Appendix For
The Evolution of Access to Public Accommodations in the United States

Last Updated: February 13, 2022

Lisa Cook
Michigan State & NBER

Maggie Jones
University of Victoria

Trevon Logan
The Ohio State University & NBER

David Rosé
Wilfrid Laurier University

Contents

I Data Sources and Variable Construction

I.A Main Controls

We draw on a number of supplementary data sources to construct the set of controls included in our main difference-in-differences results table. The following section outlines these sources and data construction in more detail.

Logan and Parman Residential Segregation:
We control for three measures of residential segregation, each of which was constructed by Logan and Parman (2017) using the complete-count 1940 Census of Population. These measures include the Dissimilarity and Isolation indices, as well as the Logan-Parman index.

1940 Full Count Census of Population
We use the full-count 1940 Census of Population, available through Ruggles et al. (2020), the share of the Black population who migrated within the state within the last 5 years, the share of the Black population who migrated between states within the last 5 years, the level of educational attainment for working age (15-64 years old) White and Black residents separately (grouped by low, high school, and no school), the share of Black residents living in a rural area, the labour force participation rate for the working age population (by race), and the number of Black employees who work in the postal service.

Southern Poverty Law Centre Confederate Symbols
Using the list of Confederate symbols and monuments, which includes schools named after prominent Confederate generals, as well as monuments and other notable symbols, we construct the number of Confederate symbols located in each county in the U.S. by 1940. These data were acquired from the Southern Poverty Law Centre's project, *"Whose Heritage? Public Symbols of the Confederacy,"* available for download here: `https://www.splcenter.org/20190201/whose-heritage-public-symbols-confederacy`.

New National Lynching Database
Using a newly constructed nation-wide dataset on the location and instance of all lynchings in the U.S. that builds on Cook (2012), we construct the number of Black lynchings and the number of White lynchings that occurred up to 1936 by county.

Haines' County Data Books

We use the Historical, Demographic, Economic, and Social Data: The United States, 1790-2002, by Michael Haines to control for manufacturing activity, war contracts, the percent with electricity, the percent with a radio, and the percent with a refrigerator, and the population of Black and White residents in 1930, 1940, 1950, and 1960 (we use linear interpolation to construct annual county-level population by race). This is also used to construct the percent of homes that are owner occupied.

I.B Census of Business

We digitized entries from the 1935 and 1948 U.S. Census of Business to obtain county-level establishment counts for formal accommodations (hotels and motels), eating and drinking businesses, and gasoline service stations. In 1935 the disclosure criteria for formal accommodations are not explicitly stated; instead, establishment counts are available for counties that were deemed to have adequate coverage by the census enumerators (i.e. reliable data) and enough hotels or hotel chains to prevent disclosing the identity of a particular operator. In 1948, all counties with a population of at least 2,500 inhabitants are listed. Establishment counts for eating and drinking businesses and gasoline service stations are published in the Retail Trade Area Statistics volumes of the 1935 and 1948 Censuses of Business. In both publication years all counties with a population of at least 2,500 inhabitants are listed (in Table 10 of Volume 3 in 1935 and in Table 103 of Volume 3 in 1948).

The 1935 Census of Business only lists hotels explicitly (in a separate Hotels volume); however, it is our contention that the 1935 definition would have included both categories of establishment. This is supported by the fact that the inclusion criteria are quite detailed and many other types are explicitly excluded (boarding houses, Y.M.C.A.s, and tourist camps, for example). Moreover, the decision to list them separately most likely reflects the increased prominence of tourist courts following the end of WWII gasoline rationing and the return of troops following the cessation of hostilities. The heyday of motels was during the 1940s and 1950s (Jackson, 1993). The 1948 Census of Business lists hotels and tourist courts (also known as motels) separately in the Service Trade Area Statistics (in Tables 103C and 103D of Volume 7, respectively).

I.C Time-Consistent Counties

We use 1940 counties as the basic unit of observation for our analysis. In order to account for county border changes over time and the formation of new counties (or disappearance of old ones) we use Ferrara et al. (2021)'s county crosswalks with population based weights (model 1). These weights are based on a 1×1km grid-cell population distributions produced by Fang and Jawitz (2018). An alternative is to use area based weights, these were popularized

by Hornbeck (2010) which imposes the assumption that the variables being transformed are uniformly spread across the area of a county; however, for most demographic variables this does not hold (most counties have urban and rural areas). Ferrara et al. (2021) show that weights that account for patterns of population distribution within counties outperform area-based weights.

