# New Findings on the Gender Wage Gap: Evidence from Kenya

Uyanga Byambaa, Edward Miguel, Michael Walker and Samuel Zicheng Wang*

This Version: March 20, 2025
*Conference version, please do not cite*

**Abstract**

The gender wage gap remains persistent across the world, especially in low- and middle-income countries. We estimate the magnitude of the gender wage gap and its key contributors in Kenya utilizing the Kenya Life Panel Survey (KLPS), with its unusually rich set of individual characteristics. Using both regression and causal machine learning methods to estimate the gender wage gap, we find that women earn 78 log points less than men (54%) without adjustments, and 41 log points less (34%) after controlling for education, experience, demographics, occupation, cognition, personality, economic preferences, and job task indices. The latter four account for 18% of the residual gap unexplained by education and occupation. The presence of a large gender wage gap even after accounting for a rich set of potential confounders suggests that discrimination and social norms remain important in hindering women's access to labor market opportunities in Kenya.

*Byambaa: University of California, Berkeley (u.byambaa@berkeley.edu); Miguel: University of California, Berkeley, NBER, and CEGA (emiguel@berkeley.edu); Walker: University of California, Berkeley, and CEGA (mwwalker@berkeley.edu); Wang: University of California, Berkeley (samzwang@berkeley.edu). Acknowledgements: We thank seminar participants at UC Berkeley for helpful comments and suggestions, and Eric Ochieng, Brenda Ochieng, and Innovations for Poverty Action Kenya for data collection. We gratefully acknowledge funding from the U.S. National Institutes of Health (NIH) (#R01-TW05612, #R01-HD044475, #R01-HD090118, #R03-HD064888, #R01-HD108281, and #R01-AG077001), the U.S. National Science Foundation (#SES-0418110, #SES-0962614), the Dioraphte Foundation, Givewell, and the Berkeley Population Center. The study received IRB approval from Maseno University and the University of California, Berkeley.

# 1 Introduction

Women tend to earn less than men in both rich and poor countries, and the presence and persistence of this gender wage gap is a core empirical fact in labor economics. To illustrate, in the US the median salary for women working full-time, year-round is 84% of that of men (U.S. Department of of Labor 2024), while 2022 European Union data shows that women's gross hourly earnings were on average 12.7% below those of men (Eurostat, 2022). In low and middle income countries (LMICs), gender gaps in the labor market are often even larger and empowering women has been a key policy focus; for example, one of the United Nation's Sustainable Development Goals is to "achieve gender equality and empower all women and girls".[1] Various interventions have been implemented by governments and policy organizations to promote gender equality in the labor market, including initiatives to increase women's educational attainment, lower barriers to their entrepreneurship, improve access to jobs and training opportunities, and reduce workplace discrimination.

Yet there remain many gaps in our understanding of the factors contributing to the gender wage gap, especially in poor countries. Many studies have found that the gender wage gap remains large even after controlling for widely available demographic, socioeconomic and educational measures (e.g., Blau and Kahn (2017)). In many LMIC settings, such analyses have been more limited due to the frequent lack of detailed individual-level data on workers, raising the possibility that the observed gender gap could be driven by omitted variables. And even in wealthy countries, many of the factors that could plausibly affect labor market outcomes – including information on individual skills, job content, and economic preferences – are rarely available in commonly used administrative datasets, raising similar concerns.

This present study aims to make progress on this issue by leveraging both highly detailed individual-level human capital and labor market data in Kenya and a combination of econometric and machine learning methods. To start, we document the extent of the gender wage gap in a sample of thousands of both urban and rural Kenya workers. We find that gender is a key predictor of wages: it is the third-most important predictor out of 104 variables in a flexible machine learning model. The raw gender wage gap between women and men in this Kenyan sample is 78 log points, or 54 percent. Echoing existing work, controlling for individual education, work experience, demographics, and occupation, the gender wage gap falls somewhat but it remains large at 50 log points (39 percent).

The first main innovation of the present study is the use of a greatly expanded set of covariates on individual cognitive ability, economic preferences, personality, and job task indices. We find that the women in this sample on average score lower on the cognitive performance index, and also have job tasks with lower analytical and interpersonal complexity. There are also statistically significant and economically meaningful differences between women and men on the Big Five personality

[1] This is Goal 5 of the UN SDGs, `https://sdgs.un.org/goals/goal5`.

measures. Inclusion of these variables as controls accounts for 18% of the remaining gender wage gap, yielding an estimated difference between men and women of 41 log points (34 percent). A Oaxaca-Blinder decomposition delivers a similar punchline: with the full set of both standard individual controls and the expanded set of characteristics, roughly half of the raw gender gap remains (39 log points).

Beyond the gender wage gap, we also document a sizeable a gender gap in labor supply, with women on average working 12.1 fewer hours per week than men (who work 47 hours per week on average), while accounting for the full set of individual controls (including the expanded variables) narrows the gap but it remains large at 6.8 hours. Taken together, the wage and hours worked findings indicate that the gender gap remains persistent and large even when accounting for a far richer set of individual and job covariates than existing research.

The study setting is Kenya, a lower middle income country with a GDP per capita of $1,952 in 2023. Kenya is viewed as a relative success story in terms of gender equality among its regional neighbors, which makes the size and persistence of the gender wage gap (and hours worked gap) even more striking. Women in Kenya achieve educational attainment close to their male peers: in recent years, 74.6% of women and 83.5% of men completed primary education, and 60.5% of women and 70.1% of men completed lower secondary education (The World Bank, 2022). Women's labor force participation rate in 2023 was 72.2%, which is higher than the average among high-income countries (54.4%), Sub-Saharan Africa (60.5%), and South Asia (31.6%) (Ortiz-Ospina et al., 2024).

The present analysis uses the Kenya Life Panel Survey (KLPS), a longitudinal panel dataset of individuals who participated in a public school deworming program in Western Kenya starting in 1998 and have been tracked in multiple survey rounds ever since. The fifth round of the survey (KLPS-5) includes detailed data on over 5,800 prime-age adult individuals living and working in both urban and rural Kenya. Individuals in the KLPS-5 sample range from roughly 34 to 41 years of age (the 10-90 range) and so are broadly at the midpoint of their prime working years; 44% live in urban areas (with 24% in the major cities of Nairobi or Mombasa), and slightly more than half (56%) are female. A major strength of the sample is the fact that all individuals originally attended primary school in the same region of rural western Kenya, and then are tracked and surveyed regardless of where they move over time, helping to alleviate some concerns about selection of workers into particular labor markets. The survey gathers information on individual labor market earnings, labor supply, occupation, education, and demographic information. Beyond these standard measures, KLPS-5 also gathers detailed modules on individual cognitive performance, economic preferences (e.g., risk aversion, time preferences), Big Five personality measures, and the type of tasks that individuals do on the job.

In terms of statistical methods, we combine classic selection-on-observables regressions with machine learning approaches. We use machine learning methods to first predict wages using a number of covariates to examine the importance of gender as a predictor, and subsequently use

Double/Debiased Machine Learning (DML) to allow for more flexibility in the regression model and to confirm the robustness of the results. Following the labor economics literature, we also perform a Oaxaca-Blinder decomposition of the gender wage gap into a component explained by covariates and an unexplained component. We further conduct multiple sensitivity analyses to examine robustness of the main finding that the gender gap is large and persists even when conditioning on a unusually wide set of covariates.

So what can explain the large and persistent gender wage gap in Kenya? We examine the potential roles of discrimination and social norms, leveraging survey questions in KLPS-5 on those topics. Yet perhaps surprisingly, individual survey responses do not provide strong evidence that discrimination or social norms are key drivers. For instance, women in KLPS-5 are no more likely than men to report frequently experiencing discrimination, and in some cases even report fewer instances of mistreatment by colleagues. Neither women nor men cite their gender as a common reason for why they experience discrimination in their daily lives. Furthermore, both men and women in KLPS-5 – who, recall, are young to middle aged Kenyan adults – overwhelmingly express egalitarian views regarding gender roles in work, politics, and family life. For example, 84% of men and 80% of women state that they believe the husband should help the wife with household chores if the wife is working outside the home, and 91% of men and 92% of women say that women should have equal rights and receive the same treatment as men. Of course there is the possibility of some experimenter demand effects when it comes to survey responses but we view these findings as only deepening the puzzle, as the survey patterns argue against a simple discrimination-based explanation for gender gaps in the Kenya labor market, and point to the need for a deeper investigation.

The current study aims to contribute to several strands of research. First, we extend existing work on the gender wage gap. The results here largely affirm a common finding in this literature that conventional human capital variables—educational attainment and labor market experience—together explain relatively little of the gender wage gap in the aggregate. This is in part due to the "reversal" of gender gaps in education, where in many countries (including the US) women have actually overtaken men in terms of years of education, as well as to the substantial reduction in the gender work experience gap in recent decades (Goldin, 2014; Blau and Kahn, 2017). Evidence from rich countries suggests that occupational sorting remains a significant contributor to gender wage gaps (US Department of Labor, 2024), although a substantial portion of the gap remains unaccounted for even after controlling for occupation. For example, Blau and Kahn (2017) find that in 2010 PSID data, 38 percent of the gender wage gap cannot be explained by education, experience, region, race, unionization, industry, and occupation variables. Most recently, Flinn et al. (2024) point to the important role of Big Five personality traits on the gender wage gap in Germany, where such traits account for as much of the gap as labor market experience.

The gender wage gaps is also a leading labor market issue in LMICs (Verick, 2014; Klasen, 2019),

where gaps may be larger due to stronger traditional gender norms and weaker legal protections against discrimination. Appleton et al. (1999) estimate the gender wage gap in Cote D'Ivoire, Ethiopia, and Uganda, and find that education, experience, and marital status help explain a large part of the gender wage differences. Nordman et al. (2011) estimate the gender earnings gap from seven major West African countries with data from 2001-2003, and find large gaps in all the cities in their sample while noting that 46-67% of the gender earnings gap cannot be explained by education, experience, and marital status. In a related study, Nordman et al. (2015) estimate the gender wage gap among a sample of formal urban workers in Bangladesh and are able to measure cognition through literacy and numeracy tests and utilize survey questions about workers' personality; they find that the gender wage gap narrows somewhat but persists when these additional covariates are included. A 2023 UN Women study on Kenya estimates an adjusted gender pay gap of 9.5% after accounting for education, occupation, marital status, and labor market informality, reducing the raw gap by nearly half but highlighting persistent inequities (Alarakhia et al., 2023). Although their sample is nationally representative, wage data is available for only 20% of adults, with less than one-third of these wages imputed, which likely contributes to their estimate of a smaller gender wage gap compared to this study.

The present study also contributes to the literature studying the roles of cognitive and non-cognitive skills in labor market outcomes. Studies from high-income countries have shown that cognitive skills are rewarded in the labor market (Cawley et al., 2001; Heckman et al., 2006; Lin et al., 2018; Hermo et al., 2022). Evidence from LMICs also points to a similar direction. Boissiere et al. (1985) study this with Raven's Matrices test scores and reading and math tests among a sample of urban wage workers in Nairobi, Kenya and Dar es Salaam, Tanzania, and find that cognitive ability is large and statistically significant predictor of earnings. Hanushek and Woessmann (2008) review the evidence from Ghana, Kenya, Morocco, South Africa, Tanzania, Pakistan, and conclude that the available estimates of the impact of cognitive skills on outcomes suggest strong economic returns in LMICs. A recent literature also highlights the returns to non-cognitive skills (Deming, 2017; Weidmann and Deming, 2021; Edin et al., 2022). Starting from Acemoglu and Autor (2011) and Autor and Handel (2013), economists have focused on the role of not only occupations but also worker's task assignments within occupations, and Dicarlo et al. (2016) and Lo Bello et al. (2019) apply this approach using World Bank skill surveys across 11 LMICs.

This paper also contributes to a recent literature in labor economics that applies machine learning techniques to study related questions. Bonaccolto-Töpfer and Briel (2022), Böheim and Stöllinger (2021), Vafa, Athey, and Blei (2023), and Bach et al. (2024) have used causal machine learning methods to estimate the gender wage gap, with varying findings on how different the coefficients are between regression-based versus ML methods. Most existing papers use post-double LASSO (Belloni et al., 2014), and our application of Double/Debiased Machine Learning (DML) (Chernozhukov et al., 2016, 2018) builds on post-double LASSO.

## 2 Data

The Kenya Life Panel Survey (KLPS) began in 2003 to track a representative sample of approximately 7,500 students that were enrolled in grades 2 to 7 at the start of the Primary School Deworming Program (PSDP) from 1998 to 2003.[2] KLPS survey rounds are conducted roughly every five years, and collect detailed socioeconomic, demographic, cognitive and health measures. A notable feature of the KLPS is the commitment to tracking all respondents selected at baseline regardless of whether they have relocated within Kenya or beyond, resulting in high overall effective tracking rates, with 86.5% ever surveyed across all rounds (Walker et al., 2023). A major strength of the sample is the fact that all individuals originally attended primary school in the same region of rural western Kenya, and then are tracked and surveyed regardless of where they move over time, helping to alleviate some concerns about selection of workers into particular labor markets.

This study leverages new data collected as part of the fifth round of KLPS (KLPS-5) in combination with detailed measures collected as part of past survey rounds. KLPS-5 began in 2023 with a visit that included modules on earnings and labor supply; job tasks, complexity and job stress; a detailed cognitive battery; and questions on perceived discrimination. The KLPS-5 analysis sample includes 5,878 individuals aged 31-47 who are in their prime working years. 44% are male and 56% are female. Additionally, 56% of the sample reside in rural areas, 18% reside in Nairobi, and 26% reside in urban areas outside of Nairobi.

Each KLPS survey round has collected information on respondent earnings. In KLPS-5, the survey collects earnings from three sources: agricultural profit, self-employment profit and earnings from wage employment over the past 30 days (12 months for agriculture). We also ask about hours worked on each activity over the past 7 days. We calculate the hourly wage by dividing total weekly earnings by the hours worked, where weekly earnings are the sum of monthly earnings from all three employment sources divided by 30/7=4.29. We exclude individuals who report zero earnings from all three sources or report zero hours worked in the past 7 days. This leaves us with a sample of 5,063 (86.5%) of individuals with a non-zero wage. Using the exchange rate of 1 USD = 140 Kenyan Shillings, an average worker in the KLPS-5 sample works 40.3 hours per week, earns $89 per month, and has an hourly wage of $0.72. The median worker works 40 hours per week, earns $36 per month, and has an hourly wage of $0.24.[3]

In addition to broad employment sectors (agriculture, self-employment and wage work), we also record each individual's primary occupation at the time of the KLPS-5 survey, based on their report of the most important occupation they have worked in over the past 12 months. We standardize these reports into 49 occupation codes, which we can compare to the mix of occupations in the

[2]The PSDP involved an experimentally-assigned rollout of school-based deworming treatment to students Miguel and Kremer (2004).

[3]Unless noted otherwise, all results in U.S. dollars in this paper use the exchange rate of 1 USD = 140 Kenyan Shillings, the exchange rate on July 15, 2023.

most recent (2022) Kenyan census, Demographic and Health Survey (KDHS).[4] We find that the distribution of occupations at the 1-digit level is similar across both datasets (Appendix Table A.2), indicating that our sample captures a broadly representative snapshot of individuals in the Kenyan labor market. The collection of occupation information will also allow for exploring occupation patterns by gender, and for the inclusion of occupation fixed-effects in the analysis.

KLPS data also provides numerous standard demographic variables often included in analyses of the gender wage gap, including age/experience, education, and urban vs. rural status. We can also make use of household rosters to account for potential differences in home care responsibilities between men and women. These factors can significantly impact women's ability to participate in the labor force, as caregiving responsibilities often fall disproportionately on them. We focus on the number of children and the number of older adults over 65 years old in the household in our analyses.[5]

Furthermore, the KLPS-5 survey measures a wide range of typically unobserved variables that could differ between genders and affect individuals' labor market outcomes. First, multiple rounds of KLPS have collected measures of individual cognitive ability, which typically affects earnings and is correlated with educational outcomes. In KLPS-5, cognitive performance was measured using the approach of the Harmonized Cognitive Assessment Protocol (HCAP), which is the largest concerted global effort to date to conduct harmonized large-scale population-representative studies of cognitive aging and evaluates multiple domains: general cognition, memory, executive functioning, language, orientation, and visuospatial processing (Gross et al., 2023). Tests included items related to orientation to place and time; measures of memory including the CERAD 10-Word List Learning, East Boston Memory Test (EBMT), and Logical Memory Test; measures of executive functioning/attention including Raven's Progressive Matrices, Similarities and Differences, Token Test, Go-No-Go, Trail Making, and Symbol Cancellation; measures of language/fluency including animal fluency and various tests of object naming; and measures of visuospatial functioning including CERAD Constructional Praxis, interlocking pentagons, and clock drawing. In total, 27 sections of questions were administered. Responses were aggregated into a general cognition score and five domain-specific scores (executive functioning, orientation, memory, language, visuospatial) following the Confirmatory Factor Analysis (CFA) approach in Gross et al. (2024), which is a common method in the psychology literature that studies cognition for summarizing cognitive performance measures into a few domain-specific scores that are easily interpretable. Details of these cognition tests (and validation that the tests and cognitive scores work well for this sample) are described in Gross et al. (2024). Appendix B.1 includes more details.

[4]Specifically, in the KDHS data, occupation levels were provided at both the 1-digit and 3-digit levels. Using the ISCO-08 classification, we aggregated these to the 2-digit level, enabling a consistent mapping to KLPS occupations. The initial sample size was 134,555, representing individuals aged 31 to 47—comparable to the KLPS sample age range. After excluding individuals who had not worked in the past 12 months, the final sample was reduced to 88,805.

[5]Specifically, the question asks: How many adults from age 66 (other than yourself)/children under the age of 18 are in your household, "eat from the same pot" and spend 4 nights or more in an average week sleeping in your home?

A second area of innovation comes through detailed measurement of job tasks and job characteristics. KLPS-5 administered a job task content module for their primary job; this allows us to investigate if males and females are performing systematically different tasks within the same occupations. These job task questions follow the approach used in the Skills Toward Employment and Productivity (STEP) surveys commissioned by the World Bank (Dicarlo et al., 2016; Lo Bello et al., 2019; De La Rica et al., 2020) and US-based measures of job task content (Acemoglu and Autor, 2011; Autor and Handel, 2013). These measure three dimensions of the task content of a worker's job: analytical tasks, which measure the degree of analytic reasoning skills used at work; interpersonal tasks, which measure the degree of interactive, communication, and managerial skills used at work; routine/manual tasks, which measure the degree of physical movement and repetitive work. We sum up the scores from questions in each category following the approach taken by Lo Bello et al. (2019) and standardize each index weighted by sampling weights. Appendix B.2 includes more details on the specific questions asked and the formula used to construct the three task indices.

KLPS-5 also measures other job characteristics, including the stress level of an individual's primary job. Here, we follow the approach used in the Health and Retirement Surveys (HRS). Respondents were assessed with a 15-item questionnaire taken from the widely used Karasek Job Content Questionnaire (Karasek et al., 1998). They were asked to rate their degree of agreement, on a scale from 1 to 4, with items assessing job demands (e.g., time pressure, amount of work, little job security) and a sense of control or decision-making capacity. We calculate a standardized job stress index from these questions.

We utilize data from previous KLPS rounds to incorporate relevant measures that remain relatively stable over time. As individuals' traits could affect their labor market outcomes, we also bring in personality measures and measures of economic preferences collected as part of KLPS Round 4 (KLPS-4, 2018-2022). In particular, we have data on individuals' Big Five Personality traits—extroversion, agreeableness, openness, conscientiousness, and neuroticism—from 15 questions commonly used in psychology studies that ask about whether they agree or disagree with statements such as "I tend to feel depressed" and "I assume the best of people" on a 1 to 5 scale. We also ask respondents two additional questions—"Is it impossible to escape a predetermined fate?" and "How satisfied are you with life?"—to measure their attitude toward fatalism and their life satisfaction. The economic preferences module elicited individuals' attitude toward taking risks, reciprocity, out-group altruism, time preferences, and ambiguity preferences. We include 9 variables that measure economic preferences in our analysis. Appendix B.3 and B.4 describes the construction of these variables in detail. We also incorporate Raven's test scores from previous KLPS rounds to analyze occupational sorting. Appendix A.5 describes this data. Lastly, we also incorporate the gender norms module from KLPS-4 to provide insights into individuals' attitudes toward gender roles and expectations

In summary, the KLPS data allows for the estimation of mean differences in labor market outcomes – particularly the gender wage gap – overall and while controlling for both a "standard" set of covariates often used in the literature, as well as a much more detailed set of individual- and job-specific factors. We are also able to explore how these covariates vary by gender.

