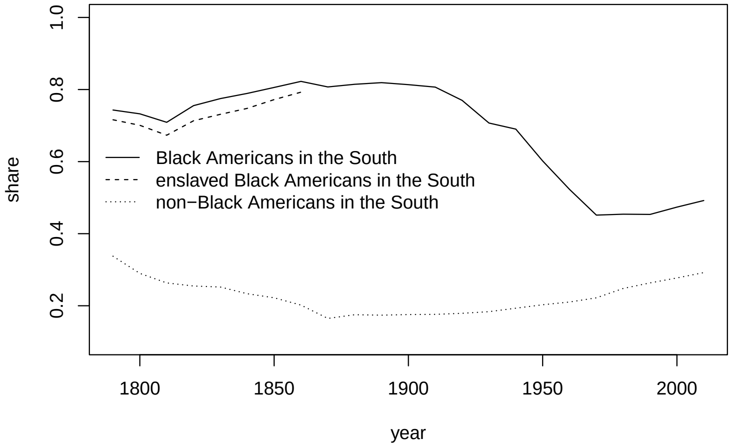
The Aggregate Effects of the Great Black Migration

Motoaki Takahashi

Population Shares in the South by Race



South: confederate states

Outline

- Four million Black Americans moved from the South to the North of the US between 1940 and 1970.
- How did it impact aggregate US output and the welfare of cohorts of Black and non-Black Americans?
- I quantify a dynamic general equilibrium model that comprises migration behavior of Black and non-Black Americans.

Preview

- Shutting down the migration of Black Americans across the North and the South between 1940 and 1970
  - decreases aggregate US output in 1970 by 0.7%,
  - decreases the welfare of Black Americans born in the South in the 1930s by 2.2%,
  - increases the welfare of Black Americans born in the North in the 1930s by 0.1%.
- Shutting down the migration of non-Black Americans across the North and the South for the same period
  - decreases aggregate US output in 1970 by 0.3%.

Contribution to Literature

1. Economic geography of Black Americans
   - Myrdal (1944)
   - **Boustan (2009**, 2010, 2017), Derenoncourt (2022), Althoff and Reichardt (2022)
2. Dynamic spatial models
   - Caliendo, Dvorkin, and Parro (2019), **Allen and Donaldson (2022)**, Kleinman, Liu, and Redding (2022)

- This paper is the first to quantify the aggregate, general equilibrium effects of the great Black migration.

Empirical Facts

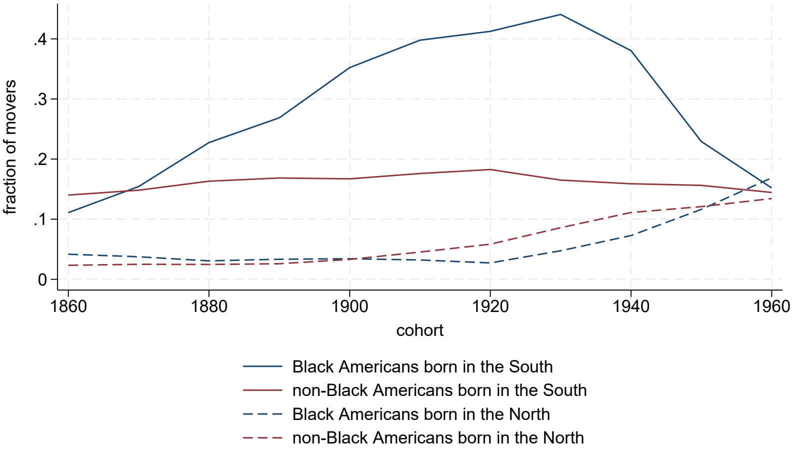
Movers and Stayers

For each cohort c, birthplace (the North or the South), race (Black or non-Black Americans),

- stayers live in the birthplace as of year $c + 50$,
- movers live in the other place than the birthplace as of year $c + 50$.

Fractions of Movers

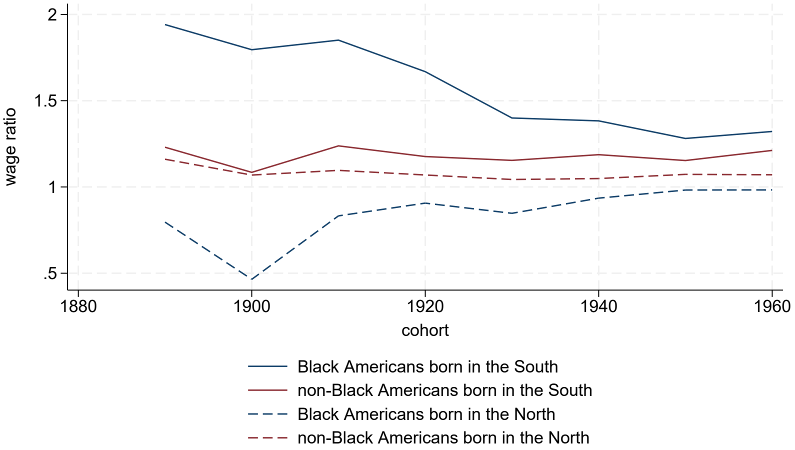
for Cohort c as of Year $c+50$



$c+40$

Mover-Stayer Wage Ratios

Cohort c as of year $c + 50$



gaps in dollar value | ratios in per capita payrolls

Wage and Rent Gaps between Movers and Stayers

for Black Americans from the South

gaps in per capita payrolls and rent

Summary of Facts

1. The migration rate of Black Americans from the South was higher than any other group of people.

2. Black Americans who moved from the South to the North earned much higher wages than Black Americans who stayed in the South.

3. The mover-stayer rent gap was about one-fourth of the mover-stayer wage gap for Black Americans from the South.

Model

Environment

- Time $t = 0, 1, \cdots$
- There are J locations.
- Individuals of cohort c are born in period c and live through at most period $c + \bar{a}$.
- Ages range from 0 to $\bar{a}$.
- A race is either Black or non-Black ($r = b, n$).

