NBER WORKING PAPER SERIES

BEYOND EDUCATION AND OCCUPATION:
UNPACKING THE LARGE GENDER WAGE GAP IN KENYA

Uyanga Byambaa
Edward Miguel
Michael W. Walker
Samuel Zicheng Wang

Working Paper 34375
http://www.nber.org/papers/w34375

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2025

We thank seminar participants at UC Berkeley for helpful comments and suggestions, Michelle Layvant, Eric Ochieng and Aditi Poduri for research assistance, and 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. Wang worked on the paper prior to joining Amazon. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Beyond Education and Occupation: Unpacking the Large Gender Wage Gap in Kenya
Uyanga Byambaa, Edward Miguel, Michael W. Walker, and Samuel Zicheng Wang
NBER Working Paper No. 34375
October 2025
JEL No. J16, O12

ABSTRACT

Gender wage gaps persist globally, particularly in poor countries. Using Kenya Life Panel Survey data, we first document a raw gender wage gap of 79 log points (55%). We show it remains large, at 39 log points (32%), even controlling for a novel set of individual characteristics – cognitive performance, personality traits, economic preferences, and job tasks – in addition to standard covariates. These novel factors account for only 20% of the residual gap unexplained by education and occupation. Though most Kenyans report egalitarian gender views, these patterns suggest that barriers still hinder women's labor outcomes.

Uyanga Byambaa
University of California, Berkeley
u.byambaa@berkeley.edu

Edward Miguel
University of California, Berkeley
Department of Economics
and NBER
emiguel@econ.berkeley.edu

Michael W. Walker
University of California, Berkeley
mwwalker@berkeley.edu

Samuel Zicheng Wang
University of California, Berkeley
samwang755@gmail.com

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 16% less than men (U.S. Department of U.S. Department of Labor 2024), while women's gross hourly earnings were on average 12.7% below those of men in the European Union (Eurostat, 2022). In low- and middle-income countries (LMICs), gender gaps in the labor market are often even larger (Nopo et al., 2011; Alarakhia et al., 2023; Heath et al., 2024).

However, interpreting these differences in mean (or median) earnings by gender is more challenging. The presumption that these may be due to discrimination, social constraints, or lack of opportunity is reasonable, due to historical gender biases and social norms. Yet individual characteristics, attributes and preferences may also differ by gender and play a role in labor market choices and worker productivity. Even if men and women work in the same occupation, they may – by their own choice or due to their employers – have different day-to-day job tasks, which could contribute to the gender wage gap.

Relatively few studies have detailed productivity information on female and male workers undertaking the same task. Thus, most empirical studies take the approach of estimating the gender wage gap controlling for key demographics – education, age/experience, and occupation, which are often available in labor market administrative datasets – and most find that the gender wage gap remains large (Blau and Kahn, 2017). Such analyses have been more limited (or focused on selected samples, such as urban wage workers) in many LMIC settings due to data availability. 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 the possibility that omitted variables could (at least partially) account for the observed gender wage gap.

The present study aims to make progress on this issue by leveraging highly detailed individual-level data in Kenya covering both the "standard" human capital and labor market variables noted above, as well as a greatly expanded set of individual covariates from the Kenya Life Panel Survey (KLPS). The KLPS is a longitudinal 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 5,300 prime-age adult individuals living and working in both urban and rural Kenya. Most individuals in the KLPS-5 sample range are between 33 and 40 years of age (the 10-90 range) and so are broadly at the midpoint of their prime working years, and about half are female. The survey gathers information on individual labor market earnings, labor supply, occupation, education, and demographic information. Beyond these standard measures, KLPS surveys have also gathered detailed modules on individual cognitive performance, economic preferences (e.g., risk aversion, time preferences), Big Five personality measures, the type of tasks that individuals do on the job, and caregiving responsibilities.

In addition to the rich set of covariates, the tracking and data collection strategy of the KLPS mitigates some selection concerns relative to typical labor force surveys. While all individuals in the KLPS attended rural primary schools in childhood, the KLPS tracks individuals that have migrated throughout Kenya (and neighboring countries), providing data on both rural and urban workers. At the time of KLPS-5, 45% lived in urban areas (with 24% in the major cities of Nairobi or Mombasa). The earnings data covers wage work, self-employment and agricultural activities, thus minimizing concerns about selection into urban areas or the formal employment sector.

Kenya offers an interesting study setting as a lower middle income country that is viewed as a relative success story in terms of gender equality among its regional neighbors. 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 (World Bank, 2025). 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).[1]

The analysis employs a combination of classic selection-on-observables linear regressions as well as machine learning approaches. Machine learning methods allow us to first predict wages using a vector 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 estimation model

[1] Based on data from the International Labour Organization. Labor force participation rate is the proportion of the population ages 15 and older that is economically active. Note that part of the lower participation rate in high-income countries reflects the fact that more people study longer in high-income countries.

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 start by documenting that the raw gender wage gap (e.g. without covariates) between women and men in this Kenyan sample is 79 log points, or 55 percent, far larger than in rich countries. Echoing existing work, controlling for individual education, work experience, demographics, and occupation, the gender wage gap falls somewhat but it remains large at 49 log points (39 percent). This is consistent with observed gender differences along these dimensions: women have an average of around one less year of education, and are more likely to be working in agriculture or small retail (lower-paying occupations), and less likely to be working as professionals. We find that gender is a key predictor of wages: it is the third-most important predictor out of the 104 variables (some described further below) employed in a flexible machine learning model.

A main innovation of the present study is the use of more detailed individual and job-specific covariates than are typically available in existing studies. We begin by characterizing gender differences in these midlife measures of cognition, preferences, personality, job tasks, and caregiving responsibilities, and document differences across multiple domains. We find that the women in this sample on average score somewhat 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 some of the "Big Five" personality measures. This raises the possibility that including these factors as controls could account for a meaningful share of the gender wage gap.

However, inclusion of these additional variables as controls only accounts for 20% of the remaining gender wage gap, yielding an estimated difference between women and men that remains large at 39 log points (32 percent). The Oaxaca-Blinder decomposition delivers a similar punchline: with the full set of both standard individual controls and the expanded set of characteristics noted above, roughly half of the raw gender gap remains (39 log points).

Beyond the gender wage gap (i.e., the gap in earnings per hour), we also document a sizable gender gap in labor supply, with women on average working 12.22 fewer hours per week than men (who work 47 hours per week on average). While accounting for the full set of individual controls (with the expanded variables, including caregiving responsibilities) narrows the gap, it remains large at 6.98 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 most existing research.

A strength of using the KLPS dataset for this analysis 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. The availability of wage and employment data for the vast majority of individuals in our sample (nearly 90%), regardless of their place of residence or occupation, helps to alleviate leading concerns about selection of workers into particular labor markets, which is a common limitation in existing studies that primarily focus on urban and/or formal-sector workers. While the study sample is not designed to be nationally representative, it offers a unique advantage in understanding how selection may affect estimates of the gender wage gap. Indeed, we show that if one simply restricts the sample to only urban wage workers, the raw gender wage gap narrows considerably down to 52 log points from a raw gap of 79 log points, a reduction of nearly one third.

So what can explain the large and persistent gender wage (and hours worked) gaps 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 these factors 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.[2] Yet, despite these egalitarian attitudes, we find that cognitively high-performing women have far fewer analytical tasks in their jobs (even conditional on broad occupational category), suggesting that some components of the

[2]Of course, experimenter demand effects are possible when it comes to survey responses but we view these findings as only de epening the puzzle, as the survey patterns argue against a simple discrimination-based explanation for gender gaps in the Kenya labor market.

labor market and/or social norms may be preventing high-ability women from fully utilizing their skills. Taken together, these results point to the need for more investigation on the drivers of persistent gender gaps in the Kenyan labor market.

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 (U.S. 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. More 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 gap is also a leading labor market issue in LMICs (Nopo et al., 2011; 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 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, similar to our findings. Delecourt et al. (2023) find that women-run businesses in western Kenya make 47% fewer profits than male-run businesses, and women are more likely to locate their business close to home. A 2023 UN Women study on Kenya estimates that the gender pay gap persists 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 the sample in this latter study is nationally representative, wage data is available for only 20% of adults, which is an important difference relative to the present study, in which earnings, hours and wage data are gathered for nearly the entire sample.

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 in a similar direction. Boissiere et al. (1985) study this with Raven's Matrices test scores as well as 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 a 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 cognitive skills generate large 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, though these surveys only cover individuals living in urban areas.

This paper also contributes to a more 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. Many existing papers use post-double LASSO (Belloni et al., 2014), and we apply the related Double/Debiased Machine Learning (DML) approach (Chernozhukov et al., 2016, 2018).

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.[3] 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 survey rates of 82% of original respondents in KLPS-5, among those that were non-deceased.[4]

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 Cognitive-Plus (C+) Module). The KLPS-5 analysis sample used in this paper includes 5,338 individuals aged roughly 31-47 who are in their prime working years, and 51% are male and 49% are female. Additionally, there are sizable shares of respondents residing in rural (56%) and urban (44%) areas, including about a quarter residing in the major cities of Nairobi or Mombasa. Appendix Figure A.1 shows the county-level number of respondents in the KLPS-5 analysis sample among respondents residing in Kenya, for men and women separately. The maps show a relatively similar geographical distribution for men and women, with Busia County (the original study area) and Nairobi being the two locations with the largest number of respondents.

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). Respondents also report 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 644 individuals who do not

[3]The PSDP involved an experimentally-assigned rollout of school-based deworming treatment to students (Miguel and Kremer, 2004). The KLPS has also tracked female participants of a girls scholarship program (GSP) (Kremer et al., 2009); since male participants in GSP were not tracked in recent KLPS survey rounds, we restrict attention to individuals that participated in the PSDP.

[4]KLPS utilizes a two-stage tracking approach (see Hamory et al., 2021). Appendix Table A.1 shows the effective tracking and survey statistics for both the fourth and the fifth rounds of KLPS.

report positive earnings from all three sources or report zero hours worked in the past 7 days and 5 individuals with missing urban/rural residency status. This leaves us with a sample of 4,694 (87.9%) of individuals with a non-zero wage. Using the relevant exchange rate of 1 USD = 140 Kenyan Shillings, workers in the KLPS-5 sample on average work 41.4 hours per week, earn $96 per month, and have an hourly wage of $0.75, while the median worker works 40 hours per week, earns $42 per month, and has an hourly wage of $0.26.[5]

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 15 occupation categories, which we can compare with the mix of occupations in the most recent (2022) Kenyan Continuing Household Survey (KCHS), implemented by the Kenya National Bureau of Statistics. The 2022 KCHS includes a representative sample of 24,000 households across urban and rural Kenya with a sample frame from the 2019 Kenyan Population and Housing Census (Kenya National Bureau of Statistics, 2024).[6] We find that the distribution of occupations at the 1-digit level is similar in both datasets (Appendix Table A.2), though KLPS-5 includes a smaller share of individuals (30%) in agriculture compared to KCHS (48%), and higher shares of individuals in wholesale and retail trade, professional activities, public administration, and education sectors. The occupation comparison shows that KLPS-5 captures a broadly representative snapshot of individuals in the Kenyan labor market. The collection of occupation information allows for the inclusion of occupation fixed effects in the analysis.

KLPS data also provides standard demographic variables often included in analyses of the gender wage gap, including age/experience, education, and urban/rural status. In addition, we can make use of household rosters to account for potential differences in home care responsibilities between men and women, which may disproportionately affect women's labor force participation if caregiving responsibilities mainly fall on them. Specifically, in the analysis we focus on the number of children (under 18) and the number of older adults over 65 years old in the household as proxies for care responsibilities.

[5] 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, which is around the midpoint of KLPS-5 data collection.

[6] Specifically, in the KCHS data, occupation levels were provided at the International Standard Industrial Classification of All Economic Activities (ISIC) Revision 4 level. We map occupations in KLPS-5, which are more detailed, to ISIC Rev.4 occupations.

Central to the contribution in this study, the KLPS-5 survey collects a wide range of variables that are less commonly measured and that could plausibly differ by gender. First, multiple rounds of KLPS have collected measures of individual cognitive ability. 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 (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 modules of cognitive questions were administered. Responses were aggregated into a general cognition score as well as five domain-specific scores (executive functioning, orientation, memory, language, visuospatial) following the Confirmatory Factor Analysis (CFA) approach in Gross et al. (2025), a common method in the psychology literature. Details of these cognition tests and their validity in this sample are described in Gross et al. (2025) and Appendix B.1.

A second area of innovation comes through detailed measurement of job tasks and job characteristics. KLPS-5 administered a job task content module for each respondent's primary job; this allows us to investigate if females and males report performing systematically different tasks overall and even within the same occupation. 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). They capture three dimensions of job task content: analytical tasks (which measure the degree of analytic reasoning skills used at work); interpersonal tasks (the degree of interactive, communication, and managerial skills used at work); and routine/manual tasks (capturing the degree of physical movement and repetitive work). We sum up the scores from questions in each category following the approach in Lo Bello et al. (2019) and standardize each index weighted by sampling weights; Appendix B.2 includes details.

KLPS-5 also measures other job-related characteristics, including the reported stress level of an individual's primary job. Here, we largely follow the approach used in the Health and Retirement Surveys (HRS, Sonnega et al., 2014). Respondents were assessed with a 15-item questionnaire taken from the widely used Karasek Job Content Questionnaire (Karasek et al., 1998). Respondents 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.

Finally, we utilize data from previous KLPS rounds for relevant measures that arguably remain relatively stable over time. These include personality measures and measures of economic preferences collected as part of KLPS Round 4 (KLPS-4, 2018-2022). In particular, KLPS-4 gathered data on individuals' Big Five Personality traits—namely, extroversion, agreeableness, openness, conscientiousness, and neuroticism—from 15 questions commonly used in psychology research. We also ask respondents two additional questions—"Is it impossible to escape a predetermined fate?" and "How satisfied are you with your life?"—to measure their attitude toward fatalism and life satisfaction, respectively. The economic preferences module elicited individuals' preferences toward taking risks, reciprocity, out-group altruism, time preferences, and ambiguity aversion; Appendix B.3 and B.4 describe the construction of these variables in more detail. From earlier KLPS rounds, we are also able to incorporate Raven's Matrices test scores as well as survey questions regarding gender norms, roles and expectations, as noted below.

In summary, the long-term KLPS panel dataset allows for the estimation of mean differences in labor market outcomes by gender overall and while controlling for both a "standard" set of socio-demographic covariates as well as a much more detailed set of individual- and job-specific characteristics that are not commonly collected, especially in LMICs.

3 Disparities in labor market outcomes by gender

This section estimates differences in earnings by gender using a range of empirical models and controls, including both linear regression and machine learning approaches. The central finding is that gender gaps are large, significant and robust in this Kenyan population.

3.1 Summary statistics of labor market 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, 34% more than women, who work 35.3 hours on average. Men also earn significantly more than women: on average, men's monthly earnings are 141.6 USD, whereas women's monthly earnings are 49.8 USD. Part of the differences in earnings reflects differences in labor supply, but the hourly wage also differs substantially. On average, men earn 1.02 USD per hour, which is twice as much as women, who earn who earn 0.50 USD per hour. The gender wage gap remains stark if we look at median wages, at 0.40 USD per hour for men versus 0.17 USD per hour for women. Compared to women, men are slightly more likely to engage in self-employment (46.0% vs 41.5%), much more likely to engage in wage employment (53.1% vs 33.9%), and much less likely to have no work hours (4.7% vs 15.1%).[7]

One obvious 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 relatively modest (Table 1, row 2). Women have slightly lower mean levels of education (8.7 years) than men (10 years) in our sample. Given the lack of a pronounced gender gap in education, it is unsurprising that there is also not a noticeable gender gap in years of experience[8]: women on average have 22 years of experience and men 21 years.

Another common explanation for the gender wage gap is occupational sorting, if women disproportionately work in occupations with lower earning potential due to factors such as preferences, social norms and/or discrimination. Table 2 presents data on primary occupations for respondents, in which we list the top 20 (out of 49) occupations, sorted by their employment shares in the full sample and report shares by gender.[9] Women are heavily represented in agriculture and fishing, with 33.4% of them working as farmers compared to 18.3% of men. Retail and commercial activities, such as being a seller or street vendor, are also disproportionately female-dominated, with 18.3% of women and only 5.2% of men engaged in these roles. Similarly, women are much more likely to be employed in domestic work, which accounts for 7.4% of their employment compared to only 1.2% of men's. These occupations are often characterized by lower earning potential and flat gradients over time, reinforcing income disparities between the genders.

[7] Some of the reported self-employment could be thought of as disguised unemployment, as is typical of labor markets in LMICs (Donovan et al., 2023).

[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 as controls in the regressions below.

[9] The KLPS-5 survey asks the respondent their primary occupation regardless of whether they are currently employed.

In contrast, men dominate occupations such as fishing (6.5% of men, compared to only 0.4% of women), and bicycle and motorbike taxi work is almost exclusively male, with 5.7% of men participating and no women recorded in this occupation in the KLPS sample. Construction-related jobs also exhibit stark gender imbalances, i.e., for masons (5.6% for men versus 0.1% for women) and unskilled construction laborer (3.5% for men versus 0.3% for women). These patterns of occupational sorting by gender could reflect underlying social norms that prescribe which occupations and what kind of work women and men should perform. Such social norms can themselves be an additional contributor to the gender wage gap, as we discuss more in Section 6.

