Online Appendix for "Task-Based Discrimination"
by Erik Hurst, Yona Rubinstein, and Kazuatsu Shimizu

Appendix A Data Description

In our empirical work, we primarily use data from three sources: cross-sectional labor market data from the Census/ACS, occupational task measures from DOT and O*Net, and panel micro data from the NLSY79 and NLSY97 that contain measures of worker pre-labor market skills.

Appendix A.1 Census/ACS Sample

To access the Census/ACS data, we download the micro data directly from the IPUMS USA website (Ruggles et al. (2021)). We use data from the 1960, 1970, 1980, 1990, and 2000 US Censuses. Additionally, we pool together data from the 2010-2012 and the 2016-2018 American Community Surveys. We refer to the former as the 2012 ACS sample and the latter as the the 2018 ACS. We restrict our Census and ACS samples to those between the ages of 25 and 54 (inclusive), those who report their race as "White" (race = 1) or "Black" (race = 2), and those born within the United States. We exclude from our sample anyone who is living in group quarters (keep gq = 1) and those who are self employed (classwkr = 2). Finally, we exclude any employed worker whose occupation has missing task values. This last restriction reduces the overall sample by less than one percent.

Appendix A.2 NLSY Data

Our primary data sources for individual skills are the 1979 and the 1997 National Longitudinal Survey of Youth, NLSY79 and NLSY97, respectively. The NLSY79 is a representative survey of 12,686 individuals who were 15-22 years old when they were first surveyed in 1979. Individuals were interviewed annually through 1994 and biennially since then. The NLSY97, which follows a nearly identical structure to the NLSY79, is a nationally representative panel survey of 8,984 individuals who were 12-16 years old when they were first surveyed in 1997. Individuals were interviewed annually through 2011 and biennially since then.

The NLSY79 and the NLSY97 waves provide detailed demographic information, such as age, gender, race, and educational attainment. The files also contain measures of cognitive ability, personality traits, and sociability. We follow a large body of research, including Neal and Johnson (1996), Heckman et al. (2006), Altonji et al. (2012) and Deming (2017), and aggregate these measures into three categories (i) cognitive, (ii) non-cognitive, and

(iii) social skills. These measures are taken directly from (Deming, 2017). Specifically, we downloaded these variables from Deming's replication files at https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/CYPKZH.

Cognitive skills are proxied using the Armed Forces Qualifications Test (AFQT). This measure is available for both the NLSY79 and the NLSY97 waves. Altonji et al. (2012) developed a mapping of the AFQT score across the NLSY79 and NLSY97 waves that accounts for differences in age-at-test and test format. Deming (2017) normalized these to have mean zero and standard deviation one. We use his measures for all of our analysis.

While both the NLSY79 and the NLSY97 include AFQT scores, these waves contain different measures of non-cognitive and social traits. Deming (2017) provides a set of unified measures of non-cognitive and social skills which we adopt. Specifically, the Deming definition for non-cognitive skills uses (i) the normalized average of the Rotter Locus of Control and the Rosenberg Self-Esteem Scale (see Heckman et al. (2006)) in the NLSY79 wave and (ii) the Big-5 factor conscientiousness, normalized and standardized, in the NLSY97 wave. The Deming definition for social skills uses (i) an average of four self-reported normalized and standardized measures, including sociability at age 6, sociability in 1981, number of clubs each respondent participated in high school, and participation in high school sports in the NLSY79 wave and (ii) an average of two items, normalized and standardized, that capture the extroversion factor from the Big-5 personality test in the NLSY97 wave.

We restrict our primary sample to Black and White men only. We exclude observations with missing demographics or missing measures of cognitive, non-cognitive, or social skills. Our wage and employment sample focuses on prime-aged male who are full-time and full-year workers. We exclude observations that report less than 1,750 annual worked hours or hourly wages lower than 2 or higher than 500 in 2010 CPI prices. We further exclude observations with missing occupation codes. When comparing over time and across cohorts of birth, we restrict the NLSY79 sample to individuals aged 25-37 for comparability to the NLSY97 wave.

Appendix A.3 Task Measures Creation

To assess the extent to which Black and White workers sort into different occupations, perform different tasks and consequently earn different amounts, we use data from the following to measure the skills demanded in each occupation: (i) the U.S. Department of Labor's Dictionary of Occupational Titles (DOT) and (ii) the Occupational Information Network (O*NET) sponsored by the U.S. Department of Labor/Employment and Training Administration (USDOL/ETA). The DOT was constructed in 1939 to help employment offices match

job seekers with job openings. It provides information on the skills demanded of over 12,000 occupations. The DOT was updated in 1949, 1964, 1977, and 1991, and replaced by the O*NET in 1998.

The DOT and the O*NET measure task requirements associated with many detailed occupations. For example, one O*Net question asks whether the occupation requires dealing with external customers - survey respondents provide responses on an ordinal scale of 0 to 5 where the higher values signify that the job requires more of that task. Different questions have answers that range on different ordinal scales (e.g., 0-5, 1-7, 0-10, etc.). We again downloaded the tasks measures directly from the replication package for Deming (2017). For all questions we use from both surveys, we follow Deming (2017) and re-scale the answers so they range from zero to ten to ensure consistency in units when we combine questions. We convert the answers into z-score units after combining them into different tasks.

We focus on four occupational task measures that are relevant for our study: (i) *Abstract*; (ii) *Routine*; (iii) *Manual* and (iv) *Contact*. The first three measures were created following the definitions in Autor and Dorn (2013) using the DOT data while the last measure builds on Deming (2017) using the O*Net data. Our goal is to stay as close to possible to the definitions of task measures developed by others to focus our analysis on the racial differences in these measures. Throughout the main paper, we define the key task measures as follows:

Abstract: indicates the degree to which the occupation demands (i) analytical flexibility, creativity, reasoning, and generalized problem-solving, and (ii) complex interpersonal communications such as persuading, selling, and managing others. Following Dorn (2009) and Autor and Dorn (2013), we measure *Abstract* tasks in practice by using the 1977 DOT data using the average scores from questions measuring *General Educational Development in Math (GED-Math)* and *Direction, Control, and Planning of Activities (DCP)*. Higher levels of *GED-Math* are associated with higher quantitative abstract tasks. Occupations with high measures of *GED Math* include various medical professionals, various engineers, accountants, and software developers. Higher levels of *DCP* are associated with higher levels of abstract thinking associated with management, organizational, and teaching tasks. Occupations with high measures of *DCP* include various managers, high school teachers, college professors and judges. To create our measure of the *Abstract* task content of an occupation, we follow Autor and Dorn (2013) and Deming (2017) and take the simple average of *GED-Math* and *DCP* for each occupation.

Routine: measures the degree to which the task requires the precise attainment of set standards and/or repetitive manual tasks. Following Dorn (2009) and Autor and Dorn (2013), we measure *Routine* task using the 1977 DOT data taking the average scores from questions measuring *Finger Dexterity (FINGDEX)* and *Set Limits, Tolerances, or Standards*

(STS). *FINGDEX* measures the ability to move fingers and manipulate small objects with fingers and serves as a proxy for repetitive routine manual tasks. Occupations with high measures of *FINGDEX* include secretaries, dental hygienists, bank tellers, machinists, textile sewing machine operators, dressmakers, and x-ray technology specialists. *STS* measures the adaptability to work situations requiring setting of limits and measurements and serves as a proxy for routine cognitive tasks. Occupations with high measures of *STS* include meter readers, pilots, drafters, auto mechanics, and various manufacturing occupations. To create our measure of the *Routine* task content of an occupation, we follow Autor and Dorn (2013) and Deming (2017) and take the simple average of *FINGDEX* and *STS* for each occupation.

Manual: measures the degree to which the task demands eye, hand, and foot coordination. Following Dorn (2009), Autor and Dorn (2013) and and Deming (2017), we measure *Manual* using the 1977 DOT data using the question *EYEHAND* which measures the ability to coordinately move hand and foot in accordance with visual stimuli. Occupations with high measures of *EYEHAND* include athletes, police and fire fighters, drivers (taxi, bus, truck), skilled construction (e.g, electricians, painters, carpenters) and landscapers/groundskeepers. To create our measure of the *Manual* task content of an occupation, we just use the *EYEHAND* measure for that occupation.

Contact: measures the extent that the job requires the worker to interact and communicate with others whether (i) within the organization or (ii) with external customers/clients or potential customers/clients. For this measure of *Contact* tasks we use two 1998 O*NET work activity variables taken from Deming (2017). Specifically, we use the variables *Job-Required Social Interaction (Interact)* and *Deal With External Customers (Customer)*.[34] *Interact* measures how much workers are required to be in contact with others in order to perform the job. *Customer* measures how much workers have to deal with either external customers (e.g., retail sales) or the public in general (e.g., police work). To make our measure of the *Contact* task content of an occupation, we take the simple average of *Interact* and *Customer* for each occupation. Occupations with high measures of *Contact* tasks include various health care workers, waiter/waitress, sales clerks, lawyers, various teachers, and various managers.

[34] Deming (2017)'s focus is creating a measure of occupational tasks that require social skills and document how the returns to social skills have increased over time. His measure of social skills include measures of whether the job requires the worker to have social perceptiveness and the ability to coordinate, persuade and negotiate with others. For example, his measure of social skills includes O*NET questions assessing whether the job requires "adjusting actions in relation to others' action", "being aware of others' reactions and understanding why they react the way they do", and "persuading others to approach things differently". His measure of social skills do not include measures for whether the task requires interactions with other co-workers or customers. He uses the measures of customer (*Customer*) and broader social interactions (*Interact*) as controls in some of his specifications. These questions are much more suited to our purpose of trying to measure taste-based discrimination. We explore the relationship between Deming's *Social Skills* task measure and our *Contact* task measure in the online appendix.

The data we use from Deming (2017) are available at the 3-digit occupational code level. We use Deming (2017)'s crosswalk to merge these measures to (i) the Census and the American Community Surveys (ACS) and (ii) the National Longitudinal Survey of the Youth (NLSY 1979 and 1997 waves) which we use for our analysis. Again, we download these data directly from Deming's replication file at https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/CYPKZH.

Appendix B Robustness of Racial Task Gaps: Alternate Specifications

In this section of the appendix, we show the robustness of our results pertaining to racial task gaps. We start by showing the raw task trends separately for Black and White men (in the main text, we only show the racial gaps). We then show the robustness of the racial task gaps documented in the main text to an alternate specifications. Finally, we show the racial task gaps separately for different education groups.

Appendix B.1 Occupational Task Sorting, Separately by Race

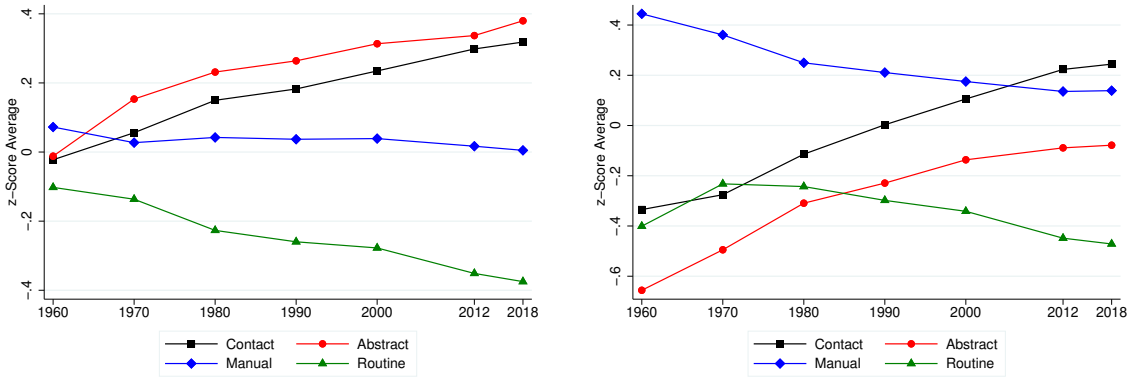
Appendix Figure A1 plots the raw trends in occupational tasks separately for White (Panel A) and Black (Panel B) men since 1960 using the Census/ACS data. As in the main text, we restrict our sample to native born men between the ages of 25 and 54 who are not self employed and who report currently working full time (e.g., at least 30 hours per week). Specifically, Appendix Figure A1 reports the coefficients on the year dummies (β_t) from the following regressions using our individual Census/ACS data:

$$\tau^k_{ijgt} = \sum_t \beta^g_t D_t + \epsilon_{ijgt} \tag{A1}$$

where τ^k_{ijgt} is the task content of task k for individual i from group g working in occupation j in period t. Task contents are expressed in z-score units. We run this regression separately for two groups g: White men and Black men. As a result, all coefficients have g superscripts. We explore the four tasks k highlighted in the main text. D_t is a vector of dummies that take the value of 1 if the year is, respectively, 1960, 1970, 1980, 1990, 2000, 2012, or 2018. The coefficient on the year dummies from these regressions, β^g_t are plotted in the figure.