II Additional Descriptive Work

II.A Frequency of Establishments

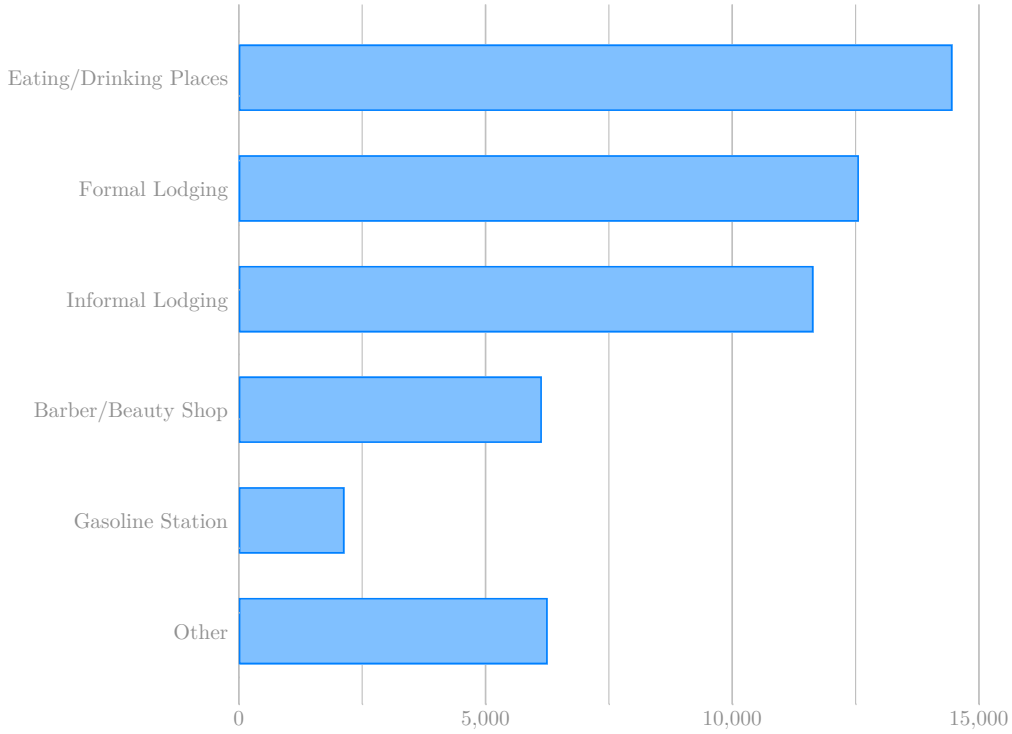

Figure I: Frequency of establishments between 1938 and 1955 listed by type of establishment.

II.B Wisconsin Black Business Directory

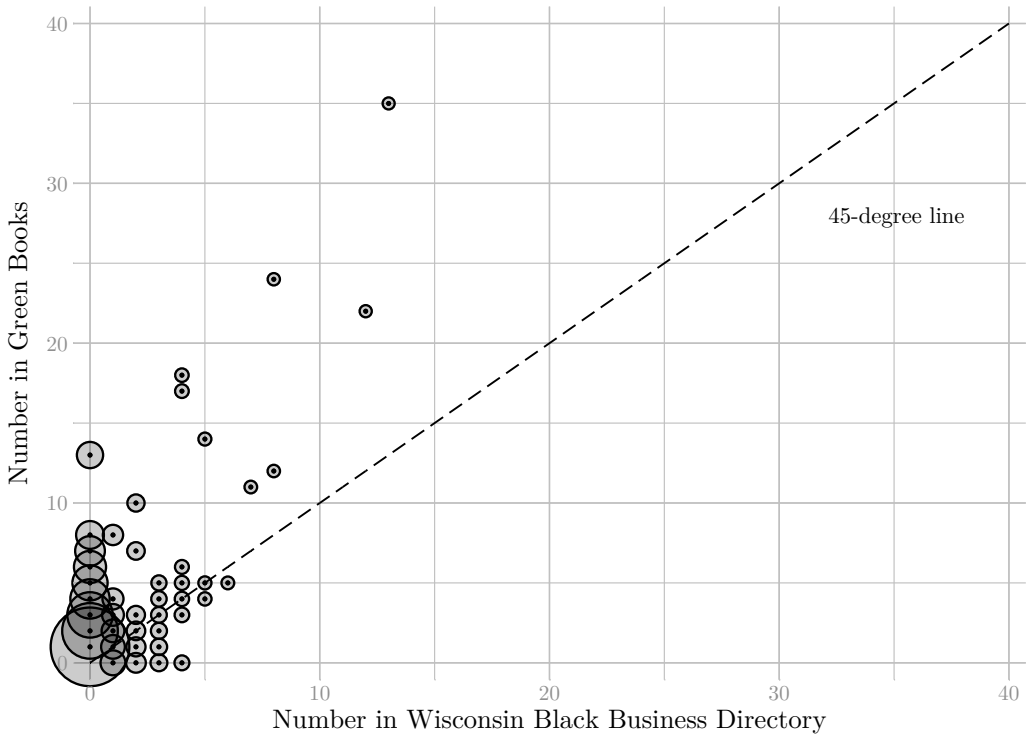

Figure II: City-level relationship between the number of formal accommodations in the Green Books and the number of hotels listed in the Wisconsin Black Business Directory

Notes: Each dot has been weighted by the frequency this observation is observed in the Green Book & Wisconsin Black Business Directory (WBBD) data. A 45 degree line has been drawn to compare the number of instances that the Green Books list more establishments than the WBBD and vice-versa.

