*ets research institute

2024

POLICY REPORT

Another Look at U.S. Gender-based Wage Disparities:

Insights from PIAAC

AUTHORS

Henry Braun,
Boston College

Gulsah Gurkan,
Independent Scholar

Table of Contents

This report was written by:

Henry Braun
Gulsah Gurkan

July 2024

ETS Center for Research on Human Capital and Education

Research and Development
Educational Testing Service
Rosedale Road
Princeton, NJ 08541-0001

Suggested citation:
Henry Braun and Gulsah Gurkan, *Another Look at Gender-Based Wage Disparities: Insights from PIAAC* (Princeton, NJ: Educational Testing Service, 2024).

Preface

In this important, new policy report from the ETS Center for Research on Human Capital and Education, researchers Henry Braun and Gulsah Gurkan, of Boston College, explore and evaluate a range of background and labor market data, including details on occupation, industry, and work experience drawn from a nationally representative sample of adults aged 16 to 65 in the United States, to shed light on the complexities surrounding gender-based wage disparities in the US labor market.

At the outset, the authors engage in a thoughtful dialogue concerning key issues explored in the literature regarding gender wage gaps. The highlighted research suggests that although there has been noticeable reduction in gender wage disparities during the 20th century, persistent structural economic factors and societal norms continue to hinder full equity. As an example, research summarized in the report demonstrates that while there has been a narrowing of the gender wage gap, the gap between men and women at the top of the income distribution remains comparatively wide. Other research reported in this paper notes that at least some of these disparities are likely to stem from tradeoffs for job flexibility or to accommodate work "stop-outs," where individuals temporarily exit the workforce for various reasons but intend to rejoin later. The range of findings informs and contextualizes the discourse around gender wage disparities while also underscoring the intricate challenge researchers face when trying to quantify the size and nature of gender earnings gaps.

To help advance the field and scholarship on this issue, Braun and Gurkan turn to the OECD's Programme for the International Assessment of Adult Competencies (PIAAC). The PIAAC survey is somewhat unique in that it provides nationally representative data that includes both demographic and background information on adults along with comparable measures of cognitive skills - a key and often missing component in analyses that consider human capital. Adding cognitive skills to the analyses is critical, according to the research reviewed by the authors, because educational attainment alone may provide a less comprehensive measure of human capital -- a key factor in labor earnings outcomes.

In addition, the authors use logistic regression in their analyses to determine the probability of earners falling into the lower or upper tails of the income distribution. This approach was pursued for its ability to address nuanced income relationships across various income levels; to highlight pronounced disparities at higher incomes; and to mitigate gender-based wage differences often underestimated by conventional regression models. Each of these is a known challenge in previous research and thoughtfully discussed in this paper.

At every level of education and every category of race/ethnicity, the authors report that women earn less than men. Further, while there is a "cognitive skills premium," meaning that higher skills lead to greater earnings for each group studied, the payoff for these skills is consistently greater for men than for women. Additionally, the analyses revealed that women with children are four times less likely to attain the highest earnings category compared to men with children after accounting for a broad set of relevant factors, such as education, occupation, industry, hours worked, and skills. While the substantial disparities at the upper tail of the income distribution found here are in line with previous research, the authors caution that "...it has proven to be very difficult to disentangle the various factors that may contribute to these wage disparities, principally because some ... factors are difficult to measure and because they often interact in complex ways over time," they continue, "...a 'one size fits all' explanation is not very likely to exist..."

The complexity around this issue may be at the core of the mixed results from public policy efforts enacted since World War II. The authors highlight several significant laws including the Equal Pay Act (1963, 1965), Title VII of the Civil Rights Act (1964), and Title IX of the amended Civil Rights Act (1972) and note that research on the impacts of these efforts reveal varied results. They also note that policies like the Family and Medical Leave Act (1993), aimed to balance work and family needs, had complex effects on female employment. More positive outcomes highlighted from research on affirmative action indicate that as women advance in corporate hierarchies due to these initiatives, they are well-placed to mentor and support other women through diverse networks and help acclimate male colleagues to gender diversity. While legislation can play a part, the authors suggest evolving social norms around, for example, work allocation and job redesign may play an important role in reducing earnings differentials

From Braun and Gurkan's solid review of existing scholarship to the sophisticated analyses of the PIAAC data and their discussion of notable legislative efforts aimed at reducing gender wage disparities, readers will encounter deep and nuanced insights to advance their understanding of this critical topic. This report is recommended reading for policymakers and others seeking to better understand the complexity of the gender wage gap, as well as for those interested in exploring how shifting social norms, employer flexibility, and additional legislative adjustments may serve as critical levers for mitigating gender disparities in the American labor market.

Irwin Kirsch (retired) and Anita Sands
ETS Research Institute

Acknowledgements

The authors wish to acknowledge the thoughtful comments and suggestions received from our reviewers: Laura Halderman, Sara Haviland and Timothy Davey from ETS, as well as the ETS associate editor, Heather Buzick. A special thank you to K. Lang (Boston U.) for invaluable advice on the relevant econometric literature. Although those who reviewed the paper provided valuable comments and feedback, all errors of fact or interpretation are those of the authors. The authors are also grateful for the editorial assistance from Kim Fryer and Ayleen Gontz, whose careful attention to details improved the flow of the paper. We also wish to thank Nicole Fiorentino and Phillip Leung and their team for their help in producing the report. Finally, we would like to thank Meghan Moore and her team for publication support throughout the development of this report.

Introduction: Setting the Stage

There is ongoing (and understandable) interest in labor market outcomes and their relationships to individuals' characteristics. That interest is driven, in part, by questions of how labor markets operate and the extent to which they advantage or disadvantage certain groups of workers. The most common outcomes are employment status and income; individual characteristics comprise various combinations of demographics, family background, human capital, and employment-related variables.[1] The present study employs data from the Programme of International Assessment of Adult Competencies (PIAAC).[2] PIAAC is nearly unique among surveys of adults in that its database contains credible measures of adults' cognitive skills. Thus, it enables more refined analyses in response to some of the questions arising in the econometric literature. In this regard, the purpose of the study is two-fold: first, to identify patterns in the data relating individuals' incomes to a range of characteristics, including cognitive skills, and second, to find and evaluate evidence of systematic inequalities related to gender and to race/ethnicity. The results of these analyses are interpreted and their implications for policy are considered in view of the limitations inherent in this and other studies.

Generally speaking, demographics comprise gender, race/ethnicity, country of birth, and disability status. Family background comprises parental education, family socioeconomic status, home language, and immigration status. Human capital comprises the collection of a person's knowledge and experience, cognitive skills, as well as interpersonal skills such as collaboration and teamwork. In some cases, it also includes traits such as motivation, persistence, and risk-taking. Finally, employment-related variables can include work experience and "stop-outs" as well as industry and occupational categories. To our knowledge, no large-scale surveys that incorporate data on all these variables. Thus, any study based on data from a particular survey suffers from a "missing data problem" that limits the strength of the conclusions that can be drawn from the analyses. Nonetheless, each well-conducted study contributes to the cumulation of knowledge that is the hallmark of science.

With regard to income, population-level studies reveal substantial variation across strata defined by combinations of demographic characteristics and family background factors. In view of the U.S. sociocultural and labor market history, there has been a particular focus on observed differences by gender and by race/ethnicity. These differences are sometimes interpreted as evidence of bias in the labor market. However, such interpretations are premature, inasmuch as there are many circumstances that are associated with differential earnings.

One problem in most published studies is that they rely on years of schooling (or levels of educational attainment) and years of experience as proxies for the full set of foundational human capital factors.[3] Consequently, most studies of the "skills premium" compare individuals with different levels of educational attainment (e.g., high school diploma and college degree). Although educational attainment and years of experience certainly capture some of the relevant differences among individuals, they are certainly incomplete. As many authors have argued, augmenting these variables with one or more measures of cognitive skills strengthens the predictive model, thereby making interpretations of between-stratum differences more credible.[4]

From a methodological perspective, cross-tabulations are limited in the number of variables they can comfortably represent. To address this difficulty, it is common in the econometric literature to construct regression models with log(income) as the outcome. The suite of predictors includes sets of variables representing as many factors as are available. They are often added in blocks, with the predictors in a block having some commonality (e.g., measures of family background). The basic idea in regression modeling is to track the signs and magnitudes of the fitted regression coefficients corresponding to the focal factors (e.g., gender or race/ethnicity) as more sets of predictors are added to the model. Typically, the coefficients do not change sign, but their magnitudes may diminish, often substantially so. Nonetheless, in the final model, many of the coefficients remain statistically significant and of practical importance. The variance in the outcomes explained by the models is also relevant to the interpretation of the results.

In many respects, these regression models represent an improvement over simple cross-tabulations; however, they are subject to important limitations as will be elucidated below. Thus, the interpretation of the coefficient of a focal variable (e.g., gender) as a credible measure of bias against the corresponding group is subject to multiple caveats. Further, the amount of variation in the outcome accounted for by these models rarely exceeds 50%, and is often substantially lower. Nonetheless, this approach can provide important insights and should be pursued when possible.

This study is noteworthy in a number of respects. To begin, it is the first study to employ the full set of U.S. data collected over three rounds of PIAAC administrations. As such, it incorporates a range of predictors, including educational attainment and measures of cognitive skills to represent human capital. Second, it does not employ ordinary multiple regression models to the full range of income data. Rather, using logistic multiple regression, it focuses on the nature of the relationships in the lower and upper tails of the income distribution.

We present three rationales for this choice of modeling strategy. First, there is no guarantee that a single regression model can adequately capture the relationships of interest over the full range of incomes. This is an open question inasmuch as econometricians typically fit sequences of regression models using log(income) as the criterion. They report the variance accounted for ($R^{2)}$ by the various models and remark/interpret the magnitudes and signs of the coefficients but rarely examine their goodness-of-fit. Second, there is substantial evidence that wage disparities are concentrated at the high end of the income distribution,[5] suggesting that a particular focus on the upper tail of the income distribution is called for. Finally, with log(income) as the criterion, the coefficients of the predictor variables represent relative rates of return. Because males and females have different baselines, the fitted models typically underestimate the absolute wage differentials.[6] This last difficulty is circumvented by employing the logistic regression models with outcomes defined in terms of the observed income distributions.

The present study has yielded a number of interesting findings. For example, for all groups defined by combinations of gender and race/ethnicity, we find consistent evidence of a skills premium by level of educational attainment (holding cognitive skills approximately constant) and positive trends with cognitive skills (holding educational attainment constant). In addition, in the upper tail of the income distribution among full-time workers we find substantial differences by gender even after controlling for the full set of measured factors. Differences by race/ethnicity are less consistent and will be discussed after the findings are presented. In the lower tail of the income distribution, employment status accounts for most of the explanatory power of the model. However, there are still meaningful differences by gender. Our results are broadly in line with those in the econometric literature, as well with those reported by Neeta Fogg, Paul Harrington, and Ishwar Khatiwada,[7] who employed different statistical methods and had access only first round of PIAAC data.

For many reasons, the interpretation of gender-based wage disparities for policy purposes must be made cautiously, if only because the models explain less than a third of the variance in the outcomes. Moreover, in the case of gender, a number of studies have attempted to disentangle the impact of such factors as occupational choice, work experience, hours worked, the "child penalty," and social norms.[8] Unless these and other factors are taken into account, they confound simple comparisons and can lead to erroneous policy conclusions. These and related issues are discussed more fully in the final section.

This next section offers a brief review of the literature and a discussion of the construction of the analytic database, as well as a number of methodological considerations. We then present a variety of descriptive statistics, followed by a summary of the results for the skills premium. Next is a presentation of the results of fitting a nested sequence of logistic regression models. The final sections contain a summary of the results, interpretations based on studies from labor economics, implications, and limitations.

Literature Review

Economists have had a longstanding interest in examining the circumstances and causes of disparities in labor market success. In view of the importance of the issue, there is continuing attention to obtaining estimates of wage disparities after accounting for those factors that are available in databases derived from large-scale surveys. Unsurprisingly, study results vary with respect to the proportion of the gender gap explained by these variables. Typically, there is a substantial residual remaining. As it is beyond the scope of this article to present a full review of the voluminous literature, we will only cite a few particularly noteworthy contributions.

