Speech and Wages

Jeffrey Grogger

University of Chicago Harris School of Public Policy
1155 E. 60th Street
Chicago, IL 60637
jgrogger@uchicago.edu
February 2, 2018
Version 8

Abstract: Speech is a fundamentally human phenomenon that has been widely studied. Nevertheless, relatively little is known about how a worker's speech, in his native language, is related to his wages, or what explains the observed relationship. To address these questions I collected audio data, which I coded and merged to respondents from the NLSY97. Wages are strongly associated with speech patterns among both African Americans and Southern whites. For Southern whites, this is largely explained by family background and residential location. For blacks, it is explained by sorting: workers whose speech is racially indistinct sort toward occupations that involve intensive interpersonal interactions and earn a sizeable wage premium in those jobs. They do not sort toward other types of skill-demanding jobs.

For comments and assistance of many different kinds, the author thanks two anonymous reviewers, Kerwin Charles, Amy Claessens, Holly Craig, Dan Hamermesh, Magne Mogstad, Derek Neal, Brett Naul, Charles Pierret, Chenyu Qiu, Chuck Safiran, Ursina Schaede, Nola du Toit, Brandon Waldon, seminar participants at the University of Chicago, the University of Zurich, the Center for European Economic Research, and especially Dan Black. This work has been supported by the National Institute of Child Health and Human Development and the Russell Sage Foundation. I am grateful for their support but am solely responsible for the opinions and conclusions expressed herein.

I. Introduction

Language is a fundamentally human phenomenon that separates us from other forms of life. It has been studied in such detail that linguists can explain why we say "razzle dazzle" instead of "dazzle razzle" (Pinker 1999). Nevertheless, we know relatively little about two important questions. The first is how a worker's speech, in his native language, is related to his wages. The second is what explains the observed relationship.

To address these issues I make use of the fact that different speakers of the same language may speak different dialects. In the United States, many African Americans speak African American Vernacular English (AAVE) and many Southerners speak Southern American English (SoAE). Both of these dialects are widely recognized and differ from Standard American English (SAE) in ways that are well understood (Nagle and Sanders 2003).

Evidence that speech is related to wages comes from two recent studies. Rickford et al (2015) found a significant negative correlation between earnings and a measure of AAVE usage among adults in the Moving to Opportunity (MTO) study. Grogger (2011) showed that black workers who spoke mainstream English, rather than a racially distinctive dialect, had wages similar to those of equally skilled whites.

This paper extends my previous work in several ways. First, working with a much larger sample than before, I show that the same basic patterns between speech and wages are robust to different characterizations of speech. I then show that, for Southern whites, speech-related wage differences are largely an urban-rural phenomenon. In

contrast, among African Americans, I show that the wage premium for mainstream speech is robust to controls for a large number of variables that one might expect to explain the relationship between speech and wages.

Moving well beyond my earlier work, I see to explain the mainstream speech premium for blacks in terms of a model of occupational sorting. An occupational sorting model posits that there is one set of occupations in which a worker trait, such as mainstream speech, is particularly productive, and another in which it is not. Workers who possess the trait sort disproportionately into the productive sector and earn a wage premium there.

One reason for differential productivity between sectors may be consumer discrimination. In their model of worker beauty, Hamermesh and Biddle (1994) argued that beauty may be productive in jobs involving customer contact due to preferences among consumers for interacting with more attractive people. If consumers are willing to pay to interact with more attractive employees, employers should be willing to pay such workers higher wages. One could similarly imagine a consumer taste for mainstream speech. Indeed, the social psychology literature, discussed in more detail below, shows that listeners express strong views about the speech of others.

Deming (2015) provides another reason why speech may be productive. In his model, there is one sector that involves extensive interpersonal interaction in which workers effectively trade skills with one another to produce output. In that sector, social skills reduce the cost of trade, making workers more productive and raising their pay. If mainstream speech reflects a social skill, one would expect it to be productive in

occupations that involve extensive interpersonal interactions. The idea that speech patterns may reflect social skill gets further support from Fan, et al. (2015), who find that bilingual children are more effective communicators because they more readily perceive the point of view of their speaking partner. Likewise, Lang (1986) provides numerous examples of communications breakdowns that arise between the individuals who know the grammar and vocabulary, but not the social cues, of their interlocutor.

This paper contributes to a number of literatures within economics. One is discrimination, in terms of both race and other traits such as beauty, height, and body weight. (Hamermesh and Biddle 1994; Biddle and Hamermesh 1998, Persico, et al, 2004; Cawley 2004). Another is the burgeoning literature on non-cognitive skills and their link to labor market outcomes (Borghans, et al. 2008). Since speech has been linked to social identity (Trudgill 1972, Labov 1966; Fordham and Ogbu 1986; Baugh 1992; Pattillo 1999; Schneider 2003), the paper also relates to work that analyzes how identity influences behavior (Akerlof and Kranton 2000; Austin-Smith and Fryer 2004; Kim and Loury 2012; Bertrand and Mullainathan 2004, Fryer and Levitt 2004, Fryer and Torelli 2005).

Speech has been shown to affect other aspects of economic behavior as well. Audit studies indicate that landlords are more likely to show apartments to renters who sound white over the phone (Purnell, Idsardi, and Baugh 1999; Massey and Lundy 2001). Falck et al (2012) show that migration is higher between areas with more similar dialects, holding distance constant. Chen (2013) argues that grammatical differences across languages can explain differences in savings behavior between their speakers.

Also related to this study is work on language and the success of immigrants. Immigrants generally fare better in the labor market of their host country, the better they speak that country's language (McManus, Gould, and Welch 1983; Chiswick 1991; Trejo 1997, and Bleakley and Chin 2004; Chiswick and Miller 2001). This is consistent with Lang's (1986) model, which predicts that bilingual immigrants should be compensated in a competitive market for the cost of acquiring the host-country language.

Yet the link between speech and wages may be very different for immigrants speaking a second language than for non-migrant workers speaking their native tongue. The target language skills of immigrants vary from rudimentary to near-native. At the bottom of this spectrum, the speaker may struggle to convey basic meaning. In this study, in contrast, the speakers are generally native English speakers, whose basic meaning is clear to all other native English speakers. Thus the question is not about basic meaning, but rather about other reactions by listeners to non-mainstream dialects.

In the next section of the paper I provide some background on speech and listener responses to it. Section III discusses collection and processing of the speech data. Section IV presents the main regression results and a number of robustness checks. In Section V I present the model of occupational sorting. Section VI takes up the question of whether speech adjusts over time and section VII concludes.

II. Background

The linguistics literature addresses two questions that are important for the analysis to follow. The first is, where does variation in speech come from? The second

is, what influences the way we speak our native language? A broader social science literature analyzes listeners' responses to speakers' voices.

A. Variation in speech

As mentioned above, variation in speech comes from the fact that different speakers of the same language speak different dialects. Like the mainstream standard to which they are related, dialects follow rules implicitly known by all speakers of the dialect. What distinguishes dialects from the standard (and each other) is that some of the rules are different.

The rules in question can involve different aspects of speech, including syntax (e.g., negation rules), morphology (e.g., subject-verb agreement), and phonology (e.g., the way vowels are produced). Furthermore, rules may be variably applied. Few speakers apply only the rules of their native dialect. Most incorporate both dialect and mainstream rules at different times (Labov 1972).

Linguists have shown that roughly 40 features differentiate AAVE from SAE (Bailey and Thomas 1998; Clopper and Pisoni 2004; Green 1998; Labov 1972; Martin and Wolfram 1998; Mufwene 1998; Rickford and Rafal 1996; Thomas and Reaser 2004; Walton and Orlikoff 1994; Washington and Craig 2002; Wolfram 1969, 1991). Thomas and Reaser (2004) conclude that listeners rely not only on these features, but on the speaker's vowel quality, voice quality, cadence, and intonation, to identify race on the basis of speech. Importantly for the data collection effort described below, linguists have shown that listeners can identify the race of a speaker based on very short audio clips (Thomas 2002; Thomas and Reaser 2004).

B. Dialect acquisition

The literature on second language learning shows evidence of a "sensitive period" for native dialect acquisition. Before the sensitive period ends, children are capable of acquiring native-sounding speech in whatever language they are exposed to. Once the sensitive period ends, it is more difficult to acquire a native-sounding accent in a second language.

There is some debate as to when the sensitive period ends, but a fair amount of agreement that it is over before puberty concludes (Johnson and Newport 1989; Hyltenstam and Abrahamsson 2003; Granena and Long 2013). There is some evidence that different aspects of language acquisition have different sensitive periods. For example, the sensitive period for the acquisition of native-sounding phonology may end as early as age seven or eight, whereas the sensitive period for morphology may extend into the teens (Granena and Long 2013; Siegel 2010). The evidence indicates further that one tends to acquire one's native accent from one's linguistic peers during the sensitive period, rather than from one's parents or other sources such as broadcast media (Labov 1972). Second dialect acquisition is similar to second language acquisition, in that it is difficult to acquire a native-sounding accent in a second dialect after the sensitive period ends (Siegel 2010).

C. Listeners' reaction to speakers' voices

Listeners react strongly to speech. Social psychologists have shown that both black and white listeners routinely rate AAVE speakers lower than SAE speakers in terms of socioeconomic status, intelligence, and even personal attractiveness (Bleile,

McGowan and Barnthal 1997; Doss and Gross 1992, 1994; Johnson and Buttny 1982; Koch, Gross, and Kolts 2001; Rodriguez, Cargile, and Rich 2004). This evidence seems consistent with the notion that people have preferences over the speech of others. This is important for the analytical approach adopted here, since consumer and coworker preferences underlie the occupational sorting model.

III. Data

A. General information

Data come from the 1997 cohort of the National Longitudinal Survey of Youth (NLSY97), which is a large, nationally representative panel survey of the labor market behavior of youths who were aged 12 to 16 in 1997. NLSY interviews cover topics such as schooling, employment, earnings, sources of income, and crime, among others. The response rates for recent interviews have been roughly 80 percent. About 85 percent of the interviews are carried out in person.

B. Audio data collection

This section covers essential information about the collection and processing of the audio data. I refer readers interested in more detail to the Data Appendix.

Audio data were collected during Round 15 of the NLSY97, which was fielded between September 2011 and June 2012. The data were collected in response to two speech prompts, which were designed to capture both formal and informal speech. The prompt for formal speech involved a job-search (JS) role-playing exercise that was administered during the employment section of the interview. The prompt for informal

speech was administered at the end of the interview, when respondents were asked to recount their happiest moment (HM) since their last interview.

All respondents who conducted in-person interviews and gave consent to be recorded were eligible to be assigned at least one speech prompt. For budgetary reasons, not all respondents were assigned both prompts. African-American respondents were assigned both prompts, since they were the primary focus of the study. Since Southern white dialects share some of the features of AAVE, Southern whites were also assigned both prompts. I define Southern whites as non-Hispanic whites who resided in the South Census region at age 12. Non-southern whites were randomly assigned to only one speech prompt, since few were expected to produce AAVE features.[1]

Among black respondents who were interviewed in-person, 83 percent gave consent to be recorded. The share was 80 percent for whites, both Southern and non-Southern. Appendix Table 1 shows that 6,080 respondents were interviewed in person and gave consent.