A few things are noticeable from Figure A1. First, focusing on Panel A, there were large shifts in the task content of occupations to which White men sorted into over the last six decades. During this time period, White men have sorted more into *Abstract* occupations and

Figure A1: Raw Task Trends: White and Black Men



Panel A: White Men

Panel B: Black Men

Notes: Figure shows the raw trend in the task content of jobs for White and Black men using Census and ACS data. Sample restricted to native born individuals between the ages of 25 and 54 who are not self-employed but who are working full time. Tasks are expressed as z-scores across occupations. Regressions are run separately for White men (Panel A) and Black men (Panel B) and were weighted using Census/ACS individual sampling weights.

sorted away from *Routine* occupations. Likewise, White men have sorted into occupations which require meaningful *Contact* tasks during the last half century. These patterns mimic the patterns for *Abstract*, *Routine* and *Manual* tasks highlighted in Dorn (2009), Autor and Dorn (2013), and Deming (2017). This is not surprising given we are using their task measures in this analysis. The main text highlights the racial gaps in these measures by comparing the patterns between Panels A and B with and without additional demographic controls.

Appendix B.2 Alternate Specification of Racial Task Gaps

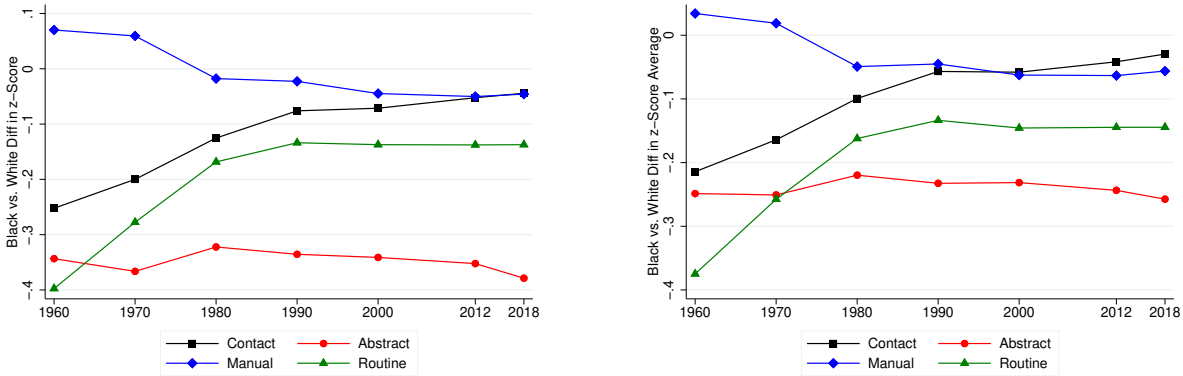
We next provide an alternate specification of racial gaps in the task content of occupations. This is an alternate way to display the data shown in Figure 2 of the main text. Particularly, we estimate the following:

$$\tau_{ijt}^{k} = \alpha_t^k + \beta_t^k Black_{it} + \sum_{s \neq k} \omega_{st}^k \tau_{ijt}^s + \Gamma^k X_{it} + \epsilon_{ijt}^k \tag{A2}$$

where τ_{ijt}^k, $Black_{it}$ and X_{it} are defined as in the main text. In this specification, we isolate the partial effect of the racial gap in the task content of occupations by controlling directly for the other tasks in the regression ($\sum_{s \neq k} \tau_{ijt}^s$). We run this regression separately for each

year and for each task yielding 28 estimates of β_t^k. These coefficients are plotted in Appendix Figure A2. Panel A shows the results excluding the X vector of demographic controls while Panel B shows the results including the additional controls. The racial gaps are expressed in z-score units.

Figure A2: Race Gap in Tasks Without and With Demographic Controls, Alternate Specification



Panel A: No Demographic Controls Panel B: With Demographic Controls

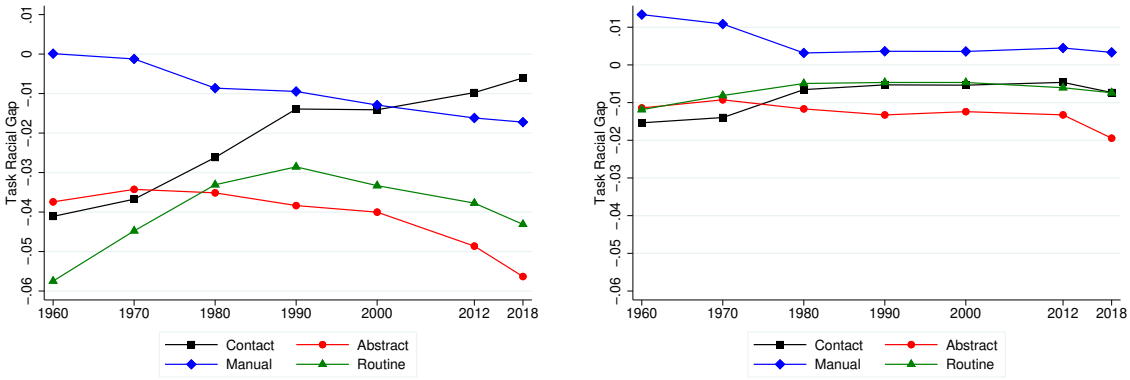
Notes: Figure shows the racial gap in task trends using our alternate specification in the appendix without (Panel A) and with (Panel B) demographic controls. Sample restricted to native born individuals between the ages of 25 and 54 who are not self-employed but who are working full time. Tasks are expressed as z-scores across occupations.

This figure is the analog to Figure 2 in the main text. The time series patterns from this specification are nearly identical to the patterns in the main text. The only difference is the units in which the racial task gaps are expressed. The racial gap in *Abstract* tasks was essentially constant between 1960 and 2018 at a level a 0.25 standard deviation gap (with controls). However, the racial gap in *Contact* tasks converged from roughly a 0.25 standard deviation gap in 1960 to a close to zero gap in 2018.

Appendix B.3 Racial Task Gaps, by Education Levels

Finally, we show robustness of the time series patterns in racial task gaps within different education groups using our main specification described in the text. Panel A of Appendix Figure A3 redoes the main results of Figure 2 of the main text (with demographic controls) but segmenting the sample to only those individuals with education less than a bachelor's degree. Panel B shows the same specification but restricting the sample to those individuals with a bachelors degree or more. These figures shed light on whether our time series patterns

Figure A3: Race Gap in Tasks: By Educated Groups



Panel A:
Less than a Bachelors Degree

Panel B:
Bachelors Degree of More

Notes: Figure re-estimates Panel B of Figure 2 of the main text separately by those with less than a bachelors degree (Panel A) and those with a bachelors degree or more (Panel B).

of the changing racial task gaps that we highlight in the main paper are found in both higher and lower education samples. For both education groups, there was a convergence in *Contact* tasks and a slight divergence in *Abstract* tasks. The magnitude of the *Contact* convergence is much larger for less educated individuals, but given selection (Panel A represents between 70 and 75 percent of the sample depending on the year), it is not surprising that the convergence in *Contact* tasks is smaller for higher educated individuals.

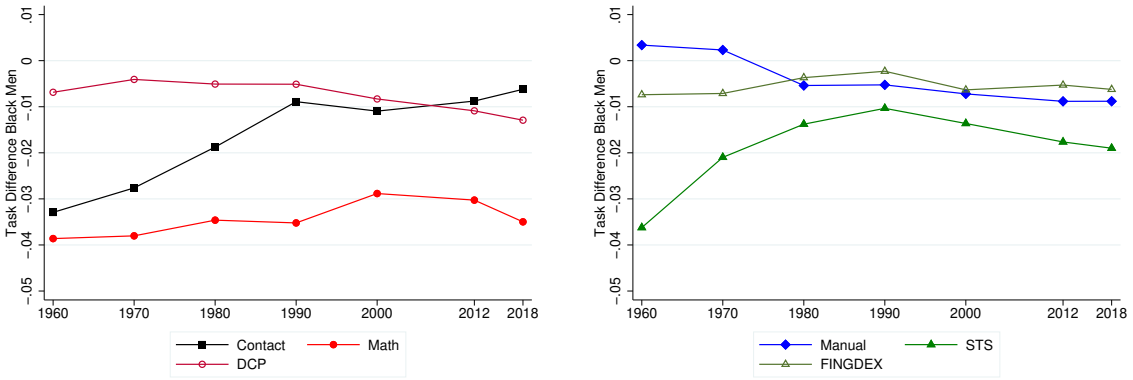
Appendix C Robustness of Racial Task Gaps: Alternate Task Definitions

In this section, we explore the robustness of our results to alternate task definitions. We begin by disaggregating our current task measures into their separate task components. We then explore the racial gaps in alternate definitions of four main task categories. Finally, we compare our *Contact* tasks measure to Deming's *Social* task measure. As seen in this section, our results are quite robust to alternate task definitions.

Appendix C.1 Decomposing Task Measures into Sub-Components

We used three task measures emphasized in the recent literature using DOT data: *Abstract*, *Routine* and *Manual* tasks. As discussed above, these three measures of tasks were created using five separate questions from the DOT data. *Abstract* task is a combination of $GED-$

Figure A4: Race Gap in Tasks: Disaggregated Task Measures



Panel A Panel B

Notes: Figure re-estimates Panel B of Figure 2 of the main text with six task components instead of four. In particular, we disaggregate *Abstract* tasks into its (1) *Math* and (2) DCP sub-components. Likewise, we disaggregate *Routine* tasks into its (1) STS and (2) Finger subcomponents.

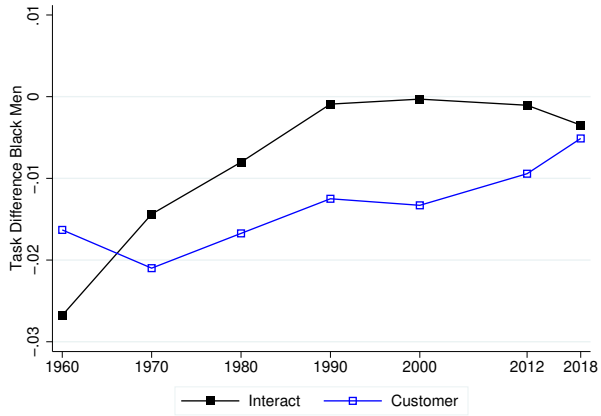
Math and *DCP*. *Routine* task is a combination of *FINGDEX* and *STS*. In this subsection of the appendix, we move from using four tasks measures (*Abstract*, *Routine*, *Manual*, and *Contact*) to six tasks measures (*GED-Math*, *DCP*, *FINGDEX*, *STS*, *Manual* and *Contact*). In particular, we re-estimate the results in Panel B of Figure 2 using six task measures instead of four. The sample used is the same as in Panel B of Figure 2 of the main text. The coefficients on the task measures from these yearly regressions are plotted in Appendix Figure A4. We plot the coefficients in two panels instead of one for readability.

