2013

Sources of the Persistent Gender Wage Gap along the Unconditional Earnings Distribution: Findings from Kenya

Richard U. Agesa
Marshall University, agesa@marshall.edu

Jacqueline Agesa
Marshall University, agesaj@marshall.edu

Andrew Dabalen

Sources of the Persistent Gender Wage Gap along the Unconditional Earnings Distribution: Findings from Kenya

RICHARD U. AGESA, JACQUELINE AGESA & ANDREW DABALEN

ABSTRACT *Past studies on gender wage inequality in Africa typically attribute the gender pay gap either to gender differences in characteristics or in the return to characteristics. The authors suggest, however, that this understanding of the two sources may be far too general and possibly overlook the underlying covariates that drive the gender wage gap. Moreover, past studies focus on the gender wage gap exclusively at the conditional mean. The authors go further to evaluate the partial contribution of each wage-determining covariate to the magnitude of the gender pay gap along the unconditional earnings distribution. The authors' data are from Kenya, and their empirical technique mirrors re-centered influence function regressions. The authors' results are novel and suggest that while gender differences in characteristics and the return to characteristics widen the gender pay gap at the lower end of the wage distributions, gender differences in characteristics widen the gender wage gap at the upper end of the wage distributions. Importantly, the authors find that the underlying covariates driving gender differences in characteristics and the return to characteristics are the industry, occupation, higher education and region covariates. In the middle of the distributions, however, the authors find that gender differences in the return to characteristics, fueled by education and experience covariates, exert the strongest influence on the magnitude of the gender pay gap.*

I. Introduction

While the literature on gender wage differentials in advanced economies of the West is large and well developed, the literature on gender wage differentials in the developing economies of Asia, Latin America, and Africa is relatively modest, though burgeoning. Weichselbaumer & Winter-Ebmer (2005) provide evidence that the largest share of the literature on gender wage differentials in developing countries examines Asia (e.g. Ashraf & Ashraf, 1993; Horton, 1996). Latin America has the second largest share (e.g. Psacharoupoulos & Tzannatos, 1991; Montenegro, 2001). Africa has the smallest share (e.g. Appleton *et al.*, 1999; Siphambe & Bakwena-Thokweng, 2001; Hinks, 2002; Temesgen, 2006; Nordman & Roubaud, 2009; Nordman & Wolff-Francois, 2009). Furthermore, studies of Africa account for only 3% of all global studies on gender wage differences since the 1990s.

In estimating gender wage differences, previous studies typically employ the canonical Oaxaca (1973)/Blinder (1973) technique, or variants of the technique, to decompose the gender pay gap into two portions: one due to gender differences in measured or observed productivity-enhancing attributes, such as education and experience, and another due to gender differences in the treatment of otherwise equally qualified male and female workers, i.e. gender differences in unmeasured or unobserved characteristics. The latter is commonly referred to as discrimination, where discrimination is measured as gender differences in the return to attributes.

Studies from Africa, and Kenya in particular, concur that males earn higher wages than females; however, findings from Kenya are divided over the relative impact of gender differences in measured and unmeasured attributes on the magnitude of the gender pay gap (e.g. Agesa, 1999; Mariara-Kabubo, 2003). Findings from Kenya are hence inconclusive on the specific sources driving the gender pay gap. Moreover, studies from Africa (including Kenya) focus almost exclusively on the gender pay gap at the conditional mean, rather than along the wage distribution. Only recently have Ntuli (2009) and Agesa *et al.* (2009) separately examined the gender wage gap along the entire wage distribution.[1]

Without downplaying the significance of the above studies, this paper contends that the issues surrounding gender wage inequality in Kenya deserve further inquiry. We add two seemingly obvious yet surprisingly overlooked layers of research to this literature. First, we suggest that previous studies, by explaining the gender pay gap in terms of either gender differences in attributes or the return to attributes, identified general rather than specific sources of the gap. Here, we argue that traditional wage-determining covariates impact on the magnitude of the gender pay gap differently at different quantiles. In other words, the partial contribution of each individual wage-determining covariate to the magnitude of the gender pay gap may vary along the earnings distribution. Past studies, however, overlook the effect of each wage-determining covariate on the magnitude of the gender pay gap and hence may be silent on the underlying covariates (forces) enlarging the gender pay gap at different quantiles. We fill this void in the literature by considering the partial effect of wage-determining covariates on the magnitude of the gender pay gap at each quantile and thus identify explicit rather than implicit sources fueling the gender wage gap.

Second, past studies from Kenya use ordinary least squares (OLS) to estimate the gender pay gap either at the conditional mean (e.g. Agesa, 1999; Mariara-Kabubo, 2003), or along the entire wage distribution using reweighting methods (e.g. Agesa *et al.*, 2009). Ntuli (2009) extends the latter by using conditional quantile wage regressions (Autor *et al.*, 2005; Machado & Mata, 2005; Melly, 2005), generated from workers with specific characteristics to examine the gender pay gap along the entire wage distribution in South Africa. Our study takes this literature further; specifically, we exploit re-centered influence function (RIF) regressions to consider the gender pay gap along the entire unconditional wage distribution, generated from workers with different characteristics (Firpo, Fortin, and Lemieux, 2009a, 2009b, hereafter FFL, 2009). This approach has two advantages: first, the procedure allows us to evaluate the marginal impact of covariates on targeted unconditional quantiles. Second, stemming from this, since RIF wage distributions are generated from workers with different characteristics, our estimates may have more practical policy applications.

The RIF empirical approach entails decomposing the gender pay gap in Kenya into two components: one due to gender differences in productivity-enhancing attributes (i.e. composition effects) and another due to gender differences in the return to attributes (i.e. wage structure effects). We then account for the partial contribution of each covariate on wage structure and composition effects and identify covariates with the most impact on the magnitude of the gender pay gap along the unconditional distribution. Also, since our data may be vulnerable to issues of omitted variable bias, tied to unmeasured aspects of finding work for male workers, we correct for potential endogeneity in the data.

Such an analysis is justified because traditional societies typically regard women as subservient and dependent on their husbands (Wanjala & Were, 2009). For this reason, women are likely to be associated with the "domestic" sphere and men with the "marketplace." It is hence conceivable that women might be sorted into industries and occupations with a domestic orientation and men into industries and occupations with a market orientation.[2] Consequently, wage-determining covariates may impact gender earnings differently, with some covariates exerting a relatively stronger influence on the gender pay gap than others. For example, human capital covariates such as primary schooling, secondary schooling, or university education may impact the gender wage gap differently at different quantiles. Likewise, factors such as occupation (e.g. professional and administration) and industry (e.g. agriculture and transportation) may impact the gender pay gap differently at different quantiles.

We explore the empirical validity of this hypothesis.

Our results are novel and may uncover the fundamental forces driving the gender pay gap along the unconditional wage distribution. We find that gender differences in both characteristics and the return to characteristics widen the gender pay gap at the lower end, while gender differences in characteristics widen the gap at the upper end of the earnings distribution. Importantly, we find that the underlying covariates fueling gender differences in characteristics and the return to characteristics are industry (e.g. community and social services, wholesale and retail trade, and transportation), occupation (e.g. administration, professionals, farm, fisheries and wildlife, and services), higher education (e.g. university and postgraduate education), and regional covariates (e.g. Nyanza, Coast, and Western). In the middle of the distributions, however, we find that gender differences in the return to characteristics driven by higher education and experience covariates exert the strongest influence on the magnitude of the gender pay gap.

The rest of the paper is organized as follows. Section 2 discusses the data and descriptive statistics. Section 3 provides a theoretical background and methodology. Section 4 presents and analyzes the results. Section 5 concludes.

2. Data

2.1 Data Sources

Our data are drawn from the 2004–2005 Kenya Integrated Household Budget Survey (KIHBS), a nationally representative cross section data-set collected by the Kenya National Bureau of Statistics. The KIHBS was designed to provide regular updates on key economic indicators such as wages, poverty, unemployment, and consumption in an integrated way. The KIHBS has three relevant survey instruments: the household questionnaire, community questionnaire, and the market price questionnaire. The household survey instrument consists of several modules such as household information, education, health, energy, labor, housing, water, and sanitation. Using the household survey instrument, we create a subsample of 7521 urban and rural workers aged between 15 and 65. This subsample consists of 4834 male and 2687 female workers with reported wages, including the self-employed.

2.2 Choice of Covariates

The covariates consist of two continuous variables, experience and square of experience, and four categorical variables, education, region, industry, and occupation. The covariates reflect the standard human capital theory of wage determination formulated by Mincer (1974). The dependent variable is the log of hourly wages.

2.3 Descriptive Statistics

The complete list of all covariates, including their average values, is reported in Table 1. Interestingly, from Table 1, we find that average values for most covariates accentuate gender differences in observed characteristics as far as placement in various categories is concerned.[3] For example, while the averages for the education dummy variables suggest no significant differences in the proportion of male and female workers with a general education[4] (i.e. primary, secondary, and undergraduate university education), the averages for the higher education variables, i.e. 4 years of university education plus postgraduate education, suggest that there are significant differences in the proportions for male and female workers. The former reflects recent gains in general education for women, the latter implies barriers to higher education for women.

The averages for the regional dummy variables suggest a more-or-less balanced proportion of male and female workers across all regions. The exception is Nairobi and Central provinces where there are disproportionately more female workers. The averages for the industry dummy variables suggest a somewhat uneven representation of male and female workers across various industries. For instance, the averages for community and social services, wholesale and retail trade, restaurants, and hotels (which require relatively fewer skills) indicate a disproportionately larger share of female workers in these industries. By contrast, the averages in industries such as manual and transportation (which require somewhat more specialized skills) show that the proportion of male workers is substantially higher.

The averages for the occupation dummy variables also suggest a fairly unbalanced representation of male and female workers across various occupations. For example, the most conspicuous imbalance is in administration and professional occupations, which in general require relatively more skills and have disproportionately more male workers.

Overall, the averages for the industry and occupation dummy variables reported in Table 1 shed light on an important link between the possession of skills, gender, and labor allocation in Kenya. Industries and occupations, which require relatively more skills and hence are likely to pay higher wages, have a greater proportion of male workers; industries and occupations, which require relatively fewer skills and hence are likely to pay lower wages, have a greater proportion of female workers. It should be noted, however, that what we observe here is merely a correlation and does not necessarily imply causation.

