Migration and Segregation

Matz Dahlberg

Institute for Housing and Urban Research and Department of Economics
Uppsala University

Winter School in Canazei, January, 2018

Changes in the wake of immigration

- Immigration ⇒ ethnically more heterogeneous population
- ⇒ Discussion on potential effects following this change:
  - effects on welfare attitudes/preferences for redistribution
    (Alesina and Glaeser, Book 2004; Dahlberg et al, JPE 2012)
  - ethnic segregation (residential, workplace, schools)
    (Saiz and Wachter; Sá; Fairlie and Betts)
  - effects on political landscape and political preferences
    (Dustmann et al, CrEAM WP 2016)
  - labor market effects (employment, wages, ...)
    (Borjas; Card – many papers!)
  - effects on income inequality
    (Card, NBER WP 2009)

Immigration and residential segregation

- Increased immigration does not need to imply increased (residential) segregation
- But it can be so, if
  1. immigrants face discrimination in the housing market
  2. new immigrants are drawn to previous immigrants
  3. new immigrants induce the native population to leave or avoid moving into increasingly ethnically diverse neighborhoods

Question of interest

Does immigration into a n'hood affect natives' migration behavior?

$$outflow_{i,t+s} = \alpha^{out} + \beta^{out} im_{i,t} + \epsilon^{out}_{i,t+s} \quad (1)$$

$$inflow_{i,t+s} = \alpha^{in} + \beta^{in} im_{i,t} + \epsilon^{in}_{i,t+s} \quad , \quad (2)$$

1. Flight
2. Avoidance

Outline

1. Why would immigration affect native migration? (Mechanisms)
2. How to identify a causal effect?
3. Empirical evidence
4. Tipping points

Why would immigration affect native migration?

Why would immigration affect native migration?

- Racial/ethnic preferences (preferences for neighborhood homogeneity)
- School quality
- Crime or social unrest
- Socio-economic characteristics
- Economic/financial (via house prices)

Farley et al, 1994, Am J of Sociology

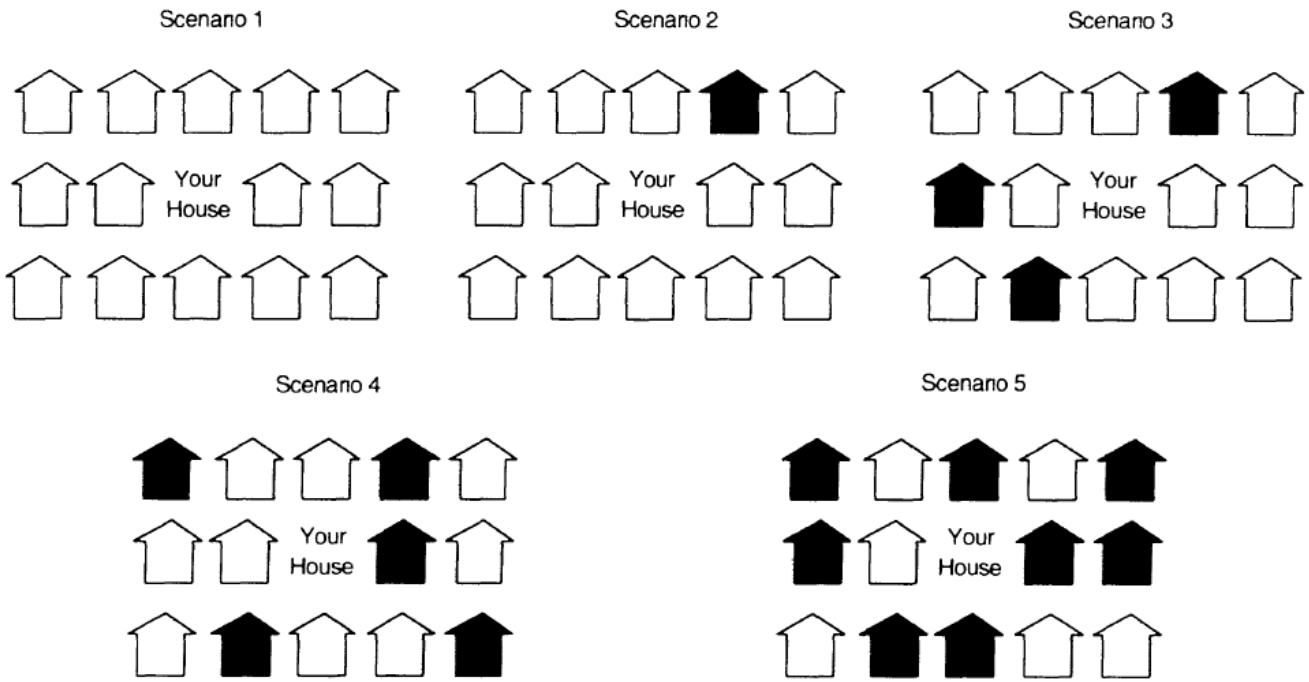


FIG. 2.—Neighborhood diagrams used for white respondents, DAS 1976 and 1992.