Claudia Goldin presents a comprehensive review of the gender wage gap over the course of the twentieth century.[9] She argues that despite some evident convergence, structure of the economy along with social norms represent major obstacles to full equality. In a much-cited paper, Francine D. Blau and Lawrence M. Kahn examine wage gap over the period (1988–2010).[10] They find that there was some convergence, particularly during the 1980s. However, as they note, "The gender pay gap declined much more slowly at the top of the wage distribution than at the middle or the bottom."[11] In a survey of the relevant literature, they conclude that traditional human-capital factors do not account for much of the gap, while other factors (e.g., work history, hours worked, industry, and occupation) are more predictive. Blau and Kahn display the results[12] of fitting a log wage equation to data derived from the Panel Study of Income Dynamics (2020). The analytic sample included individuals aged 25 to 64 who worked at least 26 weeks in the previous year. With a full specification, the mean female residual from the male log wage equation was 38%, corresponding to a substantial disadvantage. Note that this specification did not include a measure of cognitive skills. Blau and Kahn also discuss some of the tradeoffs that many women make, namely, time flexibility for lower pay, which has the greatest impact in those professions that reward long hours and continuous employment.[13]

In the same vein, Marianne Bertrand[14] describes results employing U.S. Census data (1970–2000) and data from the American Community Survey (2008–2011). The analytic sample included women between 25 and 64 years, with a college degree, and working full-time, full-year. Taking as the reference distribution the earnings among men working full-time, full year, the share of women at or above the 80th percentile rose from 1.6% in 1970 to 6.2% in 2010. Similarly, in comparison to men working full-time, full year *in the same occupation*, the share of women at or above the 80th percentile rose from 5.8% in 1970 to 10.2% in 2010. Evidently, there was considerable progress over that 40-year period, but very significant disparities remain. These disparities are even more pronounced at the top 10%, top 1 %, and top 0.1 % of the labor income distribution.[15] In her article, Bertrand surveys the literature comprising studies examining "the challenges that women may face when trying to juggle competing demands on their time in the workplace and the home, particularly when the home includes children."[16] Assaf Rotman and Hadas Mandel (2023) investigate the relative contributions of different factors in accounting for the wage gap.[17] In particular, they consider both gender differences in work-related characteristics and gender-based differences in returns to skills. As differences of the first type diminish, differences of the second type assume greater importance.

Fogg et al. employed first-round PIAAC data to examine the relationship of human capital to labor market success.[18] They focused on the mean monthly earnings of full-time workers ages 25 to 54. They document that within each category of educational attainment, the distributions of foundational skills (literacy and numeracy) are quite broad and, moreover, that mean monthly income (MMI) is approximately monotone increasing in skill levels. They point out that these results speak to the limitations of relying solely on educational attainment as a measure of human capital.

Fogg et al. also fit several families of multiple regression wage equations using log (mean monthly earnings) as the outcome with a broad range of controls.[19] Generally speaking, the partial regression coefficients corresponding to indicators of race/ethnicity are small in magnitude and not statistically significant. By contrast, the coefficients for gender are large in magnitude and statistically significant. The coefficients can be interpreted as indicating that, other things being equal, males on average earned about 25% more than women.

With regard to gender-related wage disparities, there have been a number of international studies utilizing PIAAC data. Lorenzo Cappellari et al. (2016) employed first-round PIAAC data from European Organisation for Economic Co-operation and Development (OECD) countries.[20] Their main focus was on comparing the relative "impact" of schooling and cognitive skills on log (hourly wages). They found that the relative impact differed by family socioeconomic status and student ability. Most relevant to the present study, they demonstrated gender disadvantage (i.e., favoring males) in all countries and, moreover, that the nature of the relationships differed in the two tails of the earnings distribution.

Using essentially the same PIAAC data, Henry Braun employed logistic regression to model the probabilities that an individual's annual income falls either in the highest quartile (Q4) or the lowest quartile (Q1) in the national income distribution.[21] With respect to the former, only individuals employed full-time were included in the analytic data set. Analyses were conducted for each OECD country. After controlling for family background, age, educational attainment, cognitive skills, and occupational category, there was evidence of a substantial gender disadvantage in each country for both Q1 and Q4; that is, holding other factors constant, in comparison to men's incomes, women's incomes were more likely to be located in Q1 and less likely to be located in Q4. It is noteworthy that the structure of the regression relationship differed between Q1 and Q4. Finally, the magnitudes of the gender disadvantages in the United States were nearly the smallest among OECD countries.

Harry J. Holzer and Robert I. Lerman employed first-round PIAAC data using individuals in the age range of 25 to 65.[22] They also argued for the importance of incorporating measures of cognitive skills in wage equations. They employed measures of literacy, numeracy, and problem-solving. Their approach was to segment each proficiency scale into three intervals (low proficiency, proficiency, high proficiency). Prior to fitting wage equations, they presented a number of tables of descriptive statistics. For example, their Table 2 displays the distribution of each cognitive skill by the level of educational attainment. Again, it makes evident the wide distribution of skills within each level. By way of illustration, among high school graduates, nearly 75% had low proficiency in numeracy. Among those with a 4-year degree, it was nearly 37%. They demonstrated that at all levels of educational attainment, even with a large number of control variables, differences in proficiency levels in literacy and/or numeracy were strongly associated with earnings' differentials for both males and females. They did not explicitly estimate gender-based wage disparities.

In a highly cited article, Derek A. Neal and William R. Johnson presented a methodological critique of then-current practices in the estimation of wage equations, with a particular focus on estimating labor market discrimination against minority groups.[23] The two main issues were indicators of human capital and endogeneity. With regard to the former, they argued that educational attainment was a poor proxy for human capital and should be augmented by some measure of cognitive skills. With regard to the latter, they argued that many of the control variables typically employed in such equations could themselves have been affected by discrimination. Consequently, including them in the equation could lead to an underestimation of wage disparities. Their solution was to construct an analytic database from the National Longitudinal Study of Youth that included the Armed Forces Qualification Test (AFQT) scores. In addition, they employed "reduced form" equations that only incorporated variables measured before entry into the labor market. They found that differences in AFQT scores, along with other control variables, accounted for nearly all of the Black–White wage gap for women and most of the gap for men. They did not directly address gender-based wage disparities. It is worth noting that the R^2 for the fitted equations did not exceed 0.2, so that policy implications of the findings should be made with due caution.

Kevin Lang and Michael Manove carried out theoretical and empirical analyses of the Black–White income gaps.[24] They found that for given levels of cognitive skills (as measured by select components of the AFQT), on average Blacks obtained higher levels of education, presumably in order to present a stronger signal of competence to prospective employers. The authors argued that the perceived need for a stronger signal was in anticipation of a certain degree of labor market discrimination. From a methodological perspective, Lang and Manove concluded that not including educational attainment in the prediction model would lead to biased estimates, notwithstanding the endogeneity of that variable. Although Lang and Manove focus more on males (of all races), we believe their analysis pertains to the present case, inasmuch as women also expect to face discrimination in the labor market. Accordingly, we will incorporate measures of both educational attainment and cognitive skills in our analyses while recognizing that there is necessarily some ambiguity in the interpretations of the results.

It bears mentioning that it is precisely the very substantial dispersion of cognitive skills within each level of educational attainment that makes it feasible to estimate the "cognitive skills premium." In this regard, Franziska Hampf, Simon Wiederhold, and Ludger Woessmann focused on methodological issues; principally, on how to establish causal relations between cognitive skills on the one hand and employment and wages on the other.[25] Their data were drawn from the first two rounds of PIAAC (32 countries), with particular attention to numeracy. Acknowledging the challenges in using cross-sectional data for this purpose, they employed a number of different approaches, including instrumental variables. Further, they systematically investigated threats to the validity of the estimates due to measurement error, reverse causality, and omitted variables. They conducted a number of robustness checks and concluded that there is strong evidence of a causal relationship between numeracy skills and labor market outcomes. On the basis of their results, they further asserted that the usual least squares estimates likely underestimate the true impact of skills on income.

This review of the literature makes clear that wage disparities are an ongoing concern that merits further study. The use of the PIAAC data enables adding a measure of cognitive skills to the usual suite of predictors, thereby addressing a lacuna in much of the relevant literature. Linking our findings to the labor economics literature, particularly explanatory studies of gender disparities, provides a more nuanced view of those disparities that can inform policymakers and other stakeholders.

Programme for the International Assessment of Adult Competencies (PIAAC)

PIAAC is the third in a sequence of periodic surveys conducted under the auspices of the OECD to examine the relationships among adults' demographic and family characteristics, cognitive skills, educational attainment, work experiences, and labor market outcomes. It is conducted as a household survey that targets nationally representative samples of adults ages 16 to 65. The utility of PIAAC is that it offers a common framework for comparing patterns of relationships across countries and, in particular, the contribution(s) of differences in family background, cognitive skills, and educational attainment in accounting for the variation in labor market outcomes. Unlike the administrative databases that are often used to address similar questions, PIAAC provides direct measures of foundational skills such as literacy and numeracy as well as problem solving in technology-rich environments. However, due to the limitations of time, the administration of the cognitive instruments follows a complex strategy according to which each respondent receives only a fraction of the item pool in two or three of the domains. Combining item response theory and latent regression modeling, each respondent is then associated with ten plausible values for each of the three domains. The plausible values can be combined to provide a mean score and an estimate of measurement error associated with the score. For further details, consult the *OECD Technical Report of the Survey of Adult Skills (PIAAC)*.[26] In the first cycle of PIAAC, there were three separate rounds of data collection: 2011–2012, 2013–2014, and 2017. The United States participated in all three rounds.

Analytic Database Construction and Methodological Considerations

Sample Size

This study utilized the restricted-use file that contained data collected in the United States in all three PIAAC rounds combined. Public use files, as well as instructions for how to obtain access to the restricted use file, are provided on the Institute of Education Sciences and National Council on Measurement in Education webpage.[27] The original sample contained records for 12,330 respondents of whom 5,045 were employed full-time.

For the purposes of this study, the sample was limited to respondents with a proper reported age and whose reported ages were between 25 and 54. This led to a reduced sample comprising 6,651 respondents, of whom 3,621 were employed full-time. For each country, the reduced samples should be representative of the target populations for this study, that is, either all adults in the focal age range of 25 to 54 or full-time workers in that age range. However, to obtain the final analytical sample, further filters were employed to remove respondents who were missing income or any of the background variables or who had extreme outlying values. This resulted in a final sample of 4,234 respondents, of whom 3,276 were employed full-time. Note that the filtering resulted in a reduction of about 33% in the sample of all respondents but less than 10% in the employed full-time sample.

Variables

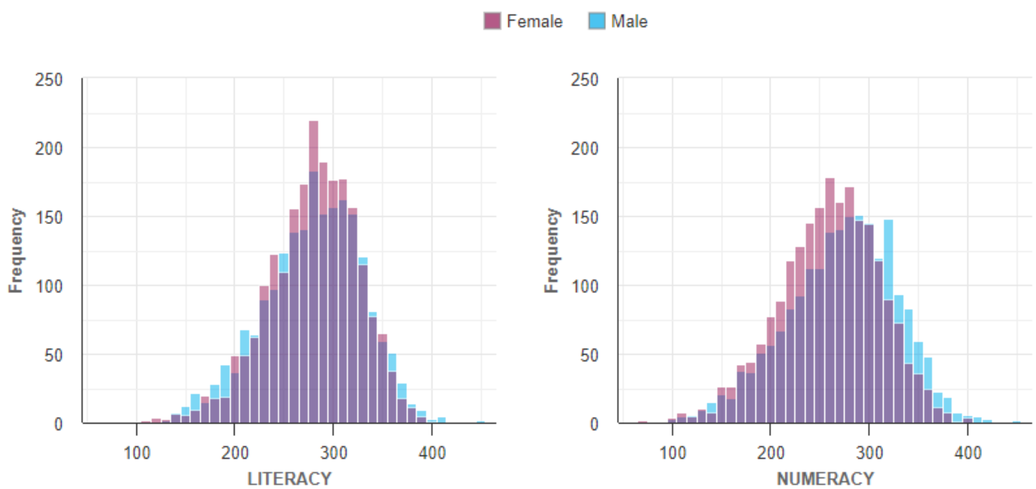
Self-reported data on respondents' monthly earnings, including bonuses, were top-coded at a maximum of $68,553. A yearly income percentile rank variable (YEARLYINCPR) derived from this data and country-level annual income distributions provided in the database was used to create two dummy variables indicating whether the respondent's yearly annual income fell in the lowest quartile (Q1) or the highest quartile (Q4) of the national income distribution. These variables served as criterion variables in the analyses described below.

Information on employment status, age, race, educational attainment, gender, being born in the United States, the U.S. region where the respondent lives, if the respondent has children, and occupational category were employed as predictors. The employment status, which initially had ten categories, was simplified into four categories: full-time employed, part-time employed, unemployed, and other. The "other" employment status category contained respondents who originally responded with one of the following categories: pupil/student, apprentice/internship, in retirement/early retirement, permanently disabled, in compulsory military or community service, fulfilling domestic tasks or looking after children/family, and other.

The educational attainment variable that originally had six categories was coarsened into three categories: lower secondary or less and upper secondary categories were grouped into secondary, nontertiary postsecondary was renamed as nontertiary, and tertiary-professional degree and tertiary-bachelor degree were grouped into tertiary. Additionally, the gender variable was recoded into a dummy variable indicating whether the respondent was a female. Furthermore, the first plausible values for literacy and for numeracy for each respondent were summed and standardized as a proxy for respondents' cognitive skills. The standardization procedure was performed in base R using the scale function which centers the values on the mean (i.e., column average is subtracted from each value) and scales the values by dividing them by the column standard deviation.

Distribution of Skills by Gender

The distributions of literacy scores and numeracy scores for both males and females, for all respondents indicated that scores for males are slightly shifted to the right relative to those for females. The shifts are greater for numeracy than for literacy.



The distributions of literacy scores and numeracy scores for both males and females, for full-time respondents indicated that scores for males are slightly shifted to the right relative to those for females. The shifts are greater for numeracy than for literacy.