Table 1. Average values of explanatory variables

	Male	Female
Variable		
Experience	22	20
Square of experience	484	400
Education categories		
Standard 1	0.004	0.007
Standard 2	0.014	0.017
Standard 3	0.026	0.027
Standard 4	0.048	0.043
Standard 5	0.039	0.043
Standard 6	0.056	0.059
Standard 7	0.131	0.122
Standard 8	0.197	0.189
Form 1	0.015	0.022
Form 2	0.046	0.044
Form 3	0.019	0.015
Form 4	0.308	0.338
Form 5	0.001	0.001

Form 6	0.029	0.019
One year university	0.002	0.001
Two years university	0.003	0.003
Three years university	0.008	0.006
Four years university	0.027	0.017
Postgraduate	0.015	0.009
Other	0.011	0.014
None	0.001	0.004
Regional categories		
Nairobi	0.085	0.103
Central	0.118	0.149
Coast	0.123	0.112
North-eastern	0.018	0.010
Nyanza	0.164	0.183
Rift valley	0.240	0.211
Western	0.089	0.074
Eastern	0.163	0.158
Industry categories		
Community	0.298	0.422
Agriculture	0.259	0.249
Manual	0.152	0.028
Wholesale	0.159	0.256
Transportation	0.103	0.020
Finance	0.029	0.025
Occupation categories		
Administration	0.037	0.022
Professionals	0.075	0.069
Technicians	0.104	0.108
Clerical	0.027	0.056
Service	0.089	0.147
Farm, fisheries and wildlife	0.147	0.171
Craftsmen	0.082	0.067
Elementary occupations	0.356	0.352
Machine operators	0.083	0.008
Number of observations	4834	2687

Table 2. Gender mean and median wages and the gender wage gap at all quantiles

	Male wages		Female wages		
Quantile	Mean	Median	Mean	Median	Wage gap, pay gap
10th	7.365	7.32	6.925	6.96	0.440
20th	7.813	7.77	7.316	7.28	0.497
30th	8.081	8.00	7.614	7.49	0.467
40th	8.263	8.19	7.817	7.66	0.446
50th	8.433	8.37	8.024	7.84	0.409
60th	8.595	8.52	8.218	8.02	0.377
70th	8.768	8.71	8.411	8.24	0.357
80th	9.009	8.92	8.607	8.44	0.402
90th	9.339	9.21	8.914	8.77	0.425

Table 2 reports the mean, the median, and the gender pay gap for male and female workers along the wage distribution. Columns 1 and 2 report the mean and median wages for male workers. Columns 3 and 4 report the mean and median wages for female workers. Column 5 reports the gender pay gap. From column 5 in Table 2, we find that the gender pay gap is largest at the lower end of the earnings distributions, between the 20th and 30th quantiles. The gap then narrows in the middle of the distributions, between the 40th and 70th quantiles, and then widens dramatically at the upper end of the distributions, between the 80th and 90th quantiles. These are important findings and may have two implications: first, at the top of the wage distributions, there may be a "glass ceiling" for female workers; at the bottom of the wage distributions, there may be a "sticky floor" for female workers. The glass ceiling and sticky floor refer to the phenomenon of female workers with the same attributes as men receiving lower wages.

Moreover, our finding of a relatively larger gender pay gap at the lower end of the earnings distribution suggests that the sticky floor phenomenon has a stronger effect than the glass ceiling. Our finding of a stronger sticky floor effect is also consistent with a similar finding in the literature (Agesa *et al*., 2009); however, we extend the Agesa *et al*. (2009) study by examining the fundamental forces (covariates) driving the gender pay gap at the upper, middle, and lower ends of the distributions.

3. Estimation Methods

Our empirical approach is analogous to RIF quantile regression methods formalized by FFL (2009). RIF regressions have two important properties: first, RIF regressions are semi-parametric and make no prior assumptions about the functional form of the wage distributions. Second, RIF regression estimates are generated from unconditional wage distributions and hence may have an advantage over OLS. The following example illustrates why. Suppose a policy maker wants to compute the impact of a 1-year increase in education on earnings for a subgroup of workers with different characteristics (unconditional effects). This contrasts with the impact of the extra year of schooling on earnings for workers with a specific set of covariates (conditional effects). Regarding the mean, the unconditional properties of wages W for the subgroup of workers with different characteristics can be computed by simply averaging over the covariates X. The averaging over X is possible because OLS models rely on the classical linear assumption of the expectations operator, i.e. $E(W|X) = X\beta$, which leads to $E(W) = E(X)\beta$.

The linearity assumption, however, cannot be generalized to nonlinear operators such as quantiles (the distributional statistic of interest in this paper). Therefore, conditional quantile regression models using the Autor *et al*. (2005), Machado & Mata (2005), Melly (2005) approach may not answer questions regarding unconditional properties of wages W. RIF regressions, by contrast, yield estimates generated from unconditional distributions and hence would have more practical interest to economists and policy makers.

3.1 RIF Regressions: Theoretical Background

At the core of RIF regressions is the ability to generate the average effects of all explanatory variables at a particular earnings quantile with the original dependent variable (log of monthly earnings) replaced by the RIF. In particular, the RIF for a quantile qt has the following specification:

$$\mathrm{RIF}(W;\, q_\tau) = q_\tau + \frac{\tau - \mathrm{I}(\mathrm{W} \le q_\tau}{f_W(q_{W\tau})}, \qquad (1)$$

where f_W is the marginal density function of earnings W and $I(\cdot)$ is an indicator function. FFL

(2009) show that if the RIF regression E[RIF($W; q_\tau$)|X] is well modeled by the familiar linear regression model E[RIF(W; q_τ)|X] = β, then the estimated coefficients represent the mean marginal effects of explanatory variables on the earnings quantiles.

However, since the true RIF(W; q_τ) is unobservable, we use its sample analogy R$\hat{\text{I}}$F ($W;\hat{q}_\tau$) by replacing the unknown quantities by the corresponding estimators as follows:

$$\widehat{RIF}(W;\hat{q}_\tau) = \hat{q}_\tau + \frac{\tau - I(W \leq \hat{q}_\tau))}{\hat{f}_w(\hat{q}_\tau))} \tag{2}$$

where $\hat{q}_\tau$ is the τth sample quantile and $\hat{f}_w$ is the kernel density estimator. FFL (2009) show that after averaging out, the coefficient estimates $\hat{\beta}$ generated from RIF regressions provide the average effect of the explanatory variables on earnings.

Central to the RIF unconditional quantile method is an influence function. The influence function is a widely used tool in robust statistics. The RIF function represents the influence of an individual observation on a distributional statistic of interest such as a quantile. The RIF is a linear approximation (the leading terms of the von Mises expansion) to the nonlinear function of distributional statistics, such as a quantile, and it essentially captures the change of the distributional statistic, such as a quantile, in response to a change in the underlying distribution.[5]

The RIF regression is a function E[RIF($Y; v$)\$X = x$]. By taking iterated expectations, the derived marginal effects of the covariates on the statistic of interest are obtained by averaging the RIF function with respect to changes in the distribution of the covariates. Here, like OLS regressions, RIF regressions typically assume a linear specification E[RIF($Y;\hat{q}_\tau$|X] = $X\beta$, where the coefficient β represents the marginal effect of X on the distributional statistic, the quantile qt. Also, FFL (2009) provide proof of the unconditional property of RIF regressions. As such, one can compare RIF regressions to OLS regressions, and while RIF regressions have the same nice unconditional properties as OLS, RIF regressions are more general as they apply to any distributional statistic such as the quantile and not just the mean. Moreover, a simple proof shows that RIF regressions associated with the mean statistic are identical to OLS regressions (FFL, 2009).

3.2 RIF Regressions: Empirical Estimation

The empirical estimation has two steps. The first is a reweighting procedure where we estimate three weighting functions v defined as follows:

$$\widehat{w}_m M = \frac{1}{\hat{p}}, \tag{3}$$

$$\widehat{w}_f F = \frac{1}{(1-\hat{p})}, \tag{4}$$

And $$\widehat{w}_c C = \frac{1}{\hat{p}}, * \frac{\hat{p}(x)}{1-\hat{p}(x)}, \tag{5}$$

where $\widehat{w}_m M$ is the weight for the distribution of male workers, $\widehat{w}_f F$ is the weight for the distribution of female workers, and $\widehat{\omega}_c C$ is the female counterfactual weighting function that would prevail if female workers have the same distribution of observed and unobserved characteristics as males. The variable x is the distribution of covariates and $\hat{p}$ is the probability that an individual i is male. The coefficient $\hat{p}(x) = p\boldsymbol{r}(m|\mathrm{x})$ is the propensity score, i.e. the conditional probability that individual i is male, given a set of observed covariates x. The propensity score has added significance because, as with most cross section work, our data may be vulnerable to endogeneity, i.e. omitted variable bias correlated with unmeasured aspects of finding work for male workers. The propensity score adjusts for the potential endogeneity. We estimate the propensity score using probit analysis. The results from the probit model are reported in Table 3.

The independent variables used in the probit model are age; marital status (married), not married is the base group; education (i.e. 4 years of university education and postgraduate education), workers without schooling and with 1, 2, and 3 years of university education.

Table 3. Probit estimates of the probability of being male in the labor market

Variable	Estimated coefficient
Age	0.016 (6.67)
Married	0.381 (4.95)
Four years of university education	0.239 (3.01)
Postgraduate university education	0.174 (2.11)
Manufacturing	0.143 (7.33)
Construction	0.112 (12.89)
Transportation	0.239 (10.96)
Machine operators	0.188 (8.44)

Note: T-statistics are in parentheses.

and postgraduate education constitute the base group; occupation (machine operators, construction, and manufacturing), the rest of occupation categories constitute the base group; and an industry covariate (transportation), the rest of the industry categories constitute the base group. The covariates selected in the probit model have disproportionally more male workers, consistent with the information gathered from Table 1. The findings from Table 3 suggest that all covariates in the probit model (i.e. age, marital status, industry, occupation, and human capital) are significant at the 1% level.

A word of caution is necessary on the estimation of the propensity score. We recognize that in the probit model, the independent variables are a subset of x, and identification of the participation terms would have been more satisfactorily resolved if one had variables that shifted the probability of male employment without affecting wages—this indeed is a perennial problem in the literature. The very limited nature of the data, however, precluded such attempts by the authors. And while the equations are technically identified, the dearth of instruments may suggest fragile estimates. Despite this possibility, our choice of instruments represents the probability of being male in the Kenyan labor market.