The figure shows that the main take-aways highlighted in the text are unaltered when using the six task measures. Specifically, there have been no relative gains by Blacks with respect to either component of *Abstract* tasks; Blacks were underrepresented in both *GED Math* and *DCP* in 1960 and the race gap was constant through 2018. However, Blacks made large gains in *Contact* tasks over this time period.

Appendix Figure A5 shows the results from the regression but with seven tasks measures. We still include *GED-Math*, *DCP*, *FINGDEX*, *STS* and *Manual*. But, we now disaggregate *Contact* into its two sub-components: *Interact* and *Customer*. The former measures the extent to which the job requires social interactions with others while the latter measures whether the job requires individuals to deal with external customers. Instead of showing all seven coefficients, we only show the coefficients on *Interact* tasks and *Customer* tasks.[35] There was racial convergence in both tasks requiring contact within the firm (*Interact*) and tasks requiring contact with external customers (*Customer*). These results highlight that

[35]The coefficients on the other five tasks were essentially unchanged relative to Appendix Figure A4.

Figure A5: Race Gap in Disaggregated *Contact* Task Measures



Notes: Figure re-estimates Panel B of Figure 2 of the main text with seven task components instead of four. In particular, we disaggregate *Abstract* tasks into its (1) *Math* and (2) *DCP* sub-components. Likewise, we disaggregate *Routine* tasks into its (1) STS and (2) *FINGDEX* subcomponents. Finally, we disaggregate *Contact* tasks into (1) *Interact* and (2) *Customer* sub-components. Only the coefficients on the *Interact* and *Customer* task measures from these yearly regressions are plotted in the figure.

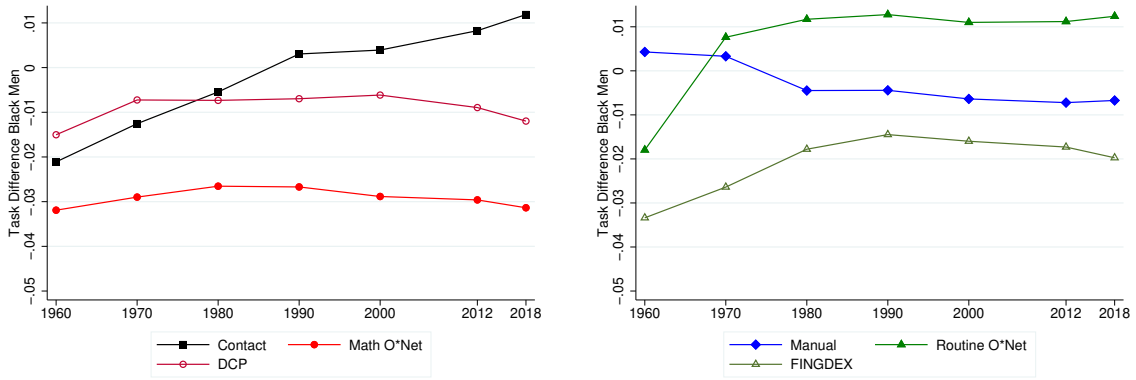
Blacks were moving into occupations (relatively) that require both forms of contact with others.

Appendix C.2 Robustness to O*Net Measures of *Math* and *Routine* Tasks

Deming (2017) used data from 1998 O*Net survey to make two alternate measures of *Math* and *Routine* occupations. For his alternate *Math* task measure, he combines O*Net questions measuring (i) the extent to which an occupation requires mathematical reasoning, (ii) whether the occupation requires using mathematics to solve problems, and (iii) whether the occupation requires knowledge of mathematics. The measure of the *GED-Math* task content of an occupation created using DOT data is highly correlated with Deming's *Math* task content of an occupation created using the O*Net data; the correlation between the two series (weighted by 1990 population in each occupation) is 0.81.

For his alternate *Routine* task measure, Deming again uses the 1998 O*Net and combines the questions measuring (i) how automated is the job and (ii) how important is repeating the same physical activity (e.g. key entry) or mental activities (e.g., checking entries in a ledger over and over, without stopping to perform the job). This measure is highly correlated with the *STS* portion of *Routine* tasks within the DOT data. However, conditional on controlling for the *STS* content of a job, the Deming *Routine* task measure using the O*Net data is

Figure A6: Race Gap in Tasks: Alternate Measures of *Routine* and *Math* Task Measures



Panel A Panel B

Notes: Figure re-estimates Panel B of Figure 2 of the main text with six task components instead of four. In particular, we disaggregate *Abstract* tasks into its (1) *Math* and (2) *DCP* sub-components. For this figure, we use Deming's measure of occupational *Math* task measures using the O*Net data. Likewise, we disaggregate the DOT *Routine* tasks into its (1) *STS* and (2) *Finger* subcomponents. However, we replace the DOT *STS* measure with Deming's *Routine* task measure using O*Net data.

uncorrelated with the occupations *FINGDEX* task content.[36] Given this, we treat Deming's *Routine* task measure created using the 1998 O*Net data as being an alternative for the *STS* task measure within the DOT data.

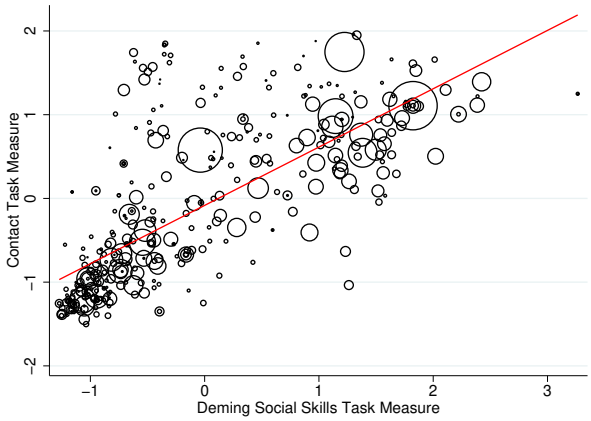
With this in mind, we explore the sensitivity of our results to using Deming's *Math* and *Routine* measure using the O*Net data as alternative task measures for the *GED-Math* and *STS* measures using the DOT data. We re-estimate the patterns in Appendix Figure A4 with the six task measures but we use the alternate Deming measures for *Math* and *STS*. The results of this regression are shown in Appendix Figure A6. Again, we display the results over two panels for readability. Our main results are unchanged with these two alternative task measures. Primarily, there has still been no racial progress in the *Math* task content of an occupation over the last 60 years. However, there have been a large convergence in the racial gap in occupational *Contact* tasks.

Appendix C.3 Alternate Measures of *Contact* Tasks

The key finding in our paper is the racial convergence in *Contact* tasks relative to *Abstract* tasks in the U.S. over the last half century. In this sub-section, we explore the sensitivity of

[36]Regressing the Deming *Routine* task content of an occupation on the occupation's *STS* and *FINGDEX* task content (weighted by 1990 population counts in each occupation) yields a coefficient on STS of 0.50 (standard error = 0.05) and a coefficient on *FINGDEX* of -0.06 (standard error = 0.06).

Figure A7: Correlation Between *Contact* Task and Social Task, Cross-Occupation Variation



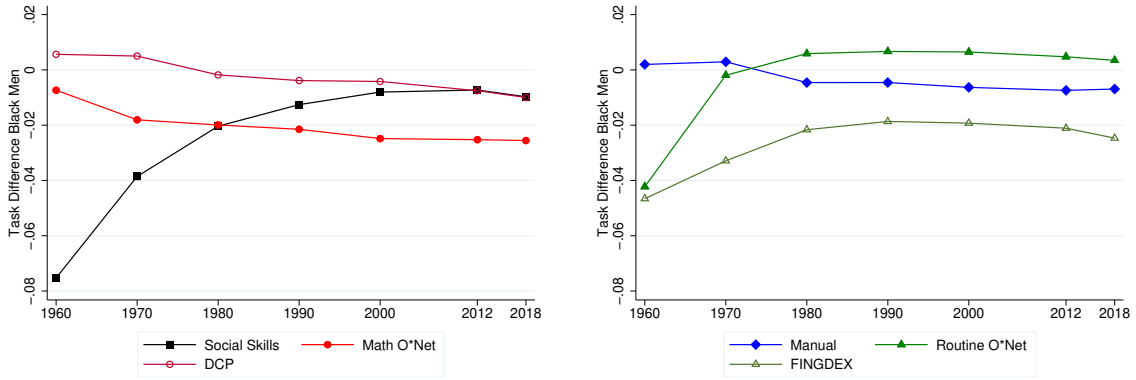
Notes: Figure shows a scatter plot of the correlation between the *Contact* task content of an occupation and Deming's *Social Skills* task content of an occupation. Each observation in the figure is an occupation. *Contact* and *Social Skills* tasks are measured in z-score space. The size of the circle represents the number of prime age men working in that occupation in 1990. Figure also includes the weighted simple regression line through the scatter plot. The coefficient on the z-score for *Social* tasks is 0.70 (standard error = 0.03) and an adjusted R-squared of 0.65.

our results to using other measures of *Contact* tasks. Deming's *Social Skills* task measure is highly correlated with our *Contact* task measure. This is not surprising given that Deming's measure of *Social Skills* tasks measures whether the occupation requires skills associated with the ability to coordinate, negotiate, and persuade. The ability to coordinate, negotiate, and persuade is needed when the job requires workers to come into contact with other co-workers, clients and customers. The simple correlation between Deming's *Social Skills* task measure and our *Contact* task measure is about 0.7 (weighted by 1990 population counts within in each occupation). We show the simple scatter plot by occupation of the two measures in Appendix Figure A7.

Appendix Figure A8 is the analog to Appendix Figure A6 except we replace our *Contact* task measure with Deming's *Social Skills* task measure. As highlighted in Deming (2017), the *Social Skills* task content of an occupation is highly correlated with the *Math* and the *DCP* task content of an occupation. As a result, the racial gap in *Abstract* Skills is smaller and the racial gap in *Social Skills* is larger in 1960. Despite that, our key patterns remain. There was a substantial narrowing of the racial gap in the *Social Skills* task content of an occupation since 1960. When we use this measure, there is also a slight divergence in the task content of the two components of *Abstract* skills. Black men are gaining relative to White men not because of a convergence in *Abstract* tasks but a convergence in tasks that require interactions with others.

As one final robustness exercise, we created a combined *Contact*/Social task measure by

Figure A8: Race Gap in Tasks: Social Skills Tasks



Panel A Panel B

Notes: Figure re-estimates Appendix Figure 2 with six task measures further replacing Deming's *Social Skills* task measure for our *Contact* task measure. As with Appendix Figure A6, we use Deming's measure of occupational *Math* and *Routine* task measures along with our measures of DCP, FINGDEX, and *Manual* tasks.

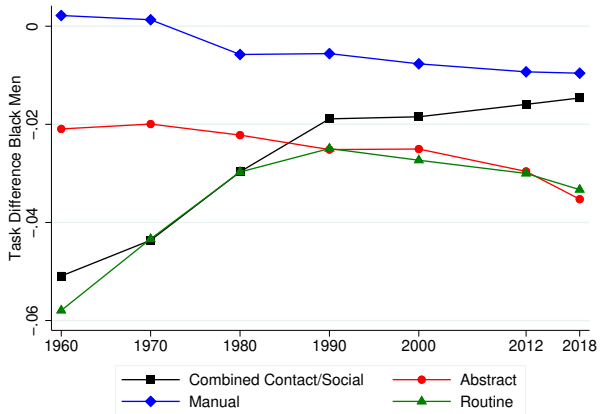
taking the simple average of our *Interact* task measure, our *Customer* task measure, and Deming's *Social Skills* task measure for each occupation. Appendix Figure A9 shows the analog of Panel B of Figure 2 of the main text but replacing our main *Contact* task measure with the combined *Contact*/Social Skills task measure. As seen from the figure, our key findings remain with this broader measure of the social interactions needed in occupations.

