



Sorting the Gender Earnings Gap: Heterogeneity in the South African Labor Market

Dorrit Posel, Dambala Gelo, Daniela Casale & Adeola Oyenubi

To cite this article: Dorrit Posel, Dambala Gelo, Daniela Casale & Adeola Oyenubi (2024) Sorting the Gender Earnings Gap: Heterogeneity in the South African Labor Market, *Feminist Economics*, 30:1, 106-133, DOI: [10.1080/13545701.2023.2279227](https://doi.org/10.1080/13545701.2023.2279227)

To link to this article: <https://doi.org/10.1080/13545701.2023.2279227>



Published online: 15 Dec 2023.



Submit your article to this journal [↗](#)



Article views: 339



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 1 View citing articles [↗](#)



SORTING THE GENDER EARNINGS GAP: HETEROGENEITY IN THE SOUTH AFRICAN LABOR MARKET

Dorrit Posel, Dambala Gelo, Daniela Casale, and Adeola Oyenubi

ABSTRACT

This study explores inequality among Africans in the South African labor market by investigating heterogeneity in the gender earnings gap. The article adds to the existing literature by applying an innovative sorting method, which provides a complete mapping of heterogeneity in the gender earnings gap, ordering the full distribution of the partial effects from largest to smallest with respect to the underlying characteristics of the population of interest. This makes it possible to identify the extent of the dispersion in earnings inequality as well as those characteristics that are associated with very large, as well as very small (or even reversed), gender gaps in earnings. The study also compares heterogeneity in the hourly and monthly gender earnings gap to assess how adjusting for working hours affects the correlates of earnings inequality.

KEYWORDS

Gender wage gap, sorted partial effects, heterogeneity, South Africa

JEL Codes: J16, J31

HIGHLIGHTS

- There is significant heterogeneity in the size of the gender earnings gap in South Africa.
- This is illustrated using a technique that maps the full distribution of earnings differences.
- Marriage, children, and geography are important sources of heterogeneity.
- There is more heterogeneity in the monthly, than hourly, gender earnings gap.

INTRODUCTION

South Africa has been identified recently as the country with the highest rate of inequality among the 164 countries tracked by the World Bank (based on Gini coefficients for consumption or income; World Bank 2022: 1). Given the legacy of apartheid,¹ much of the early empirical research which described and analyzed this inequality focused on its inter-racial dimensions (Allanson, Atkins, and Hinks 2000; Rospabé 2001). However, in recent decades, inequality has increasingly been influenced by intra-African² trends, deriving from rising inequality in the labor market among the majority population group (Leibbrandt et al. 2010).

In this article we explore inequality among Africans in the South African labor market by investigating heterogeneity in the gender earnings gap. We focus on working-age Africans (ages 16–64) with either wage or self-employment and apply an innovative sorting method, developed in Victor Chernozhukov, Ivan Fernandez-Val, and Ye Luo (2018). This method permits a complete mapping of the gender earnings gap with respect to the underlying distribution of the covariates in the population of interest. Previous studies of the gender earnings gap in South Africa, and elsewhere, have typically only presented average partial effects in their analysis, or they have relied on quantile regression to map how the partial effects vary at certain points along the earnings distribution (Paul, Zaw, and Darity 2022). The “sorted effects method,” in contrast, involves estimating the partial effects for each observation in the population of interest and ordering them from biggest to smallest. This method makes it possible to identify the extent of heterogeneity in the gender earnings gap as well as which groups of African women (and men) are the most, and the least, affected by gender earnings inequality.

Chernozhukov, Fernandez-Val, and Luo (2018) illustrated their method by describing the gender earnings gap in the US, while very recent studies have applied the method to explore other topics in economics. For example, Martin Kroczeck and Philipp Kugler (2022) investigate heterogeneity in job attractiveness, and Hemant K. Pullabhotla and Mateus Souza (2022) probe the heterogeneous health effects of agricultural fires.

The sorted effects method is particularly apposite for a country like South Africa, which has such extreme inequality in the labor market. Mapping the complete range of partial effects (in a highly intuitive visual plot) draws attention both to the degree and pattern of heterogeneity in earnings inequality. In addition, the method lends itself to a very useful (but easy to understand) classification analysis that compares the characteristics of the individuals at the tails of the sorted partial effects plot. This analysis makes visible not only characteristics that are associated with an earnings disadvantage among women (which is typically the focus of the gender gap research, both in South Africa and elsewhere), but also those attributes

that are correlated with an earnings advantage among women relative to comparable men.