# 3 Descriptive Facts

We begin by establishing some descriptive facts about labor market outcomes by gender, establishing large differences in mean and median earnings by gender. We then look descriptively at several common potential explanations for these patterns: differences in educational attainment and cognitive abilities, occupational sorting, and differences in job tasks (even within an occupation). We also explore the interplay between these factors.

## 3.1 Labor market and education outcomes by gender

We first examine gender differences in labor market outcomes. Table 1 shows the average outcomes for women and men in the KLPS-5 sample. We see striking differences in labor supply, earnings, and wages. On average, men work 47.4 hours per week, 35% higher than women, who work 35.1 hours on average. Men also earn significantly more than women: On average, men's monthly earnings are 141.5 USD, whereas women's monthly earnings are 50.2 USD. Part of the differences in earnings reflects differences in labor supply, but the hourly wage also differs substantially. On average, men earn 1 USD per hour——more than twice as much as women, who earn who earn 0.46 USD per hour on average. Compared to women, men are slightly more likely to engage in self-employment (46.0% vs 41.0%), much more likely to engage in wage employment (53.1% vs 34.1%), and much less likely to experience unemployment over the past 30 days (4.7% vs 15.1%).

## 3.2 Gender differences in education and cognitive performance

One factor that could explain the striking gender differences in labor market outcomes is a difference in educational attainment. However, the gender gap in educational attainment is small (Table 1, row 1). Women on average have slightly lower levels of education than men in our sample: The average woman has attained 8.8 years of education, while the average man has attained 9.9 years of education. This parallels findings in high-income countries where women have closed the education gap with men, sometimes even exceeding men (Goldin, 2014).

However, educational attainment may not be fully capturing all relevant aspects of cognitive performance that could matter in the labor market. Are there substantial gender differences in cognition? Figure 1 shows the average differences, in standard deviations, between men and women for the general cognition score as well as the five component-specific cognition scores. Men outperform women on the general cognition score by 0.21 standard deviations, with notable variation across

specific cognitive skills. The largest disparities are observed in executive function and visuospatial skills, where men score 0.31 and 0.28 standard deviations higher, respectively. Spatial orientation and language show smaller gaps, with men scoring 0.06 and 0.10 standard deviations higher, while memory exhibits the smallest difference at just 0.04 standard deviations. These findings suggest that gender differences in cognitive performance are persistent, though not uniform, across domains.

In addition to looking at midlife cognitive peformance, we can also examine differences in cognitive performance at younger ages. The rightmost column in Figure 1 shows the differences in the "lagged" Raven's Matrices test score collected as part of earlier KLPS rounds, and which were administered about 10-15 years prior to KLPS-5.[6] The sample for this measure includes 5,264 individuals (out of 5,878) who were surveyed in both the earlier and KLPS-5 rounds. This measure reveals the largest gender disparity, with men outperforming women by 0.41 standard deviations. As the Raven's Matrices test corresponds to components of the executive functioning abilities, the observed difference here is consistent with persistent differences in executive functioning over time.

The fact that we document differences in measured cognitive performance raises the possibility that controlling for cognitive performance could reduce the gender wage gap.

### 3.3 Occupational sorting by gender

Another common source of the gender wage gap is occupational sorting. Women may disproportionately work in occupations with lower earning potential due to factors such as preferences, social norms and/or discrimination. Table 2 shows data on primary occupations for respondents. We list the top 20 (out of 49) occupations, sorted by their employment shares in the full sample for both men and women, and report the share of men and women working in these occupations. Women are heavily represented in agriculture and fishing, with 33.2% of them working as farmers compared to 17.1% of men. Retail and commercial activities, such as being a seller or street vendor, are also disproportionately female-dominated, with 17.9% of women and only 5.0% of men engaged in these roles. Similarly, women are much more likely to be employed in domestic work, which accounts for 8.1% of their employment compared to only 0.8% of men's. Some of these occupations are often characterized by lower earning potential, reinforcing income disparities between the genders. In contrast, men dominate several occupations that tend to have higher physical demands or are socially perceived as male-oriented. For instance, 7.7% of men work in fishing, compared to only 0.6% of women. Similarly, bicycle and motorbike taxi work is almost exclusively male, with 6.0% of men participating and no women recorded in this occupation. Construction-related jobs also exhibit stark gender imbalances, with roles such as mason (4.9% for men versus 0.1% for women) and unskilled construction laborer (4.0% for men versus 0.3% for women) heavily skewed toward men.

[6] See Appendix B.5 for more details.

### 3.4 Gender differences in job task contents

Previous discussions have shown that men and women tend to work in different occupations. On top of that, the contents of the jobs that men and women engage in can also differ significantly, even for the same occupation. Figure 2 shows the average differences between men and women in the analytical, interpersonal, and routine/manual task indices. Compared to the average job performed by women, the average job performed by men scores 0.52 SD higher on the analytical task index, 0.67 SD higher on the interpersonal task index, and 0.57 SD higher on the routine/manual task index. The tasks performed by men tend to be high in abstract thinking, more intensive in interacting with others, and feature a larger share of repetitive and manual work. Appendix Table A.1 shows the gender differences in each of the task content measures that are inputs to the indices. The difference in the analytical task index stems from substantial disparities in tasks such as reading long documents (0.39 SD), performing math at work (0.40 SD), and engaging in activities that involve learning new things (0.33 SD). Among activities that demonstrate and develop interpersonal skills, men are much more likely to supervise others (0.48 SD), collaborate with coworkers (0.65 SD), and make presentations at work (0.43 SD). For the components of the routine/manual task index, while men and women are equally likely to perform repetitive tasks at work, men are more likely to operate heavy machinery at work (0.43 SD), repair or maintain equipment at work (0.27 SD), and find their work physically demanding (0.18 SD). These results potentially reflect differences in physical strength and endurance between men and women, again suggesting that accounting for these task-based factors could reduce the gender wage gap.

Figure 6 presents a Radar chart that visualizes the gender differences across five skills and human capital categories—educational attainment, cognition, analytical task index, interpersonal task index, and routine/manual task index—for females and males. The figure displays, in standard deviations, the average male and female outcomes across the above mentioned categories, allowing for easy comparison across domains of skills and human capital measures. The blue line (Male) consistently extends further outward compared to the red line (Female) across all five dimensions, suggesting that men score higher on average in these skill measures. The red line (Female) remains within a lower range across all dimensions, indicating lower measured skill levels or human capital accumulation in these areas compared to males. Among the categories, the largest gender gap is observed in interpersonal skills, where the difference between males and females is 0.67 SD. On the other hand, the smallest gap appears in Cognition, where the difference is 0.21 SD. This indicates that gender differences in cognitive abilities are less pronounced compared to job skill measures and educational attainment.

### 3.5 Gender Differences in Relation of Cognition and Job Task Content by Occupation

One of the unique features of our data is the inclusion of lagged Raven's score, which captures cognitive ability from roughly a decade ago,when respondents were young adults. Appendix A.5 describes these data. This measure enables us to explore the interplay between occupational sorting, cognitive ability and job task content skills—specifically, analytical and interpersonal skills. In other words, we examine how past measures of cognitive ability (the lagged Raven's score) predicts current occupational choices.

Figures A.5 and A.6 demonstrate several key points. For analytical skills among males, the concurrent measure revealed a strong positive association with a coefficient of 1.51 (SE = 0.31, $R^2 \approx 0.58$), while the lagged measure maintained a similarly robust relationship with a coefficient of 1.42 (SE = 0.24, $R^2 \approx 0.66$). These results indicate that cognitive ability—whether measured concurrently or earlier—strongly predicts analytical skills in men. In contrast, among females, the concurrent measure showed a lower coefficient of 0.63 (SE = 0.28, $R^2 \approx 0.23$), but when using the lagged measure, the coefficient increased to 1.06 (SE = 0.21, $R^2 \approx 0.60$). This suggests that earlier cognitive ability may exert a stronger influence on the development or display of analytical skills among women. A similar pattern is observed for interpersonal skills. Among males, the concurrent measure yielded a coefficient of 1.21 (SE = 0.23) with high explanatory power ($R^2$ around 0.5–0.6), while the lagged measure showed a comparable coefficient of 1.13 (SE = 0.18), indicating that both current and past cognitive abilities robustly predict interpersonal skills for men. For females, however, the concurrent measure produced a lower coefficient of 0.43 (SE = 0.28, $R^2 \approx 0.13$), whereas the lagged measure increased the predictive strength to 0.84 (SE = 0.24, $R^2 \approx 0.43$). These findings imply that for women, earlier cognitive performance is a much stronger predictor of later interpersonal abilities compared to current cognitive performance. For routine/manual skills, the associations are weak for both genders. Males display a consistently slight positive association, whereas for females the relationship shifts from a negative association in the concurrent measure to a nearly negligible effect in the lagged measure.

Overall, for both analytical and interpersonal skills, lagged Raven's scores are at least as predictive—and, in the case of females, notably more so—than concurrent scores. This suggests that earlier cognitive performance has a lasting or even increasing influence on these task content skills over time. The differences in predictive strength between concurrent and lagged measures are more pronounced for females than for males, hinting at possible gender differences in how cognitive ability translates into task content skills over time.

Additionally, the figures illustrate a consistent occupational ordering: in every scatter plot—whether based on the concurrent or the lagged Raven's Overall Score—the occupations are listed in the same order (e.g., Farmer, Hawking/selling, Teacher, Domestic work, Fishing, etc.). This stability in occupational sorting indicates that while the magnitude of the cognitive–task content skill

associations may vary with the timing of the cognitive measure, the relative ranking of occupations by their task content skill demands remains unchanged.

The analysis supports the notion that cognitive abilities measured a decade ago can effectively forecast current occupational choices, particularly in jobs that depend on complex cognitive and social skills. Individuals with higher cognitive scores from a decade ago are more likely to develop robust analytical and interpersonal skills, making them better suited for high-skill, professional occupations today, whereas the selection into routine/manual jobs appears less influenced by cognitive ability. This long-term perspective emphasizes the value of early cognitive development, which can shape future occupational trajectories and continues to influence skill development, even after additional on-the-job training.

### 3.6 Gender differences in other observables

We close this section with a discussion on the gender differences in other measures that could affect wages: economic preferences, personalities, and job stress. Personality traits have the potential to explain individual labor market outcomes; studies in the psychology literature tend to find a positive association between earnings and the traits of openness, conscientiousness, and extraversion, and a negative association between earnings and the traits of agreeableness and neuroticism (Alderotti et al., 2023). In light of those findings, we examine gender difference in Big Five personalities, shown in Figure 3 in standard deviations. Men are more likely than women to demonstrate extraversion, agreeableness, conscientiousness, and openness, while women are more likely to than men to demonstrate neuroticism. Among economic preferences measures, women are more likely than men to display altruistic tendencies, both for members of the in-group and out-group.[7] Women are 0.78 SD more likely than men to believe that it is impossible to escape a predetermined fate, when presented with that statement. There are no statistically significant gender differences on other economic preference variables, such as ambiguity, risk aversion, reciprocity, and patience. We do not find gender differences in life satisfaction (diff = 0.002 SD). We also do not find gender differences in the stress level experienced on the job between men and women (diff = 0.01 SD) or the number of older adults over 65 years old in the household (diff = 0.00). However, women on average have 0.34 more children under 18 years old than men in their household (p-value $< 0.001$). Figure 4 illustrates these results. Even though the differences in the number of children is not causal and likely include differences in household characteristics between women and men in the sample, such difference may point to a higher burden for women in the sample to engage in home production, especially because a large majority of the sample (73% of women and 86% of men) are married. In subsequent empirical analyses, we control for these variables along with the others mentioned in this section.

[7] We ask respondents how much they would donate to charity if they unexpectedly received 3,200 shillings ($22.9) today. We ask the same question again, but specifying that the respondent knew that it was a charity helping people in their ancestral home area, and again specifying that the respondent knew that it was a charity helping people in other parts of Kenya other than their ancestral home area.

# 4 Machine Learning Prediction of Wages

Does an individual's gender predict their wages? In this section, we document the role of gender and other observables in predicting wages through a series of machine learning exercises. We use two sets of explanatory variables (i.e., features): the full set of 104 variables that contain education, experience, demographics, occupation, cognition, task indices, personality, preferences, home care, and job stress measures, and a "basic" set of 55 education, experience, demographics, and occupation variables.[8] All exercises use the log wage as the outcome and 5-fold cross-validation with a random forest regressor, which allows features to affect the outcome in a non-linear manner and avoids overfitting by creating random subsets of the features and building smaller trees using those subsets. The results are robust to using other machine learning algorithms, such as LASSO and XGBoost.

We first predict the log wage for the entire sample of individuals. The random forest algorithm achieves an average test R-squared of 0.3 when we use all features, and an R-squared of 0.28 when we use only the basic set of features. Without any covariate adjustment, women earn 80 log points less than men, which represents a 55% difference. On average, the predicted log wage for women is 69 log points less than men, which represents a 50% difference. If we remove gender as a predictor, the R-squared becomes 0.28 with all other features included. While the R-squared only declines modestly, the predicted gender wage gap shrinks substantially if we do not include gender as a predictor: The predicted gender wage gap becomes 41 log points, which is equivalent to women earning 33% less than men. Importantly, removing gender as a predictor substantially lowers the predicted wage for men and increases the predicted wage for women. The average predicted wage becomes 13 log points higher for women and 15 log points lower for men. This highlights the importance of gender as a predictor of wages: even though on average a model without accounting for gender still maintains predictive power, it cannot accurately predict gender-specific wages. On the other hand, if gender does not matter for wages conditional on other observables, a model without gender as a predictor should not produce drastically different results for the average male and female wages, compared to one with gender as a predictor.

Even though random forest models lack the interpretable coefficients of regression models, which allow us to easily observe which explanatory variables affect the outcomes the most, SHAP (SHapley Additive exPlanations) provides a way to examine the influence of each feature on the model's predictions. It builds on an approach from coalitional game theory that measures each player's contribution to the final outcome (Molnar, 2023). SHAP values decompose a prediction into the contributions of each feature for that observation, which are analogous to marginal effects in regression but account for feature interactions and non-linearities inherent in random forests. SHAP has been widely used in machine learning to increase the transparency and interpretability of otherwise complex, non-linear models.

[8] Following standard approaches in the labor literature, years of experience is calculated by (age - years of education - 6) (Card et al., 2018). We include both years of experience and its squared term.

Figure 5 shows the SHAP beeswarm plot for the features with the highest predictive power in predicting log wages. The plot visualizes the relative importance of features in predicting log wages, along with the direction and magnitude of their effects. Features are ranked by their average SHAP value magnitude, with the most important predictors at the top. Residing in an urban area (urban_binary) is the most important predictor, with higher SHAP values (pink points) for individuals in urban areas pushing predicted wages upward, while lower values (blue points) for rural residents decrease predictions. Following urban residency, gender emerges as the second most important predictor. Its SHAP values reveal a consistent negative impact of being female (blue points) on wages, as the majority of points cluster on the left side of the plot. In contrast, being male (pink points) has a positive effect, pushing predictions toward higher wages. This highlights the substantial role of gender in wage disparities, even after accounting for occupation, education, job characteristics, and individual traits. Other influential predictors include the interpersonal task index, educational attainment, analytical index, executive functioning score, and general cognition score. Consistent with our intuition, all of those predictors have a positive effect on predicted log wages. As a robustness exercise, Appendix Figure A.1 shows the SHAP beeswarm plot from a model with only the education, experience, demographics, and occupation variables as predictors. Gender still emerges as the second most important predictor of wages.

We also perform the same exercises to predict hours worked. Recall from Table 1 that men on average work 46.6 hours per week, but women work only 35.3 hours per week. Using the full set of 104 covariates to predict hours worked, the random forest model achieves an R-squared of 0.22. If we remove gender as a predictor, the R-squared becomes 0.21 with all other features included. Without gender as a predictor, the model cannot tailor its prediction to males and females quite well: The average predicted hours worked becomes 1.26 hours higher for women and 1.01 hours lower for men. Appendix Figure A.2 shows the SHAP beeswarm plot that summarizes the contribution of top features to the prediction of hours worked. Gender emerges as the third most important feature, behind the interpersonal index and self-employment (self-employed workers work substantially more hours). Appendix Figure A.3 shows the SHAP beeswarm plot from a model that uses only education, experience, demographics, and occupation as features. Again, gender emerges as the second most important feature, behind only self-employment.

# 5 Regression Analyses

## 5.1 Main results: wages

Having established that many individual characteristics both differ by gender and affect wages, we now use a regression framework to study how the observed gender wage gap once we account for different sets of observable characteristics. We implement the following regression model:

$$\ln Y_i = \beta_0 + \beta_1 Female_i + \beta_2 Sch_i + \beta_3 Exp_i + \beta_4 Exp_i^2 + \beta_5 X_i + \gamma_i + \epsilon_i \quad (1)$$

Here, $Y_i$ is the log hourly wage of individual $i$, $Sch_i$ is the years of schooling completed, $Exp_i$ is the years of experience, $\gamma_i$ includes 49 occupation indicators, and $X_i$ is a vector of person-level controls that include different sets of covariates depending on the specification. Table 3 shows the results. Column 1 shows that the raw, unadjusted gender wage gap is 78 log points, which is equivalent to a 54 percent difference. Column 2 adds controls for years of educational attainment, experience, location, and any types of agricultural work or self-employment. These are standard control variables that, due to lack of available data, studies examining the gender wage gap in a developing context typically include. The inclusion of these variables reduces the gender wage gap to 62 log points, suggesting that some of the gap is explained by observable demographic characteristics. Urban residence has a positive and significant association with log wages (coefficient = 0.62) and being in Nairobi also has a compounding effect, both of which highlight the wage premium associated with urban locations. Column 3 adds 49 occupation fixed effects but without the controls in Column 2. Column 4 includes both occupation fixed effects and controls in Column 2. The coefficient in Column 4 reduces to 50 log points, suggesting that occupation sorting contributes to the gender wage gap. Column 5 introduces controls for cognitive ability, the number of children, and the number of older adults in the household. The gender wage gap narrows slightly to 48 log points. The general cognition score is positively associated with wages: a one standard deviation increase in the general cognition score increases wages by 19 log points. Once the general cognition score is included, the coefficient on education drops by half and becomes no longer statistically significant, which is unsurprising since cognition and educational attainment are positively correlated. Having more children in the household is associated with a small but statistically significant reduction in wages with a coefficient of 5 log points, which potentially points to a childcare penalty on wages.

Column 6 further introduces the three task indices—abstract, interpersonal, and routine/manual—as controls to account for task differences within the same occupation. There is a positive wage premium of 10 log points associated with analytical tasks and a 12 log point premium associated with interpersonal tasks. These suggest that within the same occupations, workers whose on-the-job activities involve more abstract thinking and interacting with others earn higher wages. The gender wage gap falls further to 41 log points, suggesting that part of the reason why men earn more than women can be explained by the differences in the tasks that men and women perform. As shown in Figure 2, women tend to work in jobs that involve a lower degree of analytical and interpersonal tasks. In Column 7, we include the rest of observables that could potentially affect wages: the job stress index, economic preferences and personality variables. We also include the five component-specific cognition scores along with the general cognition score: executive functioning, visuospatial, language, memory, and orientation, as well as the lagged Raven's score from KLPS-3. The coefficient on general cognition score becomes significantly smaller and not statistically significant, as the addition of component-specific cognition scores absorbs some of the effects loaded on the general cognition score in previous columns. The gender wage gap remains at 41 log points, with little movement. Interestingly, high-stress jobs are associated with lower wages, suggesting a

positive correlation between wages and amenities. This kind of positive correlation has also been observed by studies on the US labor market, such as Lamadon et al. (2024) and Sockin (2021). Comparing Column 7 with Column 4, we see that cognition, task indices, personality, and preferences explain 18% of the gender wage gap that was left unexplained by occupation, education and demographic variables. This shows that even after accounting for a comprehensive set of potential confounders, there still remains a large portion of the gender wage gap unexplained.

