Supplementary Appendix

01 December, 2020

Mortality Records

We use public administrative records from the Social Security Administration as provided in the Berkeley Unified Numident Mortality Database (BUNMD) for this analysis. The BUNMD was constructed from the most informative parts of the Numerical Identification System (Numident). The Numident—for every person with a Social Security number—tracks date of birth, date of death (if applicable), and other background information including birthplace, race, sex, and names of parents. The Social Security Administration transferred a large number of these records to the National Archives and Records Administration (NARA), who publicly released the records in 2019. The BUNMD is a cleaned and harmonized version of the records publicly released by NARA (Goldstein and Breen 2020). For this analysis, we use records which consist of individuals who die over age 65 between 1988 and 2005.

Identifying Siblings

To establish two men in the BUNMD as brothers, we identify men who share a common mother and father's first and last name. We have confidence in the accuracy of sibling matches because of the high rate of being born in the same state (90.6%), although we did not use this as a matching criterion.

Sibling			**Father**		**Mother**	
First (standardized)	Last	Birthplace	First	Last	First	Last
Ernest	Cottman	Ohio	Royal	Cottman	Leona	Jones
Royal	Cottman	Ohio	Royal	Cottman	Leona	Jones
James	Mason	Georgia	Jonas	Mason	Nettie	Jackson
Arthur	Mason	Georgia	Jonas	Mason	Nettie	Jackson
Oscar	Watson	Louisiana	Louis	Watson	Marinda	White
Spencer	Watson	Louisiana	Louis	Watson	Marinda	White

Table 1: Establishing siblingship based on exact matches on parents' names. For exact matches, birthplace was not used as a matching field.

We identify an additional set of siblings by allowing for some flexibility on the exact spelling of parent's first names. This allows for minor misspellings and transcription errors. For these additional siblings, we (conservatively) require birthplace to match. We use the following procedure:

1. Identify potential pairs of brother (restricted to 2) who reported the same place of birth, mother's last name, father's last name, but **different** father's first names and/or mother's first name.
2. If both potential siblings reported the same father's first name but different mother's first names, establish as sibling pair if string distance between the two discrepant mother's first names names > 0.7 (based on the Levenshtein distance).
3. If both potential siblings reported the same mother's first name but different father's first names, establish as sibling pair if string distance between the two discrepant father's first names names > 0.7 (based on the Levenshtein distance).
4. If neither mother nor father's first name match exactly, establish as a sibling pair if both string distance between the two discrepant father's first names names and between the two discrepant mother's first names > 0.8 (based on the Levenshtein distance).

Sibling			**Father**			**Mother**	
First (Standardized)	Last	BPL	First*	Lev Distance	Last	First	Last
Harrison	Jones	Tennessee	Ernest	0.857	Jones	Carrie	Fields
James	Jones	Tennessee	Earnest	0.857	Jones	Carrie	Fields
George	Pittman	Georgia	Deillie	0.714	Pittman	Mattie	Henry
Dwellie	Pittman	Georgia	Dwellie	0.714	Pittman	Mattie	Henry
Arthur	Mitchell	Mississippi	Matthew	0.875	Mitchell	Beatrice	Johnson
Jarvis	Mitchell	Mississippi	Matthews	0.875	Mitchell	Beatrice	Johnson

Table 2: Illustration of establishing sibling pairs allowing for flexibility in father's first name. The same procedure was conducted for mother's first name. A minimum Levenshtein string distance of > 0.7 was used as a threshold to establish a sibling pair.

Name Standardization and BNI

We pre-process first names to account for nicknames or minor spelling and transcription errors. To clean first names by (i) removing any non-alpha characters, (ii) selecting the first word for compound string names, and (iii) restricting to first names great than one character in length. After cleaning the names, we standardize first names to account for nicknames (e.g., "Bill" $\rightarrow$ "William") and misspellings (e.g., "Williaam" $\rightarrow$ "William"). We construct a master name standardization dictionary by combining name standardization dictionaries from the IPUMS and ABE research teams.

We use the Black Name Index (BNI), a summary measure of how distinctly black a first name proposed by Fryer and Levitt (2004). We use the standardized names to compute the Black Name Index:

$$BNI_{name_i} = \frac{P(name_i|Black)}{P(name_i|Black) + P(name_i|White)} \tag{1}$$

To compute the BNI, we use BUNMD birth cohorts of 1895-1940 with information on race. This larger sample gives more stability. We restrict to names that occur over 500 times to avoid the noise introduced by BNI.