In the second step, we estimate RIF unconditional quantile wage regressions for male, female, and counterfactual female earnings specified as follows:

$$\widehat{RIF}(\omega_k; \hat{q}_\tau) = X_k \widehat{\beta}_k, \tag{6}$$

where $k = m, f, c$ and $\widehat{RIF}(\omega_k; \hat{q}_\tau)$ is the RIF estimate at the τ^{th} quantile $\hat{q}$ the coefficient $\hat{\beta}$ is the estimate of the unconditional quantile partial effect. Using the unconditional quantile regression estimates from Equation (6), if $v(W)$ is a quantile of the earnings distribution W (in logs), we can obtain the male–female pay gap $[v(W)m - v(W)f]$ at selected quantiles and decompose the pay gap into portions attributable to differences in characteristics (composition effects) and the return to characteristics (wage structure effects).

The decomposition is generalized as follows:

$$[v(W)_m - v(W)_f] = [v(W)_m - v(W)_c] + [v(W)_c - v(W)_f], \qquad (7)$$

where the first component on the right-hand side $[v(W)m - v(W)c]$ represents the composition effects, i.e. the gender earnings difference due to differences in labor market characteristics. The second term on the right-hand side $[v(W)c - v(W)_f]$ represents the wage structure effect, i.e. the gender earnings differences due to differences of men and women in the return to labor market characteristics.

Equation (7) can specifically be expressed as follows:

$$\hat{q}_\tau(W_m) - \hat{q}_\tau(W_f) = \{\bar{X}_f(\hat{\beta}_c - \hat{\beta}_f) + \hat{R}^S_\tau\} = \left\{\bar{X}_m\hat{\beta}_m - (\bar{X}_f\,\hat{\beta}_c) + \hat{R}^C_\tau\right\} \qquad (8)$$

where $\hat{q}_\tau(W_m) - \hat{q}_\tau(W_f)$ represents the raw gender earnings difference at the τth quantile. $\bar{X}_f$ and $\bar{X}_m$ represent the vector of covariate averages for female and male workers, respectively. The coefficient $\hat{\beta}_c$ is the estimate from the counterfactual distribution, which assumes the female distribution that would prevail if female workers had the same distribution of observed and unobserved characteristics as males. The gap $(\hat{\beta}_c - \hat{\beta}_f)$ measures male and female differences in the return to labor market characteristics, and the magnitude $\hat{X}_f(\hat{\beta}_c - \hat{\beta}_f)$ represents the wage structure effect, i.e. the gender earnings difference at the τth quantile attributable to different returns in labor market characteristics for male and female workers.

The difference $(\bar{X}_m\hat{\beta}_m - (\bar{X}_f\,\hat{\beta}_c)$ represents the composition effect, i.e. the gender earnings differential at the τth quantile attributable to gender differences in labor market characteristics. The magnitudes $\hat{R}^S_\tau$ and $\hat{R}^c_\tau$ are estimates of approximation errors corresponding to the wage structure and composition effects, respectively. These approximation errors result from the linear specification assumed by RIF regression functions. Consistent with FFL (2009), the approximation errors are specified as follows:

$$\hat{R}^S_\tau = [(q)(W_c) - \hat{q}(W_f)] - [\bar{X}_f(\hat{\beta}_c - \hat{\beta}_f)]$$

And

$$\hat{R}^c_\tau = [(\hat{q})(W_m) - \hat{q}(W_c)] - [\bar{X}_m(\hat{\beta}_m - \bar{X}_f(\hat{\beta}_c].$$

Moreover, in all the regression specifications, we multiply the relevant reweighting functions with the KIHBS sample weights and normalize each of the three weights to sum to 1, consistent with FFL (2009).

4. Findings

Visual representations of male and female wage distributions are reported in Figure 1. The wage distributions are generated using the Epanechnikov kernel density with a bandwidth of 0.0480. From Figure 1, both the male and female kernel wage densities are more-or-less bell shaped, with the male density showing a decisive rightward translation throughout the entire distribution. The rightward translation implies a relatively higher wage for male workers along the earnings distribution. The gap between the two densities represents the gender pay difference. Unsurprisingly, the visual evidence presented in Figure 1 is consistent with the findings from Table 2, which also suggests that the gender pay gap is not even (constant) along the entire wage distribution.

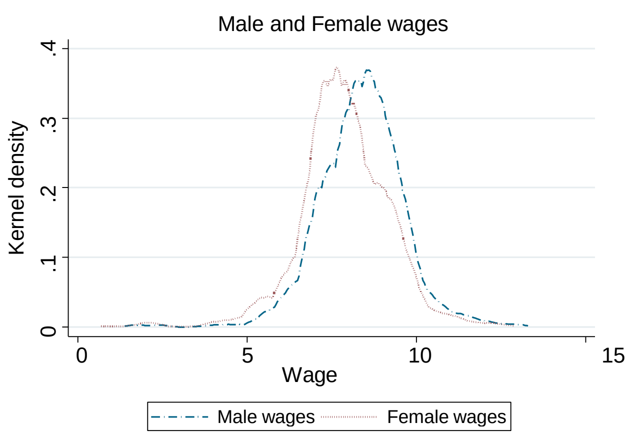

Figure 1. Male and female kernel wage densities.

4.1 Separate RIF Unconditional Quantile Wage Regressions for Male and Female Workers

The estimated coefficients from the RIF quantile wage regressions, together with the estimated (robust) standard errors (SE) for male and female workers, are reported in Tables 4 and 5, respectively. In general, most of the findings are unsurprising and are consistent with *a priori* expectations. For example, the estimated coefficient 0.035 for the human capital variable standard 1 in the 10th quantile of Table 4 suggests that male workers with standard 1 level of education earn about 4% more than male workers without formal education in the base group. This percentage is computed by taking the exponential of the coefficient 0.035 subtracting 1 and multiplying by 100 to express it as a percentage. Likewise, percentages for all other estimated coefficients in Tables 4 and 5 would be computed and interpreted in the same way. Additionally, the findings from Tables 4 and 5 mirror human capital theory of wage determination as formalized by Mincer (1974). The Mincerian theory contends that higher skilled groups typically earn relatively higher wages. Our findings support this theory. In particular, we find that the returns to different categories of human capital skills are relatively higher for more skilled workers of either gender. Tables 4 and 5 also suggest that the returns to other wage-determining covariates (i.e. regional, industry, and occupation) show no consistent pattern for separate male and female earnings.

Table 4. Unconditional RIF regression estimates for males

	10th		20th		30th		40th		50th		60th		70th		80th		90th	
Variable	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Constant	6.787	0.104	6.627	0.173	6.464	0.203	6.399	0.064	7.891	0.243	7.942	0.066	8.120	0.159	7.983	0.148	7.316	0.392
Experience	0.058	0.003	0.050	0.004	0.047	0.004	0.050	0.001	0.056	0.003	0.061	0.001	0.060	0.003	0.071	0.003	0.063	0.005
Square of experience	-0.001	0.001	-0.001	0.001	-0.001	0.002	-0.001	0.001	-0.001	0.001	-0.001	0.001	-0.001	0.001	-0.001	0.001	-0.001	0.001
Standard 1	0.035	0.086	0.101	0.193	0.236	0.215	0.310	0.066	-0.933	0.269	-0.909	0.074	-0.812	0.174	-0.381	0.168	-0.110	0.419
Standard 2	0.185	0.087	0.689	0.151	1.016	0.191	1.117	0.062	-0.320	0.240	-0.279	0.064	-0.053	0.157	0.149	0.147	0.519	0.389
Standard 3	-0.024	0.101	0.472	0.155	0.738	0.189	0.877	0.060	-0.523	0.235	-0.249	0.062	-0.038	0.148	0.095	0.132	0.246	0.384
Standard 4	0.233	0.085	0.778	0.143	1.011	0.182	1.228	0.059	-0.266	0.231	-0.247	0.061	-0.077	0.144	0.393	0.130	0.449	0.370
Standard 5	0.001	0.096	0.600	0.151	0.967	0.186	1.173	0.060	-0.262	0.232	-0.194	0.061	0.057	0.145	0.294	0.128	0.492	0.381
Standard 6	-0.057	0.084	0.715	0.144	0.966	0.182	1.220	0.059	-0.123	0.231	-0.067	0.061	0.174	0.145	0.408	0.129	0.561	0.379
Standard 7	0.096	0.081	0.706	0.141	0.946	0.179	1.197	0.059	-0.149	0.229	-0.064	0.060	0.092	0.141	0.411	0.124	0.500	0.377
Standard 8	0.355	0.080	0.902	0.139	1.224	0.179	1.367	0.059	-0.043	0.228	0.015	0.060	0.230	0.140	0.466	0.122	0.670	0.377
Form 1	-0.050	0.097	0.170	0.158	0.772	0.197	1.168	0.063	-0.011	0.241	0.016	0.065	0.106	0.160	0.379	0.145	0.371	0.388
Form 2	0.164	0.085	0.842	0.147	1.234	0.185	1.481	0.060	0.076	0.233	0.198	0.062	0.411	0.148	0.808	0.129	0.734	0.381
Form 3	0.469	0.104	0.986	0.156	1.362	0.196	1.588	0.062	0.181	0.239	0.201	0.064	0.374	0.151	0.471	0.141	0.806	0.389
Form 4	0.528	0.081	1.112	0.141	1.438	0.179	1.657	0.059	0.292	0.228	0.408	0.060	0.579	0.140	0.813	0.123	0.913	0.377
Form 5	0.884	0.142	1.244	0.246	2.807	0.271	2.813	0.073	1.144	0.260	1.032	0.069	1.012	0.161	0.909	0.144	1.444	0.364
Form 6	1.224	0.100	1.609	0.167	1.827	0.196	1.962	0.062	0.591	0.238	0.706	0.063	0.823	0.150	1.184	0.134	1.363	0.385
One year varsity	1.696	0.095	1.946	0.161	2.051	0.183	2.075	0.075	0.493	0.266	0.462	0.071	0.460	0.175	0.532	0.203	1.287	0.496
Two years varsity	1.117	0.092	1.413	0.158	1.535	0.203	1.552	0.066	0.452	0.318	0.483	0.087	0.542	0.221	0.765	0.219	1.214	0.454
Three years varsity	1.484	0.114	1.884	0.188	2.174	0.219	2.440	0.067	0.973	0.249	1.014	0.070	1.291	0.184	2.311	0.195	1.977	0.424
Four years varsity	1.273	0.095	1.893	0.159	2.232	0.198	2.496	0.062	1.104	0.240	1.244	0.064	1.316	0.154	1.532	0.135	1.713	0.391
Postgraduate	1.289	0.105	1.755	0.177	2.393	0.211	2.736	0.065	1.418	0.249	1.614	0.065	1.632	0.163	2.034	0.139	1.962	0.400
Other	0.568	0.093	1.032	0.197	1.412	0.214	1.534	0.066	0.265	0.247	0.365	0.067	0.440	0.153	0.546	0.142	0.831	0.396
Nairobi	0.624	0.037	0.414	0.049	0.387	0.047	0.406	0.012	0.382	0.042	0.380	0.013	0.324	0.039	0.432	0.425	0.480	0.049
Central	0.539	0.029	0.301	0.042	0.286	0.041	0.275	0.010	0.221	0.037	0.243	0.122	0.186	0.035	0.210	0.038	0.291	0.048
Coast	0.506	0.033	0.489	0.045	0.480	0.044	0.470	0.011	0.426	0.040	0.416	0.013	0.344	0.037	0.373	0.042	0.413	0.049
North-eastern	0.121	0.087	0.181	0.108	0.093	0.108	0.024	0.027	0.023	0.094	0.140	0.029	0.171	0.079	0.158	0.092	0.041	0.197
Nyanza	-0.254	0.029	-0.312	0.040	-0.182	0.039	-0.154	0.010	-0.108	0.036	-0.041	0.011	-0.086	0.033	-0.064	0.037	-0.104	0.0473
Rift valley	0.178	0.027	0.069	0.038	0.086	0.036	0.091	0.009	0.084	0.033	0.083	0.010	0.050	0.031	0.066	0.033	0.084	0.042
Western	-0.860	0.040	-0.631	0.049	-0.481	0.049	-0.374	0.012	-0.356	0.043	-0.264	0.013	-0.268	0.038	-0.231	0.041	-0.423	0.053
Community	-0.526	0.044	-0.321	0.083	-0.277	0.076	-0.324	0.019	-0.309	0.065	-0.439	0.020	-0.523	0.057	-0.456	0.065	-0.445	0.088
Agriculture	-0.859	0.049	-0.579	0.088	-0.523	0.082	-0.536	0.021	-0.513	0.072	-0.631	0.023	-0.780	0.065	-0.689	0.074	-0.737	0.096
Manual	-0.421	0.048	-0.181	0.088	-0.156	0.081	-0.193	0.020	-0.128	0.070	-0.247	0.022	-0.384	0.062	-0.416	0.070	-0.307	0.094