For reasons still undetermined, audio files were obtained from only 4,907 of these respondents. Although this represents a substantial loss, loss rates varied little by race or region of origin. Some of the files obtained proved to be empty or inaudible, and for budgetary reasons I limited the processing of files from respondents who were neither black nor white. As a result, speech data are available for 4,225 respondents.

To check whether respondents with speech data are similar to those without, I tabulated respondent characteristics that appear in the regressions below. The results

[1] A few hundred other, mostly non-white, non-black respondents, were assigned both speech prompts for purposes of a separate study.

appear in Appendix Table 2. Respondents with speech data were similar in many ways to those without. Blacks with speech data were significantly less likely to have missing location data, but more likely to have grown up with two parents and to have missing maternal education data.

C. Producing numerical data from the audio files

To generate data that could be used in a regression analysis, I recruited anonymous listeners to listen to the audio files in a secure environment and answer questions about the speakers.[2] After listening to each audio file, listeners were asked to specify the speaker's sex, race/ethnicity, and region of origin. Three listeners were assigned to each audio file. Thus speakers who responded to both the HM and JS prompts have six listener reports, whereas speakers who responded to only one of the prompts have three.

Table 1 reports the distribution of listener reports that the speaker was black (first two columns) or Southern (last two columns). This table is restricted to males, as is the rest of the analysis below. The vast majority of blacks and Southern whites responded to both speech prompts and thus had six listener reports (tabulations for those who responded to only one prompt appear in Appendix Table 4). Thus zero to six listeners could have reported the speaker to be black, and likewise, zero to six listeners could have reported the speaker to be Southern. Since non-Southern whites responded to only one speech prompt, the corresponding range for them is zero to three.

[2] Listeners were employees of NORC, the firm that conducts the NLSY97 interviews. Listener demographic characteristics are summarized in Appendix Table 3. Below I show that they do not affect the regression results.

Column (2) shows that at least one listener reported all but 5.3 percent of the black speakers to be black. At the same time, all six listeners concurred that 32.6 percent of the black speakers were black. Seventy-one percent of the black speakers were perceived to be black by at least four listeners.

Moving down the column, we see that 39.8 (=100-60.2) percent of Southern white speakers were reported to be black by at least one listener. Considering the similarities between AAVE and SoAE as discussed above, this level of confusion is not surprising. There is considerably less such confusion when it comes to non-Southern whites, of whom only 11.7 percent were reported to be black by any of the listeners.

Column (4) shows the distribution of listeners reporting that the speaker was Southern. For 15 percent of the blacks, none of the listeners perceived the speaker to be Southern. All six listeners reported 9.7 percent of the black speakers to be Southern. The distribution between these extremes was roughly constant. About 58 percent of blacks lived in the South at age 12, so the rough uniformity of listener reports that black speakers sound Southern may have to do with similarities between AAVE and SoAE.

The distribution for Southern whites shows that 76.2 percent were reported to be Southern by at least one listener. At the same time, 32.2 percent of Southern whites were reported to be Southern by four or more listeners. This compares well to findings by Baily and Tillery (1996), who report that 32 percent of Southern residents sound "strongly Southern."

Although non-Southern whites have only three listener reports, the data show nonetheless that the distribution of Southern-sounding speech is very different between

the two groups. To verify this, I randomly sampled three of each Southerner's six listener reports. This showed that only about 35 percent of Southerners were not reported to be Southern by any of the listeners. This is less than half of the 73.9 percent of non-Southerners who were not reported to be Southern by any of the listeners.

E. Speech Patterns and Worker Characteristics

I conclude this section by presenting data in Table 2 on human capital, family income, residential location and wages by the worker's speech pattern. Human capital measures include years of education and Armed Forces Qualifying Test (AFQT) scores. Years of education are the highest level of schooling observed by 2013, when the NLSY97 respondents were 28 to 32 years old. The AFQT is a test administered by the Department of Defense for the purpose of screening recruits. It was taken by NLSY97 respondents in 1997. It has been interpreted as a measure of ability or pre-market skills in a number of previous wage studies (Altonji and Pierret 2001; Cameron and Heckman 1993; Farber and Gibbons 1996; Neal and Johnson 1996).[3] Family income was collected from respondents' parents in 1997. Residential location may vary from year to year, as do wages.

There is a strong relationship between speech patterns, education, and test scores. Among blacks, there is a difference of almost three years of education, and nine-tenths of a standard deviation of AFQT score, between the extremes of the speech-pattern distribution. A similar pattern appears among Southern whites. Among both groups,

[3] Throughout this study I use the AFQT scores provided by Altonji, Bharadwaj, and Lange (2009). Although these scores were originally constructed to match the distribution of AFQT scores from a previous cohort, preliminary work showed them to have a more significant effect on wages than the percentile scores distributed with the NLSY97 public-use files. I have standardized these scores to have mean zero and standard deviation one in the full sample.

family income tends to be higher, the fewer the number of listeners who reported the speaker to be black or Southern. For blacks, there is a slight tendency for more workers with distinctive speech to reside in rural areas. For Southern whites, this tendency is quite strong. Likewise, although there is variability due to the small sample sizes, wages tend to be lower for workers with more distinctive speech. A key issue in the next section will be distinguishing whether the link between speech and wages persists after accounting for observable characteristics of the worker.

V. Regression analysis of wages

A. The regression model

The first goal of this section is to estimate speech-related wage differences in the context of a Mincer-type wage regression. I estimate regressions where the log of the hourly wage is the dependent variable, restricting attention to workers who are either black or white in order to focus on those who are likely to be native English speakers.[4] I restrict attention to males to avoid the sample selection issues that arise among women of prime childbearing age. I further limit the sample to workers who have spent at least two consecutive years out of school, in order to focus on those whose primary activity is likely to be employment, rather than education.[5] I pool data over the years 2005-2013, which means that the mean age of workers in the estimation sample is 27.

[4] The NLSY97 has questions about languages spoken by other people in the respondent's childhood home, but no questions that ask directly what languages are spoken by the respondent.

[5] I also eliminate observations of wages that are less than $1/hour or greater than $60/hour, where wages are expressed in terms of 2008 dollars. I also drop wage observations during periods when the worker reports himself to be enrolled or self-employed and when his potential experience is negative.

The key explanatory variables in the regression are dummy variables for race/region at age 12 (i.e., black and Southern white dummies, where non-Southern whites constitute the omitted category) and variables capturing the worker's speech pattern. Also included are a basic set of human capital variables, including dummy variables for educational attainment, experience, and experience squared. The basic regressor set also includes dummies for whether the worker currently resides in the South, currently resides in an urban area, and is currently married.[6]

The first column of Table 3 reports for comparison purposes estimates from a wage equation that does not include any measures of speech. The coefficient on the black dummy is -0.113 and significant. The coefficient indicating that the worker resided in the South at age 12 is 0.032 and insignificant. The other coefficients are largely as one might expect.

B. Alternative representations of speech

An important question is how to represent workers' speech patterns in a regression model. Table 3 shows results from a number of different specifications. The first, in column (2), is the simplest. For workers with six listener reports, the second row reports the coefficient on the number of listener reports that the worker is black. For workers who responded to only one of the speech prompts, and who therefore have only three listener reports, this variable equals two times the number of actual listener reports that the speaker was black. The coefficient is -0.037 and statistically significant,

[6] In this and all other regressions reported below, in addition to variables shown, the regressions include year dummies and missing value flags for region, educational attainment, and urbanicity. Missing value flags equal one when the corresponding variable is missing and equal zero otherwise. Missing values of the corresponding variable are recoded to zero.

indicating that the wage of a worker with six such reports on average is 18.5 percent lower than that of a worker with one such report. The coefficient on the black dummy indicates that black workers not reported to be black by any of the listeners have wages 3.1 percent less than those of similarly skilled whites.

In the 13th row of the table is the coefficient on the number of listener reports that the worker was Southern.[7] This coefficient is -0.046 and statistically significant. Workers reported to be Southern by six listeners earn wages roughly 23 percent lower than workers reported to be Southern by only one listener. Including the speech variable raises the South-at-age-12 coefficient a bit, but it remains insignificant.

Column (3) replaces the simple speech-pattern variables above with analogous variables designed to check whether listener characteristics affect the results. To construct it, I estimated two regressions for which the unit of observation was the speaker-listener pair. In the first, the dependent variable was equal to one if the listener reported the speaker to be black and was equal to zero otherwise. The explanatory variables were dummy variables for each listener and a dummy equal to one for responses to the job-search prompt. The second regression was similar, except the dependent variable equaled one if the listener reported the speaker to be Southern and equaled zero otherwise. I then averaged the residuals for each speaker and standardized them to have mean zero and standard deviation one. This makes their coefficients comparable to those of the number of listener reports, since the number of listener reports has a standard deviation close to one.

[7] As above, each report for workers who responded to only one speech prompt was treated as if it represented two reports.

These residualized listener reports were included in the regression that appears in column (3). Purging the speech variables of listener characteristics in this way has no effect on the estimates. Presumably, aggregating over multiple listener reports already neutralizes the effect of any individual listener's characteristics, so that explicitly eliminating the effects of listener characteristics has little bearing on the regression results.

The regressions discussed so far impose linearity, whose main virtue is simplicity. The regression in column (4) relaxes that constraint. Here I replace the previous speech measures with two sets of dummies. One set contains one dummy variable for each possible number of listener reports that the speaker is black. The other contains one dummy variable for each possible number of listener reports that the speaker is Southern. In both cases, speakers with no such listener reports constitute the omitted group. As above, for speakers who responded to only one of the speech prompts, I classify the speakers as if they had twice the number of actual listener reports indicating that they were black or Southern.

The estimates show a strong and somewhat non-linear relationship between speech patterns and wages. The coefficients for the dummies indicating that one, two, or three listeners reported the speaker to be black have mixed signs and are insignificant. The coefficients for the dummies indicating that four, five, or six listeners reported the speaker to be black are all negative and sizeable and two are statistically significant. The F-statistic for the joint significance of all six dummies is 3.59, with a p-value of 0.002.

Turning to the estimates for Southern speech patterns, all six coefficients are negative. The coefficients for the dummies indicating that one, two, or three listeners reported the speaker to be Southern are insignificant; those indicating that four, five, or six listeners reported the speaker to be Southern are all sizeable and statistically significant. The F-statistic for the joint significance of all six dummies is 2.5, with a p-value of 0.022.

The results from this flexible specification show a strong and non-linear relationship between wages, black-sounding speech, and Southern-sounding speech. At the same time, the large number of coefficients makes the specification unwieldy. To ease the discussion to follow, I adopt a simplified specification that nonetheless preserves the main dimension of the non-linearity.

