4. Example results

Table 10 summarizes results. Results from estimating a log-linear model without controlling for any fixed effects are shown in column (1). It suggests that FTAs would increase trade by $e^{0.412} = 1.51$ times. Column (2) introduces country-pair fixed effects, leaving the distance coefficient unfilled.[16] It shows that, controlling for country-pair fixed effects reduces FTAs' trade-creating effect to a factor of $e^{0.229} = 1.26$.

Table 10: Effects of FTAs on international trade

	Log-linear model				PPML	
	(1)	(2)	(3)	(4)	(5)	(6)
FTA dummy	0.412***	0.229***	0.228***	0.228***	0.083***	0.083
	(0.032)	(0.020)	(0.022)	(0.058)	(0.032)	(0.051)
ln(Exporter GDP)	0.965***	0.665***				
	(0.005)	(0.012)				
ln(Importer GDP)	0.674***	0.539***				
	(0.005)	(0.011)				
ln(Distance)	−1.263***					
	(0.016)					
Sample size	754,582		841,975		1,344,568	
Country-pair fixed effects		Yes	Yes	Yes	Yes	Yes
Exporter × year fixed effects			Yes	Yes	Yes	Yes
Importer × year fixed effects			Yes	Yes	Yes	Yes
Three-way robust standard errors				Yes		Yes

Note: Author's estimation using Conte et al. (2022)'s dataset. Standard errors are in parentheses. *** indicates significance at the 1% level. The sample period is 1962–2020.

Column (3) shows the results of further controlling for "exporter × year fixed effects" and "importer × year fixed effects," but the trade creation effect is almost the same as in column (2). The results shown in column (4) are based on robust standard errors clustered by exporters, importers, and year. Although the estimated coefficients are, of course, the same as in column (4), standard errors are over twice as large as in column (3). Nevertheless, the results remain significant at the 1% level.

Column (5) shows PPML results, which increases the sample size from 841,975 to 1,344,568 due to the inclusion of zero trade flows. It shows that FTAs' trade creation effect is $e^{0.083-1} = 0.086$ (8.6%), and statistically significant at the 1% level. Column (6) is based on robust standard errors clustered by exporters, importers, and year. The coefficients are the same as in column (5), but the standard errors are larger and are no longer statistically significant.

Overall, these results suggest that (1) log-linear models overestimate FTAs' trade-creating effects, (2) failing to control for fixed effects introduces bias, and (3) standard errors may be misestimated if not properly clustered. Even this analysis does not address FTA endogeneity from reverse causality or other confounders, which requires further methods—such as the instrumental variable approach—to mitigate bias.

[16] Given that Conte et al. (2022)'s distance variable, "distcap," is time variant, the coefficients can be estimated even after introducing country-pair fixed effects. The change in distance may be due to changes in capital cities, etc., and observing a change in the variable does not imply an actual change in distance between the countries. For this reason, it is excluded from the regression equation in the analysis with fixed effects.

References

Anderson, James E. (2011) "The gravity model". *Annual Review of Economics*, 3: pp. 133–160.

Anderson, James E. and Eric van Wincoop. (2003) "Gravity with gravitas: a solution to the border puzzle". *American Economic Review*, 93(1): pp. 170–192.

Baldwin, Richard and Daria Taglioni. (2007) "Trade effects of the euro: a comparison of estimators". *Journal of Economic Integration*: pp. 780–818.

Cameron, A. Colin, Jonah B. Gelbach and Douglas L. Miller. (2011) "Robust inference with multiway clustering". *Journal of Business & Economic Statistics*, 29(2): pp. 238–249.

Campbell, Douglas L. (2013) "Estimating the impact of currency unions on trade: solving the Glick and Rose puzzle". *The World Economy*, 36(10): pp. 1278–1293.

Cipollina, Maria and Luca Salvatici. (2010) "Reciprocal trade agreements in gravity models: a meta-analysis". *Review of International Economics*, 18(1): pp. 63–80.

Coe, David T., Arvind Subramanian and Natalia T. Tamirisa. (2007) "The missing globalization puzzle: evidence of the declining importance of distance". *IMF Staff Papers*, 54(1): pp. 34–58.

Conte, Maddalena, Pierre Cotterlaz and Thierry Mayer (2022). "The CEPII Gravity Database". CEPII Working Paper 2022-05, Paris. URL: https://ideas.repec.org/p/cii/cepidt/2022-05.html.

Glick, Reuven and Andrew K. Rose. (2002) "Does a currency union affect trade? the time-series evidence". *European Economic Review*, 46(6): pp. 1125–1151.

Glick, Reuven and Andrew K. Rose. (2016) "Currency unions and trade: a post-EMU reassessment". *European Economic Review*, 87: pp. 78–91.

Head, Keith and Thierry Mayer. (2014) "Gravity equations: Workhorse, toolkit, and cookbook". *Handbook of International Economics*, 4: pp. 131–195.

Larch, Mario, Joschka Wanner, Yoto V. Yotov and Thomas Zylkin. (2019) "Currency unions and trade: a PPML re-assessment with high-dimensional fixed effects". *Oxford Bulletin of Economics and Statistics*, 81(3): pp. 487–510.

Larch, Mario and Yoto V. Yotov. (2024) "Estimating the effects of trade agreements: lessons from 60 years of methods and data". *The World Economy*, 47(5): pp. 1771–1799.

McCallum, John. (1995) "National borders matter: Canada-US regional trade patterns". *American Economic Review*, 85(3): pp. 615–623.

Rose, Andrew K. (2000) "One money, one market: the effect of common currencies on trade". *Economic Policy*, 15(30): pp. 08–45.

Rose, Andrew K. (2017) "Why do estimates of the EMU effect on trade vary so much?" *Open Economies Review*, 28(1): pp. 1–18.

Rose, Andrew K. and Tom D. Stanley. (2005) "A meta-analysis of the effect of common currencies on international trade". *Journal of Economic Surveys*, 19(3): pp. 347–365.

Shepherd, Ben (2016). "The gravity model of international trade: a user guide (an updated version)". United Nations ESCAP. URL: https://www.unescap.org/resources/gravity-model-international-trade-user-guide-updated-version.

Shepherd, Ben (2019). “The gravity model of international trade: a user guide (an updated version)”. United Nations ESCAP. URL: https://www.unescap.org/resources/gravity-model-international-trade-user-guide-r-version.

Silva, JMC Santos and Silvana Tenreyro. (2006) “The log of gravity”. *Review of Economics and Statistics*: pp. 641–658.

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

Yotov, Yoto V. (2022). “Gravity at sixty: the bijou of trade”. Drexel School of Economics, Working Paper 2022-01.

Yotov, Yoto V., Roberta Piermartini, José-Antonio Monteiro and Mario Larch (2016). *An Advanced Guide to Trade Policy Analysis: The Structural Gravity Model*. World Trade Organization. URL: https://www.wto.org/english/res_e/booksp_e/advancedwtounctad2016_e.pdf.