Urban Economics
2022-1
Paula Torres Higuera
February 18th, 2022

Replication project for *The Wrong Side(s) of the Tracks: The Causal Effects of Racial Segregation on Urban Poverty and Inequality*

The Wrong Side(s) of the Tracks is a paper written by Professor Elizabeth Oltmans Ananat with the purpose of contributing to the literature regarding the negative relationship between segregation and several population outcomes. The major contribution of Ananat's paper is her instrumental variables (IV) approach to establish causality between segregation and various city-level characteristics using the arrangements of railroad tracks in the nineteenth century to identify exogenous variation in a city's exposure to segregation. The instrument's validity is proved through several falsification checks and analyses that corroborate the fulfillment of IV assumptions.

Data used to estimate the causal effect between segregation and population outcomes comes primarily from the US Census Bureau reports on metropolitan demographics, individual Census microdata from the Integrated Public Use Microdata Series, measures of metropolitan segregation and information for proximity of the city to the nearest former slave state. The sample consists of 121 cities that were not slave-owning at the time of the Civil War since those cities had a small African American population prior to the Great Migration which guarantees that railroad design in those areas was not arranged to segregate.

The primary econometric method is an Instrumental Variables specification to account for omitted variable bias (for example, unmeasured political attributes that directly affect both segregation and city outcomes) to correct endogeneity and therefore to identify the causal effect of segregation on net population outcomes. Ananat uses the railroad division index (RDI), defined as a measure of division of a city's land into subunits generated by railroads, as the instrument. The specification for the two-stage least squares (2SLS) estimation is defined as:

$$Seg = \alpha_1 RDI + \alpha_2 \mathbf{X} + \mu$$

$$Y = \beta_1 Seg + \beta_2 \mathbf{X} + \varepsilon,$$

Where *Seg* is the level of segregation, *RDI* is the instrument, *Y* is the outcome of interest and *X* is a vector of control variables.

For RDI to be a valid instrument it must satisfy four major assumptions. First, there should exist a strong first stage which means that the index must generate a significant level of racial segregation. This is known as the relevance assumption. Second, a higher RDI index induces more segregation on cities affected by the instrument; this is known as the monotonicity assumption. In this case, RDI can affect segregation since a city with more subunits has more clear boundaries that serve to enforce segregation, and therefore implies RDI has a non-

negative effect on segregation, which validates the previous two assumptions. Third, the effect of RDI on population outcomes for each city must only be induced through segregation, which means that the exclusion restriction must be satisfied. The evidence presented in the paper shows that RDI did not alter city characteristics prior to the Great Migration, and only influences population outcomes in cities that experienced substantial black migration. Lastly, people's city choices during the Great Migration cannot be correlated with railroad division.

Ananat presents various tests, including first stage estimation, falsification checks and historical facts to demonstrate that her chosen instrument satisfies all four assumptions and is thus a valid instrument. Intuitively, railroads cover enough land to offer a clear boundary that allows for communal consensus on neighborhood borders by the community. Furthermore, railroads were constructed to optimize placement and variation in ground slope before 1900, when most African Americans lived in the South, and more than 20 years before the Great Migration, which means that they were not constructed to segregate or influence city outcomes.

Since it is demonstrated that RDI is a valid instrument, the results derived from estimating the model allow to identify casual relationships between segregation and city-level outcomes. Particularly, the estimation is a LATE for cities where higher RDI induced more segregation (the compliers). The estimations use 121 cities close enough to the South to have experienced a major population shift during the Great Migration, but far enough to guarantee a very small African American population prior to this era. The main results from the 2SLS model evidence the negative impact that race segregation has on population outcomes. It is shown that exogenous segregation leads cities to have higher poverty rates among African Americans and white communities with lower poverty rates. These results extend to other outcomes such as migration, housing demand and human capital; in all cases, segregation is seen as a disamenity. Therefore, segregation raises inequality between black and white communities in a city, mainly by lowering black peoples' outcomes, while inducing more equal outcomes within its white community.

