Spatial Interdependence and Instrumental Variable Models – Supplementary Online Appendix

Timm Betz *   Scott J. Cook†   Florian M. Hollenbach‡

November 6, 2018

*Assistant Professor of Political Science, Department of Political Science, Texas A&M University, College Station, TX 77843. Email: timm.betz@tamu.edu, URL: www.people.tamu.edu/~timm.betz

†Assistant Professor of Political Science, Department of Political Science, Texas A&M University, College Station, TX 77843. Email: sjcook@tamu.edu, URL: scottjcook.net

‡Assistant Professor of Political Science, Department of Political Science, Texas A&M University, College Station, TX 77843. Email: fhollenbach@tamu.edu, URL: fhollenbach.org

A Additional Plots Simulation: Misspecified W

What if the researcher is unsure about the spatial network underlying the modeled processes? In the simulations above, we estimated the S-2SLS model based on the correct connectivity matrix. Effectively, this assumes the researcher has complete information on the spatial network, which is often an unrealistic assumption in applied research. Therefore, we perform the same set of simulation experiments as above but we also vary the level of misspecification of the spatial network in the estimation. To do so, we draw a second set of random spatial locations and its corresponding $\mathbf{W}$ matrix. We then create the $\mathbf{W}$ matrix for the model estimation based on binary draws from either the correct $\mathbf{W_c}$ or the false $\mathbf{W_f}$ matrix. The probability of each cell value being drawn from the false matrix is the misspecification parameter. We set this parameter to three different values: 0, 0.5 and 1. The results presented above assumed no misspecification, such that the probability of drawing from the false $\mathbf{W_f}$ matrix is 0.

As Figure A.1 shows, S-2SLS generally outperforms or is equivalent to 2SLS. In this scenario we consider positive correlation between the first and second stage, i.e. sufficient non-spatial endogeneity. In the worst case (bottom row in Figure A.1), when the $\mathbf{W}$ matrix is completely misspecified, S-2SLS parallels 2SLS in performance. As the median absolute error in 2SLS increases, so does the median absolute error for S-2SLS. Similarly, as Figure A.2 shows, as the coverage for 2SLS worsens, so does the coverage for S-2SLS. This demonstrates what we articulated earlier: because S-2SLS nests 2SLS, it only suffers minor efficiency losses when it is the incorrect model.[1]

Put differently, even if researchers have no knowledge of the spatial network in their data and chose a spatial matrix at random, S-2SLS does not perform significantly worse than 2SLS. Conversely, where there is spatial interdependence and some knowledge of the connectivity matrix, the gains from S-2SLS are considerable – S-2SLS provides the more robust and conservative modeling strategy.

[1]For some of the simulations we were unable to estimate the spatial model. This only occurred when the spatial matrix was drawn from two different W matrices. We drop these observations before calculating the performance statistics. The relative performance of S-2SLS does not change if we also drop the corresponding results for OLS and 2SLS.