(*Continued*)

Table 4. *Continued*

Variable	10th Estimate	10th SE	20th Estimate	20th SE	30th Estimate	30th SE	40th Estimate	40th SE	50th Estimate	50th SE	60th Estimate	60th SE	70th Estimate	70th SE	80th Estimate	80th SE	90th Estimate	90th SE
Wholesale	0.486	0.049	-0.204	0.087	-0.166	0.081	-0.213	0.021	-0.154	0.070	-0.289	0.022	-0.403	0.063	-0.388	0.072	-0.332	0.094
Transportation	0.194	0.044	0.242	0.089	0.231	0.083	0.167	0.021	0.140	0.074	0.203	0.023	-0.108	0.065	-0.141	0.074	0.044	0.096
Administration	0.250	0.067	0.400	0.081	0.433	0.075	0.473	0.021	0.535	0.071	0.669	0.022	0.608	0.063	0.658	0.077	0.547	0.088
Professionals	0.669	0.046	0.562	0.061	0.596	0.063	0.575	0.017	0.490	0.060	0.529	0.019	0.395	0.058	0.465	0.069	0.497	0.079
Technicians	0.422	0.047	0.461	0.057	0.470	0.058	0.481	0.015	0.376	0.052	0.344	0.017	0.200	0.051	0.081	0.062	0.341	0.065
Clerical	0.454	0.064	0.376	0.083	0.285	0.084	0.201	0.020	0.075	0.081	0.021	0.023	-0.113	0.068	-0.105	0.069	0.126	0.090
Service	-0.251	0.046	-0.318	0.054	-0.333	0.056	-0.252	0.015	-0.208	0.053	-0.163	0.017	-0.291	0.051	-0.213	0.062	-0.288	0.065
Farm, fisheries	-0.267	0.043	-0.578	0.055	-0.487	0.055	-0.465	0.015	-0.465	0.052	-0.323	0.017	-0.406	0.050	-0.390	0.059	-0.510	0.067
Craftsmen	0.251	0.044	-0.381	0.061	-0.205	0.059	-0.157	0.016	-0.142	0.053	0.020	0.017	-0.124	0.049	-0.156	0.055	-0.238	0.065
Elementary occupation	-0.259	0.034	-0.333	0.043	-0.328	0.044	-0.324	0.012	-0.354	0.041	-0.308	0.013	-0.387	0.039	-0.454	0.046	-0.411	0.051

Notes: For education categories, workers with no formal education (none) constitute the base group; for regional categories, eastern province is the base group; for occupation categories, finance workers are the base group; and for industry categories, machine operators are the base group.

Table 5. Unconditional RIF regression estimates for females

	10th		20th		30th		40th		50th		60th		70th		80th		90th	
Variable	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Constant	5.649	0.302	5.042	0.129	5.020	0.148	7.382	0.163	7.834	0.068	7.947	0.037	7.855	0.106	7.768	0.072	7.000	0.106
Experience	0.022	0.009	0.034	0.004	0.023	0.003	0.017	0.003	0.023	0.001	0.036	0.004	0.037	0.002	0.044	0.002	0.065	0.003
Square of experience	-0.001	0.002	-0.001	0.001	-0.002	0.001	-0.001	0.001	-0.001	0.006	-0.003	0.001	-0.003	0.004	-0.001	0.002	-0.001	0.001
Standard 1	1.731	0.276	2.668	0.116	2.692	0.129	0.256	0.149	0.489	0.061	0.333	0.032	0.529	0.096	0.439	0.059	0.719	0.117
Standard 2	1.597	0.222	1.932	0.108	2.265	0.129	0.157	0.128	0.255	0.052	0.379	0.027	0.389	0.069	0.369	0.049	0.580	0.086
Standard 3	1.669	0.221	2.301	0.106	2.498	0.126	0.286	0.123	0.302	0.048	0.331	0.024	0.490	0.068	0.552	0.048	0.739	0.869
Standard 4	1.494	0.220	2.107	0.103	2.299	0.121	0.191	0.120	0.358	0.048	0.375	0.024	0.525	0.069	0.639	0.050	0.844	0.086
Standard 5	1.773	0.221	2.279	0.102	2.559	0.122	0.387	0.121	0.367	0.048	0.468	0.025	0.517	0.068	0.562	0.050	0.695	0.089
Standard 6	1.524	0.206	2.022	0.101	2.365	0.120	0.092	0.118	0.214	0.047	0.385	0.024	0.436	0.063	0.551	0.044	0.926	0.082
Standard 7	1.920	0.198	2.497	0.099	2.820	0.119	0.595	0.117	0.667	0.046	0.676	0.023	0.793	0.062	0.812	0.044	1.080	0.081
Standard 8	1.843	0.196	2.441	0.098	2.670	0.118	0.400	0.115	0.512	0.045	0.617	0.022	0.694	0.060	0.678	0.042	0.909	0.078
Form 1	1.554	0.261	2.025	0.114	2.323	0.128	0.156	0.125	0.415	0.050	0.435	0.026	0.524	0.070	0.733	0.049	1.385	0.092
Form 2	1.334	0.256	2.181	0.110	2.577	0.125	0.371	0.123	0.505	0.049	0.558	0.024	0.705	0.068	0.930	0.049	1.138	0.081
Form 3	2.220	0.231	2.533	0.114	2.738	0.130	0.433	0.129	0.415	0.051	0.346	0.027	0.466	0.075	0.805	0.054	1.371	0.097
Form 4	2.200	0.196	2.694	0.099	3.017	0.119	0.749	0.116	0.871	0.046	1.045	0.022	1.275	0.061	1.315	0.043	1.577	0.082
Form 5	1.177	0.261	1.241	0.133	2.221	0.189	-0.355	0.194	0.942	0.062	1.010	0.030	1.122	0.078	1.022	0.051	1.035	0.086
Form 6	2.971	0.257	3.488	0.115	3.639	0.130	1.416	0.134	1.425	0.054	1.491	0.028	1.748	0.076	1.765	0.057	2.354	0.106
One year varsity	3.760	0.275	4.027	0.136	4.351	0.147	1.687	0.151	1.797	0.068	1.751	0.038	1.699	0.131	1.309	0.071	1.295	0.105
Two years varsity	3.601	0.330	4.023	0.122	3.963	0.142	1.430	0.146	1.888	0.074	1.825	0.041	1.952	0.099	1.776	0.060	1.556	0.098
Three years varsity	2.804	0.219	3.184	0.107	3.027	0.128	1.141	0.182	1.142	0.073	1.349	0.379	1.879	0.096	1.697	0.069	2.911	0.107
Four years varsity	3.132	0.264	3.582	0.113	3.780	0.132	1.750	0.133	1.833	0.055	1.999	0.029	2.360	0.084	2.509	0.052	2.638	0.095
Postgraduate	3.148	0.224	4.052	0.128	4.336	0.138	1.868	0.144	2.333	0.059	2.618	0.030	2.864	0.088	3.004	0.065	3.245	0.101
Other	2.755	0.211	3.078	0.114	3.092	0.130	0.726	0.144	0.761	0.056	0.894	0.029	1.179	0.079	1.165	0.054	2.001	0.087
Nairobi	0.441	0.090	0.264	0.033	0.249	0.030	0.330	0.032	0.338	0.015	0.340	0.008	0.283	0.027	0.336	0.021	0.411	0.029
Central	0.425	0.074	0.280	0.029	0.223	0.028	0.258	0.028	0.260	0.013	0.238	0.009	0.137	0.025	0.122	0.018	0.115	0.031
Coast	0.108	0.074	0.050	0.035	0.121	0.034	0.289	0.036	0.282	0.017	0.296	0.010	0.228	0.033	0.231	0.027	0.454	0.044
North-eastern	0.377	0.119	-0.015	0.058	0.032	0.102	-0.727	0.092	0.054	0.054	0.136	0.029	0.449	0.084	0.860	0.061	1.268	0.046
Nyanza	-0.812	0.080	-0.736	0.029	-0.253	0.028	-0.345	0.028	-0.263	0.014	-0.177	0.008	-0.083	0.026	-0.109	0.019	0.027	0.029
Rift valley	-0.001	0.069	-0.004	0.027	0.109	0.026	0.173	0.028	0.171	0.013	0.174	0.007	0.077	0.024	0.034	0.018	0.169	0.028
Western	-0.757	0.102	-0.013	0.041	-0.079	0.038	-0.849	0.038	-0.847	0.017	-0.691	0.010	-0.729	0.031	-0.503	0.023	-0.217	0.036
Community	-0.152	0.147	-0.079	0.051	-0.351	0.054	-0.371	0.057	-0.748	0.026	-0.923	0.016	-0.745	0.052	-0.615	0.042	-0.476	0.044
Agriculture	-0.236	0.179	-0.859	0.060	-0.969	0.062	-0.874	0.064	-0.122	0.029	-0.139	0.018	-0.935	0.057	-0.833	0.045	-0.606	0.052
Manual	-0.209	0.217	0.047	0.075	-0.113	0.072	-0.426	0.086	-0.438	0.040	-0.415	0.024	-0.211	0.059	-0.092	0.045	-0.195	0.051