Preferences and Location Choices

- The period utility of individuals is

$$u^i_{r,a,t} = \begin{cases} 0 & \text{for } a = 0, \\ \log\left(\frac{w^i_{r,a,t}}{(r^i_t)^\gamma}\right) + \log B^i_{r,a,t}, & \text{for } a = 1, \cdots, \bar{a}. \end{cases}$$

- For $a \leq \bar{a} - 1$, the value is

$$v^i_{r,a,t} = u^i_{r,a,t} + \max_{j=1,\cdots,J} \left\{ s_{r,a,t} E[v^j_{r,a+1,t+1}] - \tau^{j,i}_{r,a,t} + \nu \varepsilon^j_{r,a,t} \right\}.$$

- For $a = \bar{a}$, the value is

$$v^i_{r,a,t} = u^i_{r,a,t}.$$

- Assuming $\varepsilon^j_{r,a,t}$ draws a type-I extreme value, for $a \leq \bar{a} - 1$, the expected value is

$$V^i_{r,a,t} = u^i_{r,a,t} + \nu \log \left(\sum_{j=1}^{J} \exp(s_{r,a,t} V^j_{r,a+1,t+1} - \tau^{j,i}_{r,a,t})^{1/\nu} \right). \tag{1}$$

Migration Flows and Populations

- The migration share of (r, a, t) from i to j is

$$\mu^{j,i}_{r,a,t} = \frac{\exp(s_{r,a,t} V^{j}_{r,a+1,t+1} - \tau^{j,i}_{r,a,t})^{1/\nu}}{\sum_{k=1}^{J} \exp\left(s_{r,a,t} V^{k}_{r,a+1,t+1} - \tau^{k,i}_{r,a,t}\right)^{1/\nu}}. \tag{2}$$

- Population in each demographic group next period is

$$L^{j}_{r,a+1,t+1} = \sum_{i=1}^{J} \mu^{j,i}_{r,a,t} s_{r,a,t} L^{j}_{r,a,t} + I^{j}_{r,a+1,t+1}. \tag{3}$$

Production

- Output is
$$Y_t^i = A_t^i L_t^i.$$
- L_t^i aggregates labor of various ages
$$L_t^i = \left(\sum_{a=1}^{\bar{a}} (\kappa_{a,t}^i)^{\frac{1}{\sigma_0}} (L_{a,t}^i)^{\frac{\sigma_0 - 1}{\sigma_0}} \right)^{\frac{\sigma_0}{\sigma_0 - 1}}.$$
- $L_{a,t}^i$ aggregates labor of different races
$$L_{a,t}^i = \left(\sum_{r=b,n} (\kappa_{r,a,t}^i)^{\frac{1}{\sigma_1}} (L_{r,a,t}^i)^{\frac{\sigma_1 - 1}{\sigma_1}} \right)^{\frac{\sigma_1}{\sigma_1 - 1}}.$$
- Wages are priced at the marginal product of labor
$$w_{r,a,t}^i = A_t^i (L_t^i)^{\frac{1}{\sigma_0}} (\kappa_{a,t}^i)^{\frac{1}{\sigma_0}} (L_{a,t}^i)^{-\frac{1}{\sigma_0} + \frac{1}{\sigma_1}} (\kappa_{r,a,t}^i)^{\frac{1}{\sigma_1}} (L_{r,a,t}^i)^{-\frac{1}{\sigma_1}}. \tag{4}$$

wage ratios across races

Fertility

- Newborns in period t are

$$L^i_{r,0,t} = \sum_{a=1}^{\bar{a}} \alpha_{r,a,t} L^i_{r,a,t}. \tag{5}$$

- $\alpha_{r,a,t}$: how many babies are born per one person of (r, a, t).

Rent

- Rent depends on a location-specific shifter and local income

$$r_t^i = \bar{r}^i \left(\gamma \sum_{r=b,n} \sum_{a=1}^{\bar{a}} L_{r,a,t}^i w_{r,a,t}^i \right)^{\eta}. \tag{6}$$

- Absentee landlords receive rent (or rent is dumped).

Equilibrium

Given $\{L^i_{r,a,0}\}$, an equilibrium is

- $\{V^i_{r,a,t}\}$ such that (1),
- $\{w^i_{r,a,t}\}$ such that (4),
- $\{L^i_{r,a,t}\}$ such that (3) and (5),
- $\{\mu^{i,j}_{r,a,t}\}$ such that (2),
- $\{r^i_t\}$ such that (6).

Steady State

A steady state is an equilibrium in which all endogenous variables are time-invariant:

- $\{V^i_{r,a}\}$ such that (1),
- $\{w^i_{r,a}\}$ such that (4),
- $\{L^i_{r,a}\}$ such that (3) and (5),
- $\{\mu^{i,j}_{r,a}\}$ such that (2),
- $\{r^i\}$ such that (6),

dropping time subscripts t from the equations.

Quantification

Data and Units of Observations

- I obtain wages, populations, and migration shares from US censuses 1940-2000 and American Community Survey 2010.

- Races are Black or non-Black.
- Age bins are:

model	0	1	$\cdots$	6
data	1-10	11-20	$\cdots$	61-70

- Locations are 36 US states, DC, and the constructed rest of the North.
  - The rest of the North accounts for
    - 0.1% of the Black population in 1940.
    - 1% of the Black population in 2010.

Elasticity of Substitution across Races

- For location n, age a, period t, the CES production function implies

$$\frac{w^i_{b,a,t}}{w^i_{n,a,t}} = \frac{\left(\kappa^i_{b,a,t}\right)^{\frac{1}{\sigma_1}}\left(L^i_{b,a,t}\right)^{-\frac{1}{\sigma_1}}}{\left(\kappa^i_{n,a,t}\right)^{\frac{1}{\sigma_1}}\left(L^i_{n,a,t}\right)^{-\frac{1}{\sigma_1}}}.$$

- Taking logs of both sides,

$$\log\left(\frac{w^i_{b,a,t}}{w^i_{n,a,t}}\right) = -\frac{1}{\sigma_1}\log\left(\frac{L^i_{b,a,t}}{L^i_{n,a,t}}\right) + \frac{1}{\sigma_1}\log\left(\frac{\kappa^i_{b,a,t}}{\kappa^i_{n,a,t}}\right).$$

Estimation

Following Card (2009)

- The main specification is

$$\log\left(\frac{w^i_{b,a,t}}{w^i_{n,a,t}}\right) = -\frac{1}{\sigma_1}\log\left(\frac{L^i_{b,a,t}}{L^i_{n,a,t}}\right) + f_a + f_t + f_{a,t} + \varepsilon^i_{a,t}.$$

- Construct an IV using shift-share predicted populations

$$\hat{L}^i_{r,a,t} = \sum_{j=1}^{J} \mu^{i,j}_{r,a-1,t-1-X} \cdot s_{r,a-1,t-1} L^j_{r,a-1,t-1}.$$

- I set $X = 2$: the migration shares 20 years before.