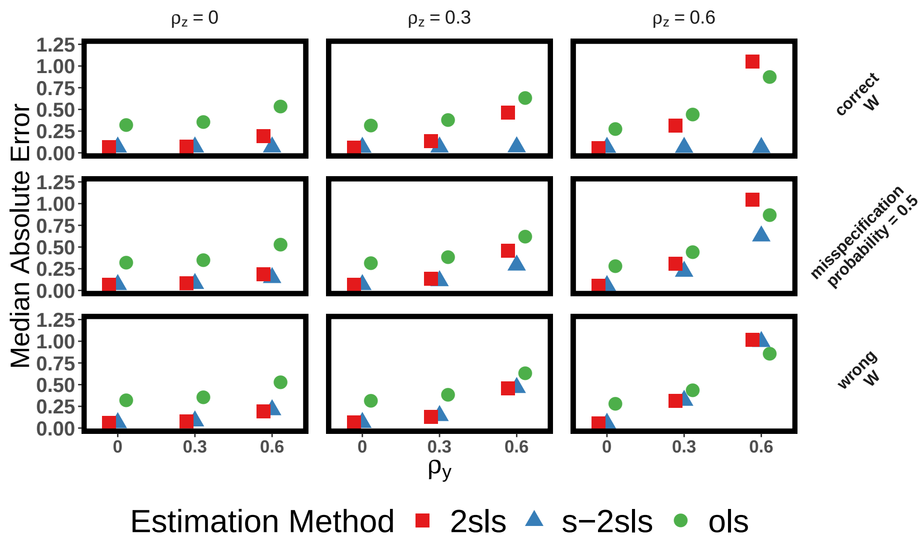

Figure A.1: Median Absolute Error over Misspecification of W ($\lambda = 1.5$ & $\delta = 0.5$ & $N = 200$)

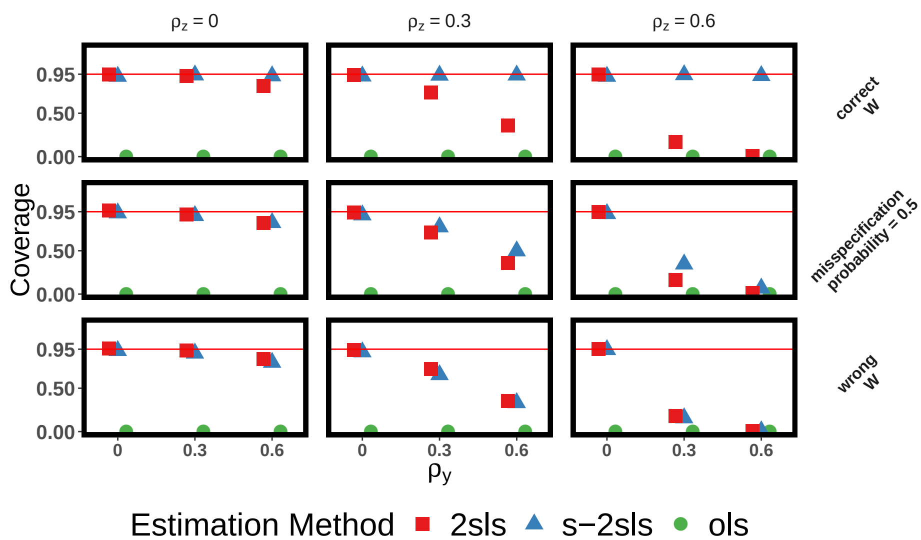

Figure A.2: Coverage over Misspecification of W ($\lambda = 1.5$ & $\delta = 0.5$ & $N = 200$)

B Additional Plots Simulation: MedAE & Coverage

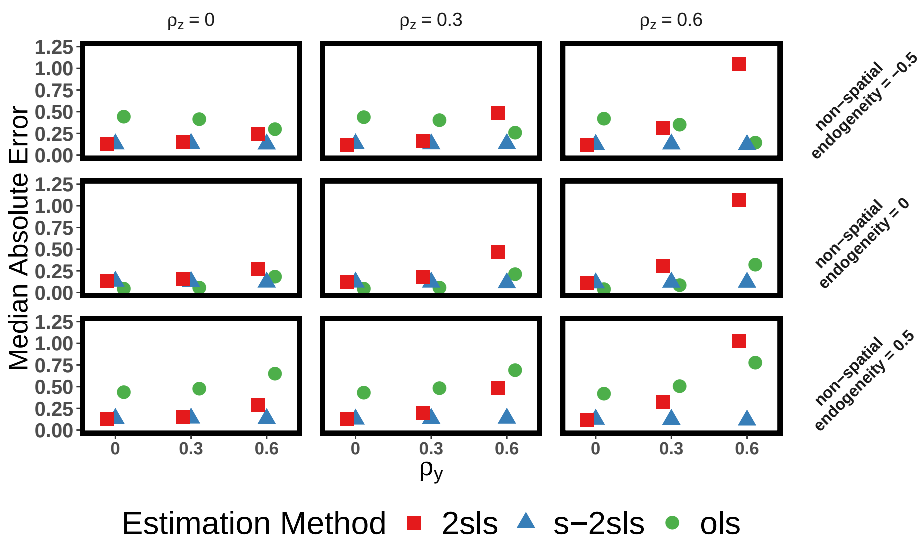

Figure B.3: Median Absolute Error over δ ($\lambda = 0.75$ & $N = 200$)