Farley et al, 1994, Am J of Sociology

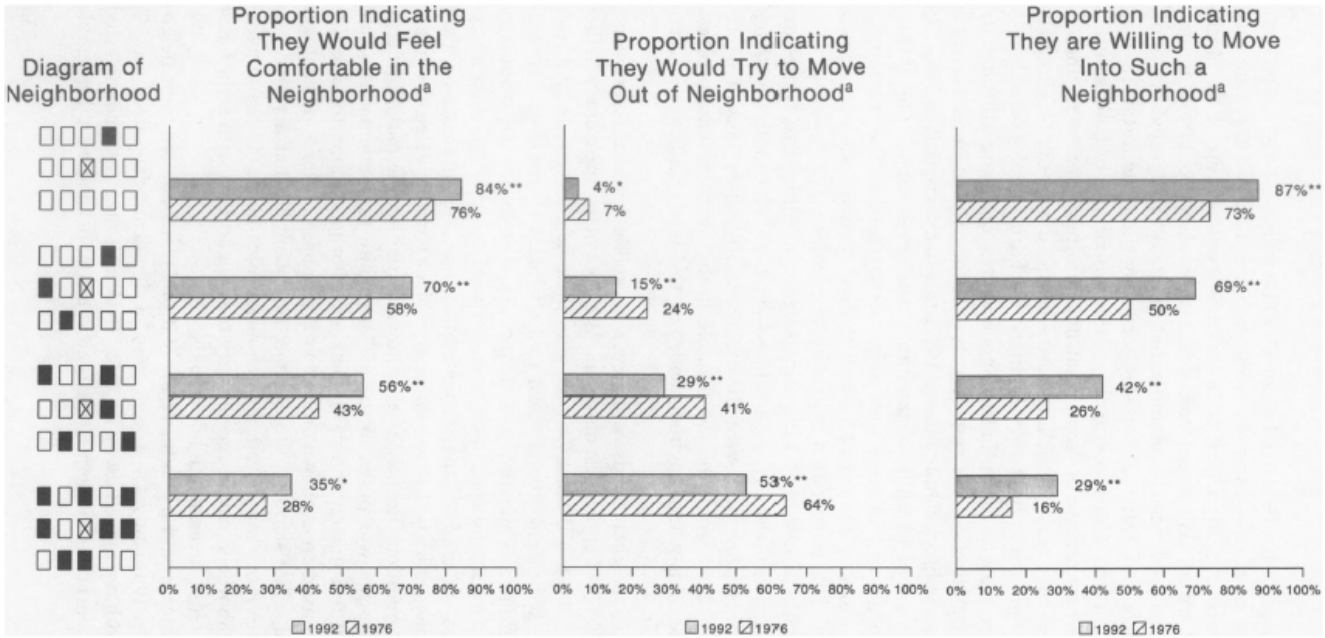


FIG. 3.—Attractiveness of neighborhoods of varying racial compositions for white respondents, DAS 1976 and 1992 (significance shown for change from 1976 to 1992; $*P < .05$; $**P < .01$; the denominator for % is the total white population).

Identification: Shift-share instrument

$$\tilde{im}_{i,t} = \sum_c \tilde{im}_{c,i,t} = \sum_c \left(\phi_{c,i,t^0} \times im_{c,SWE,t} \right), \tag{3}$$

where

$$\phi_{c,i,t^0} = \frac{im_{c,i,t^0}}{im_{c,SWE,t^0}} \tag{4}$$

- How to define c?
- How to define t^0?

UK Evidence: Sá (Economic Journal, 2014)

- Years: 2003-2010 (yearly)
- Base year for shift-share instrument: 2001
- Geographic area: 170 local authorities in England and Wales
- Data: UK Labour Force Survey (rotating panel household survey)
- Nine foreign regions of origin:
  1. India
  2. EU
  3. Americas and Caribbean
  4. Africa
  5. Other Middle East and Indian sub-continent
  6. Asia
  7. Antarctica and Oceania
  8. Republic of Ireland
  9. Other countries

Sá: Econometric specification

First stage:

$$\frac{\triangle FB_{it}}{Pop_{i,t-1}} = \alpha \frac{\sum_c \lambda_{ci,t_0} \triangle FB_{ct}}{Pop_{i,t-1}} + \phi_t + \rho_i + \eta_j + \varepsilon_{it} \tag{5}$$

- $\triangle FB_{it}$: annual change in stock of immigrants
- $\sum_c \lambda_{ci,t_0} \triangle FB_{ct}$: Shift-share instrument (predicted $\triangle FB$ in n'hood i year t)
- η_j: Education FE (four groups of education level)

Outcome variables:

1. Native population change
2. Native out-migration

Sá: Results

Table 7

Immigrant Inflows and Native Population

	Independent variable: $\Delta FB_{it}/Pop_{it-1}$	
	OLS	IV
Panel (*a*): by LA and year		
$\Delta N_{it}/Pop_{it-1}$	−0.266**	−0.868***
	(0.118)	(0.097)
R^2	0.858	0.848
Native out-migration rate	0.111***	0.123***
	(0.024)	(0.019)
R^2	0.637	0.642
Native in-migration rate	0.032**	0.053**
	(0.014)	(0.026)
R^2	0.627	0.680
Native net out-migration rate	0.080***	0.071**
	(0.025)	(0.030)
R^2	0.388	0.449
Observations	1,360	1,360
Panel (*b*): by LA, year and education group		
$\Delta N_{it}/Pop_{it-1}$	0.221*	−0.968***
	(0.113)	(0.107)
R^2	0.556	0.741
Native out-migration rate	0.062***	0.078***
	(0.022)	(0.016)
R^2	0.279	0.440
Native in-migration rate	0.015	0.029**
	(0.010)	(0.013)
R^2	0.262	0.312
Native net out-migration rate	0.047**	0.048***
	(0.019)	(0.017)
R^2	0.142	0.221
Observations	5,440	5,440

US Evidence: Saiz and Wachter (AEJ: Policy, 2011)

- Years: 1980, 1990, 2000 (3 years)
- Data: Census data (decennial)
- Geographic area: 34,835 Census tracts (in 122 MSA-year groups)
- All immigration to the US
- Suggest a variant of the shift-share instrument

Saiz and Wachter: Econometric specification

Instrument: Geographic diffusion model (gravity pull)

$$Pull_{i,t} = \sum_{j \neq i, j \in M} \frac{(IMSHARE_{j,t-10}) \times Area_j}{(d_{ij})^{\beta}} \tag{6}$$

interacted with immigration shock to MSA and initial immigrant share

- $Area_j$: Area (square miles) of census tract j
- (d_{ij}): Euclidean distance between tracts i and j

Outcome variable:

