

What's in a Name?

Saku Aura

University of Missouri and CESifo

and

Gregory D. Hess

*Claremont McKenna College and CESifo **

revised: September 2007

Abstract

This paper analyzes two broad questions: Does your first name matter? And how did you get your first name anyway? Using data from the National Opinion Research Centers (NORC's) General Social Survey we find evidence that, even after controlling for a myriad of exogenous background factors, first name features are predictors of many lifetime economic outcomes that are related to labor productivity such as education, happiness and early fertility. However, we also find evidence, based on the differential impacts of gender and race on the 'blackness' of a name that identity could be an important channel for linking first name to lifetime economic outcomes.

JEL Codes: D1, J1, J7.

Keywords: Names, Identity, Discrimination.

*We are extremely grateful to Tom Smith for help in obtaining the data. We also thank Sandy Darity, Hirsh Kasper, Steve Levitt and Ken Troske as well as seminar participants at Korea University and Claremont McKenna College for their comments. Address correspondence to Saku Aura, Department of Economics, University of Missouri, Columbia, MO 65211. Tel: (573)-882 6073. E-mail: auras@missouri.edu.

1 Introduction

Recently, economists have turned to the analysis and impact of an individual's first name.¹ This paper contributes to this literature by trying to answer two questions. First, does a person's name convey information about their background? Second, does a person's name have an impact on their long run economic outcomes such as income, education, fertility, social standing, happiness or prestige?

To date, a central question of the economic analysis of names has been on the impact of having an African American Name (i.e. the 'blackness' of a name).² More formally, these studies contrast the outcomes of an individual who has a 'black' name with the outcomes of an individual who has a 'white' name, where the presumed racial characteristics of a name can be inferred from the probability that an individual is of a certain race, conditional on having a particular name. In our dataset, this would mean comparing the outcomes of otherwise similar individuals with a name like Mark (exclusively a white name in our data) with Marcus (exclusively a black name in our data) or comparing Alice (a white name) with Tanisha (a black name).

We extend this analysis in two ways. First, we analyze additional factors of a name including both pure linguistic factors (e.g. whether a name ends in an 'oh'-sound) and factors reflecting the pure societal standing of the name (e.g. such as a name's popularity).³ These additional control variables for name characteristics can be thought of as control variables for other information that a name can convey, such as parental education and/or socio-economic background more broadly.

¹ Note that, in earlier work, names have also been a topic of research for other social scientist and linguistics. Sociologist Stanley Lieberson, in work with co-authors, provides an extended descriptive discussion on what influences a parent's choice of their children's first names, and extensively documents static and dynamic features of the distribution of first-names in the US across time, e.g. Lieberson and Bell (1992), Lieberson and Mikelson (1995) and Lieberson (2000). While it does not do full justice to their rich analysis, their framework is mostly concentrated on the question on what influences pure parental preference for first names as opposed to the more structural economic mechanisms considered here.

² This question has been most notably asked in Bertrand and Mullainathan (2004) and in Fryer and Levitt (2004).

³ Similar factors have also been used in Figlio (2004).

Second, by using a sample from the General Social Surveys (GSS), we analyze a broader set of both an individual's background and outcome variables than in the previous literature. Our central finding is that first name features are correlated with several economically relevant lifetime outcomes even after controlling for exogenous background factors.

The rest of paper proceeds as follows: Section 2 briefly reviews the existing literature. Section 3 presents our methodology and data, while Section 4 provides our main findings. We conclude in Section 5 and include some suggestions for further work.

2 Previous Literature

In a compelling paper, Bertrand and Mullainathan (2004) undertake a carefully controlled experimental study of the extent to which an African American first name affects one's job prospects. Constructing a set of synthetic resumes that differ based on whether the applicant has an African American name, the researchers send these synthetic resumes to actual job openings listed in the newspapers of a number of major cities. Importantly, they find that in order to receive one job interview, a resume with a black name needs to be sent to 15 openings. In contrast, to receive one job interview, a resume with a white name needs to be sent to only 10 job openings. Furthermore, by using auxiliary information from birth certificates from Massachusetts that also list a mother's education, Bertrand and Mullainathan construct measures of expected maternal education level for each name in their sample. Their results imply that it is the racial information conveyed in the name and not the parental background factor signaled by the name (i.e. parental education) that potential employers are using as the basis to select between resumes.⁴ They conclude that this shows that racial discrimination is a factor in the job market.

A similar set of issues is addressed in another recent study by Fryer and Levitt (2004). They make both theoretical and empirical contributions to our understanding of the economic role of first

⁴ Bertrand and Mullainathan also found little evidence that names as a social background marker have an effect on the probability that a resume leads to a callback from a potential employer.

names. With regards to the former, they outline four approaches to the economic analysis of first names: namely, ignorance, price theory model, a signaling model, and an identity model. The first model speaks for itself: it assumes that parents ignore or are unaware of any consequences of their child's first name. The price model, assumes that parents choose their child's name to maximize their expected utility. In other words, if indeed an individual of social group A is more likely to be in a society dominated by group A (B) individuals, then their parents will choose names that are more in keeping with names from group A (B), as such a choice is likely to increase their child's economic opportunities. Third, the signaling model suggests that parents signal their type (i.e. are they a member of group A or B) by choosing names that are consistent with the features of those in type A or B. In doing so, individuals can learn about the parent's type through their choice of child's name. Finally, the identity model attempts to address the presumed shortcoming of the signaling model: namely, that sending signals is generally assumed to be costly. Accordingly, if ethnic names are not costly, then economic models should incorporate a positive benefit to ethnic names. Such an approach is used by Akerloff and Kranton (2000). They provide a model and set of examples whereby an individual takes on actions that may even appear to be detrimental, in order to improve their sense of self as a member of an identified group. Applied to name choice situation, their model is closer to the framework of Lieberman and coauthors, since the choice of the name is driven by a direct argument of utility function (identity) and not by an indirect effect through the outcomes of social or market interactions like in the models of Fryer and Levitt.

Fryer and Levitt also provide an important empirical investigation as to whether the blackness of a respondent's name (that is, the conditional probability that a respondent's name is held by someone that is black) affects a respondent's fertility and other adult outcomes. Using birth records from the State of California, they analyze the effect of the blackness of a respondent's name on unwelcome birth outcomes – low birth weight babies, unmarried parents, lack of private health insurance, total children born, as well as birth mother's and father's educational attainments at the time of birth, and zip-code specific (e.g. income) and hospital specific (e.g. the racial make-up of the hospital's patients) data are the dependent variables. In general, they find that the blackness

of the mother's name is related to unwelcome birth outcomes. Using the broad time series of their data, however, Fryer and Levitt can match a mother's information to her own birth, and in doing so they demonstrate that the blackness of a mother's names is associated with her own parent's outcomes (i.e. was she born out of wedlock, was her own mother a teenager, etc ...). In particular they show that the blackness of a woman's name is related to the poor birth outcomes of her parents. They conclude from these findings that since there is evidence that the blackness of a name reflects in part the individual's background that they grew up in, it is likely to be related to factors that could realistically affect their labor productivity. As such, they reconcile their work with Bertrand and Mullainathan by noting that while a black name may get fewer interviews, interviewers are likely to be inferring something about the individuals background and perhaps labor productivity.