3.2 Raw and "basic" covariate-adjusted gender wage gaps

We estimate the gender wage gap with the following regression model:

$$\ln Y_i = \beta_0 + \beta_1 Female_i + \beta_2 Sch_i + \beta_3 Exp_i + \beta_4 Exp_i^2 + \beta_5 X_i + \gamma_i + \epsilon_i \tag{1}$$

Here, Y_i is the hourly wage of individual i, Sch_i is the years of schooling completed, Exp_i is the years of experience, X_i is a vector of person-level controls that includes different sets of covariates depending on the specification, and γ_i includes 49 occupation indicators. β_1, the coefficient on the indicator for female, thus denotes the estimated gender wage gap, conditional on the other covariates.

Table 3, Panel A presents the raw (unadjusted) gender wage gap for the entire sample (with non-zero earnings/hours) and for subgroups of urban workers, wage workers, and urban wage workers.[10] The raw gender wage gap is quite large: women earn 79 log points less than men on average, which corresponds to a 55 percent difference. Among urban workers, the difference is similar at 75 log points. Restricting the sample to wage workers, however, leads to substantially lower estimates of the gender wage gap: the gap among all wage workers is just 38 log points, and 52 log points among urban wage workers.

Panel B controls for years of education, years of experience (first and second order terms), urban residence, Nairobi residence, working in agriculture, working in self-employment as well as the 49 occupation indicators; these are standard covariates that are similar to those commonly available in household or labor force surveys used to produce existing gender wage gap estimates. With these controls, the gender wage gap narrows to 49 log points in the main estimation sample,

[10]Wage work is defined as being employed and working for pay in the last 30 days. This includes individuals working for pay in the informal sector.

which corresponds to a (still substantial) 39 percent difference. Among urban workers, the adjusted gender wage gap is 42 log points. The difference between wage work and non-wage work is even more pronounced: among wage workers, the adjusted gender wage gap is 29 log points, and 25 log points among urban wage workers. Such results demonstrate that only focusing on wage workers (as is the case in many existing labor force surveys) will likely significantly underestimate the actual gender wage gap in a low- and middle-income context like this one. This highlights a strength of the KLPS data which has followed the same individuals over time gathering detailed labor market data regardless of the respondent's current residential or employment sector.

We also examine the gender gap in hours worked in Table 4.[11] In the full sample, women work 12.2 fewer hours per week than men, while among urban workers the gap is similar at 13.6 hours. Similar to the wage results, the gap in hours worked is noticeably smaller among wage workers (5.2 hours) and urban wage workers (7.3 hours). This pattern also still persists after controlling for education, demographics, and occupation variables: among all workers women work 8.9 fewer hours than men, but 4.5 few hours among wage workers. Again, this highlights the importance of including both wage workers and other workers in the sample to produce a more accurate estimate of the overall gender gap in labor market outcomes.

The data also indicate that urban wage workers in Kenya are a selected group: Appendix Table A.4 shows that they achieve higher levels of education on average (by 1.3 year), as well as in cognitive performance (by 0.22 SD) and analytical task complexity of their jobs (by 0.20 SD), and these differences are statistically significant. Migration also appears to be selective by gender (Table 1, row 4): 41% of women and 48% of men reside in urban areas (despite all individuals starting out in the same rural primary schools), a finding that has also been documented in previous KLPS studies (Hamory et al., 2020, 2021).

3.3 Machine learning prediction of wages and hours worked

The previous section showed that gender remains a significant predictor of wages and hours worked, even after adjusting for individual education, experience, and occupation. It is now well-known that traditional regression models may be mis-specified, imposing linear relationships that may not fully capture complex interactions between explanatory variables. To further assess the

[11] The sample (5330) is larger than the wage sample (4687) because individuals who report working zero hours have their wage as missing, whereas we do include these individuals when estimating effects on hours worked.

role of gender in shaping labor market outcomes, we thus apply machine learning techniques that allow for more flexible and non-linear relationships. Specifically, we use random forest models to predict wages and hours worked, comparing models that estimate outcomes separately by gender to those that do not. We use two sets of explanatory variables (i.e., features): a "basic" set of 55 education, experience, demographics, and occupation variables (the large number is mainly due to the occupation indicators), and the full set of 104 variables that contain education, experience, demographics, occupation, cognition, task indices, personality, preferences, home care, and job stress measures. All exercises use the log wage as the outcome and five-fold cross-validation with a random forest regressor, which seeks to avoid overfitting by creating random subsets of the features and uses them to build smaller trees. (We have also confirmed that the central results are robust to using other common machine learning algorithms, including LASSO and XGBoost, not presented.)

We first predict the log wage for the entire individual sample separately by gender (Appendix Table A.3, Panel A). The random forest algorithm achieves an average test R-squared of 0.22 to 0.23 when using the basic set of features, and an R-squared of 0.25 when using the full set. On average, the predicted log wage for women is 80 log points less than men, which represents a 55% difference. If we remove gender as a predictor (Panel B), the R-squared remains similar, but the predicted gender wage gap shrinks substantially to 29 log points, which is equivalent to women earning 25% 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 30 log points higher for women and 21 log points lower for men), even though a model without accounting for gender still maintains broadly similar overall predictive power. This highlights the importance of gender as a predictor of wages, and suggests that while women and men have some similarities in terms of covariates, the actual gender wage gap is larger than predicted in the absence of gender.

Even though random forest models lack the interpretable coefficients of regression models, SHAP (SHapley Additive exPlanations) provides a way to examine the influence of each feature on the model's predictions, which also reinforces the role of gender.[12] Figure 1 shows the SHAP beeswarm plot for the features with the highest predictive power in predicting log wages. The plot

[12]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.

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 is the most important predictor of wages, with higher SHAP values (pink points) for individuals in urban areas pushing predicted wages upward, with lower values (blue points) for rural residents. Following urban/rural status, gender emerges as the third 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. This highlights the substantial role of gender in wage disparities, even after accounting for occupation, education, job characteristics, and other individual traits. Other influential predictors include years of educational attainment, the job interpersonal index, the job analytical index, executive functioning score, and general cognition score. Consistent with standard intuition, all of those features are positively correlated with predicted log wages.[13]

We also perform the same exercises to predict hours worked. Using the full set of 104 covariates to predict hours worked, the random forest model achieves a test R-squared of 0.26 for females and 0.09 for males, while if we remove gender as a predictor, the test R-squared is 0.21. As with wages, without gender as a predictor, the model does far worse at tailoring its prediction to females and males: the average predicted hours worked becomes 2.98 hours higher for women and 2.54 hours lower for men. Appendix Figure A.3 shows the SHAP beeswarm plot that summarizes the contribution of top features to the prediction of hours worked, where gender emerges as the third most important feature, in this case behind the interpersonal index and self-employment (self-employed workers work substantially more hours); Appendix Figure A.4 shows the SHAP beeswarm plot using only the basic set of features.

4 Gender differences in other personal characteristics

We have demonstrated that there is a sizable gender wage gap that cannot be fully explained by education, demographics, and occupation variables. Can the richer set of additional individual measures gathered in the KLPS help account for it? To answer this question, we begin by examining whether there are gender differences in those novel dimensions: cognitive abilities, differences in

[13] As a robustness exercise, Appendix Figure A.2 shows the SHAP beeswarm plot from a model with only the education, experience, demographics, and occupation variables as predictors, and here gender still emerges as the second most important predictor of wages.

job tasks (even within an occupation), personality measures, economic preferences, and childcare and elder care responsibilities. We also explore the interplay between these factors.

4.1 Gender differences in cognitive performance

Women and men in the sample achieve broadly similar levels of education but educational attainment may not fully capture all relevant aspects of cognitive performance that could matter in the labor market. We thus also assess whether there are gender differences in cognition in this population. Figure 2a shows the average gender difference in terms of the general cognition score (in standard deviations units) as well as for the five cognitive domains that make up the overall score. Men outperform women on the general cognition score by 0.26 standard deviations on average in this population, with notable variation across specific cognitive skills. The largest disparities are observed in executive function and visuospatial skills, where men score 0.34 and 0.29 standard deviations higher, respectively. Memory and language show smaller gaps, with men scoring 0.07 and 0.13 standard deviations higher, while spatial orientation exhibits the smallest difference (at 0.06). These findings suggest that gender differences in cognitive performance in this population are moderate in size, and present (though not uniformly) across various cognitive domains.

In addition to looking at midlife cognitive performance as we do here in KLPS-5, we can also examine differences in cognitive performance at younger ages. The rightmost column in Figure 2a presents the gender difference in the lagged Raven's Matrices test score collected in earlier KLPS rounds about 10-15 years prior to KLPS-5 (see Appendix B.5 for more details). The sample for this measure includes 4,780 individuals (out of 5,338) who were surveyed both earlier and in KLPS-5. This measure reveals the largest gender disparity, with men outperforming women by 0.35 standard deviations on average. As the Raven's Matrices test corresponds to components of the executive functioning domain, the observed difference here is consistent with patterns in KLPS-5 and indicates that there appear to be persistent differences in executive functioning.

These gender differences in average measured cognitive performance raise the possibility that controlling for individual cognitive performance could reduce the gender wage gap.

4.2 Gender differences in job task content

Beyond gender differences in job occupations, the contents of the jobs that men and women engage in could also differ meaningfully even within the same occupation. Figure 2b 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 0.53 SD higher on the analytical task index (indicating greater analytical complexity), 0.64 SD higher on the interpersonal task index, and 0.59 SD higher on the routine/manual task index. This pattern implies that the tasks performed by men tend to be higher in abstract thinking, more intensive in interacting with others, and feature a larger share of repetitive and manual work. The difference in the analytical task index stems from substantial disparities in tasks such as reading long documents (0.41 SD), performing math at work (0.39 SD), and engaging in activities that involve learning new things (0.33 SD) (Appendix Table A.5, Panel A). Among activities that demonstrate and develop interpersonal skills, men are much more likely to respond that they supervise others (0.46 SD), collaborate with coworkers (0.63 SD), and make presentations at work (0.39 SD) (Appendix Table A.5, Panel B). For the components of the routine/manual task index, while men and women are equally likely to report performing repetitive tasks at work, men are more likely to operate heavy machinery at work (0.44 SD), repair or maintain equipment at work (0.30 SD), and find their work physically demanding (0.20 SD) (Appendix Table A.5, Panel C).

These results again suggest that accounting for task-based factors could help account for the observed gender wage gap.

4.3 Gender differences in other dimensions

We close this section with a discussion on the gender differences in other measures in the KLPS data that could potentially affect wages, namely, personality traits, economic preferences, job stress, and caregiving responsibilities. The fact that there are meaningful differences by gender along several of these dimensions suggests that they could also potentially account for some of the observed gender wage gap.

Existing research in psychology and labor economics indicates that the Big 5 personality traits are associated with individual labor market outcomes: they tend to find a positive association between earnings and individual openness, conscientiousness, and extraversion, and a negative association between earnings and agreeableness and neuroticism (Alderotti et al., 2023). In light of those findings, we examine gender difference in these characteristics (Figure 2c). In the KLPS-5 data, men are more likely than women to demonstrate extraversion, agreeableness, conscientiousness, and openness, while women are more likely to have higher levels of 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.[14] Women are also more likely to be ambiguity averse. There are no statistically significant gender differences in terms of other economic preferences including risk aversion, patience, or reciprocity. In terms of other attitude measures, there are no meaningful mean gender differences in life satisfaction (diff = 0.03 SD), but women are 0.08 SD more likely than men to believe that it is impossible to escape a predetermined fate when presented with that statement. We also do not find gender differences in the stress level experienced on the job between men and women (diff = 0.01 SD, see Figure 2d). While there are no gender differences in the number of older adults over 65 years old living in the household (diff = 0.00), women on average live in households with 0.32 more children under age 18 (p-value < 0.001), suggesting that they face a greater childcare burden than men on average.

5 Regression Analyses

5.1 Main results: wages

Having established above that many individual characteristics both differ on average by gender and are associated with wages, we next employ a regression framework to estimate the gender wage gap once we account for the full set of observable characteristics available in KLPS. Specifically, we implement the same regression model as in Equation (1) with varying sets of covariates.

As noted, the raw, unadjusted gender wage gap is 79 log points, which is equivalent to a 55 percent difference (Table 5, Column 1). We next add standard controls for years of educational attainment, experience, location, and any types of agricultural work or self-employment, and their inclusion reduces the gender wage gap to 59 log points (Column 2), suggesting that some of the gap is explained by commonly observable demographic characteristics. Urban residence has a positive and significant association with log wages (coefficient = 0.58) and being in Nairobi also has a positive effect, both of which highlight the urban wage premium, although it remains likely that some of this gap is due to selective individual residential sorting. Column 3 adds 49 occupation fixed effects but excludes the individual controls from Column 2, to isolate the role that occupation may play. When both occupation fixed effects and the standard controls are included the coefficient on the gender indicator falls to 49 log points (Column 4).

[14] To capture this, the survey asked respondents how much they would donate to charity if they unexpectedly received money today. We ask the same question again twice, first specifying that it was a charity helping people in the respondent's ancestral home area, and next specifying that it was helping people in other parts of Kenya, providing the in-group versus out-group contrast.

After introducing the more novel KLPS controls for cognitive ability, the gender wage gap slightly falls to 48 log points (Column 5). The general cognition score is positively associated with wages: a one standard deviation increase in the general cognition score increases wages by 18 log points. Once the general cognition score is included, the coefficient on education drops substantially, suggesting that some of the commonly documented association between education and wages could be a proxy for underlying individual cognitive ability.

We next introduce the three job task indices—in the dimensions of abstract analytical tasks, interpersonal tasks, and routine/manual tasks—as controls to account for task differences within the same occupation (Column 6). There is a positive wage premium of 10 log points associated with each standard deviation increase in analytical tasks and a 11 log point premium associated with the analogous increase in interpersonal tasks. The implication is that, within the same occupation, workers whose on-the-job activities involve more abstract thinking and interacting with others earn substantially higher wages. The gender wage gap falls further with their inclusion to 42 log points, suggesting that part of the reason why men earn more than women may be explained by the differences in the job tasks that men and women perform.

Finally, we include the rest of the novel measures from KLPS, namely, the job stress index, economic preferences and personality variables, cognitive performance overall in KLPS-5 (the general score) and in five domains (executive functioning, visuospatial, language, memory, and orientation), the lagged Raven's score from an earlier KLPS round, as well as the the number of children and the number of older adults in the household.[15] The coefficient on the general cognitive score predictably becomes smaller and less statistically significant with the inclusion of the domain specific scores (that together make up the general score). In this specification, the gender wage gap is now at 39 log points. Note that high-stress jobs are associated with lower wages, suggesting a positive correlation between wages and at least some job 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). Having more children in the household is associated with a small but statistically significant reduction in wages, with a coefficient of 5 log points per child. Speculatively, this potentially points to a childcare penalty on wages although there are many potential confounders.

[15]We report coefficients on economic preferences, Big Five personality measures, and lagged Raven's matrices score in Appendix Table A.6.

Comparing Column 7 with Column 4 indicates that the novel cognitive measures, task indices, personality measures, economic preferences and other covariates together explain 20% of the gender wage gap that was left unexplained by standard occupation, education and demographic variables. This implies that a large share of the gender wage gap remains unexplained even after accounting for the unusually rich set of individual and job characteristics collected in the KLPS data.

5.2 Main results: hours worked

Beyond wage differences, women in the KLPS-5 data also work substantially fewer hours than men, as noted above. As boosting women's participation in the labor market has become a more prominent policy focus of many governments and development organizations in recent years, examining the gender gap in hours worked is of intellectual and practical interest in its own right, and we estimate it using the same regression model specified in Equation (1), where the dependent variable is weekly hours worked (Table 6). As noted above, the sample size increases as those who do not report working any hours worked in agriculture, self-employment, and wage employment are coded as working zero hours.

Without any covariate adjustment, women work 12.2 fewer hours per week than men (Column 1), and including standard demographic and education controls narrows the gap to 10.7 hours (Column 2). As expected, more educated individuals work more on average, and those engaging in self-employment also work significantly more hours. The gender gap in hours falls to 8.87 hours with the inclusion of the occupation fixed effects together with the standard control (Column 4).

The introduction of the KLPS general cognition score reduces the gender gap in hours worked only slightly to 8.85 hours (Column 5). Notably, this additional control does not significantly affect the coefficients on the other variables. Considering the job task measures, the analytical and interpersonal task indices are positively associated with hours worked (coefficients of 1.29 and 3.16, respectively), suggesting that workers performing more cognitively demanding or interpersonal tasks (like managing others) tend to work longer hours, and their inclusion further reduces the gender gap to 7.4 hours (Column 6). Finally, including the personality, economic preferences, and job stress measures as covariates, along with the number of children in the household, the number of older adults in the household, the domain-specific cognitive scores and the lagged Raven's score leads to a slight drop in the gender wage gap to 6.98 hours (Column 7).

Again comparing Column 7 with Column 4, we observe that the additional set of novel KLPS control variables explains only approximately 20% of the gender gap in hours worked that remains unexplained by standard occupation, education, and demographic measures, similar to the share they explain of the gender wage gap, as noted above.