Results

Dependent variable:	$\log(w^n_{b,a,t}/w^n_{o,a,t})$	
Model:	OLS	IV
$\log(L^n_{b,a,t}/L^n_{o,a,t})$	-0.1154***	-0.1108***
	(0.0120)	(0.0127)
fixed effects:		
year-age	✓	✓
Observations	1,368	1,368
First-stage F-statistic		91.24

Block bootstrap standard errors are in parentheses. ***: 0.01.

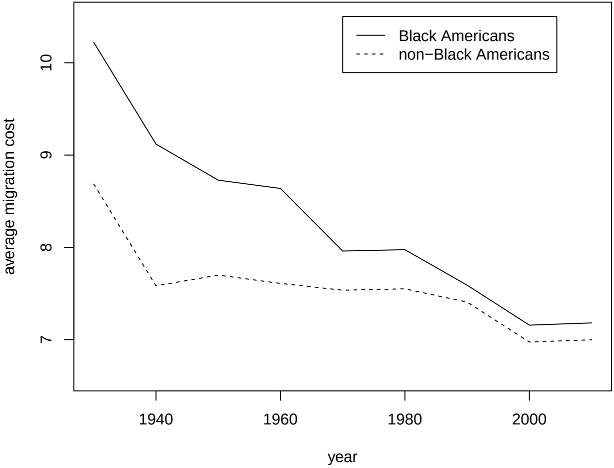
Elasticities

$1/\nu = 0.8$	migration elasticity	details
$\sigma_1 = 9.0$	substitutability across races	details
$\sigma_0 = 2.9$	substitutability across ages	details
$\eta = 0.4$	rent elasticity	details

Other Parameters

- Given the elasticities, inverting the model yields productivity, amenities, and migration costs.
- Fertility $\alpha_{r,a,t}$ is directly observed in census/ACS data.
- Survival probabilities $s_{r,a,t}$ are taken from life tables of CDC.