Figlio (2004) provides a fascinating investigation of first names that uses data containing test scores, gifted status and grade promotion information from a large Florida School District to analyze the effect of first names on these student outcomes. He carefully exploits within sibling-pair (and even within twin-pair) and within school variation in the first name characteristics to control for potential confounding factors to estimate the name effect. His main finding is that children with either a name associated with a low socio-economic status (and to lesser extent, students having a African American name) fare worse in these outcome measures. He suggest that the mechanism is that the teachers expect less from students with such name characteristics, suggesting a subtle form of discrimination.

A study very close in methodology to Bertrand and Mullainathan is Ball *et al.* (2004). However, rather than conduct a field experiment, they conduct a laboratory experiment using undergraduates to test the effect of revealing the first name of the recipient to the behavior of the dictators in a dictator game.⁵ Their central finding is that knowing that the recipient had an African American name led the dictators to give smaller shares to the recipients. Interestingly in their study

⁵ In a dictator game, the first subject (dictator) decides how to share a given sum of money with a second subject (recipient). While not rational in a one-shot game without altruism, typically the dictator gives significant share of the money to the other subject.

Arabic names were not subject to similar discrimination.

Other recent relevant papers in this literature include Foster (2004), Disdier *et al.* (2006), Woodward (2005), and Arai and Skogman Thoursie (2006). Foster (2004) tests the effects of first names in student outcomes in college setting finding moderate effects of first names on student outcomes. Disdier *et al.* (2006), Woodward (2005) and Arai and Skogman Thoursie (2006) all focus on the choice aspect of names. Disdier *et al.* (2006) analyze the effect of popular culture (songs, television shows and movies) on the name choices made by French parents. While their findings emphasize the endogeneity of name frequencies in popular culture, they conclude that popular culture also drives the popularity of names. Woodward (2005) analyzes the evolution of concentration of the most popular names in the US over the last century. He finds no evidence that concentration (homogeneity) has increased over this time period. Arai and Skogman Thoursie (2006) analyzes the effect of a Swedish immigrant's change of surname on their labor earnings.⁶ Their key finding is that earnings are 26% higher for immigrants who changed their name to a Swedish-sounding 'neutral' name as compared to immigrants who kept their non-Swedish sounding surnames.

3 Methodology, Overview of the Results and Data

Our study differs from most these previous studies in the following three important ways. First, rather than focus exclusively on the issue of African American names, we explore broader linguistic features of names that may convey meaning and/or may bring out discrimination. Generally speaking, names may affect lifetime outcomes or indicate something about a respondent's background that may work through channels other than race. Hence we include in our analysis of the impact of first names a broader, though simple, class of first name characteristics which will help us to avoid this omitted variable issue.

Second, unlike the aforementioned studies that are limited in the literal outcomes they can

⁶ While this analysis is not directly about first names, it is relevant here since it addresses two themes in the first name literature – the relevance of names for ethnic discrimination and name as a choice.

consider (i.e. interviews and early births), our use of the General Social Survey (GSS) allows us to consider a broader set of individual economic outcomes.⁷ More specifically, we investigate the extent to which a respondent's first name features affect his or her years of formal education, family income, self reported financial relative position as well as social class and happiness, probability to have a child before age of 25, and occupational prestige.⁸ Third, because we have data for both men and women, we can examine the gender differences between lifetime outcomes and first name features. Together, our use of broader linguistic features of names coupled with a more comprehensive set of lifetime outcomes for both male and female respondents makes our study an important contribution to the understanding of the role that first names play in our economic lives.

3.1 Overview of the Results

Our analysis produces three main empirical findings. First, there is a strong empirical relationship between an individual's first name and their background. Second, there is a weaker empirical relationship between an individual's lifetime outcomes and their first names. Taken together, these first two findings are consistent with those of Fryer and Levitt (2004) – names may convey information about an individual's labor market productivity.

However, our third empirical finding is that non-black non-whites females with 'blacker' names have significantly worse financial outcomes. This provides a potentially interesting insight on how a subtle form of discrimination might possibly work: namely, while black names come with discrimination and identity costs and benefits for black individuals, non-black non-whites with 'blacker' names face the costs of such names though not the benefits. This final point is

⁷ To be fair, Fryer and Levitt (2004) provide a similar type of analysis, though a number of their dependent variables are at the level of an individual's zip code, e.g. per-capita income in the individual's zip code. As well, we have a substantially larger set of control variables for background as well as outcomes for both males and females.

⁸ Other economists have also considered such a variety of lifetime economic outcomes. For example, Fryer and Levitt (2004) analyze the outcome of having a child early, Stokey (1996) has evaluated occupational prestige, and Oswald (1997) has investigated happiness outcomes.

consistent with the findings of Bertrand and Mullainathan (2004) of a link between first names and discrimination.

The usual caveats about the causal interpretation of regression coefficients as evidence of discrimination naturally apply here, primarily because we do not have exogenous instruments for the name choice in our study. Thus we can only report a statistical correlation consistent with a particular discrimination theory. Indeed, results that are attributed to discrimination but that are just proxies for unobserved labor market characteristics could make the conclusion of discrimination to be unwarranted. Nevertheless, our study points to the following direction: if discrimination is connected to an individual's first name, then an identity channel is very likely part of the explanation.

3.2 The Data

The data employed in the paper is from the General Social Survey (GSS), an annual U.S. survey that asks respondents specific questions about their financial status, standard demographic variables (e.g., age, education, etc...), family information and background, as well as many more areas. The respondents are not re-interviewed across years so that data are repeated cross sections rather than a panel. In two years, 1994 and 2002, the respondents were asked their first names which was coded by the interviewers.⁹ There are a total number of just under six thousand respondents who reported names in the 1994 and 2002 surveys. Summary statistics for the key variables are reported in the Table 1.