Appendix D Task Gaps by Gender

Appendix Figure A10 shows the occupational task differences between White men and White women (panel A) and between White women and Black women (panel B) using data from the Census/ACS. This figure uses the same specification as Panel B of Figure 2 in the main text. Panel A of this appendix figure restricts the sample to native born White men and White women between the ages of 25 and 54. Panel B restricts the sample to native born White women and Black women between the ages of 25 and 54. Both panels also restrict the sample to those individuals working full time and excludes the self-employed. As with the figures in the main text, we condition on education and age when we measure the gaps in the task content of jobs.

As seen from Panel A, women are much more likely to be in *Contact* and *Routine* tasks and are much less likely to be in *Manual* and *Abstract* tasks. Unlike the gaps between Black and White men, the gaps between White men and White women were fairly stable over the

Figure A9: Predicted Race Gap in Task Content, with Combined *Contact*/Social Skills Task Measure



Notes: Figure is analogous to Panel B of Figure 2 in the main text. The only differences is that in this figure we combine Deming's *Social Skills* task measure and our *Contact* task measure into one combined task measure.

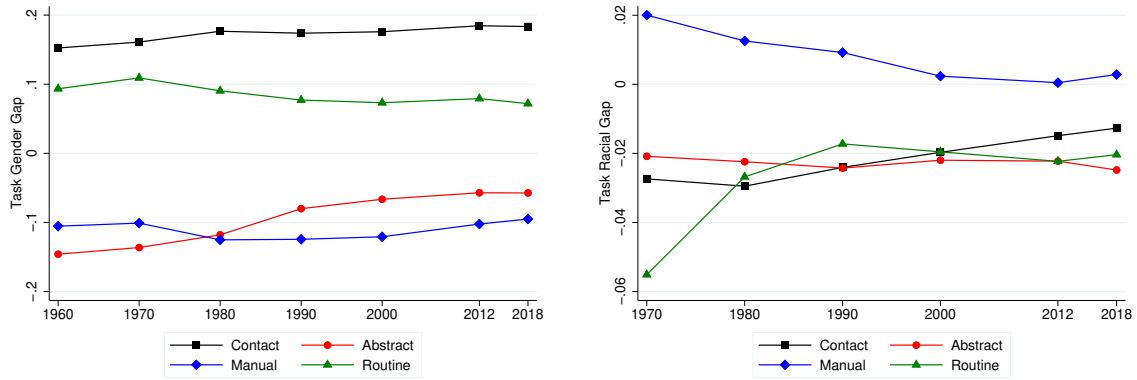
last 60 years. One exception is the gap in *Abstract* tasks. In the 1960, White women were 16 percentage points less likely to work in occupations that require 1 standard deviation higher *Abstract* tasks relative to White men (conditional on age and education). By 2018, that gap fell to 7 percentage points.

The time series patterns in Panel B between White women and Black women mirror the patterns in Panel B of Figure 2 of the main text showing differences between White men and Black men although the level gaps are smaller. The gap in the *Abstract* task content of jobs between White and Black women was roughly constant between 1960 and 2018. However, Black women converged to White women in the *Contact* task content of jobs over this period.

Appendix E The Relationship Between Log Wages and Skills, NLSY Data

In this section of the appendix, we assess the extent to which there are racial differences in the responsiveness of wages to individual skill measures. The coefficients in Panel A of Appendix TableA1 comes from a regression of log individual wages of NLSY respondents on NLSY cognitive, non-cognitive and social skill measures and those skill measures interacted with a race dummy. The regression also includes age, education and occupation fixed effects and those fixed effects interacted with a race dummy. Finally, the regression also includes year and NLSY97 sample fixed effects and those fixed effects interacted with a race dummy. For this regression, we pool together both the NLSY79 and NLSY97 samples. As with the

Figure A10: Task Differentials between White Men and White Women and between White Women and Black Women



Panel A: White Men vs White Women

Panel B: White Women vs Black Women

Notes: Figure shows the extent to which the task content of an occupation can predict whether an individual employed in that occupation is a White woman (Panel A) or a Black woman (Panel B). For the regressions in Panel A, we use the Census/ACS sample pooling together prime-age White men and women. Panel shows the coefficients from a regression of a dummy variable equal to one if the individual is a White woman on the four task measures and controls for individual education, age and Census division, separately by year. For the regressions in Panel B, we use the Census/ACS sample pooling together prime-age White women and Black women. Panel shows the coefficients from a regression of a dummy variable equal to one if the individual is a Black woman on the four task measures and controls for individual education and age, separately by year. All samples for both regressions are also restricted to full time workers who are not self employed and who are native born.

rest of the paper, we only include in our sample Black and White men between the ages of 25 and 54.

Panel A of the table highlights the labor market returns to cognitive, non-cognitive and social skills for White men (first three rows). As seen from the table, having more of any of the three skill measures raises labor market earnings for White men (even conditional on individual education and occupation). Furthermore, we find no differential labor market returns for Black men for non-cognitive and social skills. However, similar to the findings in Neal (2006), the coefficient on AFQT in a regression of log wages on AFQT scores is larger for Blacks than for Whites. This is consistent with the conjecture that Black men who receive the same AFQT test score relative to White men (conditional on education and occupation) may be positively selected in traits not measured in the NLSY that are rewarded in the labor market.

Panel B of the regression runs a similar regress as Panel A but replaces the three NLSY skill measures with individual measures of *Abstract* and *Contact* skills using the prediction

Table A1: Racial Differences in the Relationship between Log Wages and Skill Measures, Model vs. Data

Panel A: NLSY Skill Measures	
Cognitive	0.070
	(0.009)
Non-Cognitive	0.034
	(0.007)
Social	0.014
	(0.007)
Black $*$ *Cognitive*	0.037
	(0.015)
Black $*$ *Non-Cognitive*	-0.004
	(0.011)
Black $*$ *Social*	0.001
	(0.011)
Panel B: Model Imputed Skill Measures	
Abstract	0.990
	(0.121)
Contact	0.459
	(0.100)
Black $*$ *Abstract*	0.474
	(0.217)
Black $*$ *Contact*	-0.122
	(0.164)

Note: Panel A of table shows key coefficients from a regression of log wages on a cognitive, non-cognitive, and social skills and those skill measures interacted with a Black dummy using the NLSY micro data. Panel B of table shows key coefficients from a regression of log wage on model predicted Abstract and Social skills and those skill measures interacted with a Black dummy using NLSY micro data. All regressions include controls for individual age, education, and occupation and those controls interacted with a race dummy. Additionally, the regression also includes year and sample fixed effects plus those fixed effects interacted with a race dummy. The sample includes all Black and White men in both waves of the NLSY data between the ages of 25 and 54. Robust standard errors clustered at the individual level are shown in parentheses.

equation from equation (10) of the main text. Otherwise, the sample and demographic controls are the same as in Panel A. Specifically, to the get the model imputed task specific skill measures, we simply multiply the coefficients shown in Table 5 by the individual's reported measures of cognitive, non-cognitive and social skills.[37] We then multiply the imputed task specific individual skills by the respective task requirement of the individual's occupation. This ensures that that the skills are in consistent model units.[38] Three sets of results can be seen from Panel B. First, like in our calibrated model, the return to *Abstract* skills for White men is roughly twice the return to *Contact* skills for White men. Second, the return to model implied *Abstract* skills is estimated to be larger for Black men relative to White men. Finally, the estimated return to model implied *Contact* skills is not statistically different between Black and White men.

Appendix F Additional Results on Estimated Model Fit and Model Validation

In this section of the appendix, we show additional results on how well our calibrated model matches both additional targeted and non-targeted moments.

Appendix F.1 Model Fit

Figure A11 compares the key model moments (solid lines) against the corresponding data targets (dashed lines). As seen from the various panels of the figure, our model generally fits the data quite well. The model fit for the racial gap in the *Manual* task content of jobs – the moment we do not target – is naturally less tight (not shown), but nonetheless the model is able to match the fact that the racial gap in *Manual* tasks is close to zero. This makes us confident that our estimate of $\beta_{Manual,t}$ being equal to zero (which means that racial barriers in *Manual* tasks have no effect on sorting or wages) has little impact on our key paper results.

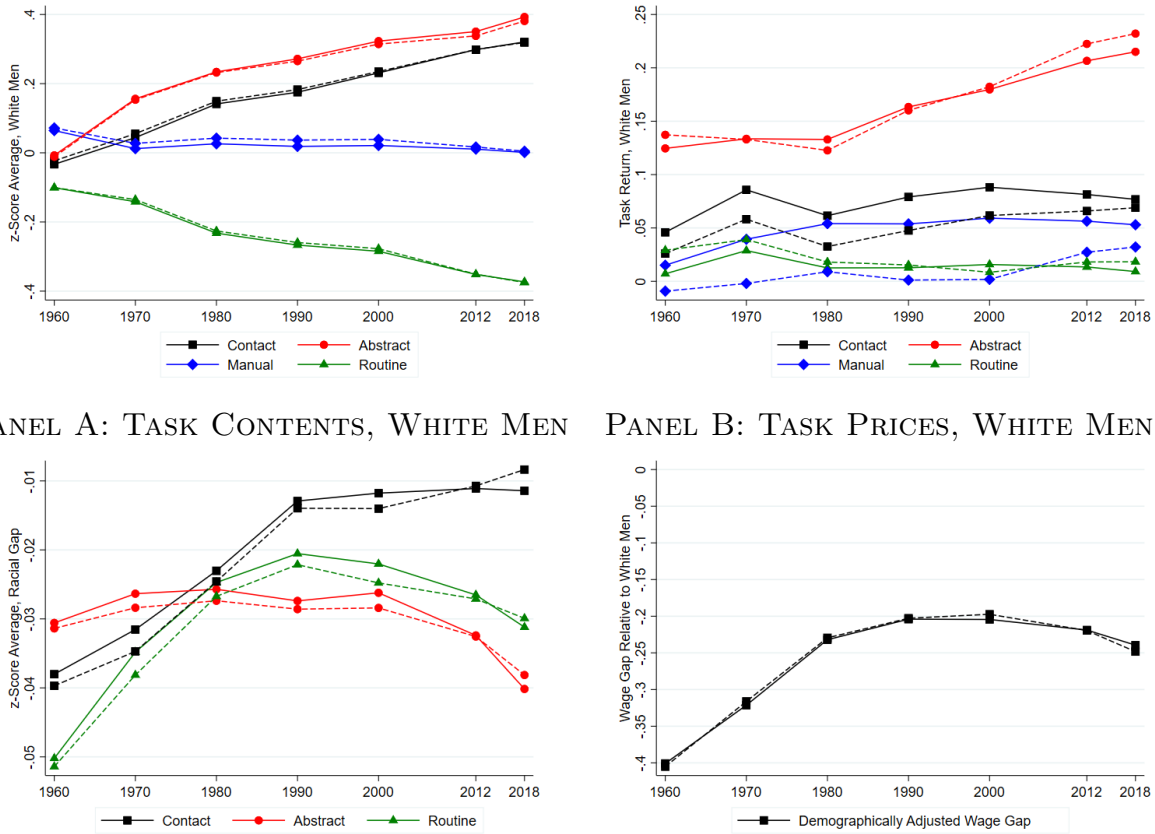
Appendix F.2 Model Validation

The counterfactuals we explore in the paper rely on the functional form assumptions we made for the various distributions from which individuals draw task specific skills or preferences.

[37] Given the low first stage power for *Routine* skills, we do not include a prediction of model implied *Routine* skills in this regression.

[38] Specifically, according to equation 8, the model structure implies that an individual's log wages are linear in $\tilde{\tau}_{jk}\phi_{ik}$. Note the δ_k's and η_k's are absorbed in the occupation-race fixed effects.