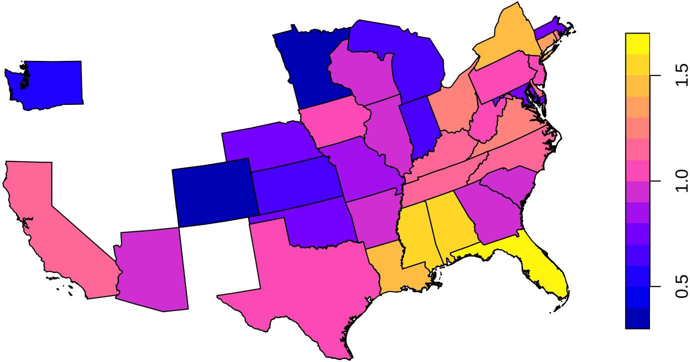
Migration Costs by Year



migration costs by ages

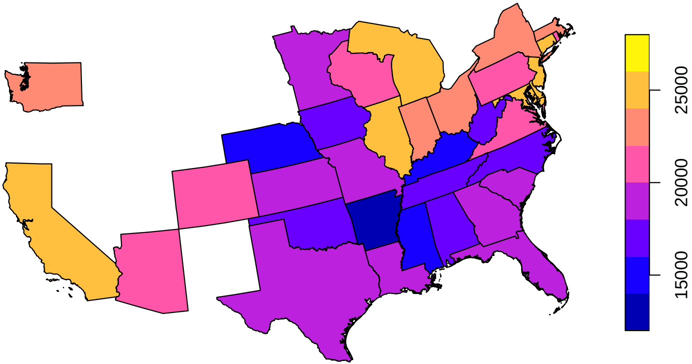
Amenities in 1960

Black Americans



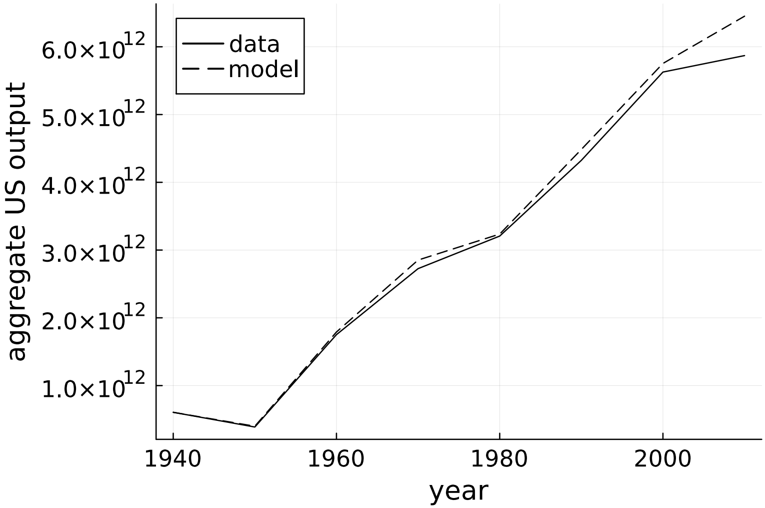
amenities for others in 1960

Productivity in 1960

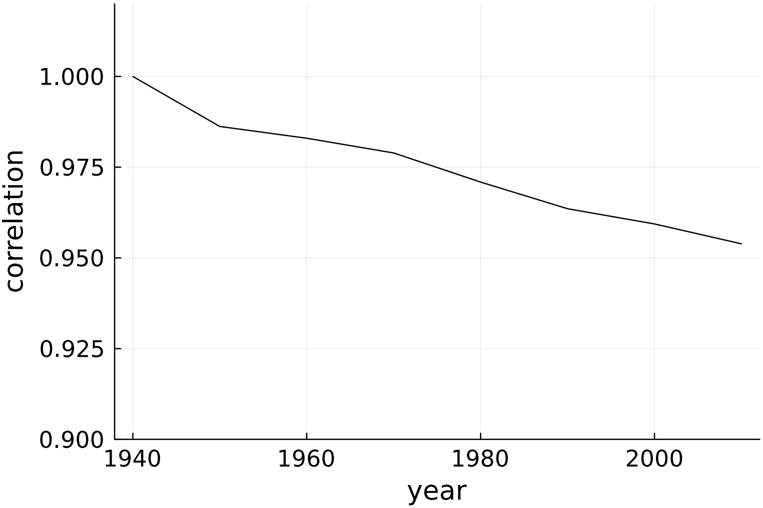


Model Fit

US output: Model vs Data



Populations of Race-Age-Locations: Model vs Data



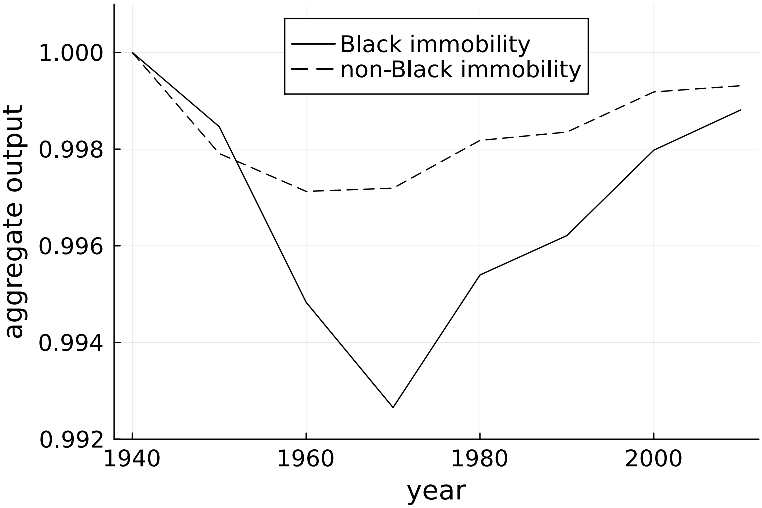
Counterfactuals

Counterfactuals

1. Black Americans cannot move across the North and the South from 1940 to 1960.

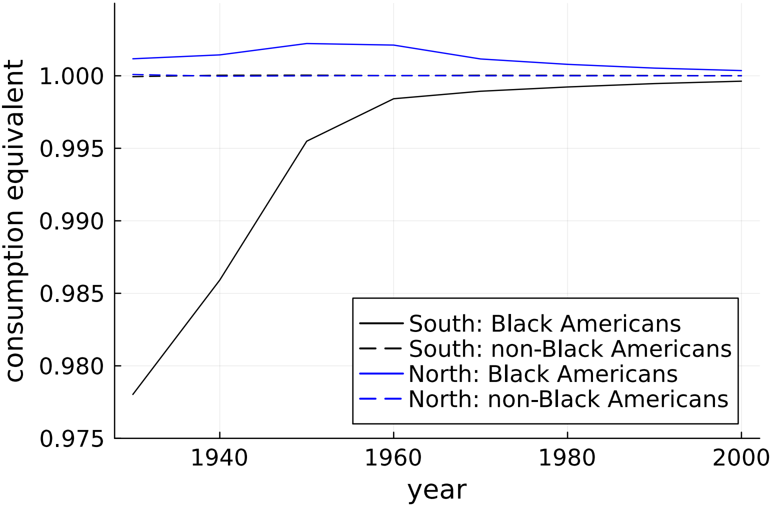
2. Non-Black Americans cannot move across the North and the South for the same period.

US output relative to the Baseline Equilibrium



Welfare

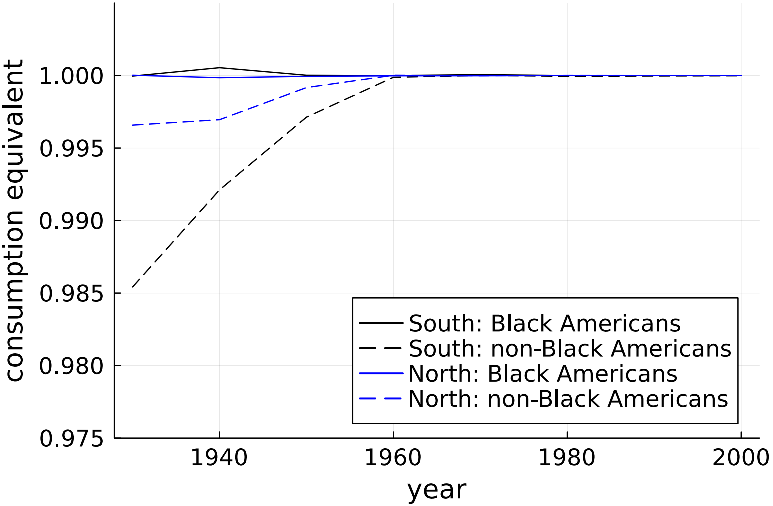
Black Immobility Relative to the Baseline



