Replication Project

BUSN41903 Applied Econometrics

Seung Chul Lee

Paper chosen: Feyrer (2019), "Trade and Income—Exploiting Time Series in Geography", *American Economic Review: Applied Economics*, 11(4): 1–35. doi.org/10.1257/app.20170616

1 Summary of Paper

This paper attempts to estimate a robust causal relationship between trade and national income. In this literature, there have been efforts to exploit geographic features as instruments (Frankel and Romer, 1999, among others) to overcome the endogenous nature of trade and income. However, the physical distance between countries is a stationary variable, which only allows for the analysis of a single cross section. This paper is one of many attempts at forming a time-varying instrument based on geography to allow applications to a panel. The author argues that, with the advancement of aircraft technology and insuing decrease in air freight costs, countries that were traditionally "far" in terms of sea distance become more accessible by air. Thus, the trade between countries with shorter air distance will grow, whereas it may not be a first-order determinant of gross domestic product. Figure 1 provides a schematic diagram of the causal relationship that the author posits.

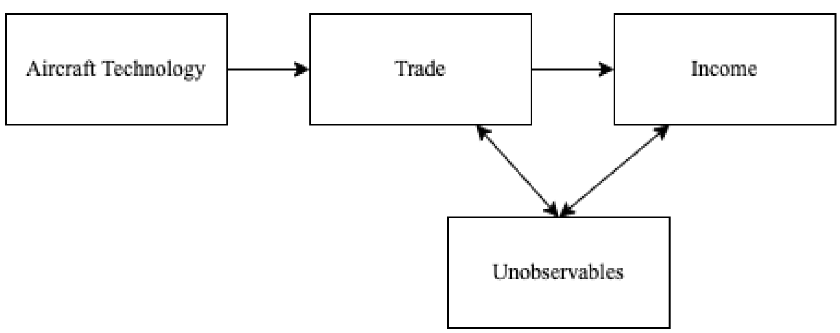

Figure 1: Schematic Graph of Causal Relationship

The paper's empirical model is as follows.

$$\ln GDP_{it} = \gamma_i + \gamma_t + \beta_1 \ln Trade_{it} + \epsilon_{it} \tag{1}$$

$$\ln Trade_{it} = \tilde{\gamma}_i + \tilde{\gamma}_t + \beta_0 \ln \widehat{Trade}_{it} + \nu_{it} \tag{2}$$

where Equation 1 is the structural equation and Equation 2 is the first stage. The proposed instrument

$\ln \widehat{Trade}_{it}$ is estimated using a gravity model[1] as follows.

$$\ln Trade_{ijt} = \alpha + \gamma_i + \gamma_j + \gamma_t + \beta_{sea,t} \ln SeaDist_{ij} + \beta_{air,t} \ln AirDist_{ij} \tag{3}$$

$$\ln Trade_{ijt} = \alpha + \gamma_{ij} + \gamma_t + \beta_{sea,t} \ln SeaDist_{ij} + \beta_{air,t} \ln AirDist_{ij} \tag{4}$$

$$\widehat{Trade}_{it} = \sum_{j \neq i} \exp\{\hat{\gamma}_j + \hat{\beta}_{sea,t} \ln SeaDist_{ij} + \hat{\beta}_{air,t} \ln AirDist_{ij}\} \tag{5}$$

where Equation 3 is the gravity model specification predicting bilateral trade using country-specific dummies and Equation 4 is the specification with country-pair dummies. For each of the specifications, a country's predicted trade is estimated by summing across possible partners as denoted in Equation 5. This predicted trade value, $\ln \widehat{Trade}$, is the instrument.

The author's conclusion is that increase in trade has a positive causal effect on a country's gross domestic product.

2 Replication of Main Results

Table 1 replicates the results of the IV regression with fixed effects (Table 4 in the paper). The STATA code and data sets are available on Open ICPSR[2].

[1] The gravity model:

$$\ln Trade_{ijt} = \ln GDP_{it} + \ln GDP_{jt} - \ln GDP_{wt} + (1 - \sigma)(\ln \tau_{ijt} + \ln P_{it} + \ln P_{jt})$$

where $Trade_{ijt}$ is the amount of trade between country i and country j at time t, $\ln GDP_{it}$ is country i's gross domestic product at time t, $\ln GDP_{wt}$ is the gross domestic product of the world at time t, τ_{ijt} is the bilateral resistance term between country i and j, and P_{it} is the global resistance term of country i at time t. The author believes that the τ_{ijt} term is determined by sea distance and air distance. In the paper, the author tries to avoid including GDP directly in the gravity model and instead uses fixed effects.