Our study also adds to the literature by estimating and comparing the sorted partial effects for hourly earnings (the typical metric) and monthly earnings. In so doing, we consider how differences in working hours affect heterogeneity in the gender earnings gap, and among which groups of women specifically.

REVIEW: HETEROGENEITY IN THE GENDER EARNINGS GAP IN SOUTH AFRICA

South Africa has a long history of racial oppression and exclusion, where laws controlled where people of color could live, with whom they could live, and what kind of work they could undertake. At the start of the democratic transition, one study estimated that the “typical white male worker . . . received 126 per cent more than the overall geometric mean wage . . . and the typical black (African) 32 per cent less” (Allanson, Atkins, and Hinks 2000: 116). However, apartheid restrictions also intersected with patriarchal systems of control to disadvantage African women more than African men (Walker 1990).

During the apartheid years, women were considerably less likely than men to participate in the labor market; and those who found employment were clustered in occupations typically seen as “women’s work,” such as clerical, sales, and service occupations (Maconachie 1989; Casale 2004; Casale, Posel, and Mosomi 2021). Together with apartheid constraints, which limited access to the labor market for people of color, African women were further marginalized, and occupied far more precarious and low-paying jobs than white women or African men (Maconachie 1989; Posel and Casale 2019; Casale, Posel, and Mosomi 2021). Paid domestic work in private households was, and remains, a very important source of employment for African women, and historically, this work was largely unprotected, precarious, and poorly paid (Cock 1980).

Although much of the early research on earnings inequality focused on the race gap in (men’s) earnings, the gender gap in South Africa has come to be relatively well studied, particularly in recent decades (Rospabé 2001; Hinks 2002; Casale 2004; Grün 2004; Ntuli 2007; Posel and Muller 2008; Shepherd 2008; Muller 2009; Casale and Posel 2011; Bhorat and Goga 2013; Mosomi 2018, 2019a, 2019b; Roberts and Schoer 2021).

At the start of the democratic transition in South Africa, the gender earnings gap was large (comparable to the gender earnings gap in the US in the 1960s; Blau and Winkler 2018), with an overall average gap of approximately 40 percent (Mosomi 2018). The gap was highest between white men and women (Rospabé 2001; Grün 2004), with some research even estimating a gap in favor of women among Africans (Hinks 2002; Grün

2004). Later studies, however, showed how these findings were polluted by under-sampling, specifically in the first post-apartheid national surveys, of low-wage work in which African women predominated (Muller 2009; Mosomi 2018).³

Initial research on gender earnings inequality in South Africa mostly explored the hourly gender earnings gap at the mean of the distribution, using methods such as the Oaxaca–Blinder decomposition (Blinder 1973; Oaxaca 1973), and later its generalization in the John DiNardo, Nicole M. Fortin, and Thomas Lemieux (DFL; 1996) method, to measure the explained and unexplained portions of the average gap. However, subsequent studies probed heterogeneity in the hourly gender earnings gap using quantile regression analysis to study the gap at various points along the earnings distribution. This research found that the median earnings gap was greater than the mean gap, and in contrast to trends in the average gender gap, the median gap had largely stagnated (at between 23 and 35 percent; Mosomi 2018; Mosomi 2019a; Casale, Posel, and Mosomi 2021). These findings reflected earnings differences between women and men in South Africa that were consistently larger at the lower tail of the distribution, or a “sticky floor effect” (Mosomi 2018: 90), indicating that even amongst the lowest paid workers, women earned far less than men (Ntuli 2007; Bhorat and Goga 2013; Mosomi 2018). Among Africans specifically, Haroon Bhorat and Sumayya Goga (2013) estimated a gender earnings gap at the bottom decile that was almost nine-fold higher than the gap at the top decile in 2007.

