Public Infrastructure and Economic Development: Evidence from Postal Systems

A Replication Study of Rogowski et al. (*American Journal of Political Science*, 2022)

Florian Neubauer [*] Julian Rose [†] Jörg Ankel-Peters[‡]

Journal of Comments and Replications in Economics, Volume 4, 2025-5, DOI: 10.18718/81781.44

JEL: A1, O18, L87
Keywords: Reproduction, Infrastructure, Postal Services
Data Availability: The Stata code and the data to reproduce the results of this replication can be downloaded at JCRE's data archive (DOI: 10.15456/j1.2024066.1359789444).

Please Cite As: Neubauer, F., Rose, J. and Ankel-Peters, J. (2025). Public Infrastructure and Economic Development: Evidence from Postal Systems A Replication Study of Rogowski et al. (*American Journal of Political Science*, 2022). *Journal of Comments and Replications in Economics*, Vol.4 (2025-5). DOI: 10.18718/81781.44

Abstract

Rogowski et al. (2022) use secondary data to study the impact of historic postal infrastructure on economic development, both cross-country and within the US. Their results suggest a large positive effect of post offices on economic development that is robust across various sensitivity checks. We successfully computationally reproduce all results. In a robustness assessment, we find the results to be robust to simple changes in the analysis but observe some sensitivity to accounting for spatial trends in the long-term cross-country analysis. Despite this, we find the results to be overall robust given the numerous analyses and robustness checks in the original paper.

[*]RWI - Leibniz Institute for Economic Research and University of Connecticut florian.neubauer@rwi-essen.de
[†]RWI – Leibniz Institute for Economic Research and University of Passau julian.rose@rwi-essen.de
[‡]RWI – Leibniz Institute for Economic Research and University of Passau joerg.peters@rwi-essen.de
Declaration: Funding: This work was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG), Grant No. 3473/1-1 within the DFG Priority Program META-REP (SPP 2317).

Received March 6, 2024; Revised April 18, 2025; Accepted July 15, 2025; Published October 17, 2025.

1 Introduction

Rogowski et al. (2022) – henceforth RGMC - investigate the impact of postal systems on economic development from the 19th century onwards. Positive impacts are expected because the expansion of postal services facilitated communication, commerce, advertising, remittances, job searches, investment monitoring, technology dissemination, and integration of remote areas. The study employs four types of analyses (see Figure 1): short-term cross-country (STCC), long-term cross-country (LTCC), short-term US county-level (STUS), and long-term US county-level (LTUS).

In this reproduction report, we conduct a computational reproduction and a robustness [1] reproduction using the provided reproduction package from the Harvard Dataverse. We successfully reproduce all results published in the original paper. Despite this successful reproduction, we note that the authors did not provide the full raw data, as some datasets contain pre-processed data. Moreover, some datasets lack variable labels or a codebook, which impedes conducting a robustness reproduction. We acknowledge that providing all raw data with comprehensive documentation is not yet the norm in the social sciences. Still, we conduct a series of robustness checks and find that the results hold for most of them.

RGMC highlight a positive relationship between postal systems and economic development in all four analyses. The LTCC results indicate that a 20% increase in post offices per capita in 1900 is associated with a 3% increase in GDP in 2000, while STCC analyses show a positive but modest effect. At the U.S. County level, a 25% increase in post offices in a county is linked to a 5-6% increase in farm values, manufacturing outputs, and capital investments in manufacturing in the short term (STUS). In the long run, a 25% increase in post offices in 1896 corresponds to a 2% higher median income in 2000 for U.S. counties (LTUS). The authors conduct robustness tests for each analysis, all supporting the main results.

For the LTCC analysis, the treatment variable is the logged number of post offices (+1) per country in 1900, sourced from the Universal Postal Union (UPU) spanning from 1875 to 2007. RGMC examine the relationship between post offices and logged GDP growth for 77 countries over time. For the STCC analysis, RGMC employ panel data to study the association between postal systems, measured as the logged stock of per capita post offices (+1), and growth at 5-year intervals. Importantly, the authors note significant endogeneity issues in the cross-country analysis (STCC and LTCC) since assignment of post offices is not random and varies across countries due to geographical, political, and other circumstances. To tackle these endogeneity issues, they conduct analyses focusing solely on counties in the U.S. (STUS and LTUS).