Figure A11: Model versus Data Moments



Panel A: Task Contents, White Men

Panel B: Task Prices, White Men

Panel C: Task Contents, Gap

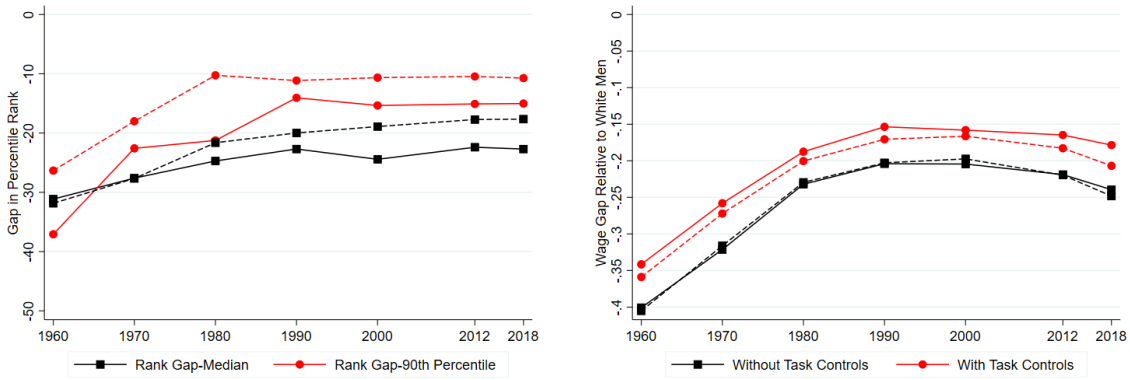
Panel D: Aggregate Wage Gap

Notes: Figure shows how selected model moments (solid lines) compare to their corresponding data moments (dashed lines). The data moments are the ones used to discipline the model. Panels A and B are data for White Men and are unconditional on education. Panels C and D are the racial gaps in wages and task content of occupations conditional on age and education as highlighted in Figures 1 and 2 to account for these demographic differences between Black and White men.

In this subsection of the appendix, we explore whether such distributional assumptions are grossly at odds with the data by assessing the extent to which our estimated model matches other non-targeted moments.

When calibrating our model, we targeted the mean wage gap between Black and White men as one of our key moments. We now explore how our model performs in matching the trends in racial wage rank gaps for different percentiles as documented by Bayer and Charles (2018). Specifically, we compute (separately by year) the median and 90th percentile of the Black wage distribution, and find out the positions of these Black wages in the White wage distribution. The differences in positions of these Black wages in Black and White distributions constitute the "wage rank gaps" at the median and 90th percentile, respectively. For example, a relative wage rank gap of -30 for the median series implies that the median

Figure A12: Model Performance Against Non-Targeted Empirical Moments



Panel A: Racial Gap in Percentile Rank of Wages

Panel B: Racial Wage Gaps Conditional on Tasks

Notes: Panel A shows the model implied racial rank gaps for different percentiles against their empirical analogs. In particular, the solid black line (with squares) shows the relative rank gap. Panel B shows model based estimates (solid lines) and data estimates from the Census/ACS (dashed lines) of demographically adjusted racial wage gaps with and without controlling for the task content of occupations. For these figures, we show model results assuming $\sigma^2 = 0$. Model results in Panels A and B are robust to alternate values of σ^2.

wage of Black men is at the 20th percentile of the White men wage distribution or 30 percentage points lower than the median. Likewise, a relative rank gap of -30 for the 90th percentile series implies that the 90th percentile in the Black man wage distribution is at the 60th percentile of the White man wage distribution. For this analysis, we follow Bayer and Charles (2018) and include both working and non-working individuals in our analysis with the wages of non-working individuals set to zero.

Panel A of Appendix Figure A12 shows our results. The dashed black line (with squares) represents the relative racial rank gap for the median series while the dashed red line (with circles) represents the relative rank gap for the 90th percentile, both using our Census/ACS data. The black and red solid lines, respectively, show the analogs from the model. It should be noted that the empirical findings from the Census/ACS data in Panel A are similar to those documented in Bayer and Charles (2018). The median Black man in 1960 had a wage that was equal to the 20th percentile of the White wage distribution. Between 1960 and 2018, the relative rank gap of the median Black made little progress. Both in 1980 and 2018, the median Black man had wages that was equal to about the 25th percentile of the White wage distribution. Conversely much more relative progress was made for Blacks at the top of the wage distribution. In 1960, the 90th percentile of the Black wage distribution was at about the 60th percentile of the White wage distribution. By 2018, the 90th percentile of the

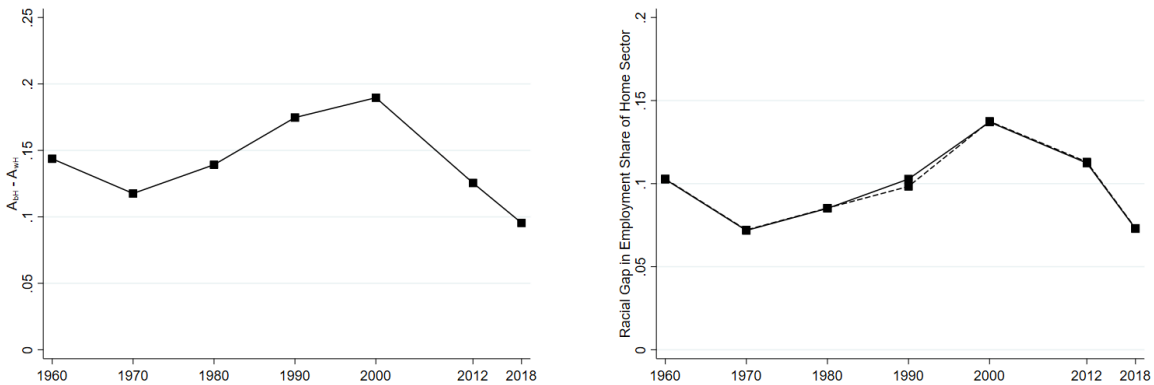
Black wage distribution had a value that was equal to roughly the 80th percentile of White distribution. However, even for the 90th percentile, little progress was made in the racial rank gap since 1980. Notice, our model (in solid lines) roughly matches these patterns even though they were not targeted. This suggests that model driving forces and racial sorting that we estimate can explain relative racial wage patterns throughout the wage distribution.

Panel B of Appendix Figure A12 shows the demographically-adjusted racial wage gap (Black lines with squares) and the racial wage gap conditional on task controls (red lines with circles), where the solid lines are model-implied and the dashed lines are their data analogs using the Census/ACS samples. Specifically, to get the red lines we regress the log wages on a race dummy and the τ_{jk}'s for each of the four tasks, separately for each year, first with the model-generated data and then with the Census/ACS data. As the comparison of the black and red solid lines reveals, the model predicts that controlling for occupational tasks only has a small effect on the estimated racial wage gap. This model finding closely matches what we find in the data. Again, these results were not targeted when calibrating the model. The similarity stems from the fact that the sorting on skills in the model is close to the sorting on skills in the data. Collectively, the fact that our estimated model matches a variety of non-target moments gives us confidence in the counterfactuals we highlight next.

Appendix G Model Estimates of Home Sector Preferences

In this section, we report the model estimates of the racial gap between preferences for the home sector in each year: $A_{bHt} - A_{wHt}$. The racial gap in the A_{gH}'s ensures that the model matches labor force participation of Black and White men in each year. The results are shown in Appendix Figure A13. For the most part, Panel A shows that the racial gap in the A_{gH}'s are relatively constant over time. However, it should be noted that the model does generate a slight increase in the preference for the home sector between Black and White men between 1980 and 2000 and a slight decline in the relative home sector preference thereafter. The relatively preferences are essentially unchanged in 1970, 1980, 2012 and 2018 relative to 1960. Panel B shows that the racial gap in non-employment rates between Black and White men both in the model (solid line) and the data (dashed line). It is not surprising that our model is matching the empirical racial gap in employment rates because we are targeting the moment. Our model estimates of $A_{bHt} - A_{wHt}$ basically just tracks the racial gap in non-employment rates between Black and White men over time.

Figure A13: Racial Differences in Home Sector Preferences



Panel A: Difference in A_{gH} Panel B: Gap in Home Sector Shares

Notes: Panel A of Figure shows the estimated differences in race-specific home sector preference parameters, A_{bH} and A_{wH}. Panel B shows the racial difference in non-employment rates in the model (solid line) and Census/ACS data (dashed line) for prime age Black and White men.

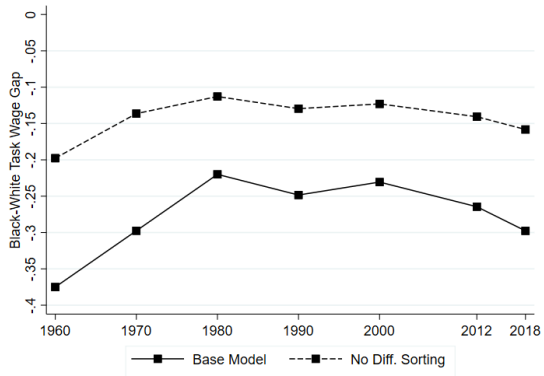
Appendix H Additional Results on Model Selection

Our task-based model captures the extent to which Black and White workers sort into different tasks in response to racial barriers. A simple statistical model with multiple skills but without a sorting mechanism will miss this force. In this section, we highlight the quantitative importance of allowing for differential sorting into tasks when we consider the effect of changing task returns on the racial wage gap.

To better understand the importance of differential selection into tasks, it would be useful to break down wages into components by task. Specifically, we define the task wage on k to be the part of the wage earned from performing task k, namely $\beta_{kt}\tilde{\tau}_{jk}(\phi_{ikt} + \eta_{gkt} + \delta^{taste}_{gkt})$; the log wage of worker i in occupation j in period t (w_{ijt}) is the sum of the wage on each task plus the occupational return A_{jt}. The Black-White gap in the *wage associated with a given task* can thus reflect three forces: (i) task racial barriers $\eta_{bkt} + \delta^{taste}_{bkt}$, (ii) differential selection on ϕ_{ikt} between Black and White men into the task, and (iii) the size of the task return β_{kt}. Importantly, the Black-White gap in task wages approximately equals the elasticity of the racial wage gap to the task price, β_{kt}. That is, a one-percent increase in β_{kt} for *Abstract* will widen the racial wage gap roughly by the size of the *Abstract* task wage gap.[39] We can therefore compare how large the racial task wage gaps in the model would be with and

[39] If workers did not re-optimize their occupational sorting in response to task return changes, the racial gap in task wages exactly equals the elasticity of the racial wage gap to β_k. But the effect of re-optimization is quantitatively small because of a logic similar to the envelope theorem. Hence, the racial gap in task wages roughly equals the elasticity.

Figure A14: *Abstract* Task Wage Gaps, with and without Sorting



Notes: Figure shows the Black-White gap in wage associated with *Abstract* tasks, namely $\beta_{kt}\tilde{\tau}_{jk}(\phi_{ikt} + \eta_{gkt} + \delta^{taste}_{gkt})$ for $k = Abstract$, in the baseline scenario (solid line) and in the counterfactual scenario where we assume Black workers sort exactly as White workers do (dotted line).

without the differential sorting into tasks and infer the quantitative difference our task-based model makes relative to a multi-skill model without differential sorting.

Figure A14 shows the Black-White gap in *Abstract* task wages over time with and without differential sorting into tasks. We focus on *Abstract* task wages because the rising *Abstract* task return post-1980 is the driving force of our model result on the stagnation of racial wage convergence. Specifically, the solid black line plots the racial gap in task wages in the base model (which captures differential sorting), whereas the dotted black line plots the counterfactual task wage gap when Black men sort exactly as White men do. Generally, even in absence of differential sorting, changes in aggregate tasks returns can disproportionately raise the wages of one group relative to the other when one group suffers from skill gaps or discrimination. For example, if Black workers sorted exactly as Whites do, a one-percent rise in $\beta_{Abstract}$ would widen the racial wage gap by about 0.1 log points in 1980 (dotted black line). Differential sorting into tasks, however, further amplifies this effect. Notably, when Black and White workers sort differentially, the effect of a one-percent rise in $\beta_{Abstract}$ on the racial wage gap in 1980 is doubled to about 0.2 log points (solid black line). Since Black workers face high racial barriers in *Abstract* tasks, they tend to sort away from occupations with high *Abstract* task requirements. But this means that Black workers will miss out relative to Whites when the return to *Abstract* task rises. Consequently, when the *Abstract* task return rises, the racial wage gap in our task-based model widens over and beyond what the composite racial barrier $\eta_{bkt} + \delta^{taste}_{bkt}$ alone would imply in absence of differential sorting.