II.C Pauli Murray without outliers

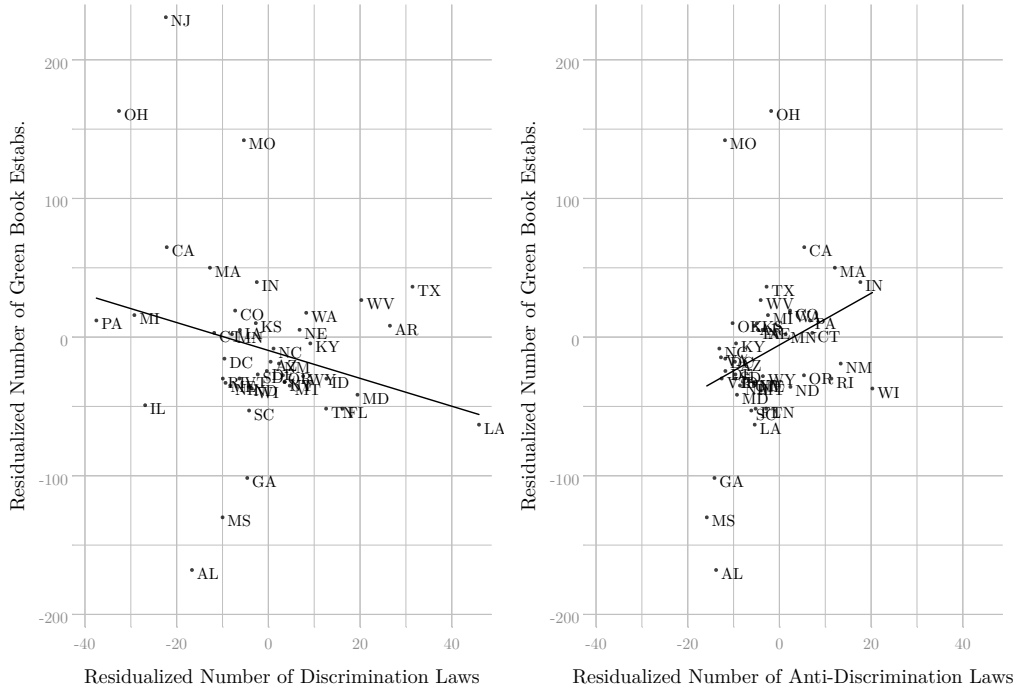

Figure III: Correlation between the number of Green Book establishments and the number of laws upholding discrimination by 1950 (left panel) and the number of anti-discrimination laws by 1950 (right panel).

Notes: All variables have been residualized using the Black population and outliers have been excluded. New York, Oklahoma, and Virginia were dropped for the discrimination laws plot; Illinois, New Jersey, and New York were dropped for the anti-discrimination laws plot. Discrimination and anti-discrimination laws come from Murray (1950).