### 5.2 Main results: hours worked

In our sample, women also work fewer hours than men. How much of the gender difference in labor supply can be explained by observable factors? As increasing women's participation in the labor market has been a key policy focus of many governments and organizations, examining the gender gap in hours worked is of interest in its own right. Table 4 shows the results from implementing the regression model as specified in Equation (1), but changing the dependent variable to weekly hours worked. All individuals are included (except for 1 observation missing gender and 102 observations missing urban/rural status); individuals who do not report working any hours in agriculture, self-employment, and wage employment are coded as working zero hours. Column 1 shows that without any adjustment, women work 12.1 fewer hours per week than men. Column 2 shows that including demographics and education controls narrows the gap to 10.59 hours. As one may expect, higher-educated individuals work more, and engaging in self-employment significantly increases one's hours worked. Column 4 shows that including occupation fixed effects together with demographics and education controls, women on average work 8.73 fewer hours than men. Column 5 introduces additional controls for cognition, the number of children, and the number of older adults in the household, which reduce the gender gap to 8.58 hours. Notably, these additional controls do not significantly affect the coefficients of the other variables, and none of the household composition variables are statistically significant. Column 6 adds controls for three task indices. Analytical and interpersonal task indices are positively associated with hours worked (coefficients of 1.36 and 3.30, respectively), suggesting that workers performing more cognitively demanding or interpersonal tasks tend to work longer hours. This inclusion reduces the gender gap further to 6.98 hours. Column 7 includes personality, economic preferences, and job stress as additional controls, along with component-specific cognition scores and the lagged Raven's score. The gender gap decreases slightly to 6.84 hours, showing that these additional factors explain a small portion of the remaining gap. Comparing Column 7 with Column 4, we observe that the addition of this set of control variables explains approximately 22% of the gender gap in hours worked that remains unexplained by occupation, education, and demographic variables.

### 5.3 Double/Debiased Machine Learning results

As the set of variables that could be correlated with wages and gender is very large—104 to be precise—and many of them are continuous, imposing linearity may bias our estimates. To ac-

count for potential non-linearities, we used a recently developed method, Double/Debiased Machine Learning (DML), from the causal machine learning literature (Chernozhukov et al., 2016, 2018). This method is useful in a context where many factors may affect both the treatment (here, being a woman) and the outcome. DML builds on previous causal machine learning methods, such as post-double LASSO, which have been applied by several labor studies examining the gender wage gap in high-income countries (Bonaccolto-Töpfer and Briel, 2022; Böheim and Stöllinger, 2021; Bach et al., 2024). As the name suggests, DML reduces the problem to first estimating two predictive tasks: predicting the outcome from the controls, and predicting the treatment from the controls. Then DML combines these two predictive models in a final stage estimation. The approach allows for arbitrary machine learning algorithms to be used for the two predictive tasks, while maintaining many favorable statistical properties related to the final model, such as small mean squared error and asymptotic normality (Econ ML User Guide 2023).

Formally, DML makes the following assumptions about the data generating process:

$$
\begin{aligned}
Y &= \theta(X) \cdot T + g(X, W) + \epsilon \\
T &= f(X, W) + \eta \\
&\mathbb{E}[\epsilon | X, W] = 0,\ \mathbb{E}[\eta | X, W] = 0,\ \mathbb{E}[\eta \cdot \epsilon | X, W] = 0
\end{aligned}
$$

Here $Y$ is the outcome, $T$ is the treatment (being female), $X$ is a set of covariates across which there are potentially heterogeneous treatment effects, and $W$ are potential confounders. What is particularly attractive about DML is that it makes no further structural assumptions on $g$ and $f$ and estimates them non-parametrically using arbitrary non-parametric Machine Learning methods (Econ ML User Guide 2023). Specifically, we rewrite the structural equation as

$$
Y - \mathbb{E}[Y | X, W] = \theta(X) \cdot (T - \mathbb{E}[T | X, W]) + \epsilon \tag{2}
$$

and estimate the following two conditional expectation functions with non-parametric machine learning algorithms

$$
\begin{aligned}
q(X, W) &= \mathbb{E}[Y | X, W] \\
f(X, W) &= \mathbb{E}[T | X, W]
\end{aligned}
$$

and compute the residuals

$$
\begin{aligned}
\tilde{Y} &= Y - q(X, W) \\
\tilde{T} &= T - f(X, W) = \eta
\end{aligned}
$$

These are related by the equation

$$
\tilde{Y} = \theta(X) \cdot \tilde{T} + \epsilon
$$

Since $\mathbb{E}[\epsilon \cdot \eta | X] = 0$, we can run a linear regression to estimate $\theta(X)$. In the implementation, we use a random forest regressor for the first-stage model predicting the outcome and a random forest classifier for the first-stage model predicting the treatment (i.e., being female). We start by imposing homogeneous treatment effects, i.e., $\theta(X) = \theta$, but in the next section we include

interaction terms to study how the gender wage gap varies across various segments of the sample.

Table 5 shows the results. The top half of the table illustrates the gender gap in log wages and weekly hours worked using DML with both the basic set of education, experience demographics, and occupation controls (Columns 1 and 3) and a full set of controls that include all observables mentioned in Section 2 (Columns 2 and 4). The gender wage gap remains substantial and similar to those predicted by the linear regression: accounting for the basic controls, the gender wage gap estimated by DML is 57 log points, versus 50 log points estimated by linear regression. Accounting for the full set of covariates, the gender wage gap estimated by DML is 41 log points, versus 40 log points estimated by linear regression. Cognition, task indices, personality, and preferences explain 28% of the gender wage gap that was left unexplained by occupation, education and demographics variables. Because our sample is slightly more female (56%) than male (44%), Columns 3 and 4 account for class imbalance by under-sampling women to achieve gender balance in the analysis sample. This is done to improve the accuracy of the final-stage DML estimates, as the first-stage model predicting being female could tend to over-predict female because it is the more common class. The results are similar to the ones generated from the full sample.

The bottom half of the table shows the DML results in a model where the outcome variable is weekly hours worked. The gender gap in hours worked is 7.07 hours with the basic set of covariates and 4.43 hours with the full set of covariates, which are smaller than the 8.73 and 6.69 hours estimated from linear regression, suggesting that non-linearities likely affect the estimation of the gender gap in hours worked more than the gender wage gap. Columns 3 and 4 perform the same under-sampling exercise to account for class imbalance; the coefficients change slightly but are similar to the ones obtained from the full sample. Taken together, the DML results provide us with more confidence that (1) cognition, job tasks, personality, and economic preference measures help explain part of the gender gap in wages and hours worked, and (2) there remain substantial differences between men and women in their wages and hours worked even after accounting for the full set of covariates observed in KLPS-5.

### 5.4 Oaxaca-Blinder Decomposition

Through regression and DML methods, we have examined how observable factors can explain part of the gender gap in wages and hours worked. Another approach commonly applied in labor economics is the Oaxaca-Blinder decomposition, in which we estimate the wage equations separately for men and women and decompose the raw or unadjusted gender gap into one part explained by observed variables and another unexplained part, following Blinder (1973) and Oaxaca (1973). Specifically, we run the following regressions separately for groups $g = \{m, f\}$, where $Y_g$ denotes log wages and $X_g$ denotes observed covariates:

$$Y_m = X_m B_m + u_m$$
$$Y_f = X_f B_f + u_f$$

where $X_g$ is a vector of coefficients. Let $b_m$ and $b_f$ be respectively the OLS estimates of $B_m$ and $B_f$, and denote mean values with a bar. Then, since OLS with a constant term produces residuals with a zero mean, we have:

$$\overline{Y}_m - \overline{Y}_f = b_m\overline{X}_m - b_f\overline{X_f} = \underbrace{b_m(\overline{X}_m - \overline{X}_f)}_{\text{Explained Part}} + \underbrace{\overline{X}_f(b_m - b_f)}_{\text{Unexplained Part}} \tag{3}$$

The first term on the far right-hand side of Equation (3) is the impact of gender differences in the explanatory variables evaluated using the male coefficients. The second term is the unexplained differential and corresponds to the average female residual from the male wage equation. This residual corresponds to an experiment where we compare a woman's actual wage with her predicted wage from the male equation, and one can think of the predicted wage as the outcome of a discrimination case in which a firm that previously was found to have discriminated against women is now required to treat women the same as it treats men Blau and Kahn (2017). The decomposition is typically performed using men as the reference group, but we perform the decomposition using women as the reference group as well in the appendix. Compared to the pooled-sample approach that we carried out with regressions and DML, Oaxaca-Blinder decomposition structurally assumes that observables affect male and female wages differently, whereas the pooled-sample approach can be considered in a potential outcomes framework where one's gender is as good as randomly assigned and observables should not impact one's labor market outcome differently by gender. As a result, it is expected that the output from the two approaches will differ.

Table 6 shows the decomposition results for the gender wage gap. The part that is explained by the basic set of covariates (education, experience, demographics, occupation) is 30.7 log points, or 39% of the raw gender wage gap; the unexplained part is 47.3 log points, or 61% of the raw gender wage gap. Using the full set of controls, 51% (39.4 log points) of the raw gender wage gap can be explained by covariates, leaving 49% (38.6 log points) of the gap unexplained. These estimates are very close to the pooled-sample regression results that show a 50 and 41 log point gender wage gap unexplained by the basic and full set of controls, as shown in Table 3. Table 7 shows the decomposition results for the gender gap in hours worked. The part of the gap explained by the basic set of covariates (education, experience, demographics, occupation) is 1.35 hours, or 11% of the raw gender hours gap; the unexplained part is 10.75 hours, or 89% of the raw gap. Using the full set of controls, 30% (3.64 hours) of the raw gender hours gap can be explained by covariates, leaving 70% (8.46 hours) of the gap unexplained. These estimates are slightly different from the ones obtained from pooled-sample regressions shown in Table 4, but still within the same range and similar. To complete the analysis, Appendix Tables A.3 and A.4 show the Oaxaca-Blinder decomposition results using female as the reference group, which differ from those presented above as it is known that estimated components may change depending on the choice of the reference group (Oaxaca and Ransom, 1994).

### 5.5 Interaction terms

How does the gender gap vary across locations, education levels, cognitive abilities, and job task contents? In this section we examine the heterogeneity by including interaction terms with the gender dummy. Specifically, we interact the gender dummy with the following variables along which the gender wage gap could differ: urban/rural residency, living in Nairobi, any agricultural work, any self-employment, educational attainment, cognition (general cognition score, each one of the five component-specific cognition scores, lagged Raven's score), experience, experience squared, analytical task index, interpersonal task index, and routine/manual task index. Table 8 shows the results from both the linear regression model (Columns 1 and 2) and the DML model (Columns 3 and 4). Columns 1 and 3 show the interaction terms with the basic set of controls, and columns 2 and 4 show those with the full set of controls. Although many of the interaction terms are not statistically significant, some patterns emerge. Most notably, in both the linear regression and the DML models, the interaction term on the analytical index is negative (-0.14 and -0.15), whereas the interaction term on the interpersonal index is positive (0.18 and 0.19). These magnitudes are practically significant, given that the gender wage gap (adjusted for all covariates) is 0.41. These results may indicate that the gender wage gap is larger among individuals involved in high analytical tasks, but smaller among individuals involved in high interpersonal tasks. One possible explanation is that gender norms are more entrenched in jobs that primarily involve abstract thinking and mathematical skills, but less so in jobs that primarily involve interpersonal relationships and interacting with customers. To address the concern that there may be limited common support between male and female on these task indices, thereby making the interaction terms hard to interpret, in Appendfix Figure A.4 we display the kernal density plots of each one of the three task indices by gender. There is substantial overlap between men and women. Interestingly, the density graph for the interpersonal index shows evidence of bunching for both male and female, but the density graphs for the analytical and routine/manual indices shoe less signs of bunching and are closer to bell-curve shapes.

The gender wage gap is larger among individuals who engage in self-employment, which echoes findings in the literature that female-run microenterprises tend to underperform their male-run counterparts on many indicators in low- and middle-income countries (Jayachandran, 2021). The gender gap is larger for workers in urban areas, though residing in Nairobi compensates for part of the additional gap. The interaction terms with cognition variables show mixed results: the coefficients on general and component-specific cognition scores are generally insignificant, except for language and visuospatial skills, which exhibit a small but significant negative interaction in some models. Due to their being borderline significant, we do not aim to draw definitive conclusions from them.

Furthermore, we also examine how the gender gap in labor supply varies across demographics and human capital measures. Appendix Table A.5 presents the interaction terms in regression and DML models following the same specification as those implemented for the study of the wage

gap. Most coefficients are either close to zero or noisily estimated; however, one observation that stands out is that among individuals whose job involves a high degree of analytical tasks, the gap between women and men in hours worked is smaller. A 1 SD increase in the analytical task index is associated with a 3.15-3.39 smaller gender gap in hours worked, which is a significant magnitude given that the gender gap in hours worked (adjusted for all covariates) is 6.83 hours. This potentially suggests that among high-skilled individuals, women work similar hours to men, though the gender wage gap is larger.

### 5.6 Robustness

We conduct a few robustness exercises. First, in Table 9 we only use wage from employment as the dependent variable, which excludes agricultural and self-employment income. The gender wage gap remains substantial: on average, among people who engage in wage employment, women earn 41 log points less than men, 34 log points less after adjusting for basic controls, and 30 log points less after adjusting for the full set of controls. The sample is substantially smaller—consisting of 2,260 individuals—because the majority of individuals do not engage in wage employment. Cognition, task indices, personality, and preferences explain about 12% of the gender wage gap that is unexplained by education, demographics, and occupation variables. In Table 10, we use log monthly earnings as the dependent variable. The gender gap remains persistent across specifications: on average, in terms of total earnings, women earn 111 log points less than men, 61 log points less after adjusting for basic controls, and 60 log points less after adjusting for the full set of controls. Interestingly, the coefficients on the female dummy in Columns 4 and 7 are almost identical, suggesting that cognition, task indices, personality, and preferences do not help further explain the gender gap in earnings that is unexplained by education, demographics, and occupation variables.

In Appendix Table A.6, we conduct the same regression on hourly log wage as in Table 3, but winsorize the dependent variable at the 95th percentile, which is 5.75 (equivalent to an hourly wage of 313.42 Kenyan Shillings or 2.24 USD), to examine the impact of outlier responses. In Appendix Table A.7, we conduct the same exercise for hours worked, winsorizing the dependent variable at the 95th percentile, which is 89 hours. For both the wage and hours regressions, the coefficients are very similar in magnitude to those obtained from the full sample, which should increase our confidence in the results. Another concern is that the gender gap in wages could be explained by the gender gap in labor supply—the fact that women work fewer hours might lead to women having less time to build up skills on the job, or supervisors might not appreciate female workers who work part time as much as male workers who work full time. Appendix Table A.8 examines this by including weekly hours worked as an additional covariate in both the regression and DML models. We find that the gender wage gap adjusted for the basic and full set of controls are similar to those in the main specification: 59 vs 50 log points with the basic set of controls and 48 vs 41 log points with the full set of controls, for example. The additional control variables explain 19% of

the gender wage gap unexplained by education, experience, demographics, and occupation, similar to the 18% from the main specification.

# 6 Discussion

We have shown that the gender wage gap remains persistent across specifications. The inclusion of cognition, job task context, economic preferences and personalities explain 18% of the gender wage gap left unexplained by education, experience, demographics, and occupation, but it still leaves a large gap of 41 log points. This suggests that other factors must be affecting the labor market outcomes of men and women differently. Recent studies in development economics have examined the role of social norms and discrimination (e.g., Bursztyn et al. 2020; McKelway 2019; Buchmann et al. 2024; see Jayachandran (2021) for a review), safety (Siddique, 2022; Becerra and Guerra, 2023), commuting (Cheema et al., 2023; Chen et al., 2024) and work location (Jalota and Ho, 2024) in affecting female labor supply. It is possible that some of those factors also contribute to lower wages earned by women relative to men.

The KLPS-5 survey contains several questions on discrimination and social norms and attitudes. Table 11 shows how individuals' experience of discrimination differs by gender. We ask individuals if they are treated with less courtesy than other people are, if they are treated with less respect than other people are, if they receive poorer service than other people at restaurants or stores, if people act as if they think the respondent is not smart, if people act as if they are afraid of the respondent, and if they are threatened or harassed. Each variable is coded as 1 if the respondent reports that the experience has happened to them more than a few times a year or more (almost every day, at least once a week, a few times a month), and 0 if they report less than once a year or never. Panel A shows both a difference in means between men and women (third column) and a full regression model including all covariates used in the wage analysis (fourth column). Across all variables, we do not discern any patterns that suggest women are more likely to report having experienced discrimination. Perhaps puzzlingly, women are less likely than men to say that they receive poorer service at restaurants or stores or being threatened or harassed.

We also ask respondents about the main reason they believe these experiences of discrimination occurred; Panel B shows the share of men and women who cite each reason as the main one. "My education or income level" emerges as the most cited reason, with 48% of men and 49% of women who experienced discrimination choosing that option; social group or family background comes next with 18% of men and 19% of women choosing that option. 12% of men and 11% of women believe aspects of their physical appearance is the main reason for discrimination. Only 1.78% of women and less than 1% of men chose gender as the main reason for discrimination. If we include other reasons that respondents offered, only 2.73% of women and less than 1% of men cite their gender as a reason, suggesting that even among women, gender is not a factor they commonly report as contributing to their experiences of discrimination. That said, none of these questions ask

specifically about discrimination experiences in the labor market or hiring process. These results could also reflect differences in self-reporting rather than differences in experiencing discrimination, but further analysis is needed to examine how the experiences of discrimination vary by gender, and whether that affects the labor market outcomes of women and men differently.

We also examine individuals' attitudes toward gender norms in Table 12, using questions administered in the fourth round of KLPS that took place between 2018 and 2021 on 5,103 out of 5,878 individuals in our analysis sample. In general, the vast majority of men and women agree with egalitarian gender norms, with little difference between men and women in their responses. 95% of men and 92% of women believe it is OK for a woman to be a mechanic; 84% of men and 80% of women believe the husband should help the wife with household chores if the wife is working outside the home; 92% of men and women believe women can be good politicians and should be encouraged to stand in elections; 91% of men and 92% of women believe that women should have equal rights and receive the same treatment as men in Kenya; 90% of men and women do not believe the husband has a right to beat his wife if she misbehaves. The only exception is the responses to the statement "Important decisions in the family should not be made by men": only 37% of men and 42% of women agree with that.[9]

At least from individuals' stated responses, the vast majority of men and women are comfortable with women performing 'masculine-coded' jobs, becoming politicians, and receiving equal rights and treatment as men. However, despite the reported comfort level with women in 'masculine-coded' jobs, recall that relatively small shares of women actually worked in these types of positions (Table 2). One potential explanation for this could be the role of social desirability bias in driving the reporting of egalitarian norms. Yet these patterns would seem to rule out a simple story of restrictive social norms for women limiting their labor market opportunities, as a social desirability bias towards equality for women is also indicative of equality norms.

# 7 Conclusion

This paper highlights the existence of a large gender wage gap among a sample of prime working-age Kenyans in the Kenya Life Panel Survey (KLPS), with a raw gap of 78 log points. A key contribution of the current analysis is showing that the gender wage gap persists and remains large after controlling for an extensive set of individual human capital and job-related characteristics, and across different econometric and statistical approaches. While these innovative measures – including individual cognitive performance, job task complexity, economic preferences, and personality traits, which are typically unobserved – help explain 18% of the unexplained gap, a substantial 41 log point difference remains. This central finding suggests that factors beyond observable individual and job characteristics continue to shape labor market disparities between men and women in Kenya.

[9]In Appendix Table A.9, we examine the responses to same set of questions administered in the third round of KLPS which took place between 2011 and 2014; the patterns are very similar.