Estimated Effect on BNI

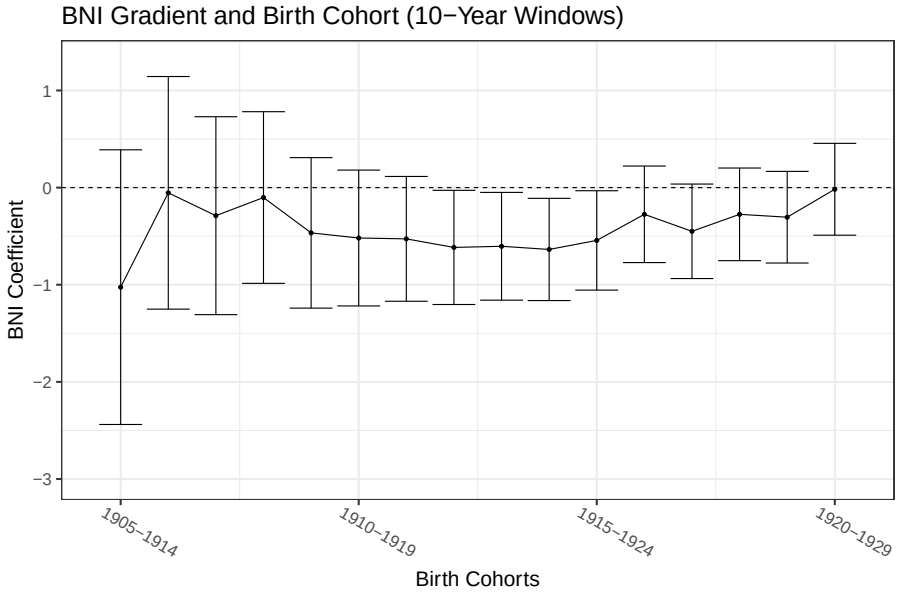
The estimated BNI gradient for additional siblings (model 2) is smaller than for the exact-matched siblings (model 1). While this may reflect incorrectly-identified sibling pairs biasing results toward population-wide estimates (i.e., our pooled OLS model coefficient), the small sample precludes any definitive conclusions. For the combined sample, both the BNI gradient and standard error decreased slightly.

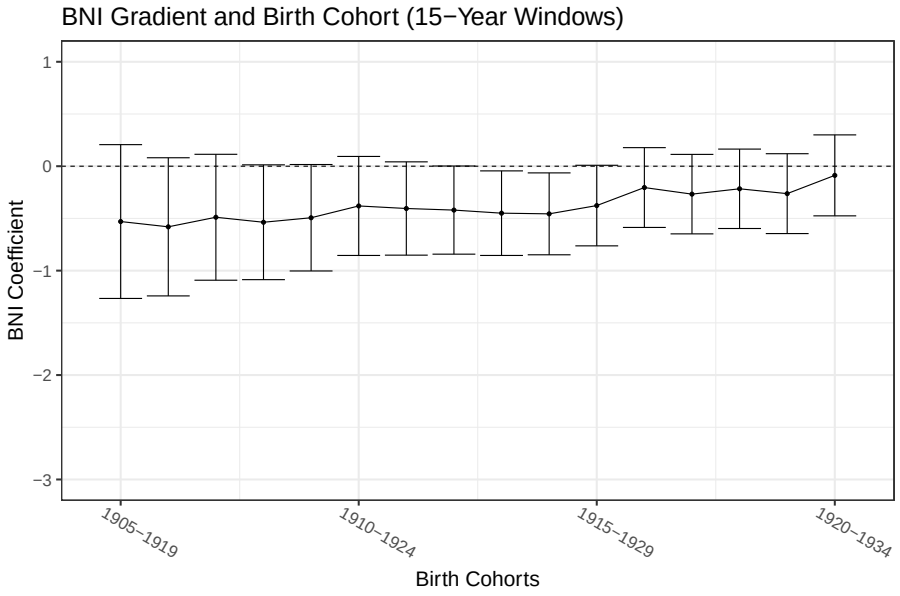
	Dependent variable:		
	death_age		
	(1)	(2)	(3)
	bni.exact.sibs	bni.add.sibs	bni.all.sibs
bni_std	−0.582**	−0.385	−0.544**
	(0.286)	(0.639)	(0.261)
Observations	24,278	4,894	29,155
R^2	0.595	0.618	0.598
Adjusted R^2	0.214	0.232	0.217
Residual Std. Error	4.879 (df = 12519)	4.867 (df = 2435)	4.878 (df = 14954)
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 3: Models include fixed effects for birth year, birth order, and family. Pooled cohorts of 1915-1925.