1. Native population (decennial) change

Saiz and Wachter: Results

TABLE 3— IMMIGRANT INFLOWS AND NATIVE MOBILITY WITHIN MSA

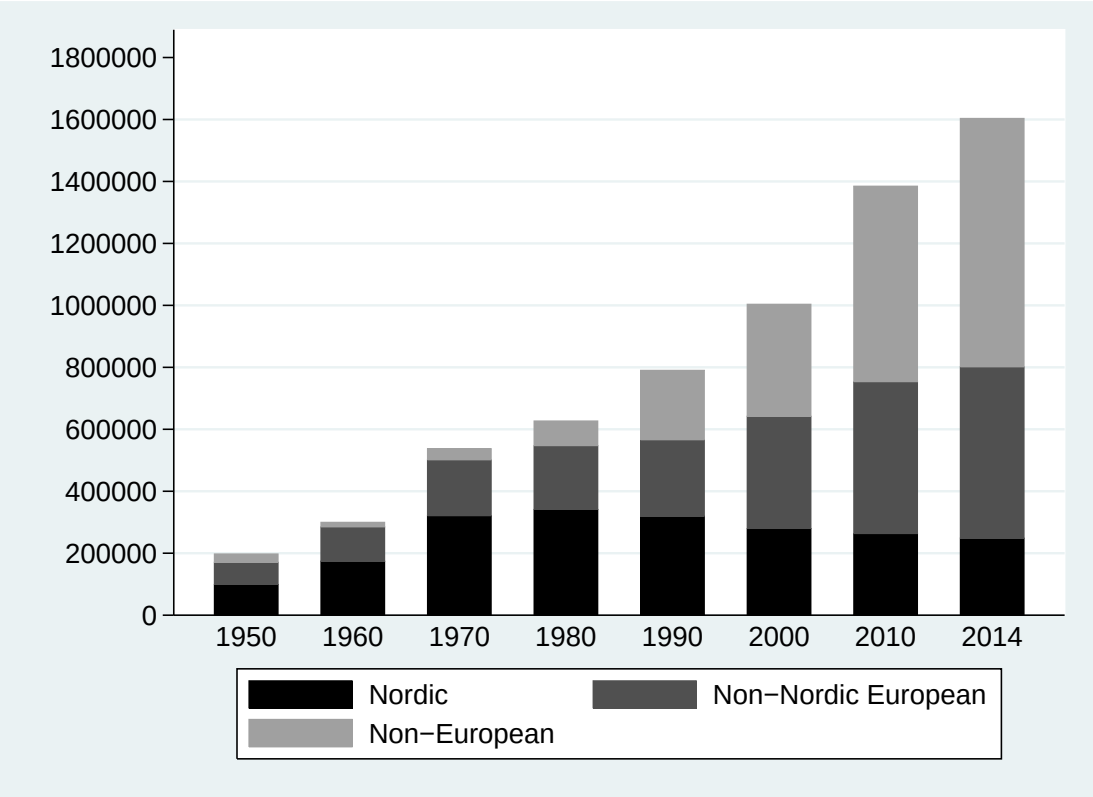
	(ΔNative population)/population at T-10				(ΔNon-Hispanic-white population)/population at T-10			
	All-OLS	Excludes new developments	Median regression	IV	All-OLS	Excludes new developments	Median regression	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(Δforeign born)/ population at T-10	1.243 (0.107)***	−0.045 (0.012)***	−0.096 (0.003)***	−0.134 (0.043)***	0.767 (0.105)***	−0.329 (0.011)***	−0.366 (0.003)***	−0.678 (0.045)***
Other variables in Table 5, column 5	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls for immigrant "gravity pull"	No	No	No	Yes	No	No	No	Yes
Method	OLS	OLS	ML	IV	OLS	OLS	ML	IV
Instruments	None	None	None	Gravity pull × MSA immigration, gravity pull × share foreign born at T-10	None	None	None	Gravity pull × MSA immigration, gravity pull × share foreign born at T-10
Observations	36,847	35,105	36,847	36,847	36,847	35,105	36,847	36,847

Swedish Evidence: Andersson, Berg, Dahlberg

Migrating Natives and Foreign Immigration: Is there a Preference for Ethnic Residential Homogeneity?

Immigration to Sweden: Number of foreign-born in Sweden by origin region, 1950–2014

- Drastic change over the last decades
- In 1950, <3% foreign-born, majority born in Nordic and other European countries
- Today >16% foreign-born, majority born outside Europe
- Among 10 most common countries of origin: Iraq, Iran, former Yugoslavia, Turkey, Somalia, Syria



Questions of interest

1. Does immigration into a n'hood affect natives' migration behavior?

$$outflow_{i,t+s} = \alpha^{out} + \beta^{out} im_{i,t} + \epsilon^{out}_{i,t+s} \tag{7}$$

$$inflow_{i,t+s} = \alpha^{in} + \beta^{in} im_{i,t} + \epsilon^{in}_{i,t+s} \quad , \tag{8}$$

2. Is the mechanism ethnically based?
   Examine if natives prefer *ethnically* homogeneous neighborhoods

Value-added

1. Data enables us to examine the support for an ethnically based mechanism for a native migration response
2. Data and Swedish institutional details enable us to improve on the instrument typically used in the “effects-of-immigration” literature
3. Data enables us to identify households with high possibility to move

Maintained hypothesis

- Preferences of the native population à la Sá (2014):

$$U_{n,i} = V_{n,i} + f(h,x) - \delta I \tag{9}$$

- Socio-economic preferences? (E.g., Boustan, 2010; Saiz and Wachter, 2011)

- Our maintained hypothesis:

$$\delta^{Ethnicity}_{Swedish\_Parents} \geq \delta^{Ethnicity}_{Western\_Parents} > \delta^{Ethnicity}_{Non-Western\_Parents} \tag{10}$$

Other potential mechanisms

1. Mechanical (1 in, 1 out)
2. School quality
3. Economic/financial (via house prices)

The instrument

$$\tilde{im}_{i,t} = \sum_c \tilde{im}_{c,i,t} = \sum_c \left(\phi_{c,i,t^0} \times im_{c,SWE,t} \right), \tag{11}$$

where

$$\phi_{c,i,t^0} = \frac{im_{c,i,t^0}}{im_{c,SWE,t^0}} \tag{12}$$

1. How define c and t^0?

The instrument

$$\tilde{im}_{i,t} = \sum_c \tilde{im}_{c,i,t} = \sum_c \left(\phi_{c,i,t^0} \times im_{c,SWE,t} \right), \tag{11}$$

where

$$\phi_{c,i,t^0} = \frac{im_{c,i,t^0}}{im_{c,SWE,t^0}} \tag{12}$$

1. How define c and t^0?
2. **Push-related migration**
   (Know reason for migration to Sweden; work, education, tied family member, **refugee**)

The instrument

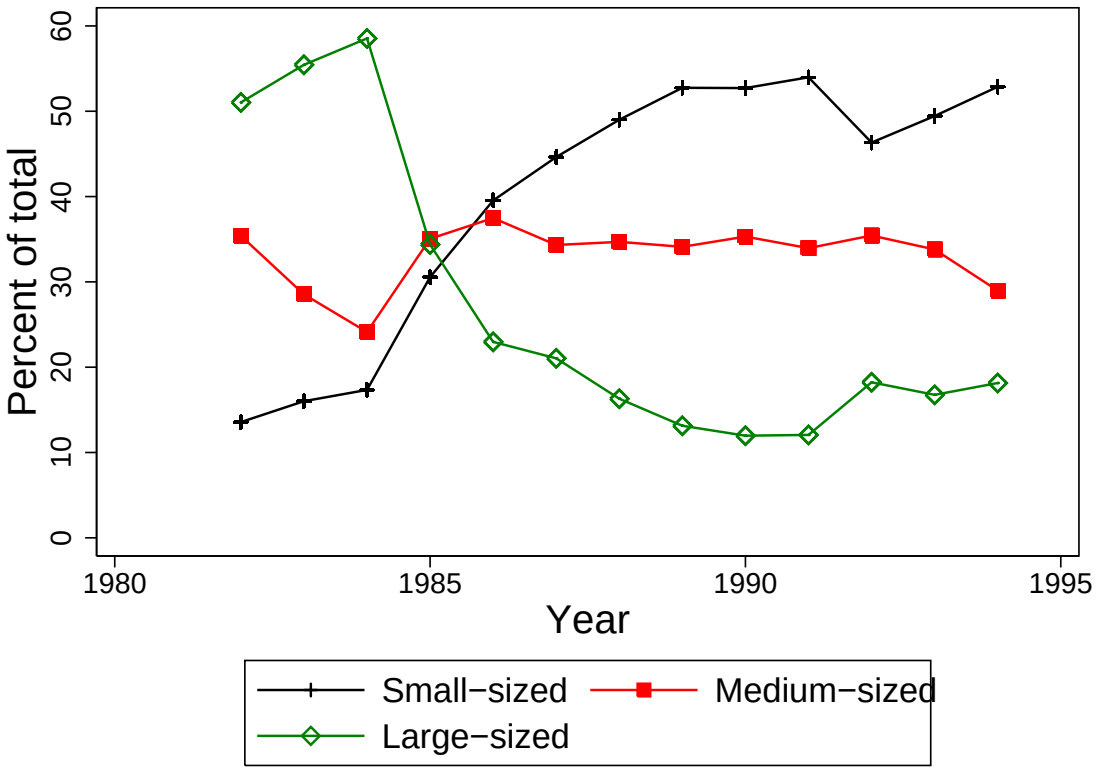
$$\tilde{im}_{i,t} = \sum_{c} \tilde{im}_{c,i,t} = \sum_{c} \left(\phi_{c,i,t^0} \times im_{c,SWE,t} \right), \tag{11}$$