map: African Americans
welfare formula

Welfare

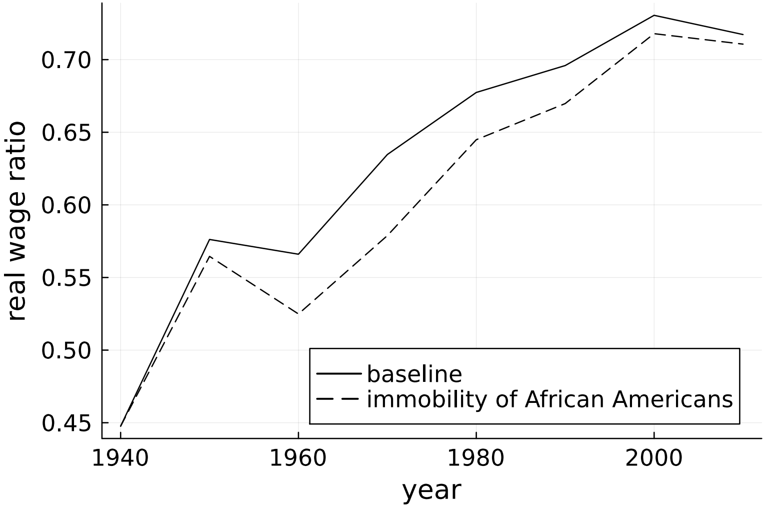
Non-Black Immobility Relative to the Baseline



map: others

Average Real Wage Ratios

between Black and non-Black Americans

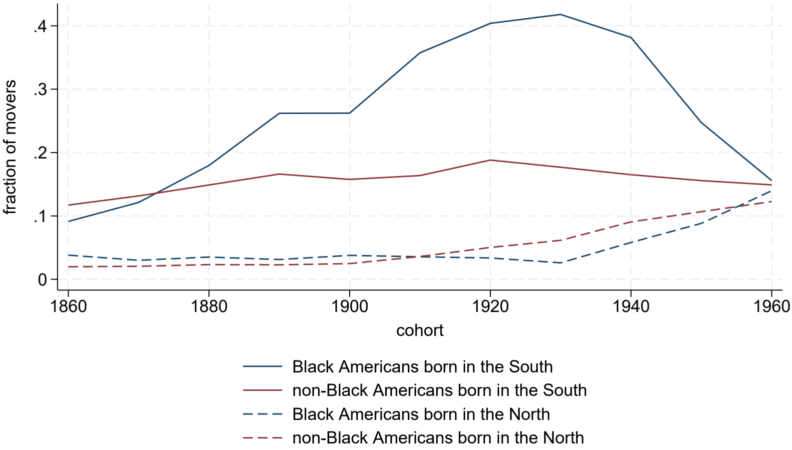


Conclusion

- I quantify the aggregate effects of the great Black migration with a dynamic spatial model.
- Black Americans migrated from the South to the North for higher wages despite their high migration costs and low amenities in the North.
- The mobility of Black and non-Black Americans increased aggregate output in 1970 by 0.7 and 0.3%, respectively.
- The mobility of Black Americans induced
  - a 2.2 percent increase in the welfare of Black Americans in the South,
  - a 0.1 percent decrease in the welfare of Black Americans in the North.

Fractions of Movers

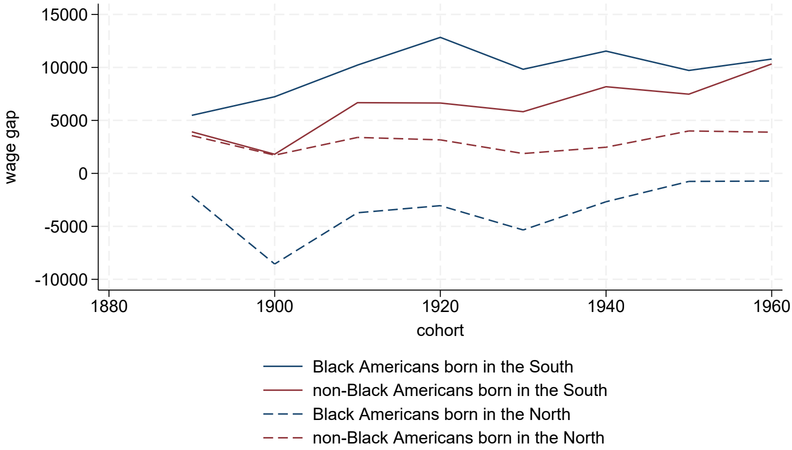
for Cohort c as of Year $c + 40$



back

Mover-Stayer Wage Gaps

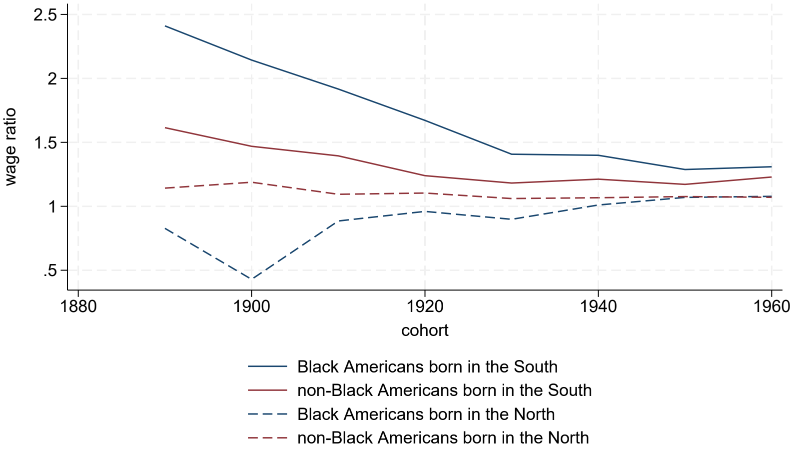
Cohort x as of year $x + 50$



Back

Mover-Stayer Ratios in Per Capita Payroll

Cohort x as of year $x + 50$



Back

Gaps in Per Capita Payroll and Rent

for Black Americans from the South

Back

Relative Wages of Races within Ages

- The relative wages within cohorts are

$$\frac{w^n_{b,a,t}}{w^n_{o,a,t}} = \frac{\left(\kappa^n_{b,a,t}\right)^{\frac{1}{\sigma_1}} \left(L^n_{b,a,t}\right)^{-\frac{1}{\sigma_1}}}{\left(\kappa^n_{o,a,t}\right)^{\frac{1}{\sigma_1}} \left(L^n_{o,a,t}\right)^{-\frac{1}{\sigma_1}}}$$

back

Migration Elasticity

- If real $\text{wage}^{j}_{r,a,t+1}$ increases by 1% *ceteris paribus*, $\mu^{j,i}_{r,a,t}$ increases by $\frac{1}{\nu}$%.