[2] https://www.openicpsr.org/openicpsr/project/116365/version/V1/view

Table 1: IV with Fixed Effects (Table 4)

Panel A. Second Stage					
	$\ln GDP_i$				
		Country Dummies		Pair Dummies	
	OLS	Trade Weight	Pop Weight	Trade Weight	Pop Weight
	(1)	(2)	(3)	(4)	(5)
$\ln Trade_i$	0.446***	0.578***	0.611***	0.459***	0.716***
	(0.0412)	(0.0818)	(0.131)	(0.0974)	(0.128)
N	774	774	774	774	774
R^2	0.965				

Panel B. First Stage				
	$\ln Trade_i$			
	(2)	(3)	(4)	(5)
$\ln \widehat{Trade}_i$	0.993***	0.731***	1.385***	1.353***
	(0.144)	(0.187)	(0.251)	(0.296)
F Stat	47.22	15.29	30.47	20.92
N	774	774	774	774
R^2	0.975	0.972	0.973	0.972

Panel C. Reduced Form				
	$\ln GDP_i$			
	(2)	(3)	(4)	(5)
$\ln \widehat{Trade}_i$	0.573***	0.446***	0.636***	0.968***
	(0.116)	(0.130)	(0.185)	(0.195)
N	774	774	774	774
R^2	0.947	0.943	0.943	0.945

Standard errors in parentheses, clustered by country.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2 replicates the results of the IV regression with first differences (Table 5 in the paper).

Table 2: IV with First Differences (Table 5)

Panel A. Second Stage					
	$\Delta \ln GDP_i$				
		Country Dummies		Pair Dummies	
	OLS	Trade Weight	Pop Weight	Trade Weight	Pop Weight
	(1)	(2)	(3)	(4)	(5)
$\Delta \ln Trade_i$	0.229***	0.739***	0.698**	0.540***	1.201+
	(0.0300)	(0.149)	(0.234)	(0.152)	(0.722)
N	673	673	673	673	673
R^2	0.186				

Panel B. First Stage				
	$\Delta \ln Trade_i$			
	(2)	(3)	(4)	(5)
$\Delta \ln \widehat{Trade}_i$	0.548***	0.342*	0.640**	0.361
	(0.118)	(0.150)	(0.201)	(0.278)
F Stat	21.68	5.182	10.15	1.686
N	673	673	673	673
R^2	0.470	0.463	0.465	0.461

Panel C. Reduced Form				
	$\Delta \ln GDP_i$			
	(2)	(3)	(4)	(5)
$\Delta \ln \widehat{Trade}_i$	0.404***	0.238*	0.345*	0.434**
	(0.0955)	(0.104)	(0.133)	(0.139)
N	673	673	673	673
R^2	0.0798	0.0595	0.0602	0.0613

Standard errors in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

3 Additional Analysis

I propose the following as my additional analysis using the data for this paper.

- Weak instrument correction
- Estimation of interactive fixed effects

Weak Identification Robust Infernece

Some columns in Table 2, namely 2 and 4, exhibit first-stage F statistics below 10, which is indicative of weak identification. Although it may not be of material difference, I wanted to apply weak identification robust inference that we have covered in class. I closely follow Chernozhukov and Hansen (2008) and implement the following steps:

1. Select $\mathcal{B} = [-1, 4]$ as my desired set of potential values for β
2. Partial out the fixed effects from the $\ln GDP$ (the response), $\ln Trade$ (endogenous regressor), and $\ln \widehat{Trade}$ (the instrument).
3. For each β_0 spaced equally at 0.001 unit intervals in $\mathcal{B}$, regress $\ln GDP - \ln Trade\beta_0$ on the instrument $\ln \widehat{Trade}$ to obtain $\hat{\alpha}$.
4. Test the null that $H_0 : \alpha = 0$ and construct the confidence interval as the region that H_0 cannot be rejected using the F test.

The only difference is that I employed the F test, which is essentially equivalent to the Wald test in this context. Table 3 reports the results of this procedure. I also include a robust confidence interval for the main specification in this paper (column 2 of Table 1) for comparison. Not surprisingly, the robust confidence intervals also largely support the conclusion of this paper. I believe it is worth noting that column 2 of Table 3, which is the FD specification with population weights and country-pair dummies, actually becomes significant with the robust interval. It is interesting to observe that robust inference can sometimes be more supportive of one's conclusions.

Table 3: Weak Identification Robust Inference

	(1) FD Column 3	(2) FD Column 5	(3) FE Column 2
N	774	774	774
F	5.1819	1.6865	47.2193
$\hat{\beta}_{2SLS}$	0.6978	1.2008	0.5777
Asymptotic C.I	(0.2389, 1.1566)	(-0.2143, 2.6159)	(0.4174, 0.7381)
Robust C.I.	(0.208, 3.778)	(0.501, 4]	(0.494, 0.665)

N refers to sample size; F refers to first-stage F statistic; $\hat{\beta}_{2SLS}$ refers to the 2SLS point estimate.