where

$$\phi_{c,i,t^0} = \frac{im_{c,i,t^0}}{im_{c,SWE,t^0}} \tag{12}$$

1. How define c and t^0?
2. **Push-related migration**
   (Know reason for migration to Sweden; work, education, tied family member, **refugee**)
3. **A refugee placement policy in effect 1990–1993**
   $\Rightarrow$ makes the selection problem of initial location of refugee's countrymen less problematic

Location of refugees



Groups with high and low possibility to react on increased immigration

- Divide sample into:
  - Owners (high possibility to react/move)
  - Renters (low possibility to react/move)

- Refugee immigration increases competition for public rental apartments:
  - Municipality-provided rental housing make up majority of rental market
  - Access to these apartments requires queuing (in many municipalities for years, or even decades)
  - True also for those already living in public rentals
  - Municipalities responsible for finding apartments for refugees

Data: The GeoSweden database

- Yearly register data (1990–2014) covering all individuals permanently living in Sweden
- Geo-coded (100×100 meters) (where they live, work; know which school they go to)
- Information on date and from which country an individual immigrates to Sweden and the reason for immigration (*Grund för bosättning*; 1997–2010)
- Information on migration patterns within Sweden (from one year to another – from the geo-coded information we also know exactly from and to a household has moved)
- Rich on registered-based background characteristics for the individuals:
  - year and country of birth
  - marital status
  - number of children in the household
  - individuals' level and type of education
  - pre-tax income from different sources
  - disposable income
  - employment
  - school grades after finishing compulsory school (grade 9) and finishing high school
  - information on parents

Descriptives

Variable	Obs	Mean	Std. Dev.	Min	Max
Key variables:					
Outflow	114 477	85.2	118	0	2352
Inflow	114 470	85.2	121	0	2716
Immigration (main)	114 478	0.82	4.7	0	313
Predicted immigration (instrument)	113 503	0.81	3.6	0	251
Control variables:					
Population	114 478	1019	1236	1	20 285
Students	114 478	53.1	107.5	0	2642
Disposable income	114 478	155 838	538	-107 050	5 688 067
Social assistance	114 478	8700	22 500	0	108 200
Other non-OECD immigration	114 478	2.4	9.4	0	590
Other:					
Public rentals	113 681	143	447	0	9647

8,723 neighborhoods over the time period 1997-2010

Refugees' country of origin

Iraq	40537	43.31%	43.31%
Somalia	11597	12.39%	55.70%
Yugoslavia	8345	8.92%	64.61%
Bosnia	6727	7.19%	71.80%
Iran	5105	5.45%	77.26%
Afghanistan	4347	4.64%	81.90%
Syria	3954	4.22%	86.12%
Russia	2676	2.86%	88.98%
Lebanon	2563	2.74%	91.72%
Thailand	1225	1.31%	93.03%
Ethiopia	1142	1.22%	94.25%
Croatia	887	0.95%	95.20%
Colombia	736	0.79%	95.98%
India	683	0.73%	96.71%
Peru	520	0.56%	97.27%
Bangladesh	469	0.50%	97.77%
Pakistan	468	0.50%	98.27%
China	269	0.29%	98.56%
Uganda	187	0.20%	98.76%
Romania	165	0.18%	98.93%
Bolivia	164	0.18%	99.11%
Vietnam	160	0.17%	99.28%
Algeria	125	0.13%	99.41%
Poland	108	0.12%	99.53%
Moroco	86	0.09%	99.62%
Tunisia	78	0.08%	99.70%
Latvia and Lithuania	71	0.08%	99.78%
Bulgaria	49	0.05%	99.83%
Estonia	38	0.04%	99.87%
Phillipines	36	0.04%	99.91%
Gambia	31	0.03%	99.94%
Argentina	29	0.03%	99.98%
Slovenia	12	0.01%	99.99%
Brazil	11	0.01%	100.00%

Econometric model

First stage (**X** incl. non-refugee immigration and socio-economic charac.):

$$im_{i,t} = \gamma \tilde{im}_{i,t} + \sum_{p=1}^{3} \phi^p pop^p_{i,t} + \Gamma \mathbf{X} + \mu_i + \tau_t + \epsilon_{i,t} \tag{13}$$

Prediction $\widehat{im}_{ij,t}$ then used in equations of interest:

$$outflow_{i,t+1} = \beta^{out} \widehat{im}_{i,t} + \sum_{p=1}^{3} \delta^p pop^p_{i,t} + \Pi \mathbf{X} + \mu_i + \tau_t + \varepsilon^{out}_{i,t+1} \tag{14}$$

and

$$inflow_{i,t+s} = \beta^{in} \widehat{im}_{i,t} + \sum_{p=1}^{3} \delta^p pop^p_{i,t} + \Pi \mathbf{X} + \mu_i + \tau_t + \varepsilon^{in}_{i,t+s}, \tag{15}$$

First stage estimates

VARIABLES	(1) No controls	(2) Baseline specification
im	0.674***	0.607***
	(0.0808)	(0.0859)
Observations	113,503	104,772
Number of sams	8,731	8,731

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered on sams-level. Covariates: linear, quadratic and cubic controls for population size in $t-1$, all non-refugee immigration from the refugees' source countries in year $t-1$, time and neighborhood fixed effects, time-varying socio-economic characteristics of neighborhood (measured in $t-1$); average disposable income, the number of students, and the per capita cost of social assistance.

