Supplemental Digital Content
Does segregation lead to lower birth weight? An instrumental variable approach

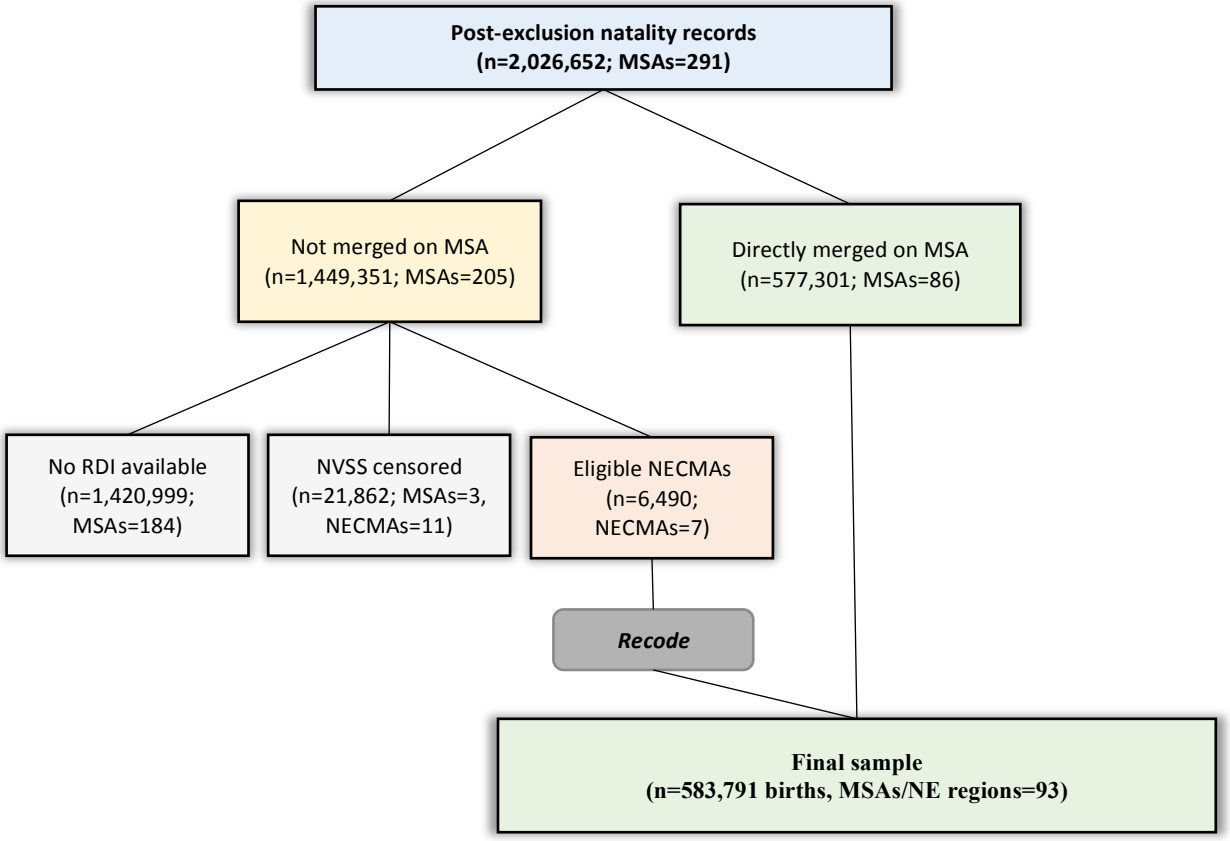
eAppendix 1: Selection of included MSAs

Data sources were merged by maternal MSA of residence at the time of delivery. Each US MSA is assigned a unique 4-digit Federal Information Processing Standards (FIPS) code. These codes were used to link the data sets for all locations except New England, as the NVSS natality files use New England Consolidated Metropolitan Area (NECMA) codes in lieu of MSAs to identify New England statistical areas. The merging process for New England locations is described below. For confidentiality purposes, the NVSS does not disclose information on a mother's MSA of residence if their place of residence contained fewer than 100,000 residents at the time of the preceding Census (which, in this case, was 1990).