However, there have also been significant changes over the post-apartheid period: the gender gap at the lower tail has fallen by more than the average gap, a trend that is attributed largely to the effects of new minimum wage legislation, which has offered some wage protection to jobs in which women predominate (such as paid domestic work; Mosomi 2018). Moreover, although the hourly gender gap is smaller at the upper end of the earnings distribution, decomposition analysis suggests that the unexplained part of the gap (that which is not explained by differences in observable characteristics) is larger. This finding, which indicates that women at the upper end of the distribution likely experience more labor market discrimination than women at the lower end (Mosomi 2018), is consistent with research globally of a “glass ceiling” experienced by women at the top of the labor market (Blau and Kahn 2017).

A recent study for South Africa showed that women working in occupations dominated by men earned significantly more than other women, with earnings highest, and the average gender earnings gap lowest, among self-employed women working in occupations dominated by men (Roberts and Schoer 2021). These findings indicate that job sorting by gender remains a key driver of the gender earnings gap in South Africa. This occupation segregation is not attributable to women having lower

levels of education: in contrast to many developing countries, women in South Africa are now more educated than men on average, and this is particularly the case among younger cohorts (Van Broekhuizen and Spaull 2017; Spaull and Makaluza 2019). However, women still tend to acquire different types of education to men, and particularly at the post-secondary level, where women remain under-represented in science, technology, engineering, and mathematics (STEM) fields and over-represented in fields more closely related to the “care economy,” such as education, psychology, and social work (Van Broekhuizen and Spaull 2017; Spaull and Makaluza 2019).

Decomposition analysis of the gender earnings gap has also revealed life cycle and marriage effects, pointing to both age and marital status as contributors of the earnings differences. Jacqueline Mosomi (2019a, 2019b), for example, estimates that on average, men’s earnings in South Africa are higher than women’s earnings across the whole age distribution; but the gender gap increases with age, until about age 45, corresponding to the period when women are most likely to experience labor market discontinuity because of childcare responsibilities. The returns to marriage are also higher for men than for women, and although this is evident across the wage distribution, the gap is widest at the upper tail of the earnings distribution (Mosomi 2019a, 2019b).

In all these studies on the gender gap for South Africa (and in most other countries), average partial effects or partial effects at specific points along the earnings distribution (as in quantile regression) are presented, and sometimes for certain sub-groups (such as white or African workers). However, partial effects can vary substantially with respect to the underlying distribution of the covariates in the model. The sorting method used in this study analyzes heterogeneity in the gender earnings gap in South Africa by estimating the partial effect for each observation in the population of interest and sorting them from largest to smallest, or put differently, from women most affected by the gender earnings gap to women least affected (or in some cases even advantaged relative to observably comparable men). The method also lends itself to a classification analysis in which the characteristics of the most and least affected groups of women are compared, thereby revealing those covariates that are associated with the greatest disadvantage as well as those associated with the greatest advantage. We describe this method more formally below.

METHODS, DATA, AND DESCRIPTIVE STATISTICS

Modeling the gender earnings gap

We start off by specifying a basic earnings equation by gender as follows:

$$\ln w_{if} = \alpha + \beta \text{ gender} + \gamma x + \mu x \text{ gender} + \epsilon_{if} \quad (1)$$

$$\ln w_{im} = \alpha + \beta \text{ gender} + \gamma x + \mu x \text{ gender} + \epsilon_{im} \quad (2)$$

where $\text{gender} = 1$ if female (f) and $\text{gender} = 0$ if male (m), β is a change in intercept due to being female, x is a vector of demographic, employment, and household level covariates, γ is a vector of parameters representing the main associations between the covariates, and w (earnings), and μ is a vector of parameters representing the associations between w and the interaction terms (gender interacted with the covariates).⁴ Note that with $\text{gender} = 1$, (1) and (2) can be represented as follows respectively:

$$\ln w_{if} = \alpha + \beta + (\gamma + \mu)x + \epsilon_{if} \quad (1')$$

$$\ln w_{im} = \alpha + \gamma x + \epsilon_{im} \quad (2')$$

We can derive a heterogeneous gender earnings gap by taking the difference between the expectation of (1') and (2') as follows:

$$\Lambda(x) = [\alpha + \beta + (\gamma + \mu)x] - [\alpha + \gamma x] = \beta + \mu x = \beta + \sum_{j=1}^p \mu_j x_j \quad (3)$$