The key variable of interest for this study is the extraction of important *first name features* from a respondent's name. The names data are recorded by interviewers and only clear keyboard errors were fixed, such as if the number '1' was struck instead of an 'l' or if the word 'respondent' was recorded as the first name. The *first name features* we create are as follows. Following Fryer

⁹ This data was made available to us based on a joint understanding between the authors and the National Opinion Research Center that the data would remain strictly confidential.

and Levitt (2004), the *blackness index* is the probability of being black conditional on having the name in our sample. This was calculated only for names that have at least two occurrences in the sample.¹⁰

In addition, we also allow for a rich set of more linguistically based features of first names. As noted above, this broader set of first name features will allow us to reduce the likelihood of omitted variable bias if the effect of first names on lifetime outcomes works through a channel other than through race. To keep things simple, however, we consider well-known linguistic (i.e. ending and beginning sounds, how complicated the name is, etc ...) and social characteristics (i.e. popularity, etc....) factors in our analysis of first name features. First, the *Popularity* of a name was calculated from the respondent's name rank in gender specific Social Security Administrations' 'Top 1,000 popularity list' for the birth decade of the individual.¹¹ The series runs in ascending from 0 to 1, and has the value 0 if the name does not appear on the list, and 1 if it has the top name, namely $1 - ((rank - 1)/1000)$. To account for *unusual spellings* of the name, we create a dummy variable equal to one if the name is spelled unusually, and zero otherwise. A name was determined to be unusually spelled if multiple spellings were pronounced the same way (e.g. Jeffrey and Geoffrey), the more popular name appeared at least four times, and the less popular spelling occurred half or fewer times as the most popular spelling. Also, for each name we calculate the *number of syllables* in a name, coded by following the rules of syllabication.¹² Finally, we code a variable for *nickname*, which is a dummy variable if the name was listed as a nickname by

¹⁰ So as not to introduce correlation between our measured *blackness index* and the error term, the variable is calculated for each individual's name omitting their own name observation in the calculation. To note, the results presented below are very similar if this adjustment is not made. For unique names, the *blackness index* is set equal to the probability that the respondent is black, conditional on having a unique name. This treatment differs from that in Fryer and Levitt (2004), though the findings below do not depend on this data construction decision.

¹¹ The Social Security Administration lists the top 1,000 names by decade and by gender on their website, <http://www.ssa.gov/OACT/babynames/>.

¹² For an example of the rules of syllabication see <http://english.glendale.cc.ca.us/syllables.html>. We also coded the number of letters in a name but do not include it in our analysis as it is very highly correlated (approximately 0.9) with the number of syllables.

Wallace (1992).¹³

We also include a number of additional potentially important linguistic features of names that have currently not been explored and yet may convey information to employers and/or reflect aspects of an individual's upbringing and background. The following dummy variables are defined to be equal to one if their criteria is met and zero otherwise: *ah-ending* if the name ends with an 'ah' sound (e.g. Laura), *oh-ending* if the name ends with an 'oh' sound (e.g. Mario), *ee-ending* if the name ends with an 'ee' sound (e.g. Brittany), *vowel-beginning* if the name begins with a vowel (e.g. Abigail), and *consonant-ending* if the name ends with a consonant (e.g. Meredith). Finally, we adopt *scrabble* as a measure of name's linguistic complexity. It is measured by the simple word score the name would receive from the Scrabble, the popular board game. Taken together, these linguistic feature provide a broader scope for discerning potential information that may be projected from and/or onto a name.

Next, we now turn to a description of the seven lifetime outcomes that we examine in this study. Our measures of lifetime outcomes are a wide-ranging scope of economic and social standing outcomes that have clear economic content and implications and which may be affected by a respondent's first name. First, to measure perceived wealth, we analyze a respondent's *household financial status relative to others*, which is measured on a self reported increasing scale of 1 to 5.¹⁴ Second is *self-reported family income*.¹⁵ Third, we use the *respondent's occupational prestige score* based on 1980 occupational classifications. It is measured on an increasing scale from 17 to

¹³ Our paper differs from Fryer and Levitt (2004) in that we do not have an individual's legal name (e.g. as on his or her birth certificate). Hence, some of our names in our sample could be endogenous variations of an individuals legal name. For this reason we control for nicknames in our estimation results. Moreover, though not shown, the results below are not affected by removing all nicknames from the sample.

¹⁴ Financial relative status' is measured on a self reported scale of 1 to 5, with 1 being that your financial status is well below the mean, 2 being that it is below the mean, 3 being that it is about the mean, 4 being that it is above the mean, and 5 being that it is far above the mean.

¹⁵ Income is reported in discrete intervals in the GSS. We use the midpoints of those intervals and follow the Pareto-extrapolation procedure suggested in Ligon (1994) for the last open-ended category. All incomes are converted into 2002 dollars.

86. Fourth, we use the *respondent's self-reported social class* relative to others. It is measured on a scale of 1 to 4, with 1 being lower class, 2 being working class, 3 being middle class and 4 being upper class. Fifth, since Fryer and Levitt point to the importance of birth outcomes in relation to their black names index, we define *children before 25* (abbreviated as CHILD25 in the tables) to be a dummy variable equal to 1 if an over-25 year old respondent reports having a child before 25, and zero otherwise.¹⁶ This variable is defined for both male and female respondents. Our final two measures of lifetime outcomes are *happiness*, measured on an increasing scale of 1 ('not too happy'), 2 ('pretty happy') or 3 ('very happy'), and *education* which is the number of years of education that the respondent has obtained.^{17,18}

We also have additional explanatory variables to control for observable factors that can systematically influence a respondent's lifetime outcomes. For example, we have data on the respondent's age and gender. We also code the dummy variables *black* if the respondent lists this as their race and *other* is for those who report a race other than white or black. In addition, *mother's education* and *father's education* are the number of years of the respondent's parent's formal education. If either of a respondent's *parents were born outside of the U.S.*, then an indicator is set equal to one, and is zero otherwise (this is abbreviated ABROAD in the tables).¹⁹

¹⁶ We also examined variants of this variable for years less than the 25 years of age cutoff. The results are very similar if this number is decreased. However, due to the small sample size, the precision of this estimate falls as the cut-off is reduced. As our results for this variable generally conform to those in Fryer and Levitt (2004), our choice of 25 years of age does not appear to bias our results.

¹⁷ The actual question is: "Taken all together, how would you say things are these days would you say that you are:" very happy (3), pretty happy(2), or not too happy (1)? Note that the GSS codes happiness so that an increase in the score makes one unhappy. To ease the presentation, we changed the order of the scale.

¹⁸ We also experimented with measuring education by the highest degree obtained. They are not included as the results were broadly similar.

¹⁹ The results below are robust to separately including dummy variables for whether the respondent was born abroad or just their parents were born abroad.