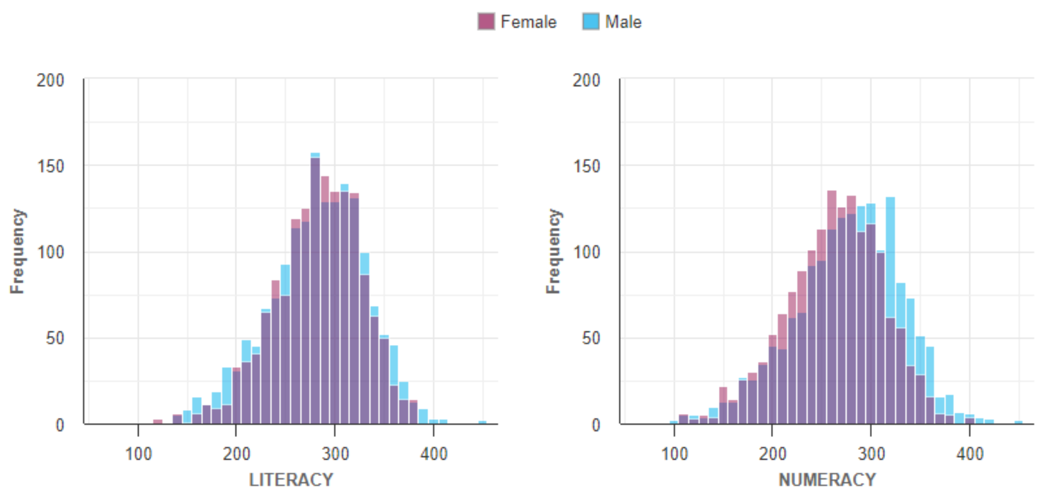


Table 1. Factors employed in analyses, the number of categories corresponding to each factor, and the reference group used in regression analyses

FACTOR NAME	NUMBER OF CATEGORIES	REFERENCE GROUP
Employment status	4	Unemployed
Age (Set 1)	3	25–34
Race/ethnicity (Set 1)	4	White
Educational attainment (Set 1)	3	Secondary
Gender (Set 2)	2	Male
U.S. born (Set 3)	2	Foreign born
U.S. region (Set 3)	4	Northeast
Children	2	No
Occupational category	4	Elementary

Data Quality

Participants' income information was collected by asking their actual income directly. If the respondents were unwilling or unable to report this information, they were provided with an option to report their income in broad ranges.[28] Among the study sample (6,651 respondents), 58% reported their actual income directly and about 7% of the study sample reported their income in broad ranges when followed up. For these cases, the method for reporting earnings including bonuses (EARNFLAG) was listed as either "Reported directly" (4,088 cases) or "Earnings and/or bonuses imputed" (177 cases). Therefore, yearly income percentile rank information, which was derived from the collected income information (in both broad ranges and actual amounts), was only available for 65% of the study sample. Thirty-five percent of the study sample (2,345 respondents) with no income information was removed from the analyses. For these cases, earnings including bonuses reporting method (EARNFLAG) was reported as either "Valid skip" or "Neither reported nor imputed." From the remaining sample, 40 respondents for whom the information on a subset of predictors (race, educational attainment, U.S.-born, children, and occupational category) was not available were also removed from the study sample.

Among the remaining sample of 4,266 respondents, there were cases with extremely high monthly earnings with bonuses reported. In order to identify outliers systematically, the interquartile range (IQR) for the monthly earnings with bonus was calculated within each stratum defined by the intersection of three factors: educational attainment, race/ethnicity, and gender. Any data point that was more than 4 IQR above the 75th percentile was designated as an outlier. With this procedure, 32 cases were removed from the sample.

Moreover, respondents who reported their race/ethnicity as other and their highest educational attainment level as nontertiary were merged into the tertiary group due to the small cell size (23 respondents). Lastly, for those respondents who did not report their actual income but provided their income in broad ranges instead, missing values for their monthly income with bonus (478 observations) were replaced with the averages given their reported yearly income percentile rank category. The final study sample contained 4,234 respondents of whom 3,276 of them were employed full-time (either employed or self-employed), 611 were part-time employed, 73 were unemployed, and 274 had other employment statuses such as internship or compulsory military or community service. The respondents employed full-time constituted 77% of the study sample and were almost evenly split between females (48%) and males (52%).

Methodological Considerations

The full sample (unscaled) sampling weights are provided in the PIAAC database and they were used in both descriptive analyses and logistic regression models. (Note: No adjustment was made for nonresponse or deleted cases.) Analyses were carried out in R using the *survey* package,[29], [30] which accommodates sampling weights. The *svydesign* function was used to specify the complex survey design for the data. For the descriptive analyses, *svybys, svymean,* and *svyvar* functions were utilized to calculate the weighted group means and variances. In order to fit logistic regression models, the *svyglm* function was used. R scripts for data preprocessing, modeling, and visualizations can be accessed via authors' GitHub page.[31]

Analytic Strategy

The analysis begins with an extensive set of descriptive statistics for the all respondents and full-time respondents samples. These are displayed in both tabular and graphical forms. The goal is to facilitate comparisons among different subgroups. The next section focuses on an examination of the skills premium. This is followed by two sets of logistic regression analyses using location in the Q1 and Q4 income quartiles as criteria, respectively. We employ stage-wise addition of predictors in order to capture the impact of additional predictors on the fitted model. Tue Tjur's D-statistic[32] is used to quantify the models' goodness-of-fit.

Results

Descriptive

Sample size distributions for the full sample and for full-time respondents are contained in Tables 2, 3, 4, and 5. In Tables 2 and 3, the data are disaggregated by gender, educational attainment, and race/ethnicity. The percentages are calculated by race/ethnicity within each stratum defined by Gender x Educational Attainment. The smallest strata are defined by Educational Attainment = Nontertiary. In both tables, for both genders, the largest cells are Whites with Educational Attainment = Tertiary, containing approximately one-third of the sample by gender. Note that the distribution by race/ethnicity within each stratum is very similar in the two tables.

Table 2. Sample size distribution by gender, educational attainment, and race/ethnicity—All respondents

EDUCATIONAL ATTAINMENT	GENDER	WHITE	BLACK	HISPANIC	OTHER	TOTAL
Secondary	Female	55% (446)	19% (155)	20% (163)	6% (46)	100% (810)
Secondary	Male	60% (573)	15% (144)	19% (180)	6% (53)	100% (950)
Nontertiary	Female	63% (127)	21% (43)	15% (31)	—	100% (201)
Nontertiary	Male	73% (128)	14% (25)	13% (22)	—	100% (175)
Tertiary	Female	69% (791)	12% (136)	9% (106)	10% (116)	100% (1,149)
Tertiary	Male	72% (681)	8% (80)	7% (63)	13% (125)	100% (949)
Total	—	65% (2,746)	14% (583)	13% (565)	8% (340)	100% (4,234)

— Not applicable.

Table 3. Sample size distribution by gender, educational attainment, and race/ethnicity—Full-time respondents

EDUCATIONAL ATTAINMENT	GENDER	WHITE	BLACK	HISPANIC	OTHER	TOTAL
Secondary	Female	58% (314)	20% (108)	17% (94)	5% (30)	100% (546)
Secondary	Male	63% (464)	12% (92)	19% (143)	5% (37)	100% (736)
Nontertiary	Female	62% (85)	21% (29)	16% (22)	—	100% (136)
Nontertiary	Male	72% (105)	15% (22)	13% (19)	—	100% (146)
Tertiary	Female	69% (624)	12% (110)	9% (81)	9% (84)	100% (899)
Tertiary	Male	74% (603)	7% (58)	6% (51)	12% (101)	100% (813)
Total	—	67% (2,195)	13% (419)	13% (410)	8% (252)	100% (3,276)

— Not applicable.

In Tables 4 and 5, the data are disaggregated by gender, educational attainment, and occupational category. The percentages are calculated by occupational category within each stratum defined by Gender x Educational Attainment. Occupational Category = Elementary was observed to be the smallest group in the sample, followed by Semi-Skilled, Blue Collar.

Table 4. Sample size distribution by gender, educational attainment, and occupational category—All respondents

EDUCATIONAL ATTAINMENT	GENDER	ELEMENTARY	SEMI-SKILLED, BLUE-COLLAR	SEMI-SKILLED, WHITE-COLLAR	SKILLED	TOTAL
Secondary	Female	13% (102)	8% (62)	51% (414)	29% (232)	100% (810)
Secondary	Male	18% (169)	38% (364)	21% (197)	23% (220)	100% (950)
Nontertiary	Female	6% (12)	5% (10)	45% (90)	44% (89)	100% (201)
Nontertiary	Male	6% (11)	39% (69)	19% (33)	35% (62)	100% (175)
Tertiary	Female	1% (13)	2% (19)	18% (209)	79% (908)	100% (1,149)
Tertiary	Male	3% (26)	8% (80)	14% (129)	75% (714)	100% (949)
Marginal	—	8% (333)	14% (604)	25% (1,072)	53% (2,225)	100% (4,234)

— Not applicable.

Table 5. Sample size distribution by gender, educational attainment, and occupational category—Full-time respondents

EDUCATIONAL ATTAINMENT	GENDER	ELEMENTARY	SEMI-SKILLED, BLUE-COLLAR	SEMI-SKILLED, WHITE-COLLAR	SKILLED	TOTAL
Secondary	Female	8% (44)	10% (52)	50% (272)	33% (178)	100% (546)
Secondary	Male	13% (99)	40% (295)	21% (152)	26% (190)	100% (736)
Nontertiary	Female	4% (5)	3% (4)	39% (53)	54% (74)	100% (136)
Nontertiary	Male	4% (6)	40% (58)	19% (28)	37% (54)	100% (146)
Tertiary	Female	0% (4)	1% (12)	14% (128)	84% (755)	100% (899)
Tertiary	Male	2% (19)	9% (70)	12% (95)	77% (629)	100% (813)
Marginal	—	5% (177)	15% (491)	22% (728)	57% (1,880)	100% (3,276)

— Not applicable.

Income data are contained in Table 6 (all respondents) and Table 7 (full-time respondents). The samples are disaggregated by gender, race/ethnicity, and educational attainment. In both tables, for strata defined by Gender x Race/Ethnicity, mean income increases with higher educational attainment with the largest gap between nontertiary and tertiary. For every combination of Gender x Educational Attainment, Whites had the highest mean income, with Black and Hispanic respondents having approximately similar mean incomes.

As expected, for each distribution the mean exceeds the median, which is consistent with a long-tailed, right-skewed distribution. Both the mean and the median convey useful information. Note that logistic regression modeling as employed below is not sensitive to outliers.

Table 6. Descriptive statistics of top-coded monthly income with bonus (EARNMTHBONUSUS_C) reported in hundreds of dollars, by subgroups—All respondents

EDUCATIONAL ATTAINMENT	RACE/ ETHNICITY	GENDER	MEAN (WEIGHTED)	*SD* (WEIGHTED)	MINIMUM	MEDIAN	MAXIMUM	N
Secondary	White	Female	24	15	0	20	88	446
Secondary	White	Male	38	24	2	31	128	573
Secondary	Black	Female	23	13	0	19	64	155
Secondary	Black	Male	25	17	2	19	82	144
Secondary	Hispanic	Female	20	15	1	15	102	163
Secondary	Hispanic	Male	29	18	0	23	102	180
Secondary	Other	Female	25	14	1	24	56	46
Secondary	Other	Male	28	18	2	26	88	53
Nontertiary	White	Female	27	17	1	24	88	127
Nontertiary	White	Male	45	25	3	43	133	128
Nontertiary	Black	Female	21	13	3	20	54	43
Nontertiary	Black	Male	38	22	9	32	102	25
Nontertiary	Hispanic	Female	21	16	1	19	75	31
Nontertiary	Hispanic	Male	38	20	9	32	75	22
Tertiary	White	Female	44	28	2	39	158	791
Tertiary	White	Male	68	41	0	54	262	681
Tertiary	Black	Female	41	30	2	34	154	136
Tertiary	Black	Male	50	39	5	41	192	80
Tertiary	Hispanic	Female	41	30	2	34	158	106
Tertiary	Hispanic	Male	59	38	9	43	208	63
Tertiary	Other	Female	44	25	4	37	113	116
Tertiary	Other	Male	69	52	2	51	292	125
Total	—	—	—	—	—	—	—	4,234

— Not applicable.

Table 7. Descriptive statistics of top-coded monthly income with bonus (EARNMTHBONUSUS_C) reported in hundreds of dollars, by subgroups—Full-time respondents

EDUCATIONAL ATTAINMENT	RACE/ ETHNICITY	GENDER	MEAN (WEIGHTED)	*SD* (WEIGHTED)	MINIMUM	MEDIAN	MAXIMUM	N
Secondary	White	Female	28	14	3	24	77	314
Secondary	White	Male	42	24	2	34	128	464
Secondary	Black	Female	26	13	6	22	64	108
Secondary	Black	Male	30	16	5	26	82	92
Secondary	Hispanic	Female	26	17	5	20	102	94
Secondary	Hispanic	Male	30	17	5	26	102	143
Secondary	Other	Female	32	11	15	28	56	30
Secondary	Other	Male	34	17	8	31	88	37
Nontertiary	White	Female	32	17	5	27	88	85
Nontertiary	White	Male	50	23	5	50	133	105
Nontertiary	Black	Female	28	12	13	23	54	29
Nontertiary	Black	Male	38	22	10	32	102	22
Nontertiary	Hispanic	Female	27	17	7	25	75	22
Nontertiary	Hispanic	Male	40	19	9	32	75	19
Tertiary	White	Female	50	27	5	42	158	624
Tertiary	White	Male	71	41	0	56	262	603
Tertiary	Black	Female	47	28	11	38	154	110
Tertiary	Black	Male	59	41	14	47	192	58
Tertiary	Hispanic	Female	45	27	2	39	149	81
Tertiary	Hispanic	Male	60	39	9	43	208	51
Tertiary	Other	Female	51	23	5	44	113	84
Tertiary	Other	Male	77	52	6	58	292	101
Total	—	—	—	—	—	—	—	3,276

— Not applicable.