In sum, differential sorting into tasks amplifies the impact of changing task returns on the racial wage gap. We will miss this quantitatively significant amplification mechanism even in a model with multiple skills so long as it ignores sorting. To appreciate the true impact

of changing task returns, we need a structural model with a meaningful sorting mechanism. The next section of the appendix compares the results of our structural model to other statistical decomposition methods.

Appendix I Comparison of Model Based Decomposition Method to Juhn-Murphy-Pierce Statistical Decomposition Method

Our estimated model yields quantitatively different conclusions about the extent to which race-specific factors (like a narrowing of racial skill gaps or a decline in discrimination) have improved in the United States during the last forty years relative to what would have been concluded using a popular statistical decomposition method developed by Juhn et al. (1991). Juhn et al. (1991) attribute the slowdown of convergence in the racial wage gap to rising skill prices. Central to their analysis is the racial wage *rank* gap, i.e., the position (percentile rank) of Black workers in the White earnings distribution. Specifically, they decompose trends in racial wage gaps into Black workers changing their position in the White distribution ("positional" convergence) and a change in the variance of the (White) earnings distribution ("distributional" convergence).[40] Their key insight is that changes in the level of inequality within the White earnings distribution can impact the racial wage gap even if Blacks maintained the same position, simply because Blacks and Whites occupy different initial positions in the earnings distribution. In their attempt to distinguish race-specific forces from general forces such as skill price changes, they perform the statistical decomposition of the racial wage gap trends into distributional and positional convergence, and then interpret the former as stemming from trends in skill prices and the latter as proxies for trends in race-specific forces. Such an interpretation is valid in a univariate skill model. In such a model, two workers with the same earnings will have the same underlying levels of aggregate skills, so changes in aggregate skill returns will affect them in the same way. Said differently, White men of a given wage is a good control group to proxy for the unobserved skills of Black men in a model with one aggregate skill price. This means trends in skill prices cause distributional convergence but not positional convergence; hence, when there is only one aggregate skill price, it is correct to attribute positional convergence to trends in race-specific forces.

However, in a multivariate skill model like ours, the distributional convergence fails to

[40] The terms positional convergence and distributional convergence were introduced in Bayer and Charles (2018).

capture the full effects of relative skill return changes. This is because White workers with identical initial wages are not a good control group for Black workers. A change in one skill price (relative to other skill prices) can affect two workers with the same initial earnings differently depending on the exact mix of skills they possess ($\eta_{gk}+\phi_{ik}$'s), the level of discrimination they face (δ_{gk}^{total}'s), as well as the task requirements ($\tilde{\tau}_{jk}$'s) in the occupations they have sorted into. Hence, changes in relative skill (or task) returns can shift the percentile ranks of Black workers in the White earnings distribution therefore also causing positional convergence (or divergence). As we have documented throughout the paper, Black workers have lower *Abstract* skills, face higher discrimination in *Abstract* tasks, and as a result are less likely to be in occupations with high *Abstract* task content. Given that, a rising *Abstract* task return will on average benefit Black workers less than White workers with the same initial earnings and will therefor shift down their relative positions in the earnings distribution. To the extent that this force is ignored, measured distributional convergence understates the impact of the rising *Abstract* task return on the racial wage gap. By the same token, the shifting down of Black percentile ranks in the earnings distribution (due to the rising *Abstract* task return) will dampen any estimated gains Blacks have made in reducing racial wage rank gaps, so the positional convergence will also understate the effects of declining discrimination and narrowing racial skill gaps.

Table A2: Model Decomposition vs Juhn-Murphy-Pierce Decomposition

	Change in Wage Gap	Task Model Decomposition		JMP Decomposition	
Time Period	Total	β's/A's	$(\delta+\eta)$'s	Distributional Convergence	Positional Convergence
$1960-1980$	0.169	0.056	0.120	0.035	0.134
$1980-2018$	-0.007	-0.066	0.080	-0.034	0.027

Notes: Table shows counterfactual wage gaps using our task-based model and then separately the Juhn-Murphy-Pierce (JMP) decomposition. The first column shows the actual change in the Black-White wage gap during the given time period. The next two columns decomposes how much of the change in the wage gap is due to the changing β_{kt}'s and A_{jt}'s and how much is due to the changing $\delta_{bkt}+\eta_{bkt}$. The final two columns show the JMP decomposition where the distributional convergence refers to how much of the racial wage gap is due to the changing aggregate price of skill throughout the wage distribution. Positional convergence refers to the sifts in the relative positions of Blacks and Whites within the earnings distribution.

To illustrate the quantitative difference between our model and a model with one aggregate skill price, we compare the estimated effects of changing β's and A_j's and changing

δ's and η's presented in Figure 7 to what we would find if we did a Juhn-Murphy-Pierce (JMP) decomposition on the same model-generated data. We perform this comparison during two time periods: 1960-1980 and 1980-2018.[41] The results of this comparison are shown in Table A2. During the early period, our model based decomposition and the JMP decomposition yield very similar results.[42] This is not surprising given the results in Panel B of Figure 6 showing that there was no differential trend in task prices during the 1960-1980 period. When relative task prices evolve similarly, the implications of a one-skill model and a multi-task model are similar. However, during the post-1980 period, the JMP decomposition dramatically understates the importance of skill price changes in widening the racial wage gap relative to our model. In particular, we find that the changing task prices caused the racial wage gape to increase by 6.5 log points during this period while the JMP decomposition concludes that changing skill prices increased the racial wage gap by half that amount. Because the distributional convergence is understated relative to our model, the JMP model also substantially understates the importance of declining discrimination and the narrowing of racial skill gaps in improving relative Black wages during the last forty years. Collectively, the results in Table A2 highlight that analyzing racial wage gaps in a multi-skill task model can lead to quantitatively different conclusions relative to a standard JMP decomposition, particularly in periods when relative task prices are changing.

Appendix J Additional Model Wage Decompositions

In Table 6, we show estimates of racial skill differences η_{bkt} and taste-based discrimination intensities δ_{bkt}^{taste}. However, it is hard to see just from the table how much each parameter contributes to the overall racial wage gap. Appendix Table A3 uses the model to decompose the racial wage gap in 1960, 1990, and 2012 into various additional components. To do this, we re-solve the model setting the δ_{bkt}^{taste} for both *Contact* and *Abstract* tasks to zero (row 1 of all panels) or setting η_{bkt} for both *Contact* and *Abstract* tasks to zero (row 2 of all panels). In

[41] On the model-side, we fix β_{kt}'s (and A_{jt}'s) or $\delta_{bkt}^{total}+\eta_{bkt}$'s at the levels at the beginning of the period, and report the differences between the counterfactual racial wage gap thus computed and the actual racial wage gap at the end of the period as the estimated effects of changing β_{kt}'s (second column) and $\delta_{bkt}^{total}+\eta_{bkt}$ (third column), respectively. As for the JMP decomposition, we use the model-generated earnings distributions to compute the changes in the percentile rank of each Black worker in the White earnings distribution over each period, and estimate what their wages would have been at the end of the period if the White earnings distribution were fixed at the beginning of the period; the difference between the counterfactual racial wage gap thus computed and the racial wage gap at the beginning of the period gives an estimate for positional convergence (fifth column), while the difference between the actual racial wage gap at the end of the period and the counterfactual wage gap gives an estimate for distributional convergence (fourth column).

[42] Our decomposition does not exactly sum to the empirical change in the wage gap because of an unreported covariance term.

all cases, we hold all other model parameters fixed during these respective counterfactuals. We do note, however, that when we set $\eta_{bkt} = 0$, δ_{bkt}^{stat} will also be set to zero; by definition, there can be no statistical discrimination when the average racial skill gap is set to zero. We again estimate the model under three different assumptions about the extent to which employers noisily observe worker skills. The percent explained does not sum to 100 percent because we are not doing counterfactuals on the $\eta_{bkt} + \delta_{bkt}^{total}$ for *Routine* tasks nor are we reporting the covariance terms between the various η_{bkt}'s and δ_{bkt}'s.

A few results can be seen from Appendix Table A3. First, in a world where skills are measured without error ($\sigma = 0$), tasted-based discrimination in *Contact* and *Abstract* tasks explain 28 percent of the racial wage gap in 1960 and roughly 0 to 10 percent of the racial wage gap in 1990 and 2012. Conversely, racial differences in skills associated with these tasks explain 30 percent of the racial wage gap in 1960 and 70 to 85 percent of the racial wage gap in 1990 and 2012. Second, as skills are measured with more error, taste-based discrimination explains a smaller share of the racial wage gap in all years and differences in skills explain a larger fraction in all years. Third, regardless of our assumptions about the extent to which skills are measured with error, our model finds that a substantial amount of the current racial wage gap is due to racial differences in skills. Even when skills are perfectly measured, our model concludes that half of the current racial wage gap can be explained by a remaining racial skill gap in primarily *Abstract* tasks.

Appendix K Counterfactual Robustness to Alternate θ's and ψ's

We chose our base estimate of θ in part to match the racial wage gaps at different parts of the distribution. In this section of the appendix, we show the fit of different θ's in matching Figure A12 of the main text. We then show the robustness of many of our key counterfactual findings to alternate values of θ and ψ.

In Appendix Figure A15, we show how our model matches the racial gap in percentile ranks of wages with $\theta = 4$ and $\theta = 8$. As seen from the figure, lower levels of θ fit the distributional racial wage gaps less well. However, choosing a high θ (like $\theta = 8$) does not improve fit on this dimension in any meaningful way. This is partly why we chose $\theta = 6$ as our baseline parameterization.