4 The Empirical Analysis

Our empirical analysis section is divided into two main parts. In sub-section 4.1 we provide the simple empirical regularities about first name features. In sub-section 4.2 we provide an empirical analysis of how first name features affect a respondent's lifetime outcomes and how they are influenced by exogenous background factors. In particular, we analyze the following questions: first, to what extent can an individual's background information be inferred from the features of their first name? Second, can lifetime outcomes be explained by an individual's first name features? Third, is the ability of first name features to help predict lifetime outcomes robust to the inclusion of a myriad of exogenous background factors that include information on the respondent's upbringing and their parents?

4.1 Empirical Regularities

Table 1 provides a number of basic statistics for the main data of interest. In particular, we report the mean, standard deviation, median, minimum, maximum and number of observations for each of the 11 *first name features*, the seven measures of *lifetime outcomes*, and the eight *exogenous background factors*.²⁰

The name variables are somewhat unusual and so deserve our initial attention. Table 1 reveals a number of interesting features about the data. First, the average popularity of names in our sample is .61 with a median of .75, suggesting a lot of clustering at the top of the popularity index. Second, the linguistic features of names suggest that, on average, a name has 2 syllables, a scrabble score of 10, with about 6 percent having a non-standard or unusual spelling. The average probability of having a black name is estimated to be 14 percent, suggesting that our method of estimating this feature has some reliability in that it exactly reproduces the sample's probability of being black. Finally, the median value for the *blackness index* is much lower at .09, consistent

²⁰ We also include a respondent's religion, region, and city size at age 16 as additional variables in Tables 4 and 5. The inclusion or exclusion of these variables does not affect our reported results.

with a skewed distribution for the *blackness index*, where some names have very high values for the *blackness index*.²¹

A simple analysis of the raw correlations between our variables confirms that there is significant correlation between first name features.²² Perhaps more interestingly, most first name features are significantly correlated with most exogenous background factors and there are several significant simple correlations between first name features and life outcomes.²³

4.2 Estimation Results

In this sub-section we adopt a standard regression format for disentangling the role that names play in predicting both an individual's background as well as their outcomes. We begin by analyzing the extent to which a respondent's exogenous background factors can be predicted from an individual's first name features. To do so, we regress each of the exogenous background factor on the set of

²¹ Note that, though not shown, the mean value of the *blackness index* among whites is .11 and the standard deviation is .14. Among blacks the corresponding values are .20 and .22.

²² A table of correlations is available from the authors upon request.

²³ First name features are also correlated across time within a family. Though not shown, the GSS data also contains first name information for the other household members. We used this information to form parent child pairs to study intergenerational transmission of the first name features considered here using intergenerational correlations or contingency tables as indicators of intergenerational persistence. The analysis was performed for the whole sample, for the African American sub-sample, and for a matched gender sub-sample (where parent and child were of same gender). The overall result of this analysis is that most first name features are statistically significantly correlated across generations, but that this correlation is a moderate one. The only exception to this moderate correlation is the correlation of the Blackness index across generations in the entire sample (and matched gender sub sample), where the intergenerational correlation is approximately .20. However, the intergenerational correlation in the African American sub sample is only .08, implying that the high correlation in the overall sample is driven by the presence of the African American minority in that sample.

first name features.²⁴ These regression results are presented in Table 2.^{25,26}

The results in Table 2 suggest that first names are significant predictors of a respondent's background. For instance, respondents who are black, as reported in the first column of results, have names that are less popular, more syllables and higher Scrabble scores, fewer 'oh' endings, are more likely to end in a vowel, and have a higher blackness index. This same pattern also exists for those that are neither white nor black (i.e. other), except that they have names with more 'ah' endings and more 'oh' endings, not fewer ones. Also, as indicated in the bottom rows of the table, in both of these equations the R^2 is .04 and the p-value from an F-test that the coefficients on First Name Features are all equal to zero is well below .001. As well, even if the 'blackness index' is excluded from the F-test, the remaining first name features are jointly significant at below the .001 level. The remaining rows indicate a number of key features of the relationship between background factors and first names. First, names indicate a great deal of information on gender and the age of the respondent. In both of the regressions, the R^2 's are relatively large (over .35 for Male).²⁷ Also, a respondent's higher parental education background can be partially inferred from higher popularity, fewer syllables, more standard spellings, more nicknames, fewer 'oh' endings, not starting with a vowel, ending with a consonant and having a lower *blackness Index*. This latter

²⁴ We also estimated the regressions in this analysis by including quadratic terms for the relevant first name features as well as allowing for asymmetric effects by measuring these same variables as the absolute deviation from the median. In general, these effects were not significant and so were excluded from the results reported below.

²⁵ Of course, the dependent variables in many of these regressions are discrete. However, rather than estimate probit or ordered probit specifications, we provide estimates using the simple OLS estimator. To note, when one estimates these empirical models using probit or ordered probit estimators, the results of the linear effects are similar to those presented in the following tables. Moreover, the pattern of statistical significance is largely unaffected by using this alternative estimation strategy. Hence, to keep matters simple we present the OLS estimates below though the alternative estimates are available upon request.

²⁶ The standard errors were estimated allowing for clustering based on the individuals names that occur more than twice and by assigning all unique names to the same cluster.

²⁷ The point we are making here, is that purely linguistic features (non-gender specific) of the names do convey information about the gender. Naturally, most names are gender-specific and thus convey gender information.

result for the *blackness index* is actually quite large: moving from a purely non-black name to a fully black name is associated with a Father having 2 fewer years of formal education and a Mother having 1 year less.

While it has been important to establish that first names confer a great deal of information about a person's background, it is also essential to understand whether any linkages exist between *first name features* and an individual's *lifetime outcomes*. In the remainder of this section, we explore these linkages after controlling for alternative sets of information. This proceed as follows: in Table 3 we regress *lifetime outcomes* on first the *blackness index* alone (columns 1-7) and then on all *first name features*. In Table 4, investigate whether first name features can help predict lifetime outcomes when we also include additional *exogenous background factors* as explanatory variables. Finally, in Table 5, we re-estimate the regression specification in Table 4 separately for males and females. Our aim is to analyze here how the *blackness index* and other *first name features* are associated with these outcome variables, and how much of this association survives the inclusion of a broader set of background information.²⁸

Our first set of results is presented in the first seven columns of Table 3. These results show that, taken alone, the blackness of a name is associated with worse *life outcomes*: less education, occupational prestige, happiness, social class and income, and an increased likelihood of having a child before 25. All these results are highly significant.

