Contours and Consequences of Black First Names in the Historical United States

Extended Abstract for PAA 2014 Submission

Joshua R. Goldstein (UC Berkeley)
Guy Stecklov (Hebrew University)

September 27, 2013

Abstract

The distinctiveness of African American first names has become a topic of much interest in recent decades. In this paper, we document the historical pattern of distinctive naming using census microdata from 1900 to 1930. We show (1) Certain first names were given disproportionately to African Americans (2) The distinctiveness of naming was related to the social position of the parents and (3) Even controlling for parental social position, more distinct names were associated with lower occupational achievement. This third finding suggests that even in environments of high racial segregation and where people differ by skin color that the use of cultural symbols like names can influence socio-economic achievement.

1 Introduction

Interest in distinctive first names among African-Americans has grown in recent years, due in large part to the increasing practice of giving newly-invented, unique names. Audit studies have found that the use of recognizably "black" names disadvantages job applicants. Even as the color line in American society has become less rigid, there is still a clear effect of appearing "black".

1

In this paper, we use historical census records from 1900 and 1930 to look at the naming practices of African-Americans before World War II. We are interested in three research questions:

1. Were there recognizably "black" names in the past?

2. Was there a socio-economic gradient in naming, such that higher status African-Americans gave "whiter" names?

3. In a highly segregated society characterized by the one-drop rule, does it appear to have been an economic disadvantage of carrying a "blacker" name while having a "whiter" made it easier to succeed?

All of these questions aim to give a deeper understanding of the nature of racial discrimination and differentiation in the highly segregated context of Pre-War United States. The use of names provides, we argue, a new source for measuring gradations in group membership. In a society characterized by the one-drop rule, in which any African-American ancestry was sufficient to make some one wholly "black" in the eyes of the law, was there nonetheless a detectable continuum within the African-American population?

We make use of a continuous quantitative measure of how "black" or "white" a first name is based on its statistical distribution in the population. A name that is disproportionately found among African-Americans is considered a "blacker" name than one that is found equal proportions among blacks and whites.

2 Literature

Naming practices have long been a topic of sociological study. Beginning with Rossi (1965), names have been studied for what they reveal about group membership. The changing patterns of naming has itself become a topic of study, notably by Lieberson (2000).

More recently, economists have begun to use names as both an experimental variable in discrimination studies Bertrand and Mullainathan (2003); Neumark (2012).

Studies on naming of African Americans include Fryer and Levitt (2004) and Lieberson and Mikelson (1995) for contemporary populations. Historical studies include London and Morgan (1994) and the recent papers by papers by Cook et al. (2012, 2013).

The contributions of this paper are to introduce the large-scale use of census data for the study of naming practices. This enables comparisons across space and time – and importantly also includes individual measures of socio-economic status. The census also allows the study of the presentation of self by respondents. For example, we can study the use of nicknames as a measure of informality and the use of multiple first names as a measure of formal presentation of full names. Each of these it turns out are strongly related with socio-economic status. Finally, our measurement of gradients in name giving and in achievement by name allow us to better identify the potential causal effect of carrying names.

3   Data and Methods

Our analysis uses the publicly-available IPUMS micro-samples of the U.S. censuses of 1900 and 1930. These samples, no longer covered by confidentiality safeguards, include full transcriptions of first and last names. In addition to names, the census records also include individual-level information on age, race, father and mother's birthplace, individual occupation, and basic individual demographics, such as age and sex.

Our primary data set is the census of 1930, for which a 5 percent sample is available. In order to take advantage of occupational information, we limit our analysis to males aged 20 to 50 in 1930, giving a sample size of slightly more than 100,000 black respondents and about 10 times this number of white respondents. We exclude the population of individuals with foreign born fathers because of the additional complexities that ethnic origins introduce into how names are seen and used to discriminate.[1]

To measure occupational attainment, we use the synthetic IPUMS variable, OCCSCORE. OCCSCORE is based on the average income observed in the 1950 census, in hundreds of dollars, and adjusted to account for changing occupational categories over time (Ruggles et al., 2010). Advantages of OCCSCORE include comparability over time, ease of interpretation, and ranking of occupations along a single dimension. A disadvantage, shared with all occupational measures, is that it cannot detect variation within occupations. We use OCCSCORE thus not as a proxy for direct measurement of income at the individual level, but more broadly as a measure of occu-

---

[1]The children of foreign born are the object of study by Goldstein and Stecklov (2013).

pational achievement, in line with the long tradition of occupational status (Stevens and Featherman, 1981).