[1] See Dreber and Johannesson (2024)

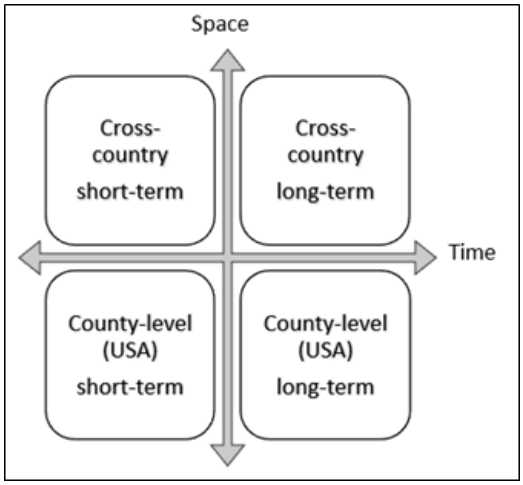

Figure 1: Overview of the different analyses in Rogowski et al. (2022)

For the panel analysis in the U.S. (STUS), RGMC utilize comprehensive panel data on the number of post offices encompassing approximately 2,700 U.S. counties, recorded at ten-year intervals from 1846 to 1896. They combine this dataset with economic data containing information on farm values, value of manufacturing output, and capital investment in manufacturing that RGMC use as dependent variables (parameterized as log(1+x)), obtained from nineteenth-century census reports spanning from 1850 to 1900. The STUS evaluation involves regressing the outcome variables on a lagged measure of the logged number of post offices (+1), clustering the standard errors at the county level, allowing for an analysis of immediate impacts.[2] In contrast, the LTUS analysis explores the relationship between post offices in 1896 (again, parameterized as log(1+x)) and logged median household income in the year 2000, incorporating additional covariates such as population, percentage of the population born outside the United States, and population density, each measured from the 1890 Census and in log form.

This background of a multi-contextual empirical analysis is important when assessing our only problematic finding. In line with Conley & Kelly (2025)'s recommendation, we introduce dummy variables for World Bank regions to the LTCC analysis. This is to account for spatial trends in persistence regressions given that spatial data tends to be highly autocorrelated. We find that the inclusion of those dummy variables renders the LTCC results statistically insignificant, suggesting some sensitivity of this finding. Additionally, we implemented several simple robustness checks, which mostly support the original results, such as applying the inverse hyperbolic sine transformation in all four analyses, clustering the standard errors at the state level instead of the county level as well as using bootstrapping to calculate standard errors in the STUS analysis, and using the Poisson pseudo-maximum likelihood estimator in the STUS and LTUS analyses. Overall, given the multitude of supporting robustness checks in RGMC and our robustness checks, we affirm the overall robustness of the results.

[2]RGMC use a lagged measure of post offices to address endogeneity in their US county analysis. However, while there is no good fix to this problem, it has been shown that using lags does not address endogeneity sufficiently (Bellemare et al., 2017)

2 Computational Reproduction

We assess the computational reproducibility of RGMC based on the original data and code ('reproduction package'), accessible from the Harvard Dataverse (Rogowski et al., 2020). RGMC provide a 'read-me' document, outlining the contents of the reproduction package, including datasets, code, a codebook, and computing specifications. Unfortunately, the codebook only covers the final datasets used for the analyses but not the preliminary datasets, which RGMC use to construct the analysis dataset. Additionally, missing variable labels in many datasets impede the identification of variables.

RGMC employ 18 preliminary datasets to generate six final datasets for various aspects of the analysis. The lack of information regarding the preliminary datasets makes it challenging to understand how imputations were carried out and whether the data is raw data (or a scientific use file) or already manipulated. For instance, the dataset "postal_imputation" contains already imputed data, and it is unclear how the authors did this. Moreover, the authors do not provide the full code or data on how they created the stock variable on the number of post offices (treatment variable). Following the BITSS' Guide for Accelerating Computational Reproducibility in the Social Sciences, the lack of raw data leads to a reproducibility level seven out of ten.[3]

Despite this, thanks to the well-documented reproduction package, we exactly reproduce RGMC's main results from Tables 1-3, except for a small difference in the constant in model (1) in Table 1 (RGMC: 22.790; Reproduction: 22.365).