II.D Correlations with the Share of Non-Discriminatory Establishments

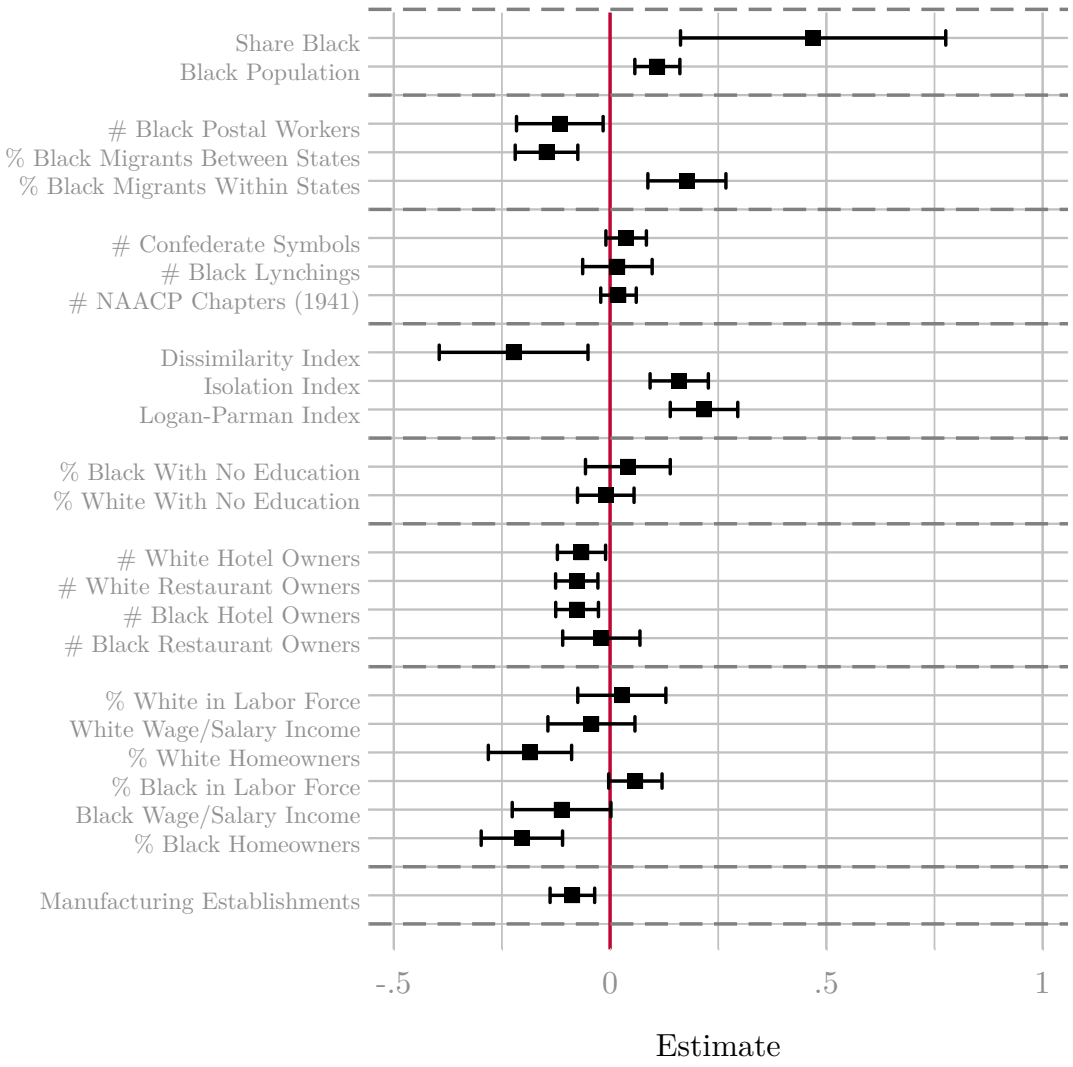

Figure IV: Correlates of the share of hotels listed in the Green Books, conditional on the Black population

Notes: Each point represents the magnitude of the coefficient estimate on the regressor associated with the label on the vertical axis. Estimates are from separate regressions of the share of hotels listed in the Green Books on each correlate, conditional on the Black population. Error bands represent 95% confidence intervals computed using robust standard errors.

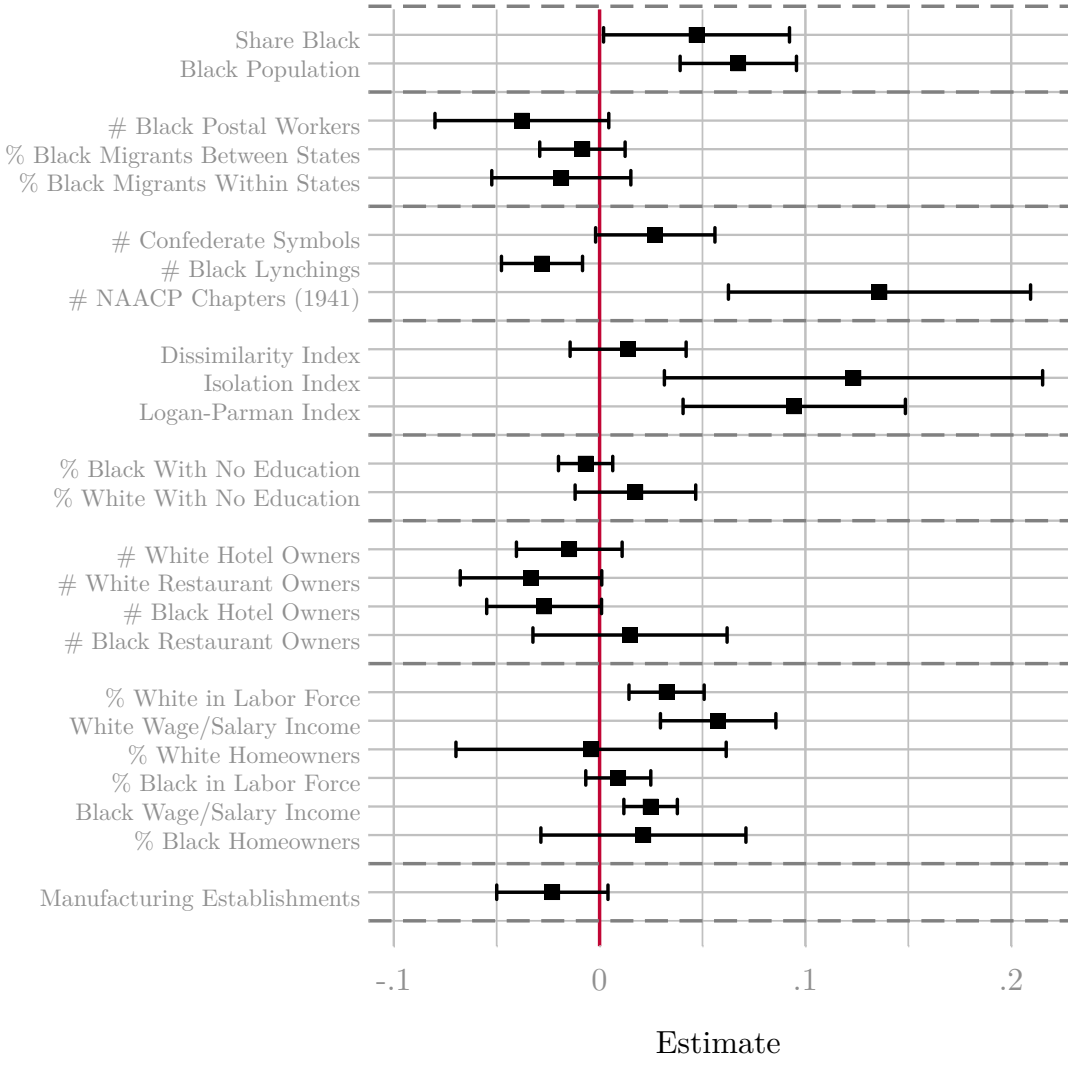

Figure V: Correlates of the share of eating and drinking establishments listed in the Green Books, conditional on the Black population

Notes: Each point represents the magnitude of the coefficient estimate on the regressor associated with the label on the vertical axis. Estimates are from separate regressions of the share of eating and drinking establishments listed in the Green Books on each correlate, conditional on the Black population. Error bands represent 95% confidence intervals computed using robust standard errors.