(*Continued*)

Table 5. *Continued*

	10th		20th		30th		40th		50th		60th		70th		80th		90th	
Variable	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Wholesale	−0.199	0.158	−0.091	0.054	−0.298	0.057	−0.248	0.060	−0.607	0.028	−0.779	0.017	−0.535	0.055	−0.322	0.044	−0.243	0.047
Transportation	0.303	0.167	0.487	0.072	0.077	0.070	0.203	0.075	−0.258	0.033	−0.534	0.020	−0.221	0.064	−0.050	0.050	0.313	0.075
Administration	0.224	0.201	0.414	0.084	0.786	0.084	0.765	0.109	0.959	0.044	1.060	0.026	0.794	0.079	1.226	0.052	1.855	0.048
Professionals	0.021	0.201	0.458	0.072	0.723	0.076	0.751	0.104	0.566	0.044	0.440	0.024	0.319	0.070	0.254	0.041	0.526	0.052
Technicians	−0.162	0.195	0.139	0.069	0.572	0.072	0.595	0.100	0.529	0.043	0.343	0.025	0.223	0.067	0.215	0.039	0.417	0.051
Clerical	−0.259	0.224	0.092	0.075	0.396	0.076	0.516	0.102	0.335	0.044	0.158	0.024	0.057	0.070	0.067	0.042	0.195	0.051
Service	−0.825	0.191	−0.499	0.068	−0.247	0.071	−0.258	0.099	−0.384	0.042	−0.340	0.023	−0.352	0.067	−0.340	0.038	0.081	0.050
Farm, fisheries	−0.821	0.208	−0.561	0.072	−0.246	0.075	−0.266	0.101	−0.376	0.044	−0.488	0.024	−0.459	0.069	−0.366	0.040	−0.001	0.055
Craftsmen	−0.731	0.201	−0.216	0.074	−0.027	0.078	0.062	0.104	0.052	0.045	−0.108	0.025	−0.230	0.071	−0.365	0.044	0.091	0.051
Elementary occupation	−0.797	0.302	−0.687	0.065	−0.451	0.069	−0.410	0.097	−0.525	0.068	−0.553	0.023	−0.550	0.065	−0.596	0.037	−0.257	0.106

Notes: For education categories, workers with no formal education (none) constitute the base group; for regional categories, eastern province is the base group; for occupation categories, finance workers are the base group; and for industry categories, machine operators are the base group.

However, Tables 4 and 5 reveal a notable difference in the return to human capital skills between male and female workers. Specifically, we find that the returns to all categories of human capital skills, i.e. primary schooling, secondary schooling, and university education, are generally higher for female workers reported in Table 5 than for male workers reported in Table 4. These are puzzling results and warrant further discussion.

A possible explanation for this may be differences in returns to human capital for skilled labor relative to unskilled labor in the base group for either gender. In particular, a comparison of Tables 4 and 5 suggests that, relative to male workers in Table 4, the returns to human capital are relatively higher for skilled female workers reported in Table 5 than for unskilled female workers in the base group. A likely explanation for this may be traditional attitudes, which in the past have steered women toward spending more time in household production than in wage employment (Wanjala & Were, 2009). Fewer women would hence join the labor force, and, with a smaller proportion of women in the labor force, it is conceivable that the returns to human capital would be higher for the few skilled female workers relative to the unskilled female workers in the base group.[6]

However, while a discussion on the size of the coefficients on the wage-determining covariates (i.e. age, education, industry, occupation, and region) and their relative impact on separate wages for male and female workers is important and warrants detailed discussion, such an analysis falls outside the scope of this paper. Moreover, the previous studies cited in Section 1 address these issues.

Our principal focus in this paper is to explore the partial impact of covariates and identify the covariates with the strongest influence on the magnitude of the gender pay gap along the entire unconditional wage distribution. Here, we decompose the gender pay gap at each quantile into two components: composition and wage structure effects. Relative differences in the two effects and the impact of each covariate on composition and wage structure effects allow us to identify the impact of each covariate on the magnitude of the gender pay gap across the earnings distribution.

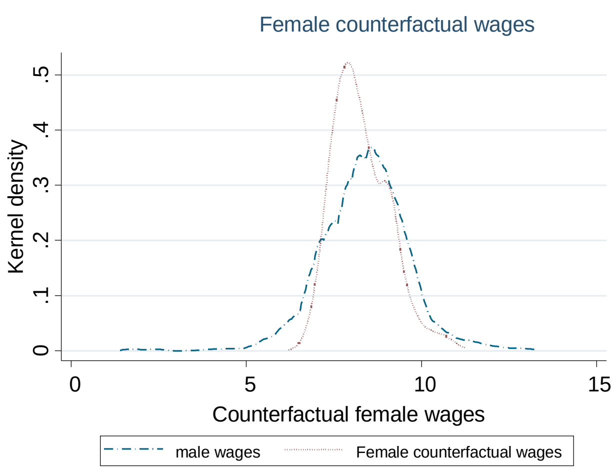

Figure 2. Composition effects.

4.2 Female Counterfactual Wage Distributions

Before we decompose the gender pay gaps into portions attributable to composition and wage structure effects, we first compute a female counterfactual wage that would prevail if female workers possessed the same distribution of observed and unobserved characteristics as males at all quantiles. The female counterfactual wage is then used to compute the male–female wage gap. Kernel density wage estimates of the female counterfactual wage distribution alongside the kernel density wage estimate for male workers are reported in Figure 2.

From Figure 2 we can see that the female counterfactual wage density is more or less "bell-shaped" and has a relatively higher peak than the male wage density. The female counterfactual wage density also has relatively less of a rightward translation, except at the very bottom of the wage distributions. The gap between the two densities captures gender differences in observed characteristics or composition effects.

By contrast, visual estimates of the kernel density wage estimate for female workers and the kernel density wage estimate for female counterfactual earnings are reported in Figure 3.

From Figure 3, we can see that the female counterfactual wage density has a relatively rightward translation throughout much of the lower, middle, and upper half of the earnings distributions. The gap between the two densities reflects gender differences in the return to characteristics or wage structure effects.

Care should be exercised in interpreting the counterfactual wage densities reported in Figures 2 and 3, derived from the Machado & Mata (2005) approach and from which the FFL (2009) model is drawn. This is because the relative position of female workers in the wage distributions may have changed, compared with the original distribution, and the estimates may not accurately predict discriminatory occurrences for females (Del Río *et al.*, 2011). From a distribution point of view, therefore, it is conceivable that Figures 2 and 3 may not truly depict discriminatory experiences for female workers and should be interpreted with caution. Nonetheless, our use of the FFL (2009) approach provides a close approximation of gender wage discrimination in Kenya.

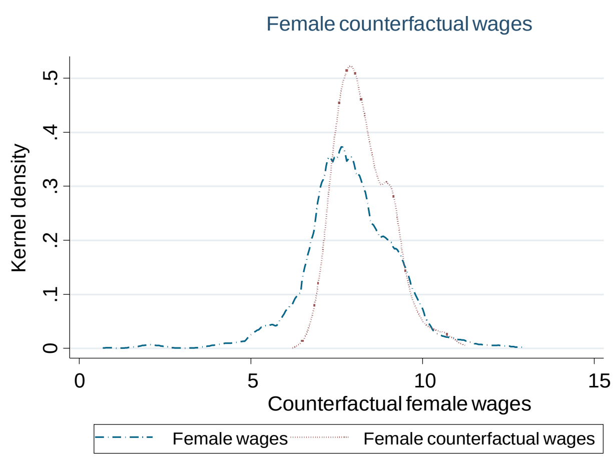

Figure 3. Wage structure effects.

4.3 Decomposing the Gender Wage Gap Into Composition and Wage Structure Effects

The wage structure and composition effects associated with the X variable correspond to $\bar{X}_f(\hat{\beta}_c - \hat{\beta}_f)$ and $\bar{X}_m(\hat{\beta}_m - \bar{X}_f(\hat{\beta}_c)$, respectively, as specified in Equation (8). The coefficients $\hat{\beta}_m$ and $\hat{\beta}_f$ at each quantile are unconditional quantile regression estimates for male and female workers reported in Tables 4 and 5, respectively.