5.3 Double/Debiased Machine Learning results

The set of variables that could be correlated with wages and gender in the dataset is quite large (at 104 in all) and many of them are continuous, so the standard linear regression model may be mis-specified. We use a method, Double/Debiased Machine Learning (DML), from the causal machine learning literature (Chernozhukov et al., 2016, 2018) to account for potential non-linearities. This method is particularly useful in a context where many factors may be correlated with both the "treatment" (here, being a woman) and the outcome (wages or hours worked). DML builds on earlier causal machine learning methods, such as post-double LASSO, which have been used in several gender wage gap studies in rich 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 (Chernozhukov et al., 2016).

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

$$Y = \theta(X) \cdot T + g(X, W) + \epsilon$$

$$T = f(X, W) + \eta$$

$$\mathbb{E}[\epsilon|X, W] = 0,\ \mathbb{E}[\eta|X, W] = 0,\ \mathbb{E}[\eta \cdot \epsilon|X, W] = 0$$

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 using arbitrary non-parametric Machine Learning methods (Econ ML User EconML User Guide 2023). Specifically, we can 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

$$gX, W) = \mathbb{E}[Y|X, W]$$

$$f(X, W) = \mathbb{E}[T|X, W]$$

and then compute the residuals

$$\tilde{Y} = Y - g(X, W)$$

$$\tilde{T} = T - f(X, W) = \eta$$

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 choose the random forest model for both its established accuracy and ability to deal with relatively small sample sizes and high-dimensional feature spaces without overfitting (Breiman, 2001; Biau and Scornet, 2016). We start by imposing homogeneous treatment effects, i.e., $\theta(X) = \theta$, but in the next section include interaction terms to study how gender gaps vary across segments of the sample.

Table 7 presents 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 54 log points, versus 49 log points estimated by linear regression (Table 5, column 4). Accounting for the full set of covariates, the gender wage gap estimated by DML is 40 log points, which is nearly identical to the 39 log points estimated by linear regression (Table 5, column 7). Cognition, task indices, personality, and preferences explain 26% of the gender wage gap that was left unexplained by occupation, education and demographics variables. As a robustness check, columns 3 and 4 account for class imbalance (i.e. differences in the number of male and female respondents within the sample) by over-sampling men to achieve gender balance in the analysis sample; the results are similar to the main estimation sample.

The bottom half of the table presents the DML results in a model where the outcome variable is weekly hours worked. The gender gap in hours worked is 8.19 hours with the basic set of covariates and 6.37 hours with the full set of covariates, versus 8.87 and 6.98 hours, respectively, estimated via linear regression (Table 6, columns 4 and 7). 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 main estimation sample.

Taken together, the DML results provide us with more confidence, first, that cognition, job tasks, personality, and economic preference measures do help explain part of the gender gap in wages and hours worked, but second, that there remain substantial differences between women and men in both their wages and hours worked even after accounting for the full set of KLPS covariates.

5.4 Oaxaca-Blinder Decomposition

The linear regression and DML methods used so far examine the extent to which observable factors can explain the gender gap in wages and hours worked. Another approach commonly applied in labor economics is the Oaxaca-Blinder decomposition, which estimates the wage equation separately for women and men and then decomposes the raw unadjusted gender gap into one part explained by observed variables and another unexplained part (Blinder (1973), Oaxaca (1973)). Specifically, we run the following regressions separately for groups $g = \{m, f\}$, where Y_g denotes log wages and X_g is a vector of observed covariates:

$$Y_m = X_m B_m + u_m$$

$$Y_f = X_f B_f + u_f$$

where B_g is a vector of coefficients. Let b_m and b_f denote, respectively, the OLS estimates of B_m and B_f, and mean values are denoted with a bar. Then, since OLS with a constant term produces residuals with a zero mean, this yields:

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

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 corresponds to an experiment where a woman's actual wage is compared with her predicted wage from the male equation (see Blau and Kahn (2017) for a discussion). 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. The Oaxaca-Blinder decomposition thus structurally assumes that observables affect male and female wages differently, whereas the pooled-sample regression and DML approach can be thought of in a potential outcomes framework where gender is as good as randomly assigned and observables need not impact outcomes differently by gender. As a result, the output from the two approaches will likely differ.

The top panel of Table 8 presents the decomposition results for the gender wage gap. The part that is explained by the basic set of covariates (education, experience, demographics, occupation) is 32 log points, or 41% of the raw gender wage gap; the unexplained part is 47 log points, or 59% of the raw gender wage gap. Using the full set of controls, 51% (40.5 log points) of the raw gender wage gap can be explained by covariates, leaving 49% (38.7 log points) of the gap unexplained. These estimates are similar to the pooled-sample regression results that show a 49 and 39 log point gender wage gap unexplained by the basic and full set of controls, as shown in Table 5.

The bottom panel of Table 8 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 1.5 hours (12% of the raw gender hours gap); the unexplained part is 10.7 hours(88% of the raw gap). Using the full set of controls, 32% (3.9 hours) of the raw gender hours gap can be explained by covariates, leaving 68% (8.4 hours) of the gap unexplained. These estimates are somewhat different from the ones obtained from pooled-sample regressions shown in Table 6, but still broadly similar.[16]

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 indicator variable. Specifically, we interact the gender indicator with the following variables along which the gender wage gap could differ: urban/rural residence, living in Nairobi, any agricultural work, any self-employment, educational attainment, cognition (general cognition

[16] Appendix Tables A.7 present the Oaxaca-Blinder decomposition results using females as the reference group, which are qualitatively similar but differ somewhat from those presented above; it is known that estimated components may change depending on the choice of the reference group (Oaxaca and Ransom, 1994).

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 9 presents 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 include the full set of controls. Some notable patterns emerge. In both the linear regression and the DML models, the interaction term on the analytical index is negative (-0.14 and -0.13), whereas the interaction term on the interpersonal index is positive (0.17 and 0.04). These magnitudes are substantively meaningful given that the gender wage gap (adjusted for all covariates) is 0.39. These results may indicate that the gender wage gap is larger among individuals involved in more complex analytical tasks but smaller among individuals involved in more sophisticated interpersonal tasks. One speculative 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 potential concern that there may be limited common support between female and male workers on these task indices, in Appendix Figure A.5 we display the kernel density plots of each one of the three task indices by gender, and show that there is substantial overlap between men and women.

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 in LMICs (Jayachandran, 2021) and Kenya specifically (Delecourt et al., 2023). The gender gap is also 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 not statistically significant, except for language and visuospatial skills, which exhibit a small but significant negative interaction in some models.

Furthermore, we examine how the gender gap in labor supply (hours worked) varies across demographic and human capital measures following the same analytical approach (Appendix Table A.8). Most coefficients are either close to zero or not statistically significant. One notable exception is the finding that, among individuals whose jobs involve 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 2.6-4.6 smaller gap in hours worked, which is a meaningful magnitude given that the adjusted gender gap in hours worked is 6.98 hours. Thus while the gender wage gap remains large among more individuals whose jobs feature analytically complex tasks, the hours worked of such women is relatively similar to men.

5.6 Gender Differences in Cognition and Job Task Content by Occupation

The analysis indicates that occupation, cognitive ability, and job task complexity all hold some explanatory power for the gender wage gap in Kenya. Combining these three aspects, we next examine how women and men of different cognitive abilities sort into occupations with different analytical and interpersonal task complexity. One of the unique features of the KLPS data is the inclusion of a lagged Raven's score, which captures a dimension of individual cognitive ability from roughly 10-15 years before KLPS-5 when respondents were young adults. This measure, alongside current cognitive assessments, enables us to explore the interplay between occupational sorting, cognitive ability and job task content skills—specifically, analytical and interpersonal skills. In particular, we examine how both current and past measures of cognitive ability (the current and lagged Raven's score) predict current occupational choices, and examine the patterns by gender.

We first construct a ranking based on the standardized Raven's scores. Standardization was performed by subtracting the round-specific sample mean from each respondent's raw score and dividing the difference by the corresponding standard deviation. We find that individuals with higher lagged cognitive scores are more likely to have jobs requiring robust analytical and interpersonal skills (including many higher-skill professional occupations), whereas the selection into routine/manual jobs appears less influenced by cognitive ability. More notably, the gradient linking cognitive ability to job task complexity is steeper for males than for females, suggesting that higher-ability males are better able (or more willing) to enter occupations that are more intensive in analytical and interpersonal tasks. Appendix Figures A.6 and A.7 present these findings. Among males, there is a strong positive association between the analytical skills in the individual's current job and the current Raven's score, with a coefficient of 1.4 (SE = 0.3) (Appendix Figures A.6, upper left panel), and a similar identical relationship with the lagged Raven's score (1.1, SE = 0.23, Appendix Figure A.7, upper left panel).[17] In contrast, among females, the current Raven's score shows a smaller association with current job analytical skills, with a coefficient of 0.44 (SE = 0.27) though the lagged score has a somewhat larger coefficient of 0.88 (SE = 0.21). When we

[17] Note that there is a correlation of 0.37 between the current (KLPS-5) and lagged Raven's Matrices score.

use the current Raven's score, the differences in the slope between men and female is statistically significant at the 10 percent level (p-value = 0.099), though the difference in slope between men and women when we use the lagged Raven's score is less statistically significant (p-value = 0.135). A similar pattern is observed for interpersonal skills, in which there is a much stronger correlation between both the current and lagged Raven's score measure among males (with coefficients of 1.16 (SE = 0.20) and 0.94 (SE = 0.17), respectively) than among females (0.46 (SE = 0.27) and 0.80 (SE = 0.23)), though the difference is not statistically significant. For routine/manual skills, the associations with cognitive scores are weaker for both genders.

Taken together, these results show that cognitively high-ability women do not sort into jobs featuring higher-complexity tasks as much as high-ability men do. This seems likely to hinder women's ability to earn higher wages in the labor market, given that many of the occupation with the highest analytical and interpersonal task complexity in our data (e.g., teachers, police offers, salaried professionals, etc.) are also the ones that are the best paid. Understanding the causes of this differential job sorting by gender remains an important area for future investigation.

5.7 Robustness

This subsection present several robustness checks. First, the analysis in Appendix Table A.9 only includes urban wage workers (as is common in labor force surveys), and shows that the gender wage gap remains substantial: on average, among urban workers who engage in wage employment, women earn 52 log points less than men, 25 log points less after adjusting for basic controls, and 24 log points less after adjusting for the full set of controls. The sample is substantially smaller in this case—at 1,213 individuals—because most individuals do not engage in wage employment as their primary occupation. In this case, the cognition, task indices, personality, and preferences measures together explain only about 4% of the gender wage gap that is unexplained by education, demographics, and occupation variables.

We next examine log total monthly earnings – rather than the hourly wage – as the dependent variable (in Appendix Table A.10), and the gender gap remains persistent across specifications: on average, in terms of total earnings, women earn 112 log points less than men, 75 log points less after adjusting for basic controls, and 59 log points less after adjusting for the full set of controls.

The next set of robustness checks addresses outliers in various ways. The analysis in Appendix Table A.11 conducts the same main analysis on the hourly log wage (as in Table 5) but winsorizes the dependent variable at the 95th percentile (which is equivalent to an hourly wage of 330.3 Kenyan Shillings or USD 2.40), and Appendix Table A.12 winsorizes the hours worked variable at the 95th percentile. For both, the results are similar to the main tables.

Another concern is that the gender gap in wages could in part be explained by the gender gap in labor supply, namely, the fact that women working fewer hours might lead them to build up fewer on the job skills or be less appreciated. The analysis in Appendix Table A.13 examines this issue by including weekly hours worked as an additional covariate and finds that the gender wage gap is similar to the main specification, at 58 vs 49 log points with the basic set of controls and at 46 vs 39 log points with the full set of controls. The additional control variables again explain 21% of the gender wage gap unexplained by education, experience, demographics, and occupation, nearly identical to the main specification.

Because not all individuals in the KLPS-5 sample were surveyed for KLPS-4, in the main regressions we had imputed the missing personality and economic preferences covariates with the sample mean and introduce missing indicators, and we similarly imputed missing job task indices with the same approach (for the N=292 individuals that did not provide valid responses to the full set of job task content questions). Appendix Table A.14 runs the wage regressions on the observations without any missing covariates: the estimated gender gap for both basic and full covariates are within 0.01 log points of the main results. Appendix Table A.15 examines weekly hours worked using this approach and again yields similar results.

We further confirm the robustness of the wage results by estimating the gender wage gap on the subsample of 2,140 individuals who were interviewed in both the fourth and fifth rounds of KLPS with a non-zero wage in both rounds (KLPS-4 measurement of labor market outcomes took place between 2017 and 2019). Similar to the earnings and hours module in KLPS-5, respondents were asked about their earnings and hours worked in agriculture, self employment, and wage employment. The analysis in Appendix Table A.16 first runs the main wage regression using KLPS-5 data on this subsample of individuals for comparability; the estimated gender gap is slightly smaller than but similar to that estimated with the main estimation sample. In Appendix Table A.17 we estimate the gender wage gap as measured in KLPS-4.[18] The raw gender wage gap in KLPS-4 data (75

[18]We use data on experience, urban, Nairobi, agriculture, and self-employment from KLPS-4. We use occupation, cognition, task indices and job stress measured in KLPS-5 because those questions were either more detailed in the fifth round (the first two) or only asked in the fifth round (the last two). Personality and economic preference

log points) is similar to that in KLPS-5, and the adjusted gender wage gaps controlling for basic covariates (56 log points) and all covariates (53 log points) are somewhat larger than those in KLPS-5. KLPS-4 also asks about the number of hours individuals spend on household chores and childcare per week (which were not collected in the KLPS-5 visit). Interestingly, when we add these two variables as controls in Column 8 of Appendix Table A.17, the adjusted gender gap narrows meaningfully from 53 to 38 log points. This suggests that uneven burdens of household work and childcare might be depressing women's wages in the Kenyan labor market. In fact, women in KLPS-4 spend on average 15 hours on childcare and 28 hours on household chores per week, substantially higher than the 6 hours and 10 hours, respectively, spent by men. That said, the 38 log point gap even in this case remains large and not qualitatively different from the main findings.

6 Discussion and Implications

The analysis has shown that the gender wage gap remains economically large and statistically significant across many different analytical specifications (including both linear regression and machine learning models), including with the inclusion of detailed measures of cognition, job task context, economic preferences and personality, in addition to more standard socio-demographic and occupation controls. This suggests that other unmeasured factors must be affecting the labor market outcomes of women and men differently.

A leading explanation for the remaining gap is the role of gender-related social norms and discrimination, and several recent studies in development economics have examined their importance in driving labor outcomes across genders (e.g., Bursztyn et al. 2020; McKelway 2019; Buchmann et al. 2024; Kala and McKelway 2025; see Jayachandran (2021) for a review). Other work has explored different factors that may influence female labor supply, including the interrelated roles of job safety, commuting costs, work location and childcare (Siddique, 2022; Becerra and Guerra, 2023; Cheema et al., 2023; Chen et al., 2024; Delecourt et al., 2023; Jalota and Ho, 2024; Bjorvatn et al., 2025).

The KLPS-5 survey contains several questions on discrimination and social norms and attitudes, including how individuals' experience of discrimination differs by gender (Table 10). In particular, the survey asks individuals if they are treated with less courtesy than other people are, if they are

variables were already obtained from KLPS-4 data, though those were collected in a separate visit between 2018 and 2021 (the KLPS-4 I Module).

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, as well as the frequency of this treatment.[19] Panel A presents both a difference in means between men and women (the third column) and a full regression model including all covariates used in the wage analysis above (in the fourth column). Throughout, women do not report that they have experienced much discrimination in these ways. Perhaps puzzlingly, women are even somewhat less likely than men to say that they receive poorer service at restaurants or stores, or have been threatened or harassed.

The KLPS-5 survey also explicitly asks respondents about the main reason(s) they believe that these experiences of discrimination have occurred, and Panel B presents the share of women and men who cite each reason. Socioeconomic status emerges as the most common reported cause, with far more predictive power than gender or other demographic characteristics. For instance, "My education or income level" emerges as the most cited reason, with 48% of men and 50% of women who experienced discrimination choosing that option; the perhaps closely related social group or family background comes next with 17% of men and 19% of women choosing that option. A further 12% of men and 11% of women believe aspects of their physical appearance is the main reason for discrimination. Only 2% of women and less than 1% of men chose gender as the main reason for discrimination, and if we include "other" (non-main) reasons, a perhaps surprisingly low 3% of women (and under 1% of men) cite their gender. That said, an important caveat is that the questions were quite general and none explicitly focused on discrimination experiences in the labor market. There may also be some concern about biased reporting of discrimination experiences, although we see that as a second-order concern in this context, given the general openness of respondents to discuss issues of gender and other social categories.

The KLPS-4 data also allows us to examine individuals' stated attitudes toward gender norms (Table 11, for 4,597 out of 5,338 individuals in the main analysis sample). The vast majority of both women and men agree with statements articulating egalitarian gender norms, with little difference by gender: to illustrate, 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

[19]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, and 0 if they report less than once a year or never.

chores if the wife is working outside the home; 92% of women and men believe women can be good politicians and should be encouraged to stand in elections; 92% of women and 91% of men believe that women should have equal rights and receive the same treatment as men in Kenya; and both 90% of men and women do not believe the husband has a right to beat his wife if she misbehaves. The only partial exception is the responses to the statement "Important decisions in the family should be made by men": 37% of men and 42% of women disagree with that statement, although even that difference is muted.[20]

This data indicates that the vast majority of men and women state that they are comfortable with women performing 'masculine-coded' jobs, becoming politicians, and receiving equal legal rights and treatment in society as men. Yet recall that relatively small shares of women actually work in these types of male-dominated positions in our Kenyan sample (Table 2), and that high-ability women do not sort into high-task complexity, high-earning occupations as much as men do (Figures A.6 and A.7). One immediate explanation for this pattern could be the role of social desirability bias in driving the survey reports regarding egalitarian norms. Yet this in itself would seem to rule out the simplest stories of highly restrictive social norms for women limiting their labor market opportunities in Kenya, as widespread social desirability towards supporting equality for women also appears indicative of emerging gender equality norms.