Any cursory examination of the names reported in the census reveals, particularly for blacks, reveals a large number of variations of spelling for common names as well as a wide range of nicknames. Nicknames, as we discus further below, are interesting in and of themselves. However, there are many names that are quite clearly misspelled as well. In the case of the 1930 census, census officials visiting the homes filled in census data. The census takes themselves might have misspelled names and it is possible that less literate census takers were used in black residential areas. We use a list of name-modifications developed by IPUMS staff to transform recorded names to their standard forms (e.g., "JIM" → "JAMES"). This translation file still needs improvement as some common nick-names (e.g., "EDDIE") are left un-standardized.

4   Methods

4.1   Measurement of "blackness" and "whiteness"

We adopt the black name index ("BNI") from contemporary research by Fryer and Levitt based on the relative frequency of each name in the black population compared to the frequency in the entire population (Fryer and Levitt 2004). The BNI measure for name j is defined as

$$\text{BNI}(\text{name}_j) = \frac{p(\text{name}_j|\text{black})}{= p(\text{name}_j|\text{black}) + p(\text{name}_j|\text{white})}$$

where p is the fraction holding given name j among either blacks or whites. For example, the name Samuel is held by about 2 percent of blacks in our sample and it is held by about 0.8 percent of whites. This gives a BNI(Samuel)=0.02/(0.02+0.008) = 0.71. This metric ranges from 0 to 1. A value of zero for the index means the name is absent among the black group, and present in the white population. A value of 1.0 means the name is exclusive to blacks. The name Samuel is clearly more popular among blacks but not exclusive.

5 Preliminary Results

5.1 The distinctiveness of names

In figure 1, we show all of the names ranking in the top 100 among either whites or blacks in terms of their relative ranking in the two groups. Thus, "MOSES" is in the top 100 black names but ranks above 300 among white names, whereas "KENNETH" ranks within the top 100 white names but near 200 among white names. The figure allows us to see that there are clear divergences in naming, even among names that are relatively common. The "whiter" and "blacker" names are not exclusive, but are disproportionately held by one group.

The names that rank higher among blacks than whites include religious names ("MOSES", "ISAIAH", "ISAAC", "LEVI", "ELIJAH") and president names ("CLEVELAND", "ROOSEVELT") as well as nick-names that have not been standardized using the IPUMS file ("EDDE", "MOSE", "BENNIE").

The names that rank higher among whites than blacks include Germanic names ("EMIL", "AUGUST", "OTTO", "RALPH") as well as some more common Catholic names (e.g., "FRANCIS").

Table 1 shows selected names in order to illustrate how BNI scores are closely related to the difference in ranking. Here we have shown names that are not at the extremes in order to illustrate the graded nature of our index of distinctive black naming.

Our answer to our first question about the existence of distinct naming patterns is that indeed there appears to be a clear gradient in naming, with some names being disproportionately held by blacks and others being disproportionately held by whites.

5.2 The status gradient in name giving

In Latin America there has been a great deal of study as to how "money whitens" (Loveman and Muniz, 2007; Schwartzman, 2007). The same issues are of interest in the contemporary United States with the advent of multiracial classification. Our methods allow us to study a continuum of racial affiliation within the black population by looking at whether higher-status African American parents gave their children less distinctively black names.

Figure 2 shows the patterns of naming by the BNI score of the name

and father's status in 1900. In the figure we can see a clear tendency for higher income African-American fathers to give their sons "whiter" names like "EARL", "RAYMOND", and "CLARENCE", whereas more distinctive black names like ("ISAIAH" and "WASHINGTON") tend to be given by father's with lower earnings.

Multivariate analysis controlling for region, and rural-urban residence, and including more names, confirms this pattern of "blacker" names coming from lower earning African-American families.