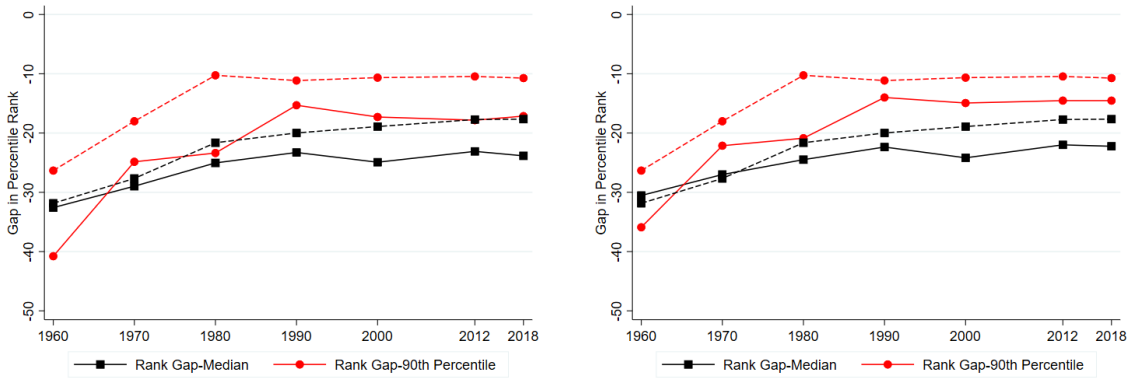
Appendix Table A4 highlights that many of our key findings are quite robust to our choice of θ and ψ. The table shows the robustness of the results in Table 7 (row 1) and Panels A and B of 7 (rows 2 - 6). In column 1 of Appendix Table A4, we re-report our baseline results. In

Table A3: Decomposition of Racial Wage Gaps

	Counterfactual Wage Gap			Percent Explained		
	1960	1990	2012	1960	1990	2012
	(1)	(2)	(3)	(4)	(5)	(6)
Base Model	-0.40	-0.20	-0.22	–	–	–
Panel A: Signal To Noise = 1.0 ($\sigma = 0$)						
Setting $\delta_{bkt}^{taste} = 0$ ($k = Abstract, Contact$)	-0.29	-0.20	-0.19	28%	0%	12%
Setting $\eta_{bkt} = 0$ ($k = Abstract, Contact$)	-0.28	-0.03	-0.07	30%	84%	68%
Panel B: Signal To Noise = 0.9 ($\sigma = 0.17$)						
Setting $\delta_{bkt}^{taste} = 0$ ($k = Abstract, Contact$)	-0.30	-0.21	-0.20	26%	-5%	9%
Setting $\eta_{bkt} = 0$ ($k = Abstract, Contact$)	-0.27	-0.02	-0.06	32%	92%	72%
Panel C: Signal To Noise = 0.75 ($\sigma = 0.30$)						
Setting $\delta_{bkt}^{taste} = 0$ ($k = Abstract, Contact$)	-0.31	-0.22	-0.21	23%	-11%	3%
Setting $\eta_{bkt} = 0$ ($k = Abstract, Contact$)	-0.26	0.00	-0.04	36%	102%	80%

Note: Table shows the racial wage gap in our base model (row 1) as well the racial wage gap in various counterfactuals were we separately set the η_{bkt}'s and δ_{bkt}^{taste}'s for *Abstract* and *Contact* tasks to 0 for Black men. When setting the η's to zero, δ_{bkt}^{stat} will also equal zero by definition. The three panels show the counterfactuals under different assumptions about the which skills are noisily observed. Columns (1)-(3) show the level of the racial wage gap setting the various parameters to 0 in 1960, 1990, and 2012, respectively. Columns (4)-(6) show the share of the wage gap explained in counterfactual. We compute the share explained by computing the counterfactual wage gap relative to the base model wage gap and dividing that difference by the base model wage gap.

Figure A15: The Role of Theta in Matching the Racial Gap in Percentile Ranks of Wages



Panel A: $\theta = 4$ Panel B: $\theta = 8$

Notes: Figure reproduces the results from Figure A12 of the main text with different values of θ.

columns 2 and 3, we show the robustness of results when we set $\theta = 4$ and $\theta = 8$, respectively. In columns 4 and 5, we show the results for $\psi = 3.5$ and $\psi = 5.5$. Importantly, across all the robustness exercises, declining taste-based discrimination in *Abstract* and *Contact* tasks explains between 30 and 50 percent of the decline in the racial wage gap between 1960 and 2018. Likewise, across all the robustness specifications, the increasing returns to tasks post-1980 exacerbated the racial wage gap while declining δ_{bkt}'s and η_{bkt}'s narrowed the gap. In all cases, the two effects roughly offset each other such that racial wage gaps were constant during the 1980 to 2018 period.

Appendix L Counterfactual Robustness to Alternate First Stage Decomposition Projection

For our main decomposition of the composite gap in $\delta_{kt}^{total}+\eta_{kt}$ into the individual components of δ_{kt}^{taste}, η_{kt} and δ_{kt}^{stat}, we used the first stage projections in from Table 5 from the main text. Specifically, by estimating equation 10 for each task-specific skill, we produced a weighting (the b's) of each NLSY individual skill measure for each of the model task-specific skills. Our baseline projection, however, restricted the estimated b's from the first stage regression to be the same across all years. The purpose of this restriction was to increase the power in estimation by pooling the data for all three years. Yet, it is possible that the importance of each individual skill in performing certain tasks – the b's – changed over time, for example due to changing technology. As a robustness exercise, we re-estimate the first stage regressions separately using the 1979 NLSY data (which we map to model year 1990) and using the

Table A4: Key Result Robustness to Alternate Values of θ and ψ

		θ		ψ	
	Base	4	8	3.5	5.5
Share of 1960-2012 Wage Gap Explained					
Holding δ^{taste}_{bkt} fixed ($k = Abstract, Contact$)	48%	49%	47%	33%	48%
Counterfactual: Wage Change 1980-2018					
Holding $\beta_{Abstract}$ at 1980 Levels	0.100	0.073	0.119	0.106	0.096
Holding all β_k's and A_i's at 1980 Levels	0.066	0.049	0.073	0.058	0.062
Holding $\delta^{total} + \eta$ for *Contact* at 1980 Levels	-0.057	-0.055	-0.058	-0.027	-0.058
Holding $\delta^{total} + \eta$ for *Abstract* at 1980 Levels	-0.031	-0.028	-0.032	-0.034	-0.027
Holding $\delta^{total} + \eta$ for all tasks at 1980 Levels	-0.080	-0.067	-0.086	-0.047	-0.082

Note: Table shows the robustness of many of our key results to alternate values of θ and ψ. In our baseline model, we use $\theta = 6$ and $\psi = 5$. We replicate the results from our baseline model in column 1. We show the robustness of our results to values of $\theta = 4$ and $\theta = 6$ (columns 2 and 3) and values of $\psi = 4$ and $\psi = 6$ (columns 4 and 5). The first row replicates results results from row 1 of Table 7 showing the percent of the actual decline in the racial wage gap between 1960 and 2012 explained by the falling δ^{taste} for *Contact* and *Abstract* tasks. The remaining rows show the counterfactual changes in the racial wage gap between 1980 and 2018 from the various counterfactual exercises done in Figure 7.

1997 NLSY data (which we map to model year 2012).

Appendix Table A5 shows the coefficients in the alternate first stage regressions where we allow the b's to vary by year.[43] Three things are of note in the table. First, the power of the alternate first-stage regressions is lower than that of the baseline regressions. For example, the F-statistics in the baseline specification are 21 for *Abstract* and 10 for *Contact*; the respective F-stats in the alternate specification are roughly around 12 and 9 in 1990 and 2012, and they are even lower in 1960. Second, however, the estimated b's are roughly stable over time. In all years, cognitive skills strongly predict *Abstract* task efficiency and social skills strongly predict *Contact* task efficiency. The stability of coefficients over time justifies our restriction in the baseline specification that the b's are constant across years. Last, that said, the alternate first stage regressions seem to suggest that the weighting on cognitive skills for *Abstract* task efficiency increased between 1990 and 2012 (from 0.13 to 0.19). Of course, given the lower power of the alternate regressions, it is hard to tell whether the estimated trends reflect actual changes in the structure of the economy or they solely capture noise. Nevertheless, we can proceed with the decomposition with the alternative

[43] We omit the results for *Routine* task skills because the power is too low for the results to be credible.

Table A5: First Stage Regression of Average Model Task Skills on Average NLSY Individual Skills, Cross-Occupation Variation

	Pane A: 1960		Pane B: 1990		Panel C: 2012	
	Abstract	*Contact*	*Abstract*	*Contact*	*Abstract*	*Contact*
Cognitive	0.14	0.03	0.13	0.02	0.19	0.05
	(0.09)	(0.03)	(0.04)	(0.03)	(0.03)	(0.01)
Non-Cognitive	0.04	0.05	0.09	0.05	0.11	-0.02
	(0.07)	(0.03)	(0.07)	(0.04)	(0.04)	(0.02)
Social	-0.01	0.11	0.02	0.18	-0.04	0.11
	(0.11)	(0.04)	(0.09)	(0.06)	(0.08)	(0.03)
Constant	1.10	1.11	1.03	1.09	1.04	1.10
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Adj. R-Squared	0.30	0.34	0.46	0.45	0.54	0.43
F-Stat	4.95	3.00	10.5	8.2	15.0	9.9

Notes: Table shows estimate coefficients from first stage regression equation (10) for White men. Each column is a separate regression exploiting cross-occupation variation. We use 66 broad occupation categories. For these regressions, we estimate these regressions separately for 1960 (using data in the South region from the 1979 NLSY), 1990 (using data from the 1979 NLSY) and for 2012 (using data from the 1997 NLSY). See the text for additional details.

weightings and assess their implications in case that these trends are real.

Appendix Table A6 shows estimates of δ_{kt}^{taste} and η_{kt} in different years calculated using our alternate first-stage specification. For illustrative purposes, we only show the results assuming employers observe worker skills without noise ($\sigma = 0$). Importantly, the estimated decline in η for *Abstract* tasks between 1990 and 2012 is much smaller under the alternate first-stage specification. In the baseline specification, η for *Abstract* tasks declines from -0.26 to -0.18 over the period, reflecting the narrowing cognitive skill gap observed in the NLSY data; in the alternate specification, however, it declines from -0.24 to -0.22 only. This is because an increase in the importance of cognitive skills in performing *Abstract* tasks – which the alternative first-stage regressions imply over the 1990-2012 period – disadvantages Blacks relative to Whites with respect to efficiency in performing *Abstract* tasks given that Black men have a relative deficit in cognitive skills. Hence, despite the significant decline in the cognitive skill gap between 1990 and 2012, the racial gap in *Abstract* task-specific skills declined little over the period. Put differently, the increase in the importance of cognitive skills for *Abstract* tasks might be another force that offset the decline in race-specific barriers and caused the stagnation of the racial wage gap over the last couple of decades, along with

the rising return to *Abstract* tasks we discuss in Section 6. If this is the case, then declining taste-based discrimination accounts for almost all the decline in $\delta + \eta$ for *Abstract* tasks over the 1990-2012 period.

Table A6: Task Decomposition of Racial Skill Gap and Task-Based Discrimination, Alternate First Stage Regression ($\sigma = 0$)

	Contact				*Abstract*			
	1960	1990	2012	$\Delta(12-60)$	1960	1990	2012	$\Delta(12-60)$
δ^{taste}, $\sigma = 0$	-0.18	-0.00	-0.05	0.13	-0.14	0.02	0.03	0.16
η, $\sigma = 0$	-0.11	-0.08	-0.03	0.08	-0.24	-0.24	-0.22	0.02

Notes: Table shows model decomposition of racial differences in δ_{bk}^{taste}, and η_{bk} for *Abstract* and *Contact* task in 1960, 1990, and 2012 using our alternate first stage regression where the b's are allowed to differ over time. We do the decomposition assuming employers can accurately observe the skill levels of their employees ($\sigma = 0$).

Appendix M Additional Model Validation Using Cross-Region Analysis

To further highlight the importance of declining taste-based discrimination in explaining the racial wage convergence, we show one additional set of results in this appendix section exploiting regional variation in the micro data from the Census/ACS. Appendix Figure A16 compares the cross-regional difference in racial wage gaps to the cross-regional differences in various racial task gaps. Specifically, the solid black line (with squares) shows the convergence of racial wage gaps (conditional on education) in the South relative to the non-South regions since 1960 using the Census/ACS data (left axis). This line is just the difference between the two lines in Figure 8 discussed above. The other two lines in Figure A16 show the difference in the racial gaps in *Contact* and *Abstract* tasks between the South and non-South regions in each year (right axis). This data is just the difference between the patterns shown in Panels A and B of Figure 3.