[3] See: https://bitss.github.io/ACRE/

Table 1: Main results from RGMC

	Cross-country		U.S. Counties	
	Short-term	Long-term	Short-term	Long-term
Name of display item	Table 1	Table 1	Table 2	Table 3
Column	2	4	2	9
Outcome variable	Per capita GDP growth (ln)	Per capita GDP (ln, 2000)	Farm value (ln)	Median income (ln, 2000)
Variable of interest	# post offices (ln, stock)	# post offices (ln, 1900)	# post offices (ln)	# post offices (ln, 1896)
Coefficient	0.02	0.174	0.314	0.071
Standard error	0.006	0.082	0.028	0.015
Number of observations	1,540	77	12,254	1,674
Econometric method	OLS	OLS	OLS	OLS
Fixed Effects	Country, Year	No	County, Year	State
Standard Error type	Clustered at country-level	Robust	Clustered at county-level	Robust
Control variables	GDP per capita (ln), Population (ln), Urbanization, Regime type	GDP per capita (ln, 1900), Population (ln, 1900), Urbanization (1900), Regime type, Years since independence (1900)	Population (ln), Foreign born (%, ln)	Population in 1890 (ln), Population density in 1890 (ln), % Foreign born in 1896 (ln), Economic development vars (1890–1860), Rail and water access (1860), Slave population (1860), Latitude and longitude

Notes: The short-term U.S. County-level analysis is also done for manufacturing output and manufacturing capital with similar results. The long-term U.S. County-level analysis is also done for the number of manufacturing establishments with similar results.

Table 2 summarizes the results of the computational reproduction showing that the results are not fully reproducible from raw data.

Table 2: Assessment of computational reproduction

	Fully	Partial	No
Raw data provided		✗	
Cleaning code provided	✗		
Analysis data provided	✗		
Analysis code provided	✗		
Reproducible from raw data		✗	
Reproducible from analysis data	✗		

3 Robustness Reproduction

When constructing the main explanatory variable in all four analyses, RGMC add the value 1 to the number of post offices before calculating the logarithm. The method to add a small positive number to zero values has been criticized in the literature for its arbitrariness and alteration of the original data structure (Aihounton & Henningsen, 2021; Bellemare & Wichman, 2020). To test the robustness of RGMC's results to other specifications, we apply the inverse hyperbolic sine transformation (IHST) in all four analyses and the Poisson pseudo-maximum likelihood (PPML) estimator following Santos Silva & Tenreyro (2011) in the US county-level analyses.[4] Implementing IHST, Table A1 shows that the effect in the LTCC analysis decreases in magnitude (0.174 to 0.158) and turns insignificant at the 5% level. The STCC results are unchanged, and Table A2 depicts that the magnitude of the estimated effects decreases slightly in the STUS and LTUS analyses. Implementing PPLM, Table A2 shows that the effect size increases slightly in both the STUS and LTUS analyses.

We also investigate RGMC's decision of clustering standard errors at the county level in their STUS analysis. This choice assumes that errors are uncorrelated across counties, which seems questionable. We test two alternatives: First, we cluster standard errors at the next higher level, namely the state, and second, we use bootstrapping with 200 repetitions as an additional robustness check. Results remain largely unchanged and are shown in Table A2. We also combine these alternative ways to cluster the standard errors with the inverse hyperbolic sine transformation and the Poisson pseudo-maximum likelihood estimator. The coefficients remain statistically significant but decrease slightly when using the inverse hyperbolic sine transformation and increase when using the Poisson pseudo-maximum likelihood estimator.