7 Conclusion

The current study highlights the empirical existence of a large gender wage gap among a sample of prime working-age adults in the Kenya Life Panel Survey (KLPS), with a raw gender wage gap of 79 log points. A key contribution of the analysis is the presentation of evidence showing that the gender wage gap persists and remains large even when controlling for an unusually extensive set of individual human capital and job-related characteristics, and across multiple econometric and statistical approaches. The innovative individual and job measures – including individual cognitive performance, job task complexity, economic preferences, personality traits, and caregiving responsibilities, which are typically unobserved in most datasets – help explain only 20% of the

[20]In Appendix Table A.18, we examine the responses to these questions separately by urban or rural residence. We find that urban respondents are slightly more likely to agree with the statements in favor of egalitarian gender norms. Appendix Table A.19 presents the responses to same set of questions administered in the KLPS 3 round (2011-2014) and patterns are similar, indicating that the KLPS respondents have expressed largely egalitarian gender views for most of their adult lives.

unexplained gap (beyond that captured by standard socio-demographic and occupational controls), and a substantial 39 log point difference remains. This central finding suggests that factors beyond readily observable individual and job characteristics continue to shape labor market disparities between men and women in Kenya.

Yet despite these pronounced wage gaps, self-reported measures of experiences with everyday discrimination and gender attitudes are not strongly biased against women in this population, and instead point towards norms promoting gender equality. Taken together, then, these findings present a puzzle: if observed characteristics cannot account for the gender gap in wages (and hours worked), and we take the survey evidence to indicate that discrimination against women does not appear to be a leading explanation for it, then what is driving these large and persistent differences? Heath et al. (2024) review existing research and find that policies increasing childcare availability, improving female empowerment within households, and increasing job flexibility can increase female labor force participation, though increased participation may not address disparities in earnings. Future research could usefully explore alternative explanations, including gender differences in structural barriers in the labor market, discrimination in firm hiring practices, access to professional networks, the shaping of occupational preferences by gender, and the role of the informal labor sector. Better understanding of the underlying drivers of these gender differences in labor market outcomes may help facilitate the development of more effective public policy or firm interventions to enhance labor market opportunities for women in Kenya and beyond.

References

Acemoglu, Daron and David Autor, "Skills, Tasks and Technologies: Implications for Employment and Earnings," in David Card and Orley Ashenfelter, eds., *Handbook of Labor Economics*, Vol. 4, Elsevier, January 2011, pp. 1043–1171.

Alarakhia, Mehjabeen, 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.

Alderotti, Giammarco, Chiara Rapallini, and Silvio Traverso, "The Big Five personality traits and earnings: A meta-analysis," *Journal of Economic Psychology*, January 2023, *94*, 102570.

Appleton, Simon, John Hoddinott, and Pramila Krishnan, "The Gender Wage Gap in Three African Countries," *Economic Development and Cultural Change*, January 1999, *47* (2), 289–312.

Autor, David H. and Michael J. Handel, "Putting Tasks to the Test: Human Capital, Job Tasks, and Wages," *Journal of Labor Economics*, 2013, *31* (2), S59–S96.

Bach, Philipp, Victor Chernozhukov, and Martin Spindler, "Heterogeneity in the US gender wage gap," *Journal of the Royal Statistical Society Series A: Statistics in Society*, January 2024, *187* (1), 209–230.

Becerra, Oscar and José-Alberto Guerra, "Personal safety first: Do workers value safer jobs?," *Journal of Economic Behavior & Organization*, August 2023, *212*, 996–1016.

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

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen, "Inference on Treatment Effects after Selection among High-Dimensional Controls," *The Review of Economic Studies*, April 2014, *81* (2), 608–650.

Biau, Gérard and Erwan Scornet, "A random forest guided tour," *TEST*, June 2016, *25* (2), 197–227.

Bjorvatn, Kjetil, Denise Ferris, Selim Gulesci, Arne Nasgowitz, Vincent Somville, and Lore Vandewalle, "Childcare, Labor Supply, and Business Development: Experimental Evidence from Uganda," *American Economic Journal: Applied Economics*, April 2025, *17* (2), 75–101.

Blau, Francine D. and Lawrence M. Kahn, "The Gender Wage Gap: Extent, Trends, and Explanations," *Journal of Economic Literature*, September 2017, *55* (3), 789–865.

Blinder, Alan S., "Wage Discrimination: Reduced Form and Structural Estimates," *The Journal of Human Resources*, 1973, *8* (4), 436–455.

Boissiere, M., J. B. Knight, and R. H. Sabot, "Earnings, Schooling, Ability, and Cognitive Skills," *The American Economic Review*, 1985, *75* (5), 1016–1030.

Bonaccolto-Töpfer, Marina and Stephanie Briel, "The gender pay gap revisited: Does machine learning offer new insights?," *Labour Economics*, October 2022, *78*, 102223.

Breiman, Leo, "Random Forests," *Machine Learning*, October 2001, *45* (1), 5–32.

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

Bursztyn, Leonardo, Alessandra L. González, and David Yanagizawa-Drott, "Misperceived Social Norms: Women Working Outside the Home in Saudi Arabia," *American Economic Review*, October 2020, *110* (10), 2997–3029.

Böheim, René and Philipp Stöllinger, "Decomposition of the gender wage gap using the LASSO estimator," *Applied Economics Letters*, June 2021, *28* (10), 817–828.

Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline, "Firms and Labor Market Inequality: Evidence and Some Theory," *Journal of Labor Economics*, January 2018, *36* (S1), S13–S70.

Cawley, John, James Heckman, and Edward Vytlacil, "Three observations on wages and measured cognitive ability," *Labour Economics*, September 2001, *8* (4), 419–442.

Cheema, Ali, Asim I. Khwaja, Farooq Naseer, and Jacob N. Shapiro, "Glass walls: Experimental evidence on access constraints faced by women," 2023.

Chen, Yutong, 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.

Chernozhukov, Victor, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins, "Double/Debiased Machine Learning for Treatment and Causal Parameters," July 2016.

_ **, 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.

Delecourt, Solène, Anne Fitzpatrick, Layna Lowe, and Anya Marchenko, "Firm location, childcare, and the gender profit gap," Technical Report 2023.

Deming, David J., "The Growing Importance of Social Skills in the Labor Market*," *The Quarterly Journal of Economics*, November 2017, *132* (4), 1593–1640.

Dicarlo, Emanuele, 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.

Donovan, Kevin, Will Jianyu Lu, and Todd Schoellman, "Labor Market Dynamics and Development," *The Quarterly Journal of Economics*, November 2023, *138* (4), 2287–2325.

EconML User Guide, "Orthogonal/Double Machine Learning," 2023.

Edin, Per-Anders, Peter Fredriksson, Martin Nybom, and Björn Öckert, "The Rising Return to Noncognitive Skill," *American Economic Journal: Applied Economics*, April 2022, *14* (2), 78–100.

Eurostat, "Gender pay gap statistics - Statistics Explained," 2022.

Flinn, Christopher J., Petra E. Todd, and Weilong Zhang, "Labor Market Returns to Personality: A Job Search Approach to Understanding Gender Gaps," *Journal of Political Economy*, November 2024, pp. 000–000.

Goldin, Claudia, "A Grand Gender Convergence: Its Last Chapter," *American Economic Review*, April 2014, *104* (4), 1091–1119.

Gross, Alden L., 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*, October 2023, *4* (10), e573–e583.

Gross, Alden L, Madeline Duhon, Eric Ochieng, Jean N Ikanga, William H Dow, Jinkook Lee, Michael W Walker, Michelle Layvant, Anthony Ngugi, Joshua R Ehrlich, and Edward A Miguel, "Midlife cognitive testing in Africa: validity of the Harmonised Cognitive Assessment Protocol in the Kenya Life Panel Survey," *BMJ Open*, August 2025, *15* (8), e096619.

Hamory, Joan, Edward Miguel, Michael Walker, Michael Kremer, and Sarah Baird, "Twenty-year economic impacts of deworming," *Proceedings of the National Academy of Sciences*, April 2021, *118* (14), e2023185118.

_ **, Marieke Kleemans, Nicholas Y Li, and Edward Miguel**, "Reevaluating Agricultural Productivity Gaps with Longitudinal Microdata," *Journal of the European Economic Association*, nov 2020, *19* (3), 1522–1555.

Hanushek, Eric A. and Ludger Woessmann, "The Role of Cognitive Skills in Economic Development," *Journal of Economic Literature*, September 2008, *46* (3), 607–668.

Heath, Rachel, Arielle Bernhardt, Girija Borker, Anne Fitzpatrick, Anthony Keats, Madeline McKelway, Andreas Menzel, Teresa Molina, and Garima Sharma, "Female Labour Force Participation," *VoxDevLit*, February 2024, *11.*

Heckman, James J., Jora Stixrud, and Sergio Urzua, "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior," *Journal of Labor Economics*, July 2006, *24* (3), 411–482.

Hermo, Santiago, 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*, November 2022, *137* (4), 2309–2361.

Jalota, Suhani and Lisa Ho, "What Works For Her? How Work-from-Home Jobs Affect Female Labor Force Participation in Urban India," January 2024.

Jayachandran, Seema, "Social Norms as a Barrier to Women's Employment in Developing Countries," *IMF Economic Review*, September 2021, *69* (3), 576–595.

Kala, Namrata and Madeline McKelway, "The Power of Persuasion: Causal Effects of Household Communication on Women's Employment," 2025. Working paper.

Karasek, R., 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*, October 1998, *3* (4), 322–355.

Kenya National Bureau of Statistics, "The Kenya Poverty Report 2022 Based on the 2022 Kenya Continuous Household Survey," Technical Report 2024.

Klasen, Stephan, "What Explains Uneven Female Labor Force Participation Levels and Trends in Developing Countries?," *The World Bank Research Observer*, August 2019, *34* (2), 161–197.

Kremer, Michael, Edward Miguel, and Rebecca Thornton, "Incentives to learn," *Review of Economics and Statistics*, 2009, *91* (1), 437– 456.

Lamadon, Thibaut, Jeremy Lise, Costas Meghir, and Jean-Marc Robin, "Labor Market Matching, Wages, and Amenities," July 2024.

Lin, Dajun, Randall Lutter, and Christopher J. Ruhm, "Cognitive performance and labour market outcomes," *Labour Economics*, April 2018, *51*, 121–135.

McKelway, Madeline, "Experimental evidence on the effects of women's employment," *Working Paper*, 2019.

Miguel, Edward and Michael Kremer, "Worms: identifying impacts on education and health in the presence of treatment externalities," *Econometrica*, 2004, *72* (1), 159–217.

Molnar, Christoph, "Interpreting Machine Learning Models With SHAP: A Guide With Python Examples And Theory On Shapley Values," 2023.

Nopo, Hugo, Nancy Daza, and Johanna Ramos, "Gender Earnings Gaps in the World," *SSRN Electronic Journal*, 2011.

Nordman, Christophe J., Anne-Sophie Robilliard, and François Roubaud, "Gender and ethnic earnings gaps in seven West African cities," *Labour Economics*, December 2011, *18*, S132–S145.

Nordman, Christophe Jalil, 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.

Oaxaca, Ronald, "Male-Female Wage Differentials in Urban Labor Markets," *International Economic Review*, 1973, *14* (3), 693–709.

Oaxaca, Ronald L and Michael R Ransom, "On discrimination and the decomposition of wage differentials," *Journal of Econometrics*, March 1994, *61* (1), 5–21.

Ortiz-Ospina, Esteban, Sandra Tzvetkova, and Max Roser, "Women's Employment," *Our World in Data*, March 2024.

Rica, Sara De La, Lucas Gortazar, and Piotr Lewandowski, "Job Tasks and Wages in Developed Countries: Evidence from PIAAC," *Labour Economics*, August 2020, *65*, 101845.

Siddique, Zahra, "Media-Reported Violence and Female Labor Supply," *Economic Development and Cultural Change*, July 2022, *70* (4), 1337–1365.

Sockin, Jason, "Show Me the Amenity: Are Higher-Paying Firms Better All Around?," November 2021.

Sonnega, Amanda, Jessica D Faul, Mary Beth Ofstedal, Kenneth M Langa, John WR Phillips, and David R Weir, "Cohort Profile: the Health and Retirement Study (HRS)," *International Journal of Epidemiology*, April 2014, *43* (2), 576–585.

U.S. Department of Labor, "What You Need to Know About the Gender Wage Gap | U.S. Department of Labor Blog," 2024.

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

Verick, Sher, "Female labor force participation in developing countries," *IZA World of Labor*, 2014.

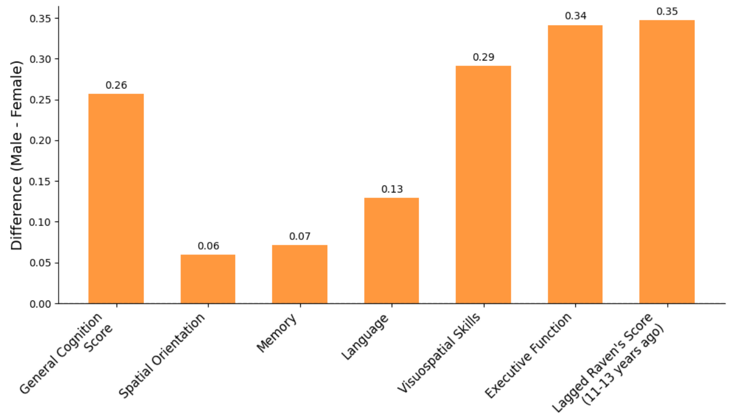
Weidmann, Ben and David J. Deming, "Team Players: How Social Skills Improve Team Performance," *Econometrica*, 2021, *89* (6), 2637–2657.

World Bank, "World Bank Gender Data Portal," 2025.

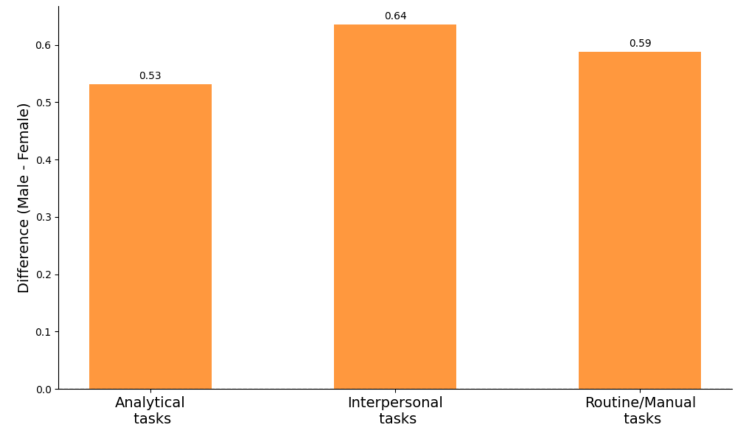
Figures and Tables

Figure 1: SHAP summary plot predicting log wages

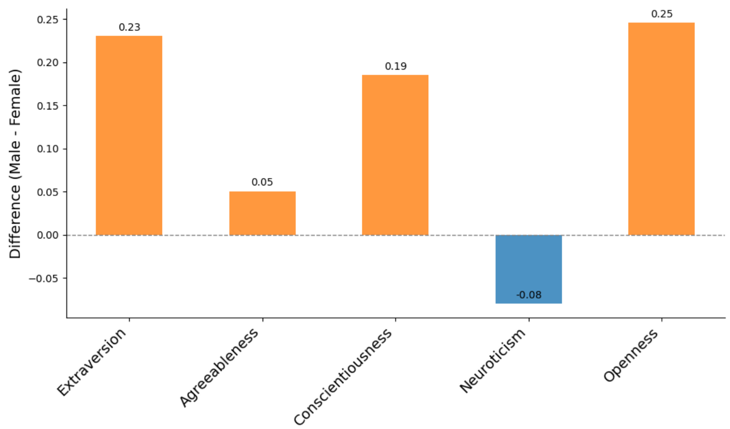
Note: Figure shows the impact of key features on predicted log wages, which are derived from a random forest machine learning model. Features are ranked by their average absolute SHAP (SHapley Additive exPlanations) values, with more important predictors appearing higher on the vertical axis. Each dot 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. The factor score variables include those for executive function (labeled "exf"), language ("lang"), orientation ("orient"), visuospatial ("vsp"), and memory ("mem"). See Appendix B for more details on variables.