Despite these pronounced wage gaps, self-reported measures of experiences with everyday discrimination and gender attitudes are not strongly biased against women.

Taken together, these findings present a puzzle: despite incorporating a rich set of typically unobserved factors, a significant portion of the gender wage (and associated labor supply gap) remains unexplained. Future research should explore alternative explanations, including gender differences in structural barriers in the labor market, hiring practices, professional networks, occupational preferences, and the role of informal labor dynamics, to further our understanding and to develop more effective interventions for reducing gender disparities in earnings and employment opportunities.

# Figures and Tables

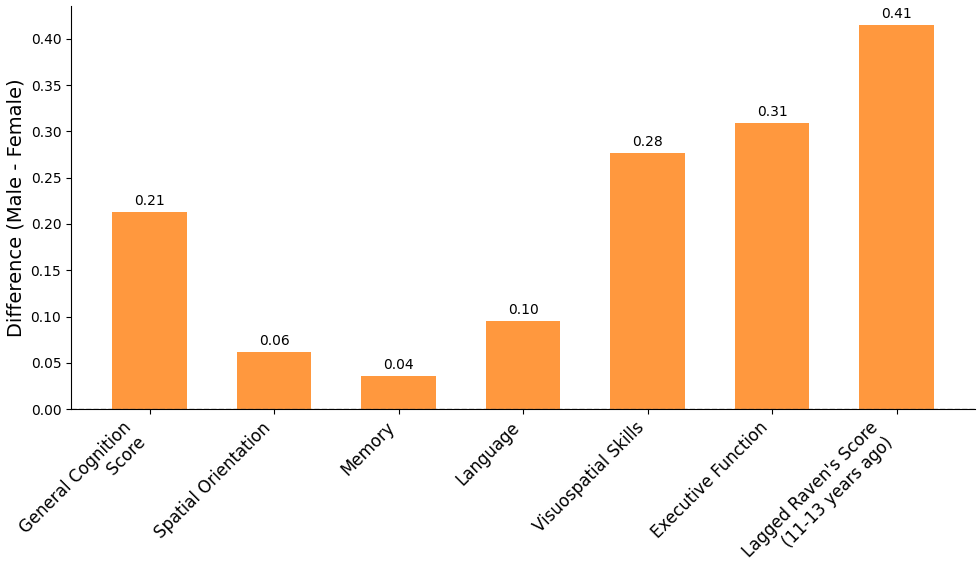



Figure 1: Gender differences in cognition

Note: Figure shows the differences in standard deviations between average male and female cognition scores. General cognition score and the five component scores are obtained from the cognition module in KLPS-5. Lagged Raven's scores are obtained from KLPS-3.

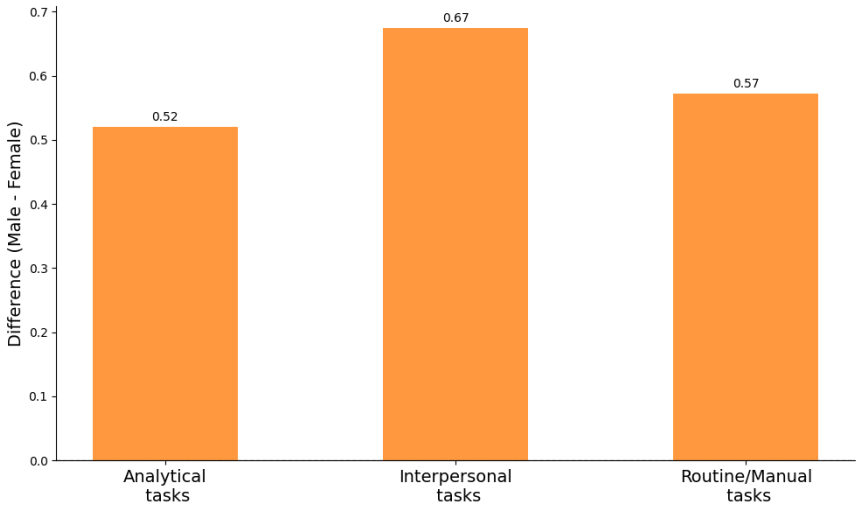



Figure 2: Gender differences in task indices

Note: Figure shows the differences in standard deviations between average male and female task index scores.

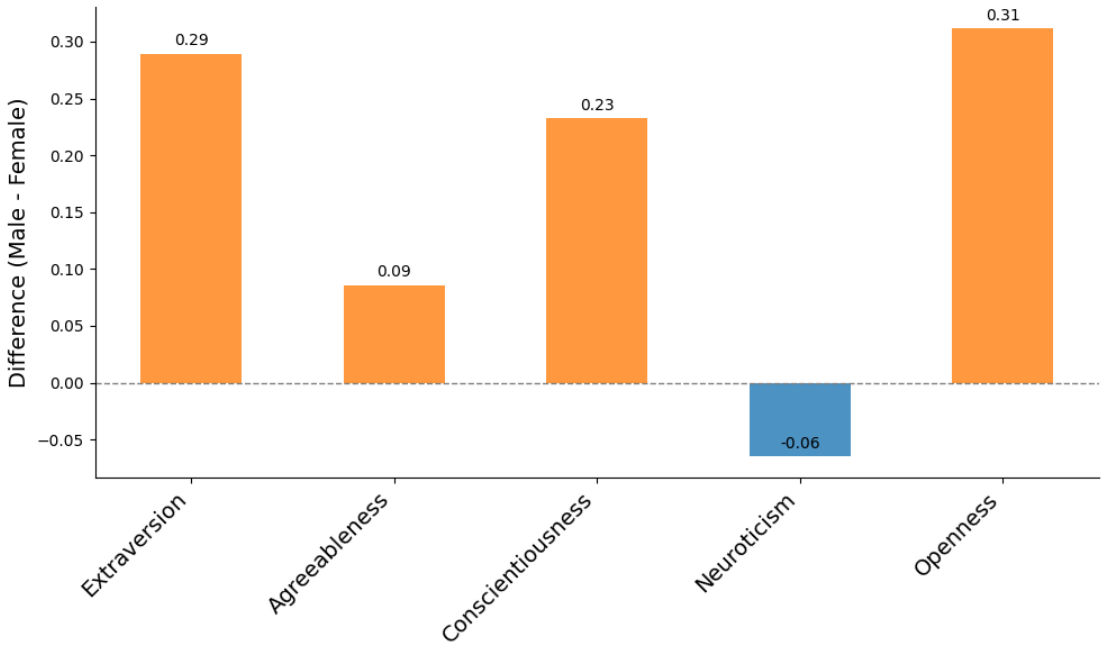



Figure 3: Gender differences in Big Five personalities

Note: Figure shows the differences in standard deviations between average male and female Big Five personality indices.

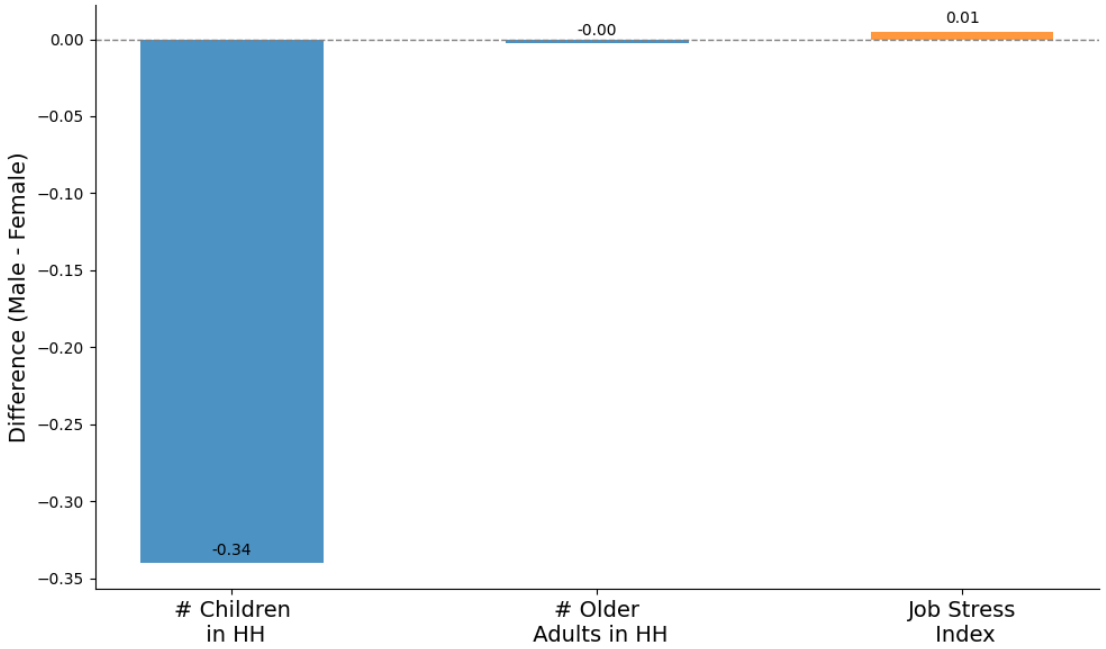



Figure 4: Gender differences in home care and job stress

Note: Figure shows the differences in standard deviations between average male and female on: the number of children in the household, the number of older adults over 65 years ago in the household, and the job stress index.




Figure 5: SHAP summary plot predicting log wages

Note: Figure shows the impact of key features on predicted log wages, which are derived from a random forest machine learning model. Features are ranked by their average absolute SHAP (SHapley Additive exPlanations) values, with more important predictors appearing higher on the vertical axis. Each dot represents a single observation, and its position on the horizontal axis reflects the SHAP value (i.e., the contribution of the feature to the prediction for that observation). Dot colors indicate feature values, with higher feature values in pink and lower feature values in blue.

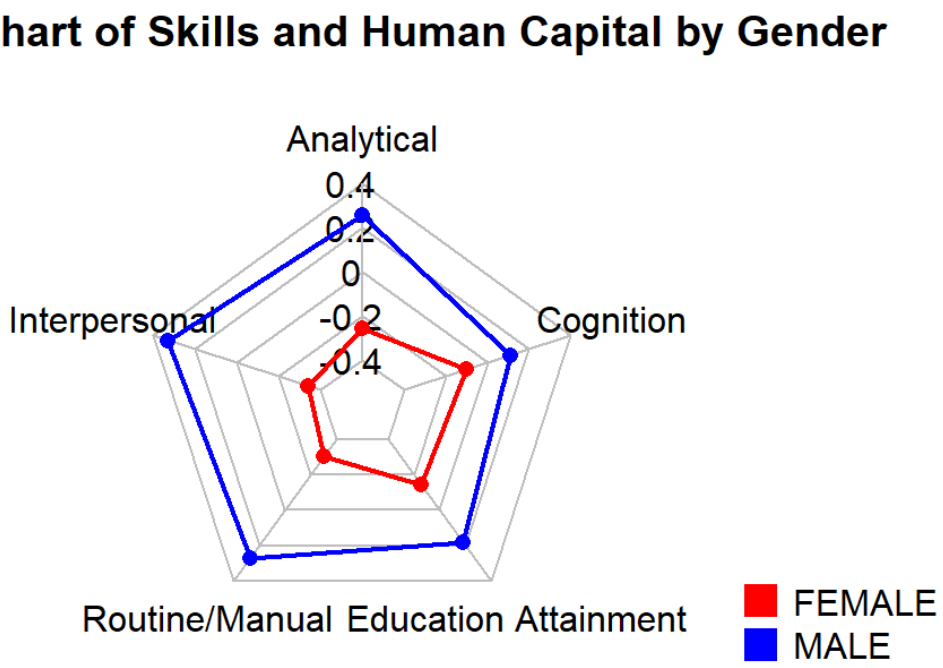



Figure 6: Gender Differences in Human Capital Measures

Note: Figure displays (in standard deviations) the average male and female outcomes on educational attainment, general cognition, and the three task indices.

Table 1: Summary of labor market outcomes and education by gender

| | Mean (Male) | Mean (Female) |
|---|---|---|
| Years of Education | 9.9 | 8.8 |
| Hourly Wage (USD) | 1.00 | 0.46 |
| Monthly Earnings (USD) | 141.5 | 50.2 |
| Weekly Hours Worked | 47.4 | 35.1 |
| Any self-employment(%) | 46.0 | 41.0 |
| Any wage employment(%) | 53.1 | 34.1 |
| Not employed past 30 days(%) | 4.7 | 15.1 |

Note: N=5,877 (excluding 1 observation with missing gender). Outcomes use 1 USD = 140 KSH.

Table 2: Top occupations by gender

| Occupation | Share Among Female (%) | Share Among Male (%) |
|---|---|---|
| Agriculture and Fishing - Farmer | 33.2 | 17.1 |
| Retail and commercial - Hawking/ selling | 17.9 | 5.0 |
| Professionals - Teacher | 4.8 | 5.1 |
| Unskilled trades - Domestic work | 8.1 | 0.8 |
| Agriculture and Fishing - Fishing | 0.6 | 7.7 |
| Skilled & semi-skilled trades - Barber | 5.1 | 0.9 |
| Unskilled trades - Cleaner (other) | 4.6 | 0.9 |
| Unskilled trades - Bicycle/motorbike taxi work | 0.0 | 6.0 |
| Retail and commercial - Own shop | 3.0 | 2.3 |
| Skilled & semi-skilled trades - Tailor | 3.9 | 0.9 |
| Skilled & semi-skilled trades - Mason | 0.1 | 4.9 |
| Unskilled trades - Watchman/ Guard | 0.7 | 3.6 |
| Unskilled trades - Unskilled construction laborer | 0.3 | 4.0 |
| Unskilled trades - Cook/ Chef/ Caterer | 2.2 | 1.6 |
| Other - Other (specify) | 1.0 | 2.9 |
| Skilled & semi-skilled trades - Other skilled construction | 0.2 | 3.8 |
| Unskilled trades - Work in other person's business | 1.2 | 2.4 |
| Professionals - Salaried professional | 1.1 | 2.6 |
| Agriculture and Fishing - Agricultural laborer | 1.9 | 1.6 |
| Professionals - Police/military officer | 0.5 | 2.7 |

Note: N= 5,878. Table shows the top 20 most common occupations among both men and women in the KLPS-5 sample, in descending order of number of workers. The right two columns show, for each occupation, the share of women and men employed in that occupation respectively.

Table 3: Gender wage gap: main regressions

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| Female | -0.78*** | -0.62*** | -0.58*** | -0.50*** | -0.48*** | -0.41*** | -0.41*** |
| | (0.05) | (0.05) | (0.06) | (0.06) | (0.05) | (0.06) | (0.06) |
| Educational attainment | | 0.08*** | | 0.04*** | 0.02 | 0.02 | 0.01 |
| | | (0.01) | | (0.01) | (0.01) | (0.01) | (0.01) |
| Experience | | -0.13*** | | -0.03 | -0.03 | -0.01 | -0.03 |
| | | (0.05) | | (0.04) | (0.04) | (0.04) | (0.04) |
| Experience$^2$/100 | | 0.00** | | 0.00 | 0.00 | 0.00 | 0.00 |
| | | (0.00) | | (0.00) | (0.00) | (0.00) | (0.00) |
| Urban | | 0.62*** | | 0.43*** | 0.42*** | 0.37*** | 0.37*** |
| | | (0.06) | | (0.06) | (0.06) | (0.06) | (0.06) |
| Nairobi | | 0.16** | | 0.13* | 0.14** | 0.13** | 0.14** |
| | | (0.07) | | (0.07) | (0.06) | (0.06) | (0.06) |
| Agriculture | | 0.08 | | -0.05 | -0.06 | -0.04 | -0.03 |
| | | (0.06) | | (0.06) | (0.06) | (0.06) | (0.05) |
| Self-employment/business | | 0.03 | | 0.27*** | 0.28*** | 0.24*** | 0.19*** |
| | | (0.05) | | (0.05) | (0.05) | (0.05) | (0.05) |
| General cognition score | | | | | 0.19*** | 0.17*** | 0.01 |
| | | | | | (0.03) | (0.03) | (0.28) |
| # Older adults in HH | | | | | 0.03 | 0.03 | 0.06 |
| | | | | | (0.06) | (0.06) | (0.06) |
| # Children in HH | | | | | -0.05*** | -0.05*** | -0.05*** |
| | | | | | (0.01) | (0.01) | (0.01) |
| Analytical Index | | | | | | 0.10*** | 0.08*** |
| | | | | | | (0.03) | (0.03) |
| Interpersonal index | | | | | | 0.12*** | 0.10*** |
| | | | | | | (0.03) | (0.03) |
| Routine/Manual index | | | | | | -0.01 | 0.02 |
| | | | | | | (0.02) | (0.02) |
| Job stress index | | | | | | | -0.14*** |
| | | | | | | | (0.02) |
| N | 5063 | 5063 | 5061 | 5061 | 5061 | 5061 | 5061 |
| r2 | 0.08 | 0.22 | 0.31 | 0.34 | 0.35 | 0.36 | 0.38 |
| Occupation_FE | | | Yes | Yes | Yes | Yes | Yes |
| Personality | | | | | | | Yes |
| Preferences | | | | | | | Yes |

Note: All columns include the female dummy as an explanatory variable. Column 2 includes standard demographics and education controls. Column 3 includes 49 occupation dummies as controls. Column 4 includes the controls in columns 2 and 3. Column 5 additionally include the general cognition score, number of children, and number of older adults over 65 as controls. Column 6 additionally controls for the three task indices. Column 7 additionally controls for the job stress index, personality measures, and economic preference measures, as well as the five component-specific cognition scores and the lagged Raven's score from KLPS-3. All standard errors are robust.

Table 4: Gender gap in hours worked

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| Female | -12.10*** | -10.59*** | -10.07*** | -8.73*** | -8.58*** | -6.98*** | -6.84*** |
| | (1.00) | (0.97) | (1.10) | (1.07) | (1.08) | (1.12) | (1.13) |
| Educational attainment | | 0.57* | | 0.40 | 0.32 | 0.28 | 0.24 |
| | | (0.31) | | (0.29) | (0.29) | (0.29) | (0.30) |
| Experience | | 1.25 | | 1.17 | 1.22 | 1.48* | 1.53* |
| | | (0.99) | | (0.89) | (0.89) | (0.89) | (0.87) |
| Experience$^2$/100 | | -0.03 | | -0.03 | -0.03 | -0.03 | -0.03 |
| | | (0.02) | | (0.02) | (0.02) | (0.02) | (0.02) |
| Urban | | 6.67*** | | 2.47* | 2.30* | 1.28 | 1.01 |
| | | (1.39) | | (1.34) | (1.36) | (1.37) | (1.37) |
| Nairobi | | -1.80 | | -1.76 | -1.69 | -1.87 | -1.77 |
| | | (1.70) | | (1.55) | (1.55) | (1.52) | (1.51) |
| Agriculture | | 0.12 | | -3.19** | -3.32*** | -2.80** | -2.98** |
| | | (1.28) | | (1.25) | (1.26) | (1.27) | (1.26) |
| Self-employment/business | | 12.86*** | | 11.65*** | 11.68*** | 10.86*** | 10.80*** |
| | | (0.95) | | (1.07) | (1.07) | (1.09) | (1.09) |
| General cognition score | | | | | 0.61 | 0.18 | 16.67*** |
| | | | | | (0.55) | (0.56) | (6.47) |
| # Older adults in HH | | | | | -0.79 | -0.77 | -0.73 |
| | | | | | (1.24) | (1.23) | (1.25) |
| # Children in HH | | | | | -0.33 | -0.37 | -0.37 |
| | | | | | (0.26) | (0.26) | (0.26) |
| Analytical Index | | | | | | 1.36** | 1.35** |
| | | | | | | (0.61) | (0.61) |
| Interpersonal index | | | | | | 3.30*** | 3.33*** |
| | | | | | | (0.66) | (0.65) |
| Routine/Manual index | | | | | | 0.37 | 0.36 |
| | | | | | | (0.59) | (0.58) |
| Job stress index | | | | | | | 0.10 |
| | | | | | | | (0.54) |
| N | 5773 | 5773 | 5772 | 5772 | 5772 | 5772 | 5772 |
| r2 | 0.04 | 0.10 | 0.21 | 0.24 | 0.24 | 0.25 | 0.26 |
| Occupation_FE | | | Yes | Yes | Yes | Yes | Yes |
| Personality | | | | | | | Yes |
| Preferences | | | | | | | Yes |

Note: Please see Table 3 for the notes, all of which apply here. The dependent variable is weekly hours worked. Individuals who do not report working any hours are included in estimation and coded as working zero hours. All standard errors are robust.