Merging was highly dependent upon RDI data, as only certain MSAs received an RDI score. From the initial, post-exclusion sample of 2,026,652 natality records (including the necessary exclusion of nonmetropolitan/rural counties not belonging to an MSA), 577,301 were directly merged with the segregation data based on maternal MSA of residence. Of the natality records not merged, 1,420,999 reflected births occurring in a location/MSA for which RDI calculation was not available (including all Southern MSAs and the 77 non-Southern locations lacking available historical railroad maps), 21,862 occurred in a censored MSA, and 6,490 occurred in a New England location for which RDI data was available, but the MSA/NECMA linkage could not be made directly. The RDI was derived using historical data; it is therefore a static measure and does not change over time. However, like many geographical delineations, MSAs are defined for administrative purposes and evolve over time to reflect regional changes and growth. Ananat applied RDI information to MSAs as defined in 1988. By 2000, 14 of these previously-defined MSAs (four of which were in New England) had either ceased to exist or had been combined with other locations. Because these regions were no longer uniquely identified, they were excluded from analysis, resulting in 107 (as opposed to 121) MSAs. Of these, three MSAs had populations of less than 100,000 as of the 1990 Census and were thus censored in the 2000 natality data; these MSAs were necessarily excluded from analysis (n=104). Another 11 New England regions were excluded due to censoring as described below (final n=93).

To account for the MSA/NECMA discordance, New England natality files were linked to segregation data by a combination of principal city and state. The rationale for this approach was that the central cities would certainly be contained within the RDI calculations, the segregation indices, and the natality files, whereas attempting to aggregate New England MSAs (or even counties) to the NECMA level would have resulted in the inclusion of areas to which the RDI and segregation indices would not apply. NECMAs were generally quite diverse in comparison to MSAs and frequently contained multiple major urban areas; for example, the Boston-Worcester-Lawrence-Lowell-Brockton NECMA captured much of eastern Massachusetts (and some of southern New Hampshire) and contained locations with a wide range of dissimilarity and railroad division indices. Most NECMAs also contained parts of many counties, and it was possible for an individual county to be in multiple NECMAs. This strategy resulted in a smaller sample size for New England locations, but it was the cleanest way to incorporate this data. It should be noted, however, that this method likely produced analytical New England locations

with disproportionately high proportions of black residents, given the higher concentration of blacks in central cities than surrounding suburbs. Of the 18 New England locations in the RDI dataset that were still in existence as of 2000, 11 had populations of less than 100,000 as of the 1990 Census and were excluded from analysis due to NVSS censoring practices. The final merged dataset contained 93 MSAs (7 of which were actually New England cities) and 583,791 birth records. The merge process is concisely illustrated in the following diagram:



Finally, a list of included and excluded MSAs, as well as the reason for exclusion, is provided below. The list of "original MSAs" refers to the sample with which RDI was originally validated. There are additional US MSAs not reflected here, as RDI calculation was dependent upon historical map availability (24).

Included/excluded MSAs

			Reason for exclusion (if applicable)	
Original MSAs (Ananat)	**Included?**	**New England?**	**NVSS censored (pop<100,000)**	**MSA not uniquely identified**
Anaheim, CA	*No*			X
Aurora, IL	*No*			X
Battle Creek, MI	*No*			X

Beaver, PA	*No*			X
Brockton, MA	*No*	X	X	
Burlington, VT	*No*	X	X	
Danbury, CT	*No*	X	X	
Elmira, NY	*No*		X	
Fall River, MA	*No*	X		X
Fitchburg, MA	*No*	X	X	
Iowa City, IA	*No*		X	
Joliet, IL	*No*			X
Kankakee, IL	*No*		X	
Lawrence, MA	*No*	X	X	
Lorain, OH	*No*			X
Manchester, NH	*No*	X	X	
Middletown, CT	*No*	X		X
Muskegon, MI	*No*			X
New Bedford, MA	*No*	X	X	
New London, CT	*No*	X	X	
Niagara Falls, NY	*No*			X
Norwalk, CT	*No*	X		X
Pittsfield, MA	*No*	X	X	
Portland, ME	*No*	X	X	
Portsmouth, NH	*No*	X	X	
Poughkeepsie, NY	*No*			X
Salem, MA	*No*	X		X
Vancouver, WA	*No*			X
Akron, OH	Yes			
Albany, NY	Yes			
Altoon, PA	Yes			
Ann Arbor, MI	Yes			
Atlantic City, NJ	Yes			
Benton Harbor, MI	Yes			
Binghamton, NY	Yes			
Bloomington, IN	Yes			
Boise, ID	Yes			
Boulder, CO	Yes			
Bridgeport, CT	Yes	X		
Buffalo, NY	Yes			
Canton, OH	Yes			
Champaign, IL	Yes			
Chico, CA	Yes			
Cincinnati, OH	Yes			
Colorado Springs, CO	Yes			
Dayton, OH	Yes			
Decatur, IL	Yes			
Des Moines, IA	Yes			
Detroit, MI	Yes			
Duluth, MN	Yes			
Erie, PA	Yes			
Eugene, OR	Yes			
Flint, MI	Yes			
Fort Collins, CO	Yes			