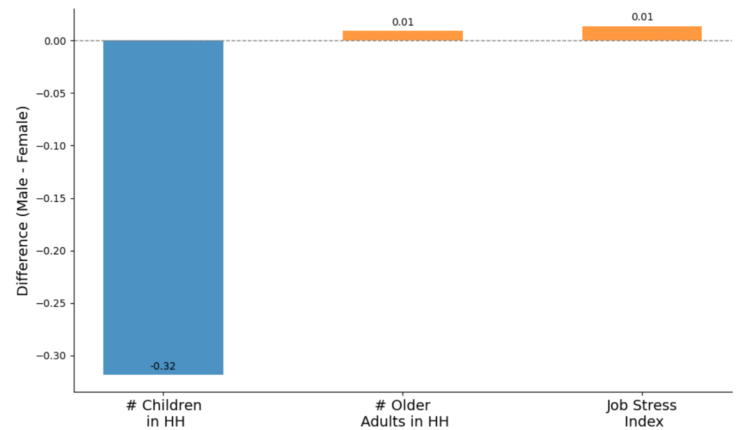
(a) Gender differences in cognition



(b) Gender differences in task indices



(c) Gender differences in Big Five personalities



(d) Gender differences in home care and job stress

Figure 2: Gender differences in novel KLPS measures

Note: These four subfigures show the differences between average male and female cognition scores, task indices, Big Five personalities, and home care and job stress index. Task indices, job stress index, general cognition score and the five component cognitive scores are obtained from the KLPS-5 C+ Module. Big Five personality indices are obtained from KLPS-4 I Module. Lagged Raven's scores are obtained from KLPS-2 and KLPS-3 data.

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

	Male			**Female**		
	Mean (SD)	Median	N	Mean (SD)	Median	N
Age of Respondent	37.2	37.0	2596	36.6	37.0	2742
	(2.7)			(2.6)		
Educational attainment	10.0	10.0	2596	8.7	8.0	2742
	(3.1)			(2.8)		
Experience	21.2	21.0	2596	21.9	22.0	2742
	(4.8)			(4.3)		
Urban residence (%)	47.7		2594	41.4		2739
Any agricultural work (%)	65.6		2595	60.6		2741
Any self-employment (%)	46.0		2596	41.5		2742
Any wage employment (%)	53.1		2595	33.9		2741
No hours worked in the last week (%)	4.7		2596	15.1		2742
Hours worked last week	47.4	48.0	2596	35.3	30.0	2742
	(28.2)			(29.4)		
Earnings last month (USD)	141.6	72.3	2596	49.8	21.4	2742
	(249.8)			(97.2)		
Hourly wage (USD)	1.02	0.40	2414	0.50	0.17	2280
	(4.96)			(2.43)		
Log hourly wage	3.97	4.04	2414	3.18	3.15	2280
	(1.35)			(1.34)		

Note: Sample consists of all individuals in KLPS-5, including those with a zero wage or work zero hours, for outcomes other than log hourly wage. Outcomes use 1 USD = 140 KSH. Years of experience is defined as (age - education - 6). All outcomes are weighted by KLPS population and tracking weights.

Table 2: Top occupations by gender

Occupation	Share Among Male (%)	Share Among Female (%)
Agriculture and Fishing - Farmer	18.26	33.36
Retail and commercial - Hawking/ selling clothes, food, other items	5.20	18.29
Professionals - Teacher	5.10	4.27
Unskilled trades - Domestic work (house boy/girl)	1.17	7.40
Agriculture and Fishing - Fishing	6.48	0.44
Unskilled trades - Bicycle/motorbike taxi work (boda-boda, piki-piki)	5.73	0.00
Unskilled trades - Cleaner (other)	0.85	4.18
Retail and commercial - Own shop	2.02	2.70
Skilled & semi-skilled trades - Barber or hairdresser	1.17	4.18
Skilled & semi-skilled trades - Mason	5.63	0.09
Skilled & semi-skilled trades - Tailor or seamstress	0.74	3.92
Unskilled trades - Unskilled construction laborer	3.50	0.35
Unskilled trades - Watchman/ Guard	3.72	0.87
Skilled & semi-skilled trades - Other skilled construction work	4.03	0.26
Unskilled trades - Cook/ Chef/ Caterer	1.49	2.26
Other - Other (specify)	2.55	0.78
Professionals - Salaried professional (manager, accountant, legal clerk)	2.23	0.96
Unskilled trades - Work in other person's commercial or financial business	2.34	0.96
Agriculture and Fishing - Agricultural laborer	2.23	1.74
Professionals - Police/military officer	2.87	0.78

Note: N= 5,338. 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 (combining both men and women). The right two columns show, for each occupation, the share of women and men employed in that occupation respectively.

Table 3: Raw and adjusted gender wage gap

Sample	(1) All	(2) Urban	(3) Wage Workers	(4) Urban Wage Workers
Panel A: No covariates				
Female	-0.79***	-0.75***	-0.38***	-0.52***
	(0.05)	(0.07)	(0.07)	(0.08)
N	4689	1959	2153	1213
Panel B: Control for education, demographics, and occupation				
Female	-0.49***	-0.42***	-0.29***	-0.25***
	(0.06)	(0.09)	(0.07)	(0.09)
Educational attainment	0.05***	0.07***	0.05***	0.05**
	(0.01)	(0.02)	(0.02)	(0.02)
Experience	-0.04	-0.06	-0.05	-0.04
	(0.04)	(0.06)	(0.05)	(0.06)
Experience^2/100	0.08	0.16	0.10	0.10
	(0.09)	(0.13)	(0.11)	(0.14)
Urban	0.41***	0.00	0.36***	0.00
	(0.06)	(.)	(0.08)	(.)
Nairobi	0.16**	0.18***	0.21***	0.21***
	(0.07)	(0.07)	(0.08)	(0.08)
Agriculture	-0.05	-0.14**	-0.06	-0.12
	(0.06)	(0.07)	(0.07)	(0.08)
Self-employment/business	0.29***	0.10	0.31***	0.29***
	(0.05)	(0.07)	(0.07)	(0.09)
N	4687	1957	2152	1211

Note: The dependent variable is log hourly wage. Panel A does not include any control variables. Panel B includes standard demographic and education controls (years of education, experience, urban, Nairobi, working in agriculture, and working in self-employment) as well as 49 occupation indicators. Wage work is defined as being employed and working for pay in the last 30 days. All standard errors are robust. * $p < 0.1$, ** $p < .05$, *** $p < 0.01$.

Table 4: Raw and adjusted hours worked gap by gender

Sample	(1) All	(2) Urban	(3) Wage Workers	(4) Urban Wage Workers
Panel A: No covariates				
Female	-12.22***	-13.64***	-5.23***	-7.31***
	(1.03)	(1.71)	(1.70)	(2.29)
N	5331	2324	2261	1288
Panel B: Control for education, demographics, and occupation				
Female	-8.87***	-9.46***	-4.49**	-4.05*
	(1.09)	(1.89)	(1.74)	(2.30)
Educational attainment	0.30	-0.26	0.32	-0.31
	(0.28)	(0.47)	(0.47)	(0.67)
Experience	1.18	2.09	0.43	0.84
	(0.85)	(1.28)	(1.13)	(1.46)
$\text{Experience}^2/100$	-2.84	-6.06**	-1.23	-3.47
	(1.90)	(2.96)	(2.69)	(3.47)
Urban	2.52*	0.00	1.52	0.00
	(1.31)	(.)	(2.02)	(.)
Nairobi	-1.88	-2.08	-2.39	-2.06
	(1.56)	(1.57)	(2.09)	(2.09)
Agriculture	-3.31***	-2.36	-3.37**	-2.82
	(1.27)	(1.64)	(1.70)	(1.99)
Self-employment/business	11.48***	10.56***	13.88***	16.62***
	(1.09)	(1.83)	(2.16)	(3.03)
N	5330	2324	2260	1286

Note: The dependent variable is weekly hours worked. Panel A does not include any controls. Panel B includes standard demographics and education controls (years of education, experience, urban, Nairobi, working in agriculture, and working in self-employment) as well as 49 occupation indicators. Wage work is defined as being employed and working for pay in the last 30 days. All standard errors are robust. * $p < 0.1$, ** $p < .05$, *** $p < 0.01$.

Table 5: Gender wage gap: controlling for novel individual and job controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female	-0.79***	-0.59***	-0.59***	-0.49***	-0.48***	-0.42***	-0.39***
	(0.05)	(0.05)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Educational attainment		0.09***		0.05***	0.03**	0.03**	0.02*
		(0.01)		(0.01)	(0.01)	(0.01)	(0.01)
Experience		-0.14***		-0.04	-0.05	-0.04	-0.04
		(0.04)		(0.04)	(0.04)	(0.04)	(0.04)
Experience^2/100		0.28***		0.08	0.11	0.10	0.09
		(0.09)		(0.09)	(0.09)	(0.09)	(0.09)
Urban		0.58***		0.41***	0.42***	0.37***	0.34***
		(0.06)		(0.06)	(0.06)	(0.06)	(0.06)
Nairobi		0.21***		0.16**	0.18***	0.16**	0.19***
		(0.07)		(0.07)	(0.07)	(0.07)	(0.06)
Agriculture		0.08		-0.05	-0.04	-0.01	-0.03
		(0.06)		(0.06)	(0.06)	(0.06)	(0.05)
Self-employment/business		0.03		0.29***	0.29***	0.25***	0.20***
		(0.05)		(0.05)	(0.05)	(0.05)	(0.05)
General cognition score					0.18***	0.16***	-0.16
					(0.03)	(0.03)	(0.29)
Analytical Index						0.10***	0.07***
						(0.03)	(0.03)
Interpersonal index						0.11***	0.10***
						(0.03)	(0.03)
Routine/Manual index						-0.00	0.02
						(0.02)	(0.02)
# Older adults in HH							0.06
							(0.07)
# Children in HH							-0.05***
							(0.01)
Job stress index							-0.15***
							(0.02)
N	4689	4689	4687	4687	4687	4687	4687
Occupation FEs			Yes	Yes	Yes	Yes	Yes
Personality Vars							Yes
Preferences Vars							Yes

Note: The dependent variable is log hourly wage. 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. In Columns 6 and 7, we impute missing job task indices, personality and economic preferences variables with the sample mean of the respective variable, and introduce missing indicators. All standard errors are robust. * $p < 0.1$, ** $p < .05$, *** $p < 0.01$.

Table 6: Gender gap in hours worked: controlling for novel individual and job controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female	-12.22***	-10.74***	-10.13***	-8.87***	-8.85***	-7.37***	-6.98***
	(1.03)	(0.99)	(1.12)	(1.09)	(1.09)	(1.13)	(1.15)
Educational attainment		0.48		0.30	0.25	0.21	0.09
		(0.29)		(0.28)	(0.28)	(0.28)	(0.30)
Experience		1.33		1.18	1.16	1.35	1.35
		(0.94)		(0.85)	(0.84)	(0.85)	(0.83)
Experience^2/100		-3.43*		-2.84	-2.76	-3.12*	-3.11*
		(2.08)		(1.90)	(1.89)	(1.89)	(1.85)
Urban		6.51***		2.52*	2.54*	1.61	0.92
		(1.38)		(1.31)	(1.31)	(1.33)	(1.34)
Nairobi		-1.88		-1.88	-1.85	-2.08	-1.62
		(1.72)		(1.56)	(1.56)	(1.54)	(1.53)
Agriculture		0.11		-3.31***	-3.29***	-2.76**	-3.28***
		(1.29)		(1.27)	(1.27)	(1.28)	(1.27)
Self-employment/business		12.83***		11.48***	11.48***	10.72***	10.52***
		(0.97)		(1.09)	(1.09)	(1.11)	(1.11)
General cognition score					0.41	0.00	9.62
					(0.56)	(0.57)	(6.48)
Analytical Index						1.29**	1.18*
						(0.62)	(0.62)
Interpersonal index						3.16***	3.16***
						(0.67)	(0.66)
Routine/Manual index						0.42	0.45
						(0.60)	(0.59)
# Older adults in HH							-1.04
							(1.28)
# Children in HH							-0.29
							(0.26)
Job stress index							0.16
							(0.55)
N	5331	5331	5330	5330	5330	5330	5330
Occupation FEs			Yes	Yes	Yes	Yes	Yes
Personality Vars							Yes
Preferences Vars							Yes

Note: 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. Please see Table 5 for details on the controls by column. All standard errors are robust. * $p < 0.1$, ** $p < .05$, *** $p < 0.01$.

Table 7: Double/Debiased Machine Learning results

	(1)	(2)	(3)	(4)
Panel A: Log Wage				
Female	-0.54***	-0.40***	-0.56***	-0.40***
	(0.05)	(0.05)	(0.06)	(0.05)
Observations	4689	4689	4554	4554
Panel B: Weekly Hours Worked				
Female	-8.19***	-6.37***	-7.29***	-6.10***
	(1.12)	(1.22)	(1.10)	(1.18)
Observations	5331	5331	5186	5186
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. * $p < 0.1$, ** $p < .05$, *** $p < 0.01$.

Table 8: Oaxaca-Blinder decomposition of wages and hours worked

	Coef.	95% CI Lower	95% CI Upper
Panel A: Log Wage			
With basic controls:			
Difference(Male-Female)	0.792	0.702	0.882
Explained	0.320	0.217	0.424
Unexplained	0.472	0.345	0.599
With full set of controls:			
Difference(Male-Female)	0.792	0.702	0.883
Explained	0.405	0.331	0.479
Unexplained	0.387	0.284	0.490
Panel B: Weekly Hours Worked			
With basic controls:			
Difference(Male-Female)	12.223	10.313	14.133
Explained	1.499	-0.692	3.691
Unexplained	10.724	8.027	13.421
With full set of controls:			
Difference(Male-Female)	12.223	10.288	14.158
Explained	3.861	2.290	5.432
Unexplained	8.362	6.030	10.695

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

Table 9: Gender wage gap: interaction terms

	(1)	(2)	(3)	(4)
Female	-0.49**	-0.53	-0.74***	-0.38
	(0.21)	(1.00)	(0.23)	(1.20)
Interaction terms with Female				
Urban	-0.09	-0.12	0.043	-0.16
	(0.12)	(0.12)	(0.18)	(0.15)
Nairobi	0.06	0.17	-0.051	0.02
	(0.13)	(0.13)	(0.18)	(0.16)
Agriculture	0.08	0.03	-0.06	-0.15
	(0.11)	(0.11)	(0.15)	(0.14)
Self-employment/business	-0.14	-0.12	-0.24***	-0.22**
	(0.09)	(0.09)	(0.11)	(0.11)
Educational attainment	-0.00	0.01	0.04*	0.04
	(0.02)	(0.03)	(0.02)	(0.03)
General cognition score		0.82		0.70
		(0.58)		(0.79)
Executive functioning		-0.32		-0.19
		(0.24)		(0.33)
Orientation		0.02		-0.06
		(0.11)		(0.16)
Memory		-0.38		-0.28
		(0.31)		(0.42)
Language		-0.24		-0.17
		(0.15)		(0.20)
Visuospatial		-0.18		-0.19
		(0.12)		(0.17)
Lagged Raven's		0.02		-0.04
		(0.05)		(0.07)
Experience		0.01		-0.31
		(0.08)		(0.09)
$Experience^2/100$		-0.01		0.11
		(0.18)		(0.21)
Analytical Index		-0.14***		-0.13*
		(0.05)		(0.07)
Interpersonal Index		0.17***		0.04
		(0.05)		(0.07)
Routine/Manual Index		0.05		0.03
		(0.05)		(0.06)
N	4687	4687	4687	4687
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 5, along with the listed interaction terms. Columns 2 and 4 include the basic set of controls a la Column 7 in Table 5, along with the listed interaction terms. Columns 3 and 4 run DML regressions following the same setup as in Table 7. All standard errors are robust. * $p < 0.1$, ** $p < .05$, *** $p < 0.01$.

Table 10: 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	0.50	-0.02	-0.03
			(0.02)	(0.02)
Treated with less respect than other people	0.52	0.52	-0.00	-0.03
			(0.02)	(0.02)
Receive poorer service than other people at restaurants/stores	0.31	0.27	-0.04**	-0.07***
			(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	0.42	0.37	-0.05**	-0.04
			(0.02)	(0.02)
You are threatened or harassed	0.29	0.26	-0.03	-0.07***
			(0.02)	(0.02)
Panel B: Main perceived reason for discrimination	Mean (Male)	Mean (Female)		
Your education or income level	0.48	0.50		
Your Social Group or Family Background	0.17	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 5. 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 = 5,287 (excluding 51 individuals who did not complete the discrimination module). * $p < 0.1$, ** $p < .05$, *** $p < 0.01$.