Table 5: Double/Debiased Machine Learning results

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dep. Var.: Log Wage | | | | |
| Female | -0.57*** | -0.41*** | -0.55*** | -0.41*** |
| | (0.04) | (0.04) | (0.04) | (0.04) |
| Observations | 5063 | 5063 | 4758 | 4758 |
| Dep. Var.: Weekly Hours Worked | | | | |
| Female | -7.07*** | -4.43*** | -6.78*** | -5.07*** |
| | (0.89) | (0.88) | (0.95) | (0.92) |
| Observations | 5773 | 5773 | 5120 | 5120 |
| Controls | Basic | Full | Basic | Full |
| Class Balance | No | No | Yes | Yes |

Note: This table reports results from Double/Debiased Machine Learning (DML). The top row reports the coefficient and standard error on the female indicator in a model where the dependent variable is log wage; in the bottom row the dependent variable is weekly hours worked. Columns 1 and 3 include occupation, education, and demographics as potential confounders; Columns 2 and 4 include the full set of controls. Columns 3 and 4 randomly under-samples women so as to achieve perfect balance between men and women in the analysis sample. All models use a random forest regressor to predict the outcome and a random forest classifier to predict being female. All models use the default 3-fold cross-validation of DML.

Table 6: Oaxaca-Blinder decomposition of wages

| | Coef. | 95% CI Lower | 95% CI Upper |
|---|---|---|---|
| With basic controls: | | | |
| Difference(Male-Female) | 0.780 | 0.692 | 0.867 |
| Explained | 0.307 | 0.202 | 0.411 |
| Unexplained | 0.473 | 0.345 | 0.601 |
| | | | |
| With full set of controls: | | | |
| Difference(Male-Female) | 0.780 | 0.691 | 0.868 |
| Explained | 0.396 | 0.321 | 0.470 |
| Unexplained | 0.384 | 0.283 | 0.486 |

Note: This table reports the Oaxaxa-Blinder decomposition results of the gender gap in log wages, using male as the reference group. "Basic" controls include eduaction, experience, demographics, and occupation variables. "Full" controls additionally include cognition, personality, economic preferences, job stress, and home care responsibilities.

Table 7: Oaxaca-Blinder decomposition of hours worked

| | Coef. | 95% CI Lower | 95% CI Upper |
|---|---|---|---|
| With basic controls: | | | |
| Difference(Male-Female) | 12.099 | 10.235 | 13.964 |
| Explained | 1.345 | -0.857 | 3.547 |
| Unexplained | 10.754 | 8.073 | 13.436 |
| | | | |
| With full set of controls: | | | |
| Difference(Male-Female) | 12.099 | 10.201 | 13.998 |
| Explained | 3.641 | 2.064 | 5.219 |
| Unexplained | 8.458 | 6.139 | 10.777 |

Note: This table reports the Oaxaxa-Blinder decomposition results of the gender gap in hours worked using male as the reference group. "Basic" controls include eduaction, experience, demographics, and occupation variables. "Full" controls additionally include cognition, personality, economic preferences, job stress, and home care responsibilities.

Table 8: Gender wage gap: interaction terms

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Female (intercept) | -0.32 | 0.44 | -0.07 | 0.67 |
| | (0.21) | (1.01) | (0.19) | (1.10) |
| **Interaction terms with Female** | | | | |
| Urban | -0.07 | -0.12 | -0.38*** | -0.28** |
| | (0.11) | (0.11) | (0.12) | (0.11) |
| Nairobi | 0.03 | 0.14 | 0.20 | 0.02 |
| | (0.13) | (0.13) | (0.13) | (0.12) |
| Agriculture | 0.08 | 0.00 | -0.24** | -0.16 |
| | (0.11) | (0.10) | (0.11) | (0.11) |
| Self-employment/business | -0.15* | -0.13 | -0.19** | -0.20** |
| | (0.09) | (0.09) | (0.09) | (0.09) |
| Educational attainment | -0.02 | -0.03 | -0.01 | -0.02 |
| | (0.02) | (0.03) | (0.01) | (0.03) |
| General cognition score | | 1.00* | | 0.80 |
| | | (0.56) | | (0.62) |
| Executive functioning | | -0.37 | | -0.32 |
| | | (0.23) | | (0.26) |
| Orientation | | -0.01 | | 0.02 |
| | | (0.11) | | (0.13) |
| Memory | | -0.47 | | -0.33 |
| | | (0.30) | | (0.33) |
| Language | | -0.26* | | -0.12 |
| | | (0.14) | | (0.15) |
| Visuospatial | | -0.22* | | -0.22 |
| | | (0.12) | | (0.12) |
| Lagged Raven's | | 0.03 | | 0.02 |
| | | (0.05) | | (0.05) |
| Experience | | -0.04 | | -0.06 |
| | | (0.08) | | (0.09) |
| Experience$^2$/100 | | 0.00 | | 0.00 |
| | | (0.00) | | (0.00) |
| Analytical Index | | -0.14*** | | -0.16*** |
| | | (0.05) | | (0.05) |
| Interpersonal Index | | 0.18*** | | 0.20*** |
| | | (0.05) | | (0.06) |
| Routine/Manual Index | | 0.04 | | -0.02 |
| | | (0.05) | | (0.05) |
| N | 5061 | 5061 | 5061 | 5061 |
| Model | Regression | Regression | DML | DML |
| Controls | Basic | Full | Basic | Full |

Note: Columns 1 and 3 include the basic set of controls a la Column 4 in Table 3, along with the listed interaction terms. Columns 3 and 4 include the basic set of controls a la Column 7 in Table 3, along with the listed interaction terms. Columns 3 and 4 run DML regressions following the same setup as in Table 5. All standard errors are robust.

Table 9: Gender gap in employment wage

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| Female | -0.41*** | -0.28*** | -0.41*** | -0.34*** | -0.33*** | -0.30*** | -0.30*** |
| | (0.06) | (0.06) | (0.07) | (0.07) | (0.06) | (0.07) | (0.06) |
| Educational attainment | | 0.10*** | | 0.04** | 0.02 | 0.01 | 0.01 |
| | | (0.02) | | (0.02) | (0.02) | (0.02) | (0.02) |
| Experience | | -0.16*** | | -0.09 | -0.09 | -0.07 | -0.09 |
| | | (0.06) | | (0.05) | (0.05) | (0.05) | (0.06) |
| Experience$^2$/100 | | 0.00*** | | 0.00* | 0.00* | 0.00 | 0.00* |
| | | (0.00) | | (0.00) | (0.00) | (0.00) | (0.00) |
| Urban | | 0.50*** | | 0.39*** | 0.37*** | 0.34*** | 0.32*** |
| | | (0.08) | | (0.08) | (0.07) | (0.07) | (0.07) |
| Nairobi | | 0.09 | | 0.12 | 0.13* | 0.12 | 0.12 |
| | | (0.08) | | (0.08) | (0.08) | (0.08) | (0.08) |
| Agriculture | | -0.09 | | -0.05 | -0.06 | -0.04 | -0.03 |
| | | (0.07) | | (0.07) | (0.07) | (0.07) | (0.07) |
| Self-employment/business | | 0.14* | | 0.19*** | 0.19*** | 0.16** | 0.11 |
| | | (0.07) | | (0.07) | (0.07) | (0.07) | (0.07) |
| General cognition score | | | | | 0.15*** | 0.13*** | 0.29 |
| | | | | | (0.03) | (0.03) | (0.34) |
| # Older adults in HH | | | | | -0.23** | -0.21** | -0.22** |
| | | | | | (0.09) | (0.09) | (0.09) |
| # Children in HH | | | | | -0.04*** | -0.04*** | -0.04*** |
| | | | | | (0.02) | (0.02) | (0.01) |
| Analytical Index | | | | | | 0.15*** | 0.12*** |
| | | | | | | (0.03) | (0.03) |
| Interpersonal index | | | | | | 0.04 | 0.03 |
| | | | | | | (0.04) | (0.04) |
| Routine/Manual index | | | | | | -0.08*** | -0.05** |
| | | | | | | (0.02) | (0.02) |
| Job stress index | | | | | | | -0.16*** |
| | | | | | | | (0.03) |
| N | 2260 | 2260 | 2260 | 2260 | 2260 | 2260 | 2260 |
| r2 | 0.04 | 0.20 | 0.29 | 0.32 | 0.34 | 0.35 | 0.39 |
| Occupation_FE | | | Yes | Yes | Yes | Yes | Yes |
| Personality | | | | | | | Yes |
| Preferences | | | | | | | Yes |

Note: Please see Table 3 for the notes, all of which apply here. The dependent variable is wage from employment. Individuals who do not report engaging in any wage employment are excluded from estimation. All standard errors are robust.

Table 10: Gender gap in earnings

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| Female | -1.11*** | -0.90*** | -0.87*** | -0.76*** | -0.72*** | -0.61*** | -0.60*** |
| | (0.05) | (0.05) | (0.06) | (0.05) | (0.05) | (0.05) | (0.05) |
| Educational attainment | | 0.11*** | | 0.06*** | 0.03** | 0.03** | 0.03** |
| | | (0.02) | | (0.01) | (0.01) | (0.01) | (0.01) |
| Experience | | -0.11** | | -0.00 | 0.01 | 0.03 | 0.02 |
| | | (0.05) | | (0.04) | (0.04) | (0.04) | (0.04) |
| Experience$^2$/100 | | 0.00* | | -0.00 | -0.00 | -0.00 | -0.00 |
| | | (0.00) | | (0.00) | (0.00) | (0.00) | (0.00) |
| Urban | | 0.72*** | | 0.43*** | 0.41*** | 0.33*** | 0.32*** |
| | | (0.06) | | (0.06) | (0.06) | (0.06) | (0.05) |
| Nairobi | | 0.11 | | 0.07 | 0.08 | 0.07 | 0.08 |
| | | (0.07) | | (0.06) | (0.06) | (0.06) | (0.06) |
| Agriculture | | 0.29*** | | 0.00 | -0.01 | 0.03 | 0.04 |
| | | (0.06) | | (0.05) | (0.05) | (0.05) | (0.05) |
| Self-employment/business | | 0.29*** | | 0.51*** | 0.52*** | 0.46*** | 0.42*** |
| | | (0.05) | | (0.05) | (0.05) | (0.05) | (0.05) |
| General cognition score | | | | | 0.22*** | 0.18*** | 0.15 |
| | | | | | (0.03) | (0.03) | (0.28) |
| # Older adults in HH | | | | | 0.02 | 0.02 | 0.04 |
| | | | | | (0.06) | (0.06) | (0.06) |
| # Children in HH | | | | | -0.06*** | -0.06*** | -0.06*** |
| | | | | | (0.01) | (0.01) | (0.01) |
| Analytical Index | | | | | | 0.15*** | 0.13*** |
| | | | | | | (0.03) | (0.03) |
| Interpersonal index | | | | | | 0.19*** | 0.19*** |
| | | | | | | (0.03) | (0.03) |
| Routine/Manual index | | | | | | 0.04* | 0.06*** |
| | | | | | | (0.02) | (0.02) |
| Job stress index | | | | | | | -0.12*** |
| | | | | | | | (0.02) |
| N | 5434 | 5433 | 5336 | 5336 | 5336 | 5335 | 5335 |
| r2 | 0.12 | 0.43 | 0.31 | 0.38 | 0.40 | 0.51 | 0.52 |
| Occupation_FE | | | Yes | Yes | Yes | Yes | Yes |
| Personality | | | | | | | Yes |
| Preferences | | | | | | | Yes |

Note: Please see Table 3 for the notes, all of which apply here. The dependent variable is log monthly earnings. All standard errors are robust.

Table 11: Experiences of discrimination by gender

| **Panel A: Share experiencing discrimination by type** | Mean (Male) | Mean (Female) | Female Coef (No controls) | Female Coef (All controls) |
|---|---|---|---|---|
| Treated with less courtesy than other people | 0.52 | 0.49 | -0.02 | -0.03 |
| | | | (0.02) | (0.02) |
| Treated with less respect than other people | 0.52 | 0.51 | -0.00 | -0.02 |
| | | | (0.02) | (0.02) |
| Receive poorer service than other people at restaurants/stores | 0.31 | 0.27 | -0.04** | -0.06*** |
| | | | (0.02) | (0.02) |
| People act as if they think you are not smart | 0.45 | 0.45 | 0.00 | -0.04* |
| | | | (0.02) | (0.02) |
| People act as if they are afraid of you | 0.41 | 0.37 | -0.04** | -0.03 |
| | | | (0.02) | (0.02) |
| You are threatened or harassed | 0.28 | 0.26 | -0.02 | -0.07*** |
| | | | (0.01) | (0.02) |

| **Panel B: Main perceived reason for discrimination** | Mean (Male) | Mean (Female) |
|---|---|---|
| Your education or income level | 0.48 | 0.49 |
| Your Social Group or Family Background | 0.18 | 0.19 |
| Aspects of your physical appearance | 0.12 | 0.11 |
| Place of origin | 0.06 | 0.05 |
| **Your gender** | $< 0.01$ | 0.02 |
| Your ancestry | 0.02 | 0.01 |
| Your language ability | 0.01 | 0.01 |
| Your religion | $< 0.01$ | $< 0.01$ |
| Your age | $< 0.01$ | $< 0.01$ |
| Your race | $< 0.01$ | $< 0.01$ |
| Your sexual orientation | $< 0.01$ | $< 0.01$ |
| Your political orientation | $< 0.01$ | $< 0.01$ |

Note: The first two columns of Panel A show the the share of individuals, among men and women respectively, who report experiencing discrimination of a specific type. Each variable is coded as 1 if the respondent reports that the experience have happened to them more than a few times a year or more, and 0 if they report less than once a year or never. The third and fourth columns report the regression coefficient of an indicator for experiencing a specific type of discrimination on gender in the full sample. The third column includes no covariates; the fourth column includes the full set of covariates used in regression analyses in Table 3. Panel B shows the share of individuals, among men and women respectively, who say that the left-most column reason is the main reason why they think discrimination happened to them. N = 5,786 (excluding one individual missing gender and 91 individuals who did not complete the discrimination module).

Table 12: Attitudes toward gender norms

| **Statement** | Agree (Male) | Agree (Female) | Female Coef (No controls) | Female Coef (All controls) |
|---|---|---|---|---|
| It is okay for a woman to be a mechanic | 0.95 | 0.92 | -0.03*** | -0.01 |
| | | | (0.01) | (0.01) |
| Important decisions in the family should be made by the men of the family (reverse-coded) | 0.37 | 0.42 | 0.05*** | 0.14*** |
| | | | (0.02) | (0.02) |
| If wife is working outside the home, husband should help her with household chores | 0.84 | 0.80 | -0.05*** | -0.04** |
| | | | (0.01) | (0.02) |
| Girls and boys have equal opportunities to get a secondary education | 0.99 | 0.99 | -0.00 | 0.00 |
| | | | (0.00) | (0.00) |
| Women and men have equal opportunities to get a job that pays a wage or salary | 0.98 | 0.99 | 0.00 | -0.00 |
| | | | (0.01) | (0.01) |
| Women can be good politicians and should be encouraged to stand in elections | 0.92 | 0.92 | 0.00 | 0.04*** |
| | | | (0.01) | (0.01) |
| In our country, women should have equal rights and receive the same treatment as men | 0.91 | 0.92 | 0.02 | 0.03** |
| | | | (0.01) | (0.01) |
| No one has the right to use physical violence against anyone else | 0.90 | 0.90 | -0.00 | 0.03** |
| | | | (0.01) | (0.01) |

Note: The first two columns of Panel A show the the share of individuals, among men and women respectively, who agree with the statement on the left-most column. The third and fourth columns report the regression coefficient of an indicator for agreeing with the statement on gender in the full sample. The third column includes no covariates; the fourth column includes the full set of covariates used in regression analyses in Table 3. The second statement, "Important decisions in the family should be made by the men of the family", is reverse-coded such that the numbers show the share of men and women who disagree with that statement. Each of the last three statements (A) was presented together with an opposing statement (B), and the respondent was asked to select "agree very strongly with A", "agree with A", "agree with B", "agree very strongly with B", or "agree with neither". Specifically, the opposing statements are "women should stay at home to take care of their children", "women have always been subject to traditional laws and customs, and should remain so", and "a married man has a right to beat his wife if she misbehaves". Data from KLPS-4 survey administered between 2018 and 2021. N = 5,103 (excluding individuals in KLPS-5 who did not complete the KLPS-4 survey).

## References


Daron Acemoglu and David Autor. Skills, Tasks and Technologies: Implications for Employment and Earnings*. In David Card and Orley Ashenfelter, editors, *Handbook of Labor Economics*, volume 4, pages 1043–1171. Elsevier, January 2011. doi: 10.1016/S0169-7218(11)02410-5. URL https://www.sciencedirect.com/science/article/pii/S0169721811024105.

Mehjabeen Alarakhia, Marjan Petreski, Zahra Sheikh Ahmed, and Tanima Tanima. Why Women Earn Less: Gender Pay Gap and Labour-Market Inequalities in East and Southern Africa. Technical report, UN Women, 2023. URL https://africa.unwomen.org/sites/default/files/2023-10/gpg_regional_un_women_final.pdf.

Giammarco Alderotti, Chiara Rapallini, and Silvio Traverso. The Big Five personality traits and earnings: A meta-analysis. *Journal of Economic Psychology*, 94:102570, January 2023. ISSN 0167-4870. doi: 10.1016/j.joep.2022.102570. URL https://www.sciencedirect.com/science/article/pii/S0167487022000812.

Simon Appleton, John Hoddinott, and Pramila Krishnan. The Gender Wage Gap in Three African Countries. *Economic Development and Cultural Change*, 47(2):289–312, January 1999. ISSN 0013-0079. doi: 10.1086/452402. URL https://www.journals.uchicago.edu/doi/abs/10.1086/452402. Publisher: The University of Chicago Press.

David H. Autor and Michael J. Handel. Putting Tasks to the Test: Human Capital, Job Tasks, and Wages. *Journal of Labor Economics*, 31(2):S59–S96, 2013. ISSN 0734-306X. doi: 10.1086/669332. URL https://www.jstor.org/stable/10.1086/669332. Publisher: [The University of Chicago Press, Society of Labor Economists, NORC at the University of Chicago].

Philipp Bach, Victor Chernozhukov, and Martin Spindler. Heterogeneity in the US gender wage gap. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 187(1):209–230, January 2024. ISSN 0964-1998. doi: 10.1093/jrsssa/qnad091. URL https://doi.org/10.1093/jrsssa/qnad091.

The World Bank. Home | Gender Data Portal. URL https://genderdata.worldbank.org/en/home.

Oscar Becerra and José-Alberto Guerra. Personal safety first: Do workers value safer jobs? *Journal of Economic Behavior & Organization*, 212:996–1016, August 2023. ISSN 0167-2681. doi: 10.1016/j.jebo.2023.06.017. URL https://www.sciencedirect.com/science/article/pii/S0167268123002238.

Alexandre Belloni, Victor Chernozhukov, and Christian Hansen. Inference on Treatment Effects after Selection among High-Dimensional Controls†. *The Review of Economic Studies*, 81(2):608–650, April 2014. ISSN 0034-6527. doi: 10.1093/restud/rdt044. URL https://doi.org/10.1093/restud/rdt044.

Francine D. Blau and Lawrence M. Kahn. The Gender Wage Gap: Extent, Trends, and Explanations. *Journal of Economic Literature*, 55(3):789–865, September 2017. ISSN 0022-0515. doi: 10.1257/jel.20160995. URL https://www.aeaweb.org/articles?id=10.1257/jel.20160995.