To capture mainstream speech among African Americans, I dichotomize the number of listener reports that the speaker was black. I construct a mainstream-speech dummy that is equal to one for speakers with fewer than four such reports and equal to zero otherwise. By this measure, 29 percent of blacks have mainstream speech. I interact this dummy with the black dummy, and refer to the interaction as "black * mainstream speech". For white speakers (regardless of region at age 12), I construct a complementary dummy that is equal to one for speakers with four or more listener reports that the speaker was black and equal to zero otherwise. I refer to this variable as "white * black speech."

I construct similar variables to capture Southern-sounding speech. One is a mainstream speech dummy that is equal to one for speakers with fewer than four listener

reports that the speaker is Southern. By this measure, 68 percent of Southerners have mainstream speech. I interact this dummy with the South-at-age-12 dummy, and refer to the interaction as "South * mainstream speech". For speakers who did not reside in the South at age 12 (regardless of race), I construct a complementary dummy that is equal to one for speakers with four or more listener reports that the speaker was Southern and equal to zero otherwise. I refer to this variable as "non-South * Southern speech."[8]

Regression results that make use of these variables appear in column (5). The black coefficient now measures the wage gap between blacks whose speech is distinctively black and non-Southern whites who sound neither black nor Southern. The coefficient is -0.152 and significant. Black workers with mainstream speech earn 13.6 percent more than blacks with distinctively black speech. Adding these coefficients shows that black workers with mainstream speech patterns earn 1.6 percent (standard error =5.1 percent) less than comparably skilled whites, on average. The results also show that whites whose speech is perceived as black earn 16.2 percent less than whites with mainstream speech.[9]

The South-at-age-12 coefficient now measures the difference in log wages between workers who resided in the south at age 12 and have Southern-sounding speech patterns and non-Southern white workers who sound neither black nor Southern. The coefficient is -0.047 but insignificant. However, the South*mainstream speech coefficient indicates that Southerners with mainstream speech patterns earn 8.6 percent

[8] These mainstream speech dummies are similar to the variables I constructed in Grogger (2011). There, I had five listener reports for each speaker and labeled the speakers as mainstream if three or fewer indicated that the speaker was black (or Southern).

[9] There are 21 such workers in the sample. Nine of them are Southern, which may indicate that features common to AAVE and SoAE are leading some Southern whites to be categorized as black.

more than Southerners with distinctively Southern speech. At the same time, non-Southerners who are perceived as sounding Southern earn wages 8.7 percent lower than non-Southerners with mainstream speech.[10]

C. Do background characteristics explain the speech-related wage gaps?

A natural question is whether these regression estimates reflect the effect of speech per se, or whether they are also picking up other characteristics of workers. The discussion in Section II suggests that it is probably impossible to isolate the effect of speech completely. Children acquire their native dialects before puberty, largely from their linguistic peers. This means that to isolate the effect of speech, one would have to assign young children to different linguistic peer groups in such a way as to influence their speech without influencing them in other ways that might affect their eventual labor market productivity. Since linguistic peer groups are probably composed of children who live in the same neighborhood or attend the same school, it's hard to see how one could implement this thought experiment.

However, one can ask a related question, which is whether the effects attributable to speech stem rather from other observable characteristics that may be related to speech. The NLSY97 provides a number of such characteristics. In addition to the variables in Table 2, these include family structure, maternal education, and information on whether the worker attended a Catholic or private school.

[10] There are 60 such workers in the sample. Of these, 23 are black, whose speech may exhibit features common to SoAE and AAVE. Moreover, 57 percent are Midwestern and 28 percent have missing values for region at age 12. The South Midlands dialect spoken in some plains states shares features with SoAE (Clopper and Pisoni 2004), and some workers with missing region data could be Southern.

Column (1) of Table 4 presents results from a regression that adds the family background measures.[11] Although I omit the additional coefficients in order to save space, the coefficients on AFQT scores, family income during childhood, and maternal education were all significant. These additional variables reduce the coefficient on the black dummy to -0.099. Comparing this to the estimate of -0.152 from the last column of Table 3, about one-third of the wage gap between black workers with racially distinctive speech and non-Southern whites can be explained by variables related to family background and school quality. However, adding these variables has much less effect on the mainstream speech premium for blacks, reducing it from 0.136 only to 0.123. Family background has a greater effect on the mainstream speech premium for Southern whites, reducing it from 0.086 to 0.063.

The regression in column (2) adds location fixed-effects. These location measures indicate which state the worker resides in, and within each state, whether he lives in an urban or rural area. Adding the location fixed-effects has little effect on the mainstream speech coefficient for blacks. However, it reduces the mainstream speech coefficient for Southern whites to 0.036 and renders it insignificant.

One explanation for this is suggested in Table 2. More Southern-sounding workers tend to live in more rural areas, whereas mainstream speakers live in more urban areas. Thus for Southern whites, the mainstream speech premium may be picking up cost-of-living differences that are reflected in wages. As a result, the rest of this paper focuses on the mainstream speech premium for blacks.

[11] All regressions that include this extended set of regressors also include missing value flags for household income and parental education.

The black-white wage gap varies by education level, with a narrower gap for more-educated workers (Neal and Johnson 1996; Lang and Manove 2011). The strong correlation between speech patterns and education observed in Table 2 raises the question of whether the variation in the racial wage gap by schooling may be related to speech. The regressions reported in columns (3) and (4) address this question.

Column (3) adds to the regression in column (2) interactions between the black dummy and the education dummies. The pattern observed here is similar to that reported elsewhere: among college-educated workers, there is no significant difference between blacks and whites (and if anything, a small though insignificant wage premium for blacks), whereas the gap is substantial among workers with less schooling. Column (4) adds three-way interactions between the black dummy, the education dummies, and the mainstream speech dummy. The three-way interactions suggest that the mainstream speech premium for blacks is larger for workers with at least some college education, although the interactions are non-monotone in schooling and only significant for workers with some college but no degree.[12] At the same time, adding the three-way interactions has little effect on the two-way interactions between the black dummy and the education dummies. Differences in the black-white wage gap by education level seem to be largely unrelated to speech.

Columns (5) and (6) report regressions that include variables designed to address whether specific measures of the worker's grammar can explain the mainstream speech

[12] These results are roughly consistent with findings that knowing a second language is more valuable for more educated workers (Berman, Lang, and Siniver 2003; Lang and Siniver 2009).

premium for blacks. The grammar variables are based on dialect density measures (DDM's) that were constructed by linguists at the University of Michigan.

To construct the DDM's, the audio files were transcribed and the transcriptions were coded for 24 grammatical features that differ between AAVE and SAE. The DDM is the number of such features produced by the speaker, divided by the number of words uttered by the speaker in response to the speech prompts. DDM's have been used to measure speech differences by race and have been shown to predict test scores among children (Craig, & Washington, 2004; Craig, Washington, & Thompson-Porter, 1998; Craig 2015). The average African American worker in the sample used roughly one such feature per 50 words in response to the job-search prompt, with a range of zero to one-in-ten. Mean AAVE production was higher in response to the happiest-moment prompt, consistent with the expectation that the HM prompt would elicit more informal speech (Craig and Grogger 2012). Due to budgetary constraints, DDMs were constructed for only a random sample of roughly half of those respondents who provided useful responses to the speech prompts.

Column (5) adds two variables to the regression from column (2): the DDM based on the job-search prompt and an interaction between that variable and the black dummy. This regression is restricted to sample members for whom the the job-search DDM is available. The coefficient on the interaction term is negative, suggesting that black speakers who used more AAVE have lower wages. However, that coefficient is

insignificant, and adding the DDM variables to the model has little effect on the mainstream speech premium.[13]

Column (6) adds to the specification from column (2) a measure of code-shifting. It equals one if the job-search DDM was weakly less than the happiest-moment DDM, that is, if the speaker used more mainstream speech in response to the job-search prompt. This regression only includes sample members for whom both DDM's are available. The regression also includes an interaction between that measure and the black dummy. The coefficient on the interaction suggests that workers who code-shift earn a wage premium. However, that coefficient is insignificant, and adding the code-shifting variables to the regression does not affect the estimated mainstream wage premium.[14]

D. Robustness and additional specification issues

In Appendix Table 5, I report regressions that include additional background variables.[15] These include the Ten-Item Personality Inventory, designed to capture personality traits that may be correlated with sociability and confidence, both of which have been shown to positively influence wages (Mobius and Rosenblatt 2006; Borghans et al 2014; Deming 2015).[16] They also include measures of the respondent's skin color, interactions between dummies for 2-digit occupation and 2-digit industry, and dummies

[13] Results based on the happiest-moment DDM were similarly insignificant, as were results based on the DDM that pooled responses to the two prompts.
[14] The coefficient of an alternative measure, equal to one if the job-search DDM was strictly less than the happiest-moment DDM, was similar.
[15] In addition to the variables shown, all the regressions also include all the variables from the specification reported in column (2) of Table 4.
[16] The TIPI is designed to capture the so-called Big Five personality traits, which include openness to exposure, conscientiousness, extraversion, agreeableness, and neuroticism. TIPI scales were coded in the manner proposed by Gosling et al. (2003).

for whether the worker was ever arrested or incarcerated.[17] None of these variables had much effect on the mainstream speech coefficients.[18] Column (6) restricts the sample to black workers, which allows for all the regression coefficients to differ by race. The resulting mainstream wage premium falls slightly as a result.

Another concern involves the young age of the sample members. One might be concerned that any results obtained from young workers might change as they age. Although there are limits to what one can do, given the basic age limits of the sample, Appendix Table 6 reports estimates from regressions that impose various minimum age thresholds. The table reports the black*mainstream speech coefficient from regressions that include all the variables from column (2) of Table 4. If anything, the mainstream speech premium rises slightly as the minimum age rises.

Another issue that may bear on the validity of the regression results involves the availability of wage data. Even after a worker leaves school, there are many periods when the NLSY97 does not observe his wages, either because he is not employed or because of item non-response. If wages are missing non-randomly in a manner that is correlated with speech patterns, then the speech coefficients could reflect the process underlying the missing data rather than speech.

Table 5 presents results from a regression that addresses this issue. The sample includes all person-years in which the worker satisfies the sample inclusion criteria discussed above. The dependent variable equals one if a wage is observed and zero otherwise. The regression includes all the variables included in the wage regression

[17] The skin-color measures were coded as in Kreisman and Rangel (2013).
[18] Additionally including interactions between these variables and the black dummy yielded similar results.

reported in the second column of Table 4. Only variables related to race and speech are included here in order to save space.

The joint F-statistic for all the variables that appear in column (1) has a p-value of 0.483. Only the coefficient for Southern-sounding speech on the part of non-Southerners is individually even marginally significant. The availability of wages is largely independent of speech.

VI. Explaining the relationship between speech patterns and wages

The analysis above shows a strong relationship between mainstream speech and wages among African Americans that cannot be explained by omitted variable bias involving location, AFQT scores, household income, family structure, parental education, personality traits, or criminal records. Indeed, the speech-related wage gap among blacks is as large as the wage gap one observes between blacks and whites, even after conditioning on this broad set of controls. I next discuss an occupational sorting model that helps explain this finding.