Figure 1 contains boxplots of the monthly income data (with bonus) for all respondents, disaggregated by race/ethnicity, highlighting the differences in the income distributions. The next three figures contain bar charts of the MMI data (with bonus) for all respondents, with different groupings of variables to emphasize different comparisons. Figure 2 emphasizes, for each gender, comparisons across levels of educational attainment within race/ethnicity categories, as well as across race/ethnicity categories within levels of educational attainment. Figure 3 displays comparisons across race/ethnicity categories within gender categories and Figure 4 highlights comparisons across gender categories within race/ethnicity categories for each level of educational attainment.

From these figures, we observe some clear patterns. For both genders, MMI increases with greater educational attainment. Furthermore:

1. For both genders, Whites earn more than the other race/ethnicity categories at each level of educational attainment. At the lower levels of educational attainment, Blacks and Hispanics earn similar amounts, but at the tertiary level, Hispanics earn more than Blacks.

2. At each level of educational attainment and each race/ethnicity category, males earn more than females.
3. At each level of educational attainment and each gender, Whites earn more than the other race/ethnicity categories. At the tertiary level, Hispanics earn more than Blacks.

Figure 1. Distribution of top-coded income with bonus (EARNMTHBONUSUS_C) reported in hundreds of dollars, by race/ethnicity—All respondents

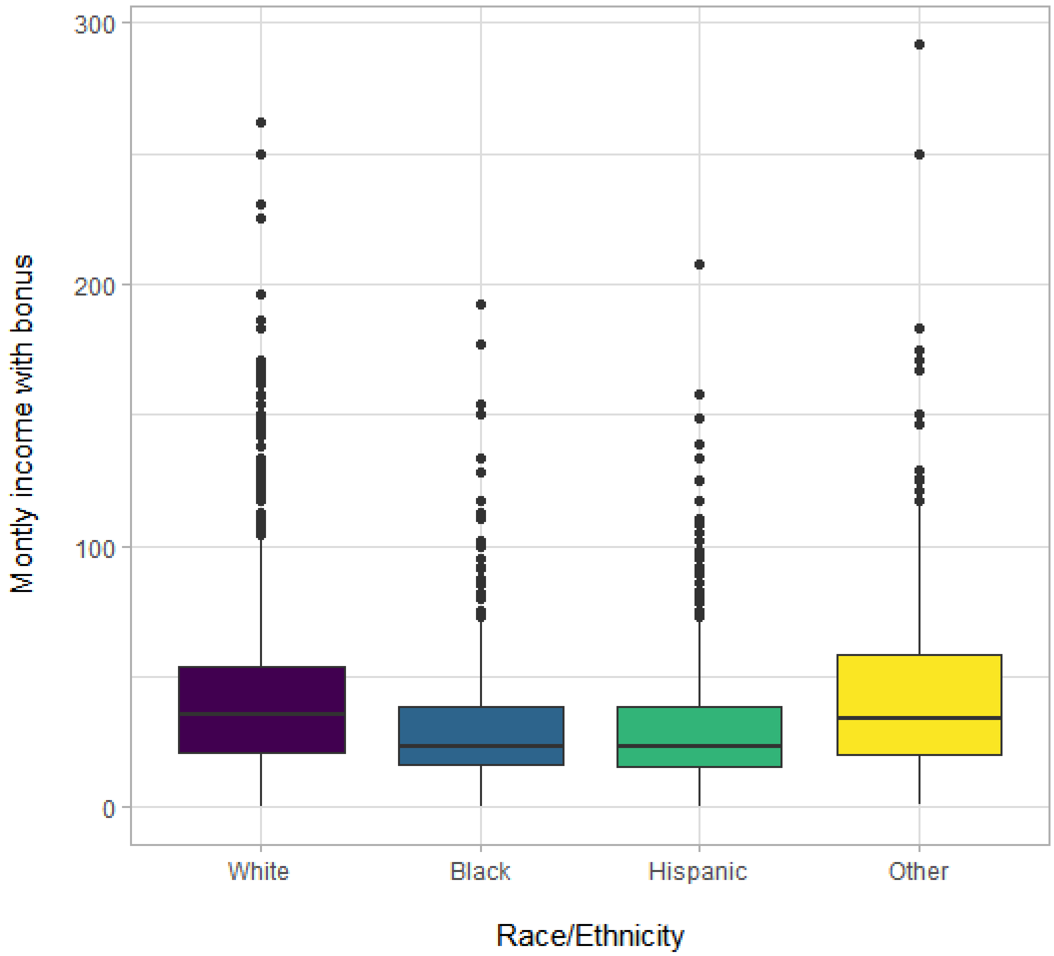


Figure 2. Weighted means of top-coded monthly income with bonus (EARNMTHBONUSUS_C) reported in hundreds of dollars, by educational attainment—All respondents

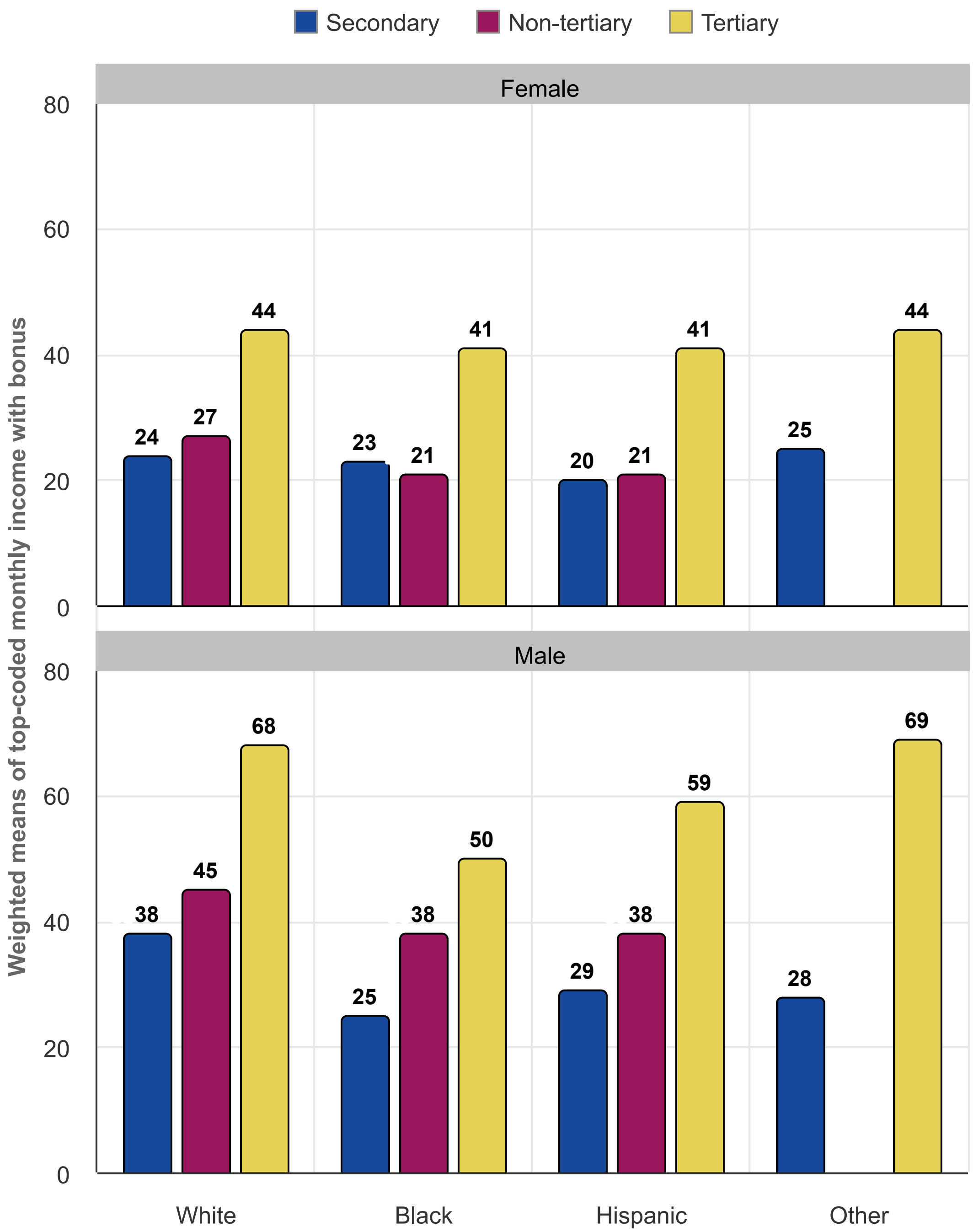


Figure 3. Weighted means of top-coded monthly income with bonus (EARNMTHBONUSUS_C) reported in hundreds of dollars, by race/ethnicity—All respondents

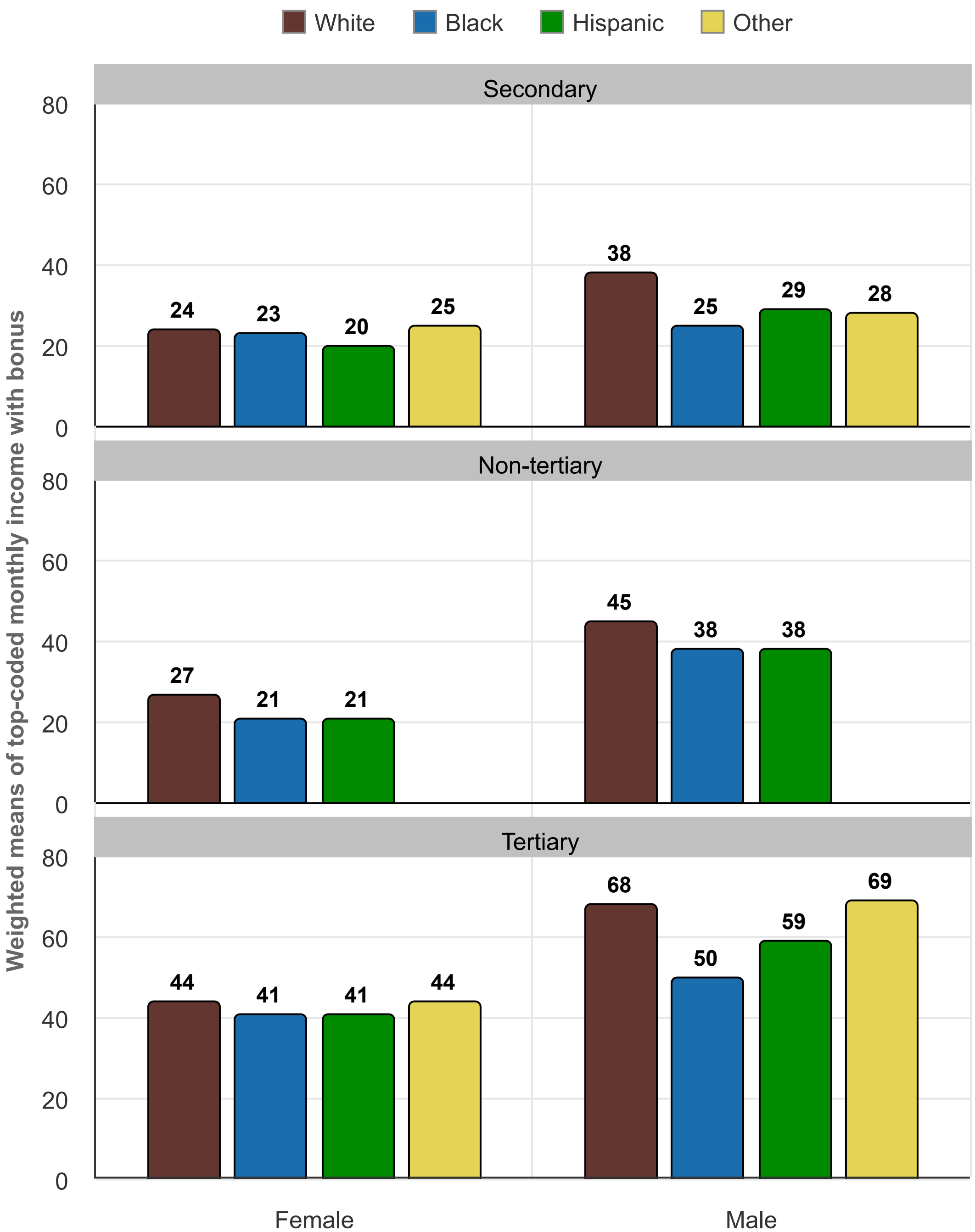
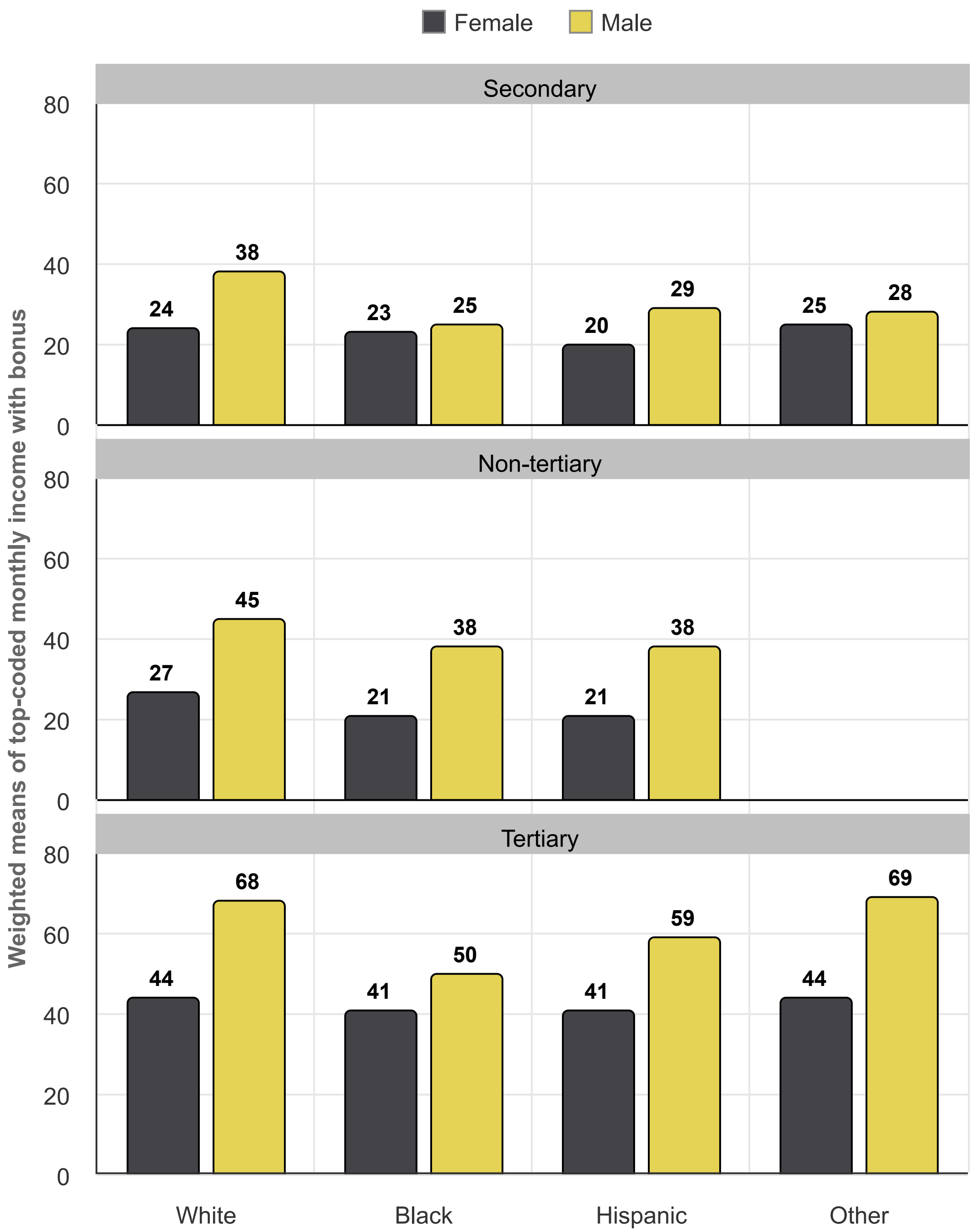