Table 6. Wage structure effects at all quantiles

	10th	20th	30th	40th	50th	60th	70th	80th	90th
Variable									
Experience	0.72	0.000	0.000	0.660	0.660	0.500	0.460	0.001	0.000
Square of experience	0	0.000	0.000	0	0	0.800	0.800	0.000	0.000
Education									
Standard 1	−0.012	−0.018	−0.017	0.000	−0.010	−0.009	−0.009	−0.006	−0.005
Standard 2	−0.024	−0.021	−0.021	0.016	−0.010	−0.011	−0.008	−0.004	−0.001
Standard 3	−0.046	−0.049	−0.047	0.016	−0.022	−0.016	−0.014	−0.012	−0.013
Standard 4	−0.054	−0.057	−0.055	0.044	−0.027	−0.027	−0.026	−0.011	−0.017
Standard 5	−0.076	−0.072	−0.068	0.034	−0.027	−0.028	−0.020	−0.011	−0.009
Standard 6	−0.093	−0.077	−0.083	0.067	−0.020	−0.027	−0.015	−0.008	−0.022
Standard 7	−0.222	−0.219	−0.229	0.073	−0.100	−0.090	−0.085	−0.049	−0.071
Standard 8	−0.281	−0.291	−0.274	0.183	−0.105	−0.114	−0.087	−0.040	−0.045
Form 1	−0.035	−0.041	−0.034	0.022	−0.009	−0.009	−0.009	−0.008	−0.022
Form 2	−0.051	−0.059	−0.059	0.049	−0.019	−0.016	−0.013	−0.005	−0.018
Form 3	−0.026	−0.023	−0.021	0.017	−0.004	−0.002	−0.001	−0.005	−0.008
Form 4	−0.565	−0.534	−0.534	0.307	−0.196	−0.215	−0.235	−0.169	−0.224
Form 5	−0.000	0.000	0.001	0.003	0.000	0.000	0.000	0.000	0.000
Form 6	−0.033	−0.036	−0.034	0.010	−0.016	−0.015	−0.017	−0.011	−0.019
One year university	−0.002	−0.002	−0.002	0.000	−0.001	−0.001	−0.001	0.000	0.000
Two years university	−0.007	−0.008	−0.007	0.000	−0.004	−0.004	−0.004	−0.003	−0.001
Three years university	−0.008	−0.008	−0.005	0.008	−0.001	−0.002	−0.004	0.004	−0.006
Four years university	−0.032	−0.029	−0.026	0.013	−0.012	−0.012	−0.017	−0.016	−0.016
Postgraduate	−0.017	−0.021	−0.017	0.008	−0.008	−0.009	−0.011	−0.009	−0.011
Other	−0.031	−0.029	−0.023	0.011	−0.007	−0.007	−0.010	−0.009	−0.016
Region									
Nairobi	0.019	0.015	0.014	0.008	0.005	0.004	0.004	0.009	0.007
Central	0.017	0.003	0.009	0.003	−0.006	0.001	0.007	0.013	0.026
Coast	0.045	0.049	0.040	0.020	0.016	0.013	0.013	0.015	−0.005
North-eastern	−0.003	0.002	0.001	0.008	0.000	0.000	−0.003	−0.007	−0.012
Nyanza	0.102	0.078	0.064	0.035	0.028	0.013	−0.001	0.008	−0.024

(*Continued*)

Table 6. *Continued*

	10th	20th	30th	40th	50th	60th	70th	80th	90th
Rift valley	0.038	0.015	−0.005	−0.017	−0.018	−0.019	−0.006	0.007	−0.018
Western	−0.008	0.028	0.044	0.035	0.036	0.032	0.034	0.020	−0.015
Industry									
Community	−0.158	−0.102	0.031	0.020	0.185	0.204	0.094	0.067	0.013
Agriculture	0.094	0.069	0.111	0.084	0.151	0.126	0.039	0.036	−0.033
Manual	−0.006	−0.006	−0.001	0.001	0.009	0.005	−0.004	−0.009	−0.003
Wholesale	0.175	−0.029	0.034	0.009	0.116	0.125	0.034	−0.017	−0.023
Transportation	−0.002	−0.005	0.003	−0.001	0.008	0.684	0.002	−0.002	−0.054
Occupation									
Administration	0.001	0.000	−0.008	−0.006	−0.009	−0.009	−0.004	−0.012	−0.028
Professionals	0.045	0.007	−0.008	−0.012	−0.005	0.006	0.005	0.015	−0.002
Technicians	0.063	0.035	−0.011	−0.012	−0.017	0.000	−0.002	−0.014	−0.008
Clerical	0.034	0.015	−0.006	−0.018	−0.015	−0.007	−0.010	−0.009	−0.038
Service	0.084	0.027	−0.013	0.001	0.015	0.026	0.009	0.018	−0.054
Farm, fisheries and wildlife	0.033	−0.002	−0.041	−0.034	−0.015	0.028	0.009	−0.004	−0.087
Craftsmen	0.066	−0.011	−0.012	−0.015	−0.013	0.009	0.007	0.014	−0.022
Elementary occupations	0.189	0.125	0.043	0.030	0.060	0.086	0.057	0.050	−0.054
Residual	0.373	−0.784	−0.799	−1.213	−0.184	−1.636	−0.601	0.229	−0.533

Notes: The results are rounded to three digits after the decimal. If the effect is less than 0.0005, it is reported as “0.000” after rounding. For education categories, workers with no formal education (none) constitute the base group; for regional categories, eastern province is the base group; for occupation categories, finance workers constitute the base group; and for industry categories, machine operators are the base group.

Table 7. Composition effects at all quantiles

	10th	20th	30th	40th	50th	60th	70th	80th	90th
Variable									
Experience	0.116	0.000	0.001	0.100	0.112	0.122	0.120	0.000	0.000
Square of experience	−0.084	0.000	0.003	−0.084	−0.084	−0.084	−0.084	0.000	0.000
Education									
Standard 1	−0.000	0.000	−0.018	−0.001	0.003	0.003	0.002	0.001	0.000
Standard 2	−0.001	−0.002	−0.024	−0.003	0.001	0.001	0.000	0.000	−0.001
Standard 3	0.000	−0.005	−0.048	−0.001	0.001	0.000	0.000	0.000	−0.002
Standard 4	0.001	0.004	−0.050	0.006	−0.001	−0.001	−0.000	0.002	0.002
Standard 5	−0.000	−0.002	−0.072	−0.004	0.001	0.001	−0.000	−0.001	−0.002
Standard 6	0.000	−0.002	−0.085	−0.004	0.000	0.000	−0.001	−0.001	−0.002
Standard 7	0.001	0.006	−0.220	0.011	−0.001	−0.001	0.001	0.003	0.005
Standard 8	0.002	0.007	−0.264	0.011	−0.000	0.000	0.002	0.003	0.005
Form 1	0.000	−0.001	−0.039	−0.008	0.000	−0.000	−0.001	−0.003	−0.003
Form 2	0.000	0.002	−0.056	0.003	0.000	0.000	0.001	0.002	0.001
Form 3	0.002	0.004	−0.015	0.006	0.001	0.001	0.001	0.002	0.003
Form 4	−0.016	−0.034	−0.576	−0.050	−0.009	−0.012	−0.017	−0.024	−0.027
Form 5	0	0.000	0.001	0	0	0	0	0.000	0.000
Form 6	0.012	0.016	−0.016	0.020	0.006	0.007	0.008	0.011	0.014
One year university	0.002	0.002	0.000	0.002	0.002	0.000	0.000	0.000	0.001
Two years university	0	0.000	−0.007	0	0	0	0	0.000	0.000
Three years university	0.003	0.004	−0.001	0.005	0.002	0.002	0.003	0.005	0.004
Four years university	0.013	0.019	−0.004	0.025	0.011	0.012	0.013	0.015	0.017
Postgraduate	0.008	0.011	−0.003	0.016	0.009	0.010	0.010	0.012	0.011
Other	−0.002	−0.003	−0.028	0.005	−0.001	−0.001	−0.001	−0.002	−0.002
Region									
Nairobi	−0.011	−0.007	0.007	−0.007	−0.007	−0.007	−0.006	−0.008	−0.009
Central	−0.017	−0.009	0.001	−0.008	−0.007	−0.008	−0.006	−0.007	−0.009
Coast	0.006	0.005	0.045	0.005	0.005	0.005	0.004	0.004	0.006
North-eastern	0.001	0.001	0.001	0.000	0.000	0.001	0.001	0.001	0.000
Nyanza	0.005	0.006	0.067	0.003	0.002	0.001	0.002	0.001	0.002

(*Continued*)

Table 7. *Continued*

	10th	20th	30th	40th	50th	60th	70th	80th	90th
Rift valley	0.005	0.002	−0.002	0.003	0.002	0.002	0.001	0.002	0.002
Western	−0.013	−0.009	0.037	−0.006	−0.005	−0.004	−0.004	−0.003	−0.006
Industry									
Community	0.065	0.040	0.066	0.040	0.038	0.054	0.065	0.057	0.055
Agriculture	−0.009	−0.006	0.106	−0.005	−0.005	−0.006	−0.007	−0.007	−0.007
Manual	−0.052	−0.022	−0.021	−0.023	−0.016	−0.031	−0.048	−0.052	−0.038
Wholesale	−0.047	0.019	0.049	0.020	0.014	0.028	0.039	0.038	0.032
Transportation	0.016	0.020	0.022	0.014	0.012	0.017	−0.009	−0.012	0.004
Occupation									
Administration	0.004	0.006	−0.001	0.007	0.008	0.010	0.009	0.010	0.008
Professionals	0.004	0.003	−0.005	0.003	0.003	0.003	0.002	0.003	0.003
Technicians	−0.002	−0.002	−0.013	−0.002	−0.002	−0.001	−0.001	0.000	−0.001
Clerical	−0.013	−0.011	−0.014	−0.006	−0.002	−0.001	0.003	0.003	−0.004
Service	0.015	0.018	0.007	0.015	0.016	0.009	0.017	0.012	0.017
Farm, fisheries and wildlife	0.015	0.014	−0.029	0.011	0.011	0.008	0.010	0.009	0.012
Craftsmen	0.038	−0.006	−0.015	−0.002	−0.002	0.000	−0.002	−0.002	−0.004
Elementary occupations	−0.001	−0.001	0.042	−0.001	−0.001	−0.001	−0.002	−0.002	−0.002
Residual	0.374	0.410	−0.704	0.330	0.292	0.238	0.232	0.330	0.340

Notes: The results are rounded to three digits after the decimal. If the effect is less than 0.0005, it is reported as "0.000" after rounding. For education categories, workers with no formal education (none) constitute the base group; for regional categories, eastern province is the base group; for occupation categories, finance workers constitute the base group; and for industry categories, machine operators are the base group.

The coefficient $\hat{\beta}_c$ is derived from unconditional regression estimates based on the counterfactual earnings of female workers. That is, $\hat{\beta}_c$ assumes that the returns to the distribution of earnings for female workers are as if they possessed the same distribution of measured and unmeasured characteristics as male workers. Since these earnings assume that male returns to labor market characteristics apply for women, then $\hat{\beta}_c$ is comparable to $\hat{\beta}_m$ and for this reason estimates of the coefficients for $\hat{\beta}_c$ are not reported.

With regard to the wage structure effects, if the return to an *X* variable is higher for males than for females, then $\hat{\beta}_c > \hat{\beta}_f$ and the wage structure effect contributed by this variable would be positive, indicating potential discrimination (or bias) against women. On the other hand, if the return to an *X* variable is higher for women than for men, then $\hat{\beta}_f > \hat{\beta}_c$ and the wage structure effect contributed by this variable would be negative. When the explanatory variable is a dummy variable, then the estimate would indicate that the contribution of the specific dummy variable is relative to the designated base group. With regard to the composition effect, $(\bar{X}_m(\hat{\beta}_m - \bar{X}_f\hat{\beta}_c$) since $\hat{\beta}_c$ is comparable to $\hat{\beta}_m$, the composition effect associated with an *X* variable captures the gender earnings gap attributed to the gender endowment differences in *X*, assuming the same returns for men and women. The residuals correspond to $\hat{R}^S_\tau$ and $\hat{R}^c_\tau$ for the wage structure and composition effects, respectively.