A. Occupational sorting

Think about classifying occupations according to the extent to which they require workers to carry out some task *T*. Denote occupations that are *T*-intensive by t=1 and those that are not *T*-intensive by t=0. Consider a worker skill *S* measured in terms of the worker's efficiency at carrying out task *T*. Denote the price paid for task *T* in each sector by λ_t, t=0, 1. Adopting a simple Roy (1951) model, the worker's log wage in sector t is given by

(1) $w_t = \alpha_t + \lambda_t S + \varepsilon_t$ t=0, 1.

Since *S* makes the worker more productive at *T*, it is more valuable in the *T*-intensive sector than the *T*-non-intensive sector, so $\lambda_1 > \lambda_0 \geq 0$.[19] The terms ε_0 and ε_1 are zero-mean disturbances assumed to be normally distributed independently of *S*.

The worker chooses the task-intensive sector if $w_1 > w_0$. This implies:

(2) $P(i\ chooses\ t = 1) = F[(\alpha_1 - \alpha_0) + (\lambda_1 - \lambda_0)S]$,

where $F(\ )$ is the normal cdf. Since $\lambda_1 > \lambda_0$, workers with high values of *S* should sort disproportionately into the task-intensive sector. A probit regression of a sector dummy equal to 1 if the worker works in the task-intensive sector on *S* provides an estimate of the sign of $\lambda_1 - \lambda_0$, which can used to test for sorting. Sector-specific wage regressions of the form in (1) provide estimates of sector-specific skill prices.[20]

B. Data on tasks

Empirically, *S* represents mainstream speech, so the key is to find a set of occupational tasks *T* where *S* is particularly productive. As discussed in the Introduction, it is natural to think that mainstream speech might raise the worker's productivity in interactive tasks involving consumers and coworkers. This could stem from consumer discrimination or arise if mainstream speech represents a social skill that makes the worker more effective in interacting with colleagues.

---

[19] Equilibrium with workers in both sectors then requires $\alpha_1 < \alpha_0$.

[20] The OLS estimate of λ_1 will be biased downward, and the OLS estimate of λ_0 will be biased upward, provided that Cov(ε_0, ε_1)<Var(ε_1). Heckman and Honore (1990) refer to this as the "normal case." To see the upward bias argument, consider a worker with low *S*. To be found in sector 1, he must have a high value of ε_1. Thus within sector 1, *S* and ε_1 will be negatively correlated. The argument for downward bias in sector 0 is similar.

I use the O*NET database to classify occupations according to their interaction intensity, specifically utilizing Deming's (2015) index of social skills. This index is a measure of the extent to which jobs require tasks involving social perceptiveness ("being aware of others' reactions and understanding why they react the way they do"), coordination ("adjusting actions in relation to others' actions"), persuasion ("persuading others to approach things differently"), and negotiation ("bringing others together and trying to reconcile others"). Details of index construction are provided in the Data Appendix and Appendix Table 7 lists the 20 occupations with the highest and lowest values of the index. I classify occupations as interaction-intensive if they fall within the top quartile of the index.

C. Regression results

The first row of Table 6 presents the black*mainstream speech coefficient from sorting probits and sector-specific log wage regressions.[21] Column (1) shows that black mainstream speakers sort strongly into interaction-intensive jobs; the t-statistic for the black*mainstream coefficient is 4.4. Column (2) shows that black mainstream speakers in that sector earn a sizeable wage premium as compared to non-mainstream speakers. In the non-interaction-intensive sector (column (3)), mainstream speech does not earn a significant wage premium. This evidence is consistent with the occupational sorting model.

One can take the occupational sorting model a step further and ask whether mainstream speakers truly sort toward occupations where interpersonal interaction is

[21] All regressions include all of the variables from column (2) of Table 4, in addition to the variables shown.

important, or whether they sort into task-intensive occupations more generally. In that case, we might conclude that mainstream speakers just had higher levels of general skill that were not being accounted for in the regressions.

To draw this distinction, I present additional sets of sorting probits and sector-specific wage regressions where the sectors are defined in terms of their task-intensity involving other types of tasks for which speech does not confer an obvious advantage. I consider six different types of tasks. These include non-routine analytical tasks, information use, inductive and deductive reasoning, facility with numbers, routine manual tasks, and routine cognitive tasks. The first four measures are taken from Deming (2015); the last two are taken from Autor and Handel (2012). As above, I constructed task indexes and then defined the task-intensive sector as occupations in the top quartile of the task index. Details are in the Data Appendix.

The remaining rows of Table 6 present mainstream speech coefficients from the six different sets of regressions corresponding to each of these tasks. Among the six tasks not related to speech, the sorting coefficients for non-routine analytical tasks and inductive/deductive reasoning are significant. Moreover, wage premia in the task-intensive and non-task-intensive sectors are generally similar. However, workers do not earn a wage premium from sorting into occupations that are intensive in these tasks. For the other four tasks, there is no significant sorting by task intensity, and for none of the tasks besides interaction does the mainstream speech premium differ significantly between task-intensive and task-non-intensive occupations.

Finally, it is worth nothing that the wage premium for mainstream speech arises in occupations with the highest levels of interaction intensity. When I defined interaction-intensive occupations to include those in the top quintile of the interaction-intensity distribution, as opposed to the top quartile as in Table 6, I obtained similar results. However, when I defined interaction-intensive occupations to include those in the top half or top tercile of the interaction-intensity distribution, or allowed interaction-intensity to enter linearly, the combination of strong sorting and differential wage premia was not apparent.

VI. Evaluating Reverse Causation

The evidence above suggests that sorting helps explain the mainstream speech premium for black workers: mainstream speakers sort into interaction-intensive occupations, where they are presumably more productive than their counterparts, all else equal. As a result, they earn a wage premium in such occupations. This section considers the evidence on an alternative explanation, which is that speech patterns are shaped by experiences on the job. There is no definitive test to distinguish between these hypotheses, but I consider three types of evidence.

The first was the literature discussed in Section II of the paper. The sensitive-period literature cited there shows that one's native accent is largely set by puberty, acquired from one's linguistic peers during childhood. However, as extensive as that literature is, some workers may nevertheless be able to adopt more mainstream speech patterns as adults if they perceive it to be in their interest to do so. Such adaptation could

exaggerate the extent of wage inequality stemming from speech if workers facing higher wage prospects were those most able to adapt.

The other types of evidence are empirical, the first of which takes the form of data on changes in mainstream speech. Such evidence comes from combining the speech data here with speech data from Grogger (2011). Data for that much smaller study were collected in 2006. As with the data here, anonymous listeners listened to short audio clips and answered short questions about the speakers. In the 2006 data, speakers who were perceived as black by at least four out of five listeners were categorized as having distinctively black speech. I refer to the remainder as having mainstream speech. The main limitation of this approach is that there are so few black NLSY97 respondents with speech data from both 2006 and 2011.

Table 7 presents cross-tabulations of 2006 speech patterns by 2011 speech patterns among black workers. If workers adapt their speech, then we would expect to observe a higher share of speakers moving from non-mainstream in 2006 to mainstream in 2011 rather than moving from mainstream to non-mainstream. Table 6 shows just the opposite. Of the nine speakers counted as mainstream in 2006, four were classified as non-mainstream in 2011. In contrast, only two of the 25 speakers who were classified as non-mainstream in 2006 were classified as mainstream in 2011. This suggests that there is little adaptation toward mainstream speech.

The final piece of evidence takes the form of additional occupational sorting regressions. I consider sorting at two jobs, the first and the last that I observe during my sample period. The idea is that, if speech responds to the workplace environment,

workers should be sorted more strongly at the end of the sample period than at the beginning of their careers.

Table 8 presents the black*mainstream coefficients from two sorting probits along the lines reported in Table 6. In column (1), the dependent variable equals one if the worker was in an interaction-intensive occupation in the first job at which he was observed. In column (2), the dependent variable equals one if the worker was in an interaction-intensive occupation in the last job at which he was observed. I restrict attention to workers whose two observations were at least five years apart, in order to provide a reasonable amount of time for speech to adapt. In both cases, interaction-intensive occupations are those that lie in the top quartile of the interaction intensity distribution.

Both coefficients are positive and significant. More importantly for my purposes, the coefficients are essentially the same for both jobs. Thus there is no evidence that mainstream-spoken workers sort more strongly into interaction-intensive occupations as their careers advance. More generally, neither piece of empirical evidence is inconsistent with the finding from the linguistics literature that speakers' accents change little after puberty.

VII. Conclusion

The results here reveal a little-known dimension of inequality. Racially and regionally distinctive speech patterns are strongly negatively correlated with wages. For Southern whites, those differences can be explained by family background and location

of residence. For African Americans, however, speech-related wage differences are not explained by family background, location, personality traits, or arrest records.

Considering what is known about language acquisition, it is impossible say that the regression analysis isolates the effect of mainstream speech. Nevertheless, these findings fit nicely with recent results from the Moving to Opportunity study. MTO offered low-income African American families housing vouchers that could be used to move into higher-income areas. Rickford et al (2015) show that children who were young shortly after their families received the vouchers reduced their use of AAVE, but children who were older did not. At the same time, Chetty et al (2015) show that the younger children, but not the older children, experienced higher earnings as young adults. Since children acquire their speech patterns while they are young, largely from the children who surround them, these results put together are consistent with the notion that speech affects earnings.

To explain the mainstream speech premium for black workers, I propose an occupational sorting model. That model predicts that workers who are productive in a particular task should sort toward occupations that utilize that task intensively and earn a wage premium in the process. This is exactly what we observe: mainstream-spoken black workers sort into jobs that involve intensive interactions with customers and coworkers, and earn a sizeable wage premium in those jobs. They do not sort as much into other types of skill-intensive jobs, nor receive wage premiums in occupations that intensively utilize other types of skill.

An important question is why mainstream speech is productive in interaction-intensive jobs. Semantic understanding seems an unlikely answer, since AAVE and SAE are fully mutually intelligible dialects of the same language. One explanation is customer and coworker discrimination. An abundance of evidence from social psychology shows that listeners prefer mainstream to non-mainstream speech, which could result in higher wages for mainstream-spoken workers in highly interactive sectors. Either way, an interesting extension for future work would be to test whether mainstream speech earns a wage premium in other multi-dialectical labor markets.

References

Akerlof, George A. and Rachel E. Kranton. "Economics and Identity." Quarterly Journal of Economics 115 (3), 2000, 715-753.

Altonji, J. G., Peasant Bharadwaj, and Fabian Lange (2009). "Constructing AFQT Scores that are Comparable Across the NLSY79 and the NLSY97." http://www.fabianlange.ca/data.html

Altonji, J. G., and Rebecca M. Blank. "Race and Gender in the Labor Market." In O. Ashenfelter and D. Cards, eds., Handbook of Labor Economics, v. 3. Amsterdam: Elsevier, 1999.

Altonji, J. G. & Pierret, C. R. (2001). Employer Learning and Statistical Discrimination. Quarterly Journal of Economics, 464 (February), pp. 313-350.

Austen-Smith, David, and Roland G. Fryer, Jr. "An Economic Analysis of 'Acting White.'" Quarterly Journal of Economics, 2005, 551-583.