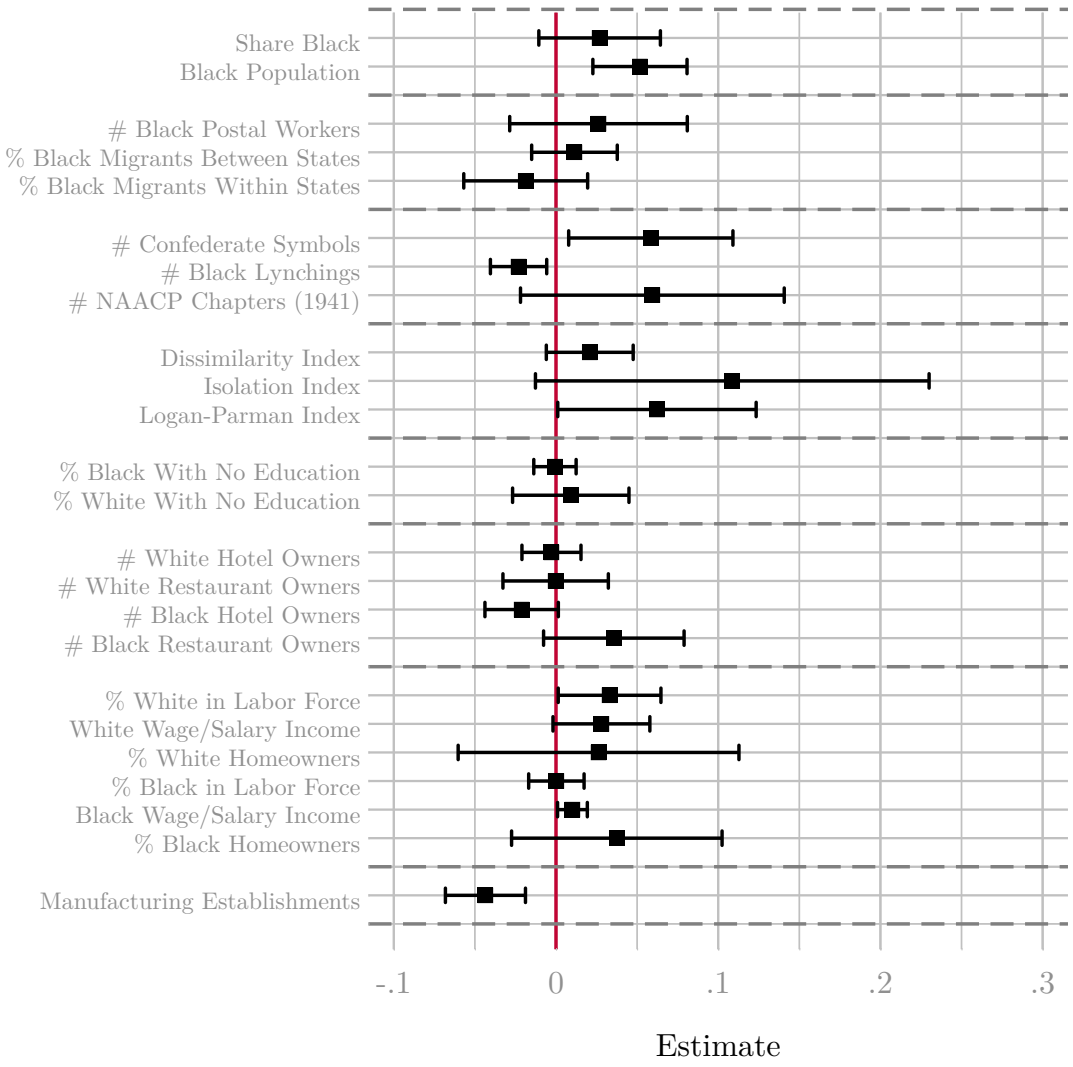

Figure VI: Correlates of the share of gas stations listed in the Green Books, conditional on the Black population

Notes: Each point represents the magnitude of the coefficient estimate on the regressor associated with the label on the vertical axis. Estimates are from separate regressions of the share of gas stations listed in the Green Books on each correlate, conditional on the Black population. Error bands represent 95% confidence intervals computed using robust standard errors.