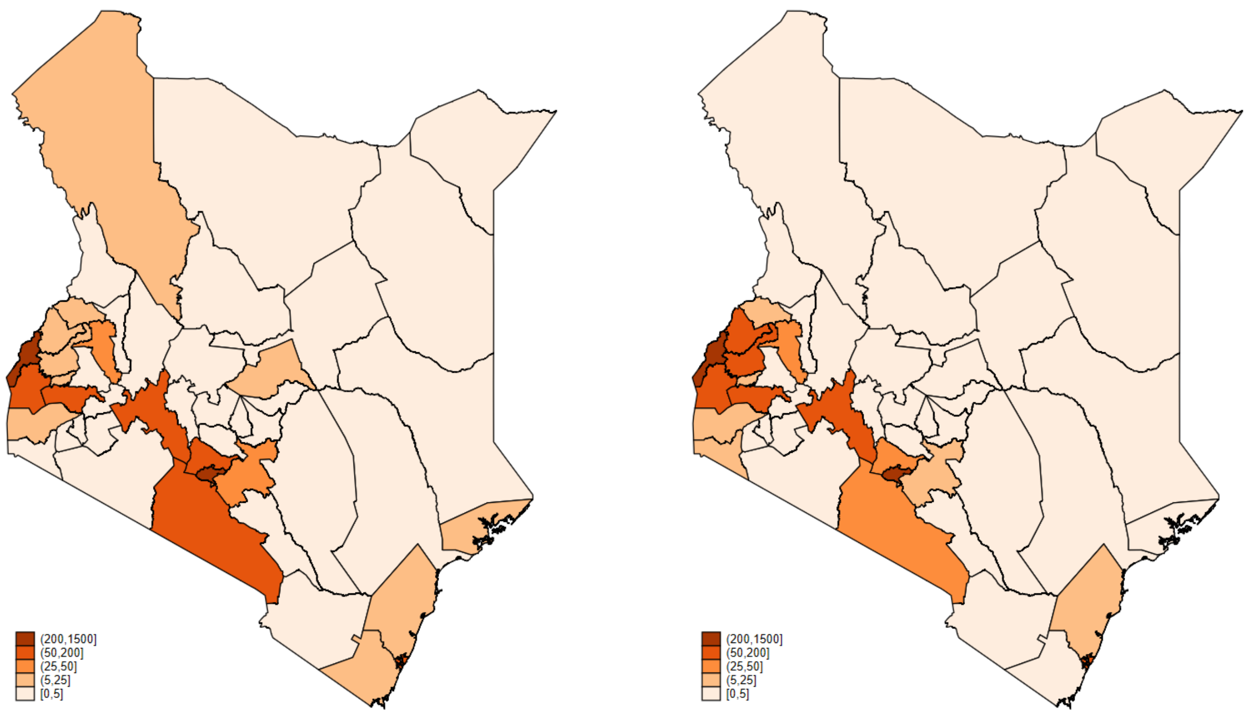
Table 11: Attitudes toward gender norms

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***	0.15***
			(0.02)	(0.02)
If wife is working outside the home, husband should help her with household chores	0.84	0.80	-0.05***	-0.05**
			(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 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 5. 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 = 4,598 (excluding individuals in KLPS-5 who did not complete the KLPS-4 survey).

Appendix (for online publication)

A Additional Exhibits



(a) Men (b) Women

Figure A.1: Residential location at the time of KLPS-5 survey

Note: This figure plots the residential location at the time of the KLPS-5 survey conducted in 2023. All respondents had originally attended primary school in 1998 in Busia County in western Kenya (the very dark shaded county all the way to the left). The figure presents the number of observations by Kenyan county that were surveyed in the KLPS-5 Module, for men and women separately. Observations are weighted to be representative of the original PSDP population.

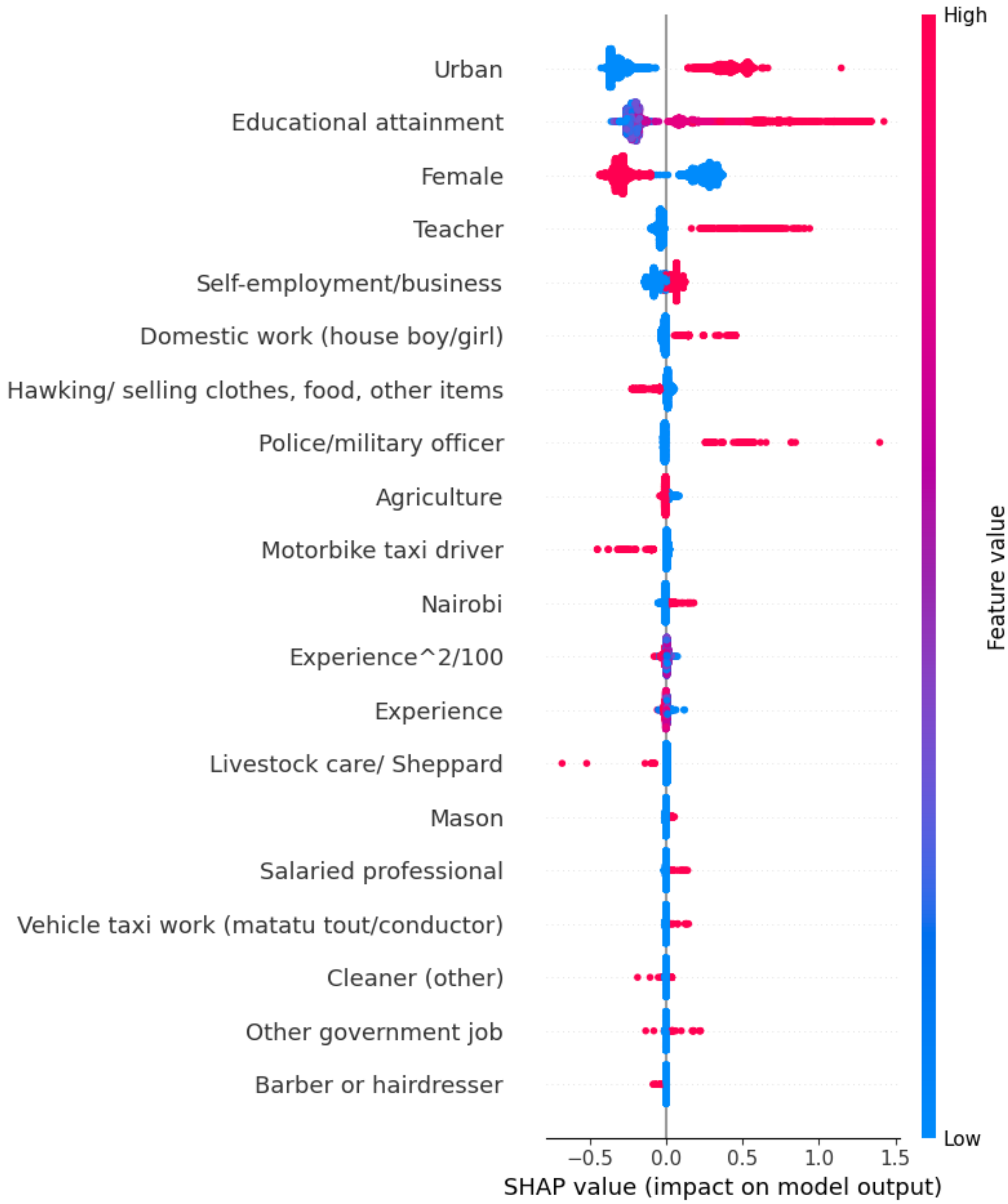


Figure A.2: SHAP summary plot predicting wages with basic set of predictors

Note: See Figure 1 notes, all of which apply here. The model includes 55 education, experience, demographics, and occupation variables as predictors of log wages.

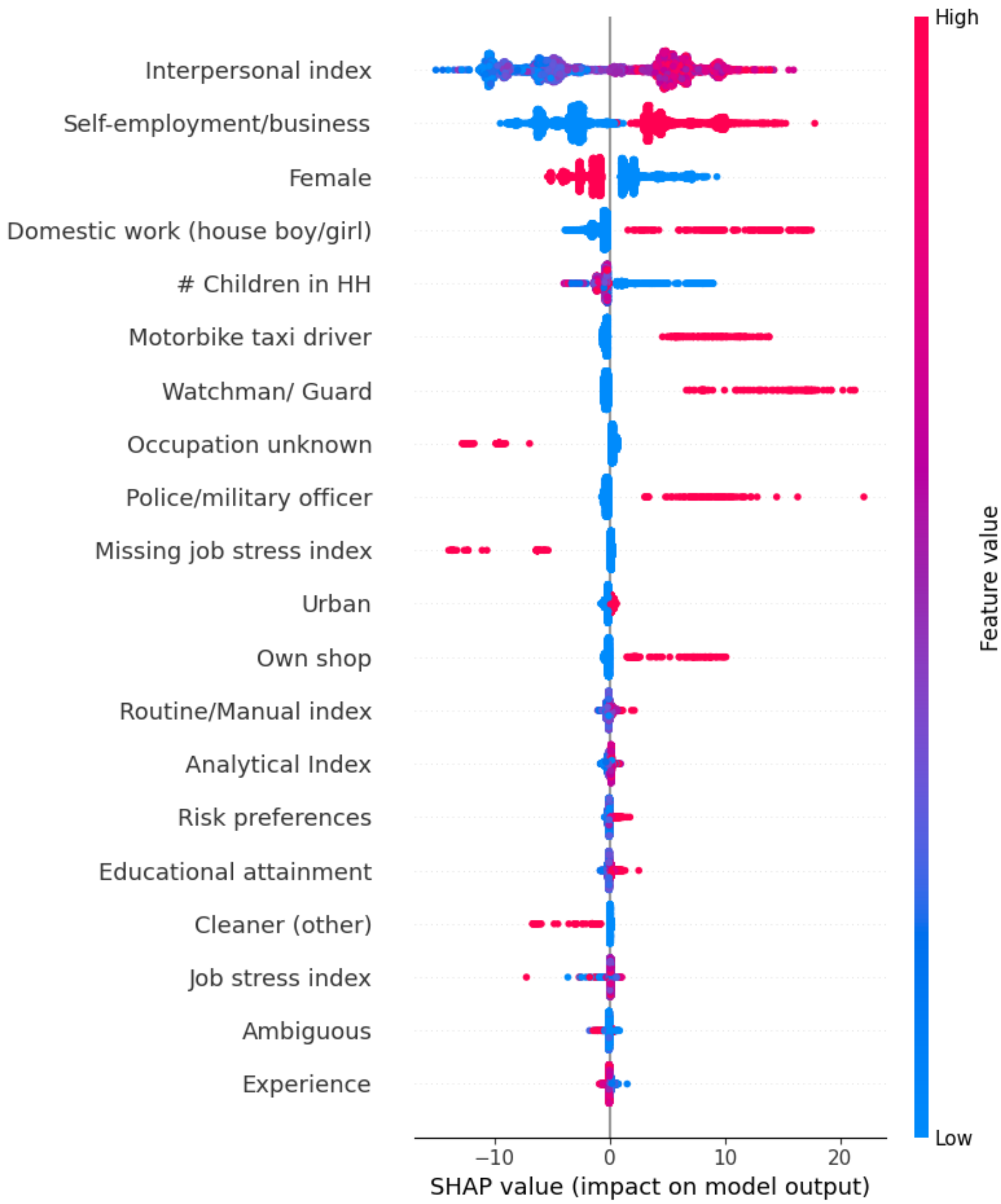
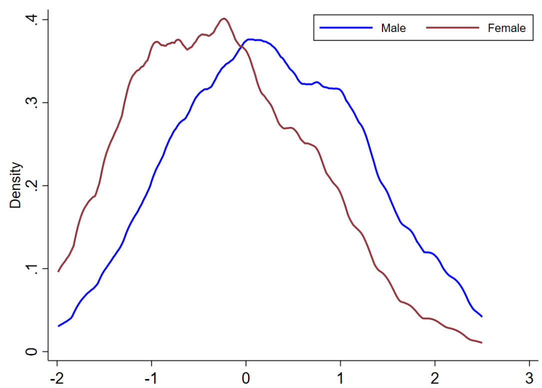


Figure A.3: SHAP summary plot predicting hours worked

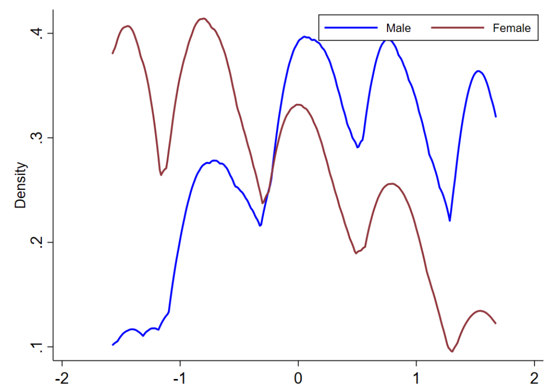
Note: Figure shows the impact of key features on predicted hours worked, which are derived from a random forest machine learning model. Features are ranked by their average absolute SHAP (SHapley Additive exPlanations) values, with more important predictors appearing higher on the vertical axis. Each 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.4: SHAP summary plot predicting hours worked with basic set of predictors

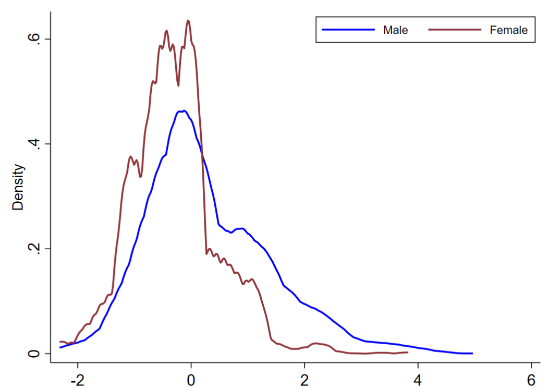
Note: See Figure A.3 notes, all of which apply here. The model includes 55 education, experience, demographics, and occupation variables as predictors of hours worked.



(a) Analytical Index



(b) Interpersonal Index



(c) Routine/Manual Index

Figure A.5: 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.

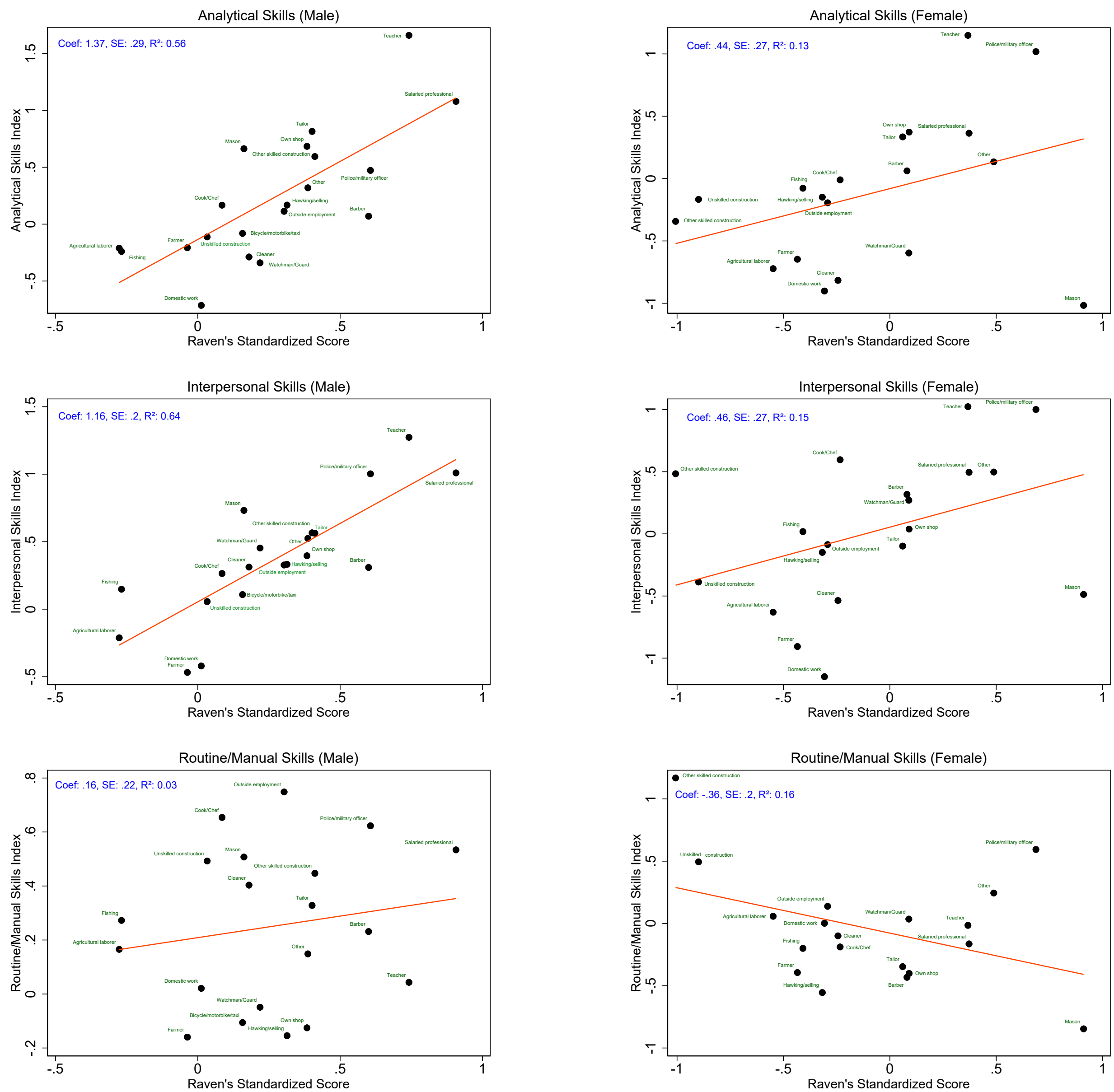


Figure A.6: Scatter plots of top 20 KLPS occupations of task indices and raven's standardized score by gender

Note: Figures plot the raven's standardized score for top 20 KLPS occupations by gender for the analytical task index, interpersonal task index, and routine/manual task index. The standardized Raven's score was calculated by subtracting the round-specific sample mean from each respondent's raw score and dividing the result by the corresponding standard deviation, with sampling weights applied in all calculations.