Bailey, Guy and Erik Thomas. "Some Aspects of African American Vernacular Phonology." In Mufwene, Salikoko S, John R. Rickford, Guy Bailey, and John Baugh, eds., African American English: Structure, History, and Use. London: Routledge, 1998.

Bailey, Guy and Jan Tillery. "The Persistence of Southern American English." Journal of English Linguistics 24(4), 1996, 305-321.

Baugh, John. "Hypocorrection: Mistakes in Pronunciation of Vernacular African American English as a Second Dialect." Language and Communication 12 (3/4), 1992, 317-326.

Becker, Gary S. The Economics of Discrimination. Chicago: The University of Chicago Press, 2nd edition, 1971.

Berman, Eli, Kevin Lang, and Erez Siniver. "Language-skill complementarity: returns to immigrant language acquisition. Labour Economics 10 (2003), 265-290.

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(4), 2004, 991-1013.

Biddle, Jeff E. and Daniel S. Hamermesh. "Beauty, Productivity, and Discrimination: Lawyers' Looks and Lucre." Journal of Labor Economics, Vol. 16, No. 1 (January 1998), pp. 172-201.

Borghans, Lex, Angela Lee Duckworth, James J. Heckman, and Bas Ter Weel. "The Economics and Psychology of Personality Traits." Journal of Human Resources 43(4), 2008, 972-1059.

Borghans, Lex, Bas Ter Weel, and Bruce A. Weinberg. “People, Skills, and the Labor Market Outcomes of Underrepresented Groups.” ILR Review 67(2), 2014, 287-334.

Bleakley, Hoyt and Aimee Chin. “Language Skills and Earnings: Evidence from Childhood Immigrants.” Review of Economics and Statistics 86 (2), 2004, 481-496.

Cameron, Steven and James J. Heckman. "The Nonequivalence of High School Equivalents." Journal of Labor Economics 11, (January 1993), 1-47.

Cawley, John. “The Impact of Obesity on Wages.” Journal of Human Resources 39, March 2004, 451-474.

Charles, Kerwin K. and Jon Guryan. “Prejudice and the Economics of Discrimination.” NBER Working Paper 13661, December 2007.

Chiswick, Barry R. “Speaking, Reading, and Earnings among Low-Skilled Immigrants.” Journal of Labor Economics 9 (2), 1991, 149-170.

Clopper, Cynthia G., and David B. Pisoni. “Some Acoustic Cues for the Perceptual Categorization of American English Regional Dialects.” Journal of Phonetics 32, 2004, 111-140.

Cook, Phillip, and Jens Ludwig. “Weighing The “Burden Of ‘Acting White’”: Are There Race Differences In Attitudes Towards Education? Journal of Public Policy and Analysis, 16, 1997, 256-278.

Craig, H. K. (2015). African American English and its links to achievement. In Pollatsek, A. & Treiman, R. (Eds.), Oxford Handbooks Online. DOI:10.1093/oxfordhb/9780199324576.013.27

Craig, H. K., & Washington, J. A. (2004). Grade-related changes in the production of African American English. Journal of Speech, Language, and Hearing Research, 47,450–463

Craig, H. K., & Grogger, J. T. (2012). Social and style constraints on adult usage of African American English variants. Journal of Speech, Language, and Hearing Research, 55,1274–1288.

Craig, H. K., Washington, J. A., & Thompson-Porter, C. (1998). Average c-unit lengths in the discourse of African American children from low-income, urban homes. Journal of Speech, Language, and Hearing Research, 41, 433–444.

Cukor-Avila, Patricia. "The complex grammatical history of African-American and white vernaculars in the South." In Stephen J. Nagle and Sara L. Sanders, eds., English in the Southern United States. Cambridge: Cambridge University Press, 2003, 82-105.

Dorrill, George. "Sounding southern: A look at the phonology of English in the South." In Stephen J. Nagle and Sara L. Sanders, eds., English in the Southern United States. Cambridge: Cambridge University Press, 2003, 17-35.

Doss, Richard C. and Alan M. Gross. "The Effects of Black English on Stereotyping in Intraracial Perceptions." Journal of Black Psychology 18 (2), 1992, 47-58.

Doss, Richard C. and Alan M. Gross. "The Effects of Black English and Code-Switching on Intraracial Perceptions." Journal of Black Psychology 20 (3), 1994, 282-293.

Falck, Oliver, Stephan Heblich, Alfred Lameli, and Jens Südekum. "Dialects, cultural identity, and economic exchange." Journal of Urban Economics 72 (2–3), 2012, 225-239.

Fan, Samantha P., Zoe Liberman, Boaz Keyser, and Katherine D. Kinzler. "The Exposure Advantage: Early Exposure to a Multilingual Environment Promotes Effective Communication." Psychological Science, 2015, DOI: 10.1177/0956797615574699

Farber, H. S. & Gibbons, R. (1996). Learning and Wage Dynamics. Quarterly Journal of Economics 111:1007-47.

Fasold, Ralph W. "The Relation Between Black and White Speech in the South." American Speech 56, 163-189.

Fordham, Signithia and John U. Ogbu. "Black Students' School Success: Coping with the 'Burden of "Acting White."' The Urban Review 18 (3), 1986, 176-206.

Fridland, Valerie. "Regional differences in perceiving vowel tokens on Southerness, education, and pleasantness ratings." Language Variation and Change 20, 2008, 67-83.

Fridland, Valerie, and Kathryn Bartlett. "Correctness, Pleasantness, And Degree Of Difference Ratings Across Regions." American Speech 81 (4), 2006, 358-386.

Fridland, Valerie, Kathryn Bartlett and Roger Kreuz. "Making Sense Of Variation: Pleasantness And Education Ratings Of Southern Vowel Variants." American Speech 80 (4), 2005, 366-387.

Fryer, Roland G. Jr., and Steven D. Levitt. "The Causes and Consequences of Distinctly Black Names." Quarterly Journal of Economics 119, 2004, 767-805.

Fryer, Roland G. Jr., and Paul Torelli. "An Empirical Analysis of 'Acting White.'" NBER Working Paper 11334, May 2005.

Gosling, S. D., Rentfrow, P. J., & Swann, W. B., Jr. (2003). A Very Brief Measure of the Big Five Personality Domains. Journal of Research in Personality, 37, 504-528.

Granena, Gisela, and Mike Long. Sensitive periods, language aptitude, and ultimate L2 attainment. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2013.

Green, Lisa. "Aspect and predicate Phrases in African American Vernacular English." In Mufwene, Salikoko S, John R. Rickford, Guy Bailey, and John Baugh, eds., African American English: Structure, History, and Use. London: Routledge, 1998.

Grogger, J. "Speech Patterns and Racial Wage Inequality." Journal of Human Resources 46, Winter 2011, 1-25.

Hamermesh, Daniel S., and Jeff E. Biddle. "Beauty and the Labor Market." American Economic Review, Vol. 84, No. 5 (Dec., 1994), pp. 1174-1194.

Hartley, Laura. "A View from the West: Perceptions of U.S. Dialects by Oregon Residents." In D. R. Preston, ed., Handbook of Perceptual Dialectology. Amsterdam: John Benjamins Publishing Company, 1999, 315-332.

Heblich, Stephan & Lameli, Alfred & Riener, Gerhard, 2014. "The Effect of Perceived Regional Accents on Individual Economic Behavior: A Lab Experiment on Linguistic Performance, Cognitive Ratings and Economic Decisions," IZA Discussion Papers 8640, Institute for the Study of Labor (IZA).

Heckman, James J. and Bo E. Honore. "The Empirical Content of the Roy Model." Econometrica 58(5), 1990, 1121-1149.

Hyltenstam, Kenneth and Niclas Abrahamsson, "Maturational constraints in SLA," in Doughty, Catherine J. and Michael H. Long, eds., The handbook of second language acquisition, 539-599. Malden, MA: Blackwell Publishing, 2003.

Johnson, Jacqueline S., and Elissa L. Newport. "Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language." Cognitive psychology 21 (1), 1989, 60-99.

Kim, Young Chul, and Glenn C. Loury. "Sterotypes and Inequality: A 'Signaling' Theory of Identity Choice." KDI Journal of Economic Policy, Vol. 34, No. 2, 2012.

Kinzler, Katherine D., and Jasmine M. DeJesus. "Northern = smart and Southern = nice: The development of accent attitudes in the United States." The Quarterly Journal of Experimental Psychology 66 (6), 2012, 1146-1158.

Koch, L., Gross, A., & Kolts, R. "Attitudes toward Black English and code switching." Journal of Black Psychology 27, 2001, 29–42.

Kurath, Hans, and Raven Ioor David. The Pronunciation of English in the Atlantic States: Based upon the Collections of the Linguistic Atlas of the Eastern United States. University of Michigan Press, 1961.

Labov, William. The social stratification of English in New York City. Washington, DC: Center for Applied Linguistics, 1966.

Labov, William. Language in the Inner City: Studies in the Black English Vernacular. Philadelphia: University of Pennsylvania Press, 1972.

Labov, William. "Co-existent Systems in African American Vernacular English." In Mufwene, Salikoko S, John R. Rickford, Guy Bailey, and John Baugh, eds., African American English: Structure, History, and Use. London: Routledge, 1998.

Lang, Kevin. "A Language Theory of Discrimination." Quarterly Journal of Economics 101 (2), 1986, 363-382.

Lang, Kevin and Michael Manove. "Education and Labor Market Discrimination." American Economic Review 101(3), June 2011.

Lang, Kevin and Erez Siniver. "The return to English in a non-English-speaking country: Russian immigrants and native Israelis in Israel." B.E. Journal of Economic Analysis and Policy 9 (1) (Topics), 2009, article 50.

Lange, Fabian. "The Speed of Employer Learning." Journal of Labor Economics 25 (1), January 2007.

Long, Michael H. "Maturational Constraints on Language Development." Studies in Second Language Acquisition 12, 251-285, 1990.

Martin, Stefan and Walt Wolfram. "The Sentence in African American Vernacular English." In Mufwene, Salikoko S, John R. Rickford, Guy Bailey, and John Baugh, eds., African American English: Structure, History, and Use. London: Routledge, 1998.

Massey, Douglas S. and Garvey Lundy. "Use of Black English and Racial Discrimination in Housing Markets." Urban Affairs Review 36 (4), 2001, 452-469.

McManus, Walter, William Gould, and Finis Welch. "Earnings of Hispanic Men: The Role of English Language Proficiency." Journal of Labor Economics 1 (2), 1983, 101-130.

Mobius, Markus M. and Tanya S. Rosenblat. "Why Beauty Matters." American Economic Review, Vol. 96, No. 1 (Mar., 2006), pp. 222-235.

Mufwene, Salikoko S. "The Structure of the Noun Phrase in African American Vernacular English." In Mufwene, Salikoko S, John R. Rickford, Guy Bailey, and John Baugh, eds., African American English: Structure, History, and Use. London: Routledge, 1998.