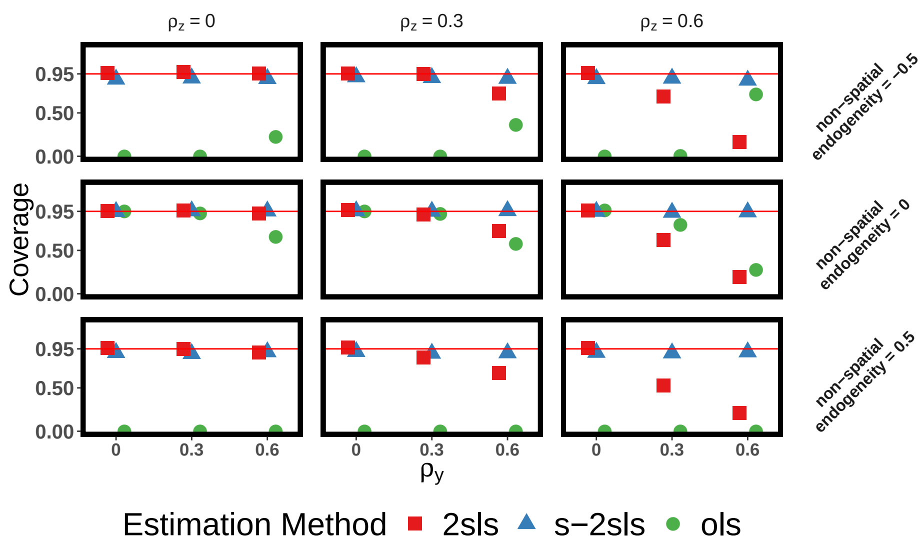

Figure B.4: Coverage over δ ($\lambda = 0.75$ & $N = 200$)

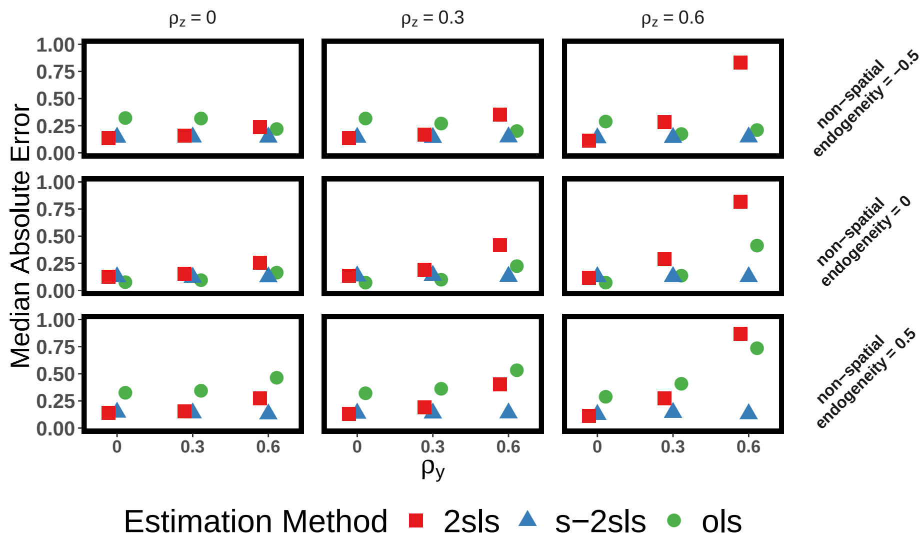

Figure B.5: Median Absolute Error over δ ($\lambda = 1.5$ & $N = 50$)