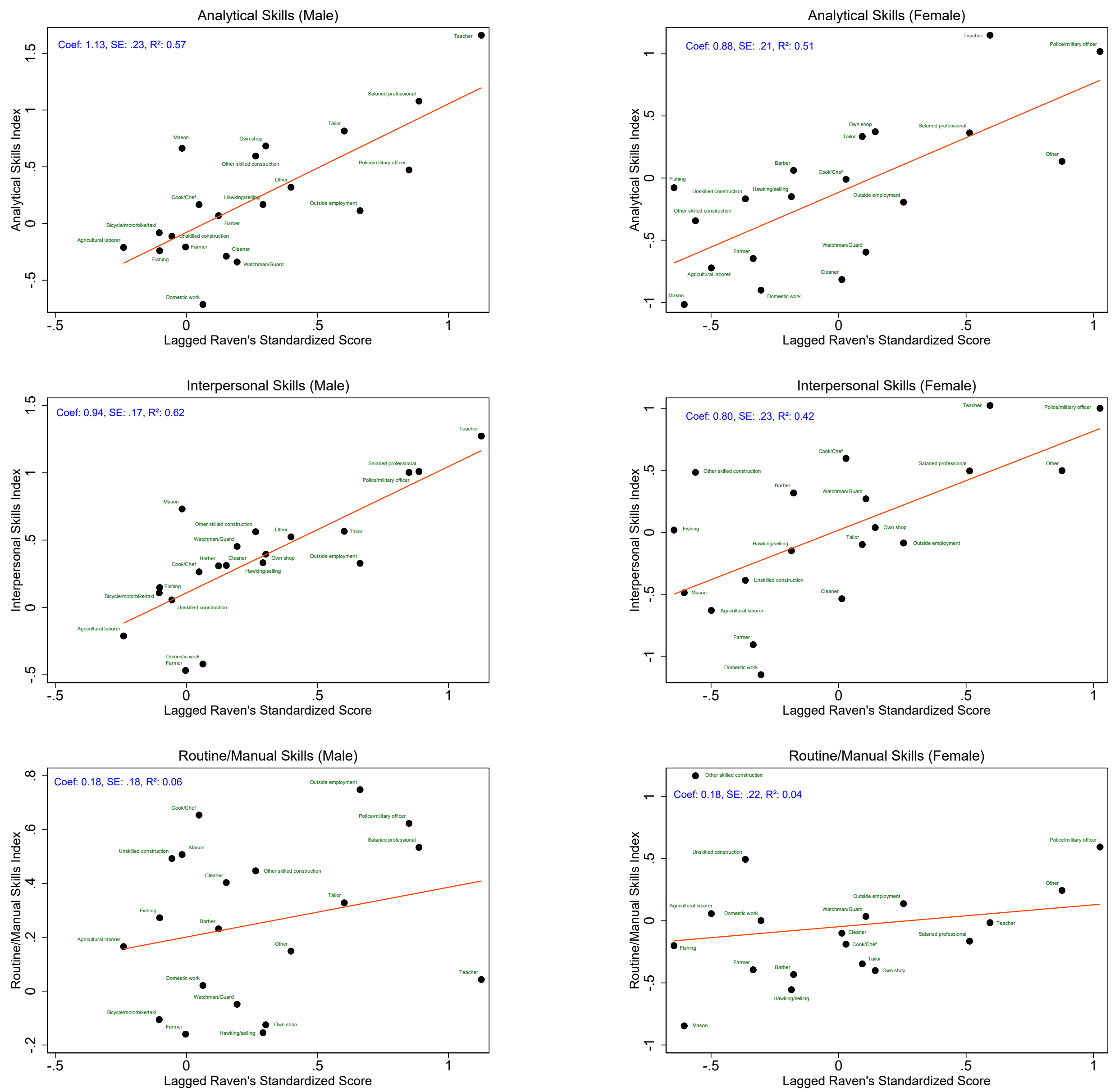


Figure A.7: Scatter plots of top 20 KLPS occupations of task indices and lagged raven's standardized score by gender

Note: Figures plot the lagged raven's standardized score for top 20 KLPS occupations by gender for the analytical task index, interpersonal task index, and routine/manual task index. The standardized Raven's score was calculated by subtracting the round-specific sample mean from each respondent's raw score and dividing the result by the corresponding standard deviation, with sampling weights applied in all calculations.

Table A.1: KLPS-5 and KLPS-4 effective tracking and survey statistics

	(1) Found	(2) Deceased	(3) Surveyed, among non-deceased	(4) N Surveyed
Panel A: KLPS-5 C+ Module				
Overall	0.910	0.056	0.820	5339
Male	0.910	0.064	0.823	2603
Female	0.910	0.048	0.817	2736
Panel B: KLPS-4 I Module				
Overall	0.909	0.054	0.876	5182
Male	0.919	0.062	0.887	2543
Female	0.898	0.047	0.865	2639
Panel C: KLPS-4 E+ Module				
Overall	0.899	0.045	0.849	5003
Male	0.901	0.053	0.855	2483
Female	0.897	0.037	0.843	2520

Note: Table shows the "effective" tracking statistics for the main survey modules used in this paper, from KLPS-5 and KLPS-4. The KLPS-5 C+ Module includes the cognitive battery, earnings, occupation, and job tasks data. The KLPS-4 I Module includes preference and personality measures, along with gender attitudes. The KLPS-4 E+ Module includes economic activities and earnings data. The first three columns show the means of indicator variables for respondent found, deceased, or surveyed (among non-deceased respondents), respectively. The last column shows the number of respondents surveyed. Observations are weighted to be representative of the original KLPS population, and include KLPS population weights and KLPS intensive tracking weights, reflecting the two-stage tracking design of KLPS (see Hamory et al. (2021) for a discussion of the construction of effective tracking rates.)

Table A.2: Comparisons of Occupations in KLPS and KCHS, ISIC Rev.4

Occupation (ISIC-Rev.4)	**Percent (2022 KCHS)**	**Percent (KLPS-5)**
Agriculture; forestry and fishing	47.46%	29.76%
Manufacturing and mining	5.86%	4.73%
Construction	6.01%	8.82%
Wholesale and retail trade; repair of motor vehicles and motorcycles	12.93%	16.76%
Transportation and storage	5.23%	4.88%
Accommodation and food service activities	2.41%	3.19%
Information and communication	0.74%	0.39%
Professional, scientific and technical activities	2.33%	5.85%
Administrative and support service activities	4.19%	5.15%
Public administration and defense; compulsory social security	0.97%	2.51%
Education	3.50%	5.19%
Human health and social work activities	1.34%	0.54%
Other service activities	2.99%	3.03%
Domestic Work	3.07%	4.56%
Other	0.97%	4.64%

Note: The table presents the distribution of occupations in the 2022 Kenyan Continuing Household Survey (KCHS) and KLPS-5 data at the International Standard Industrial Classification of All Economic Activities (ISIC) Fourth Revision (Rev.4) level, a standard classification of economic activities from the United Nations. The 2022 KCHS is a representative sample of 24,000 households across urban and rural Kenya with a sample frame from the 2019 Kenyan Population and Housing Census.

Table A.3: Machine Learning Predictions for Log Wages and Hours Worked

	(1)	(2)	(3)	(4)
	Basic covariates		Full covariates	
	Male	Female	Male	Female
Panel A: Models accounting for gender				
Log wages	3.95	3.15	3.96	3.14
R^2	0.22	0.23	0.25	0.25
Hours worked	47.06	34.87	47.21	35.04
R^2	0.10	0.22	0.09	0.26
Panel B: Models not accounting for gender				
Log wages	3.74	3.45	3.78	3.39
R^2	0.25		0.29	
Hours worked	43.41	39.25	44.67	38.02
R^2	0.17		0.21	

Note: This table presents machine learning model predictions for log wages and hours worked, including gender as a predictor (Panel A) and excluding gender as a predictor (Panel B). Models are random forests with five-fold cross-validation; basic covariate models use 55 variables (basic demographics and 49 occupational indicators, corresponding to column 4 of Table 5), while full covariate models include basic demographics plus the novel measures (104 variables, corresponding to column 7 in Table 5).

Table A.4: Correlates of being a urban wage worker

	Educational attainment	General cognition score	Lagged Raven's	Analytical Index	Interpersonal index
Indicator: urban wage worker	1.28***	0.22***	0.24***	0.20***	0.36***
	(0.14)	(0.04)	(0.04)	(0.04)	(0.04)
N	4689	4689	4689	4689	4689

Note: The table shows a coefficient of each dependent variable on an indicator for being an urban wage worker in KLPS-5 and an indicator for gender. All standard errors are robust. * $p < 0.1$, ** $p < .05$, *** $p < 0.01$.

Table A.5: Differences in job task content by gender

	Mean (male)	Mean (female)	Diff	Std. Error	t
Panel A: Analytical Index	0.27	-0.26	0.53	0.03	15.83
Length documents read as part of work	0.20	-0.21	0.41	0.04	11.65
More than 30 min thinking as part of work	0.17	-0.17	0.34	0.04	9.36
Do math as part of work	0.20	-0.19	0.39	0.04	11.02
Learning new things as part of work	0.16	-0.16	0.33	0.04	9.10
Panel B: Interpersonal index	0.33	-0.31	0.64	0.03	19.46
Supervise as part of work	0.23	-0.23	0.46	0.04	13.01
Contact with clients as part of work	0.14	-0.14	0.28	0.04	7.81
Make presentation at work	0.20	-0.19	0.39	0.04	11.00
Collaborating with co-workers at work	0.32	-0.31	0.63	0.03	18.17
Panel C: Routine/Manual index	0.30	-0.29	0.59	0.03	17.75
Freedom to decide as part of work	0.09	-0.09	0.18	0.04	4.98
Repetitive tasks as part of work	-0.01	0.02	-0.03	0.04	-0.86
Operate heavy machinery as part of work	0.22	-0.22	0.44	0.04	12.56
Drive as part of work	0.25	-0.25	0.50	0.04	14.17
Repair/maintain equipment as part of work	0.15	-0.15	0.30	0.04	8.45
Work is physically demanding	0.10	-0.10	0.20	0.04	5.46
Observations	5338				

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.6: Coefficients on additional covariates in wage and hours worked regressions

	(1) Log Wage	(2) Hours worked
Female	-0.39***	-6.98***
	(0.06)	(1.15)
Executive functioning cog score	0.19	-2.40
	(0.12)	(2.63)
Orientation cog score	0.10*	-1.44
	(0.06)	(1.21)
Memory cog score	0.12	-5.28
	(0.15)	(3.43)
Language cog score	0.11	-2.35
	(0.07)	(1.62)
Visuospatial cog score	-0.02	-2.03
	(0.06)	(1.27)
Lagged Raven's Score	-0.03	-0.61
	(0.03)	(0.58)
Big 5 Personality: Extraversion	0.04*	0.32
	(0.02)	(0.45)
Big 5 Personality: Agreeableness	0.02	-0.07
	(0.02)	(0.48)
Big 5 Personality: Conscientiousness	-0.02	0.88*
	(0.02)	(0.49)
Big 5 Personality: Neuroticism	-0.06**	0.80*
	(0.02)	(0.47)
Big 5 Personality: Openness	0.03	0.64
	(0.02)	(0.50)
Willingness to buy gift	0.00	-0.54
	(0.09)	(2.00)
Reciprocity Measure	0.00**	-0.00
	(0.00)	(0.00)
Generalized Altruism Measure	-0.00	0.00**
	(0.00)	(0.00)
In-group Altruism Measure	-0.00	-0.00
	(0.00)	(0.00)
Out-group Altruism Measure	0.00	-0.00
	(0.00)	(0.00)
Risk aversion	0.00	0.05
	(0.00)	(0.05)
Patience	0.00*	0.02
	(0.00)	(0.05)
Ambiguity aversion	0.01	1.90
	(0.05)	(1.18)
N	4687	5330

Note: This table reports the coefficients on additional covariates in Column 7 of Table 5 and Column 7 of Table 6. Please see Appendix B for details on the economic preferences, Big Five personalities, and lagged Raven's variables. All standard errors are robust. * $p < 0.1$, ** $p < .05$, *** $p < 0.01$.

Table A.7: Oaxaca-Blinder decomposition of wages and hours worked, using female as the reference group

	Coef.	95% CI Lower	95% CI Upper
Panel A: Log Wage			
With basic controls:			
Difference(Male-Female)	0.792	0.715	0.870
Explained	0.309	0.153	0.464
Unexplained	0.483	0.320	0.646
With full set of controls:			
Difference(Male-Female)	0.792	0.702	0.883
Explained	0.430	0.350	0.509
Unexplained	0.363	0.258	0.468
Panel B: Weekly Hours Worked			
With basic controls:			
Difference(Male-Female)	12.223	10.667	13.780
Explained	4.575	1.272	7.878
Unexplained	7.649	4.139	11.158
With full set of controls:			
Difference(Male-Female)	12.223	10.288	14.158
Explained	7.599	5.755	9.442
Unexplained	4.625	2.076	7.174

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

Table A.8: Gender gap in hours worked: interaction terms

	(1)	(2)	(3)	(4)
Female (intercept)	-14.17***	-55.36**	-4.19	-80.08**
	(4.00)	(21.36)	(4.64)	(41.04)
Interaction terms with Female				
Urban	-2.65	-2.32	-4.26	-1.46
	(2.58)	(2.60)	(3.03)	(3.47)
Nairobi	-0.95	-0.79	1.9	-0.78
	(3.15)	(3.13)	(4.15)	(4.83)
Agriculture	2.05	1.95	-4.79	-1.53
	(2.49)	(2.44)	(3.08)	(3.21)
Self-employment/business	3.94**	4.32*	0.53	2.24
	(1.95)	(1.95)	(2.23)	(2.44)
Educational attainment	0.23	1.21**	0.05	1.35
	(0.30)	(0.56)	(0.35)	(0.73)
General cognition score		3.77		0.34
		(12.88)		(16.59)
Executive functioning		-2.89		-3.51
		(5.21)		(6.70)
Orientation		-1.54		-1.99
		(2.42)		(3.59)
Memory		-2.27		-1.01
		(6.86)		(8.72)
Language		-1.89		-2.53
		(3.19)		(4.21)
Visuospatial		-0.72		-0.07
		(2.47)		(3.51)
Lagged Raven's		0.15		1.23
		(1.14)		(1.40)
Experience		2.61		5.32*
		(1.69)		(2.90)
$\text{Experience}^2/100$		-4.62		-10.95*
		(3.79)		(6.28)
Analytical Index		2.62**		4.56***
		(1.16)		(1.51)
Interpersonal Index		-0.16		-0.96
		(1.23)		(1.63)
Routine/Manual Index		1.80		1.13
		(1.24)		(1.59)
N	5330	5330	5330	5330
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 6, along with the listed interaction terms. Columns 3 and 4 include the basic set of controls a la Column 7 in Table 6, along with the listed interaction terms. Columns 3 and 4 run DML regressions following the same setup as in Table 7. All standard errors are robust. * $p < 0.1$, ** $p < .05$, *** $p < 0.01$.

Table A.9: Gender gap among urban wage workers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female	-0.52***	-0.28***	-0.35***	-0.25***	-0.25***	-0.22**	-0.24***
	(0.08)	(0.07)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)
Educational attainment		0.09***		0.05**	0.03	0.03	0.02
		(0.02)		(0.02)	(0.03)	(0.03)	(0.02)
Experience		-0.08		-0.04	-0.05	-0.03	-0.02
		(0.06)		(0.06)	(0.06)	(0.06)	(0.07)
Experience^2/100		0.19		0.10	0.12	0.10	0.05
		(0.14)		(0.14)	(0.14)	(0.14)	(0.15)
Nairobi		0.18**		0.21***	0.22***	0.20**	0.20***
		(0.08)		(0.08)	(0.08)	(0.08)	(0.07)
Agriculture		-0.22***		-0.12	-0.11	-0.08	-0.06
		(0.08)		(0.08)	(0.07)	(0.08)	(0.07)
Self-employment/business		0.21**		0.29***	0.29***	0.27***	0.18**
		(0.10)		(0.09)	(0.09)	(0.09)	(0.09)
General cognition score					0.16***	0.14***	-0.43
					(0.04)	(0.04)	(0.43)
Analytical Index						0.15***	0.13***
						(0.04)	(0.04)
Interpersonal index						0.01	-0.02
						(0.05)	(0.05)
Routine/Manual index						-0.06*	-0.03
						(0.03)	(0.03)
# Older adults in HH							-0.12
							(0.22)
# Children in HH							-0.03
							(0.02)
Job stress index							-0.19***
							(0.04)
N	1213	1213	1211	1211	1211	1211	1211
Occupation FEs			Yes	Yes	Yes	Yes	Yes
Personality Vars							Yes
Preferences Vars							Yes

Note: Please see Table 5 for the notes, all of which apply here. The sample is restricted to individuals who reside in an urban area and perform wage work, defined as being employed and working for pay in the last 30 days. All standard errors are robust. * $p < 0.1$, ** $p < .05$, *** $p < 0.01$.