$\Lambda(x)$, referred to as the partial effect (PE), is a measure of the gender earnings gap conditional on characteristics x (Chernozhukov, Fernandez-Val, and Luo 2018). As can be seen from (3), the heterogeneity of $\Lambda(x)$ arises from the fact that it is a function of individual characteristics rather than being a scalar (single summary). That is, $\Lambda(x)$ is a function of x and hence is different for each observational unit/individual.⁵ In contrast, the average partial effect (APE) defined as $\int \Lambda(x) d\pi(x)$, where π denotes the distribution of x in the population of interest, estimates the average effect of the predictor on the outcome variable w , holding all other predictor variables constant (Chernozhukov, Fernandez-Val, and Luo 2018).

An attractive feature of this heterogeneous gender earnings gap analysis is the ability to sort the PEs in decreasing or increasing order, and thereby rank them with respect to the distribution of the covariates, π . This option helps provide a more complete representation of the heterogeneity of $\Lambda(x)$. To be specific, the sorted effects, referred to as sorted partial effects (SPEs), each correspond to a specific percentile of the PE, defined as:

$$\Lambda(x^\tau) = \{\Delta \in \mathbb{R} : F(\Lambda(x)) \leq \tau\} \quad (4)$$

where $\tau \in [0, 1]$ and Δ is the τ th quantile of $\Lambda(x)$.

We estimate the SPEs using the sorted effects estimator of Chernozhukov, Fernandez-Val, and Luo (2018). Apart from providing estimates and

confidence bands for SPEs, this estimator supports a classification analysis, which involves (i) classifying the observational units into the most or least affected depending on whether their PEs are above or below some tail SPE (such as 10 percent or 20 percent of the sample), and then (ii) testing differences in moments or distributions of the observed characteristics of individuals across these subgroups (Chernozhukov, Fernandez-Val, and Luo 2018).⁶

To operationalize the model, we estimate linear models which include a full set of interactions between the covariates and a female dummy variable. We do not control for the possibility of selection into employment, in keeping with many studies of the gender earnings gap in South Africa (Hinks 2002; Ntuli 2007; Posel and Muller 2008; Shepherd 2008; Bhorat and Goga 2013; Mosomi 2019b; Roberts and Schoer 2021). The few studies that do adjust for selection (typically with a two-stage Heckman model) mostly do not find evidence of significant effects (perhaps because of the difficulty of finding good exclusion restrictions).

The regression equations estimate two outcome variables – the log of monthly earnings and the log of hourly earnings. We use the log transformation because the earnings distribution is skewed (with a far larger lower, than upper, tail); and we distinguish between monthly and hourly earnings to consider how differences in working hours affect variation in the gender earnings gap. In most countries including South Africa, most part-time workers are women because adjusting work hours is integral to how women negotiate unpaid work in the care economy with paid work in the labor market (Tijdens 2002; Posel and Muller 2008).

The selection of covariates in the earnings regressions is typical of Mincerian-type earnings equations for South Africa (Mincer 1974; Rospabé 2001; Ntuli 2007; Muller 2009; Bhorat and Goga 2013; Mosomi 2019b), and includes controls for human capital, location, and industry of employment. As we want to describe whether heterogeneity in the gender earnings gap is correlated with non-market work, we also add variables for marital status (Rospabé 2001; Bhorat and Goga 2013; Mosomi 2019a, 2019b), as well as household size and the presence of children (Posel and Muller 2008; Bhorat and Goga 2013). Our objective is only to map associations and not attribute causality, but we recognize that these latter variables may be endogenous to earnings (Blau and Kahn 2017). For example, Daniela Casale and Dorrit Posel (2010b) use longitudinal data for South Africa from the 2000s to show positive selection effects of higher-earning African men into marriage. As these covariates may therefore comprise “bad controls” (Angrist and Pischke 2008), we also estimate pared-down regressions (or what Blau and Kahn [2017] refer to as the “human capital specification”) without controls for marital status, household size, and the co-residence of children.