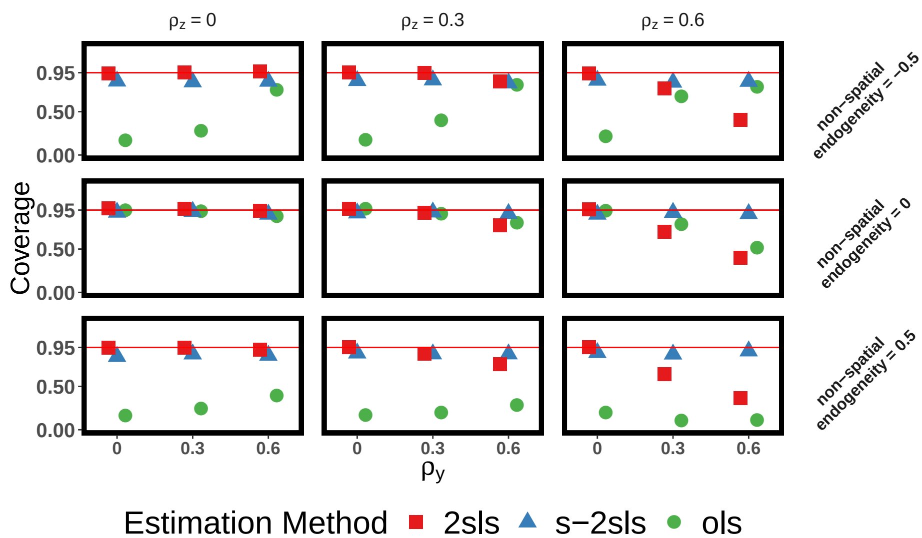

Figure B.6: Coverage over δ ($\lambda = 1.5$ & $N = 50$)

C Additional Plots Simulation: RMSE

Since the RMSE is very sensitive to outliers, we drop any simulation where the absolute error is greater than 10. The only estimation method for which this occurs is standard 2SLS, for which we drop 2217 simulated data sets. Thus, this adjustment actually improves the results for 2SLS.

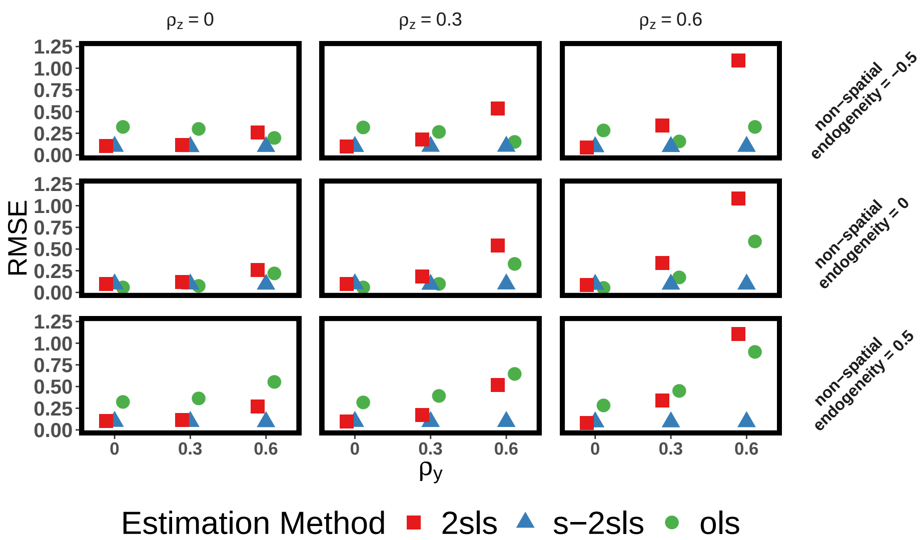

Figure C.7: RMSE over δ ($\lambda = 1.5$ & $N = 200$)