back

Rewriting Expected Values

Toward the estimation of the migration elasticity

- The expected value is the period utility plus the option value.

$$
\begin{aligned}
V^i_{r,a,t} &= u^i_{r,a,t} + \nu \log \left(\sum_{j=1}^{J} \exp(s_{r,a,t} V^j_{r,a+1,t+1} - \tau^{j,i}_{r,a,t})^{1/\nu} \right) \\
&= u^i_{r,a,t} + \Omega^i_{r,a,t}.
\end{aligned}
$$

Decomposing Migration

- Using $\Omega^j_{r,a,t}$, I can write migrants of (r,a,t) from i to j as

$$L^i_{r,a,t}\mu^{j,i}_{r,a,t} = \exp\left\{\frac{1}{\nu}(s_{r,a,t}V^j_{r,a+1,t+1} - \tau^{j,i}_{r,a,t}) - \frac{1}{\nu}\Omega^i_{r,a,t} + \log(L^i_{r,a,t})\right\}$$

- Destination and origin fixed effects capture $V^j_{r,a+1,t+1}$ and $\Omega^i_{r,a,t}$ respectively:

$$L^i_{r,a,t}\mu^{j,i}_{r,a,t} = \exp\{v^j_{r,a,t} + \omega^i_{r,a,t} + \tilde{\tau}^{j,i}_{r,a,t}\},$$

where

$$\begin{aligned} v^j_{r,a,t} &= \frac{1}{\nu}s_{r,a,t}V^j_{r,a+1,t+1}, \\ \omega^i_{r,a,t} &= -\frac{1}{\nu}\Omega^i_{r,a,t} + \log(L^i_{r,a,t}), \\ \tilde{\tau}^{j,i}_{r,a,t} &= -\frac{1}{\nu}\tau^{j,i}_{r,a,t}. \end{aligned}$$

Recovering Period Utility

- Arranging destination and origin fixed effects backs out period utilities

$$
\begin{aligned}
&\frac{v^j_{r,a,t}}{s_{r,a,t}} + \omega^j_{r,a+1,t+1} - \log(L^j_{r,a+1,t+1}) \\
=&\frac{1}{\nu} u^j_{r,a,t} \\
=&\frac{1}{\nu} \left\{ \log\left(\frac{w^j_{r,a+1,t+1}}{(r^j_{t+1})^\gamma} \right) + \log(B^j_{r,a+1,t+1}) \right\}.
\end{aligned}
$$

Two-Step Estimation of $1/\nu$

Following Artuc and McLaren (2015)

1. Regress the number of migrants on the destination and origin fixed effects and the terms capturing migration costs

$$L^{i}_{r,a,t}\mu^{j,i}_{r,a,t} = \exp\left\{ v^{j}_{r,a,t} + \omega^{i}_{r,a,t} + \tilde{\tau}^{j\neq i}_{t} + \tilde{\tau}^{\{i,j\}}_{r,G(t)} + \tilde{\tau}^{\{i,j\}}_{a,G(t)} \right\} + \varepsilon^{j,i}_{r,a,t}.$$

   - $G(\cdot)$ classifies years to groups.

2. Regress the induced period utilities times the migration elasticity on wages and the terms capturing amenities

$$\begin{aligned} &\frac{\hat{v}^{j}_{r,a,t}}{s_{r,a,t}} + \hat{\omega}^{j}_{r,a+1,t+1} - \log(L^{j}_{r,a+1,t+1}) \\ =&\frac{1}{\nu}\log(w^{j}_{r,a+1,t+1}) + \tilde{B}^{j}_{r,a+1} + \tilde{B}^{j}_{r,t+1} + \varepsilon^{j}_{r,a,t}. \end{aligned}$$

   - I instrument $w^{j}_{r,a+1,t+1}$ by $w^{j}_{r,a+1,t}$.

Estimates of Migration Elasticity

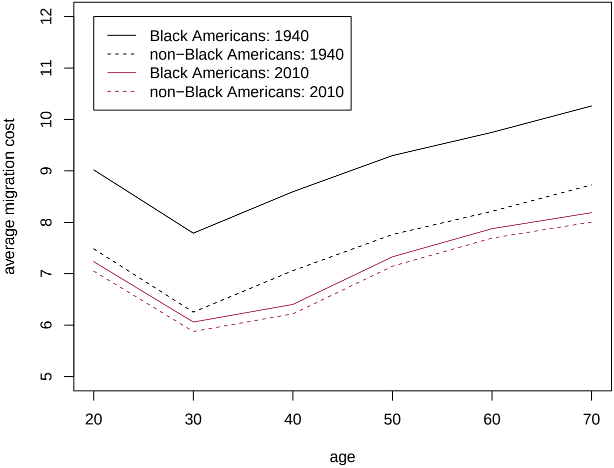
Dependent variable:	period utility $\times$ migration elasticity		
	(1)	(2)	(3)
log(real wage)	0.4976***	0.6129***	0.7676***
	(0.1323)	(0.1665)	(0.1952)
fixed effects:			
race-location	✓	✓	✓
age-location	✓	✓	✓
year-location	✓	✓	✓
age-race	✓	✓	✓
year-race	✓	✓	✓
age-race-location		✓	✓
year-race-location			✓
Observations	2,660	2,660	2,660

Robust standard errors clustered at locations. ***: 0.01.