For the LTCC analysis, RGMC document substantial positive effects of postal infrastructure on GDP (0.174, SE: 0.082, p-value$<$0.05). For our robustness check, we follow Conley & Kelly (2025)'s suggestion to include dummy variables for the World Bank regions to control for spatial trends. Conley & Kelly (2025) emphasize that fitting spatial trends can lead to misleading correlations. Moreover, they highlight the challenge posed by high spatial autocorrelation in economic variables, making it crucial to carefully consider and control for such spatial influences to avoid mistaking

[4] Since the dependent variable in the STCC analysis contains negative value and the LTCC data contains multiple imputed values for which the PPML estimator cannot account, we only implement this estimator in the STUS and LTUS analyses.

incidental correlations for significant relationships. Conley & Kelly (2025) propose two simple robustness checks: first, the introduction of the World Bank regions as dummy variables; second, introduction of longitude and latitude as control variables.

In line with findings from Conley & Kelly (2025) regarding the substantial impact of regional dummies, we observe a similar pattern when incorporating them into RGMC's analysis. Table A1 shows that the initially positive and significant long-term effect of post offices in 1900 on economic development in 2000 decreases by 30% and turns insignificant; the coefficient for the number of post offices undergoes a change from 0.174 (SE: 0.082, $p<0.05$) to 0.116 (SE: 0.089, $p=0.200$). This adjustment results in a reduction in effect size, rendering the association statistically insignificant, even at the 10% significance level, in alignment with Conley & Kelly (2025)'s observations for other studies. [5] Overall, our robustness check points to some sensitivity of the LTCC results and calls for a more careful discussion of Conley & Kelly (2025)'s critique.

With respect to the LTUS analysis, RGMC find substantial positive impacts of postal infrastructure in 1896 on both median income and the number of manufacturing establishments in 2000, across numerous specifications. One point of concern is that the expansion of post offices is correlated with other changes in the US at the time. The original authors discuss the expansion of the railroad network as one potential confounder. [6] If the railroad expansion (and other variables RGMC include as robustness checks like patents, market access, urbanization, population growth, and the share of foreign population) is a true confounder, controlling for it is necessary. However, if it is a collider, meaning it is influenced by both post offices and growth, controlling for it would introduce collider bias. Thus, RGMC in our view sufficiently explore other channels that are correlated with post offices and could influence GDP growth.

4 Conclusion

In this reproduction report, we have successfully reproduced the original results reported in Rogowski et al. (2022) using the provided reproduction package. We only highlighted some missing raw data and a lack of a clear description for some of the preliminary data in the reproduction package, posing challenges for the robustness reproduction. Subsequently, we conducted robustness checks along five dimensions:

- inclusion of dummy variables for the World Bank regions according to Conley & Kelly (2025) in the LTCC analysis,
- using the 1) inverse hyperbolic sine transformation in all four analyses and 2) the Poisson pseudo-maximum likelihood estimator in the US county-level analyses instead of the log(1+x) parameterization for variables that include zero values,
- 1) clustering the standard errors at the state level instead of the county level, and 2) using bootstrapping to calculate standard errors in the STUS analysis.

[5] A robustness check combining the dummy for World Bank regions and the inverse hyperbolic sine transformed variables, also yields a smaller and insignificant coefficient of post offices (Cofficient: 010, SE: 0.088, $p=0.252$).

[6] RGMC estimate models using different measures of railroad access and find no significant diminishing effect on their results.

Our results demonstrate that the inclusion of dummy variables for World Bank regions, as proposed by Conley & Kelly (2025), renders the LTCC effects statistically insignificant and reduces the effect size. Furthermore, using the inverse hyperbolic sine transformation reduces the effect size in the LTCC analysis and renders the coefficient insignificant at the 5% level. While these robustness checks reveal some sensitivity of the results for the LTCC analysis, it is important to emphasize that the overall conclusion of the paper remains intact. RGMC conduct a wide array of analyses and robustness checks, all of which consistently point towards positive effects of postal services on economic outcomes.

References

Aihounton, G. B. D., & Henningsen, A. (2021). "Units of Measurement and the Inverse Hyperbolic Sine Transformation:" *The Econometrics Journal*, 24(2): 334–351. DOI: 10.1093/ectj/utaa032.

Bellemare, M. F., Masaki, T., & Pepinsky, T. B. (2017). "Lagged Explanatory Variables and the Estimation of Causal Effect." *The Journal of Politics*, 79(3): 949–963. DOI: 10.1086/690946.