First stage: Robustness

VARIABLES	(1) Exclude least populated n'hoods	(2) Exclude most populated n'hoods	(3) Exclude Gothenburg	(4) Exclude n'hoods with most immigration
im	0.622***	0.251***	0.607***	0.118***
	(0.0882)	(0.0292)	(0.0895)	(0.0109)
Observations	78,997	94,156	95,676	103,312
Number of sams	6,770	7,948	7,973	8,728

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered on sams-level. (1) excludes the 10 percent least populated neighborhoods (less than 322 people). (2) excludes the 10 percent most populated (more than 2043 people). (3) excludes Gothenburg. (4) excludes neighborhoods with the ten percent highest number of immigrants, given positive immigration (more than 14 immigrants).

Flight and avoidance: Average effects

	(1) All Natives	(2) Home Owners
OUTFLOW		
im	0.0660	0.352**
	(0.158)	(0.156)
INFLOW		
im	-0.0913	0.169
	(0.185)	(0.137)
Observations	104,771	104,771
Number of sams	8,731	8,731
Mean of Dep. Variable	85	39

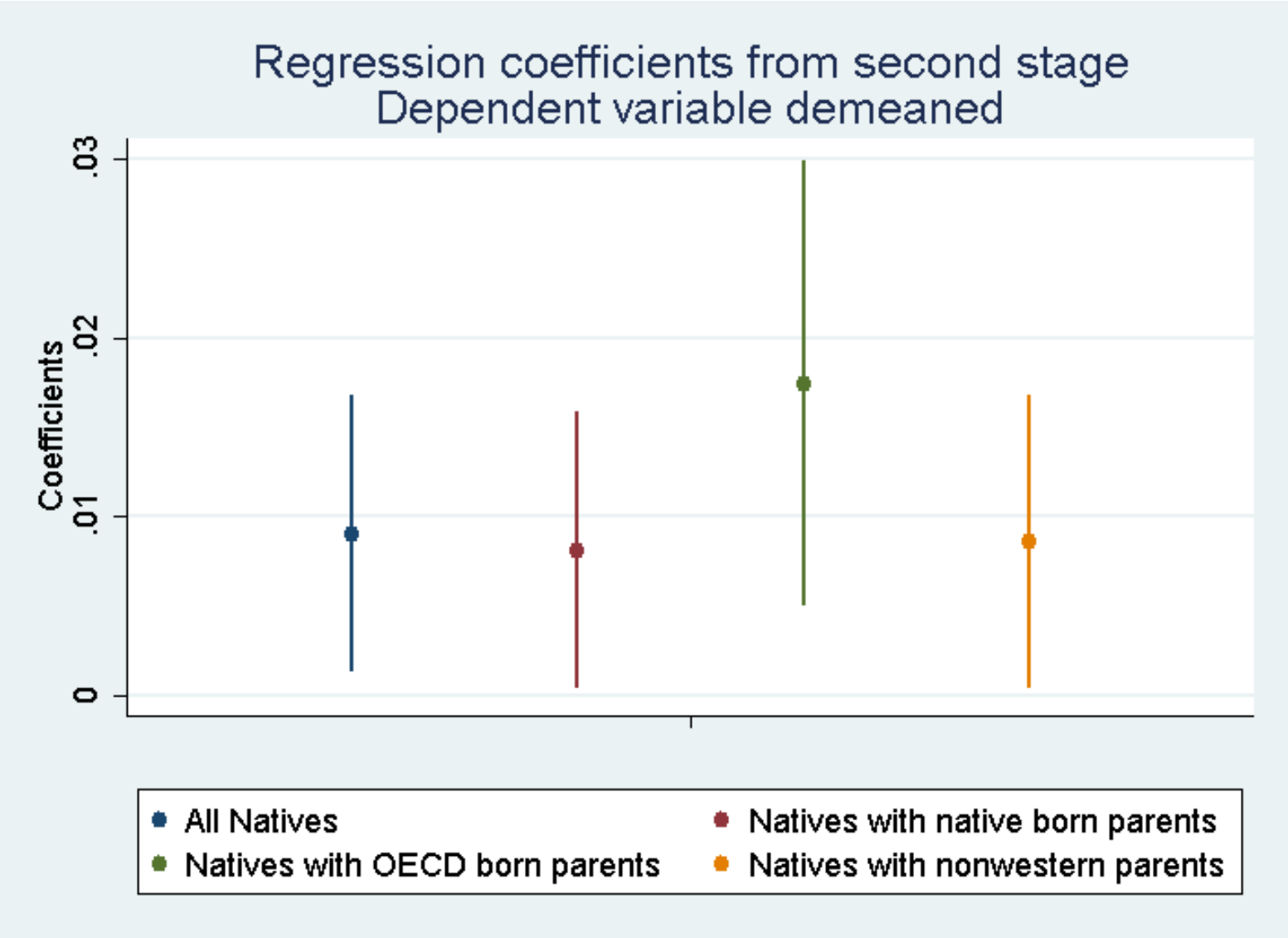
*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered on sams-level. Column (1) looks at outflow and inflow of all natives and column (2) looks at individuals owning a house or living in a condominium.

Flight and avoidance: By parents' ethnic background

VARIABLES	(1) Natives	(2) Natives with Native parents	(3) Natives with ≥ 1 OECD parent	(4) Natives with ≥ 1 non-Western parents
OUTFLOW				
im	0.357**	0.255**	0.0654***	0.0362**
	(0.156)	(0.124)	(0.0238)	(0.0177)
INFLOW				
im	0.169	0.117	0.0185	0.0332*
	(0.137)	(0.111)	(0.0190)	(0.0186)
Observations	104,248	104,248	104,248	104,248
Number of sams	8,710	8,710	8,710	8,710
Mean Dep.Var	39	31	3.8	4.2

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered on sams-level. Column 1 includes all native home owners, column 2 is restricted to native home owners with Swedish-born parents, column 3 is restricted to native home owners with at least one parent born in another OECD country, and column 4 is restricted to remaining native home owners.

Effects, normalized by mean of dependent variable



Socio-economic effects: Share highly educated

$$\frac{Educated_{i,t+1}}{Pop_{i,t+1}} = \beta^{educ}\widehat{im}_{i,t} + \sum_{p=1}^{3}\delta^{p}pop^{p}_{i,t} + \Pi\mathbf{X} + \mu_i + \tau_t + \varepsilon^{educ}_{i,t+1} \quad (16)$$

VARIABLES	(1) Owners	(2) Natives	(3) Natives with Native parents	(4) Natives with ≥ 1 OECD parent	(5) Natives with ≥ 1 non-Western parents
im	-0.000416***	-0.000356***	-0.000312***	-2.66e-05***	-1.77e-05***
	(9.88e-05)	(8.65e-05)	(7.60e-05)	(9.30e-06)	(5.89e-06)
Observations	95,551	95,551	95,551	95,551	95,551
Number of sams	8,709	8,709	8,709	8,709	8,709
Mean Dep.Var	0.12	0.11	0.09	0.008	0.01

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered on sams-level. The outcome variable is the neighborhood share of home owners in various groups that has at least some university education. Column 1 includes all home owners, column2 includes all native home owners, column 3 is restricted to native home owners with Swedish-born parents, column 4 is restricted to native home owners with at least one parent born in another OECD country, and column 5 is restricted to remaining native home owners.