Migration Elasticities in Literature

	location	value
Bryan and Morten	Indonesia	3.18
	US	2.69
Tombe and Zhu	China	1.50
Fajgelbaum, Morales, Suarez Serrato, and Zider	US	2.10
Caliendo, Opromolla Parro, and Sforza	EU	0.50
Suzuki	Japan	2.01
		(1.57$\sim$3.32)

Migration Costs by Age



back

Estimates of Elasticity of Substitution across Races

	$-1/\sigma_1$	implied σ_1
This paper	-0.111	9.0
Boustan (2009)	-0.120	8.3
	(-0.186$\sim$-0.090)	(5.38$\sim$11.11)

Elasticity of Substitution across Ages

- The nested CES production function implies

$$\frac{w^i_{a,t}}{w^i_{a',t}} = \frac{\left(\kappa^i_{a,t}\right)^{\frac{1}{\sigma_0}}\left(L^i_{a,t}\right)^{-\frac{1}{\sigma_0}}}{\left(\kappa^i_{a,t}\right)^{\frac{1}{\sigma_0}}\left(L^i_{a',t}\right)^{-\frac{1}{\sigma_0}}},$$

where

$$w^i_{a,t} = \left[\sum_{r'} \kappa^i_{r',a,t}\left(w^i_{r',a,t}\right)^{1-\sigma_1}\right]^{\frac{1}{1-\sigma_1}},$$

$$L^i_{a,t} = \left[\sum_{r'}\left(\kappa^i_{r',a,t}\right)^{\frac{1}{\sigma_1}}\left(L^i_{r',a,t}\right)^{\frac{\sigma_1-1}{\sigma_1}}\right]^{\frac{\sigma_1}{\sigma_1-1}}.$$

back

Estimation of Elasticity of Substitution across Ages

- Fix age bin a'.
- The main specification is, for any $a \neq a'$,

$$\log\left(\frac{w^i_{a,t}}{w^i_{a',t}}\right) = -\frac{1}{\sigma_0}\log\left(\frac{L^i_{a,t}}{L^i_{a',t}}\right) + f_a + f_t + f_{a,t} + \varepsilon^i_{a,t}.$$

- $\hat{L}^i_{a,t}$ aggregates the shift-share predicted populations for (r, a, t, n)

$$\hat{L}^i_{a,t} = \left[\sum_r (\kappa^i_{r,a,t})^{\frac{1}{\sigma_1}} (\hat{L}^i_{r,a,t})^{\frac{\sigma_1 - 1}{\sigma_1}}\right]^{\frac{\sigma_1}{\sigma_1 - 1}}.$$

- Construct an IV using $\hat{L}^i_{a,t}$.

Elasticity of Substitution across Ages

Dependent variable:	$\log(w^i_{a,t}/w^i_{a',t})$	
Model:	OLS	IV
$\log(L^i_{a,t}/L^i_{a',t})$	-0.2978***	-0.3401*
	(0.0672)	(0.1922)
fixed effects		
year-age	✓	✓
Weights	-	-
Observations	1,140	1,140
First-stage F-statistic		426.8

Block bootstrap standard errors are in parentheses. ***: 0.01, **: 0.05.

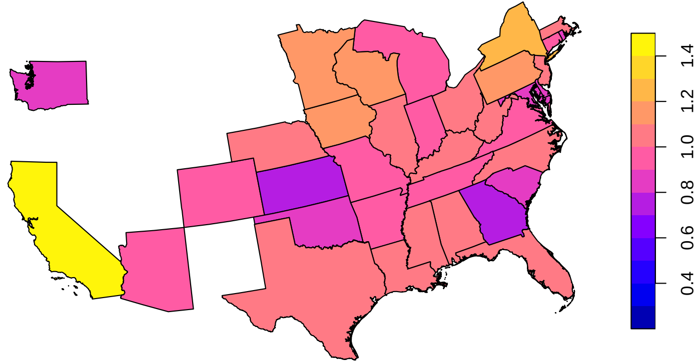
Estimates of Elasticity of Substitution across Ages

	$-1/\sigma_0$	implied σ_1
my estimate	-0.340	2.9
Card and Lemieux	-0.203	4.9
	(-0.233∼-0.165)	(4.3∼6.1)

- Ottaviano and Peri (2012) and Manacorda et. al. (2012) found estimates similar to Card and Lemieux (2001).
- My age bin is 10 years but the literature's age bin is 5 years.

Amenities in 1960

Others



back

Estimation: Rent Elasticity η

- Assume that the rent elasticity η is common in all locations.
- Taking logs of the rent equation:

$$\log r_t^i = \log \bar{r}^i + \eta \log \left(\gamma \sum_r \sum_c L_{r,c,t}^i w_{r,c,t}^i \right).$$

- Take time differences:

$$\Delta \log r^i = \eta \Delta \log(\text{income}^i).$$

- Then I can use states as a sample.
- For state i, the econometric specificaiton is

$$\Delta \log r^i = \eta \Delta \log(\text{income}^i) + \varepsilon_i.$$

  - The time differences are taken between 1970 and 2010.
- I instrument $\Delta \log(\text{income}^i)$ by the manufacturing shares and college graduates shares as of 1950.

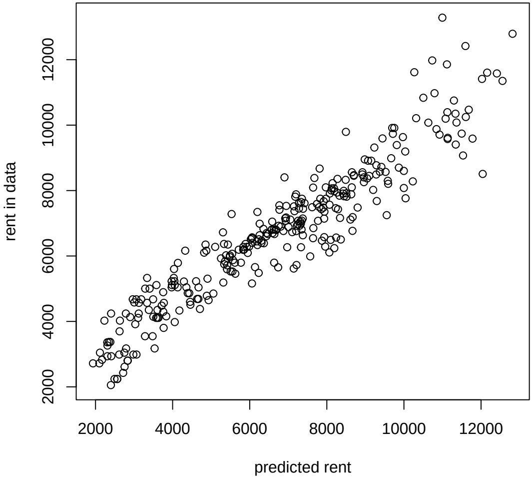
back

Estimates: Rent Elasticity η

Dependent variable:	$\Delta \log r^i$	
Model:	OLS	IV
$\Delta \log(\text{income}^i)$	0.3948***	0.4092***
	(0.0254)	(0.0264)
Weights	L^i_{1970}	L^i_{1970}
Observations	38	38
First-stage F-statistic		162.4

Robust standard errors are in parentheses. ***: 0.01.