In the remaining columns of Table 3 we include all of our *first name features* into the regression. This can be viewed as a robustness check on the previous results: namely, it helps to answer whether it is the *blackness* of the name or the other features of the name that are correlated with its *blackness* that drives the result? Our first result here is that the previous result survives the inclusion of the other *first name features* into the regression: high values of the *blackness index* are associated with worse *life outcomes*. Note that this relationship is attenuated by the inclusion of other *first name features*, though it is statistically significant at conventional levels across the

²⁸While any causal interpretation of our results (e.g. evidence for discrimination based on names) is only suggestive, the predictive interpretation of our results is more robust. These regression results answer the question 'Does knowing someone's name help you to draw an inference about their outcomes?'

columns.

Several additional results can be found in Table 3. First, in all cases, the additional *first name features* are jointly statistically significant at conventional levels. Second, the R^2 's are typically lower in these regressions as compared to those in Table 2 where exogenous background characteristics were the dependent variables. Third, more popular names are generally associated with better lifetime outcomes: that is, more education, occupational prestige and income, and a reduced likelihood of having a child before 25. Also, broadly speaking, names starting with vowels and ending in either an 'ah' or 'oh' sound are related to poorer lifetime outcomes.

In Table 4 we address the question on whether the associations pointed out above are due to the characteristics of the names or whether the name characteristics are just proxies for *exogenous background factors*. In the following discussion, we follow the literature and focus mostly on a name's racial component.

The results demonstrate that the inclusion of these exogenous background factors increase the predictive power of the regression significantly as measured by R^2 . As a group, the *exogenous background factors* are significant in all regressions. Also note that *first name features* other than the *blackness index* lose their significance in most regressions in Table 4, providing further evidence that they are proxies for background factors in Table 3.

Most interestingly, high values of the *blackness index* continue to be associated with several measures of less desirable life outcomes. The sign of the regression coefficient is always consistent with this statement and the result is statistically significant at the 5% significance level for *social class, happiness, having a child before the age of 25, and educational attainment* outcomes.

Table 5 investigates whether the association between the *blackness* of an individual's name with life outcomes could vary across gender and race categories. While it would be preferable to estimate the regressions of Table 4 separately for each category (e.g. black females), unfortunately our sample is too small to estimate these types of sub-samples. Rather, we present modified versions of the regressions in Table 4 – that is, we estimate the regressions separately for males and females and include interaction terms between racial indicators and a name's *blackness index*.

A complicated picture arises from Table 5. First, *first name features* other than the *blackness index* are generally statistically significant explanatory variables in the female sample, but not for the male sample. There is also some statistically significant evidence that having a high *blackness index* is associated with poor life outcomes for females, with one major exception to this: black females with black names have higher family incomes.²⁹ For both white and black females, having a ‘black’ name increases the probability of having a child before the age of 25. Moreover, for females that are black and ‘other’ (i.e. neither black nor white), a black name is associated with a lower social class.

For the male sample, there are fewer significant results for the name variables in Table 5. If we use statistical significance as a guide, generally the name variables are not important explanatory variables for males. The only exception to this pattern is that males of race other than black and white are less happy if they have a name with high *blackness indicator*. This and other negative findings from Table 5 should be interpreted cautiously, given that lack of statistical power due to small sample size could well be part of the reason.

Taken together, Tables 4 and 5 imply that there is evidence that having a black name is generally associated with worse lifetime outcomes even controlling for a myriad of other factors. However, this result is not uniform across all subgroups of population and these differences provide interesting speculative evidence about the economic and sociological forces driving these results.

5 Interpretations and Conclusion

Our research on the economic impact of first names demonstrates three findings. First, the features of an individual’s name reveal a great deal of information their background, even the part of their background that they inherited at birth. Second, the features of an individual’s name have a great deal of predictive power about their lifetime outcomes such as income and social status, educational

²⁹ In addition some statistically insignificant coefficients provide evidence for this exception pattern. For example, black females having a blacker name seem to be happier, though this is not a statistically significant result.

attainment, whether they have a child before 25, etc ... Third, in a great many instances these first names also continue to have predictive power for an individual's lifetime outcomes even when one controls for their exogenous background factors. Indeed, two dimensions that emerge as important determinants of the linkage between first names and lifetime outcomes are race and gender.

Ultimately, there are two potential explanations for the transmission mechanism which links names and lifetime outcomes: namely, either individuals face discrimination or preference based on their first names, or first names are correlated with factors that affect an individual's labor productivity. The former explanation has been favored by Bertrand and Mullainathan (2004) while the latter has been favored by Fryer and Levitt (2004). Overall, our findings provide evidence in favor of the latter. We also demonstrate that gender and race based components of names are independent determinants of less happy individuals, who are less educated, and who have children before 25. These outcomes are likely related to workers that are less productive, *ceteris paribus*. However, these factors do not generally affect income, though they affect the respondent's perceived social standing.

Nevertheless, while we do find evidence of the link between the race and gender qualities of a first name and factors that are related to labor productivity, we believe that our study does reveal some indirect evidence, though in no way proof, of discrimination. In particular, our findings that the blackness of a name has differential impact based on the respondent's race for females indicates that names can have an impact on lifetime financial outcomes that is disproportionate to their relationships to fundamentals. Specifically, the finding that white females and non-black non-white females with names with a high *blackness index* often have lower lifetime outcomes compared to black females potentially demonstrates a more subtle form of discrimination based on identity. Namely, while blacks may face costs and benefits from a blacker name that correspond to the groups they do and don't identify with, whites and non-black non-whites with blacker names would only realize the former and not the latter.

References

- Akerloff, George, and Rachel Kranton, "Economics and Identity," *Quarterly Journal of Economic*, 115 (2000), 715–753.
- Arai, Mahmood, and Peter Skogman Thoursie, "Giving up Foreign Names: An Empirical Examination of Surname Change and Earnings," *Research Papers in Economics* 2006:13, Stockholm University, Department of Economics, 2006.
- Ball, Sheryl, Catherine Eckel, and Adam Ferguson, "What's in a Name? Experimental Evidence on Discrimination," Presentation at Southern Economic Association Meetings, New Orleans, 2004.
- Bertrand, Marianne, and Sendhil Mullainathan, "Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination," *American Economic Review*, 94 (2004), 991–1013.
- Disdier, Anne-Clia, Keith Head, and Thierry Mayer, "Exposure to Foreign Media and Changes in Cultural Traits: Evidence from Naming Patterns in France," Working Paper 213, Centro Studi Luca d'Agliano, University of Milano, 2006.
- Figlio, David, "Names, Expectations and the Black-White Test Score Gap," 2004, mimeo.
- Foster, Gigi, "Names will never hurt me: Racially Distinct Names and Identity in the Undergraduate Classroom," Manuscript, University of South Australia, 2004.
- Fryer, Roland, Jr., and Steven Levitt, "The Causes and Consequences of Distinctively Black Names," *Quarterly Journal of Economics*, 119 (2004), 767–805.
- Lieberson, Stanley, *A Matter of Taste: How Names, Fashions and Culture Change* (Yale University Press, New Haven, CT, 2000).