5.3 Carrying a "blacker" name is associated with lower occupational achievement

Figure 3 shows the relationship between BNI and the Occupational scores of adult African-American in the 1930 Census. Here we see that those carrying the "whitest" names like "HAROLD", "CARL" and "RALPH" have the highest occupational success, whereas those carrying the most distinctively "black" names like ROOSEVELT and ISAIAH are among the lowest earning.

The figure also shows the tendency of those for whom nicknames like "EDD", "BENNIE", and "MOSE" to have low earnings. Having a nickname in the census could result from a combination of respondents giving nicknames and enumerators not expanding them to full names. We plan to study this issue in more detail since it bears on the causal question of whether the name influences status or whether the relationship is driven by reverse causality with status influencing naming.

In early multivariate analysis we find that the negative relationship between BNI and occupational income is robust to controls for region, age and rural residence. Surprisingly, it also seems to be robust for controls for the status gradient of the family background of those receiving names. It also persists after our initial standardization of nicknames (although this needs to be improved). Our preliminary conclusion is that names to appear to matter in some causal sense. Even in the one-drop society of 1930, having a "blacker" name appears to have disadvantaged African Americans, and having a "whiter" name appears to lead to higher earnings.

6   Research Plans

Our next steps include

1. Multivariate analysis of the effects of names on occupational outcomes, controlling for family background effects.

2. Improved recoding of nicknames and misspellings.

3. Small area analysis of naming patterns in order to identify both the effect of enumerators and to identify highly localized naming patterns.

References

M. Bertrand and S. Mullainathan. Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. Technical report, National Bureau of Economic Research, 2003. URL `http://www.nber.org/papers/w9873`.

L. D. Cook, T. D. Logan, and J. M. Parman. The long-term consequences of distinctively black names. 2012.

L. D. Cook, T. D. Logan, and J. M. Parman. Distinctively black names in the American past. Technical report, National Bureau of Economic Research, 2013. URL `http://www.nber.org/papers/w18802`.

R. G. Fryer and S. D. Levitt. The causes and consequences of distinctively black names. *The Quarterly Journal of Economics*, 119(3):767–805, 2004. URL `http://qje.oxfordjournals.org/content/119/3/767.short`.

S. Lieberson. *A matter of taste: How names, fashions, and culture change.* Yale University Press, 2000.

S. Lieberson and K. S. Mikelson. Distinctive African American names: an experimental, historical, and linguistic analysis of innovation. *American Sociological Review*, pages 928–946, 1995. URL `http://www.jstor.org/stable/10.2307/2096433`.

A. S. London and S. P. Morgan. Racial differences in first names in 1910. *Journal of Family History*, 19(3):261–284, 1994. URL `http://jfh.sagepub.com/content/19/3/261.short`.

M. Loveman and J. O. Muniz. How Puerto Rico became white: Boundary dynamics and intercensus racial reclassification. *American Sociological Review*, 72(6):915–939, 2007. URL `http://asr.sagepub.com/content/72/6/915.short`.

D. Neumark. Detecting discrimination in audit and correspondence studies. *Journal of Human Resources*, 47(4):1128–1157, 2012. URL `http://jhr.uwpress.org/content/47/4/1128.short`.

A. S. Rossi. Naming children in middle-class families. *American Sociological Review*, pages 499–513, 1965. URL `http://www.jstor.org/stable/10.2307/2091340`.

S. Ruggles, J. T. Alexander, K. Genadek, R. Goeken, M. B. Schroeder, and M. Sobek. Integrated public use microdata series: Version 5.0 (machine-readable database). Minneapolis: University of Minnesota, 2010.

L. F. Schwartzman. Does money whiten? Intergenerational changes in racial classification in brazil. *American Sociological Review*, 72(6):940–963, 2007. URL `http://asr.sagepub.com/content/72/6/940.short`.

G. Stevens and D. L. Featherman. A revised socioeconomic index of occupational status. *Social Science Research*, 10(4):364–395, 1981. URL `http://www.sciencedirect.com/science/article/pii/0049089X81900119`.