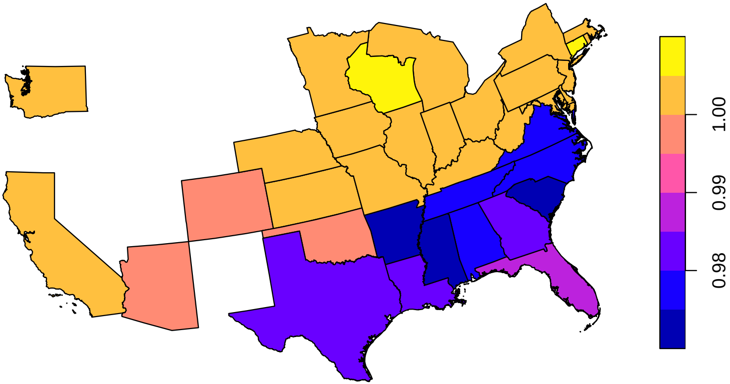
Goodness of Fit: Nation-wide Rent Elasticity



correlation: 0.944

The Welfare of Black Americans Born in the 1930s

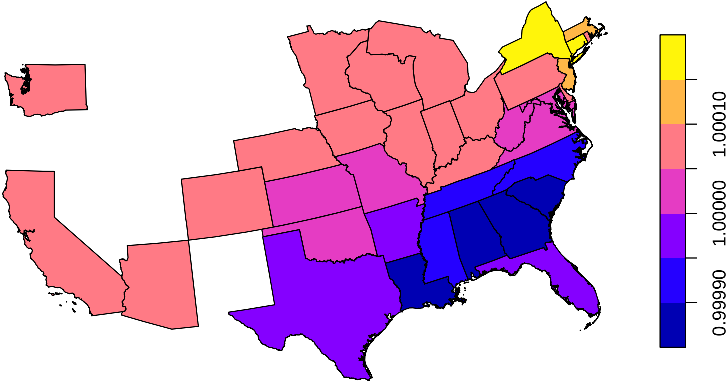
Black immobility relative to the baseline



back

The Welfare of non-Black Americans Born in the 1930s

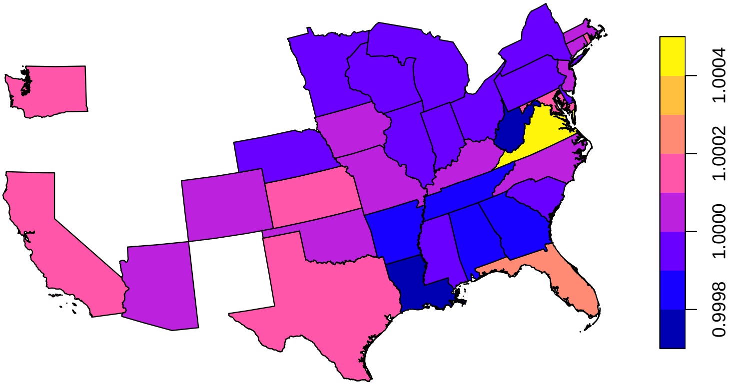
Black immobility relative to the baseline



back

The Welfare of Black Americans Born in the 1930s

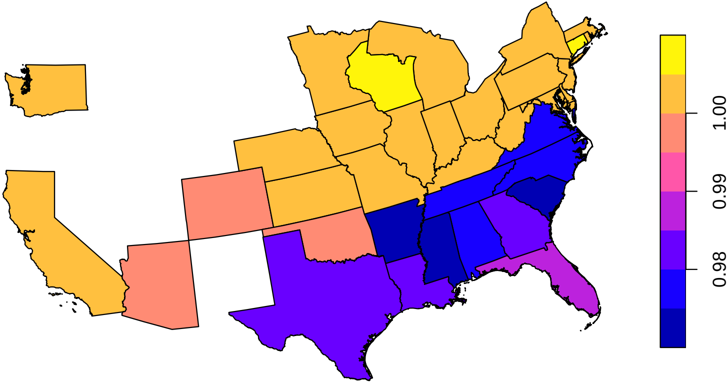
Non-Black immobility relative to the baseline



back

The Welfare of Non-Black Americans Born in the 1930s

Non-Black immobility relative to the baseline



back

Parameters in the Baseline Equilibrium

- I have parameter values from 1940 to 2010.
- From 2020 onward, I assume all parameters are as of 2010.
- But I use fertility such that the populations of Black and non-Black Americans smoothly converge from 2010 to the (final) steady state.
- So that the economy will converge to the steady state.

Value Function Iteration

1. Load the expected values of the final steady state $V^i_{r,a,\infty}$. Assume the economy converges to the steady state in period T: $V^i_{r,a,T} = V^i_{r,a,\infty}$.
2. Load the populations in the initial period $L^i_{r,a,0}$.
3. Guess the expected values from period 0 to $T-1$ $V^i_{r,a,t}$ for $t = 0, \cdots, T-1$.
4. Compute migration shares $\mu^{j,i}_{r,a,t}$ given the guessed expected values $V^i_{r,a,t}$.
5. Compute the populations $L^i_{r,a,t}$ forward given the migration shares $\mu^{j,i}_{r,a,t}$.
6. Compute wages $w^i_{r,a,t}$, rent r^i_t, and eventually period utility $u^i_{r,a,t}$ given the populations $L^i_{r,a,t}$.
7. Compute the expected values $V^i_{r,a,t}$ backward given the period utility $u^i_{r,a,t}$.

Welfare

Consumption Equivalent

- Two expected values $V^j_{r,0,t}$ (baseline) and $\tilde{V}^j_{r,0,t}$ (counterfactual).
- Define the compensating variation $\delta^j_{r,0,t}$ by

$$\tilde{V}_{r,0,t} = V^j_{r,0,t} + \sum_{a=0}^{\bar{a}} \left[\prod_{a'=-1}^{a-1} s_{r,a',t+a'} \log(\delta^j_{r,0,t}) \right].$$

  - $s_{r,-1,t-1} = 1$ for any r and t for notational convenience.
- Solving this,

$$\delta^j_{r,0,t} = \exp\left\{ \frac{\tilde{V}^j_{r,0,t} - V^j_{r,0,t}}{\sum_{a=0}^{\bar{a}} \prod_{a'=-1}^{a-1} s_{r,a',t+a'}} \right\}.$$

- Note that the welfare of the counterfactual is higher than that of the baseline if $\delta^j_{r,0,t} > 1$.

back