Results of the decomposition of the gender wage gap that yield coefficients for the wage structure and composition effects, at all quantiles, are reported in Tables 6 and 7, respectively. Indeed, Tables 6 and 7 provide useful insights into the role each individual covariate plays in determining the magnitude of the gender pay gap along the entire unconditional earnings distribution; however, the information reported may be too detailed, dense, and overwhelming.

We therefore simplify the analysis by examining the cumulative effect of all the explanatory variables within each of the categories: education, region, industry, and occupation. For example, for the industry control, we sum the total effects of the explanatory dummy variables, community and social services, agriculture, manual, wholesale, and transportation reported in Tables 6 and 7 to find the cumulative effect of the industry control on wage structure and composition effects, respectively, at each quantile. Likewise, we carry out the same procedure for all explanatory dummy variables in the controls for education, region, and occupation. By doing so, we obtain the cumulative effects of each of the dummy variables within education, region, industry, and occupation on the wage structure and composition effects at all quantiles.

For instance, the total composition effect for the industry category at the 10th quantile in Table 7 is computed by totaling the impact of each covariate in the industry category, i.e. community and social services, agriculture, manual, wholesale and retail trade, and transportation (i.e. 0.065–0.009–0.052–0.047 þ 0.016) at the 10th quantile to yield 20.027. This coefficient is reported in row 5 of column 1 in Table 8.

Table 8 also reports coefficients from all the other categories. The columns in Table 8 thus add up to the corresponding individual covariates in each of the categories of education, region, industry, and occupation reported in Tables 6 and 7 and hence do not fundamentally alter our results. Our analysis from Table 8 begins by examining covariates with the strongest influence on the magnitude of the gender pay gap where the gap is largest, i.e. at the 20th and 30th quantiles, then at the 80th and 90th quantiles, and then in the middle of the distributions at the 40th, 50th, 60th, and 70th quantiles.

Table 8. Total effects of wage structure and composition effects at all quantiles

	10th	20th	30th	40th	50th	60th	70th	80th	90th
Total change	−0.002	−2.382	−3.816	1.795	2.189	0.998	2.267	−0.101	−1.773
Wage structure	−0.068	−2.560	−1.463	1.679	2.072	0.859	2.122	−0.173	−1.858
Composition	0.066	0.178	−2.353	0.116	0.117	0.139	0.145	0.072	0.085
Composition effects									
Experience and its square	0.032	0.000	0.000	0.016	0.028	0.038	0.036	0.000	0.000
Education	0.025	0.026	−1.525	0.039	0.025	0.022	0.025	0.025	0.024
Region	−0.024	−0.011	0.156	−0.010	−0.010	−0.010	0.008	−0.010	−0.014
Industry	−0.027	0.051	0.222	0.046	0.043	0.062	0.040	0.024	0.046
Occupation	0.060	0.112	−1.206	0.025	0.031	0.027	0.036	0.033	0.029
Wage structure effects									
Experience and its square	0.720	0.000	0.000	0.660	0.660	1.300	1.260	0.001	0.000
Education	−1.615	−1.594	−1.577	0.880	0.881	−0.614	0.586	−0.372	−0.524
Region	0.209	0.190	0.167	0.092	0.061	0.044	0.040	0.065	−0.041
Industry	0.103	−0.073	0.178	0.113	0.469	−0.010	0.165	0.075	−1.000
Occupation	0.515	−1.083	−0.231	−0.066	0.001	0.139	0.071	0.058	−0.293
Total effects									
Experience and its square	0.752	0.000	0.000	0.676	0.688	1.338	1.296	0.001	0.000
Education	−1.590	−1.568	−3.012	0.919	0.906	−0.592	0.611	−0.347	−0.500
Region	0.185	0.179	0.323	0.082	0.051	0.340	0.040	0.055	−0.055
Industry	0.067	−0.022	0.400	0.159	0.512	0.052	0.24	0.099	−1.006
Occupation	0.574	−0.971	−1.491	−0.041	0.032	0.079	0.107	−0.009	−0.264

4.4 The Gender Pay Gap at the Lower End of the Wage Distributions

From Table 8, we find that that composition effects, that is gender differences in characteristics, account for the largest share of the gender pay gap at the 20th quantile: the coefficient in row 2 on the 20th quantile, i.e. 0.178, is relatively larger for composition effects than for wage structure effects (22.650). This begs the question, which covariates have the greatest influence on composition effects at the 20th quantile?

From rows 5, 6, and 3 of Table 8, we find that the occupation, industry, and education categories exert the strongest influence, respectively, on the gender pay gap at the 20th quantile. We next identify the covariates within the occupation, industry, and education categories with the strongest influence on the gender pay gap. To do so, can we use either Table 6 (wage structure effects) or Table 7 (composition effects). But since composition effects drive the gender pay gap at the 20th quantile, we identify covariates within industry, occupation, and education categories that drive composition effects at the 20th quantile from Table 7.

From Table 7, we find that the covariates within the industry category that have the greatest effect on the gender pay gap at the 20th quantile are community and social services, wholesale and retail trade, and transportation: these covariates have the largest coefficients relative to the other variables in this category, i.e. 0.040, 0.019, and 0.020, respectively.7 Occupation covariates with the largest effect on the gender pay gap are service, farm, fisheries and wildlife, administration, and professionals. Education covariates with the strongest influence on the gender pay gap are higher education covariates, 6 years of high school, 4 years of university, and postgraduate education.

Interestingly, our findings at the 20th quantile match the econometric decomposition results reported in Table 8 and the descriptive statistics reported in Table 1. In particular, industries and occupations with disproportionately more females have greater gender wage inequality due to gender differences in characteristics. Here, female-dominated industries and occupations require fewer skills and pay lower wages. Moreover, it may be that even in female-dominated occupations (service, farm, fisheries and wildlife) and industries (community and social services and wholesale and retail trade), males possess relatively more skills and would command higher wages. In addition, there are relatively more males in transportation, an industry requiring more skills, adding to the explanation of the larger gender wage gap at the 20th quantile.

At the 30th quantile, however, we find from Table 8 that gender differences in wage structure effects, i.e. gender differences in the returns to attributes, exert the strongest influence on the gender pay gap: from row 1, the coefficient -1.463 is larger for wage structure effects than for composition effects, -2.353 in row 2. We also find from rows 4 and 5 of Table 8 that the region and industry categories exert the strongest influence on the gender pay gap at the 30th quantile. To find the covariates within the region and occupation categories that have the strongest influence on the gender pay gap at the 30th quantile, we turn to Table 6. Here, we find the covariates within the region category are Coast, Nyanza, and Western provinces. The covariates within the industry category are community and social services, wholesale and retail trade, and agriculture.

4.5 The Gender Pay Gap at the Upper End of the Wage Distributions

We next identify categories, and covariates within these categories, that drive the gender pay gap at the upper end of the earnings distribution, i.e. at the 80th and 90th quantiles. The results are noteworthy. In particular, from Table 8, we find that composition effects play the most important role in widening the gender pay gap at the 80th and 90th quantiles. We also find (from Table 7) that the industry category (with covariates community and social services, wholesale and retail trade, and transportation) and the occupation category (with covariates service, farm, fisheries and wildlife, administration, and professionals) exert the largest influence on the gender pay gap at the 80th and 90th quantiles. The education covariates with the most significant impact are higher education covariates, i.e. 4 years of university education and postgraduate education.

4.6 The Gender Pay Gap in the Middle of the Wage Distributions

In the middle of the distributions, i.e. between the 40th and the 70th quantiles, we find interesting results that differ from those at the lower and upper ends of the wage distributions. Specifically, we find that wage structure effects, i.e. gender differences in the return to attributes, account for the largest share of the gender pay gap. This finding provides evidence of wage discrimination against women in the middle of the wage distributions. Also, we find a consistent pattern: the same industry, occupation, and education covariates that exert the greatest impact on the gender pay gap at the lower and upper ends of the wage distributions exert the greatest impact on the gender pay gap in the middle of the distributions as well.

5. Conclusions

The conventional literature on gender wage inequality in Africa typically attributes the causes of the male–female wage gap to either gender differences in characteristics or the return to characteristics. We go beyond the literature by using two juxtaposed arguments. First, we suggest that gender wage determination in Africa, and Kenya in particular, may be altered by a sorting mechanism that places workers into various occupations and industries according to skill and gender; consequently, wage-determining covariates (swayed by traditional factors) may impact gender earnings differently, with some covariates exerting a relatively stronger influence on the magnitude of the gender pay gap than others. We explore the partial contribution of each individual covariate on the magnitude of the gender pay gap utilizing RIF quantile regressions corrected for endogeneity. Second, and following from the first, RIF regression estimates are generated from unconditional wage distributions, raising the possibility that our coefficient estimates may have more practical policy relevance. We use data from the 2004–2005 KIHBS.

Our results are remarkable and novel. First, we find a relatively larger gender pay gap at the lower end of the wage distributions between the 20th and 30th quantiles. The gap decreases in the middle of the distributions, between the 40th and 70th quantiles, and then grows at the upper end, between the 80th and 90th quantiles. Our findings of a relatively larger gender pay gap at the lower end of the distributions match those of Agesa *et al*. (2009) and Ntuli (2009).

Importantly, our econometric decomposition results reinforce the descriptive statistics from the data in showing that a confluence of twin factors, i.e. skill level and the proportion of each gender in various industries and occupations, influences the magnitude of the gender pay gap at its largest values. Here, we find that a sorting mechanism may place workers into various industries and occupations according to skill and gender: industries and occupations which require fewer skills have disproportionately more female workers; industries and occupations which require more skills have disproportionally more male workers. Our decomposition results verify this hypothesis and shed new light on the relative significance of individual wage-determining covariates on the magnitude of the gender pay gap.