Figure 4. Weighted means of top-coded monthly income with bonus (EARNMTHBONUSUS_C) reported in hundreds of dollars, by gender—All respondents



Weighted averages of cognitive skills are presented by terciles for each level of educational attainment in Table 8 (all respondents) and Table 9 (full-time respondents). We observed that, within each tercile, mean cognitive skills increase with higher educational attainment levels. Note that the trends are very similar in the two tables. In particular, score changes across attainment levels are greatest in the low tercile and least in the high tercile.

Table 8. Weighted means of cognitive skills by terciles and educational attainment—All respondents

EDUCATIONAL ATTAINMENT	LITERACY (PVLIT1)			NUMERACY (PVNUM1)			COMBINED (STLITNUM1)		
	LOW	MEDIUM	HIGH	LOW	MEDIUM	HIGH	LOW	MEDIUM	HIGH
Secondary	201.4	258.6	305.5	179.4	240.3	292.0	-1.6	-0.5	0.5
Nontertiary	232.9	274.3	311.4	211.6	256.3	306.3	-1.0	-0.1	0.7
Tertiary	258.6	302.6	339.9	243.2	293.1	337.4	-0.4	0.5	1.3

Table 9. Weighted means of cognitive skills by terciles and educational attainment—Full-time respondents

EDUCATIONAL ATTAINMENT	LITERACY (PVLIT1)			NUMERACY (PVNUM1)			COMBINED (STLITNUM1)		
	LOW	MEDIUM	HIGH	LOW	MEDIUM	HIGH	LOW	MEDIUM	HIGH
Secondary	203.7	261.5	308.7	183.2	244.5	294.5	-1.6	-0.4	0.6
Nontertiary	232.3	275.3	310.5	212.5	258.4	306.8	-1.0	-0.1	0.7
Tertiary	260.1	303.1	340.1	244.9	294.5	337.5	-0.4	0.5	1.3

In Tables 10 and 11, weighted means of monthly income with bonus are presented for each combination of gender, race/ethnicity, and educational attainment, as well as the level of cognitive skills. Although an increasing trend is observed across cognitive skill levels within almost all strata (except for a few cases), substantial wage gaps are apparent between males and females with the same education attainment level and race/ethnicity. Although the data and trends look similar in the two tables, average monthly earnings were somewhat higher in most cases for full-time respondents.

Table 10. Weighted means of top-coded monthly income with bonus (EARNMTHBONUSUS_C) reported in hundreds of dollars, by terciles of cognitive skills (STLITNUM1) and subgroups—All respondents

EDUCATIONAL ATTAINMENT	RACE/ETHNICITY	GENDER	LOW PROFICIENCY	MEDIUM PROFICIENCY	HIGH PROFICIENCY
Secondary	White	Female	20	23	29
Secondary	White	Male	36	36	42
Secondary	Black	Female	19	24	25
Secondary	Black	Male	20	22	34
Secondary	Hispanic	Female	15	20	26
Secondary	Hispanic	Male	19	30	37
Secondary	Other	Female	23	21	31
Secondary	Other	Male	24	30	28
Nontertiary	White	Female	28	25	29
Nontertiary	White	Male	39	51	46
Nontertiary	Black	Female	20	20	23
Nontertiary	Black	Male	39	42	33
Nontertiary	Hispanic	Female	24	13	34
Nontertiary	Hispanic	Male	26	42	48
Tertiary	White	Female	40	43	48
Tertiary	White	Male	57	68	78
Tertiary	Black	Female	37	38	47
Tertiary	Black	Male	46	38	65
Tertiary	Hispanic	Female	30	45	49
Tertiary	Hispanic	Male	43	68	67
Tertiary	Other	Female	35	48	52
Tertiary	Other	Male	54	73	77

Table 11. Weighted means of top-coded monthly income with bonus (EARNMTHBONUSUS_C) reported in hundreds of dollars, by terciles of cognitive skills (STLITNUM1) and subgroups—Full-time respondents

EDUCATIONAL ATTAINMENT	RACE/ETHNICITY	GENDER	LOW PROFICIENCY	MEDIUM PROFICIENCY	HIGH PROFICIENCY
Secondary	White	Female	24	27	34
Secondary	White	Male	39	40	45
Secondary	Black	Female	23	26	28
Secondary	Black	Male	25	26	37
Secondary	Hispanic	Female	19	27	33
Secondary	Hispanic	Male	21	32	39
Secondary	Other	Female	30	28	38
Secondary	Other	Male	34	35	30
Nontertiary	White	Female	32	29	35
Nontertiary	White	Male	45	54	50
Nontertiary	Black	Female	25	29	29
Nontertiary	Black	Male	39	41	34
Nontertiary	Hispanic	Female	26	20	39
Nontertiary	Hispanic	Male	29	44	49
Tertiary	White	Female	47	47	56
Tertiary	White	Male	59	70	82
Tertiary	Black	Female	45	47	47
Tertiary	Black	Male	52	38	85
Tertiary	Hispanic	Female	36	48	51
Tertiary	Hispanic	Male	47	64	70
Tertiary	Other	Female	40	53	58
Tertiary	Other	Male	57	86	85

Skills Premium

There has been considerable interest in the economics literature regarding the extent to which the labor market rewards skills. This reward is referred to as the "skills premium."[33] Comparisons are typically made among groups of individuals possessing different levels of educational attainment, with educational attainment serving as a proxy for skills. Because there is considerable overlap in the distributions of skills between adjacent levels of educational attainment, such comparisons are necessarily somewhat crude.

As noted earlier, Fogg et al. exploited the PIAAC data to study the relationship of income to various factors, singly and in combination. Fogg et al.'s Charts 9 and 10 display mean monthly earnings (for different levels of educational attainment) as functions of measured literacy skills and numeracy skills, respectively.[34] These charts yield visual indications of the magnitudes of the skills premium. Subsequently, utilizing multivariate regressions with log (earnings) as the outcome, they control for a broad range of variables in order to obtain regression-adjusted estimates of the skills premium for literacy and for numeracy. They find that these estimates correspond to approximately 8% increases in salary associated with a one standard deviation increase in either literacy or numeracy. As noted earlier, Rotman and

Mandel point out that with log (earnings) as the outcome, the fitted coefficient of a predictor variable estimates the relative rate of return of that variable and, consequently, offers an incomplete description of the skills premium when measured in absolute dollars.[35]

In the present study, we employ data from the all respondents sample (Table 10) and for the full-time employed sample (Table 11) to carry out quantitative comparisons among levels of educational attainment (holding composite skill terciles approximately constant), as well as across composite skill terciles (holding educational attainment constant). We find consistent evidence of a substantial skills premium in each set of comparisons. To make the text manageable, we confine our comparisons to those respondents in the secondary and tertiary attainment categories. With respect to race/ethnicity, we confine our comparisons to those individuals membered White, Black, and Hispanic.

(a) All Respondents (Male)

For each combination of race/ethnicity and composite cognitive skill tercile, those with a tertiary education earn substantially more than those with only a secondary education, with the ratios ranging from about 1.5 to 2.2. For example, for Whites in the lowest skill tercile, the MMIs are 3,600 (secondary) and 5,700 (tertiary); for Blacks in the highest skill tercile, the MMIs are 3,400 (secondary) and 6,500 (tertiary); and for Hispanics in the highest skill tercile, the MMIs are 3,700 (secondary) and 6,700 (tertiary). These patterns are quite similar across categories of race/ethnicity.

In Table 8 we observe that the mean composite skill score within a tercile is substantially greater for those with higher levels of educational attainment. Thus, the skills premium referred to above is due not only to work-relevant differences between those in the secondary and tertiary educational attainment categories, but also to (modest) mean differences in cognitive skills between the two categories.

Within levels of educational attainment, those in the highest skill tercile earn more than those in the lowest tercile, with ratios ranging from 1.15 to 2. For example, for Hispanics with a secondary education, the MMIs are $1,900 (low) and $3,700 (high). For Blacks with a tertiary education, the MMIs are $4,600 (low) and $6,500 (high). Overall, the skills premium is slightly greater for Blacks and Hispanics than for Whites. Nonetheless, it is evident that for each combination of educational attainment and skill tercile, Whites earn more than Blacks and Hispanics.

(b) All Respondents (Female)

The patterns are very similar to lowest those of males, as are the skill premium ranges. For example, for Whites in the skill tercile, the MMIs are $2,000 (secondary) and $4,000 (tertiary). For Blacks in the highest skill tercile, the MMIs are $2,500 (secondary) and $4,700 (tertiary). There patterns are similar across categories of race/ethnicity.

Within levels of educational attainment, those in the highest skill tercile earn more than those in the lowest tercile, with ratios ranging from 1.2 to 1.7. For example, for Hispanics with a secondary education, the MMIs are 1,500 (low) and 2,600 (high). For Blacks with a tertiary education, the MMIs are 3,700 (low) and 4,700 (high). Overall, the skills premiums for males and females do not appear to be meaningfully different.

(c) Full-Time Respondents (Male)

For each combination of race/ethnicity and skill tercile, those with a tertiary education earn substantially more than those with only a secondary education, with the ratios ranging from about 1.25 to 2.2. Within levels of educational attainment, those in the highest skill tercile earn more than those in the lowest skill tercile, with ratios ranging from 1.15 to 1.6. In general, the skills premium is greater for Blacks and Hispanics than for Whites, though the former groups have lower baseline MMIs.

(d) Full-Time Respondents (Female)

The patterns are similar to those of males, with skills premiums for educational attainment, controlling for race/ethnicity and level of cognitive skills. For Whites, Blacks, and Hispanics, there is also a skills premium within levels of educational attainment. However, especially at the tertiary level, the skills premium is noticeably smaller for females than for males, ranging from 1.04 to 1.4. This is consistent with the findings of Rotman and Mandel.[36]

Logistic Regression Models

We now present the results of fitting two parallel, nested sequences of logistic regression models. The first sequence of six models employs the All Respondents data with the criterion being that the respondent's income is located in Q1(the lowest quartile) of the national income distribution. The second sequence of six models employs the Full-Time Respondents data with the criterion being that the respondent's income is located in Q4 (the highest quartile) of the national income distribution. Table 1 describes the (discrete) factors employed as predictors, with the exception of the measure of cognitive skills, which is the only continuous factor.