Sensitivity Analysis: Birth Cohorts

For our main analysis, we used the BUNMD birth cohorts of 1915-1925. The plot below shows the estimated BNI gradients ($\beta_1 \pm 1.96 \times SE(\beta_1)$), where the 10 or 15-year window of birth cohorts was systematically varied. The results are robust across the birth cohort windows corresponding to BUNMD high death coverage period.





References

Bertrand, Marianne, and Sendhil Mullainathan. 2003. "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination," 40.

Chae, David H., Amani M. Nuru-Jeter, Nancy E. Adler, Gene H. Brody, Jue Lin, Elizabeth H. Blackburn, and Elissa S. Epel. 2014. "Discrimination, Racial Bias, and Telomere Length in African-American Men." *American Journal of Preventive Medicine* 46 (2): 103–11. https://doi.org/10.1016/j.amepre.2013.10.020.

Cook, Lisa D., Trevon D. Logan, and John M. Parman. 2014. "Distinctively Black Names in the American Past." *Explorations in Economic History* 53 (July): 64–82. https://doi.org/10.1016/j.eeh.2014.03.001.

Cook, Lisa, Trevon Logan, and John Parman. 2015. "The Mortality Consequences of Distinctively Black Names," 29.

Gaddis, S. 2017. "How Black Are Lakisha and Jamal? Racial Perceptions from Names Used in Correspondence Audit Studies." *Sociological Science* 4: 469–89. https://doi.org/10.15195/v4.a19.

Garcia, Denia, and Maria Abascal. 2015. "Colored Perceptions: Racially Distinctive Names and Assessments of Skin Color." *American Behavioral Scientist*, October. https://doi.org/10.1177/0002764215613395.

Gerhards, J., and S. Hans. 2009. "From Hasan to Herbert: Name-Giving Patterns of Immigrant Parents Between Acculturation and Ethnic Maintenance1." *American Journal of Sociology*. https://doi.org/10.1086/595944.

Goffman, Erving. 2009. *Stigma: Notes on the Management of Spoiled Identity*. Simon; Schuster.

Goldstein, Joshua, and Casey Breen. 2020. "Berkeley Unified Numident Mortality Database: Public Administrative Records for Individual Level Mortality Research," 33.

Goldstein, Joshua R., and Guy Stecklov. 2016. "From Patrick to John F.: Ethnic Names and Occupational Success in the Last Era of Mass Migration." *American Sociological Review* 81 (1): 85–106. https://doi.org/10.1177/0003122415621910.

Hacker, J. David. 1999. "Child Naming, Religion, and the Decline of Marital Fertility in Nineteenth-Century America." *The History of the Family* 4 (3): 339–65. https://doi.org/10.1016/S1081-602X(99)00019-6.

Johfre, Sasha. 2020. "What Age Is in a Name?" *Sociological Science* 7: 367–90. https://doi.org/10.15195/v7.a15.

Jr, Roland G Fryer, and Steven D Levitt. 2004. "Understanding the Black-White Test Score Gap in the First Two Years of School." *THE REVIEW OF ECONOMICS AND STATISTICS*, 18.

Lieberson, Stanley. 2000. *A Matter of Taste: How Names, Fashions, and Culture Change*. Yale University Press.

Lindsay, Jo, and Deborah Dempsey. 2017. "First Names and Social Distinction: Middle-Class Naming Practices in Australia." *Journal of Sociology* 53 (3): 577–91. https://doi.org/10.1177/1440783317690925.

Link, Bruce G., and Jo C. Phelan. 2001. "Conceptualizing Stigma." *Annual Review of Sociology* 27 (1): 363–85. https://doi.org/10.1146/annurev.soc.27.1.363.

Obukhova, Elena, Ezra W. Zuckerman, and Jiayin Zhang. 2011. "Why Was There No Fashion in Mao's China? Effects of Politics on Culture in the Case of Chinese Naming Practices." *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1878678.

Pager, Devah, and Bruce Western. 2003. "Discrimination in a Low-Wage Labor Market: A Field Experiment," 23.

Stewart, Quincy Thomas, Ryon J. Cobb, and Verna M. Keith. 2020. "The Color of Death: Race, Observed Skin Tone, and All-Cause Mortality in the United States." *Ethnicity & Health* 25 (7): 1018–40. https://doi.org/10.1080/13557858.2018.1469735.