Socio-economic effects: Mean disposable income

$$mean(Income_{i,t+1}) = \beta^{income}\widehat{im}_{i,t} + \sum_{p=1}^{3}\delta^{p}pop^{p}_{i,t} + \Pi\mathbf{X} + \mu_i + \tau_t + \varepsilon^{income}_{i,t+1} \quad (17)$$

VARIABLES	(1) Owners	(2) Natives	(3) Natives with Native parents	(4) Natives with ≥ 1 OECD parent	(5) Natives with ≥ 1 non-Western parents
im	-3.104**	-3.245**	-2.296*	-2.497	-0.657
	(1.380)	(1.451)	(1.362)	(1.872)	(0.891)
Observations	90,825	90,748	90,555	87,025	87,946
Number of sams	8,358	8,350	8,331	8,159	8,179
Mean Dep.Var	1687	1710	1784	1687	1313

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered on sams-level. The outcome variable is the neighborhood average disposable income in various groups. Column 1 includes all home owners, column2 includes all native home owners, column 3 is restricted to native home owners with Swedish-born parents, column 4 is restricted to native home owners with at least one parent born in another OECD country, and column 5 is restricted to remaining native home owners.

Do natives with an Iraqi background react on newly arrived Iraqis?

VARIABLES	(1) First Stage	(2) Outflow (SS) Iraqi owners	(3) Outflow (SS) Native Native, owners	(4) Outflow (SS) Iraqi renters
Iraq		0.0143**	0.397**	0.0596***
		(0.00594)	(0.165)	(0.0159)
$\hat{Iraq}$	0.495***			
	(0.0575)			
Observations	104,772	104,250	104,250	104,250
Number of sams	8,731	8,710	8,710	8,710
Mean Dep.Var		0.05	31	0.13

Tipping points: Theoretical background

If whites' willingness to pay to live in a neighborhood is a decreasing function of the share of non-white neighbors, ethnically mixed neighborhoods can be dynamically unstable

- Schelling (1971, J Math Sociology): two-sided tipping model where tipping point unstable mixed equilibrium => model predicts complete segregation
- Card et al (2008, QJE): one-sided tipping model: whites are heterogeneous with respect to their intrinsic preference for a given neighborhood => mixed neighborhoods can be dynamically stable; the tipping point is the highest minority share at which a mixed neighborhood can be dynamically stable

Estimating tipping point (Card et al, QJE, 2008; Swedish example)

We estimate the following regression for candidate values of the tipping point $m^*_{1987} = 1, 2, ..., 50$:

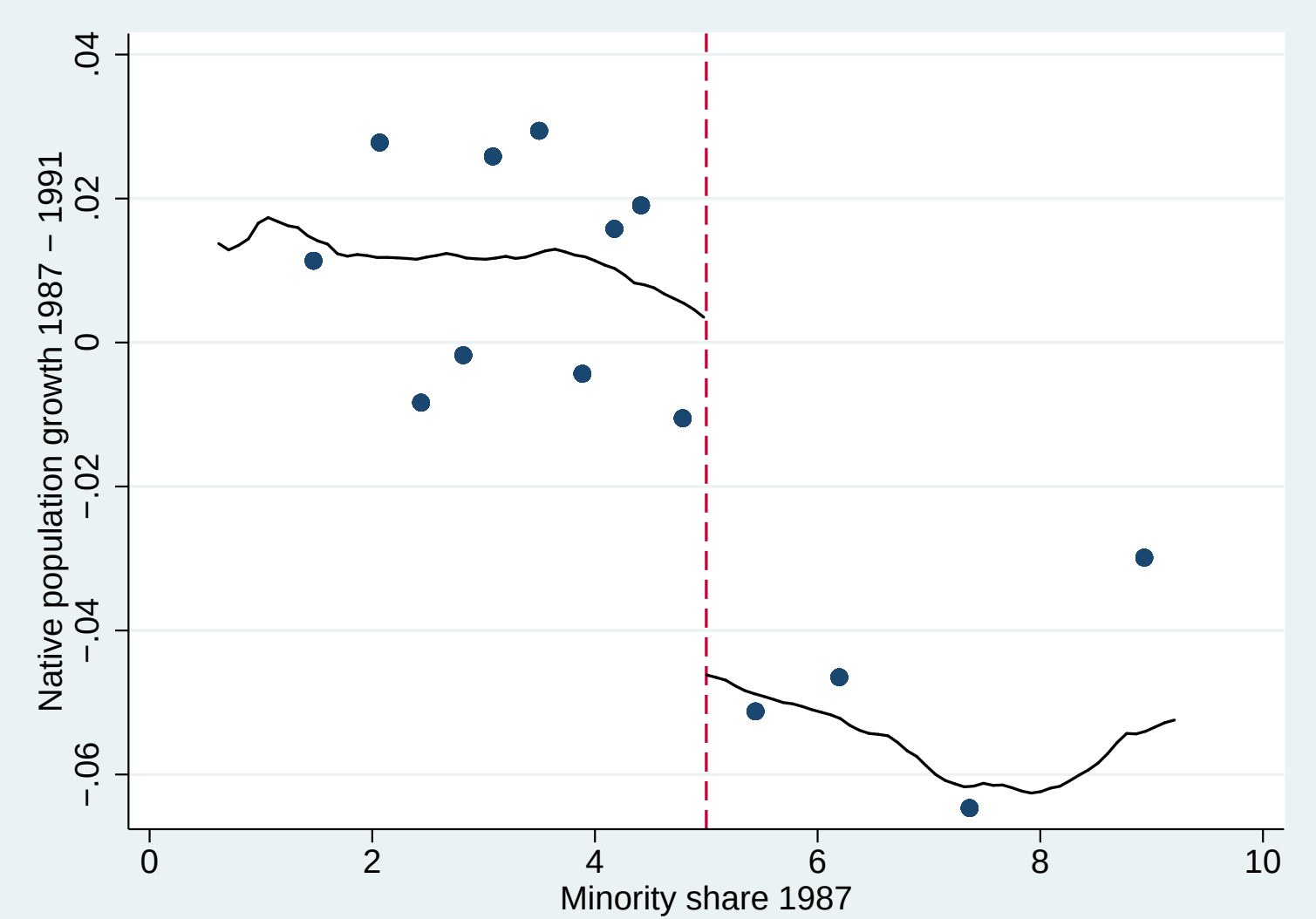
$$y_{i,1991} = \beta_0 + \beta_1(m_{i,1987} > m^*_{1987}) + g(m_{i,1985}) + \epsilon_{i,1991} \quad (18)$$