In each sequence, the baseline Model 0 incorporates the following factors: employment status (Q1 only), age, race/ethnicity, and educational attainment. Model 1 adds a measure of cognitive skills. Model 2 adds gender and Model 3 adds factors related to being U.S.-born and region of the country. Model 4 adds the factor related to having children as well as its interaction with gender. The final Model 5 incorporates variables representing occupational categories. Note that specific to the first sequence of models employing the All Respondents data, all models also incorporate respondents' employment status. For each fitted model, we also computed the goodness-of-fit D-statistic.[37] These are contained in Tables 12 and 13.

Keep in mind that the significance levels associated with the estimated odds ratios are provided for illustrative purposes only. They cannot be interpreted in "textbook" fashion in view of the substantial amount of missing data. Further, within each sequence, comparing the odds ratios for a particular variable across models must be done with caution due to the change of scale associated with the different models.[38]

With regard to the analyses for Q1 (Table 12), we note that for Model 1 most variables are not significant with the exception of full-time employment, tertiary education, and cognitive skills (STLITNUM1). Corresponding odds ratios are below 1 implying that (a) respondents with full-time employment are less likely to be located in Q1, (b) respondents with higher educational attainment are less likely to be located in Q1, and (c) the higher their cognitive skills, the less likely are respondents to be located in Q1. In Model 2, the odds ratio for gender is above 1 and highly significant. That is, with the Model 1 variables held fixed, females are substantially more likely to be located in Q1, with estimated odds being 1.5 times greater than males. None of the variables added in Model 3 are significant. In Model 4, both the factor related to having children and its interaction with gender are significant. The odds for the interaction term being below 1 implies that the female disadvantage in odds is greater for females with children. (Note that the inclusion of an interaction term, changes the interpretation of the component main effects.) Finally, in Model 5, respondents in any of the three occupational categories are less likely to be located in Q1 than the reference group (elementary), with skilled and semiskilled, blue-collar occupations being significantly so. That said, Table 12 shows that the D-statistics only slightly improved (about 0.04), indicating that very little of the variation in the criterion is accounted for by the predictors added after Model 1.

Table 13 contains the results of the analyses for Q4. In Model 1, all the variables are strongly significant with the exception of those related to race/ethnicity. When gender is added (Model 2), the corresponding odds ratio is also significant and less than 1. The odds ratio below 1 means that with the Model 1 variables held fixed, females are less likely to be located in Q4. In fact, the male advantage in odds is about 2.4 (= 1/.42). In Model 3, the odds ratios for two regions (Midwest and South) are significant and below 1, implying that with the other variables in the model held fixed, respondents residing in those regions are somewhat less likely to be located in Q4 than respondents in the reference region (Northeast). In Model 4, both the factor related to having children and its interaction with gender are significant. Similar to Q1 results but in the opposite direction; that is, the interaction term is below 1, which implies that the female disadvantage is greater for females with children. In the final model, we observe that skilled full-time respondents are highly likely to be located in Q4, with estimated odds being 8:1 relative to the base category. Semiskilled, blue-collar full-time respondents are also substantially more likely to be located in Q4. The D-statistic for the final model was 0.30, indicating that the model accounts for about 30% of the variation in the criterion. This is typical in the social sciences but low enough to suggest caution in interpretation.

Table 12. Logistic regression results for Q1 analyses

	MODEL 0		MODEL 1		MODEL 2		MODEL 3		MODEL 4		MODEL 5	
VARIABLE	EXP(B)	SIG.	EXP(B)	SIG.	EXP(B)	SIG.	EXP(B)	SIG.	EXP(B)	SIG.	EXP(B)	SIG.
Employment status												
Full-time	0.04	0.00	0.04	0.00	0.04	0.00	0.04	0.00	0.04	0.00	0.04	0.00
Part-time	1.01	0.97	1.13	0.72	0.96	0.90	0.97	0.94	0.90	0.76	0.96	0.90
Other	0.59	0.14	0.69	0.31	0.62	0.19	0.61	0.18	0.58	0.13	0.68	0.31
Age												
35–44	0.88	0.41	0.84	0.27	0.83	0.26	0.84	0.27	0.88	0.41	0.93	0.64
45–54	0.93	0.65	0.86	0.35	0.85	0.31	0.86	0.36	0.90	0.53	0.96	0.79
Race/ethnicity												
Black	1.27	0.19	1.02	0.91	1.03	0.90	0.99	0.96	1.01	0.94	0.97	0.87
Hispanic	1.24	0.28	1.00	1.00	1.00	0.99	1.08	0.75	1.06	0.79	1.04	0.88
Other	1.26	0.31	1.10	0.68	1.13	0.61	1.25	0.38	1.18	0.53	1.14	0.61
Education attainment												
Nontertiary	0.85	0.49	0.95	0.82	0.90	0.67	0.90	0.65	0.91	0.67	0.97	0.90
Tertiary	0.37	0.00	0.47	0.00	0.44	0.00	0.44	0.00	0.44	0.00	0.60	0.01
Cognitive skills (STLITNUM1)	—	—	0.77	0.00	0.77	0.00	0.77	0.00	0.76	0.00	0.81	0.03
Female	—	—	—	—	1.53	0.00	1.53	0.00	0.80	0.38	0.80	0.39
U.S.-born	—	—	—	—	—	—	1.11	0.65	1.07	0.79	1.10	0.70
U.S. region												
Midwest	—	—	—	—	—	—	1.33	0.22	1.37	0.18	1.31	0.25
South	—	—	—	—	—	—	1.37	0.15	1.39	0.13	1.38	0.14
West	—	—	—	—	—	—	1.09	0.74	1.10	0.69	1.17	0.52
Children	—	—	—	—	—	—	—	—	0.51	0.00	0.53	0.00
Female with children	—	—	—	—	—	—	—	—	2.65	0.00	2.54	0.00
Occupational category												
Skilled	—	—	—	—	—	—	—	—	—	—	0.29	0.00
Semiskilled, white-collar	—	—	—	—	—	—	—	—	—	—	0.70	0.10
Semiskilled, blue-collar	—	—	—	—	—	—	—	—	—	—	0.43	0.00
Reference group	1.29	0.45	1.10	0.78	1.01	0.97	0.73	0.49	1.15	0.76	1.71	0.28
Observation	4,234	—	4,234	—	4,234	—	4,234	—	4,234	—	4,234	—
D-statistic	0.30	—	0.31	—	0.32	—	0.32	—	0.33	—	0.35	—

— Not applicable.
NOTE: Exp(B) = Exponentiated coefficient or Odds Ratio; Sig. = significance.

Table 13. Logistic regression results for Q4 analyses

VARIABLE	MODEL 0 EXP(B)	MODEL 0 SIG.	MODEL 1 EXP(B)	MODEL 1 SIG.	MODEL 2 EXP(B)	MODEL 2 SIG.	MODEL 3 EXP(B)	MODEL 3 SIG.	MODEL 4 EXP(B)	MODEL 4 SIG.	MODEL 5 EXP(B)	MODEL 5 SIG.
Age												
35–44	2.07	0.00	2.25	0.00	2.29	0.00	2.31	0.00	2.08	0.00	2.07	0.00
45–54	2.26	0.00	2.71	0.00	2.79	0.00	2.80	0.00	2.51	0.00	2.54	0.00
Race/ethnicity												
Black	0.55	0.00	0.86	0.33	0.91	0.55	0.97	0.82	0.98	0.88	1.07	0.68
Hispanic	0.58	0.00	0.90	0.51	0.84	0.28	0.78	0.16	0.78	0.17	0.80	0.23
Other	1.08	0.62	1.39	0.04	1.31	0.12	1.22	0.31	1.31	0.17	1.40	0.08
Education attainment												
Nontertiary	2.17	0.00	1.88	0.00	2.00	0.00	2.04	0.00	2.06	0.00	1.89	0.00
Tertiary	4.99	0.00	3.15	0.00	3.76	0.00	3.78	0.00	3.94	0.00	2.37	0.00
Cognitive Skills (STLITNUM1)	—	—	1.93	0.00	1.85	0.00	1.85	0.00	1.92	0.00	1.72	0.00
Female	—	—	—	—	0.42	0.00	0.42	0.00	0.90	0.51	0.71	0.04
U.S.-born	—	—	—	—	—	—	1.03	0.00	1.05	0.51	1.06	0.04
U.S. region												
Midwest	—	—	—	—	—	—	0.70	0.00	0.70	0.00	0.80	0.10
South	—	—	—	—	—	—	0.70	0.00	0.70	0.00	0.60	0.00
West	—	—	—	—	—	—	1.10	0.70	1.10	0.70	1.00	0.90
Children	—	—	—	—	—	—	—	—	2.34	0.00	2.22	0.00
Female with children	—	—	—	—	—	—	—	—	0.33	0.00	0.34	0.00
Occupational category												
Skilled	—	—	—	—	—	—	—	—	—	—	7.97	0.00
Semiskilled, white-collar	—	—	—	—	—	—	—	—	—	—	1.42	0.31
Semiskilled, blue-collar	—	—	—	—	—	—	—	—	—	—	2.45	0.01
Reference group	0.19	0.00	0.19	0.00	0.25	0.00	0.30	0.00	0.17	0.00	0.06	0.00
Observation	3,276	—	3,276	—	3,276	—	3,276	—	3,276	—	3,276	—
D-statistic	0.16	—	0.20	—	0.23	—	0.23	—	0.24	—	0.30	—

— Not applicable.
NOTE: Exp(B) = Exponentiated coefficient or Odds Ratio; Sig. = significance.

Discussion

The motivation for this report was the desire to develop a better understanding of the factors associated with differences in labor market success as measured by income. To this end, we drew on three cycles of U.S. PIAAC data, employing both descriptive and model-based methods. Although the analyses yielded no surprises, the quantitative results provided a number of useful insights.

With regard to descriptive statistics, we presented a variety of displays that highlighted comparisons among different groups. For every combination of educational attainment and category of race/ethnicity, males earned more than females. Moreover, the differences in monthly earnings were quite striking. For example, for White respondents, the differences in favor of males ranged from $1,400 (secondary education) to $2,400 (tertiary education). For Black and Hispanic respondents, the differences were not as large, but still substantial. For another combination, at each level of educational attainment, for both genders, White respondents earned more than Black and Hispanic respondents. The differences were particularly large for males. For example, differences in mean monthly earnings between White and Black male respondents ranged from $1,200 (secondary education) to $1,800 (tertiary education). As we have learned from the logistic regressions, these differences appear to be mostly accounted for by differences in measured cognitive skills and in educational attainment.

The data displayed in Tables 10 and 11 indicate that for most combinations of educational attainment, race/ethnicity, and gender, there is a modest to substantial "cognitive skills premium;" that is, the higher the cognitive skills, the greater the earnings. We believe that some of the observed reversals are due to small sample fluctuations. In general, the skills premium is greater for males than for females, particularly for those with tertiary education. These results underscore the importance of incorporating cognitive skills into the explanatory prediction models, especially as the importance of those skills remains strong even with the addition of other explanatory variables.[39] At the same time, we note that when we added variables representing the interactions of (a) gender and cognitive skills and (b) gender and educational attainment to our final Model 5, they made negligible contributions to the explained variance.

Moving beyond descriptive statistics, we fit two sets of individual-level, nested logistic regression models with the criteria being income located in Q1 or Q4 (of the national income distribution), respectively. As noted at the outset, we employed logistic regressions rather than fitting single ordinary regressions to log(income) because of evidence that the patterns of relationships differed at the two tails of the income distribution.[40] For Q1 we employed all respondents in the analytic database; for Q4 we only employed respondents working full-time (either employed or self-employed). Logistic regression enabled us to quantify the strength of the association between each predictor and the criterion, holding all the other predictors fixed.

The outcomes of the Q4 analyses were encouraging, as the amount of variation in the criterion explained by the final model exceeded 30% (D-statistic). Notably, in Model 1, when the cognitive skills variable is added to the background variables in Model 0, the D-statistic increases from 0.16 to 0.20, an improvement of 25%. With the further addition of gender (Model 2) to the predictors in Model 1, the D-statistic increases from .20 to .23, an improvement of 15%. When the set of predictors is augmented by the variables representing occupational categories, the D-statistic increases to .30, a further improvement of 25%.

According to Model 3, controlling for all the other predictors, males were about 2.5 times more likely to be in Q4 than females. Model 4 added variables for (having) children, as well as the interaction of gender and (having) children. Model 5 added the occupational category as a predictor. It is evident that the greatest disadvantage accrues to "females with children." In comparison to this group, males with children are four times more likely to be in Q4, again holding all other factors constant. Furthermore, comparing Model 3 to Model 2, the additional predictors (native-born and regions) did not yield any improvement in the D-statistic. Again, these findings are qualitatively similar to those reported by Fogg et al.[41]

Focusing on the final Model 5, most of the variables employed in the models were strong predictors of a respondent being located in Q4. For example, with all other factors held constant, respondents in the highest educational category were about 2.5 times more likely to be in Q4 than those in the base category. Respondents in the oldest age category were about 2.4 times more likely to be in Q4 than those in the base category. Qualitatively, these comparisons are not surprising, and the magnitudes are similar to those reported by Braun.[42] We note that the coefficients for these variables do not change much from one model to the next, although direct comparisons of coefficients among nested logistic regression models are not advisable.[43]

The results for Model 5 indicate that with regard to race/ethnicity, the odds for Black respondents were not significantly different from the odds for White respondents (base category), while the odds for Hispanic respondents were significantly lower. Respondents in the Other category were 1.4 times more likely to be in Q4 than White respondents.