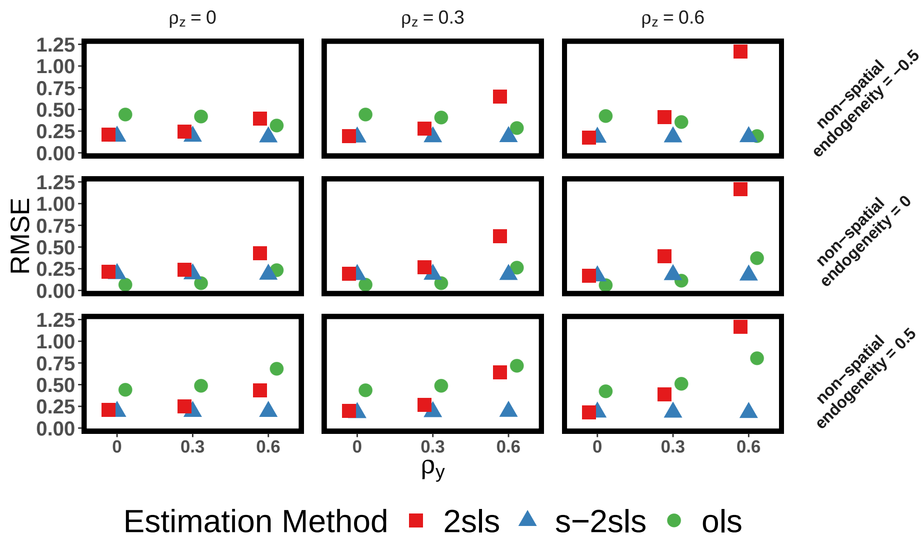

Figure C.8: RMSE over δ ($\lambda = 0.75$ & $N = 200$)

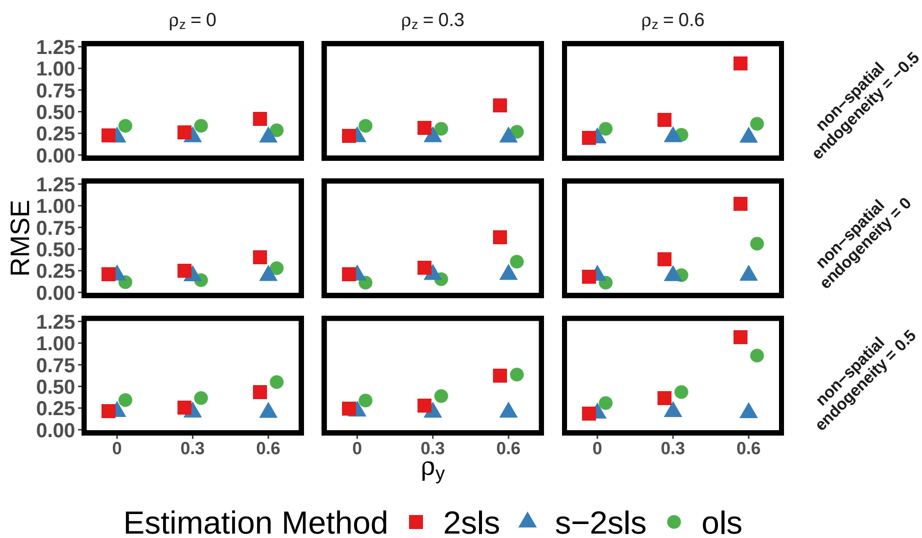

Figure C.9: RMSE over δ ($\lambda = 1.5$ & $N = 50$)

D Additional Tables for Applications

Table D.1: Replication of Robustness Checks, Table 5 Ramsay (2011)