Neal, Derek. "Black-White Labor Market Inequality in the United States." In Steven Durlauf and Lawrence Blume, eds., New Palgrave Dictionary of Economics, 2006

Neal, D. A. and Johnson, W. R. (1996). The Role of Premarket Factors in Black-White Wage Differences. The Journal of Political Economy, Vol. 104, No. 5 (Oct., 1996), pp. 869-895.

Neal, D. A., Johnson, W. R. and Kitamura. "Evaluating a Simple Method for Estimating Black-White Gaps in Median Wages." American Economic Review 90, (May 2000), 339-343.

Jencks, Christopher, and Meredith Phillips. "The Black-White Test Score Gap: An Introduction." In Jencks, Christopher, and Meredith Phillips. eds., The Black-White Test Score Gap. Washington, DC: Brookings Institution Press, 1998.

Johnson, George E. and Frank P. Stafford. "Alternative approaches to occupational exclusion." In I. Persson and C. Jonung, eds., Women's work and wages. Routledge: London, 1998.

Pattillo, Mary. *Black Picket Fences: Privilege and Peril Among the Black Middle Class.* Chicago" University of Chicago Press, 1999.

Pinker, Steven. Words and Rules: The Ingredients of Language. New York: Basic Books, 1999.

Preston, Dennis R. "Where the worst English is spoken." In Edgar Schneider, ed., Focus on the USA. Amsterdam: Benjamins, 1996, 297-360.

Preston, Dennis R. "A Language Attitude Analysis of Regional US Speech: Is Northern US English Not Friendly Enough?" Cuadernos de Filologia Inglesa 8, 1999, 129-146.

Purnell, Thomas, William Idsardi, and John Baugh. "Perceptual and Phonetic Experiments on American English Dialect Identification." Journal of Language and Social Psychology 18 (1), 1999, 10-30.

Rickford, John R. and Christine Theberge Rafal. "Preterite Had + V-Ed in the Narratives of African-American Preadolescents." American Speech 71 (3), 1996, 227-254.

Rodriguez, J., Cargile, A., & Rich, M. "Reactions to African-American vernacular English: Do more phonological features matter?" The Western Journal of Black Studies 28, 2004, 407–414.

Roy, A. D. "Some thoughts on the distribution of earnings." Oxford Economic Papers, New Series 3 (2), 1951, 135-146.

Schneider, Edgar W. 2003. "Shakespeare in the coves and hollows? Toward a history of Southern English." In Stephen J. Nagle and Sara L. Sanders, eds., English in the Southern United States. Cambridge: Cambridge University Press, 17-35.

Siegel, Jeff. Second Dialect Acquisition. 1st ed. Cambridge: Cambridge University Press, 2010.

Thomas, Erik R. “Sociophonetic Applications of Speech Perception Experiments.” American Speech 77 (2), 2002, 115-147.

Thomas, Erik R., and Jeffrey Reaser. “Delimiting Perceptual Cures Used for the Ethnic Labeling of African American and European American Voices.” Journal of Sociolinguistics 8 (1), 2004, 54-87.

Trejo, Stephen J. “Why Do Mexican Americans Earn Low Wages?” Journal of Political Economy 105 (6), 1997, 1235-1268.

Tucker, G. Richard, and Wallace E. Lambert. "White and Negro Listeners' Reactions to Various American-English Dialects." Social Forces 47 (4), 1969, 463-468.

Walton, Julie H. and Robert F. Orlikoff. “Speaker Race Identification from Acoustic Cues in the Vocal Signal.” Journal of Speech and Hearing Research 37 (4), 1994, 738-746.

Washington, Julie A. and Holly K. Craig. “Morphosyntactic Forms of African American English Used by Young Children and their Caregivers.” Applied Psycholinguistics 23, 2002, 209-231.

Wolfram, Walt. “Black-White Speech Differences Revisited.” in Wolfram, Walt and Nona H. Clarke, eds., Black White Speech Relationships. Washington, DC: Center for Applied Linguistics, 1971.

Wolfram, Walt. Dialects and American English. Englewood Cliffs, NJ: Prentice Hall, 1991.

Table 1: Percentage distribution of listener reports that speaker is black or Southern, by speaker's race/region at age 12, males

	L reports that speaker is black		L reports that speaker is Southern	
	Number	Percentage	Number	Percentage
Speaker's race/region	(1)	(2)	(3)	(4)
Black	0	5.3	0	15
	1	5.5	1	16.8
	2	5.8	2	18.3
	3	12.7	3	15.6
	4	15.6	4	14
	5	22.6	5	10.5
	6	32.6	6	9.7
Southern white	0	60.2	0	23.8
	1	24.1	1	15.3
	2	10	2	14.9
	3	3.4	3	13.8
	4	1.5	4	12.6
	5	0.8	5	10
	6		6	9.6
Non-Southern white	0	88.3	0	73.9
	1	10.1	1	19.8
	2	1.6	2	4
	3	0.1	3	2.4

Note: Excludes blacks and Southern whites with only 3 listener reports. L = listener.

Table 2: Human capital, family income, residential location and wages by listener reports of speaker's race/region at age 12, by speaker's race/region at age 12, males

		Number of listener reports that speaker is black							
Race/region at age 12	Variable	0	1	2	3	4	5	6	Total
Black	Years of education	14.37	13.14	13.31	13.45	12.77	12.22	11.46	12.44
	Standardized AFQT	0.07	0.16	-0.48	-0.42	-0.69	-0.72	-0.83	-0.6
	HH income, ($10K)	4.77	2.89	4.01	4.22	3.43	3.06	2.41	3.16
	Lives in rural area	0.14	0.21	0.06	0.1	0.19	0.21	0.25	0.19
	Log hourly wage	2.83	2.61	2.73	2.63	2.52	2.42	2.46	2.52
	N	27	28	29	64	78	114	164	504

		Number of listener reports that speaker is Southern							
Race/region at age 12	Variable	0	1	2	3	4	5	6	Total
Southern White	Years of education	15.66	14.57	13.5	12.64	12.48	12.52	12	13.7
	Standardized AFQT	0.69	0.82	0.43	-0.11	-0.3	0.22	-0.43	0.29
	HH income, ($10K)	7.68	6.69	6.86	5.15	4.92	5.6	4.75	6.22
	Lives in rural area	0.14	0.24	0.26	0.37	0.22	0.42	0.58	0.28
	Log hourly wage	2.92	2.86	2.78	2.63	2.57	2.59	2.56	2.74
	N	62	40	38	36	33	25	24	258

Note: Blacks and Southern whites with only three listener reports are excluded. N refers to the cell size for years of education.

Table 3: Log wage regressions with alternative representations of speech

Dependent variable is the log hourly wage					
Variable	(1)	(2)	(3)	(4)	(5)
Black	-0.113	-0.031	-0.038	0.008	-0.152
	(0.025)	(0.039)	(0.037)	(0.041)	(0.029)
N L reports S is black		-0.037			
		(0.017)			
N. L reports S is black, residual			-0.037		
			(0.018)		
1 L reports that S is black*				-0.001	
				(0.045)	
2 L reports that S is black*				0.065	
				(0.038)	
3 L reports that S is black*				-0.043	
				(0.058)	
4 L reports that S is black*				-0.134	
				(0.047)	
5 L reports that S is black*				-0.189	
				(0.057)	
6 L reports that S is black*				-0.094	
				(0.058)	
Black*mainstream speech					0.136
					(0.043)
White*black speech					-0.162
					(0.049)
South age 12	0.032	0.052	0.056	0.060	-0.047
	(0.044)	(0.044)	(0.044)	(0.043)	(0.049)
N L reports S is Southern		-0.046			
		(0.013)			
N L reports S is Southern, resid.			-0.046		
			(0.013)		
1 L report that S is Southern**				-0.055	
				(0.047)	
2 L reports that S is Southern**				-0.045	
				(0.030)	
3 L reports that S is Southern**				-0.077	
				(0.049)	
4 L reports that S is Southern**				-0.122	
				(0.041)	
5 L reports that S is Southern**				-0.149	
				(0.051)	
6 L reports that S is Southern**				-0.131	
				(0.043)	
South*mainstream speech					0.086
					(0.033)
Non-south*South speech					-0.087

					(0.048)
HS only	0.172	0.156	0.156	0.161	0.159
	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)
Some college	0.250	0.225	0.224	0.226	0.226
	(0.033)	(0.032)	(0.032)	(0.032)	(0.033)
BA or more	0.506	0.472	0.472	0.473	0.475
	(0.040)	(0.039)	(0.039)	(0.039)	(0.039)
Experience	0.044	0.045	0.045	0.044	0.044
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Exp. squared	-0.002	-0.002	-0.002	-0.002	-0.002
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Lives in South	-0.038	-0.002	-0.003	-0.007	-0.006
	(0.041)	(0.042)	(0.042)	(0.041)	(0.041)
Lives in urban area	0.026	0.017	0.016	0.013	0.017
	(0.022)	(0.022)	(0.023)	(0.022)	(0.022)
Married	0.166	0.167	0.167	0.167	0.166
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
Observations	6,735	6,735	6,735	6,735	6,735
Adjusted R-squared	0.208	0.216	0.216	0.222	0.218

* Labels pertain to speakers with six listener reports. Speakers with only three listener reports were coded as if two times as many listeners had reported them to be black.

** Labels pertain to speakers with six listener reports. Speakers with only three listener reports were coded as if two times as many listeners had reported them to be Southern.

Note: Standard errors, in parentheses, are clustered by worker. In addition to variables shown, the regressions include year dummies and missing value flags for region, educational attainment, and urbanicity. Missing value flags equal one when the corresponding variable is missing and equal zero otherwise. Missing values of the corresponding variable are recoded to zero

Table 4: Log wage regressions, alternative specifications

Dependent variable is the log hourly wage						
	(1)	(2)	(3)	(4)	(5)	(6)
Black	–0.050	–0.069	–0.067	–0.051	–0.075	–0.113
	(0.064)	(0.063)	(0.065)	(0.066)	(0.098)	(0.106)
Black*mainstream speech	0.109	0.119	0.110	0.057	0.127	0.126
	(0.043)	(0.044)	(0.043)	(0.080)	(0.046)	(0.047)
South age 12	–0.023	0.002	0.001	0.000	0.038	0.058
	(0.048)	(0.048)	(0.048)	(0.048)	(0.083)	(0.087)
South*mainstream speech	0.061	0.034	0.038	0.037	0.030	0.025
	(0.032)	(0.033)	(0.033)	(0.033)	(0.035)	(0.035)
White*black speech	–0.144	–0.148	–0.143	–0.144	–0.058	–0.005
	(0.049)	(0.047)	(0.047)	(0.047)	(0.068)	(0.072)
Non-south*South speech	–0.054	–0.010	–0.012	–0.010	0.091	0.101
	(0.047)	(0.047)	(0.047)	(0.047)	(0.083)	(0.084)
HS only	0.128	0.123	0.127	0.127	0.107	0.115
	(0.027)	(0.028)	(0.036)	(0.036)	(0.037)	(0.038)
Some college	0.155	0.148	0.156	0.157	0.097	0.098
	(0.034)	(0.033)	(0.041)	(0.041)	(0.048)	(0.049)
BA or more	0.350	0.344	0.314	0.316	0.438	0.434
	(0.045)	(0.045)	(0.050)	(0.050)	(0.066)	(0.067)
Black * HS only			–0.013	0.001		
			(0.054)	(0.058)		
Black * some college			–0.019	–0.083		
			(0.063)	(0.069)		
Black * BA			0.191	0.181		
			(0.078)	(0.088)		
Black * HS only * mainstream				–0.047		
				(0.106)		
Black * some col. * mainstream				0.206		
				(0.106)		
Black * BA * mainstream				0.052		
				(0.136)		
Job search DDM					1.941	
					(1.304)	
Black * JS DDM					–1.593	
					(1.595)	
Code shift						–0.050
						(0.046)
Black * code shift						0.071
						(0.062)
State-Urban FE		Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.239	0.280	0.283	0.285	0.338	0.344
Observations	6731	6731	6731	6731	3082	3004

Note: Standard errors, in parentheses, are clustered by worker. In addition to variables shown, the regressions include all variables shown in the last column of Table 3, year dummies, and missing value flags for region, educational attainment, and urbanicity. Missing value flags equal one when the corresponding variable is missing and equal zero otherwise. Missing values of the corresponding variable are recoded to zero .