To validate her findings, Ananat presents robustness checks by estimating the 2SLS model including (1) city characteristics and (2) 1920 characteristics -the beginning of the Great Migration- as controls. Results from city controls and 1920 controls are robust which means that estimates both with and without controls are not significantly different. Therefore, the selected instrument is affecting dependent variables through segregation and not through another mechanism.

The major findings, validation tests and robustness checks are presented in six tables that are replicable using the provided code. Ananat also presents some maps and two graphs that are not replicable with the provided code. An example of a replicable finding from the original paper is Table 2 – Panel 1. I present the replication in table 1 using the existing code. However, the code provided by the author does not include code to export table as .tex which is the final output in the paper. Therefore, I included outreg2 in the code to export these

results. The actual output presented in the paper will be replicated in a subsequent derivable, following provided instructions.

Table 1 – Replication of Table 2 – Panel 1 - The Effects of Segregation on Poverty and Inequality Among Blacks and Whites. 2SLS RDI as instrument for 1990 dissimilarity

	(1)	(2)	(3)	(4)
VARIABLES	Gini index - Whites	Gini index - Blacks	Poverty rate - Whites	Poverty rate - Blacks
Dissimilarity Index	-0.334***	0.875**	-0.196***	0.258**
	(0.0993)	(0.409)	(0.0648)	(0.108)
Constant	-0.730***	-1.331***	0.205***	0.121**
	(0.0547)	(0.229)	(0.0368)	(0.0613)
Observations	121	121	121	121

Robust Standard Errors in Parentheses. * p<0.01, ** p<0.05, * p<0.1.

Results presented in the previous table are based on a 2SLS estimation where RDI is the instrument for dissimilarity during 1999. Dissimilarity is defined as the degree of segregation within a city. These within-race findings evidence that higher segregation in a city reduces the Gini index and poverty between white people, while augmenting income inequality and poverty between black people. Particularly, a 0.1 increase in dissimilarity reduces the Gini index in 0.0334 for white communities and augments it in 0.0875 for black communities, while reducing the poverty rate in 0.0196 for white people and augmenting it in 0.0258 for blacks. All these results are LATE for the compliers and significant at the 5% level. The following table from the original paper shows that the previous table is in fact replicating major findings from Ananat's work, particularly the estimates in the yellow box:

TABLE 2—THE EFFECTS OF SEGREGATION ON POVERTY AND INEQUALITY AMONG BLACKS AND WHITES

	OLS: Effect of 1990 dissimilarity index		Main results: 2SLS RDI as instrument for 1990 dissimilarity		Falsification: Reduced form effect of RDI among cities far from the south	
Outcome:	Whites (1)	Blacks (2)	Whites (3)	Blacks (4)	Whites (5)	Blacks (6)
Within-race poverty and inequality						
Gini index	**−0.079**	**0.459**	**−0.334**	**0.875**	−0.110	0.167
	(0.037)	**(0.093)**	**(0.099)**	**(0.409)**	(0.066)	(0.424)
Poverty rate	**−0.073**	**0.182**	**−0.196**	**0.258**	−0.036	−0.136
	(0.019)	**(0.045)**	**(0.065)**	**(0.108)**	(0.035)	(0.094)
	White:black ratios		White:black ratios		White:black ratios	
Between-race inequality						
90 white: 90 black	0.111		−0.131		**−0.443**	
	(0.086)		(0.312)		**(0.217)**	
10 white: 10 black	**1.295**		**2.727**		−0.135	
	(0.249)		**(0.867)**		(0.532)	
90 white: 10 black	**1.172**		**1.789**		−0.449	
	(0.282)		**(0.758)**		(0.558)	
90 black: 10 white	−0.234		**−0.807**		0.130	
	(0.131)		**(0.384)**		(0.248)	
N	121		121		29	

Notes: 2SLS and reduced form estimates control for total track length. All outcomes except poverty rates are logged. Robust standard errors in parentheses.

3. GitHub Repository: https://github.com/ptorresh/Revised-reproduction-package-for-Oltmans-2011-.git

References:

Ananat, E. O. (2011). The wrong side (s) of the tracks: The causal effects of racial segregation on urban poverty and inequality. *American Economic Journal: Applied Economics*, *3*(2), 34-66.