Respective Column Table 5 in Ramsay (2011)	(1) 2SLS	(1) S-SLS	(2) 2SLS	(2) S-SLS	(3) 2SLS	(3) S-SLS	(4) 2SLS	(4) S-SLS	(5) 2SLS	(5) S-SLS	(6) 2SLS	(6) S-SLS	(7) 2SLS	(7) S-SLS
log oil income	-0.358**	-0.0867***	-0.261***	-0.0997***	-0.358**	-0.0725***	-0.358**	-0.0867***	-0.260***	-0.0897***	-0.357***	-0.177***	-0.453*	-0.0989***
per capita	(0.167)	(0.0113)	(0.0837)	(0.00970)	(0.167)	(0.00914)	(0.167)	(0.0113)	(0.0823)	(0.00883)	(0.0769)	(0.0161)	(0.252)	(0.0115)
log gdp	0.356**	0.106***	0.300***	0.129***	0.350**	0.0944***	0.356**	0.106***	0.294***	0.118***	0.371***	0.202***	0.492*	0.122***
per capita	(0.155)	(0.0113)	(0.0914)	(0.0114)	(0.151)	(0.00932)	(0.155)	(0.0113)	(0.0878)	(0.0104)	(0.0767)	(0.0172)	(0.265)	(0.0133)
gdp growth	-0.0118**	-0.00496***	-0.00851***	-0.00503***	-0.0118**	-0.00461***	-0.0118**	-0.00496***	-0.00851***	-0.00482***	-0.0112***	-0.00601***	-0.0127*	-0.00514***
	(0.00518)	(0.00102)	(0.00268)	(0.00103)	(0.00518)	(0.000975)	(0.00518)	(0.00102)	(0.00265)	(0.000998)	(0.00361)	(0.00155)	(0.00655)	(0.00103)
polity at entry	-0.00517	0.567***	0.226	0.548***	0.00684	0.595***	-0.00517	0.567***	0.239	0.567***			-0.0507	0.549***
	(0.373)	(0.0284)	(0.184)	(0.0251)	(0.363)	(0.0242)	(0.373)	(0.0284)	(0.177)	(0.0232)			(0.463)	(0.0271)
Latitude			-0.00622**	-0.00192***					-0.00630**	-0.00164***				
			(0.00247)	(0.000457)					(0.00249)	(0.000440)				
top5 oil producers					0.127	-0.0160			0.105*	-0.000276				
					(0.0948)	(0.0224)			(0.0627)	(0.0233)				
coldwar dummy							0	-0.327***						
							(.)	(0.0588)						
west dummy											-0.101	0.201***	-0.235	0.00790
											(0.165)	(0.0402)	(0.218)	(0.0238)
sub-Saharan Africa													0.462*	0.0949***
													(0.276)	(0.0233)
Constant	-0.210	-0.327***	-0.245*	-0.0670	-0.172	-0.0539	-0.210	0	-0.215	-0.0687	-0.318	-0.504***	-0.755	0
	(0.183)	(0.0512)	(0.149)	(0.0608)	(0.170)	(0.0500)	(0.183)	(.)	(0.140)	(0.0517)	(0.205)	(0.102)	(0.535)	(.)
Spatial ρ		0.119***		0.0989***		0.131***		0.119***		0.108***		0.113***		0.130***
		(0.0201)		(0.0214)		(0.0193)		(0.0201)		(0.0208)		(0.0346)		(0.0201)
Observations	1263	1263	1263	1263	1263	1263	1263	1263	1263	1263	1263	1263	1263	1263
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

E Additional Application: "Dynamics and Stagnation in the Malthusian Epoch"

In this section of the Appendix, we provide an additional application. In a 2011 article in the *American Economic Review*, Ashraf and Galor (2011) aim to test a central prediction of the famous Malthusian theory. Thomas R. Malthus (1798) argues that the main reason for stagnating incomes, prior to the industrial revolution, is that when incomes increase, population size rises as well. Since resources are limited, higher populations induce declining living standards. As a result, technological progress or the discovery of new resources only temporarily improves living standards, but does not produce sustained gains (Ashraf and Galor, 2011). As Ashraf and Galor (2011, p. 2004) outline, their article "exploits exogenous sources of cross-country variation in land productivity and technological levels to examine their hypothesized differential effects on population density versus income per capita during the time period 11500 CE." To test the Malthusian theory in pre-industrial societies, Ashraf and Galor (2011) investigate two predictions: 1) a country's improvements in productivity should lead to larger populations, but not higher living standards; and 2) countries with higher land productivity, or better technology, should have higher population densities, but again, should not be significantly richer.