Interactive Fixed Effects

A key feature of this paper is that it relies on the two-way fixed effects structure to account for any variation in the response that is not captured by trade. For instance, technological advances may appreciate the value of underlying financial assets that will also increase the country's income. For the author's model to be valid, such changes throughout time should impact all countries equally. However, it is highly likely that each country will be affected to differing degrees due to differences in the portfolios held by the national governments, etc. To allow for flexible variations of fixed effects, I attempt to estimate an interactive fixed effects structure using principal components as developed by Bai (2009). The model is

$$\ln GDP_{it} = \alpha + \beta_1 \ln Trade_{it} + \lambda_i' f_t + \epsilon_{it} \tag{6}$$

$$\ln Trade_{it} = \tilde{\alpha} + \beta_0 \ln \widehat{Trade}_{it} + \tilde{\lambda}_i' \tilde{f}_t + \nu_{it} \tag{7}$$

where f_t is the factor corresponding to time t and λ_i is country i's individual loading on this factor. Equation 6 is the newly defined structural equation and Equation 7 is the accompanying first stage. The focal point of Bai (2009) is that $\lambda_i' f_t$ can be estimated similarly to a principal component analysis (PCA), assuming the underlying factor structure to be sparse (i.e., the matrix of factors F is of rank r, for some small r). His interactive fixed effects model is implemented in R as `interFE` in the `gsynth` package. The function iteratively solves the model and has the option to estimate relevant standard errors using bootstrap.

However, there is no readily available package to estimate an interactive fixed effects model along with instrumental variables. Hence, I take on an approach similar to weak instrument correction. The steps are as follows:

1. Estimate the model: $y_{it} - x_{it}'\beta = z_{it}\gamma + \lambda_i' f_t + \epsilon_{it}$ for various values of β.

2. Obtain a confidence interval for β with values where we cannot reject $H_0 : \gamma = 0$.

This process relies on the instrument being exogenous and not having any additional explanatory power over the response given the endogenous regressor, much like the weak identification robust inference. As this procedure does not yield a point estimate, I choose the midpoint of the interval as my suggested point estimate.

I repeat these steps for $r = 2, 3, 4$. One could devise a full cross validation procedure to select an optimal r, but I could not develop something feasible as the current implementation of `interFE` is prohibitively expensive in terms of computational load. Also, for the possible values of β, I first try the set [-0.1, 0.5] with a step size of 0.01 and then narrow down the interval to [0.28, 0.31] with a step size of 0.001 after observing all three trials indicated validity around these values. I chose the two-step procedure to reduce the computational burden. The results are reported in Table 4.

Table 4: Interactive Fixed Effects Estimation

	(1)	(2)	(3)
r	2	3	4
$\hat{\beta}$	0.286	0.2955	0.2935
Robust C.I.	(0.28, 0.292)	(0.291, 0.3)	(0.287, 0.3)

I find that, with the new fixed effects structure, the estimated coefficients on trade are smaller in magnitude but still positive and statistically significant. Thus, the conclusion of the paper is still reasonable.

I also try using the function `xtivdfreg` implemented in STATA by Kripfganz and Sarafidis (2021). The theory developed for the function relies on having a large T as described in the paper. Hence, the inference may be less applicable for the data at hand, as I have only $T = 10$. However, I thought it was an interesting function that was developed very recently, so I wanted to gain some experience in using it. The estimated results are provided in Table 5.

Table 5: STATA `xtivdfreg`

	(1)
	$\ln GDP_i$
$\ln Trade_i$	0.922***
	(3.47)
const.	-11.48*
	(-2.06)
N	766

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Again, similarly to the results in Table 4, the coefficient on trade is still positive and significant. The magnitude is actually larger with this function, which may be due to having a small T not enough to induce asymptotic properties.

References

Bai, J. (2009). Panel data models with interactive fixed effects. *Econometrica*, 77(4):1229–1279.

Chernozhukov, V. and Hansen, C. (2008). The reduced form: A simple approach to inference with weak instruments. *Economics Letters*, 100(1):68–71.

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

Frankel, J. A. and Romer, D. H. (1999). Does trade cause growth? *American Economic Review*, 89(3):379–399.

Kripfganz, S. and Sarafidis, V. (2021). Instrumental-variable estimation of large-t panel-data models with common factors. *The Stata Journal*, 21(3):659–686.