Figure 1: Relative ranks of top 100 black and white names, 1930. Source: Author's calculation from IPUMS using standardized names. Population restricted to males aged 20 to 50 in 1930 with native-born fathers.
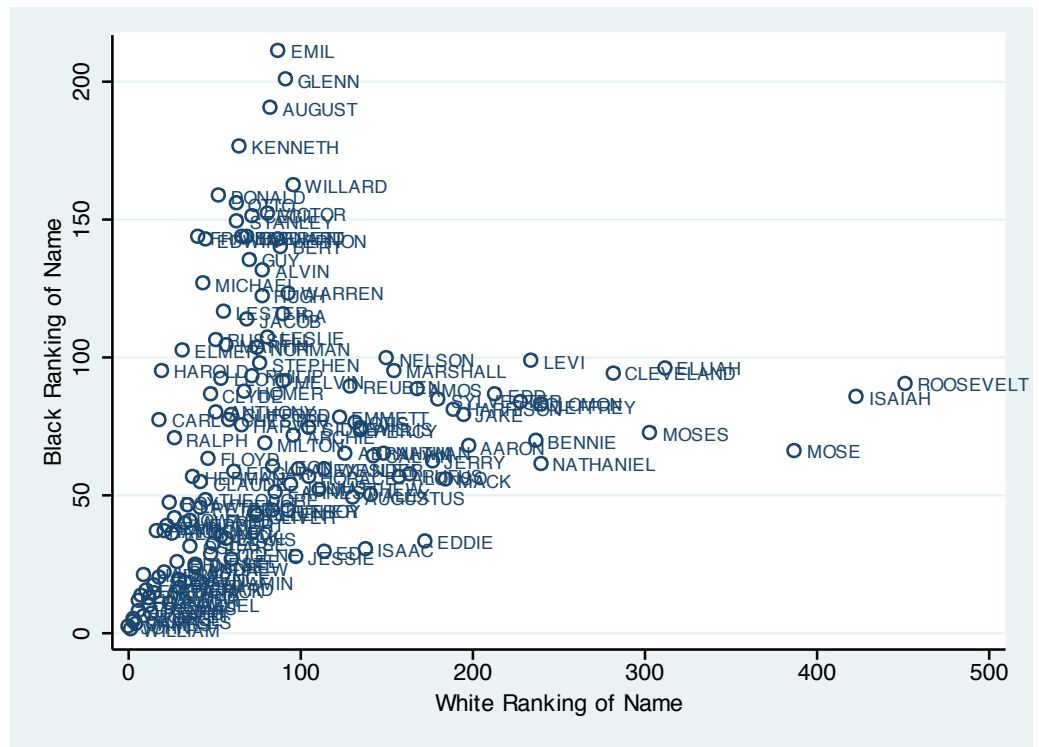
Table 1: Non Distinctively Black Names (BNI 0.1-0.2) and Distinctively Black Names (BNI 0.8-0.9) for names ranked in top 100 for either whites or blacks

Name	BNI Score	Black Name Rank	White Name Rank
AUGUST	0.104	190	83
KENNETH	0.113	176	65
FRANCIS	0.123	143	41
DONALD	0.130	158	53
HAROLD	0.131	95	20
EDWIN	0.137	142	45
CARL	0.145	77	19
OTTO	0.154	155	63
ELMER	0.164	102	32
STANLEY	0.169	149	64
MICHAEL	0.176	126	44
RALPH	0.181	70	27
ED	0.801	29	114
MACK	0.812	55	185
JEFFREY	0.817	82	241
BENNIE	0.830	69	237
ISAAC	0.838	30	138
CLEVELAND	0.847	94	282
NATHANIEL	0.851	61	240
ELIJAH	0.872	96	312
EDDIE	0.878	33	173
MOSES	0.892	72	304

Figure 2: The status gradient in name giving in 1900. The population of "sons" is restricted to black males under 10 years of age in 1900 who have native-born fathers. BNI ranges from 0 (exclusively whites) to 1.0 (an exclusively black name). OCCSCORE is annual earnings in 100s of dollars. (See text for details.)

Figure 3: The achievement gradient by name, 1930. black males aged 30-50 in 1930 with native-born fathers. BNI ranges from 0 (a name that is held only by whites) to 1.0 (an exclusively black name). OCCSCORE is annual earnings in 100s of dollars. (See text for details.)