III Difference-in-Differences Specification

III.A Balance Table

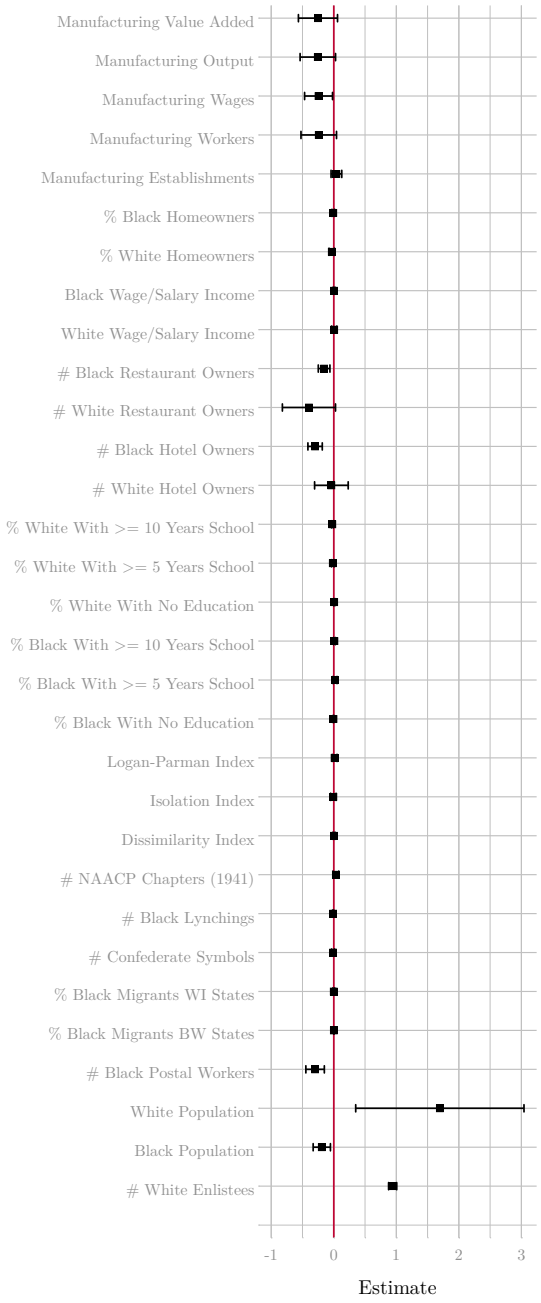

Figure VII: Balance table for White casualties

Notes: Coefficient estimates and 95% confidence bands from regressing the number of White casualties on each regressor separately. All specifications include state fixed effects and control for the total population in 1940. The dependent and independent variables of interest have been standardized to have mean 0 and variance 1.

III.B Difference-in-differences with Draftee Mortality

Here we present results analogous to those found in Table III of our paper. As noted in the text, one concern is that volunteer enlistment into the armed forces may be correlated with attitudes towards African Americans. For example, Qian and Tabellini (2021) show that Black men volunteered less for service in locations where racial discrimination was more pronounced. Alternatively, (White) ethnic or religious minorities may have volunteered to participate in the war deferentially, and they may have had different attitudes towards Black Americans than other Whites. These concerns are largely attenuated by the fact that in December, 1942 the federal government prohibited volunteering to avoid manpower shortages in sectors of the economy that were critical for the war effort (Ferrara, 2022); however, we have also estimated our difference-in-differences specifications using county-level exposure to deaths of White draftees as the treatment variable. The results are presented in Table I, restricting the analysis to drafted servicemen only does not have a big impact on results; overall, the results are very similar in both a qualitative and quantitative sense.

Table I: Effects of White Casualties on the Number of Establishments

Panel A: Main specification						
	(1)	(2)	(3)	(4)	(5)	(6)
	Total	Total	Total	Total	Total	Total
Asinh(# White Deaths) × Post-WW2	0.0605***	0.0605***	0.0196*	0.0605***	0.0880***	0.0447***
	(0.008)	(0.008)	(0.011)	(0.009)	(0.012)	(0.012)
County Controls			X			X
County F.E.				X	X	
Year F.E.			X	X		X
State F.E.		X				
State X Year F.E.					X	
County X Linear Trends						X
Observations	37248	37248	37248	37248	37248	37248
Adjusted R^2	0.170	0.248	0.624	0.906	0.909	0.958
Clusters	3104	3104	3104	3104	3104	3104
Panel B: Results by industry						
	(1)	(2)	(3)	(4)	(5)	(6)
	Barber & Beauty	Eating & Drinking	Gasoline Station	Formal Lodging	Informal Lodging	Other Retail & Service
Asinh(# White Deaths) × Post-WW2	0.0515***	0.0686***	0.0242***	0.0179***	0.00493	0.0552***
	(0.008)	(0.009)	(0.005)	(0.005)	(0.005)	(0.008)
County F.E.	X	X	X	X	X	X
Year F.E.	X	X	X	X	X	X
Observations	37248	37248	37248	37248	37248	37248
Adjusted R^2	0.817	0.844	0.766	0.865	0.878	0.787
Clusters	3104	3104	3104	3104	3104	3104
Panel C: Results by industry shares						
	(1)	(2)	(3)	(4)	(5)	(6)
	Formal Lodging	Formal Lodging	Eating & Drinking	Eating & Drinking	Gasoline Station	Gasoline Station
Asinh(# White Deaths) × Post-WW2	0.000389	0.00147*	0.000229***	0.000401***	-0.0000583	0.000104
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
County F.E.	X	X	X	X	X	X
Year F.E.	X		X		X	
State X Year F.E.		X		X		X
Observations	23953	23953	36620	36620	36684	36684
Adjusted R^2	0.668	0.677	0.744	0.744	0.638	0.637
Clusters	2957	2957	3071	3071	3070	3070

Notes: The dependent variable in panel A is the inverse hyperbolic sine of the number of Green Book establishments; in panel B it is the inverse hyperbolic sine of the number of each type of establishment; and in panel C it is the inverse hyperbolic sine of the share of all establishments that are listed in the Green Books (separately for formal lodging, eating and drinking establishments, and gasoline stations). County controls, all transformed using the inverse hyperbolic sine, include the White and Black population, the number of Black postal workers, the share of Black migrants from out of state, the share of Black migrants from in-state, the number of confederate symbols, the number of Black lynchings, the number of NAACP chapters, the dissimilarity index, the isolation index, the Logan-Parman segregation index, the share of farmland, the share of the Black and White populations with each level of education, the number of Black and White hotel owners and restaurant owners, the share of Black and White adults in the labor force, the average income of Black and White workers in the labor force, the number of Black and White homeowners, and average manufacturing establishments, workers, wages, output, and value added. We include dummy variables for missing controls and we interact all controls and dummies with a post-WWII indicator. See the online appendix for a complete list of variable definitions and sources. Standard errors clustered by county in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

IV Additional Information on the IV Specification

This section provides additional details on the IV specifications. Specifically, we discuss the construction of the Black migration instrument. This is followed by a discussion of Tables II, III, and IV which present additional IV results for a variety of additional specification checks.