Since the cognitive skills factor is represented as a single, continuous variable, its coefficient must be interpreted differently: Specifically, a one standard deviation increase in that variable corresponds to an increase of 1.72 in the odds of being located in Q4 (Model 5), other predictors held fixed. One can regard this as the "skills premium" controlling for all other predictors in the model. For the Midwest and South regions, the odds ratios are significantly different (lower) than 1 relative to the base category.

As noted above, the addition of occupational categories, though rather broadly defined, substantially improved model fit. Not surprisingly, in comparison to the base category (unskilled), individuals in the three other categories were more likely to be in Q4, especially those identified as "semiskilled, blue-collar" and "skilled." Members of the latter group are nearly 8 times more likely to be in Q4 than the base group.

Unlike the Q4 sample, which comprises only respondents working full-time, the Q1 sample is heterogeneous with respect to workforce participation. Accordingly, we began (Model 0) with a set of predictors representing different levels of work (base category is "not in workforce"), along with variables for age, educational attainment, and race/ethnicity. The D-statistic = .31. This is almost entirely due to the workforce participation variables, as a model with only the other predictors has a D-statistic = .05 (not shown). In comparison to the base category, full-time workers are 25 times less likely to be located in Q1. Notably, individuals with a tertiary education are 2.7 times less likely to be located in Q1 compared to the base category. Although the two older age groups are less likely to be located in Q1, the odds ratios are not significantly different from that of the younger age group. The odds ratios for the different race/ethnicity categories are neither substantially nor statistically different from one.

In Model 1, we added the cognitive skills variable. Its coefficient is significant with a one standard deviation increase in the score corresponding to a significant odds ratio of 1.3 of being less likely to be located in Q1. However, there was a negligible increase in the D-statistic. According to Model 2, females were 1.53 times more likely than males to be located in Q1 after other factors were held constant. Similar to the Q4 models, the additional predictors (native-born and regions) added in Model 3 were not significant and did not yield a meaningful improvement in the D-statistic. With Model 4, the predictors were augmented by the factor of having children and its interaction with gender. Particularly striking is that females with children were 2.04 times more likely to be located in Q1 than males with children. Nonetheless, the D-statistic only increased by .01.

Finally, with Model 5 we added occupational categories. In comparison to the base category (unskilled), individuals in the other three categories were all less likely to be located in Q1. Particularly salient were the odds ratios of 2.3 for the semiskilled, blue-collar category and 3.4 for the skilled category. The D-statistic increased by .02, to .35.

It is noteworthy that as new sets of predictor variables are added, the coefficients of the variables in the earlier model do not change much. Again, although direct comparisons of coefficients are inadvisable, these results suggest that the different predictor sets do not overlap much in the variance they account for. On the other hand, the addition of a predictor with a highly significant coefficient does not always lead to a substantial improvement in the D-statistic. The latter point suggests further caution in interpreting these coefficients.

Interpretations and Limitations

The analytical results reported here are chiefly of interest because of their potential to inform policy. That said, despite the wealth of information contained in the PIAAC database, there are clear limitations to the strength of the inferences that can be drawn. Recall that, due to missing data on income and some background variables, the size of the analytical sample for all respondents was substantially smaller than that of the database restricted only to the focal age groups. (For the full-time respondents sample, the reduction was only about 10%.) Moreover, we used only the first plausible value and the weights employed were the original weights and had not been rescaled to account for the missing data. Finally, recall that for both Q1 and Q4 most of the variance in the (dichotomous) outcome is not explained by the respective model.

A general issue concerns the accuracy of self-reported data. In the present context, it is difficult to find independent measures of these indicators, particularly because the samples have been selected for a particular age range. The concern is heightened with the data involves sensitive information, such as self-reported income. In this case, respondents were given a variety ways to indicate their income, thus offering them the option that made them the most comfortable communicating private information. Note that flexibility exacted a substantial cost; namely, the complex set of transformations required to place incomes on a common scale (OECD, 2019, chap. 20.4). Although, there was no discussion of evaluating the validity of the income data, that the relationships between self-reported income and other variables conformed to expectations provides some measure of validity. With these cautions in mind, we now delve more deeply into the findings.

First, we note that with regard to Black and Hispanic respondents, only the odds ratio for Hispanic respondents in Q4 was statistically significantly different from 1. These findings were consistent with those of Fogg et al.,[44] who used a different regression model. Although the nature and extent of labor market discrimination against these groups is of ongoing concern, we do not pursue the matter further here. Nonetheless, some of the issues we discuss below regarding possible labor market discrimination against females pertain *mutatis mutandis* to these groups as well.

As the review of the analytic results makes clear, while males overall are less like to be located in Q1, females with children are substantially more likely, with odds being two times greater than that of males with children. By contrast, the results of Model 5 indicate that females overall are somewhat less likely to be located in Q4 and females with children experience a very substantial disadvantage in having incomes located in Q4; specifically, males with children are four times more likely to be in Q4, even after controlling for all measured factors.

The labor economics literature has had a particular focus on wage disparities in the right tail of the income distribution. The Q4 results described above constitute a *prima facia* case for discrimination. However, it is necessary to consider a range of potential explanations that might lead to more nuanced interpretations of these findings. It is well beyond the scope of this article to review the relevant literature and we will only touch on a few key issues. Suffice it to say that there are complex dynamics at play and there remains substantial ambiguity as to the extent to which the female disadvantage is due to statistical discrimination (i.e., that decisions related to hiring and compensation that are based on perceived average characteristics of the group).

Blau and Kahn note: "Under a traditional division of labor by gender in the family, women will anticipate shorter and more discontinuous work lives as a consequence of their family responsibilities."[45] Goldin also analyzes wage differentials in high-income settings due to differences in the average number of hours worked per week and the willingness to adapt to different schedules as determined by work needs.[46] Evidently, women are more likely to experience stop-outs due to having children and more willing (or compelled) to trade time flexibility for greater pay. The former point accounts for the fact that women on average have fewer years of work experience than men of the same age. The latter point has been supported by a number of laboratory studies.[47]

Goldin also provides empirical evidence that in high-paying occupations (e.g., business and management), there is a premium to working long hours and it is precisely in those occupations where female disadvantage is greatest.[48] Bertrand et al. (2010) offers an example from a study conducted among MBA graduates of the University of Chicago Booth School of Business.[49] Ten years postgraduation, employed females earned about 50% less than employed males. Most of that difference was associated with differences in years of actual labor market experience and with average weekly hours. Notably, those differences did not pertain to women graduates without children. Another example is offered by Mary C. Noonan, Mary E. Corcoran, and Paul N. Courant's study of lawyers.[50] The authors found an earnings differential favoring males even after controlling for a number of relevant factors including detailed work history, as well as the size and type of employer. Note that such studies of more homogeneous groups of workers have the advantage of access to a broader range of predictors than is typically available in surveys of larger, more heterogeneous populations.

Both Blau and Kahn and Bertrand provide extensive discussions of the "motherhood penalty," generalizing it to females typically having greater nonmarket responsibilities, consequently leaving fewer hours for compensated work.[51] Bertrand also considers some countervailing forces that, for high-wage females, can reduce hours spent on nonmarket activities (e.g., by finding substitutes).[52] However, simultaneously, there is a trend among middle and upper middle class families with children to spend more time on enrichment activities.

Fogg et al. show that years of experience (as measured in PIAAC) does have substantial explanatory power when the outcome is log(income).[53] However, we find that although this variable has a statistically significant coefficient, it yields little incremental variance (D-statistic) in the dichotomous outcome associated with either Q1 or Q4, beyond that accounted for by the variables in Model 3. This finding is somewhat surprising and merits further investigation.

Blau and Kahn also review the literature on potential explanations of wage disparities related to average male–female differences in psychological attributes, noncognitive skills, and conformance to societal norms.[54] For example, they cite studies that indicate that, relative to males, females tend to be more risk-averse, value competition less, and are less skilled in salary negotiations. This can be particularly influential in higher-paying occupations. By contrast, females tend to have stronger noncognitive skills that are increasingly valued in the marketplace. Societal norms, as well as differential academic and peer support, surely influence (at least to some degree) the choices females make in choosing a college major, deciding on graduate education, job choice, etc.—all of which influence their labor market trajectories. Further, there is a question of how social norms may affect earnings differences between husbands and wives and, in particular, whether there is an implicit preference for wives to not earn more than their husbands. The literature is not consistent on this point. For a review, see Binder and Lam.[55] Goldin makes the case that greater equity will only be achieved with systemic changes and evolving social norms.[56]

The preceding discussion makes the case that an observed wage disparity, even one derived from a statistical model with many controls, cannot be taken at face value as a measure of discrimination. On the other hand, there is nothing to suggest that these wage disparities could be entirely explained if only we measured enough relevant variables. In fact, there is considerable evidence that statistical discrimination does operate and may be particularly powerful in higher-paying occupations. The studies of MBAs[57] and lawyers[58] cited above are certainly supportive of this hypothesis.

Correspondence studies[59] offer evidence of discrimination in decisions regarding which applicants to interview for a posted position. Blau and Kahn cite a number of field and laboratory studies that offer further evidence in this regard.[60] Notably, Claudia Goldin and Cecilia Rouse examine a natural experiment that was created when orchestras switched to blind auditions.[61] They concluded that the switch could account for about one-quarter of the 20% increase of females in the top U.S. symphony orchestras from 1970 to 1996.

After a comprehensive review of the literature, Blau and Kahn conclude that many of the factors described above have contributed (and continue to contribute) to gender-based wage disparities, but with magnitudes that vary over time and by occupation.[62] In fact, despite progress in reducing job segregation, they conclude that "given men's and women's differing skill levels and locations in the economy (by occupation, industry, and firm), overall labor market prices can have a significant effect on the gender wage gap."[63]

Policy Considerations and Conclusions

In view of the above discussion, it is natural to consider what policy initiatives could be undertaken to address the problem. Fortunately, inasmuch as gender-based wage disparities have been studied intensively, at least following World War II, the historical record documents a number of policy initiatives, and there has been considerable research on estimating their impact. As one might expect, the findings are mixed. Again, a full review of the extant research is beyond the scope of this paper and we touch on just a few salient points.

At the federal level, there have been a number of important laws outlawing discrimination in employment, including the Equal Pay Act (1963, 1965), Title VII of the Civil Rights Act (1964), and Title IX of the amended Civil Rights Act (1972). Blau and Kahn surveyed the relevant research and found mixed results.[64] For example, females appeared to make more progress during the 1980s when enforcement of these laws was noticeably weaker. On the other hand, starting in the 1970s, there was substantial growth in female enrollment in professional schools and the beginning of a decline in occupational segregation that has continued apace. During this period, there was particular pressure on federal contractors to address inequities in hiring and wages. On this point, Blau and Kahn cite the work of Fidan Ana Kurtulus, who found that over the period 1973–2003, the share of females (and minorities) in high-paying, skilled occupations grew more at federal contractors than at other employers.[65] At the same time, the impact of these laws may have been masked by other trends such as the very high growth in the numbers of women entering the labor market, which would tend to have a depressive effect on wages.

Another set of policy initiatives concerned work–family issues. The outstanding exemplar is the Family and Medical Leave Act (1993), which mandated that eligible workers be allowed to take up to 12 weeks of unpaid leave for birth, adoption, and other life circumstances, including illness (of the self or a close family member). This proved to be a two-edged sword. On the one hand, it could increase the attachment of females to the firm. On the other, mandated leaves, particularly if they were extended, would both reduce females' work experience and also make them more "expensive," leading to reduced incentives to hiring females of child-bearing age in the first place (i.e., statistical discrimination). As Bertrand notes, even in the absence of such discrimination, firms may be less likely to assign females to the most important jobs or clients with a concomitant impact on wages. She argues:

> *Overall, there very well might be, in the short-term at least, a trade-off in that these family-friendly policies might succeed on the one hand in further reducing (and maybe even flipping) the gender gap in labor force participation but might, on the other hand, increase the gender gap in earnings, especially at the top of the earnings distribution.*[66]

Bertrand also discussed the impact of affirmative action (i.e., gender quota and the like).[67] Beyond the effects on the females directly benefitting from these actions, as females rise in corporate hierarchies, they are in a position to mentor/assist other females through various networks, as well as to acclimate males to the presence of females in high-level positions.