- Lieberson, Stanley, and Eleanor O. Bell, "Children's First Names: An Empirical Study of Social Taste," *American Journal of Sociology*, 98 (1992), 511–554.
- Lieberson, Stanley, and Kelly Mikelson, "Distinctive African American Names: An Experimental, Historical and Linguistic Analysis of Innovation," *American Sociological Review*, 60 (1995), 928–946.
- Ligon, Ethan, "The Development and Use of a Consistent Income Measure for the General Social Survey," GSS Methodological Report 64, University of Chicago, 1994.
- Oswald, Andrew, "Happiness and Economic Performance," *Economic Journal*, 107 (1997), 1815–1831.
- Stokey, Nancy, "Shirtsleeves to Shirtsleeves: The Economics of Social Mobility," Nancy L. Schwartz Lecture, Northwestern University, 1996.
- Wallace, Carol, *20,001 Names for Baby* (Avon Books, New York, NY, 1992).
- Woodward, Richard, "Do Americans Desire Homogeneity? Evidence from Names from 1900–2000," *Economics Bulletin*, 4 (2005), 1–6.

Table 1: Summary Statistics

First Name Features						
Variable	Mean	Standard Dev.	Median	Min	Max	N
Popularity	0.61	0.37	0.76	0	1	5587
Syllables	2.01	0.74	2.00	0	7	5587
Scrabble	9.95	3.78	10.00	1	29	5587
Unusual	0.06	0.24	0.00	0	1	5587
Nickname	0.15	0.36	0.00	0	1	5587
Ah-ending	0.18	0.39	0.00	0	1	5587
Oh-ending	0.01	0.11	0.00	0	1	5587
Ee-ending	0.20	0.40	0.00	0	1	5587
Vowel-beginning	0.11	0.32	0.00	0	1	5587
Consonant-ending	0.48	0.50	0.00	0	1	5587
Blackness Index	0.14	0.16	0.09	0	1	5587
Exogenous Background Factors						
Variable	Mean	Standard Dev.	Median	Min	Max	N
Black	0.14	0.34	0.00	0	1	5587
Other	0.05	0.22	0.00	0	1	5587
Male	0.44	0.50	0.00	0	1	5587
Mother's Education	11.18	4.12	12.00	0	20	4065
Father's Education	11.25	3.48	12.00	0	20	4806
Siblings	3.67	3.06	3.00	0	35	5568
ABROAD	0.17	0.38	0.00	0	1	5587
Age	46.13	17.21	43.00	18	89	5570
Lifetime Outcomes						
Variable	Mean	Standard Dev.	Median	Min	Max	N
Financial Status	2.89	0.86	3.00	1	5	4227
Family Income	52,676	50,389	37,500	500	219,502	4965
Prestige	43.50	13.66	42.00	17	86	5334
Class	2.47	0.65	2.00	1	4	5552
Happiness	2.17	0.62	2.00	1	3	4261
CHILD25	0.35	0.48	0.00	0	1	1704
Education	13.25	2.97	13.00	0	20	5570

Note: The sample statistics average(Mean), standard deviation (Standard Dev.), median (Median), minimum (Min) and maximum (Max). N is the number of observations. The data were taken from the 1994 and 2002 samples from the General Social Survey (GSS). For definition of variables, see text.

Table 2: Respondent's Background Regressed on First Name Features Only

	Black	Other	Male	Age	Father's Education	Mother's Education	Siblings	ABROAD
Popularity	-0.096 ^c [0.017]	-0.052 ^b [0.024]	0.275 ^c [0.092]	-7.794 ^c [2.668]	1.025 ^c [0.320]	1.192 ^c [0.196]	-0.514 ^c [0.101]	-0.073 ^b [0.037]
Syllables	0.029 ^c [0.009]	0.013 ^b [0.006]	-0.109 ^c [0.033]	0.849 [0.755]	-0.234 ^a [0.131]	-0.234 ^b [0.103]	0.103 [0.067]	0.053 ^c [0.012]
Scrabble	0.003 ^b [0.001]	-0.001 [0.001]	0.003 [0.006]	-0.263 ^a [0.151]	0.015 [0.027]	0.021 [0.021]	-0.010 [0.012]	-0.004 ^a [0.002]
Ah-ending	0.014 [0.020]	0.022 ^a [0.011]	-0.202 ^c [0.059]	-6.566 ^c [1.540]	0.529 ^a [0.285]	0.283 [0.230]	0.019 [0.169]	-0.024 [0.027]
Oh-ending	-0.092 [0.070]	0.269 ^c [0.059]	0.688 ^c [0.083]	-5.944 ^b [2.473]	-1.617 ^a [0.855]	-2.118 ^c [0.565]	0.817 ^b [0.358]	0.413 ^c [0.060]
Ee-ending	-0.011 [0.020]	0.000 [0.009]	-0.085 [0.064]	-1.384 [1.662]	0.396 [0.292]	0.231 [0.235]	0.267 [0.184]	-0.042 ^a [0.022]
Vowel-beginning	0.025 [0.016]	0.014 [0.008]	-0.008 [0.053]	3.021 ^b [1.219]	-0.524 ^b [0.250]	-0.757 ^c [0.997 ^c]	0.311 ^a [0.132]	0.069 ^c [0.024]
Consonant-ending	-0.049 ^c [0.016]	0.003 [0.008]	0.332 ^c [0.017]	-1.299 [1.774]	0.677 ^c [0.305]	0.458 ^c [0.171]	0.016 [0.162]	-0.012 [0.021]
Unusual	-0.045 ^c [0.016]	-0.031 [0.027]	-0.017 [0.045]	-5.584 ^c [1.774]	0.856 ^c [0.305]	0.436 ^b [0.171]	-0.639 ^c [0.162]	0.003 [0.021]
Nickname	-0.040 ^b [0.017]	-0.012 [0.011]	0.210 ^c [0.062]	-5.291 ^c [1.409]	0.497 ^b [0.234]	0.436 ^b [0.178]	-0.203 ^a [0.115]	-0.026 [0.021]
Blackness Index	0.021 [0.022]	0.021 [0.022]	0.179 ^c [0.067]	-1.200 [2.296]	-1.516 ^c [0.513]	-0.732 ^a [0.379]	1.246 ^c [0.343]	-0.016 [0.035]
N	5587	5587	5587	5571	4065	4806	5568	5587
R ²	.03	.04	.36	.05	.03	.05	.02	.05
P: Names	0.000	.000	.000	.000	.000	.000	.000	.000
P: Names ex Blackness	.000	.000	.000	.000	.000	.000	.000	0.000

Note: ^c, ^b, and ^a indicate that the relationship is different from zero at or below the .01, .05 and .10 level of statistical significance, respectively. All regressions include a constant and a dummy variable for the 1994 survey. "P: Names" is the p-value from the F-test that all the coefficients on First Name Features are jointly zero. "P: Names ex Blackness" is the p-value from the F-test that all the coefficients on First Name Features, excluding the Blackness Index, are jointly zero. Standard errors, reported in brackets, were calculated based on clustering by actual first names that appear at least two times or more, with those observed once clustered together.