Table 5: Probit model for availability of wages

Dependent variable = 1 if wage is observed	(1)
Black	-0.047
	(0.184)
Black*mainstream speech	-0.101
	(0.114)
South age 12	0.178
	(0.130)
South*mainstream speech	-0.018
	(0.102)
White*black speech	0.068
	(0.209)
Non-south*South speech	0.271
	(0.145)
Observations	8865
P for joint F-statistic	0.442

Note: Standard errors, in parentheses, are clustered by worker. In addition to the variables shown, the regression includes all variables included in the regressions that appear in column (2) of Table 4.

Table 6: Mainstream speech coefficients from sorting probits and sector-specific linear wage regressions, various sector definitions

Equation:	Sorting into task-specific sector	Log wage, task-intensive sector	Log wage, other sector
Occupation in top quartile of task intensity distribution, for	(1)	(2)	(3)
Interactive tasks	0.598 (0.147)	0.228 (0.087)	0.070 (0.043)
Non-routine analytical tasks	0.418 (0.154)	0.103 (0.081)	0.099 (0.047)
Information use	0.199 (0.203)	0.105 (0.088)	0.137 (0.046)
Inductive/deductive reasoning	0.320 (0.174)	0.034 (0.111)	0.143 (0.045)
Number facility	0.177 (0.162)	0.083 (0.082)	0.125 (0.044)
Routine manual tasks	-0.100 (0.160)	0.101 (0.074)	0.123 (0.051)
Routine cognitive tasks	0.010 (0.156)	0.099 (0.072)	0.128 (0.049)

Note: Occupational tasks as defined by Deming (2015b), except routine manual and routine cognitive, which are defined by Autor and Handel (2012). Standard errors, in parentheses, are clustered by worker. In addition to the variables shown, all regressions include all variables included in the regressions reported in column (2) of Table 4. Regression in column (2) includes workers in top quartile of the task-intensity distribution. Regression in columns (3) includes workers in bottom three quartiles of the task-intensity distribution.

Table 7: Mainstream speech in 2006 versus mainstream in 2011, blacks

	Mainstream in 2011		
Mainstream in 2006	No	Yes	Total
No	23	2	25
	(92)	(8)	(100)
Yes	4	5	9
	(44)	(56)	(100)

Note: Numbers in parentheses are row percentages.

Table 8: Black*mainstream speech coefficients for sorting into interaction-intensive occupations, at beginning and end of sample period

Sorting at:	Beginning of career	End of sample
Black * mainstream speech	0.633	0.529
	(0.240)	(0.218)
No. observations	852	852

Note: Standard errors, in parentheses, are clustered by worker. In addition to the variables shown, all regressions include all variables included in the regressions reported in column (2) of Table 4.

Appendix Table 1: Round-15 response counts by respondent's race and region at age 12

Race/region:	Black	Southern white	Non-southern white	Other	Total
Original 1997 sample	2,335	1244	3169	2,236	8,984
R15 respondents	2,036	995	2524	1,868	7,423
In-person interviews	1,833	846	2220	1,680	6,579
…and consent to record	1,698	786	2034	1,562	6,080
Speech prompt assignment:					
Both questions	1,691	783	213	538	3,225
HM only	1	0	906	516	1,423
JS only	6	3	912	501	1,422
No assignment	0	0	3	7	10
At least one audio file recorded	1,402	651	1603	1,251	4,907
At least one audio file processed	1168	602	1594	861	4225
Both questions	1133	523	0	8	1664
HM only	29	36	857	802	1724
JS only	6	43	737	51	837

Notes: HM = happiest moment; JS = job search

Appendix Table 2: Characteristics of respondents by race/region and availability of speech data (for all R15 respondents)

Speaker's race/region:	Black		Southern white		Non-Southern white	
Variable	No speech data	Speech data	No speech data	Speech data	No speech data	Speech data
Black	1	1	0	0	0	0
In South at age 12	0.55	0.58	1	1	0.0	0.0
Age-12 region missing	0.104	0.095	0	0	0.105	0.078*
Less than HS	0.272	0.27	0.204	0.213	0.143	0.126
HS only	0.219	0.254	0.209	0.216	0.224	0.226
Some college	0.297	0.261	0.216	0.201	0.252	0.226
BA or more	0.195	0.203	0.369	0.365	0.376	0.417*
Educ. missing	0.017	0.012	0.003	0.005	0.005	0.005
Experience	9.196	9.242	8.645	8.654	8.529	8.308
In South	0.612	0.638	0.895	0.902	0.126	0.121
Region missing	0.006	0**	0.008	0.002	0.026	0.001***
Urban	0.833	0.837	0.621	0.64	0.737	0.754
Urban missing	0.02	0.003***	0.02	0.008	0.033	0.006***
Married	0.186	0.222	0.408	0.49*	0.374	0.48***
AFQT	-0.558	-0.527	0.163	0.297	0.328	0.457**
Missing AFQT	0.228	0.219	0.153	0.201	0.153	0.161
Cath/priv school	0.058	0.045	0.117	0.1	0.115	0.111
Two parents	0.273	0.313*	0.56	0.595	0.605	0.622
Gross HH income 97 (in $10,000s)	2.975	2.922	5.315	5.82	5.742	5.978
HH inc. missing	0.301	0.306	0.186	0.188	0.214	0.199
Mom less than HS	0.224	0.226	0.16	0.151	0.111	0.1
Mom HS grad	0.388	0.409	0.318	0.346	0.346	0.356
Mom some college	0.189	0.202	0.234	0.216	0.261	0.267
Mom college grad	0.066	0.057	0.127	0.14	0.137	0.15
Mom postgrad	0.021	0.029	0.092	0.086	0.101	0.086
Mom's ed. missing	0.113	0.077**	0.069	0.061	0.044	0.041
Observations	868	1168	393	602	930	1594

Note: Asterisks denote significance of difference between speech/no speech samples. * p<0.05, ** p<0.01, *** p<0.001

Appendix Table 3: Percentage distribution of listener characteristics, by speech prompt

Prompt: Characteristic	HM (1)	JS (2)
Sex		
Male	27	16
Female	73	84
Total	100	100
Race/ethnicity		
White	83	84
Black	13	15
Hispanic	2	1
Other	2	0
Total	100	100
Region of residence		
Northeast	21	19
Midwest	37	35
South	21	37
West	21	10
Unknown	0	0
Total	100	100
Level of education		
HS diploma or GED	5	24
HS and some college	38	33
Bachelor's degree or higher	57	43
Total	100	100
Mean age (years)	48	54

Note: Listeners are weighted by the number of speakers to whom they listened. HM = happiest moment; JS = job search.

Appendix Table 4: Frequency distribution of listener reports that speaker is black or Southern, by speaker's race/region at age 12, for blacks and Southern whites with only three listener reports, males only

	L reports that speaker is black		L reports that speaker is Southern	
Speaker's race/region	Number	Frequency	Number	Frequency
Black	0	0	0	6
	1	3	1	3
	2	2	2	2
	3	9	3	3
	Total	14	Total	14
Southern white	0	30	0	16
	1	2	1	6
	2	4	2	7
	3		3	7
	Total	36	Total	36

Appendix Table 5: Southern white*mainstream speech coefficients from sorting probits and sector-specific linear wage regressions, various sector definitions (not for publication)

	(1)	(2)	(3)	(4)	(5)	(6)
Black*mainstream speech	0.120	0.108	0.104	0.118	0.124	0.115
	(0.043)	(0.043)	(0.040)	(0.044)	(0.043)	(0.048)
TIPI scales	Yes					
Skin color dummies		Yes				
Ind*occ controls			Yes			
Ever-arrested dummy				Yes		
Ever-incarc dummy					Yes	
Black only						Yes
R-squared	0.317	0.300	0.460	0.294	0.299	0.314
Observations	6731	6731	6715	6731	6731	1977

Note: Standard errors, in parentheses, are clustered by worker. In addition to variables shown, the regressions include all variables included in the regression that appears in column (2) of Table 4. The regression in column (3) is based on fewer observations due to missing occupation/industry information.

Appendix Table 6: Southern white*mainstream speech coefficients from sorting probits and sector-specific linear wage regressions, various sector definitions (not for publication)

Dependent variable is the log hourly wage				
Minimum age:	(1)	(2)	(3)	(4)
	25	26	27	28
Black*mainstream speech	0.143	0.166	0.163	0.163
	(0.047)	(0.049)	(0.052)	(0.057)
R-squared	0.282	0.291	0.299	0.306
Observations	5228	4399	3549	2663

Note: Standard errors, in parentheses, are clustered by worker. In addition to variables shown, the regressions includes all variables included in the regressionfrom column (2) of Table 4.