In their empirical analysis, Ashraf and Galor (2011) use the timing of the onset of the neolithic revolution to proxy for technological change. Consistent with their expectations, the authors show that both the onset of the neolithic revolution and land productivity are positively (and significantly) associated with population density, but not with income per capita. In addition, Ashraf and Galor (2011) use instrumental variables to estimate the causal effect of technological progress on population density. They argue that "prehistoric biogeographical endownments," in particular the "availability of domesticable species of plants and animals," have had an important effect on technological progress and are otherwise exogenous (Ashraf and Galor, 2011, pp. 2029-2031). The use of the instrumental variable is primarily motivated by the authors to estimate the "causal impact of technology on population density" (Ashraf and Galor, 2011, p. 2031).

However, the authors ignore possible spatial interdependence in both the instrumental variables and the dependent variable. Both population density and natural wildlife are likely to be spatially clustered. In other words, it is likely that the animal and plant species found in one country are similar to those in adjacent regions. Likewise, in pre-historic times (i.e., 1000 CE), it is likely that some parts of the planet had higher population density than others, again reflecting positive spatial correlation. This does not mean that these variables are similarly clustered in space, but rather that by themselves they might exhibit (positive) spatial dependence. If correct, this would induce bias in their IV models for the reasons outlined above.

To test this, we first need to specify an appropriate weights matrix. Here, we create a row-standardized-contiguity matrix, with neighbors defined as having adjacent borders.[2] As a preliminary test of spatial autocorrelation, we estimate Moran's I based on the residuals of the original OLS model – with logged population density in 1000 CE as the dependent variable (column 2, Table 9 in Ashraf and Galor (2011)). Based on a Moran's I value of 0.375 with an associated p-value smaller than 0.001, we are able to reject the null of independence of the residuals.

Table E.2 shows the results of replicating the models with population density in 1000 CE (Table 9 in Ashraf and Galor (2011)). Column 1 replicates the original OLS model on the restricted sample (column 2 in Table 9 in Ashraf and Galor (2011)). As a first step, column 2 in Table E.2 shows the results when we estimate a spatial autoregressive (SAR) model instead of the standard OLS model. As one can see, the main coefficients of interest (technological index) have the same levels of significance as in the original OLS results. The effect estimates, however, are quite different. For the linear-additive model, the direct effect is simply the reported coefficient estimate on the log technology index (3.21). The average direct effect of the SAR model – calculated as above – is 3.52 – larger than the coefficient estimate for SAR (due to expected feedback effects), but still substantially smaller than the OLS effect estimates of 4.198.

Columns 3 and 4 replicate the instrumental variable model for population density in 1000 CE as presented in Table 9 in Ashraf and Galor (2011). The differences in results between the original

[2] We have also replicated the results with a k(=5) nearest-neighbor matrix, the results are even more striking.

2SLS model and the spatial 2SLS model are stark. The coefficient on technological progress (log of technological index) in the original 2SLS model is 14.53, almost 3.5 times as large as the OLS coefficients. Ashraf and Galor (2011) argue that the difference in estimated coefficients is "a pattern that is consistent with measurement error in the transition-timing variable and the resultant attenuation bias afflicting OLS coefficient estimates" (Ashraf and Galor, 2011, p. 2031). Column 4, however, shows the results from the model estimated with S-2SLS. Here the coefficient for technological progress is much smaller compared to 2SLS, with the average direct effect – calculated as before – being 7.03. In fact, the average effect estimate of technological progress in the spatial 2SLS model is comparable to that in the original OLS estimates. Recall, that, as we show above, the non-spatial and spatial bias in OLS can be offsetting. This may be the case here. If the non-spatial measurement bias is attenuating and the spatial bias is upward, the OLS model ends up being less biased than the 2SLS model due to the countervailing forces of both biases on the coefficient estimate.