Table 3: Respondent's Lifetime Outcomes Regressed on First Name Features

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	Financial Status	Log Family Income	Prestige	Class	Happiness	CHILD25	Education	Financial Status	Log Family Income	Prestige	Class	Happiness	CHILD25	Education
Blackness Index	-0.344 ^c [0.078]	-0.413 ^c [0.097]	-4.674 ^c [1.175]	-0.282 ^c [0.057]	-0.224 ^c [0.056]	0.376 ^c [0.082]	-1.562 ^c [0.258]	-0.218 ^c [0.078]	-0.206 ^b [0.089]	-3.487 ^c [1.168]	-0.258 ^c [0.059]	-0.195 ^c [0.060]	0.251 ^c [0.085]	-1.148 ^c [0.259]
Popularity								0.062 ^b [0.031]	0.175 ^c [0.038]	0.686 [0.567]	-0.006 [0.021]	0.008 [0.020]	-0.107 ^c [0.034]	0.393 ^c [0.121]
Syllables								-0.001 [0.022]	-0.053 ^a [0.028]	-0.236 [0.336]	0.016 [0.016]	-0.020 [0.017]	0.033 [0.024]	-0.089 [0.070]
Scrabble								-0.005 [0.003]	-0.003 [0.005]	0.022 [0.050]	-0.003 [0.003]	0.004 ^a [0.002]	-0.008 ^a [0.004]	-0.001 [0.013]
Ah-ending								-0.071 [0.048]	-0.085 [0.074]	-1.476 ^b [0.640]	-0.115 ^c [0.035]	-0.027 [0.030]	0.010 [0.060]	-0.017 [0.140]
Oh-ending								0.101 [0.111]	0.346 ^c [0.104]	-3.100 ^c [1.143]	-0.094 [0.088]	0.179 ^b [0.081]	-0.025 [0.113]	-0.423 [0.423]
Ee-ending								0.023 [0.042]	-0.008 [0.061]	-1.177 ^a [0.627]	-0.019 [0.033]	0.027 [0.031]	0.062 [0.050]	-0.119 [0.141]
Vowel-beginning								-0.065 ^a [0.038]	-0.138 ^c [0.048]	-0.277 [0.751]	-0.020 [0.030]	-0.036 [0.026]	-0.025 [0.036]	-0.328 ^b [0.146]
Consonant-ending								0.111 ^c [0.040]	0.075 [0.055]	0.362 [0.504]	-0.011 [0.024]	0.013 [0.025]	-0.039 [0.048]	0.188 [0.122]
Unusual								-0.029 [0.065]	-0.053 [0.074]	0.675 [0.748]	-0.014 [0.044]	-0.018 [0.030]	-0.005 [0.039]	0.301 ^a [0.167]
Nickname								0.072 ^a [0.040]	0.105 ^b [0.049]	-0.611 [0.626]	0.004 [0.029]	0.008 [0.028]	-0.084 ^b [0.035]	0.036 [0.119]
N	4227	4965	5334	5552	4261	1704	5570	4227	4965	5334	5552	4261	1704	5570
R ²	.01	.00	.00	.01	.00	.01	.01	.01	.02	.01	.01	.01	.04	.02
P-Names								.000	.000	.000	.000	.000	.000	.000
P-Names ex Blackness								.000	.000	.000	.131	.024	.000	.000

Note: See Table 2.

Table 4: Respondent’s Lifetime Outcomes Regressed on First Name Features and All Background Information

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Financial Status	Log Family Income	Prestige	Class	Happiness	CHILD25	Education
Blackness Index	−0.089 [0.102]	−0.023 [0.098]	−1.482 [1.424]	−0.159 ^b [0.068]	−0.192 ^b [0.085]	.254 ^b [0.107]	−0.548 ^b [0.263]
Male	0.187 ^c [0.039]	0.150 ^c [0.037]	−0.568 [0.570]	−0.031 [0.023]	0.026 [0.028]	−0.156 ^c [0.029]	0.012 [0.093]
Age	0.031 ^c [0.005]	0.097 ^c [0.007]	0.810 ^c [0.068]	0.005 [0.003]	0.002 [0.005]	0.026 ^c [0.006]	0.172 ^c [0.020]
Age ²	−0.024 ^c [0.005]	−0.094 ^c [0.006]	−0.667 ^c [0.069]	0.003 [0.004]	0 [0.005]	−0.026 ^c [0.005]	−0.161 ^c [0.020]
Black	−0.229 ^c [0.067]	−0.334 ^c [0.074]	−1.315 ^a [0.735]	−0.137 ^c [0.041]	−0.195 ^c [0.042]	0.014 [0.051]	0.068 [0.146]
Other	−0.177 ^b [0.080]	−0.054 [0.091]	1.958 ^a [1.056]	−0.057 [0.043]	−0.175 ^c [0.055]	0.010 [0.082]	0.558 ^c [0.196]
Father’s Education	0.028 ^c [0.005]	0.026 ^c [0.006]	0.599 ^c [0.070]	0.028 ^c [0.003]	−0.002 [0.003]	−0.010 ^b [0.004]	0.181 ^c [0.012]
Mother’s Education	0.019 ^c [0.006]	0.026 ^c [0.006]	0.348 ^c [0.096]	0.011 ^c [0.004]	0.009 ^b [0.005]	−0.008 ^a [0.005]	0.151 ^c [0.017]
Siblings	−0.010 [0.006]	−0.022 ^c [0.006]	−0.333 ^c [0.078]	−0.012 ^c [0.004]	0.000 [0.005]	0.011 ^b [0.005]	−0.128 ^c [0.016]
ABROAD	0.076 [0.050]	0.191 ^c [0.051]	0.805 [0.897]	0.079 ^c [0.030]	−0.009 [0.037]	−0.079 ^a [0.043]	0.310 ^b [0.139]
N	2866	3433	3661	3773	2880	1157	3778
R ²	.10	.16	.12	.12	.03	.13	.28
P-Names ex Blackness	.528	.073	.266	.004	.066	.417	.952
P-Age & Gender	.000	.000	.000	.000	.067	.000	.000
P-Race	.001	.000	.032	.002	.000	.959	.017
P-Background	.000	.000	.000	.000	.039	.000	.000

Note: See Table 2. Individual regression coefficients for first name features other than Blackness Index are suppressed in the table. Similarly individual regression coefficients for several background factors were suppressed. “P-Age & Gender”, “P-Race”, “P-Background” are the p-values from F-tests testing whether the coefficients for age and gender variables, race indicators and all other background factors (excluding age, gender and race) respectively are all jointly zero.