Appendix Table 7: 3-digit Census occupations with highest and lowest values of interaction intensity index

A. Occupations with lowest values
Packers and Packagers, Hand
Cleaners of Vehicles and Equipment
Shoe Machine Operators and Tenders
Helpers--Production Workers
Pressers, Textile, Garment, and Related Materials
Food Preparation Workers
Shuttle Car Operators
Textile Knitting and Weaving Machine Setters, Operators, and Tenders
Graders and Sorters, Agricultural Products
Molders and Molding Machine Setters, Operators, and Tenders, Metal and Plastic
Laundry and Dry-Cleaning Workers
Janitors and Building Cleaners
Proofreaders and Copy Markers
Etchers and Engravers
Production Workers, All Other
Tire Builders
Cleaning, Washing, and Metal Pickling Equipment Operators and Tenders
Food Servers, Non-restaurant
Shoe and Leather Workers and Repairers
Milling and Planing Machine Setters, Operators, and Tenders, Metal and Plastic

B. Occupations with highest values
Lodging Managers
First-Line Supervisors/Managers of Non-Retail Sales Workers
Medical and Health Services Managers
Judges, Magistrates, and Other Judicial Workers
Advertising and Promotions Managers
Sales Engineers
Psychologists
First-Line Supervisors/Managers of Police and Detectives
Construction Managers
Dietitians and Nutritionists
Counselors
Lawyers
Sales and Related Workers, All Other
Social Workers
Marketing and Sales Managers
Chief Executives
Social and Community Service Managers
Public Relations Managers
Purchasing Managers
Clergy

Appendix Table 8: O*NET elements used in constructing occupational tasks

Task	Title	File	Element
Interaction	Social Perceptiveness	Skills.txt	2.B.1.a
	Coordination	Skills.txt	2.B.1.b
	Persuasion	Skills.txt	2.B.1.c
	Negotiation	Skills.txt	2.B.1.d
Non-routine analytical	Mathematical Reasoning	Abilities.txt	1.A.1.c.1
	Mathematics	Skills.txt	2.A.1.e
	Mathematics	Knowledge.txt	2.C.4.a
Information use	Getting Information	Work Activities.txt	4.A.1.a.1
	Identifying Objects, Actions, and Events	Work Activities.txt	4.A.1.b.1
	Processing Information	Work Activities.txt	4.A.2.a.2
	Analyzing Data or Information	Work Activities.txt	4.A.2.a.4
Deductive and Inductive Reasoning	Written Comprehension	Abilities.txt	1.A.1.a.2
	Deductive Reasoning	Abilities.txt	1.A.1.b.4
	Inductive Reasoning	Abilities.txt	1.A.1.b.5
Number facility	Number Facility	Abilities.txt	1.A.1.c.2
Routine manual	Controlling Machines and Processes	Work Activities.txt	4.A.3.a.3
	Pace Determined by Speed of Equipment	Work Context.txt	4.C.3.d.3
	Spend Time Making Repetitive Motions	Work Context.txt	4.C.2.d.1.i
Routine cognitive	Importance of Repeating Same Tasks	Work Context.txt	4.C.3.b.7
	Importance of Being Exact or Accurate	Work Context.txt	4.C.3.b.4
	Structured versus Unstructured Work	Work Context.txt	4.C.3.b.8

Data Appendix

I. Speech data

A. General

The NLSY97 began as a sample of 8,984 people who were between the ages of 12 and 17 in 1997. Interviews were carried out annually until 2011 and have been carried out biennially since. The main interviews are conducted annually by NORC, a social science research organization affiliated with the University of Chicago that conducts the survey on behalf of the federal Bureau of Labor Statistics.

B. Audio data collection

Audio data were collected during Round 15 of the NLSY97, which was fielded between September 2011 and June 2012. The data were collected in response to two speech prompts, designed to capture both informal and formal speech. The informal prompt was administered at the end of the interview, when respondents were asked to recount the happiest moment (HM) in their life since the date of their last interview. The formal prompt involved a job-search (JS) role-playing exercise. Administered during the employment section of the interview, respondents were asked:

> Let's suppose you applied for a job that sounded really interesting to you and they called you and asked you to come in for an interview. How would you describe your skills, qualifications, and experience to me if I were the person interviewing you for this job? (Employed respondents heard a slightly different preamble to the question.)

All respondents who conducted in-person interviews and who gave consent to be recorded were eligible to be assigned at least one speech prompt. Their responses were

recorded by the on-board microphone in each field interviewer's (FI's) laptop. To make the recording, the CAPI interview software was programmed to turn on the FI's laptop microphone for one minute once a prompt was reached. FI's were provided with instructions designed to keep the respondent talking for as much of that minute as possible.

Because of similarities between AAVE and SoAE dialects, both stimulus questions were assigned to all African-American and Southern white respondents. Southern white respondents are defined as non-Hispanic whites who resided in the South Census region at age 12. Residence at age 12 is provided in the NLSY97. In light of what is known about language acquisition, it would be desirable to have more information about the respondent's residential history as a child. Fortunately, age 12 corresponds at least roughly to the end of the sensitive period for dialect acquisition. Non-southern whites respondents were assigned to only one speech prompt, since few were expected to produce many AAVE or SoAE features. Assignment between the job search and happiest moment prompts was made at random.

A random sample of 500 respondents who were neither black nor Southern white were also to be assigned both speech prompts, as were roughly 295 other respondents for whom speech data was collected in 2006 as part of my earlier study (Grogger 2011) but who were not included in the other categories above. All other speakers, including non-Southern white respondents and all other respondents, were randomly assigned to only one of the speech prompts.

Appendix Table 1 provides data on Round-15 speech-prompt sampling and response rates, disaggregated by race/region at age 12. Of the 8,984 original NLSY97

respondents, 7,423 were interviewed during Round 15. Among those interviews, 6,579 were carried out in-person. Among those, 6,080 provided consent and were thus eligible to be recorded. The share of Round 15 respondents providing in-person interviews and consent to be recorded was .83 for blacks, .80 for both white groups, and .84 for the other group.

The center panel of the Table shows how eligible respondents were assigned to speech prompts. For the most part, the assignments followed the sampling plan fairly closely. All but seven of the black respondents, and all but three of the Southern white respondents, were assigned both questions. Among non-Southern whites and others, 751 respondents were assigned to both stimulus questions. Ten otherwise eligible respondents were not assigned either speech question.

The bottom panel of the Table provides counts of eligible respondents for whom audio files were actually recorded in the interviews. There is a troubling discrepancy between the number of respondents from whom audio data should have been collected and the number from whom it was actually collected. Of the 2484 eligible black and Southern white respondents, audio files were obtained for 2053, corresponding to a rate of loss of 17 percent, compared to 21 percent among non-Southern whites.

The reasons for this loss of data are unclear. I have been assured by NORC staff that this is not a matter of misplaced audio files, but rather, that audio files never existed for the 431 (2484 – 2053) black and Southern white respondents, as well as the 431 (2034 – 1603) non-Southern whites, who were eligible to be recorded but for whom no audio files are available. One possibility is that a flaw in the CAPI interviewing system allowed FIs to skip the recordings. If so, any such skipping would appear to have been

unintentional, since the loss of recordings is widely distributed among FIs, rather than being concentrated among a few.

Technical and budgetary issues further limited the scope of data processing for both HM and JS files. The goals for JS file processing were to maximize the number of blacks and Southern whites for whom both HM and JS data were available, and to maximize the number of non-Southern whites for whom data from at least one of the speech prompts would be available, while meeting the project budget constraint. I thus decided to process all useable files for black and whites, but to sharply curtail processing files for the "other " race group. The bottom panel of Appendix Table 1 shows that 83 percent of the available audio files for black speakers were processed, compared to 92 percent of those for Southern whites and 99 percent of those for non-Southern whites. Speech data from at least one prompt are available for a total of 4,225 NLSY respondents.

Since speech data are unavailable for a sizable share of the sample, it is natural to ask how respondents with speech data compare to respondents without it. Appendix Table 2 provides such a comparison in terms of many variables that appear in the regression analysis. I limit attention to blacks and whites, since they are the focus of this study and since data are limited for the other race group. For blacks and non-Southern whites, respondents with speech data are less likely to have missing information on their current location. Blacks with speech data were significantly more likely to have grown up with two parents, and to have missing maternal education data. Southern whites with speech data were more likely to be married. Non-Southern whites with speech data were

less likely to have missing data about their region of residence at age 12, and more likely to be married. They also had higher AFQT scores.

C. Producing numerical data from the audio files

To generate data suitable for the regression analysis, I recruited anonymous listeners to listen to the audio files and answer questions about the speakers. After listening to each audio file, listeners were asked to specify the speaker's sex, race/ethnicity, and region of origin. Three listeners were assigned to each audio file. Thus speakers who responded to both the HM and JS prompts have six listener reports, whereas speakers who responded to only one of the prompts have three. To deal with data security issues surrounding the use of potentially identifiable voice data, listeners were recruited from the pool of NORC FIs and research assistants. Data processing was carried out remotely using specially configured laptops that provided secure connections to NORC's computer network, where the audio files resided. All listeners received confidentiality training stipulated by both NORC and BLS.

Summary characteristics of the listeners are reported in Appendix Table 3. The modal listener was white and female, reflecting the demographics of NORC's workforce. Listeners were drawn from throughout the US, with a disproportion of Midwesterners. All listeners had completed high school; most had at least some tertiary education. The 11 listeners who listened to the JS audio files tended to be older, more Southern, and less educated than the 36 listeners who listened to the HM audio files (10 listened to both).

Care was taken to ensure that speakers were not assigned to listeners who had interviewed them during Round 15.[1]

II. Occupational task data

I use data from version 19.0 of the Occupational Information Network to measure tasks performed in various occupations. I create the following composite variables based on that data. The first five are from Deming (2015b) and the last two are from Autor and Handel (2012). Appendix Table 9 provides information on item numbers and file locations.

(1) Interaction. This is constructed from four items measuring: social perceptiveness, coordination, persuasion, and negotiation. Deming (2015) refers to this as social skills.

(2) Non-routine analytical skills. Involves three items measuring mathematical reasoning, mathematical knowledge, and mathematical skills.

(3) Information use. Four items about getting information; identifying objects, actions, and events; processing information; analyzing data or information.

(4) Inductive/deductive reasoning. Three items involving written comprehension, deductive reasoning, and inductive reasoning.

(5) Number facility. One item involving facility with numbers.

(6) Routine manual tasks. Three items on controlling machines and processes, pace determined by speed of equipment, and spend time making repetitive motions.

(7) Routine cognitive tasks. Three items on importance of repeating same tasks, importance of being exact or accurate, structured versus unstructured work.

[1] Listener reports of the speaker's race and region varied by characteristics of the listener. However, listener characteristics do not have much effect on the estimated relationship between speech and wages, as I demonstrate in the paper.

These variables are all reported in terms of ordinal scales ranging from 1 to 7, except for "importance of repeating same tasks," which ranges from 1 to 5. The "structured versus unstructured work " is reverse-coded. The first step in constructing the composites used in the paper was to sum the elements and standardize them.

In the O*NET files there is one record for each so-called O*NET Standard Occupational Classification (ONET SOC) codes. Occupations in the NLSY97 are identified by 2002 Census Occupation codes. To link the NLSY data to the O*NET data, I first cross-walked the ONET SOC codes to standard 2010 SOC codes, then cross-walked the 2010 SOC codes to 2000 SOC codes, then cross-walked the 2000 SOC codes to the 2002 Census codes.[2] This resulted in the loss of two 3-digit Census occupations that appear in the estimation sample: Truck transportation (617) and Armed Forces (984: rank unspecified, last job)[3]. These occupations accounted for 1.13 percent of the observations in the wage sample. NLSY respondents whose occupation was recorded as unspecified (999) account for 0.67 percent of observations in the sample. Once the cross-walk was completed, I re-standardized the scales and constructed indicators for the occupations in the top quartile of each.

[2] The respective crosswalk files were obtained from http://www.bls.gov/emp/ep_crosswalks.htm, http://www.bls.gov/soc/soccrosswalks.htm, and http://www.xwalkcenter.org/index.php/classifications/crosswalks.

[3] http://www.bls.gov/tus/census02iocodes.pdf