To conclude, the substantial disparities at the upper tail of the income distribution revealed here are in alignment with those described by Fogg et al., as well as those described in Bertrand and Rotman and Mandel cited in the introduction to this article.[68] That substantial wage disparities remain after accounting for a range of human capital and other factors suggests that purely "benign" (i.e., nondiscriminatory) explanations are unlikely to tell the whole story. Nonetheless, it has proven to be very difficult to disentangle the various factors that may contribute to these wage disparities, principally because some of these factors are difficult to measure and because they often interact in complex ways over time. In addition, wage disparities vary systematically across different occupations and regions, so that a "one size fits all" explanation is not very likely to exist, as the economics literature referenced here makes clear. Unfortunately, communicating these more complex narratives to broader audiences is challenging, particularly for academics trained to write for their peers.[69]

Evidently, statistical discrimination plays some role in the existence of gender-based wage disparities. Employers are acting upon perceptions of the greater cost of females relative to males due to stop-outs and (perceived) lower commitment to work. This can be manifested both in hiring policies and personnel policies such as lower investments in training and slower promotion trajectories. At the same time, there is a great deal of anecdotal evidence that gender-based discrimination and sexism are pervasive in the workplace and are also contributing to the observed wage disparities in both direct and indirect ways. In this regard, legislation can help, but evolving social norms (perhaps driven in part by legislation) are probably more powerful. For example, changing norms can lead, over time, to meaningful shifts in the allocation of nonmarket work in two-partner households. They might also provide some impetus to job redesign so that the tradeoff between flexibility and compensation is not as problematic. In the meantime, there is a continuing need for both large-scale studies and a range of microstudies that focus on specific jobs and subpopulations. The former provides a broad landscape against which the results of the latter can be better contextualized and more accurately interpreted.

About the Authors

Henry Braun is the Boisi Professor Education and Public Policy in the Lynch School of Education & Human Development at Boston College. He holds a Ph.D. in mathematical statistics from Stanford University. A fellow of the American Statistical Association and of the AERA, he is an elected member of the National Academy of Education and American Academy of Arts & Sciences. Dr. Braun is a co-recipient of the 1986 Palmer O. Johnson Award of the AERA and a co-recipient of the NCME's 1999 Award for Outstanding Technical Contribution to the Field of Educational Measurement. He received the Robert L. Linn Distinguished Address award (2018) and the E.F. Lindquist award (2021), both from the AERA. In 2023 he received the NCME Career Award for Contributions to Educational Measurement. His interests include the analysis of large-scale assessment survey data and social policies related to educational opportunity, school and teacher accountability, higher education learning outcomes, and the role of testing in education policy.

Gulsah Gurkan is a data and research scientist specializing in learning analytics, operational psychometrics, and consumer research. She holds a Ph.D. in applied statistics and psychometrics from Boston College. With over ten years of experience in quantitative research and data analytics, she has had the privilege of contributing her expertise to various industries, including non-profit organizations and leading technology firms. Her current research and interests include efficacy and engagement evaluations, personalized learning research, algorithm development and implementation, and international comparative studies.

Endnotes

1 For example: Marianne Bertrand, "Coase Lecture: The Glass Ceiling," *Economica* 85, no. 338 (April 2018): 205–231, https://doi.org/10.1111/ecca.12264; Francine D. Blau and Lawrence M. Kahn, "The Gender Wage Gap: Extent, Trends, and Explanations," *Journal of Economic Literature* 55, no. 3 (September 2017): 789–865, https://doi.org/10.1257/jel.20160995.

2 Read more about the Survey of Adult Skills here: https://www.oecd.org/skills/piaac/.

3 For example: Keith Sill, "Widening The Wage Gap: The Skill Premium and Technology," *Business Review, Federal Reserve Bank of Philadelphia*, no. 4 (2002): 25-32; Assaf Rotman and Hadas Mandel, "Gender-Specific Wage Structure and the Gender Wage Gap in the U.S. Labor Market," *Social Indicators Research* 165, 585–606, https://doi.org/10.1007/s11205-022-03030-4.

4 Neeta Fogg, Paul Harrington, and Ishwar Khatiwada, *The Impact of Human Capital Investments on the Earnings of American Workers* (Princeton, New Jersey: ETS, 2017); Franziska Hampf, Simon Wiederhold, and Ludger Woessmann, "Skills, Earnings, and Employment: Exploring Causality in the Estimation of Returns to Skills," *Large-Scale Assessment in Education* 5, Article 12 (2017), https://doi.org/10.1186/s40536-017-0045-7; Harry J. Holzer and Robert I. Lerman, *Cognitive Skills in the U.S. Labor Market: For Whom Do They Matter?* (Paris, France: PIAAC, 2015), https://static1.squarespace.com/static/51bb74b8e4b0139570ddf020/t/54da7582e4b0cb4c49fc0d9d/1423603074704/Holzer_Lerman_PIAAC.pdf.

5 Blau and Kahn, "The Gender Wage Gap."

6 Rotman and Mandel, "Gender-Specific Wage Structure."

7 Fogg, Harrington, and Khatiwada, *The Impact of Human Capital Investments* (2017).

8 Martin E. Andresen and Emily Nix, "What Causes the Child Penalty? Evidence From Adoptions and Same Sex Couples," *Journal of Labor Economics* 40, no. 4 (2023): 971–1004, https://www.journals.uchicago.edu/doi/10.1086/718565; Bertrand, "Coase Lecture: The Glass Ceiling."

9 Claudia Goldin, *Career and Family: Women's Century-Long Journey Towards Equity* (Princeton, New Jersey: Princeton University Press, 2021).

10 Blau and Kahn, "The Gender Wage Gap."

11 Blau and Kahn, "The Gender Wage Gap," p.789.

12 Blau and Kahn, "The Gender Wage Gap," Table 4.

13 Blau and Kahn, "The Gender Wage Gap."

14 Bertrand, "Coase Lecture: The Glass Ceiling," Table 1.

15 Thomas Piketty, Emmanuel Saez, and Gabriel Zucman, "Distributional National Accounts: Methods and Estimates for the United States," NBER Working Paper No. 22945 (Cambridge, Massachusetts: National Bureau of Economic Research, 2016), http://www.nber.org/papers/w22945.

16 Bertrand, "Coase Lecture: The Glass Ceiling," 205.

17 Rotman and Mandel, "Gender-Specific Wage Structure."

18 Fogg, Harrington, and Khatiwada, *The Impact of Human Capital Investments.*

19 Fogg, Harrington, and Khatiwada, *The Impact of Human Capital Investments*, Table 6.

20 Lorenzo Cappellari, Paolo Castelnovo, Daniele Checchi, and Marco Leonardi, "Skilled or Educated? Educational Reforms, Human Capital, and Earnings," in *Skill Mismatch in Labor Markets*, eds. Solomon Polachek and Konstantinos Tatsiramos, vol. 45 of *Research in Labor Economics*, eds. Giovanni Russo and Konstantinos Tatsiramos (Leeds, United Kingdom: Emerald Publishing), 173–197, https://doi.org/10.1108/S0147-912120170000045005.

21 Henry Braun, "How Long Is the Shadow? The Relationships of Family Background To Selected Adult Outcomes: Results from PIAAC," *Large-Scale Assessments in Education* 6, Article 4 (2018), https://doi.org/10.1186/s40536-018-0058-x.

22 Holzer and Lerman, *Cognitive Skills in the U.S. Labor Market*.

23 Derek A. Neal and William R. Johnson, "The Role of Premarket Factors in Black–White Wage Differences," *Journal of Political Economy* 104, no. 5 (October 1996): 869–895, https://users.ssc.wisc.edu/~gwallace/Papers/Neal and Johnson (1996).pdf.

24 Kevin Lang and Michael Manove, "Education and Labor Market Discrimination," *American Economic Review* 101, no. 4 (June 2011): 1467–1496, http://www.aeaweb.org/articles.php?doi=10.1257/aer.101.4.1467.

25 Hampf, Wiederhold, and Woessmann, "Skills, Earnings, and Employment."

26 OECD, *Technical Report of the Survey of Adult Skills (PIAAC)*, 3rd ed. (Paris: OECD Publishing, 2019), https://www.oecd.org/skills/piaac/publications/PIAAC_Technical_Report_2019.pdf.

27 Visit https://nces.ed.gov/surveys/piaac/datafiles.asp to find public use files and instructions for access to the restricted use file.

28 OECD, *Technical Report.*

29 R Core Team, R: A language and environment for statistical computing. (2022). R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/, (2022)

30 Thomas Lumley, *Survey: Analysis of Complex Survey Samples, R package version 4.0*, (2020).

31 Visit https://github.com/Gulsah-G/braun_and_gurkan_2024.

32 Tue Tjur, "Coefficients of Determination in Logistic Regression Models—A New Proposal: The Coefficient of Discrimination," *The American Statistician* 63, no. 4 (2009): 366–372, https://doi.org/10.1198/tast.2009.08210.

33 For example, see Sill, 2002, and Rotman and Mandel, "Gender-Specific Wage Structure."

34 Fogg, Harrington, and Khatiwada, *The Impact of Human Capital Investments*, Charts 9 and 10.

35 Rotman and Mandel, "Gender-Specific Wage Structure."

36 Rotman and Mandel, "Gender-Specific Wage Structure."

37 Tue Tjur, "Coefficients of Determination."

38 Gulsah Gurkan, Yoav Benjamini, and Henry Braun, "Defensible Inferences from a Nested Sequence of Logistic Regressions: A Guide for the Perplexed," *Large-Scale Assessments in Education* 9, Article 16 (2021), https://doi.org/10.1186/s40536-021-00111-7.

39 Fogg, Harrington, and Khatiwada, *The Impact of Human Capital Investments*, 47–48.

40 Blau and Kahn, "The Gender Wage Gap"; Braun, "How Long Is the Shadow?"

41 Fogg, Harrington, and Khatiwada, *The Impact of Human Capital Investments.*

42 Henry Braun, "How Long Is the Shadow?"

43 Gurkan, Benjamini, and Braun, "Defensible Inferences."

44 Fogg, Harrington, and Khatiwada, *The Impact of Human Capital Investments.*

45 Blau and Kahn, "The Gender Wage Gap, 818.

46 Claudia Goldin, "A Grand Gender Convergence: Its Last Chapter," *American Economic Review* 104, no. 4 (2014): 1091–1119, https://www.aeaweb.org/articles/pdf/doi/10.1257/aer.104.4.1091.

47 Matthew Wiswall and Basit Zafar, "Preference for the Workplace, Investment in Human Capital, and Gender," *Quarterly Journal of Economics* 133, no. 1 (February 2018): 457–507, https://doi.org/10.1093/qje/qjx035.

48 Goldin, "A Grand Gender Convergence."

49 Bertrand, Marianne, Claudia Goldin, and Lawrence F. Katz, "Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors," *American Economic Journal: Applied Economics* 2, no. 3 (2010): 228-255.

50 Mary C. Noonan, Mary E. Corcoran, and Paul N. Courant, "Pay Differences Among the Highly Trained: Cohort Differences in the Sex Gap in Lawyer's Earnings," *Social Forces* 84, no. 2 (2005): 853–872.

51 Blau and Kahn, "The Gender Wage Gap"; Bertrand, "Coase Lecture: The Glass Ceiling."

52 Bertrand, "Coase Lecture: The Glass Ceiling."

53 Fogg, Harrington, and Khatiwada, *The Impact of Human Capital Investments.*

54 Blau and Kahn, "The Gender Wage Gap."

55 For a review, see Ariel J. Binder and David Lam, "Is There a Male-Breadwinner Norm? The Hazards of Inferring Preferences from Marriage Market Outcomes," *Journal of Human Resources* 57, no. 6 (2022): 1885-1914.

56 Goldin, *Career and Family.*

57 Bertrand, Goldin, and Katz, "Dynamics of the Gender Gap."

58 Noonan, Corcoran, and Courant, "Pay Differences Among the Highly Trained."

59 Patrick Kline, Evan Rose, and Christopher Walters, "Systematic Discrimination Among Large U.S. Employers," *Quarterly Journal of Economics* 137, no. 4 (2022): 1963–2036, https://doi.org/10.1093/qje/qjac024.

60 Blau and Kahn, "The Gender Wage Gap."

61 Claudia Goldin and Cecilia Rouse, "Orchestrating Impartiality: The Impact of 'Blind' Auditions on Female Musicians," *American Economic Review* 90, no. 4 (2000): 715–741, https://www.aeaweb.org/articles/pdf/doi/10.1257/aer.90.4.715.

62 Blau and Kahn, "The Gender Wage Gap."

63 Blau and Kahn, "The Gender Wage Gap," 856.

64 Blau and Kahn, "The Gender Wage Gap."

65 Blau and Kahn, "The Gender Wage Gap"; Fidan Ana Kurtulus, "Affirmative Action and the Occupational Advancement of Minorities and Women During 1973–2003," *Industrial Relations* 51, no. 2 (2012): 213–146, https://people.umass.edu/fidan/Kurtulus_IR_AA_Occ_April_2012.pdf.

66 Bertrand, "Coase Lecture: The Glass Ceiling," 224.

67 Bertrand, "Coase Lecture: The Glass Ceiling."

68 See Fogg, Harrington, and Khatiwada, *The Impact of Human Capital Investments*, as well as those described in Marianne Bertrand, "Coase Lecture: The Glass Ceiling;" and Rotman and Mandel, "Gender-Specific Wage Structure."

69 Henry Braun, "Expertise on Offer: Why Isn't Anyone Buying?" *Journal of Educational and Behavioral Statistics* 48, no. 5: (2023), 547–572, https://doi.org/10.3102/10769986231160671.

www.ets.org