Alan S. Blinder. Wage Discrimination: Reduced Form and Structural Estimates. *The Journal of*

*Human Resources*, 8(4):436–455, 1973. ISSN 0022-166X. doi: 10.2307/144855. URL https://www.jstor.org/stable/144855. Publisher: [University of Wisconsin Press, Board of Regents of the University of Wisconsin System].

M. Boissiere, J. B. Knight, and R. H. Sabot. Earnings, Schooling, Ability, and Cognitive Skills. *The American Economic Review*, 75(5):1016–1030, 1985. ISSN 0002-8282. URL https://www.jstor.org/stable/1818642. Publisher: American Economic Association.

Marina Bonaccolto-Töpfer and Stephanie Briel. The gender pay gap revisited: Does machine learning offer new insights? *Labour Economics*, 78:102223, October 2022. ISSN 0927-5371. doi: 10.1016/j.labeco.2022.102223. URL https://www.sciencedirect.com/science/article/pii/S0927537122001130.

Nina Buchmann, Carl Meyer, and Colin D. Sullivan. Paternalistic Discrimination. *mimeo*, 2024.

Leonardo Bursztyn, Alessandra L. González, and David Yanagizawa-Drott. Misperceived Social Norms: Women Working Outside the Home in Saudi Arabia. *American Economic Review*, 110(10):2997–3029, October 2020. ISSN 0002-8282. doi: 10.1257/aer.20180975. URL https://www.aeaweb.org/articles?id=10.1257/aer.20180975.

René Böheim and Philipp Stöllinger. Decomposition of the gender wage gap using the LASSO estimator. *Applied Economics Letters*, 28(10):817–828, June 2021. ISSN 1350-4851. doi: 10.1080/13504851.2020.1782332. URL https://doi.org/10.1080/13504851.2020.1782332. Publisher: Routledge _eprint: https://doi.org/10.1080/13504851.2020.1782332.

David Card, Ana Rute Cardoso, Joerg Heining, and Patrick Kline. Firms and Labor Market Inequality: Evidence and Some Theory. *Journal of Labor Economics*, 36(S1):S13–S70, January 2018. ISSN 0734-306X. doi: 10.1086/694153. URL https://www.journals.uchicago.edu/doi/10.1086/694153. Publisher: The University of Chicago Press.

John Cawley, James Heckman, and Edward Vytlacil. Three observations on wages and measured cognitive ability. *Labour Economics*, 8(4):419–442, September 2001. ISSN 0927-5371. doi: 10.1016/S0927-5371(01)00039-2. URL https://www.sciencedirect.com/science/article/pii/S0927537101000392.

Ali Cheema, Asim I. Khwaja, Farooq Naseer, and Jacob N. Shapiro. Glass walls: Experimental evidence on access constraints faced by women, 2023. URL https://mgshss.lums.edu.pk/sites/default/files/inline-files/glass_walls_draft_020222.pdf. Publication Title: Technical Report.

Yutong Chen, Kerem Coşar, Devaki Ghose, Shirish Mahendru, and Sheetal Sekhri. Gender-Specific Transportation Costs and Female Time Use: Evidence from India's Pink Slip Program, May 2024. URL https://www.nber.org/papers/w32508.

Victor Chernozhukov, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins. Double/Debiased Machine Learning for Treatment and Causal Parameters, July 2016. URL http://arxiv.org/abs/1608.00060. arXiv:1608.00060 [stat].

Victor Chernozhukov, Denis Nekipelov, Vira Semenova, and Vasilis Syrgkanis. Plug-in regularized

estimation of high dimensional parameters in nonlinear semiparametric models. Technical report, The IFS, July 2018. URL https://www.ifs.org.uk/uploads/cemmap/wps/CWP411818.pdf.

Sara De La Rica, Lucas Gortazar, and Piotr Lewandowski. Job Tasks and Wages in Developed Countries: Evidence from PIAAC. *Labour Economics*, 65:101845, August 2020. ISSN 0927-5371. doi: 10.1016/j.labeco.2020.101845. URL https://www.sciencedirect.com/science/article/pii/S092753712030049X.

David J. Deming. The Growing Importance of Social Skills in the Labor Market*. *The Quarterly Journal of Economics*, 132(4):1593–1640, November 2017. ISSN 0033-5533. doi: 10.1093/qje/qjx022. URL https://doi.org/10.1093/qje/qjx022.

Emanuele Dicarlo, Salvatore Lo Bello, Sebastian Monroy-Taborda, Ana Maria Oviedo, Maria Laura Sanchez Puerta, and Indhira Vanessa Santos. The Skill Content of Occupations Across Low and Middle Income Countries: Evidence from Harmonized Data. *SSRN Electronic Journal*, 2016. ISSN 1556-5068. doi: 10.2139/ssrn.2846315. URL https://www.ssrn.com/abstract=2846315.

Per-Anders Edin, Peter Fredriksson, Martin Nybom, and Björn Öckert. The Rising Return to Noncognitive Skill. *American Economic Journal: Applied Economics*, 14(2):78–100, April 2022. ISSN 1945-7782. doi: 10.1257/app.20190199. URL https://www.aeaweb.org/articles?id=10.1257/app.20190199.

Eurostat. Gender pay gap statistics - Statistics Explained, 2022. URL https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Gender_pay_gap_statistics.

Christopher J. Flinn, Petra E. Todd, and Weilong Zhang. Labor Market Returns to Personality: A Job Search Approach to Understanding Gender Gaps. *Journal of Political Economy*, pages 000–000, November 2024. ISSN 0022-3808. doi: 10.1086/734092. URL https://www.journals.uchicago.edu/doi/abs/10.1086/734092. Publisher: The University of Chicago Press.

Claudia Goldin. A Grand Gender Convergence: Its Last Chapter. *American Economic Review*, 104(4):1091–1119, April 2014. ISSN 0002-8282. doi: 10.1257/aer.104.4.1091. URL https://www.aeaweb.org/articles?id=10.1257/aer.104.4.1091.

Alden L. Gross, Chihua Li, Emily M. Briceño, Miguel Arce Rentería, Richard N. Jones, Kenneth M. Langa, Jennifer J. Manly, Emma Nichols, David Weir, Rebeca Wong, Lisa Berkman, Jinkook Lee, and Lindsay C. Kobayashi. Harmonisation of later-life cognitive function across national contexts: results from the Harmonized Cognitive Assessment Protocols. *The Lancet Healthy Longevity*, 4(10):e573–e583, October 2023. ISSN 2666-7568. doi: 10.1016/S2666-7568(23)00170-8. URL https://www.thelancet.com/journals/lanhl/article/PIIS2666-7568(23)00170-8/fulltext. Publisher: Elsevier.

Alden L. Gross, Madeline Duhon, Eric Ochieng, Jean N. Ikanga, William H. Dow, Jinkook Lee, Michael W. Walker, Michelle Layvant, Anthony K. Ngugi, Joshua R. Ehrlich, and Edward Miguel. Midlife cognitive testing in Africa: validity of the Harmonized Cognitive Assessment Protocol in the Kenya Life Panel Survey. *mimeo*, September 2024.

EconML User Guide. Orthogonal/Double Machine Learning,, 2023. URL https://econml.azurewebsites.net/spec/estimation/dml.html.

Eric A. Hanushek and Ludger Woessmann. The Role of Cognitive Skills in Economic Development. *Journal of Economic Literature*, 46(3):607–668, September 2008. ISSN 0022-0515. doi: 10.1257/jel.46.3.607. URL https://www.aeaweb.org/articles?id=10.1257/jel.46.3.607.

James J. Heckman, Jora Stixrud, and Sergio Urzua. The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics*, 24(3): 411–482, July 2006. ISSN 0734-306X. doi: 10.1086/504455. URL https://www.journals.uchicago.edu/doi/full/10.1086/504455. Publisher: The University of Chicago Press.

Santiago Hermo, Miika Päällysaho, David Seim, and Jesse M Shapiro. Labor Market Returns and the Evolution of Cognitive Skills: Theory and Evidence*. *The Quarterly Journal of Economics*, 137(4):2309–2361, November 2022. ISSN 0033-5533. doi: 10.1093/qje/qjac022. URL https://doi.org/10.1093/qje/qjac022.

Suhani Jalota and Lisa Ho. What Works For Her? How Work-from-Home Jobs Affect Female Labor Force Participation in Urban India, January 2024. URL https://papers.ssrn.com/abstract=4739387.

Seema Jayachandran. Social Norms as a Barrier to Women's Employment in Developing Countries. *IMF Economic Review*, 69(3):576–595, September 2021. ISSN 2041-417X. doi: 10.1057/s41308-021-00140-w. URL https://doi.org/10.1057/s41308-021-00140-w.

R. Karasek, C. Brisson, N. Kawakami, I. Houtman, P. Bongers, and B. Amick. The Job Content Questionnaire (JCQ): an instrument for internationally comparative assessments of psychosocial job characteristics. *Journal of Occupational Health Psychology*, 3(4):322–355, October 1998. ISSN 1076-8998. doi: 10.1037//1076-8998.3.4.322.

Stephan Klasen. What Explains Uneven Female Labor Force Participation Levels and Trends in Developing Countries? *The World Bank Research Observer*, 34(2):161–197, August 2019. ISSN 0257-3032. doi: 10.1093/wbro/lkz005. URL https://doi.org/10.1093/wbro/lkz005.

Thibaut Lamadon, Jeremy Lise, Costas Meghir, and Jean-Marc Robin. Labor Market Matching, Wages, and Amenities, July 2024. URL https://papers.ssrn.com/abstract=4895416.

Dajun Lin, Randall Lutter, and Christopher J. Ruhm. Cognitive performance and labour market outcomes. *Labour Economics*, 51:121–135, April 2018. ISSN 0927-5371. doi: 10.1016/j.labeco.2017.12.008. URL https://www.sciencedirect.com/science/article/pii/S0927537117303329.

Salvatore Lo Bello, Maria Laura Sanchez Puerta, and Hernan Winkler. From Ghana to America: The Skill Content of Jobs and Economic Development. *IZA Discussion Paper Series*, 2019.

Madeline McKelway. Experimental evidence on the effects of women's employment. *Working Paper*, 2019. URL https://scholar.google.com/scholar?cluster=15490416492627104447&hl=en&oi=scholarr.

Edward Miguel and Michael Kremer. Worms: identifying impacts on education and health in the presence of treatment externalities. *Econometrica*, 72(1):159–217, 2004. URL https://doi.org/10.1111/j.1468-0262.2004.00481.x.

Christoph Molnar. Interpreting Machine Learning Models With SHAP: A Guide With Python Examples And Theory On Shapley Values, 2023. URL https://leanpub.com/shap.

Christophe J. Nordman, Anne-Sophie Robilliard, and François Roubaud. Gender and ethnic earnings gaps in seven West African cities. *Labour Economics*, 18:S132–S145, December 2011. ISSN 0927-5371. doi: 10.1016/j.labeco.2011.09.003. URL https://www.sciencedirect.com/science/article/pii/S0927537111001035.

Christophe Jalil Nordman, Leopold R. Sarr, and Smriti Sharma. Cognitive, Non-Cognitive Skills and Gender Wage Gaps: Evidence from Linked Employer-Employee Data in Bangladesh. *SSRN Electronic Journal*, 2015. ISSN 1556-5068. doi: 10.2139/ssrn.2655081. URL https://www.ssrn.com/abstract=2655081.

Ronald Oaxaca. Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review*, 14(3):693–709, 1973. ISSN 0020-6598. doi: 10.2307/2525981. URL https://www.jstor.org/stable/2525981. Publisher: [Economics Department of the University of Pennsylvania, Wiley, Institute of Social and Economic Research, Osaka University].

Ronald L Oaxaca and Michael R Ransom. On discrimination and the decomposition of wage differentials. *Journal of Econometrics*, 61(1):5–21, March 1994. ISSN 0304-4076. doi: 10.1016/0304-4076(94)90074-4. URL https://www.sciencedirect.com/science/article/pii/0304407694900744.

U.S. Department of Labor. What You Need to Know About the Gender Wage Gap | U.S. Department of Labor Blog, 2024. URL https://blog.dol.gov/2024/03/12/what-you-need-to-know-about-the-gender-wage-gap.

Esteban Ortiz-Ospina, Sandra Tzvetkova, and Max Roser. Women's Employment. *Our World in Data*, March 2024. URL https://ourworldindata.org/female-labor-supply.

Zahra Siddique. Media-Reported Violence and Female Labor Supply. *Economic Development and Cultural Change*, 70(4):1337–1365, July 2022. ISSN 0013-0079. doi: 10.1086/714009. URL https://www.journals.uchicago.edu/doi/10.1086/714009. Publisher: The University of Chicago Press.

Jason Sockin. Show Me the Amenity: Are Higher-Paying Firms Better All Around?, November 2021. URL https://papers.ssrn.com/abstract=3957002.

Keyon Vafa, Susan Athey, and David Blei. Decomposing Changes in the Gender Wage Gap over Worker Careers, 2023.

Sher Verick. Female labor force participation in developing countries. *IZA World of Labor*, 2014. ISSN 20549571. doi: 10.15185/izawol.87. URL https://wol.iza.org/articles/female-labor-force-participation-in-developing-countries/v1.

Michael W. Walker, Alice H. Huang, Suleiman Asman, Sarah J. Baird, Lia Fernald, Joan Hamory, Fernando Hoces de la Guardia, Satoshi Koiso, Michael Kremer, Matthew N. Krupoff, Michelle Layvant, Eric Ochieng, Pooja Suri, and Edward Miguel. Intergenerational Child Mortality Impacts of Deworming: Experimental Evidence from Two Decades of the Kenya Life Panel Survey, April 2023. URL https://www.nber.org/papers/w31162.

Ben Weidmann and David J. Deming. Team Players: How Social Skills Improve Team Performance. *Econometrica*, 89(6):2637–2657, 2021. ISSN 1468-0262. doi: 10.3982/ECTA18461. URL https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA18461. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA18461.

# Appendix

## A Additional Exhibits

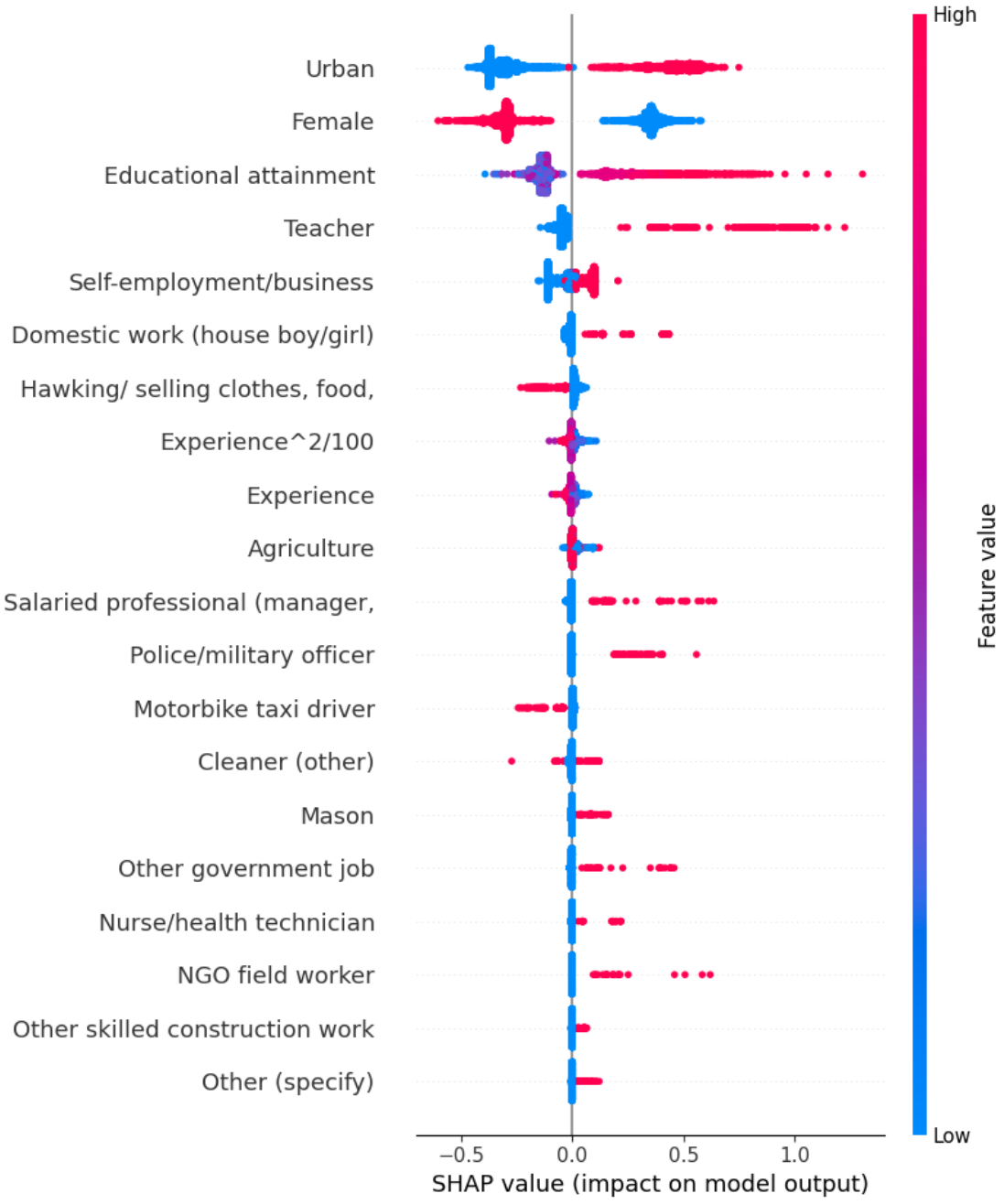



Figure A.1: SHAP summary plot predicting wages with basic set of predictors

Note: See Figure 5 notes, all of which apply here. The model includes 55 education, experience, demographics, and occupation variables as predictors of log wages.




Figure A.2: SHAP summary plot predicting hours worked

Note: Figure shows the impact of key features on predicted hours worked, which are derived from a random forest machine learning model. Features are ranked by their average absolute SHAP (SHapley Additive exPlanations) values, with more important predictors appearing higher on the vertical axis. Each dot represents a single observation, and its position on the horizontal axis reflects the SHAP value (i.e., the contribution of the feature to the prediction for that observation). Dot colors indicate feature values, with higher feature values in pink and lower feature values in blue.