Three facts emerge from the figure. First, the racial wage gap in the South fell sharply relative to the racial wage gap in the non-South region consistently during the 1960 to 2018 period. Specifically, the racial wage gap in the South was roughly 20 log points higher than that in the non-South in 1960, but by 2018, the racial wage gap in the South essentially converged to that in the non-South region. Second, the racial gap in *Contact* tasks narrowed

Figure A16: Trends in Racial Wage and Task Gaps by Region: Census/ACS Data

Notes: Figure shows *differential* racial gaps in wages and tasks between the South region and non-South region over time. The black line (with squares) is the difference between the two racial wage gap series in Figure 8. The dashed red line (with squares) shows the difference in the racial gap in *Contact* tasks over time while the dashed blue line (with triangles) shows the difference in the racial gap in *Abstract* tasks over time. The the data from the latter two series come from differencing the patterns in Panels A and B of Figure 3.

sharply in the South relative to the non-South region over the same period, from a 6 percentage point gap in 1960 to a 2 percentage point gap in 2012, closely tracking the convergence of racial wage gaps across the two regions. Last, on the other hand, there was almost no change in the racial gap in *Abstract* tasks between the South and non-South regions during this time period. Given the model results in Section 8.1 suggesting that the trends in *Contact* task gaps are a good proxy for trends in taste-based discrimination, the patterns imply that the faster decline in taste-based discrimination in the South region primarily drove the convergence of the racial wage gaps across the two regions. Overall, the regional result reinforces the model finding that the declining taste-based discrimination was the primary force behind the racial wage gap trends post-1960.

Appendix N Additional Model Results

This section of the appendix provides details on additional model results.

Appendix N.1 Conditional Wages and Statistical discrimination

In this section, we derive expressions for conditional wages and statistical discrimination when skills are observed with noise by employers. Employers set the wage of each worker at the worker's expected marginal revenue product given observed skills $(\hat{s}_1, ..., \hat{s}_K)$. Since p_j

and y_{gij} are in logs, the conditional expectation of marginal revenue product is given by

$$E\left[e^{p_j+y_{gij}} \mid (s_{g1},...,s_{gK}) = (\hat{s}_1,...,\hat{s}_K)\right] = E\left[\exp\left(A_j + \sum_K \lambda_{jk}(\phi + \delta_{gk} + \eta_{gk})\right) \mid (s_{g1},...,s_{gK}) = (\hat{s}_1,...,\hat{s}_K)\right].$$

Taking the logs and noting the independence of ϕ_{ik}'s and ϵ_{ik}'s from one another, we obtain that the conditional wage is given by

$$w_{gij}^{cond}(\hat{s}_1,...,\hat{s}_K) = A_j + \sum_K \log E\left[e^{\lambda_{jk}(\phi+\eta_{gk})} \mid s_{gk} = \hat{s}_k\right] + \sum_K \lambda_{jk}\delta_{gk}^{taste},$$

which yields the expression in the text with some rearrangements.

Next, we derive the expression for $E\left[e^{\lambda_{jk}(\phi+\eta_{gk})} \mid s_{gk} = \hat{s}_k\right]$. Define f_{ϕ_K} and f_{ϵ_k} to be the probability density functions for ϕ_{ik} and ϵ_{ik}, respectively. The joint density function for s_k and ϕ_k is then given by

$$f_k(s,\phi) = f_{\phi_k}(\phi) f_{\epsilon_k}(s - \eta_k - \phi).$$

Thus, we have

$$E\left[e^{\lambda_{jk}(\phi+\eta_{gk})} \mid s_{gk} = \hat{s}_k\right] = \frac{\int_0^\infty e^{\lambda_{jk}(\phi+\eta_{gk})} f_{\phi_k}(\phi) f_{\epsilon_k}(\hat{s}_k - \phi - \eta_{gk}) d\phi}{\int_0^\infty f_{\phi_k}(\phi) f_{\epsilon_k}(\hat{s}_k - \phi - \eta_{gk}) d\phi} \quad .$$

We evaluate the integrals numerically.

Lastly, we highlight one interesting fact about the statistical discrimination term. Note that the statistical discrimination term can be written as

$$\delta_{gk}^{stat}(\eta_{gk}, \lambda_{jk}, \hat{s}_k) = \eta_{gk} + \left[\phi_{g_0ik}^e(\lambda_{jk}, \hat{s}_k - \eta_{bk}) - \phi_{g_0ik}^e(\lambda_{jk}, \hat{s}_k)\right], \tag{A3}$$

where we defined $\phi_{g_0ik}^e(\lambda_{jk}, \hat{s}_k) = \log E\left[e^{\lambda_{jk}\phi} \mid s_{g_0k} = \hat{s}_k\right]^{1/\lambda_{jk}}$ to be the log expected efficiency of a base-group worker conditional on observing $\hat{s}_k$. The expression shows that the statistical discrimination term depends on (i) the difference in group means of effective skills (η_{gj}) and (ii) the expected within-group position (ϕ_{ik}) of each worker in the skill distribution conditional on observed skills (the term in brackets). To see this, note that the position of $\hat{s}_k$ in the observable skill distribution for group g is equivalent to the position of $\hat{s}_k - \eta_{bk}$ in the observable skill distribution of the base group. Thus, supposing $\eta_{gk} < 0$, the expected ϕ_{ik} is higher for workers of group g than for workers of the base group with the same observed skills (assuming a finite σ). The term in brackets captures how much higher the ϕ_{ik}'s of workers of group g is expected to be relative to those of base group workers with the same observed skills. For example, skills are observed perfectly by employers, then the expected ϕ_{ik} for workers of group g exceeds that of the base group workers with the same observed

skills by $-\eta_{gk}$ and completely offsets the gap in mean skill levels η_{gk}. If, on the other hand, observed skills are so noisy and give no information about true levels of worker skills, then no difference in ϕ_{ik}'s can be expected between groups even conditional on observed skills (i.e., the term in brackets equals zero) and workers are paid entirely based on group means. Said differently, the term in brackets reflects how much the information from observed skills can compensate for the known gaps in the mean human capital levels across groups.

Appendix N.1.1 Proof of Proposition 1

We prove Proposition 1 in two steps. First, it is clear from equation A3 that $\delta_{gk}^{stat}(\eta_{gk}, \lambda_{jk}, \hat{s}_k) = 0$ if $\eta_{gk} = 0$. Second, the continuity of $\phi_{g_0ik}^{e}(\lambda_{jk}, \hat{s}_k)$ in $\hat{s}_k$ implies that $\delta_{gk}^{stat}(\eta_{gk}, \lambda_{jk}, \hat{s}_k)$ is continuous in η_{gk}. Thus, $\delta_{gk}^{stat}(\eta_{gk}, \lambda_{jk}, \hat{s}_k)$ tends to zero as $\eta_{gk} \to 0$, as desired. □

Appendix N.2 Home Sector and Reservation Utility

The model allows for the possibility that workers may choose to work in the home sector, denoted as $j = H$. Specifically, we treat the home sector as another potential occupation (with task requirements $\lambda_{H1}, ..., \lambda_{HK}$ and occupational return A_{gH}) where the returns are non-pecuniary. That is, the reservation utility u_{giH} of a worker with given observable credentials equals the log of the worker's expected marginal revenue product in an occupation with task requirements $(\lambda_{H1}, ..., \lambda_{HK})$ plus the log of the idiosyncratic preference for home sector, ν_{iH}:

$$u_{giH} \equiv A_{gH} + \sum_{K} \lambda_{Hk} \left(\phi_{g_0k}^{e}(\lambda_{Hk}, \hat{s}_{ik}) + \delta_{gk}^{taste} + \delta_{gk}^{stat}(\eta_{jk}, \lambda_{Hk}, \hat{s}_{ik}) \right) + \log \nu_{giH},$$

where ν_{giH} is the idiosyncratic preference for home sector. Like the other occupational preferences ν_{gij}, the preference for home sector ν_{giH} follows a Frechet distribution with shape ψ and scale 1. Note that the term A_{gH} plays the role of a scale parameter of the Frechet distribution. We allow A_{gH} to differ by race, unlike with A_j's. The differences in A_{gH}'s across groups capture any forces other than differential task returns that may create labor supply differences between racial groups.

Appendix N.3 Employment Share of Occupations

In this section, we derive an expression for the employment share of each occupation. Recall that each worker i of race group g chooses occupation j by maximizing the utility given by

$$u_{gij} = w_{gj}^{cond}(s_{gi1}, ..., s_{giK}) + \log \nu_{ij}.$$

where the occupational preference ν_{ij} follows a Frechet distribution with scale 1 and shape ψ. Let f and F be the pdf and cdf of the Frechet distribution, respectively. Among workers with the same observed skills $(s_{gi1}, ..., s_{giK}) = (\hat{s}_1, ..., \hat{s}_K)$ and of the racial group g, the proportion of those who chooses occupation j is given by

$$\begin{aligned}\rho_{gj}(\hat{s}_1, ..., \hat{s}_K) &= \Pr\left[\exp\left(w_{gj}^{cond}(\hat{s}_1, ..., \hat{s}_K)\right)\nu_j > \exp\left(w_{gj'}^{cond}(\hat{s}_1, ..., \hat{s}_K)\right)\nu_{j'},\ \forall j' \neq j\right] \\ &= \int f(\nu) \cdot \Pi_{j' \neq j} F\left(\exp\left(w_{gj}^{cond}(\hat{s}_1, ..., \hat{s}_K) - w_{gj'}^{cond}(\hat{s}_1, ..., \hat{s}_K)\right)\nu\right) d\nu.\end{aligned}$$

Defining

$$\omega_{gj} = \sum_{j'=1,...,J,H} \exp\left(\psi w_{gj'}^{cond}(\hat{s}_1, ..., \hat{s}_K) - \psi w_{gj}^{cond}(\hat{s}_1, ..., \hat{s}_K)\right),$$

we obtain

$$\begin{aligned}\rho_{gj}(\hat{s}_1, ..., \hat{s}_K) &= \int \psi \nu^{-\psi-1} \cdot \exp\left[-\omega_g \nu^{-\psi}\right] d\nu \\ &= \frac{1}{\omega_{gj}}.\end{aligned}$$

Appendix N.4 Elasticity of Labor Supply

Lastly, we derive an expression for labor supply elasticity and show how targeting the moment can help us pin down ψ, the shape parameter for the Frechet distribution from which we draw idiosyncratic preferences for occupations.

Given the expression for the employment share in each occupation, the aggregate labor supply for race group g is given by

$$L_g = \int_{S_g} \left(1 - \rho_{gH}(\hat{s}_1, ..., \hat{s}_K)\right),$$

where the integral is taken over the support of $(\hat{s}_1, ..., \hat{s}_K)$, denoted by S_g. We want to find out the labor supply changes in response to one percent increase in wages for all occupations (excluding, of course, the home sector). To this goal, write

$$\exp\left(w_{gj}^{cond}(\hat{s}_1, ..., \hat{s}_K)\right) \equiv W \cdot \exp\left(w_{gj}^{\prime cond}(\hat{s}_1, ..., \hat{s}_K)\right)$$

for all $j' \neq H$ for some constant W. The labor supply elasticity for group g is then given by

$$\varepsilon_g = -\left[\int_{S_g} \frac{\partial \rho_{gH}(\hat{s}_1, ..., \hat{s}_K)}{\partial W}\right] \cdot \frac{W}{L_g}.$$

Using the results from the previous section, we obtain

$$\frac{\partial \rho_{gH}(\hat{s}_1, ..., \hat{s}_K)}{\partial W} = W^{-1}\psi\rho_{gH}(\hat{s}_1, ..., \hat{s}_K)\left(1 - \rho_{gH}(\hat{s}_1, ..., \hat{s}_K)\right),$$

Hence, the labor supply elasticity is given by

$$\begin{aligned}\varepsilon_g &= \psi\frac{\int_{S_g}\rho_{gH}(\hat{s}_1, ..., \hat{s}_K)(1 - \rho_{gH}(\hat{s}_1, ..., \hat{s}_K))}{\int_{S_g}(1 - \rho_{gH}(\hat{s}_1, ..., \hat{s}_K))} \\ &= \psi\left[1 - \frac{\int_{S_g}(1 - \rho_{gH}(\hat{s}_1, ..., \hat{s}_K))^2}{\int_{S_g}(1 - \rho_{gH}(\hat{s}_1, ..., \hat{s}_K))}\right].\end{aligned}$$

Note that the expression for ε_g involves the parameter ψ, the shape parameter for the Frechet distribution from which we draw idiosyncratic preferences for occupations. Hence, targeting the elasticity of labor supply helps us discipline the parameter ψ.