Glens Falls, NY	Yes		
Grand Forks, ND	Yes		
Grand Rapids, MI	Yes		
Hamilton, OH	Yes		
Harrisburg, PA	Yes		
Hartford, CT	Yes	X	
Jackson, MI	Yes		
Jamestown, NY	Yes		
Janesville, WI	Yes		
Johnstown, PA	Yes		
Kalamazoo, MI	Yes		
Lancaster, PA	Yes		
Lansing, MI	Yes		
Las Cruces, NM	Yes		
Lawton, OK	Yes		
Lima, OH	Yes		
Los Angeles, CA	Yes		
Lowell, MA	Yes	X	
Mansfield, OH	Yes		
Merced, CA	Yes		
Minneapolis, MN	Yes		
Newark, NJ	Yes		
New Haven, CT	Yes	X	
Oakland, CA	Yes		
Oklahoma City, OK	Yes		
Olympia, WA	Yes		
Omaha, NE	Yes		
Peoria, IL	Yes		
Philadelphia, PA	Yes		
Phoenix, AZ	Yes		
Portland, OR	Yes		
Pueblo, CO	Yes		
Reading, PA	Yes		
Redding, CA	Yes		
Reno, NV	Yes		
Riverside, CA	Yes		
Rochester, NY	Yes		
Rockford, IL	Yes		
Saginaw, MI	Yes		
Salem, OR	Yes		
Salinas, CA	Yes		
San Francisco, CA	Yes		
Santa Barbara, CA	Yes		
Santa Cruz, CA	Yes		
Santa Rosa, CA	Yes		
Scranton, PA	Yes		
Seattle, WA	Yes		
Spokane, WA	Yes		
Springfield, IL	Yes		
Springfield, MA	Yes	X	
State College, PA	Yes		

Steubenville, OH	Yes		
Stockton, CA	Yes		
Syracuse, NY	Yes		
Toledo, OH	Yes		
Trenton, NJ	Yes		
Tucson, AZ	Yes		
Utica, NY	Yes		
Vineland, NJ	Yes		
Visalia, CA	Yes		
Waterbury, CT	Yes	X	
Williamsport, PA	Yes		
Worcester, MA	Yes	X	
Yakima, WA	Yes		
York, PA	Yes		
Youngstown, OH	Yes		
Yuba City, CA	Yes		

eTable 1: Covariate distribution across quantiles of RDI, by race

	White				
	RDI q1	**RDI q2**	**RDI q3**	**RDI q4**	***q1-q4***
Maternal age	28.8	28.2	28.8	28.0	*0.8*
Education (years)	13.9	13.7	13.9	13.5	*0.4*
Married	77.3	75.9	77.7	74.8	*2.5*
Parity	2.3	2.2	2.2	2.1	*0.2*
Prior preterm birth(s) (%)	1.3	1.3	1.4	1.5	*-0.2*
Alcohol use (%)	1.4	1.0	1.2	1.2	*0.2*
Medical risk factors	7.7	7.1	8.1	9.3	*-1.6*
Inadequate PNC	3.2	2.9	3.9	3.0	*0.2*
Low birth weight (%)	4.3	4.8	4.2	4.8	*-0.5*
Sex	51.3	51.1	51.3	51.2	*0.1*
Gestational age	39.0	38.9	39.0	39.0	*0.0*
MSA population	1,408,714	2,633,303	2,518,572	2,056,476	*-647,762*
MSA % black	8.0%	7.4%	11.1%	12.9%	*-4.9%*