IV.A Constructing the Black Migration Instrument

The migration instrument takes the form:

$$Z_c^{1950-1940} = \frac{\sum_{j=1}^{J} \pi_{jc}^{1940} \times M_j^{1950-1940}}{\text{Total Pop}^{1940}},$$

$M_j^{1950-1940}$ represents the difference in the number of Black Southern migrants from state j between 1940 and 1950, and π_{jc}^{1940} is the share of Black migrants from state j who live in county c in 1940. We define a Black Southern migrant from state j as someone who was born in state j but who resides in another state at the time of the census. Our results are robust to defining a migrant as someone who lives in a different state of residence in 1940 from their state of residence in 1935. The objective of this exercise is to show that our IV results using White casualties are robust to using a different, but comparable, source of variation in the Black population. For this reason, we abstract from using predicted migration, as is used in other literature on the Great Migration, e.g. Boustan (2009); Derenoncourt (2022).

IV.B Variation from Other Time Periods

Unlike the majority of other industries listed in the Green Books, formal accommodations were listed consistently until the Green Books ceased publication after the Civil Rights Act of 1964. This means that while we cannot examine variation from other time periods using eating and drinking establishments or gas stations, we can evaluate whether formal accommodations respond to exogenous variation in the Black population between 1950 and 1960. Column (1) of Table II estimates the IV specification using the Bartik shock as exogenous variation in change in the share of the Black population between 1950 and 1960. Both the OLS and IV estimates are remarkably similar to those using variation between 1940 and 1950. Part of this similarity is driven by the fact that the "share" portion of the migration instrument is the same in this specification. Namely, since we are missing county of residence for a large number of counties in the 1950 public-use data, we are unable to construct the share of each county who were born in each southern state. We therefore use the 1940 values instead. Nevertheless, the fact that the estimates are so close in magnitude despite using changes in population, the share of Green Book establishments, and "shifts" from Black Southern out-migration between 1950 and 1960, is a useful reliability check on

our original estimates.

IV.C Functional Form

Our main specification uses a functional form that is consistent with the literature on the Great Migration so that we could compare our estimates using variation from White World War II casualties to variation using the Black migration instrument constructed from Black Southern out-migration. However, this specification constructs changes in the share of Black residents as a fraction of the initial population. Following this specification, we also constructed our dependent variable a chance in the share of Green Book establishments as a fraction of the total number of initial establishments. This means, that our specification does not directly account for changes in the total number of establishments throughout this time period, other than through what is captured in by differencing out county-level fixed effects. We show that our main specification is robust to a variety of different function form choices. Specifically:

1. The inverse hyperbolic sine of the difference between the share in 1950 and 1940 (column 2 in Table II and column 1 in Tables III and IV),

$$\text{Asinh}(\text{Share}_{1950} - \text{Share}_{1940})$$

2. The percent change in the share (column 3 in Table II and column 2 in Tables III and IV),

$$\frac{Share_{1950} - Share_{1940}}{Share_{1940}}$$

3. Directly controlling for the change in the total number of establishments in each industry equation (column 4 in Table II and column 3 in Tables III and IV).

Table II: The relationship between the change in the share of Black residents and the change in the share of formal accommodations listed in the Green Books

	(1)	(2)	(3)	(4)
Panel A: OLS Estimates				
Asinh(Δ Share Black$_{1950-1960}$)	0.0880*			
	(0.050)			
Asinh(Share$_{1950}$ - Share$_{1940}$)		0.0744		
		(0.046)		
$\frac{Share_{1950}-Share_{1940}}{Share_{1940}}$			0.416***	
			(0.117)	
Asinh(Δ Share Black$_{1940-1950}$)				0.0709***
				(0.025)
Adjusted R^2	0.016	0.012	0.068	0.004
Panel B: IV Estimates				
Asinh(Δ Share Black$_{1950-1960}$)	0.375			
	(0.315)			
Asinh(Share$_{1950}$ - Share$_{1940}$)		0.509		
		(0.314)		
$\frac{Share_{1950}-Share_{1940}}{Share_{1940}}$			0.722***	
			(0.256)	
Asinh(Δ Share Black$_{1940-1950}$)				0.302***
				(0.087)
IV=White Casualties		X	X	X
IV=Black Migration	X			
Observations	1235	1909	161	1909
1st stage F-stat	30.98	41.07	43.45	171.0