Bellemare, M. F., & Wichman, C. J. B. (2020). "Elasticities and the Inverse Hyperbolic Sine Transformation." *Oxford Bulletin of Economics and Statistics*, 82(1): 50–61. DOI: 10.1111/obes.12325.

Conley, T. G., & Kelly, M. (2025). "The Standard Errors of Persistence." *Journal of International Economics*, 153: 104027. DOI: 10.1016/j.jinteco.2024.104027.

Dreber, A., & Johannesson, M. (2024). "A Framework for Evaluating Reproducibility and Replicability in Economics." *Economic Inquiry*, ecin.13244. DOI: 10.1111/ecin.13244.

Rogowski, J. C., Gerring, J., Maguire, M., & Cojocaru, L. (2020). "Replication Data for: Public Infrastructure and Economic Development: Evidence from Postal Systems [Dataset]." *Harvard Dataverse*. DOI: 10.7910/DVN/33K3EF.

Rogowski, J. C., Gerring, J., Maguire, M., & Cojocaru, L. (2022). "Public Infrastructure and Economic Development: Evidence from Postal Systems." *American Journal of Political Science*, 17. DOI: 10.1111/ajps.12594.

Santos Silva, J. M. C., & Tenreyro, S. (2011). "Further Simulation Evidence on the Performance of the Poisson Pseudo-maximum Likelihood Estimator." *Economics Letters*, 112(2): 220–222. DOI: 10.1016/j.econlet.2011.05.008.

Appendix

Table A1: Results from robustness checks for the STCC and LTCC analyses

	STCC: DV = pc GDP growth		LTCC: DV = pc GDP (2000)			
	RGMC	RC: IHST	RGMC	RC: IHST	RC: WB regions	RC: IHST + WB regions
Post offices[1]	0.02	0.02	0.174	0.158	0.116	0.103
(Std. error)	(0.006)	(0.006)	(0.082)	(0.081)	(0.089)	(0.088)
p-value	0.001[2]	0.000	0.040[2]	0.058	0.200	0.252
Number of observations	1,540	1,527	77	77	77	77
Name of display item in RGMC	Table 1		Table 1			
Column	2		4			

Notes: RC = Robustness Check. IHST = Inverse hyperbolic sine transformation. SE = Standard Error. The IHST has been applied to all variables that have zero values. Standard errors in parentheses

[1] The main explanatory variable varies by analysis: STCC uses # post offices (ln, stock); LTCC uses # post offices (ln, 1900).

[2] p-values not reported in RGMC but taken from computational reproduction output in Stata.

Table A2: Results from robustness checks for the STUS and LTUS analyses

	STUS: DV = Farm value								LTUS: DV = Median income		
	RGMC	RC: IHST	RC: PPML	RC: Cluster SEs (state)	RC: Bootstrapped SEs[3]	RC: IHST + Cluster (state)	RC: IHST + Boot.[3]	RC: PPML + Cluster (state)	RGMC	RC: IHST	RC: PPML
Post offices[1]	0.314	0.272	0.418	0.314	0.319	0.272	0.272	0.418	0.071	0.067	0.080
(Std. error)	(0.028)	(0.025)	(0.039)	(0.058)	(0.027)	(0.054)	(0.024)	(0.086)	(0.015)	(0.014)	(0.015)
p-value	0.000[2]	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000[2]	0.000	0.000
Number of observations	12,254	12,254	12,254	12,254	12,254	12,254	12,254	12,254	1,674	1,674	1,674
Name of display item in RGMC	Table 2								Table 3		
Column	2								9		

Notes: RC = Robustness Check. IHST = Inverse hyperbolic sine transformation. PPML = Poisson pseudo-maximum likelihood estimator. SE = Standard Error. The IHST has been applied to all variables that have zero values. Standard errors in parentheses.

[1] Main explanatory variable: STUS uses # post offices (ln); LTUS uses # post offices (ln, 1896).

[2] p-values not reported in RGMC but taken from computational reproduction output in Stata.

[3] Bootstrapping performed with 200 repetitions.