Particularly, we find that at the 20th quantile, gender differences in characteristics enlarge the gender pay gap due to industry, occupation, and education covariates. The industry covariates are community and social services, wholesale and retail trade, and transportation. The occupation covariates are service, fisheries and wildlife, professional, administration, and farming. These occupations and industries are characterized by two features: either they have a large share of female workers, require fewer skills, and pay lower wages (i.e. community and social services, and wholesale and retail trade, services, farm, fisheries and wildlife) or they are male dominated, require more skills, and pay higher wages (i.e. transportation, professional, and administration). The human capital covariates are higher education variables (i.e. 6 years of high school, 4 years of university education, and postgraduate education), mostly characterized by a higher share of male workers. At the 30th quantile, we find that gender differences in the return to attributes widen the gender pay gap due to industry covariates similar to those at the 20th quantile and by region covariates (Coast, Nyanza, and Western provinces).

In the middle of the wage distributions, however, i.e. between the 40th and 70th quantiles, we find that gender differences in return to attributes enlarge the gender pay gap due to industry and higher education covariates similar to those at the 20th quantile. The dominance of differences in the return to characteristics in the middle of the distributions provides evidence of gender wage discrimination in this range.

At the top of the wage distributions, i.e. at the 80th and 90th quantiles, we find that gender differences in characteristics expand the gender pay gap due to the same higher education, industry, and occupation covariates at lower quantiles.

Taken together, our findings augment current studies on gender wage inequality in Africa by uncovering the underlying covariates that widen the gender pay gap along the unconditional wage distribution. Arguably, these are the most significant findings of this paper, and they also have policy implications. Here, our findings may provide an effective tool that can be used to design policies aimed at mitigating gender wage inequality. For example, to weaken the gender pay gap at the top and bottom of the wage distributions, increasing skills for women, particularly higher education skills, complemented by affirmative action policies to increase the proportion of women in male-dominated occupations and industries, may be appropriate strategies. However, a viable strategy to lessen gender wage discrimination in the middle of the distributions may be to discard laws biased against women, e.g. denying house allowance payments for married women.

Finally, a caveat: while our study considers the covariates that influence the unconditional gender wage gap at a particular point in time, a similar study that considers the covariates that may influence the unconditional gender pay gap over time offers fertile ground for future research.

Notes

[1] Ntuli (2009) uses quantile regressions which have also found widespread application in both developed and developing countries. For example, Nielsen & Rosholm (2001) and Mueller (1998) use quantile regressions to examine public–private sector wages in Zambia and Canada, respectively. Albrecht *et al*. (2007) use quantile regressions to examine urban–rural inequality in Vietnam. Quantile regressions have also been used to examine wage inequality in China (e.g. Knight & Song, 2003) and to examine gender pay gaps in Spain (e.g. Garcia *et al*., 2001), Chile (e.g. Montenegro, 2001), and the Philippines (Sakellarious, 2004).

[2] It is important to mention that reverse causality is also plausible here. In other words, male-dominated occupations could be seen as market oriented, while female-dominated occupations could be seen as domestic oriented. The point here is to avoid the arbitrary assignment of the labels “market” and “domestic.”

[3] Experience rather than age is typically used in Mincerian wage equations. However, measures of actual experience are unavailable in our data. For this reason, and consistent with Mincer (1974), we use potential experience defined as age-*S*-*6*, where *S* represents years of schooling. And since potential experience is a linear function of age and years of schooling, age and experience can be used interchangeably in the earnings equations. We choose experience.

[4] In the Kenyan education system, primary education consists of classes Standards 1–8 and secondary of Forms 1–6.

[5] Let *v* be a distributional statistic of interest such as a quantile. The influence function (IF) of *v* at a point y in robust statistics and econometrics is defined as:

$$\mathrm{IF}(y; v; F) \equiv \lim_{t0} \frac{v(F_{t,\Delta y}) - v(F)}{t} = \frac{v(F_{t,\Delta y})}{\partial t|t=0},$$

For example, RIF for a quantile q_t is given by

$$\mathrm{RIF}(Y;q_t) = q_\tau + \frac{\tau - I(Y \le q_\tau)}{f_y(q_\tau)}$$

where f_y is the marginal density function of *Y*, and $I(\cdot)$ is an indicator function. In practice, the RIF may be estimated by replacing unknown quantities by the estimators, that is $\hat{\mathrm{RIF}}(Y;\hat{q}_t) = \hat{q}_\tau + (\tau - I(Y \le \hat{q}_t))\ \hat{f}_y(\hat{q}_\tau)$, where $\hat{q}_\tau$ is the τth sample quantile and f_y is the kernel density estimator.

[6] Indeed, in our sub-samples, the female sub-sample (2687) is relatively smaller than the male sub-sample (4834). Furthermore, our results from the propensity score estimates in Table 3 suggest no *a priori* reason that would impact our results in Tables 4 and 5.

[7] The impact of either composition effects or wage structure effects and the influence of various categories and covariates on composition and wage structure effects at other quantiles will be determined and interpreted in the same way as at the 20th quantile. In this light, composition effects explain the gender pay gap at the 10th quantile and would follow the same interpretation as at the 20th quantile.

References

Agesa, R. U. (1999) The urban gender wage gap in an African country: Findings from Kenya, *Canadian Journal of Development Studies*, 20(1), pp. 59–76.

Agesa, R. U., Agesa, J. & Dabalen, A. (2009) Changes in wage distributions, wage gaps and wage inequality by gender in Kenya, *Journal of African Economies*, 18(3), pp. 431–446.

Albrecht, J., Thanh, N. B. & Susan, V. (2007) A quantile regression decomposition of urban-rural inequality in Vietnam, *Journal of Development Economics*, 38(2), pp. 466–490.

Appleton, S., John, H. & Pramila, K. (1999) The gender wage gap in three African countries, *Economic Development and Cultural Change*, 47(2), pp. 289–312.

Ashraf, J. & Ashraf, B. (1993) Estimating the gender wage gap in Rawalpindi city, *Journal of Development Studies*, 29(2), pp. 365–376.

Autor, D. H, Katz, L. F & Kearny, M. S (2005) Rising wage inequality: The role of composition and prices, Working paper No. 11628, National Bureau of Economic Research (NBER).

Blinder, A. (1973) Wage discrimination: Reduced forms and structural estimation, *Journal of Human Resources*, 8(4), pp. 436–455.

Del Río, C., Carlos, G. & Olga, C. (2011) The measurement of gender wage discrimination: The distributional approach revisited, *Journal of Economic Inequality*, 8(1), pp. 57–86.

Firpo, S., Fortin, M. N. & Lemieux, T. (2009a) Unconditional quantile regressions, *Econometrica*, 77(3), pp. 953–973.

Firpo, S, Fortin, M. N & Lemieux, T (2009b) Unconditional quantile regressions, Working Paper, http://www.eco

n.ubc.ca/nfortin/FFL_Occupations1.pdf
Garcia, J., Pedro, H. & Angel, N. (2001) How wide is the gap? An investigation of gender wage differences using quantile regression, *Empirical Economics*, 26(1), pp. 149–167.
Hinks, T. (2002) Gender wage differentials and discrimination in the New South Africa, *Applied Economics*, 34(16), pp. 2043–2052.
Horton, S. (1996) *Women and industrialization in Asia* (London: Routledge).
Knight, J. & Song, L. (2003) Increasing urban wage inequality in China, *Economics of Transition*, 11(4), pp. 597–619.
Machado, J. A. F. & Mata, J. (2005) Counterfactual decomposition of changes in Wage distributions using quantile regression, *Journal of Applied Econometrics*, 20(4), pp. 445–465.
Mariara-Kabubo, J. (2003) Wage determination and the gender wage gap in Kenya: Any evidence of gender discrimination?, Working Paper No. RP_132, *African Economic Research Consortium (AERC)*, Nairobi, Kenya, May, 2003, http://www.aercafrica.org/publications/item.asp?itemid¼146&category¼.
Melly, B. (2005) Decomposition of differences in distribution using quantile regression, *Labour Economics*, 12(4), pp. 577–590.
Mincer, J. (1974) *Schooling Experience, and Earnings* (New York: National Bureau of Economic Research, Columbia University Press).
Montenegro, C. (2001) Wage Distribution in Chile: Does Gender Matter? A Quantile Regression Approach Policy Research Report on Gender and Development, Working Paper Series No. 20, December 2001, The World Bank Development Research Group/Poverty Reduction and Economic Management Network.
Mueller, R. E. (1998) Public-private sector wage differentials in Canada: Evidence from quantile regressions, *Economics Letters*, 60(2), pp. 229–235.
Nielsen, H. S. & Rosholm, M. (2001) The public-private sector wage gap in Zambia in the 1990s: A quantile regression approach, *Empirical Economics*, 26(1), pp. 169–182.
Nordman, C. J. & Wolff-Francois, C. (2009) Is There a glass ceiling in Morocco? Evidence from matched worker-firm Data, *Journal of African Economies*, 18(4), pp. 592–633.
Nordman, C. J. & Roubaud, F. (2009) Reassessing the gender wage gap in Madagascar: Does labor force attachment really matter? *Economic Development and Cultural Change*, 57(4), pp. 785–808.
Ntuli, M. (2009) Exploring gender wage discrimination in South Africa, 1995–2004: A quantile regression approach, in: Ravi Kanbur & Jan Svejnar (Eds) *Labor Markets and Economic Development* (Routledge Studies in Economic Development), pp. 468–495.
Oaxaca, R. (1973) Male-female wage differentials in urban labor markets, *International Economic Review*, 14(3), pp. 693–709.
Psacharoupoulos, G. & Tzannatos, Z. (1991) Women’s employment and pay in Latin America, Volumes l and 2, The World Bank, Latin America and the Caribbean Region, Technical Department, Regional Studies Report No. 2, October 1991.
Sakellarious, C. (2004) The use of quantile regressions in estimating gender wage differentials: A case study of the Philippines, *Applied Economics*, 36(9), pp. 1001–1007.
Siphambe, K. H. & Bakwena-Thokweng, M. (2001) The wage gap between men and women in botswana’s formal labour market, *Journal of African Economies*, 10(2), pp. 127–142.
Temesgen, T. (2006) Decomposing gender wage differentials in Urban Ethiopia: Evidence from linked employer-employee (LEE) manufacturing survey data, *Global Economic Review*, 35(1), pp. 43–66.
Wanjala, M. B. & Were, M. (2009) Gender disparities and economic growth in Kenya: A social accounting matrix approach, *Feminist Economics*, 15(3), pp. 227–251.
Weichselbaumer, D. & Winter-Ebmer, R. (2005) A meta-analysis of the international gGender wage gap, *Journal of Economic Surveys*, 19(3), pp. 479–511.