- i: neighborhood
- y_i: outcomes, e.g. majority population growth 1987-1991
- $m_{i,1987}$: minority share in the neighborhood in 1987
- $g(m_{i,1985})$: polynomial in minority share in 1985, to make neighborhoods more comparable
- $\epsilon_{i,1991}$: random shock

Tests for structural breaks

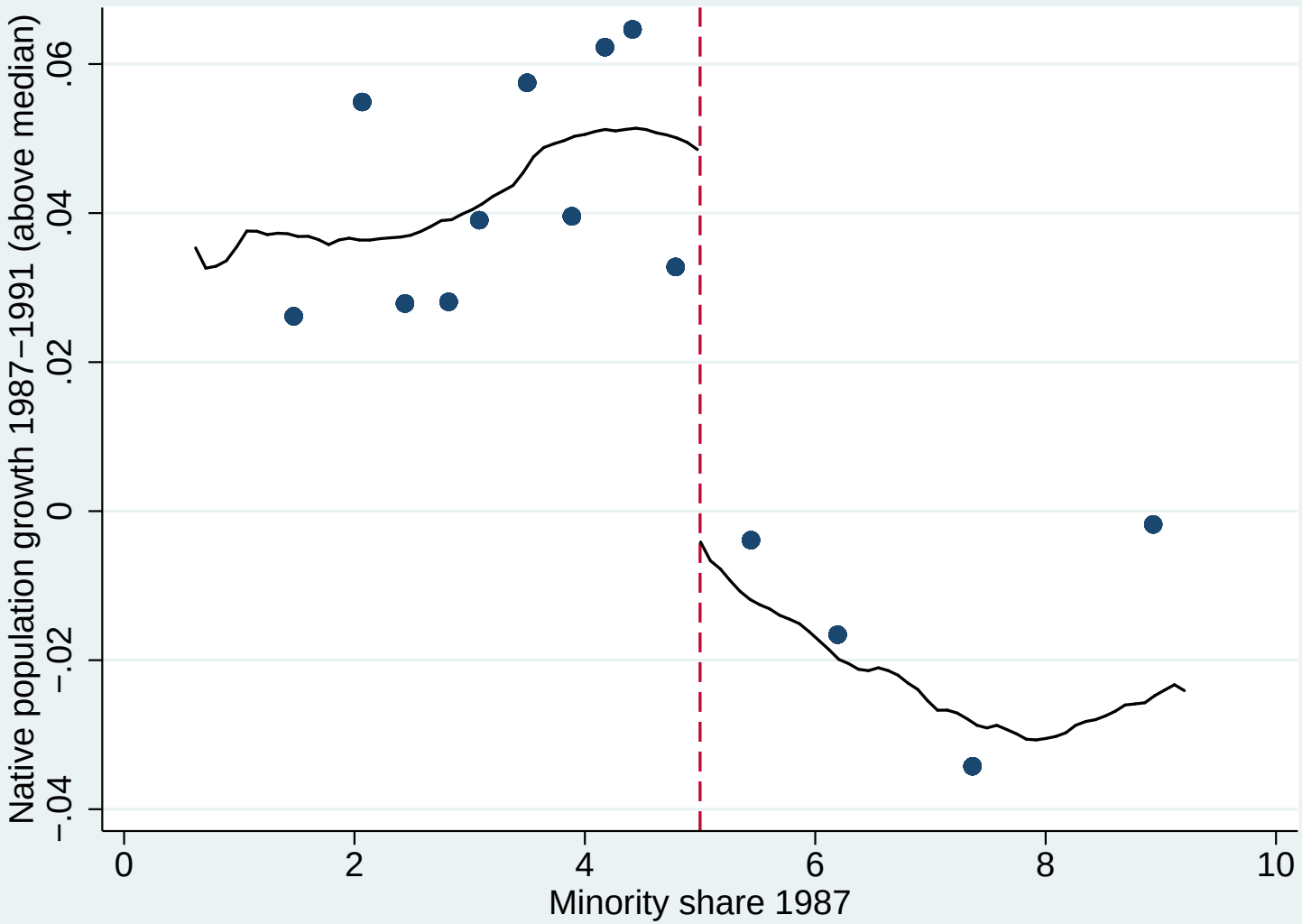
Majority population growth, 1987-1991

Figure: Majority population growth, 1987-1991

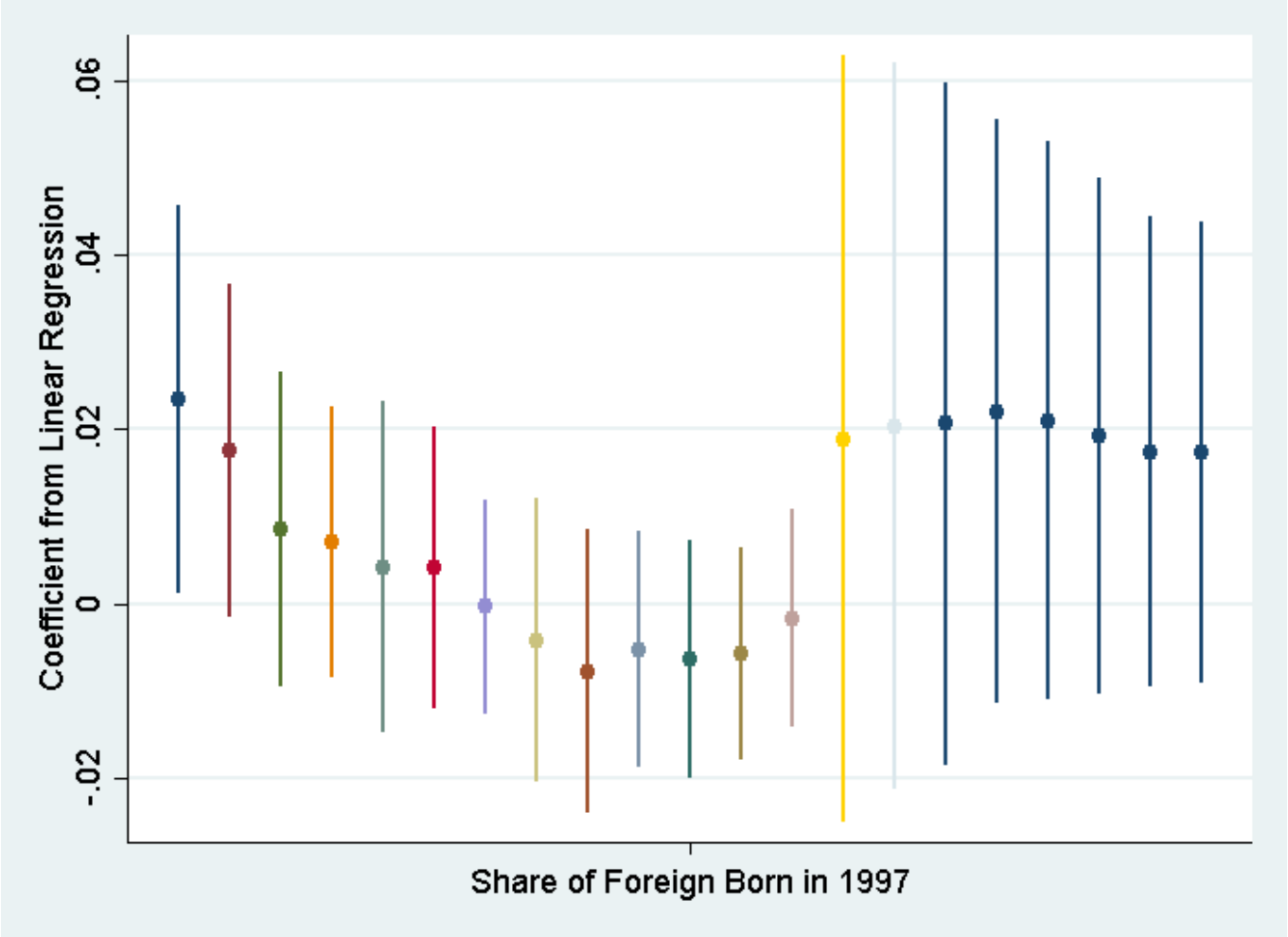


Tipping and income segregation

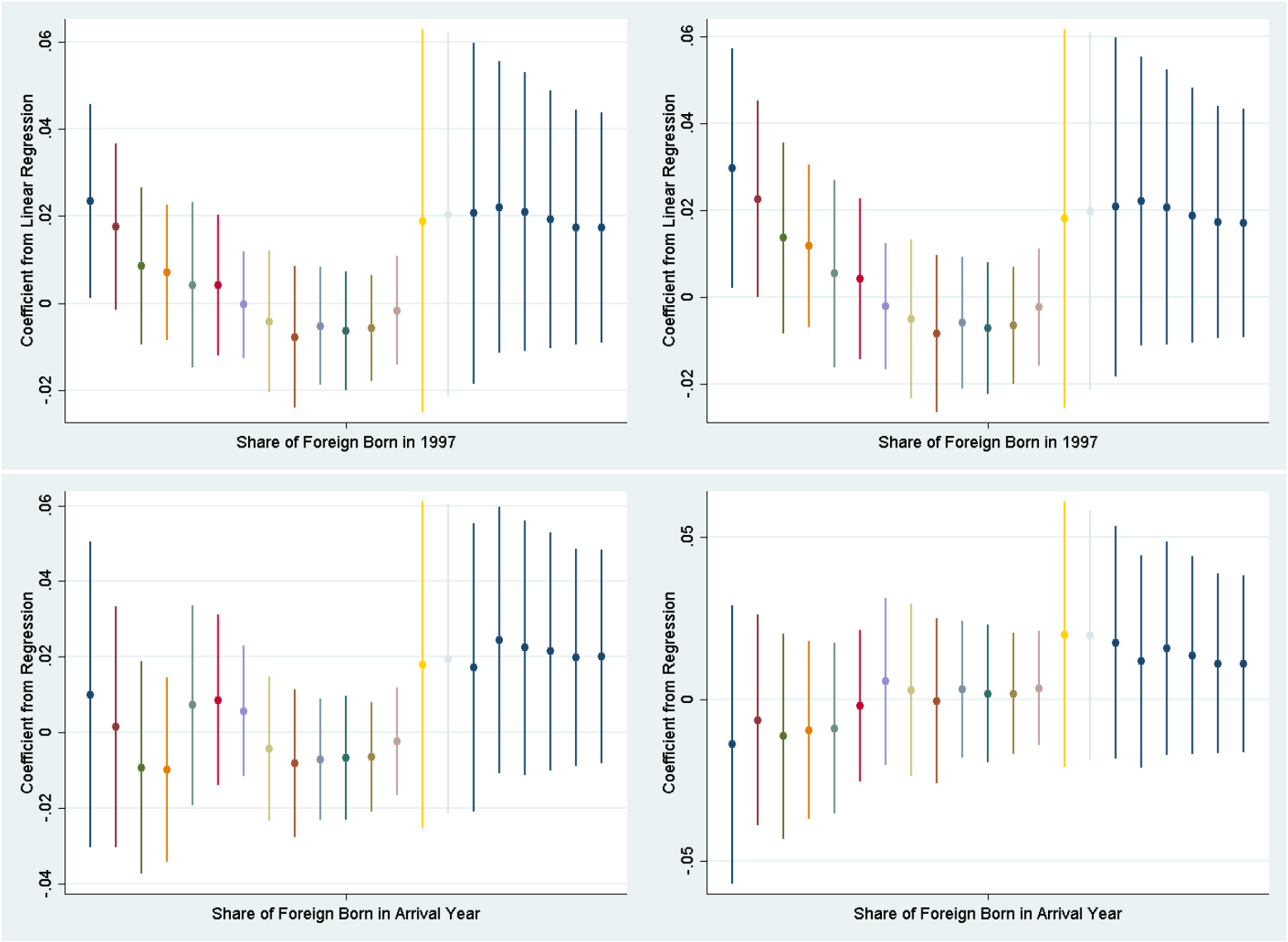
Figure: Growth rate for rich (above median), 1987-1991



Native flight and initial immigrant shares



Native flight and initial immigrant shares



Conclusions from estimations on initial immigrant shares

1. Signs of a discrete increase in the parameter estimate for an initial share of 18% or above
2. Pattern very similar for the three native-born (ethnic) groups: Difficult to find an ethnically based story that explains the results
3. The results might have important implications for the tipping point literature: If one would find significant tipping points for native-born with different ethnic backgrounds, native tipping behavior might not be based on ethnic/racial grounds but rather on socio-economic grounds

Main conclusions

Using push-driven refugee immigration to Sweden, interacted with a settlement pattern of their countrymen in the early 1990s that was partly generated by a state-run placement policy, we find:

1. Evidence of native flight, but not of native avoidance
2. Ethnic closeness between native-born individuals and the newly arrived refugees does not seem to matter for native flight
   - $\Rightarrow$ Preferences for *ethnically* homogeneous neighborhoods does not seem to be the main channel causing flight
   - Analyses of effects on socio-economic change $\Rightarrow$ indication that natives rather seem to have preferences for socio-economically homogeneous neighborhoods
3. Distinguishing between mobile/not mobile households important