	Black				
	RDI q1	**RDI q2**	**RDI q3**	**RDI q4**	***q1-q4***
Maternal age	25.3	25.3	24.7	24.9	*0.4*
Education (years)	12.5	12.5	12.3	12.2	*0.3*
Married	28.6	28.0	23.0	24.8	*3.8*
Parity	3.1	2.7	2.7	2.6	*0.5*
Prior preterm birth(s) (%)	2.1	2.0	2.4	1.5	*0.6*
Alcohol use (%)	2.2	1.8	2.1	1.4	*0.8*
Medical risk factors	8.7	6.8	8.7	8.1	*0.6*
Inadequate PNC	11.0	7.7	14.0	12.5	*-1.5*
Low birth weight (%)	11.1	10.5	11.3	12.1	*-1.0*
Sex	50.8	50.6	51.1	50.1	*0.7*
Gestational age	38.4	38.5	38.4	38.4	*0.0*
MSA population	1,725,479	4,976,287	3,507,642	2,925,026	*-1,199,547*
MSA % black	13.6%	9.8%	15.5%	17.9%	*-4.3%*

eTable 2: RDI robustness

	Individual-level birth weight			Birth weight gap[a, b]
	Overall	White	Black	Overall
Original 2SLS estimates	**-1.87**	**-0.68**	**-2.76**	**2.11**
	(-4.46, 0.72)	**(-3.48, 2.11)**	**(-6.01, 0.48)**	**(-1.78, 5.99)**
Maternal age (centered)	-1.18	-0.21	-2.63	2.12
	(-3.39, 1.04)	(-2.38, 1.96)	(-5.67, 0.40)	(-1.56, 5.81)
Marital status	-1.19	-0.49	-2.35	2.21
	(-3.36, 0.98)	(-2.88, 1.90)	(-5.10, 0.40)	(-1.71, 6.13)
Birth order	-1.87	-0.67	-2.77	2.08
	(-4.46, 0.72)	(-3.45, 2.11)	(-6.01, 0.47)	(-1.58, 5.73)
Medical risk factors	-1.84	-0.63	-2.90	2.17
	(-4.48, 0.80)	(-3.45, 2.18)	(-6.21, 0.40)	(-0.97, 5.31)
Prior preterm births	-1.80	-0.56	-3.08	1.99
	(-4.29, 0.69)	(-3.26, 2.14)	(-6.05, -0.11)	(-1.85, 5.83)
Inadequate PNC	-1.64	-0.64	-2.64	1.91
	(-4.14, 0.86)	(-3.39, 2.10)	(-5.61, 0.32)	(-1.91, 5.72)
Education	-0.85	-0.02	-2.27	2.25
	(-2.99, 1.28)	(-2.16, 2.11)	(-4.82, 0.28)	(-1.50, 6.00)
MSA proportion black	-1.26	-2.19	-4.73	1.67
	(-7.35, 4.83)	(-8.72, 4.34)	(-12.10, 2.64)	(-14.07, 17.42)
MSA population	-1.99	-1.38	-4.23	-
	(-5.19, 1.21)	(-4.67, 1.91)	(-9.01, 0.55)	-

[a]*Individual-level covariates for gap analysis aggregated by MSA & race*

[b]*MSA population weighted estimates*

eTable3: Demographics by MSA # black

	<=5000 MSA black		>5000 MSA black	
	White *n=32,978*	Black *n=459*	White *n=452,391*	Black *n=97,963*
Geographic/MSA characteristics				
Dissimilarity index	.33	.34	.62	.68
MSA population	261,342	212,732	2,287,575	3,302,918
MSA percent black	1%	2%	10%	14%
Maternal characteristics				
Age (years)	27.4	24.6	28.5	25.0
Education (years)	13.6	12.8	13.8	12.4
High school graduate (%)	88.3	79.0	89.7	74.3
Married (%)	73.5	34.6	76.7	25.9
Parity (count)	2.2	2.5	2.2	2.8
Diabetes (%)	3.0	3.1	3.0	2.9
Chronic hypertension (%)	0.6	1.5	0.7	1.3
Lung disease (%)	2.1	3.1	1.8	3.0
Prior preterm birth (%)	1.4	3.1	1.4	2.0
Alcohol use during pregnancy (%)	1.2	1.2	1.2	1.8
Inadequate PNC (%)	4.1	10.2	3.2	11.3
Infant characteristics				
Male (%)	51.0	53.6	51.2	50.6
Gestational age (weeks)	39.0	38.6	39.0	38.4
Birth weight (grams)	3413.0	3190.6	3437.4	3149.8
Low birth weight (%)	4.6	10.7	4.5	11.3