Table A.10: Gender gap in earnings

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female	-1.12***	-0.88***	-0.87***	-0.75***	-0.74***	-0.63***	-0.59***
	(0.05)	(0.05)	(0.06)	(0.05)	(0.05)	(0.05)	(0.05)
Educational attainment		0.11***		0.06***	0.04***	0.04***	0.03***
		(0.02)		(0.01)	(0.01)	(0.01)	(0.01)
Experience		-0.12***		-0.02	-0.03	-0.01	-0.00
		(0.04)		(0.04)	(0.04)	(0.04)	(0.04)
Experience^2/100		0.22**		0.02	0.06	0.03	0.01
		(0.10)		(0.08)	(0.08)	(0.08)	(0.08)
Urban		0.70***		0.41***	0.42***	0.35***	0.30***
		(0.06)		(0.06)	(0.06)	(0.06)	(0.05)
Nairobi		0.15**		0.10*	0.11*	0.09	0.11**
		(0.07)		(0.06)	(0.06)	(0.06)	(0.06)
Agriculture		0.30***		0.01	0.02	0.06	0.04
		(0.06)		(0.05)	(0.05)	(0.05)	(0.05)
Self-employment/business		0.30***		0.52***	0.52***	0.47***	0.43***
		(0.05)		(0.05)	(0.05)	(0.05)	(0.05)
General cognition score					0.20***	0.16***	-0.13
					(0.03)	(0.03)	(0.28)
Analytical Index						0.15***	0.13***
						(0.03)	(0.03)
Interpersonal index						0.19***	0.18***
						(0.03)	(0.03)
Routine/Manual index						0.04*	0.06***
						(0.02)	(0.02)
# Older adults in HH							0.02
							(0.07)
# Children in HH							-0.05***
							(0.01)
Job stress index							-0.12***
							(0.02)
N	4935	4935	4934	4934	4934	4934	4934
Occupation FEs			Yes	Yes	Yes	Yes	Yes
Personality Vars							Yes
Preferences Vars							Yes

Note: Please see Table 5 for the notes, all of which apply here. The dependent variable is log monthly earnings. All standard errors are robust. * $p < 0.1$, ** $p < .05$, *** $p < 0.01$.

Table A.11: Gender gap in wages, winsorized at the 95th percentile

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female	-0.77***	-0.58***	-0.57***	-0.48***	-0.47***	-0.42***	-0.38***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Educational attainment		0.08***		0.05***	0.03**	0.03**	0.02*
		(0.01)		(0.01)	(0.01)	(0.01)	(0.01)
Experience		-0.13***		-0.04	-0.05	-0.04	-0.04
		(0.04)		(0.04)	(0.04)	(0.04)	(0.04)
Experience^2/100		0.26***		0.08	0.11	0.10	0.09
		(0.09)		(0.08)	(0.08)	(0.08)	(0.08)
Urban		0.55***		0.38***	0.38***	0.34***	0.31***
		(0.06)		(0.06)	(0.06)	(0.06)	(0.05)
Nairobi		0.20***		0.14**	0.15***	0.14**	0.16***
		(0.07)		(0.06)	(0.06)	(0.06)	(0.06)
Agriculture		0.10*		-0.03	-0.02	0.01	-0.00
		(0.05)		(0.05)	(0.05)	(0.05)	(0.05)
Self-employment/business		0.03		0.27***	0.28***	0.24***	0.20***
		(0.04)		(0.05)	(0.05)	(0.05)	(0.05)
General cognition score					0.17***	0.15***	-0.11
					(0.03)	(0.03)	(0.27)
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)
# Older adults in HH							0.05
							(0.06)
# Children in HH							-0.04***
							(0.01)
Job stress index							-0.14***
							(0.02)
N	4689	4689	4687	4687	4687	4687	4687
Occupation FEs			Yes	Yes	Yes	Yes	Yes
Personality Vars							Yes
Preferences Vars							Yes

Note: Table follows the exact same specification and structure as Table 5, 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. * $p < 0.1$, ** $p < .05$, *** $p < 0.01$.

Table A.12: Gender gap in hours worked, winsorized at the 95th percentile

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female	-11.79***	-10.38***	-9.65***	-8.46***	-8.43***	-7.01***	-6.70***
	(0.95)	(0.93)	(1.03)	(0.99)	(0.99)	(1.03)	(1.06)
Educational attainment		0.47*		0.31	0.25	0.21	0.18
		(0.27)		(0.26)	(0.26)	(0.26)	(0.27)
Experience		1.22		1.09	1.06	1.24	1.27
		(0.90)		(0.81)	(0.81)	(0.81)	(0.80)
Experience^2/100		-3.13		-2.59	-2.50	-2.84	-2.91
		(2.00)		(1.82)	(1.81)	(1.81)	(1.78)
Urban		5.96***		2.21*	2.24*	1.34	0.73
		(1.27)		(1.21)	(1.21)	(1.23)	(1.24)
Nairobi		-1.73		-1.79	-1.75	-1.97	-1.54
		(1.58)		(1.43)	(1.43)	(1.41)	(1.40)
Agriculture		0.32		-2.78**	-2.76**	-2.24*	-2.69**
		(1.21)		(1.17)	(1.17)	(1.18)	(1.17)
Self-employment/business		12.21***		10.84***	10.84***	10.10***	10.01***
		(0.90)		(1.02)	(1.02)	(1.03)	(1.03)
General cognition score					0.50	0.09	7.23
					(0.53)	(0.53)	(5.83)
Analytical Index						1.37**	1.32**
						(0.58)	(0.58)
Interpersonal index						2.93***	2.97***
						(0.62)	(0.62)
Routine/Manual index						0.35	0.34
						(0.55)	(0.53)
# Older adults in HH							-1.04
							(1.23)
# Children in HH							-0.24
							(0.24)
Job stress index							0.42
							(0.52)
N	5331	5331	5330	5330	5330	5330	5330
Occupation FEs			Yes	Yes	Yes	Yes	Yes
Personality Vars							Yes
Preferences Vars							Yes

Note: Table follows the exact same specification and structure as Table 6, but the weekly hours worked variable is winsorized at the 95th percentile which is 89 hours. All standard errors are robust. * $p < 0.1$, ** $p < .05$, *** $p < 0.01$.

Table A.13: Gender wage gap, controlling for hours worked

	(1)	(2)	(3)	(4)
Female	-0.58***	-0.46***	-0.82***	-0.82***
	(0.05)	(0.05)	(0.04)	(0.04)
Weekly hours worked	-0.01***	-0.02***		
	(0.00)	(0.00)		
N	4687	4687	4689	4689
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 5, 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. * $p < 0.1$, ** $p < .05$, *** $p < 0.01$.

Table A.14: Gender wage gap, no imputation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female	-0.83***	-0.58***	-0.61***	-0.48***	-0.49***	-0.43***	-0.39***
	(0.05)	(0.05)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Educational attainment		0.11***		0.05***	0.02	0.02	0.02
		(0.01)		(0.01)	(0.02)	(0.02)	(0.02)
Experience		-0.15***		-0.07	-0.08*	-0.07	-0.06
		(0.05)		(0.04)	(0.04)	(0.04)	(0.04)
Experience^2/100		0.31***		0.14	0.16	0.14	0.12
		(0.11)		(0.10)	(0.10)	(0.10)	(0.10)
Urban		0.62***		0.46***	0.47***	0.42***	0.39***
		(0.07)		(0.07)	(0.07)	(0.07)	(0.07)
Nairobi		0.13*		0.08	0.10	0.10	0.10
		(0.08)		(0.07)	(0.07)	(0.07)	(0.07)
Agriculture		0.02		-0.12*	-0.11*	-0.08	-0.07
		(0.06)		(0.07)	(0.06)	(0.06)	(0.06)
Self-employment/business		0.11**		0.30***	0.29***	0.26***	0.22***
		(0.05)		(0.06)	(0.06)	(0.06)	(0.05)
General cognition score					0.18***	0.16***	0.32
					(0.03)	(0.03)	(0.32)
Analytical Index						0.09***	0.07**
						(0.03)	(0.03)
Interpersonal index						0.12***	0.10***
						(0.03)	(0.03)
Routine/Manual index						-0.00	0.02
						(0.03)	(0.03)
# Older adults in HH							0.10
							(0.07)
# Children in HH							-0.04***
							(0.01)
Job stress index							-0.13***
							(0.02)
N	3555	3555	3553	3553	3553	3553	3553
Occupation FEs			Yes	Yes	Yes	Yes	Yes
Personality Vars							Yes
Preferences Vars							Yes

Note: Please see Table 5 for the notes, all of which apply here. The sample is restricted to individuals with non-missing KLPS-4 economic preferences and personality measures as well as non-missing KLPS-5 job task complexity measures. All standard errors are robust. * $p < 0.1$, ** $p < .05$, *** $p < 0.01$.

Table A.15: Gender gap in hours worked, no imputation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female	-9.83***	-8.13***	-9.91***	-8.46***	-8.45***	-6.84***	-6.27***
	(1.18)	(1.12)	(1.27)	(1.25)	(1.25)	(1.27)	(1.32)
Educational attainment		0.70*		0.40	0.35	0.29	0.18
		(0.38)		(0.36)	(0.38)	(0.38)	(0.38)
Experience		0.55		0.76	0.74	0.99	0.97
		(0.99)		(0.91)	(0.91)	(0.91)	(0.91)
Experience^2/100		-1.24		-1.56	-1.51	-1.99	-1.90
		(2.21)		(2.06)	(2.06)	(2.06)	(2.06)
Urban		6.63***		2.14	2.16	1.12	0.77
		(1.59)		(1.58)	(1.58)	(1.58)	(1.57)
Nairobi		-0.14		-0.28	-0.24	-0.37	-0.44
		(2.05)		(1.93)	(1.93)	(1.88)	(1.87)
Agriculture		1.79		-2.02	-2.01	-1.29	-1.45
		(1.52)		(1.47)	(1.47)	(1.47)	(1.47)
Self-employment/business		12.16***		11.85***	11.86***	11.00***	10.84***
		(1.11)		(1.24)	(1.24)	(1.28)	(1.26)
General cognition score					0.33	-0.05	0.08
					(0.71)	(0.72)	(7.91)
Analytical Index						1.34*	1.36*
						(0.69)	(0.70)
Interpersonal index						3.59***	3.45***
						(0.77)	(0.76)
Routine/Manual index						0.47	0.51
						(0.66)	(0.65)
# Older adults in HH							-1.80
							(1.48)
# Children in HH							-0.08
							(0.31)
Job stress index							-0.21
							(0.54)
N	3873	3873	3870	3870	3870	3870	3870
Occupation FEs			Yes	Yes	Yes	Yes	Yes
Personality Vars							Yes
Preferences Vars							Yes

Note: Please see Table 5 for the notes, all of which apply here. The sample is restricted to individuals with non-missing KLPS-4 economic preferences and personality measures as well as non-missing KLPS-5 job task complexity measures. All standard errors are robust. * $p < 0.1$, ** $p < .05$, *** $p < 0.01$.

Table A.16: Gender wage gap, KLPS-5 sample interviewed in KLPS-4

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female	-0.69***	-0.50***	-0.49***	-0.37***	-0.37***	-0.34***	-0.32***
	(0.06)	(0.06)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
Educational attainment		0.11***		0.06***	0.05***	0.04**	0.03*
		(0.02)		(0.02)	(0.02)	(0.02)	(0.02)
Experience		-0.10*		-0.02	-0.02	-0.01	-0.01
		(0.06)		(0.05)	(0.05)	(0.05)	(0.05)
Experience^2/100		0.23*		0.05	0.06	0.04	0.03
		(0.13)		(0.12)	(0.12)	(0.12)	(0.12)
Urban		0.51***		0.44***	0.45***	0.41***	0.37***
		(0.07)		(0.07)	(0.07)	(0.07)	(0.07)
Nairobi		0.11		0.05	0.06	0.05	0.08
		(0.07)		(0.07)	(0.07)	(0.07)	(0.07)
Agriculture		-0.01		-0.05	-0.04	-0.01	-0.03
		(0.07)		(0.06)	(0.06)	(0.06)	(0.06)
Self-employment/business		-0.05		0.18***	0.20***	0.17**	0.14**
		(0.06)		(0.06)	(0.06)	(0.06)	(0.06)
General cognition score					0.13***	0.11***	0.38
					(0.03)	(0.03)	(0.37)
Analytical Index						0.06	0.04
						(0.04)	(0.04)
Interpersonal index						0.10***	0.08**
						(0.04)	(0.04)
Routine/Manual index						-0.00	0.02
						(0.03)	(0.03)
# Older adults in HH							-0.03
							(0.08)
# Children in HH							-0.03**
							(0.01)
Job stress index							-0.12***
							(0.03)
N	2140	2140	2138	2138	2138	2138	2138
Occupation FEs			Yes	Yes	Yes	Yes	Yes
Personality Vars							Yes
Preferences Vars							Yes

Note: Table follows the exact same specification and structure as Table 5, but the sample only includes individuals in KLPS-5 who were also interviewed in KLPS-4 E-module with a nonzero hourly wage in both rounds. All standard errors are robust. * $p < 0.1$, ** $p < .05$, *** $p < 0.01$.

Table A.17: Gender wage gap in KLPS-4

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	-0.76***	-0.55***	-0.72***	-0.56***	-0.56***	-0.53***	-0.53***	-0.38***
	(0.06)	(0.06)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
Educational attainment		0.13***		0.10***	0.08***	0.08***	0.09***	0.09***
		(0.02)		(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Experience		-0.09		-0.03	-0.03	-0.02	-0.00	-0.00
		(0.05)		(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Experience^2/100		0.33**		0.19	0.18	0.17	0.12	0.12
		(0.15)		(0.15)	(0.15)	(0.15)	(0.14)	(0.14)
Urban		0.45***		0.43***	0.42***	0.40***	0.36***	0.35***
		(0.07)		(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
Nairobi		0.20***		0.19***	0.21***	0.21***	0.23***	0.23***
		(0.07)		(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
Agriculture		0.01		0.01	0.01	0.01	0.01	0.03
		(0.06)		(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Self-employment/business		-0.18***		-0.12**	-0.13**	-0.14***	-0.15***	-0.14***
		(0.06)		(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
General cognition score					0.11***	0.10**	0.69*	0.68*
					(0.04)	(0.04)	(0.39)	(0.39)
Analytical Index						0.05	0.04	0.04
						(0.03)	(0.03)	(0.03)
Interpersonal index						0.08**	0.07*	0.06
						(0.04)	(0.04)	(0.04)
Routine/Manual index						-0.02	0.00	0.00
						(0.03)	(0.03)	(0.03)
# Older adults in HH							-0.05	-0.07
							(0.09)	(0.09)
# Children in HH							0.00	0.00
							(0.02)	(0.02)
Job stress index							-0.10***	-0.10***
							(0.03)	(0.03)
Chore hours last week								-0.01***
								(0.00)
Childcare hours last week								-0.00
								(0.00)
N	2140	2140	2138	2138	2138	2138.00	2138	2138
Occupation FEs			Yes	Yes	Yes	Yes	Yes	Yes
Personality Vars							Yes	Yes
Preferences Vars							Yes	Yes

Note: Table follows the same specification and structure as Table 5, but with two differences. First, the sample only includes individuals in KLPS-5 who were also interviewed in KLPS-4 E-module with a nonzero hourly wage in both rounds. Second, the outcome is the log hourly wage measured in KLPS-4 based on earnings over the past month and hours over the past week. Covariantes that changed sources include experience, urban, Nairobi, agriculture, and self-employment, which are all measured in KLPS-4. All standard errors are robust. * $p < 0.1$, ** $p < .05$, *** $p < 0.01$.

Table A.18: Attitudes toward gender norms, by urban/rural residence

	Urban		Rural	
Statement	Agree (Male)	Agree (Female)	Agree (Male)	Agree (Female)
It is okay for a woman to be a mechanic	0.96	0.94	0.95	0.91
Important decisions in the family should be made by the men of the family (reverse-coded)	0.43	0.49	0.31	0.37
If wife is working outside the home, husband should help her with household chores	0.81	0.79	0.87	0.8
Girls and boys have equal opportunities to get a secondary education	0.99	0.99	0.99	0.99
Women and men have equal opportunities to get a job that pays a wage or salary	0.99	0.99	0.98	0.98
Women can be good politicians and should be encouraged to stand in elections	0.94	0.93	0.9	0.91
In our country, women should have equal rights and receive the same treatment as men	0.92	0.94	0.9	0.91
No one has the right to use physical violence against anyone else	0.93	0.95	0.89	0.86
N	1916		2682	

Note: The first two columns show the the share of individuals in urban areas, among men and women respectively, who agree with the statement on the left-most column. The third and fourth columns show the same shares for individuals in rural areas. 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 = 4,598 (excluding individuals in KLPS-5 who did not complete the KLPS-4 survey).

Table A.19: 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	0.81	-0.11***	-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	0.08***
			(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**
			(0.02)	(0.02)
Women can be good politicians and should be encouraged to stand in elections	0.90	0.88	-0.02	0.02
			(0.01)	(0.02)
In our country, women should have equal rights and receive the same treatment as men	0.89	0.89	0.00	0.03
			(0.01)	(0.02)
No one has the right to use physical violence against anyone else	0.87	0.83	-0.04***	-0.00
			(0.01)	(0.02)

Note: The first two columns 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 5. 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. $N = 4,257$ (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., 2025). 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. The components of each index are first normalized into a variable with mean 0 and SD 1, and then added to yield the index, ensuring that all questions receive equal weights. We then normalize each one of the three indices again into a variable with mean 0 and SD 1.

The analytical Index is the sum of all the below (normalized):

- 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 the below (normalized):

- 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 the below (normalized):

- 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 in the I-module of 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 personality measures

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. 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. Thus, the KLPS-3 Raven's scores correspond to cognitive assessments conducted approximately 13 years before the KLPS-5 round, while the KLPS-2 Raven's scores represent assessments from about 16 years before KLPS-5.