Notes: The dependent variable in all specifications takes the same functional form as the independent variable of interest. For example, in the first row, the dependent variable is the inverse hyperbolic sine of Δ GB$_{1950-1960} = \frac{\#\text{ GB}_{1960}-\#\text{ GB}_{1950}}{\text{Total }\#\text{ Est}_{1950}}$ to align with the independent variable of interest, Asinh(Δ Share Black$_{1950-1960}$) = Asinh($\frac{\#\text{ Black}_{1960}-\#\text{ Black}_{1950}}{\text{Total Population}_{1950}}$). Column (2) computes the change as the inverse hyperbolic sine of the share of Green Book establishments in 1950 minus the share of Green Book establishments in 1940. Column (3) computes changes in terms of the percent growth from the original share (in 1940). Column (4) uses the original specification from the main text but conditions on the change in the total number of establishments. All columns include state fixed effects. Standard errors reported in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$

Table III: The relationship between the change in the share of Black residents and the change in the share of eating and drinking establishments listed in the Green Books

	(1)	(2)	(3)
Panel A: OLS Estimates			
Asinh(Share_{1950} - Share_{1940})	0.0166^{**}		
	(0.008)		
$\frac{Share_{1950}-Share_{1940}}{Share_{1940}}$		0.357	
		(0.221)	
Asinh(Δ Share $\text{Black}_{1940-1950}$)			0.0230^{***}
			(0.005)
Adjusted R^2	0.002	0.013	0.010
Panel B: IV Estimates			
Asinh(Share_{1950} - Share_{1940})	0.353^{***}		
	(0.121)		
$\frac{Share_{1950}-Share_{1940}}{Share_{1940}}$		1.713^{***}	
		(0.511)	
Asinh(Δ Share $\text{Black}_{1940-1950}$)			0.0778^{***}
			(0.023)
IV=White Casualties	X	X	X
Observations	3050	125	3050
1st stage F-stat	20.80	39.08	153.2

Notes: The dependent variable in all specifications takes the same functional form as the independent variable of interest. For example, in the first row, the dependent variable computes the change as the inverse hyperbolic sine of the share of Green Book establishments in 1950 minus the share of Green Book establishments in 1940. Column (2) computes changes in terms of the percent growth from the original share (in 1940). Column (3) uses the original specification from the main text but conditions on the change in the total number of establishments. All columns include state fixed effects. Standard errors reported in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$

Table IV: The relationship between the change in the share of Black residents and the change in the share of gas stations listed in the Green Books

	(1)	(2)	(3)
Panel A: OLS Estimates			
Asinh(Share_{1950} - Share_{1940})	0.00320		
	(0.008)		
$\frac{Share_{1950}-Share_{1940}}{Share_{1940}}$		0.759**	
		(0.356)	
Asinh(Δ Share $\text{Black}_{1940-1950}$)			0.00641*
			(0.003)
Adjusted R^2	-0.005	0.079	-0.004
Panel B: IV Estimates			
Asinh(Share_{1950} - Share_{1940})	0.149		
	(0.104)		
$\frac{Share_{1950}-Share_{1940}}{Share_{1940}}$		1.225*	
		(0.734)	
Asinh(Δ Share $\text{Black}_{1940-1950}$)			0.0243
			(0.015)
IV=White Casualties	X	X	X
Observations	3056	42	3056
1st stage F-stat	19.66	12.21	150.1

Notes: The dependent variable in all specifications takes the same functional form as the independent variable of interest. For example, in the first row, the dependent variable computes the change as the inverse hyperbolic sine of the share of Green Book establishments in 1950 minus the share of Green Book establishments in 1940. Column (2) computes changes in terms of the percent growth from the original share (in 1940). Column (3) uses the original specification from the main text but conditions on the change in the total number of establishments (3). All columns include state fixed effects. Standard errors reported in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$

References

Boustan, Leah (2009) "Competition in the Promised Land: Black Migration and Racial Wage Convergence in the North, 1940-1970," *The Journal of Economic History*, 69 (3), 756–783.

Cook, Lisa D. (2012) "Overcoming Discrimination by Consumers during the Age of Segregation: The Example of Garrett Morgan," *Business History Review*, 86 (2), 211–234.

Derenoncourt, Ellora (2022) "Can you Move to Opportunity? Evidence from the Great Migration," *American Economic Review*, 112 (2), 369–408.

Fang, Yu and James W. Jawitz (2018) "High-Resolution Reconstruction of the United States Human Population Distribution, 1790 to 2010," *Scientific data*, 5 (1), 1–15.

Ferrara, Andreas (2022) "World War II and Black Economic Progress," *Journal of Labour Economics*.

Ferrara, Andreas, Patrick A Testa, and Liyang Zhou (2021) "New Area-and Population-Based Geographic Crosswalks for U.S. Counties and Congressional Districts, 1790-2020,"Technical report, Competitive Advantage in the Global Economy (CAGE).

Hornbeck, Richard (2010) "Barbed Wire: Property Rights and Agricultural Development," *The Quarterly Journal of Economics*, 125 (2), 767–810.

Jackson, Kirstin (1993) "The World's First Motel Rests Upon its Memories," *The Seattle Times*.

Logan, Trevon D. and John M. Parman (2017) "The National Rise in Residential Segregation," *The Journal of Economic History*, 77 (1), 127–170.

Murray, Pauli ed. (1950) *States' Laws on Race and Color*: Cincinnati, Woman's Division of Christian Service, Board of Missions and Church Extension, Methodist Church.

Qian, Nancy and Marco Tabellini (2021) "Discrimination and State Capacity: Evidence from WWII US Army Enlistment,"Technical report, National Bureau of Economic Research.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek (2020) "IPUMS USA: Version 10.0 [dataset]," doi.org/10.18128/D010.V10.0.