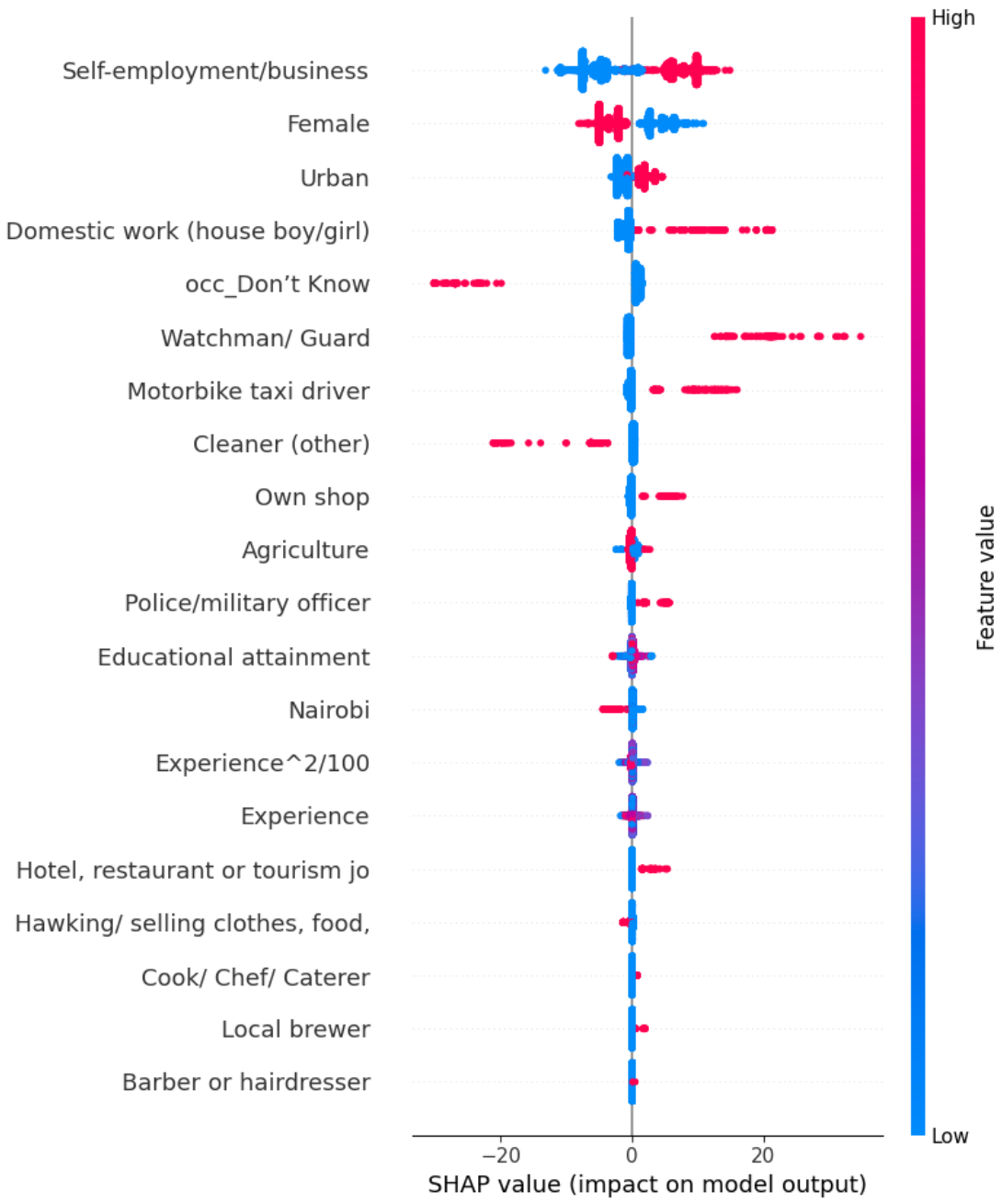
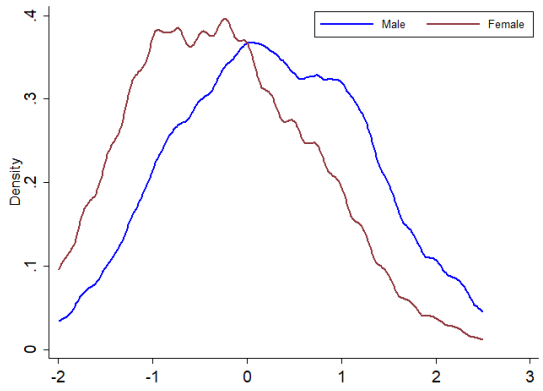
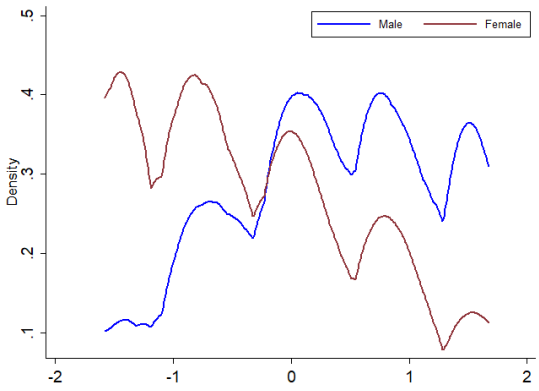



Figure A.3: SHAP summary plot predicting hours worked with basic set of predictors

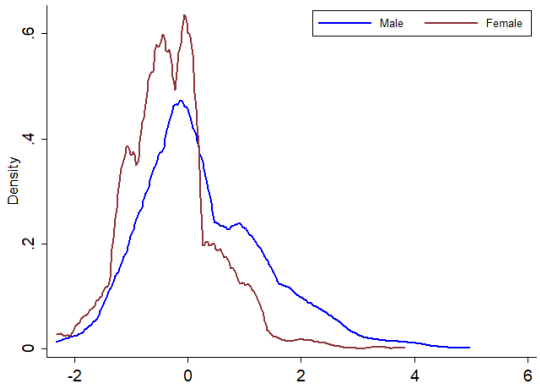
Note: See Figure A.2 notes, all of which apply here. The model includes 55 education, experience, demographics, and occupation variables as predictors of hours worked.




(a) Analytical Index




(b) Interpersonal Index




(c) Routine/Manual Index

Figure A.4: Kernal density plots of task indices by gender

Note: Figures plot the estimated kernal densities by gender for the analytical task index, interpersonal task index, and routine/manual task index.

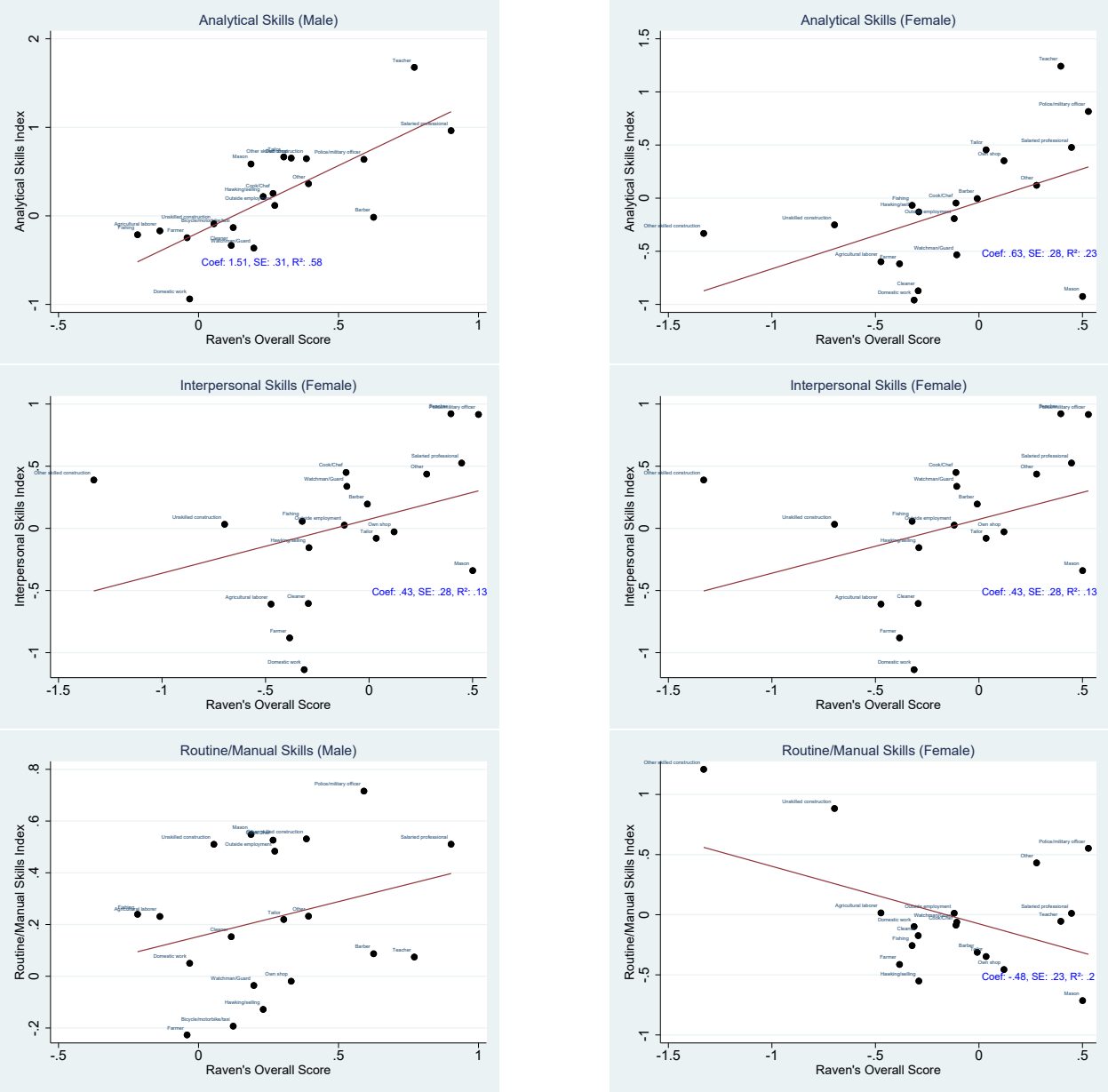



Figure A.5: Scatter plots of top 20 KLPS occupations of task indices and raven's overall score by gender

Note: Figures plot the raven's score for top 20 KLPS occupations by gender for the analytical task index, interpersonal task index, and routine/manual task index.

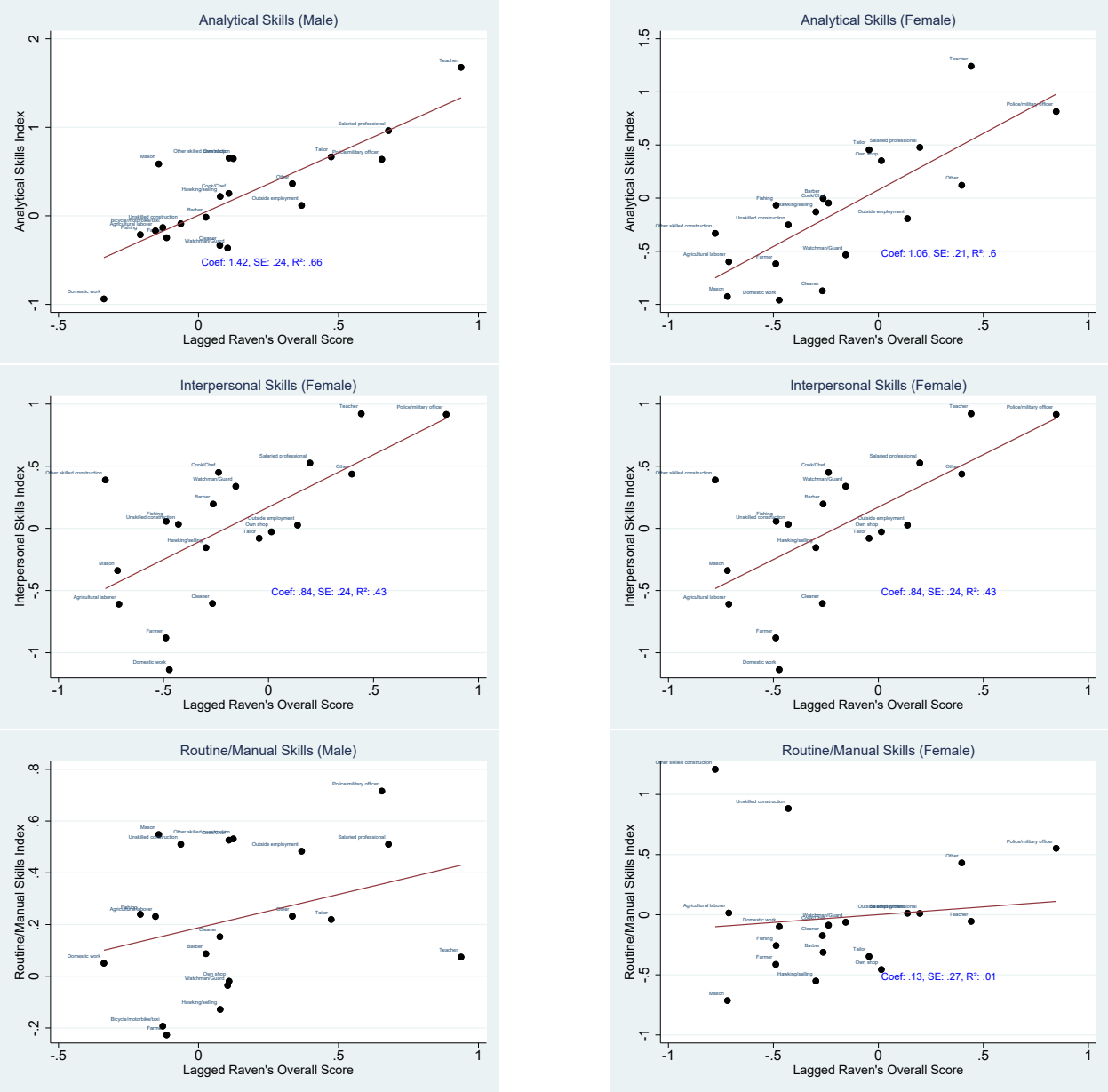



Figure A.6: Scatter plots of top 20 KLPS occupations of task indices and lagged raven's overall score by gender

Note: Figures plot the lagged raven's score for top 20 KLPS occupations by gender for the analytical task index, interpersonal task index, and routine/manual task index.

Table A.1: Differences in job task content by gender

| | Mean (female) | Mean (male) | Diff | Std. Error | t |
|---|---|---|---|---|---|
| **Analytical Index** | -0.26 | 0.26 | -0.52 | 0.03 | -20.05 |
| Length documents read as part of work | -0.20 | 0.19 | -0.39 | 0.03 | -15.13 |
| More than 30 min thinking as part of work | -0.16 | 0.16 | -0.32 | 0.03 | -12.01 |
| Do math as part of work | -0.20 | 0.20 | -0.40 | 0.03 | -14.96 |
| Learning new things as part of work | -0.16 | 0.17 | -0.33 | 0.03 | -12.16 |
| | | | | | |
| **Interpersonal index** | -0.34 | 0.33 | -0.67 | 0.03 | -26.62 |
| Supervise as part of work | -0.24 | 0.24 | -0.48 | 0.03 | -18.54 |
| Contact with clients as part of work | -0.18 | 0.15 | -0.32 | 0.03 | -12.09 |
| Make presentation at work | -0.21 | 0.22 | -0.43 | 0.03 | -16.39 |
| Collabrating with co-workers at work | -0.33 | 0.32 | -0.65 | 0.03 | -25.04 |
| | | | | | |
| **Routine/Manual index** | -0.30 | 0.27 | -0.57 | 0.03 | -22.30 |
| Freedom to decide as part of work | -0.11 | 0.09 | -0.20 | 0.03 | -7.31 |
| Repetitive tasks as part of work | -0.02 | -0.04 | 0.02 | 0.03 | 0.57 |
| Operate heavy machinery as part of work | -0.21 | 0.22 | -0.43 | 0.03 | -16.47 |
| Drive as part of work | -0.24 | 0.25 | -0.49 | 0.03 | -19.48 |
| Repair/maintain equipment as part of work | -0.13 | 0.14 | -0.27 | 0.03 | -10.12 |
| Work is physically demanding | -0.11 | 0.07 | -0.18 | 0.03 | -6.62 |
| Observations | 5609 | | | | |

Note: Table shows the gender differences and the t-test statistics for each component of the analytical, interpersonal, and routine/manual task indices.

Table A.2: Groupings of KLPS and Census at 1-digit occupational level

| Census 1-digit Grouping | Census (%) | KLPS 1-digit Grouping | KLPS (%) |
|---|---|---|---|
| Agricultural employee/self employed | 33.6 | Agriculture and fishing | 32.5 |
| Profesional/technical/managerial/clerical | 24.3 | Unskilled trades | 19.7 |
| Sales and Services | 16.7 | Retail and commercial | 18.1 |
| Unskilled manual | 13.7 | Skilled & semi-skilled trades | 16.7 |
| Household and domestic | 8.2 | Professionals | 11.2 |
| Skilled manual | 2.2 | Other occupation | 1.8 |
| Other | 1.3 | | |

Note: The table presents the distribution of occupations in the Kenyan Census (KDHS) and KLPS data at the 1-digit level.

Table A.3: Oaxaca-Blinder decomposition of wages, using female as the reference group

| | Coef. | 95% CI Lower | 95% CI Upper |
|---|---|---|---|
| With basic controls: | | | |
| Difference(Male-Female) | 0.780 | 0.705 | 0.854 |
| Explained | 0.285 | 0.135 | 0.435 |
| Unexplained | 0.495 | 0.337 | 0.652 |
| | | | |
| With full set of controls: | | | |
| Difference(Male-Female) | 0.780 | 0.691 | 0.868 |
| Explained | 0.359 | 0.285 | 0.433 |
| Unexplained | 0.421 | 0.323 | 0.519 |

Note: This table reports the Oaxaxa-Blinder decomposition results of the gender gap in log wages, using female as the reference group. "Basic" controls include eduaction, experience, demographics, and occupation variables. "Full" controls additionally include cognition, personality, economic preferences, job stress, and home care responsibilities.

Table A.4: Oaxaca-Blinder decomposition of hours worked, using female as the reference group

| | Coef. | 95% CI Lower | 95% CI Upper |
|---|---|---|---|
| With basic controls: | | | |
| Difference(Male-Female) | 12.099 | 10.593 | 13.606 |
| Explained | 5.008 | 1.763 | 8.253 |
| Unexplained | 7.091 | 3.646 | 10.537 |
| | | | |
| With full set of controls: | | | |
| Difference(Male-Female) | 12.099 | 10.201 | 13.998 |
| Explained | 7.194 | 5.479 | 8.909 |
| Unexplained | 4.905 | 2.450 | 7.361 |

Note: This table reports the Oaxaxa-Blinder decomposition results of the gender gap in hours worked using female as the reference group. "Basic" controls include eduaction, experience, demographics, and occupation variables. "Full" controls additionally include cognition, personality, economic preferences, job stress, and home care responsibilities.

Table A.5: Gender gap in hours worked: interaction terms

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Female (intercept) | -13.55*** | -46.59** | -4.12 | -48.27** |
| | (4.08) | (22.09) | (4.01) | (22.36) |
| **Interaction terms with Female** | | | | |
| Urban | -3.49 | -3.63 | -6.14** | -4.36* |
| | (2.59) | (2.62) | (2.48) | (2.54) |
| Nairobi | -0.63 | -0.18 | 6.86** | 2.44 |
| | (3.09) | (3.05) | (3.22) | (3.05) |
| Agriculture | 3.18 | 3.11 | -2.72 | -2.53 |
| | (2.43) | (2.39) | (2.48) | (2.43) |
| Self-employment/business | 3.43* | 3.47* | -0.90 | 0.20 |
| | (1.91) | (1.92) | (1.83) | (1.85) |
| Educational attainment | 0.07 | 0.64 | 0.05 | 1.24** |
| | (0.31) | (0.58) | (0.30) | (0.58) |
| General cognition score | | 8.03 | | 4.23 |
| | | (13.07) | | (12.65) |
| Executive functioning | | -4.56 | | -4.26 |
| | | (5.27) | | (5.33) |
| Orientation | | -1.55 | | 0.13 |
| | | (2.36) | | (2.39) |
| Memory | | -4.31 | | -3.73 |
| | | (6.95) | | (6.72) |
| Language | | -2.69 | | -2.49 |
| | | (3.22) | | (3.18) |
| Visuospatial | | -0.86 | | -0.54 |
| | | (2.55) | | (2.58) |
| Lagged Raven's | | 0.38 | | 0.03 |
| | | (1.13) | | (1.10) |
| Experience | | 2.47 | | 2.74 |
| | | (1.74) | | (1.78) |
| Experience$^2$/100 | | -0.05 | | -0.05 |
| | | (0.04) | | (0.04) |
| Analytical Index | | 3.16*** | | 3.44*** |
| | | (1.14) | | (1.18) |
| Interpersonal Index | | -0.33 | | 0.23 |
| | | (1.21) | | (1.17) |
| Routine/Manual Index | | 1.69 | | 1.87* |
| | | (1.21) | | (1.01) |
| N | 5772 | 5772 | 5772 | 5772 |
| Model | Regression | Regression | DML | DML |
| Controls | Basic | Full | Basic | Full |

Note: Columns 1 and 3 include the basic set of controls a la Column 4 in Table 4, along with the listed interaction terms. Columns 3 and 4 include the basic set of controls a la Column 7 in Table 4, along with the listed interaction terms. Columns 3 and 4 run DML regressions following the same setup as in Table 5. All standard errors are robust.