We note that the overall conclusion of Ashraf and Galor (2011) still stands.[3] The Malthusian theory for pre-industrial times is supported by these data. On the other hand, the causal effect of technological progress on population density is smaller than the standard 2SLS model indicates and is about the same size the original estimates in the OLS models in Ashraf and Galor (2011).

[3] In the Appendix we also replicate the results using population density in 1CE as the dependent variable (Table E.3), producing similar results.

Table E.2: Replication of Table 9 (1000 CE) in Ashraf and Galor (2011)

	(1) Original OLS	(2) SAR	(3) Original 2SLS	(4) S-2SLS
Log technology index in relevant period	4.198***	3.209***	14.53***	6.492***
	(1.164)	(0.938)	(4.437)	(1.614)
Log land productivity	0.498***	0.430***	0.572***	0.458***
	(0.139)	(0.0951)	(0.148)	(0.103)
Log absolute latitude	-0.185	-0.124	-0.209	-0.136
	(0.151)	(0.105)	(0.209)	(0.113)
Mean distance to nearest coast or river	-0.363	-0.509	-1.155*	-0.746*
	(0.426)	(0.359)	(0.640)	(0.395)
Percentage of land within 100 km of coast or river	0.442	0.306	0.153	0.224
	(0.422)	(0.340)	(0.606)	(0.366)
Constant	-1.820***	-1.311**	-5.507***	-2.493***
	(0.641)	(0.529)	(1.702)	(0.730)
Spatial ρ_y		0.455***		0.427***
		(0.0768)		(0.118)
Observations	92	92	92	92
Continent dummies	Yes	Yes	Yes	Yes

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table E.3: Replication of Table 9 (1 CE) in Ashraf and Galor (2011)

	(1)	(2)	(3)	(4)
	Original OLS	SAR	Original 2SLS	S-2SLS
Log technology index in	3.947^{***}	2.990^{***}	10.80^{***}	3.542^{***}
relevant period	(0.983)	(0.722)	(2.857)	(1.273)
Log land productivity	0.350^{**}	0.319^{***}	0.464^{**}	0.327^{***}
	(0.172)	(0.0999)	(0.182)	(0.101)
Log absolute latitude	0.0834	0.00837	-0.0521	-0.00830
	(0.170)	(0.108)	(0.214)	(0.109)
Mean distance to nearest	-0.625	-0.693^{*}	-0.616	-0.696^{*}
coast or river	(0.434)	(0.368)	(0.834)	(0.366)
Percentage of land within	0.146	-0.00297	-0.172	-0.0404
100 km of coast or river	(0.424)	(0.336)	(0.642)	(0.338)
Constant	-2.719^{***}	-1.564^{***}	-4.770^{***}	-1.676^{***}
	(0.601)	(0.476)	(0.980)	(0.624)
Spatial ρ		0.536^{***}		0.569^{***}
		(0.0725)		(0.117)
Observations	83	83	83	83
Continent dummies	Yes	Yes	Yes	Yes

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

References

Ashraf, Quamrul and Oded Galor. 2011. "Dynamics and Stagnation in the Malthusian Epoch." *American Economic Review* 101(5):2003—2041.

Malthus, Thomas R. 1798. *An Essay on the Principle of Population*. Vol. Reprint, ed. Geoffrey Gilbert,1999 Oxford, UK: Oxford University Press.

Ramsay, Kristopher W. 2011. "Revisiting the Resource Curse: Natural Disasters, the Price of Oil, and Democracy." *International Organization* 65(3):507–530.