Table 5: Respondent's Lifetime Outcomes Regressed on First Name Features and All Background Information By Gender

	Males							Females						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Financial Status	Log Family Income	Prestige	Class	Happiness	CHILDS	Education	Financial Status	Log Family Income	Prestige	Class	Happiness	CHILDS	Education	
Blackness indicator interacted with:														
White	0.251 [0.188]	0.268 [0.212]	-0.724 [2.683]	-0.117 [0.118]	-0.162 [0.160]	0.084 [0.231]	0.234 [0.441]	-0.258 [0.166]	-0.272 ^b [0.129]	-3.914 ^a [2.241]	-0.084 [0.107]	-0.270 ^b [0.133]	0.307 ^a [0.156]	-1.181 ^c [0.437]
Black	-0.318 [0.237]	0.055 [0.233]	1.352 [5.271]	-0.171 [0.199]	-0.318 [0.292]	0.472 [0.434]	0.616 [0.903]	0.339 [0.299]	0.489 ^a [0.275]	3.525 [3.699]	-0.433 ^b [0.199]	0.243 [0.275]	0.419 ^b [0.203]	-1.016 [0.830]
Other	-0.674 [0.478]	0.119 [0.474]	1.984 [5.133]	0.089 [0.254]	-0.618 ^a [0.326]	-1.228 [1.365]	0.763 [1.038]	-1.368 ^b [0.584]	-1.311 ^a [0.706]	-5.617 [6.701]	-0.684 ^a [0.353]	-0.914 ^c [0.326]	-1.252 [0.910]	-0.103 [1.181]
Age	0.032 ^c [0.007]	0.083 ^c [0.009]	0.879 ^c [0.113]	0.003 [0.005]	-0.001 [0.006]	0.022 ^b [0.009]	0.189 ^c [0.023]	0.032 ^c [0.007]	0.110 ^c [0.008]	0.775 ^c [0.096]	0.008 ^a [0.004]	0.006 [0.006]	0.027 ^c [0.007]	0.159 ^c [0.023]
Age ²	-0.028 ^c [0.008]	-0.081 ^c [0.009]	-0.701 ^c [0.116]	0.004 [0.005]	0.002 [0.007]	-0.024 ^c [0.009]	-0.176 ^c [0.024]	-0.023 ^c [0.007]	-0.107 ^c [0.008]	-0.667 ^c [0.096]	0.001 [0.004]	-0.003 [0.006]	-0.026 ^c [0.006]	-0.151 ^c [0.025]
Black	-0.144 [0.110]	-0.201 ^b [0.091]	-4.453 ^c [1.618]	-0.137 ^a [0.070]	-0.145 [0.097]	-0.015 [0.107]	-0.413 [0.310]	-0.334 ^c [0.106]	-0.577 ^c [0.133]	-0.748 [1.345]	-0.059 [0.072]	-0.291 ^c [0.071]	-0.051 [0.097]	0.324 [0.231]
Other	-0.102 [0.162]	-0.129 [0.121]	0.453 [2.230]	-0.122 [0.005]	-0.178 ^a [0.103]	0.217 [0.261]	0.097 [0.368]	0.032 [0.152]	0.196 [0.168]	3.044 ^a [1.822]	0.058 [0.084]	0.009 [0.101]	0.233 [0.202]	0.625 ^a [0.376]
Father's Education	0.027 ^c [0.008]	0.022 ^b [0.009]	0.607 ^c [0.120]	0.019 ^c [0.007]	-0.003 [0.009]	0.001 [0.006]	0.184 ^c [0.029]	0.027 ^c [0.007]	0.027 ^c [0.008]	0.557 ^c [0.090]	0.034 ^c [0.005]	-0.002 [0.004]	-0.017 ^b [0.007]	0.171 ^c [0.017]
Mother's Education	0.023 ^b [0.009]	0.023 ^b [0.010]	0.240 [0.183]	0.019 ^c [0.007]	0.006 [0.009]	-0.016 ^b [0.006]	0.124 ^c [0.029]	0.017 ^b [0.007]	0.031 ^c [0.008]	0.423 ^c [0.095]	0.005 [0.005]	0.011 ^a [0.006]	-0.001 [0.007]	0.171 ^c [0.020]
Siblings	-0.010 [0.008]	-0.023 ^b [0.010]	-0.173 [0.123]	-0.014 ^b [0.006]	0.008 [0.008]	0.000 [0.009]	-0.103 ^c [0.024]	-0.010 [0.009]	-0.022 ^b [0.010]	-0.489 ^c [0.103]	-0.010 ^a [0.005]	-0.007 [0.006]	0.021 ^c [0.007]	-0.158 ^c [0.022]
ABROAD	0.029 [0.077]	0.181 ^b [0.074]	1.160 [1.133]	0.054 [0.042]	0.056 [0.052]	-0.061 [0.056]	0.404 ^a [0.208]	0.111 [0.068]	0.201 ^c [0.073]	0.546 [1.168]	0.106 ^b [0.042]	-0.068 [0.053]	-0.092 [0.063]	0.201 [0.187]
N	1307	1567	1665	1691	1315	514	1696	1559	1866	1996	2082	1565	643	2082
R ²	.11	.14	.16	.13	.04	.12	.28	.11	.20	.12	.14	.05	.15	.30
P-Names	.357	.307	.284	.083	.703	.462	.765	.029	.000	.002	.000	.000	.692	.000
P-Age	.000	.000	.000	.000	.216	.018	.000	.000	.000	.000	.000	.000	.001	.000
P-Race	.155	.635	.961	.541	.120	.513	.742	.034	.008	.185	.030	.007	.022	.023
P-Back.	.000	.000	.000	.000	.003	.000	.000	.000	.000	.000	.000	.054	.000	.000

Note: See Table 4. The results in columns (1)-(7) are for males, while the results for females are reported in columns (8)-(14). White is the indicator variable for white race. "P-Names" and "P-Back." are shorthand for "P-Names excluding *blackness indicator* and its interactions" and "P-Background variables".