

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: Variable	Black		Southern white		Non-Southern white	
	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:		HM	JS
Characteristic		(1)	(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

Speaker's race/region	L reports that speaker is black		L reports that speaker is Southern	
	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: Log wage regressions with additional regressors

	(1)	(2)	(3)	(4)	(5)	(6)
Black*mainstream speech	0.120 (0.043)	0.108 (0.043)	0.104 (0.040)	0.118 (0.044)	0.124 (0.043)	0.115 (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: Log wage regressions, deleting observations below various minimum age thresholds

Dependent variable is the log hourly wage				
Minimum age:	(1)	(2)	(3)	(4)
	25	26	27	28
Black*mainstream speech	0.143 (0.047)	0.166 (0.049)	0.163 (0.052)	0.163 (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 regression from column (2) of Table 4.

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

---

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

---

Note: Numbers in parentheses are row percentages.

Appendix 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 ( 0.240)	0.529 ( 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 9: 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 10: 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.<sup>1</sup>

## **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 (2017) 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 (2017) 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.

---

<sup>1</sup> 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.



(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.

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.<sup>2</sup> This resulted in the loss of two 3-digit Census occupations that appear in the estimation sample: Truck transportation (617) and Armed Forces (984:

---

<sup>2</sup> The respective crosswalk files were obtained from [http://www.bls.gov/emp/ep\\_crosswalks.htm](http://www.bls.gov/emp/ep_crosswalks.htm), <http://www.bls.gov/soc/soccrosswalks.htm>, and <http://www.xwalkcenter.org/index.php/classifications/crosswalks>.

rank unspecified, last job)<sup>3</sup>. 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.

---

<sup>3</sup> <http://www.bls.gov/tus/census02iocodes.pdf>