Table A.6: Gender gap in wages, winsorized at the 95th percentile

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| Female | -0.75*** | -0.60*** | -0.56*** | -0.49*** | -0.46*** | -0.40*** | -0.40*** |
| | (0.05) | (0.04) | (0.05) | (0.05) | (0.05) | (0.05) | (0.05) |
| Educational attainment | | 0.08*** | | 0.04*** | 0.02 | 0.01 | 0.01 |
| | | (0.01) | | (0.01) | (0.01) | (0.01) | (0.01) |
| Experience | | -0.13*** | | -0.04 | -0.03 | -0.02 | -0.03 |
| | | (0.04) | | (0.04) | (0.04) | (0.04) | (0.04) |
| Experience$^2$/100 | | 0.00** | | 0.00 | 0.00 | 0.00 | 0.00 |
| | | (0.00) | | (0.00) | (0.00) | (0.00) | (0.00) |
| Urban | | 0.58*** | | 0.40*** | 0.39*** | 0.34*** | 0.33*** |
| | | (0.06) | | (0.06) | (0.06) | (0.05) | (0.05) |
| Nairobi | | 0.15** | | 0.11* | 0.12** | 0.11* | 0.12** |
| | | (0.07) | | (0.06) | (0.06) | (0.06) | (0.06) |
| Agriculture | | 0.10* | | -0.03 | -0.04 | -0.01 | -0.01 |
| | | (0.05) | | (0.05) | (0.05) | (0.05) | (0.05) |
| Self-employment/business | | 0.03 | | 0.26*** | 0.27*** | 0.23*** | 0.19*** |
| | | (0.04) | | (0.05) | (0.05) | (0.05) | (0.05) |
| General cognition score | | | | | 0.18*** | 0.16*** | 0.00 |
| | | | | | (0.02) | (0.02) | (0.26) |
| # Older adults in HH | | | | | 0.03 | 0.03 | 0.05 |
| | | | | | (0.06) | (0.06) | (0.06) |
| # Children in HH | | | | | -0.04*** | -0.05*** | -0.04*** |
| | | | | | (0.01) | (0.01) | (0.01) |
| Analytical Index | | | | | | 0.10*** | 0.08*** |
| | | | | | | (0.03) | (0.03) |
| Interpersonal index | | | | | | 0.10*** | 0.09*** |
| | | | | | | (0.03) | (0.03) |
| Routine/Manual index | | | | | | -0.00 | 0.02 |
| | | | | | | (0.02) | (0.02) |
| Job stress index | | | | | | | -0.13*** |
| | | | | | | | (0.02) |
| N | 5063 | 5063 | 5061 | 5061 | 5061 | 5061 | 5061 |
| r2 | 0.08 | 0.22 | 0.32 | 0.35 | 0.36 | 0.37 | 0.39 |
| Occupation_FE | | | Yes | Yes | Yes | Yes | Yes |
| Personality | | | | | | | Yes |
| Preferences | | | | | | | Yes |

Note: Table follows the exact same specification and structure as Table 3, but the log hourly wage variable is winsorized at the 95th percentile which is 5.75 (equivalent to 313.43 KSH or 2.24 USD). All standard errors are robust.

Table A.7: Gender gap in hours worked, winsorized at the 95th percentile

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| Female | -11.64*** | -10.21*** | -9.54*** | -8.27*** | -8.15*** | -6.62*** | -6.57*** |
| | (0.93) | (0.90) | (1.01) | (0.97) | (0.97) | (1.01) | (1.03) |
| Educational attainment | | 0.58** | | 0.42 | 0.32 | 0.28 | 0.31 |
| | | (0.28) | | (0.26) | (0.26) | (0.26) | (0.27) |
| Experience | | 1.15 | | 1.08 | 1.10 | 1.36 | 1.42* |
| | | (0.95) | | (0.85) | (0.85) | (0.85) | (0.83) |
| Experience$^2$/100 | | -0.03 | | -0.02 | -0.02 | -0.03 | -0.03 |
| | | (0.02) | | (0.02) | (0.02) | (0.02) | (0.02) |
| Urban | | 6.05*** | | 2.14* | 1.99 | 1.01 | 0.82 |
| | | (1.27) | | (1.23) | (1.24) | (1.26) | (1.27) |
| Nairobi | | -1.62 | | -1.64 | -1.57 | -1.75 | -1.66 |
| | | (1.56) | | (1.41) | (1.41) | (1.38) | (1.38) |
| Agriculture | | 0.33 | | -2.66** | -2.76** | -2.25* | -2.39** |
| | | (1.19) | | (1.15) | (1.16) | (1.17) | (1.16) |
| Self-employment/business | | 12.23*** | | 10.98*** | 11.00*** | 10.20*** | 10.26*** |
| | | (0.88) | | (0.99) | (0.99) | (1.00) | (1.01) |
| General cognition score | | | | | 0.68 | 0.26 | 14.05** |
| | | | | | (0.52) | (0.52) | (5.89) |
| # Older adults in HH | | | | | -0.91 | -0.88 | -0.81 |
| | | | | | (1.18) | (1.18) | (1.19) |
| # Children in HH | | | | | -0.26 | -0.29 | -0.32 |
| | | | | | (0.24) | (0.23) | (0.23) |
| Analytical Index | | | | | | 1.46** | 1.48*** |
| | | | | | | (0.57) | (0.57) |
| Interpersonal index | | | | | | 3.06*** | 3.13*** |
| | | | | | | (0.61) | (0.61) |
| Routine/Manual index | | | | | | 0.29 | 0.24 |
| | | | | | | (0.54) | (0.53) |
| Job stress index | | | | | | | 0.37 |
| | | | | | | | (0.50) |
| N | 5773 | 5773 | 5772 | 5772 | 5772 | 5772 | 5772 |
| r2 | 0.04 | 0.10 | 0.22 | 0.26 | 0.26 | 0.27 | 0.28 |
| Occupation_FE | | | Yes | Yes | Yes | Yes | Yes |
| Personality | | | | | | | Yes |
| Preferences | | | | | | | Yes |

Note: Table follows the exact same specification and structure as Table 4, but the weekly hours worked variable is winsorized at the 95th percentile which is 89 hours. All standard errors are robust.

Table A.8: Gender wage gap, controlling for hours worked

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Female | -0.59*** | -0.48*** | -0.65*** | -0.51*** |
| | (0.05) | (0.05) | (0.04) | (0.04) |
| hours_lastweek | -0.01*** | -0.01*** | | |
| | (0.00) | (0.00) | | |
| N | 5061 | 5061 | 5061 | 5061 |
| Model | Regression | Regression | DML | DML |
| Controls | Basic | Full | Basic | Full |

Note: Columns 1 and 2 of this table follows the exact same specification as Columns 4 and 7 of 3, but include weekly hours worked as an additional covariate. Columns 3 and 4 are the DML analogues of Columns 1 and 2, again with weekly hours worked as an additional covariate. Because DML allows the relationship between covariates and the outcome to be nonlinear, there is no coefficient on hours_lastweek in Columns 3 and 4. All standard errors are robust.

Table A.9: Attitudes toward gender norms from KLPS-3

| **Statement** | Agree (Male) | Agree (Female) | Female Coef (No controls) | Female Coef (All controls) |
|---|---|---|---|---|
| It is okay for a woman to be a mechanic | 0.92 | 0.82 | -0.10*** | -0.05*** |
| | | | (0.01) | (0.02) |
| Important decisions in the family should be made by the men of the family (reverse-coded) | 0.30 | 0.31 | 0.00 | 0.06*** |
| | | | (0.02) | (0.02) |
| If wife is working outside the home, husband should help her with household chores | 0.85 | 0.79 | -0.06*** | -0.04** |
| | | | (0.01) | (0.02) |
| Women can be good politicians and should be encouraged to stand in elections | 0.90 | 0.89 | -0.01 | 0.01 |
| | | | (0.01) | (0.02) |
| In our country, women should have equal rights and receive the same treatment as men | 0.89 | 0.90 | 0.01 | 0.03 |
| | | | (0.01) | (0.02) |
| No one has the right to use physical violence against anyone else | 0.86 | 0.83 | -0.04*** | 0.00 |
| | | | (0.01) | (0.02) |

Note: The first two columns of Panel A show the the share of individuals, among men and women respectively, who agree with the statement on the left-most column. The third and fourth columns report the regression coefficient of an indicator for agreeing with the statement on gender in the full sample. The third column includes no covariates; the fourth column includes the full set of covariates used in regression analyses in Table 3. The second statement, "Important decisions in the family should be made by the men of the family", is reverse-coded such that the numbers show the share of men and women who disagree with that statement. Each of the last three statements (A) was presented together with an opposing statement (B), and the respondent was asked to select "agree very strongly with A", "agree with A", "agree with B", "agree very strongly with B", or "agree with neither". Specifically, the opposing statements are "women should stay at home to take care of their children", "women have always been subject to traditional laws and customs, and should remain so", and "a married man has a right to beat his wife if she misbehaves". Data from KLPS-4 survey administered between 2018 and 2021. $N = 4,741$ (excluding individuals in KLPS-5 who did not complete the KLPS-3 survey).

# B Data Appendix

## B.1 Cognition data

Participants in KLPS-5 were administered a modified version of the Harmonized Cognitive Assessment Protocol (HCAP), a neuropsychological assessment battery that comprises tests and items in common with other international partner studies of the US Health and Retirement Study (HRS) (Gross et al., 2024). The standard HCAP cognitive battery was adapted to the Kenyan context and the Swahili language. Confirmatory factor analysis modeling that generates the general and component-specific scores was conducted using Mplus software, version 8.7 (41). We describe the details of the HCAP test and adaptations made below.

- **Swahili Mental State Exam (SMSE):** This test is an adaptation of other mental state exams that is adapted for Swahili-speaking populations (such as those in KLPS), evaluating domains such as temporal orientation (today's date, month, year, and season), spatial orientation (place), registration, attention, recall, and speech.
  - The question "What state are we in right now?" was replaced with "What county are we in right now?" since there are no states in Kenya. Due to recent redistricting in Kenya, it is more likely that participants would know their county rather than the province.
  - Acceptable answers for "Which season is it?" were expanded to include the "long rain season," the "short rain season," and the "January season," rather than just traditional US seasons like winter, spring, summer, and fall (autumn), which are not locally relevant in East Africa.
  - Since most surveys were conducted either outside or in homes, questions like "What floor are we on?" and "Name of building?" were replaced with "What is this place we are in used for?" and "What is the name of this city/place/town?"
  - Participants surveyed in the Innovations for Poverty Action (IPA) office were asked "What is the name of the building we are in now?" instead of "What floor are we on?"
  - The backwards spelling of "world" was omitted because the Swahili translation for world has multiple syllables. We did not replace this item as we did not find an equivalent Swahili word.
  - In the repeat a phrase task, "No ifs ands or buts" was replaced with "Tupe tupate tumpatie Taaka" because the English phrase does not have a Swahili equivalent that is a tongue twister.
- **Object Naming:** This test involved three questions on items that should be considered common for the participants.
  - "Coconut" was replaced with "watermelon" because coconut is not common in most rural parts of Kenya, and watermelon is common across all parts of Kenya.
  - "Prime Minister" was replaced with "The President of Kenya."
- **Symbol Cancellation Test (SCT):** This cognitive assessment tool evaluates visual selective

attention, requiring participants to identify and mark specific target symbols from a larger array of distractor symbols within one minute.

- **Community Screening Instrument (CSI-D, 4 items):** This is a test commonly used for detecting dementia in community settings.
  - "Point to elbow" was changed to "Point to neck" because the word for elbow is not widely used in Swahili or local languages and dialects.
  - "Point to the window and then the door/ceiling." If there was no window, then participants were asked to point to the ceiling and then at the door. If the survey was conducted outdoors, then participants were asked to point to the sky and then to the floor.
- **Raven's Progressive Matrices (RPM) Test:** This cognitive test measures abstract reasoning and non-verbal intelligence, particularly fluid intelligence, which involves reasoning and problem-solving skills that are independent of acquired knowledge.
- **Digit Span Tests:** These are cognitive assessments used to measure verbal short-term and working memory. The numbers were read to the participants in a monotonic nature (said in a continuous prose without tonal variation).
  - **Forward Digit Span:** Assesses short-term auditory memory by asking participants to repeat a sequence of digits in the same order.
  - **Backward Digit Span:** Assesses working memory and cognitive control by asking participants to repeat a sequence of digits in reverse order.
- **Similarities and Differences: Judgment and Problem Solving:** This test involves identifying similarities and differences between various concepts, objects, or ideas. The test is used to assess abstract reasoning, verbal comprehension, and conceptual thinking.
  - The "roses and daisies" comparison item was changed to "kale and spinach" because these flower names are not common in rural Kenya.
- **10 Word List Recall:**
  - "Butter" was replaced by "milk" since butter is not common in rural Kenya.
- **Found a Lost Child:** This test assesses cognitive abilities in real-world contexts, using cognitive functions like problem solving, executive functioning, and social cognition.
  - The scenario involving "found a lost child" was changed to "found a lost child that is looking for their parents" because children often walk around alone without adults in Kenya, making the original scenario less unusual.
- **Logical Memory Story #2 (East Boston Story):** Part of the Wechsler Memory Scale (WMS), this test assesses verbal memory through story recall.
  - The name and city were replaced with a localized version, "Anne Wanyonyi from East Nairobi." "Cafeteria" was replaced with "a kitchen," "State Street" with "Main Street," and dollars with shillings (the local currency unit).
- **Constructional Praxis:** This test assesses the ability to construct or assemble objects, involving motor skills and visual and spatial perception.

  - The diamond shape was replaced with a triangle since diamonds are not commonly taught in Kenyan schools.
- **Go/No-Go Trial:** This test measures the ability to control impulsive responses, assessing executive functioning, response inhibition, attention, focus, and impulse control.
- **Hand Movement Sequencing Test:** This assessment evaluates executive functions related to the frontal lobe, including testing motor planning and execution, inhibitory control, attentional flexibility, and working memory.
- **Token Test:** This assessment evaluates auditory comprehension and language processing abilities.
  - Diamonds were replaced with triangles (for reasons noted above).

### B.2 Construction of job task indices

We follow the procedure adopted by the World Bank's Skills Toward Employment and Productivity (STEP) surveys, which is an initiative to generate internationally comparable data on skills and tasks available in developing countries, described in Dicarlo et al. (2016). Specifically, we construct the following three job task indices. Each one of the following questions is standardized into a variable with mean 0 and SD 1. Then the components of each index are added to yield the index, ensuring that all questions receive equal weights.

The analytical Index is the sum of all below:

- Length documents read as part of work. Measured from 1-5. 1= One page or less, 2= 2 to 5 pages, 3= 6 to 10 pages, 4= 11 to 25 pages and 5= More than 25 pages.
- Do tasks that involve more than 30 min thinking as part of work. Measured from 1-5. 1= Never, 2= Less than once a month, 3= Less than once a week but at least once a month, 4= At least once a week but not every day and 5= Every day.
- Do math as part of work. Measured from 1-5. Sum of answer choices of a) Measure or estimate sizes, weights, distances, etc., b) Calculate prices or costs, c) Use or calculate fractions, decimals, or percentages, d) Perform any other multiplication or division and e) Use more advanced math, such as algebra, geometry, trigonometry.
- Learning new things at work. Measured from 1-5. 5= Every day, 4= At least once a week, 3= At least once a month, 2= At least every 2-3 months, 1= Rarely or never. These measures are reverse-coded to be consistent with other measures that contribute to this index.

The interpersonal index is the sum of all below:

- Supervise as part of work. Measured from 0-1. 1= yes.
- Contact with clients as part of work. Measured from 1-10. Specifically the survey asks: "Using any number from 1 to 10, where 1 is little involvement or short routine involvements, and 10 means much of the work involves meeting or interacting for at least 10-15 minutes at

a time with a customer, client, student or the public, what number would you use to rate this work?"

- Making a presentation at work. Measured from 0-1. 1= yes.
- Collaborating with co-workers at work. Measured from 1-5. 1= Never, 2= Less than once a month, 3= Less than once a week but at least once a month, 4= At least once a week but not every day and 5= Every day.

The Routine/Manual index is the sum of all below:

- Freedom to decide as part of work (reverse-coded). How much freedom do you (did you) have to decide how to do your work in your own way, rather than following a fixed procedure or a supervisor's instructions? Use any number from 1 to 10, where 1 is no freedom and 10 is complete freedom. These answers are reverse-coded for constructing the index.
- Frequency of repetitive tasks as part of work (reversed). Measured from 1-4. 4= Almost all the time, 3= More than half the time, 2= Less than half the time, 1= Almost never. These answers are reverse-coded for constructing the index.
- Operate heavy machinery as part of work. Measured from 0-1. 1= yes.
- Drive as part of work. Measured from 0-1. 1= yes.
- Repair/maintain electronic equipment as part of work. Measured from 0-1. 1= yes.
- Work is physically demanding. Measured from 1-10 where 1 is not at all physically demanding (such as sitting at a desk answering a telephone) and 10 is extremely physically demanding (such as carrying heavy loads, construction worker, etc). Higher number means more physically demanding.

### B.3 Economic preferences

Economic preference data was collected during KLPS-4, which was administered between 2018 and 2021. This module assesses individuals' responses across four key domains: risk aversion, time preferences, altruism, and trust:

- Risk preferences: Participants were offers a 50% chance of receiving some amount for sure versus a lottery. Following standard approaches, we keep the 50% lottery amount fixed at 900 Shillings, but we ask a series of questions varying the amount of payment for sure to pin down the individual's certainty equivalent.
- Social preferences: Participants were asked whether they would buy a thank-you gift for stranger who spent an hour of their own time to help them (reciprocity), give to a charity helping people in their ancestral home area, (in-group altruism ), and give to a charity helping people in Kenya but not in their ancestral home area (out-group altruism)
- Time preferences: Participants were asked to choose between a payment of a certain amount today versus a higher amount in one month. We keep the amount today fixed at 300 Shillings but vary the amount in the future across a series of questions.

- Ambiguity: Participants were asked to choose between sure probabilities versus unsure ones by playing a game where they drew a ball out of a bag without looking. There are two bags with 10 balls each. In bag 1, out of 10 balls there are 4 red balls and 6 yellow balls. In bag 2, there are also 10 balls, but the number of red and yellow balls is unknown. They can choose a bag from which they are to draw the ball. If they choose bag 1, to win 50 shillings they need to draw a red ball. If they choose bag 2, to win 50 shillings they need to decide a color and draw a ball of that color.

We code up risk and time preferences into continuous variables based on the amount that the participant was trading off for. In total, 9 variables are used from the economic preferences module in our analyses.

### B.4 Big Five personalities

Big Five personality data was collected during KLPS-4, administered between 2018 and 2021. Each of the five personality dimensions was assessed using three questions.

- Extraversion: Tends to be quiet (reverse-coded), act as a leader, is full of energy
- Agreeableness: Compassionate, rude to others (reverse-coded), assumes the best of people
- Conscientiousness: Is disorganized (reverse-coded), has difficulty getting started on tasks (reverse-coded), is reliable
- Neuroticism: Worries a lot, tends to feel depressed, is emotionally stable
- Openness: Is fascinated by art, music, or literature, interested in abstract ideas, comes up with new ideas

All questions are measured on a 5-point scale from "disagree strongly" (1) to "agree strongly" (5). The score on each personality is the sum of its three measures. We also include questions about fatalism—"is it impossible to escape a predetermined fate" and life satisfaction—"how satisfied are you with life". These yield seven personality variables that we use in our analyses.

### B.5 Lagged Raven's score

The 'lagged' Raven's scores were measured during the KLPS-2 and KLPS-3 survey rounds. The KLPS-2 Integrated Module (KLPS-2-I), conducted between 2007 and 2009, included the Raven's Progressive Matrices test, which was administered to 5,081 respondents aged 18 to 34. During the KLPS-3 Integrated Module (KLPS-3-I), conducted between 2011 and 2014, the Raven's Progressive Matrices test was administered to 3,032 individuals. This assessment was only given to respondents who had not participated in KLPS-2, as their scores were unlikely to change significantly between survey rounds due to their age. The KLPS-3-I respondents were between 15 and 31 years old. Thus, the KLPS-3 Raven's scores correspond to cognitive assessments conducted approximately 13 years ago, while the KLPS-2 Raven's scores represent assessments from about 16 years ago. More

precisely, the lagged Raven's scores used in this analysis were collected 13 to 16 years before the KLPS-5 round.