Data and descriptive statistics

The data used for the analysis come from the Post-Apartheid Labour Market Series (PALMS)⁷ version of the Quarterly Labour Force Survey (QLFS) from the first quarter of 2017. The QLFS is a nationally representative survey conducted by the official statistical agency in South Africa (Statistics South Africa) and publicly available at the open data portal, <https://www.datafirst.uct.ac.za>. In the unweighted sample, there are 12,354 African adults ages 16–64 years with either wage or self-employment who have non-missing information for income earned, hours worked, and the other covariates included in the estimations. The age range is selected because minimum and maximum ages for school-leaving and retirement in South Africa are 15 and 65 years, respectively.⁸

The QLFS asks respondents to report on gross earnings for the wage employed and earnings after expenses for the self-employed. Earnings information typically suffers from non-response and reporting errors, and as in other surveys, the QLFS allows for earnings bracket responses as a way of overcoming problems of disclosure. However, this raises the difficulty of how to work with earnings that are reported in brackets. Using the mid-points of the bracket responses “creates ‘spikes’ in the distribution which distorts other statistics, such as dispersion” (Wittenberg 2017: 303). We therefore follow an approach identified in Martin Wittenberg (2017), and frequently used by other researchers working with these data, of including only individuals who provided point estimates of earnings but reweighting the sample to account for the employed who provided bracket responses. As there are concerns about the reliability of income reported by the self-employed in particular (Mosomi 2018), we also test the robustness of our results for the sample of the wage employed only, which we report on later.

Table 1 presents descriptive statistics for the outcome variables, namely the log of monthly earnings and the log of hourly earnings, and the human capital, employment, and other covariates included in the full regression specification. Post-stratification weights, which adjust for individuals who reported earnings in brackets, and for differences in response rates across demographic groups, are applied in this table and throughout the analysis (Kerr, Lam, and Wittenberg 2019).

Employed African men earn significantly more per month on average than employed African women (R8110.14 versus R6071.10),⁹ with an average gender earnings gap of 25.1 percent. This is partly because compared to men, women work fewer hours a week (39.6 versus 44.8 average hours per week) and are significantly more likely to work part-time (23 percent versus 13 percent). However, a large and statistically significant gender gap in earnings persists even after adjusting for the number of hours worked. Average hourly earnings for African men and women are R47.02 for men and R36.73 for women, representing an average gender wage gap of 21.9 percent.

SORTING THE GENDER EARNINGS GAP

Table 1 Descriptive statistics, African women and men with wage or self-employment

<i>Averages or proportions</i>	<i>Women</i>	<i>SE</i>	<i>Men</i>	<i>SE</i>
Monthly wage (Rands)	6071.10***	229.97	8110.14	256.56
Hourly wage (Rands)	36.73***	1.43	47.02	1.76
Log of monthly wage	7.96***	0.02	8.28	0.02
Log of hourly wage	2.86***	0.02	3.05	0.02
Weekly hours worked	39.60***	0.21	44.78	0.20
Works part-time (< 35 hrs/week)	0.23***	0.01	0.13	0.00
Self-employed	0.14***	0.00	0.18	0.00
<i>Individual characteristics:</i>				
Age	39.25***	0.15	37.24	0.13
Less than grade 12	0.51***	0.01	0.57	0.01
Grade 12	0.28	0.01	0.28	0.01
Grade 12 + post-secondary	0.21***	0.01	0.15	0.00
Married or cohabiting	0.40***	0.01	0.51	0.01
Widowed	0.07***	0.00	0.01	0.00
Divorced	0.04***	0.00	0.02	0.00
Never married	0.49***	0.01	0.46	0.01
<i>Length of current job tenure:</i>				
< 5 years	0.60	0.01	0.61	0.01
≥ 5 years, < 10 years	0.20	0.01	0.20	0.01
≥ 10 years, < 20 years	0.13	0.00	0.13	0.00
≥ 20 years	0.06	0.00	0.06	0.00
<i>Industry:</i>				
Agriculture	0.04***	0.00	0.07	0.00
Mining	0.01***	0.00	0.05	0.00
Manufacturing	0.08***	0.00	0.12	0.00
Utilities	0.00***	0.00	0.01	0.00
Construction	0.03***	0.00	0.16	0.00
Wholesale and retail trade	0.24***	0.01	0.18	0.01
Transport	0.02***	0.00	0.09	0.00
Finance	0.11***	0.00	0.13	0.00
Community services	0.29***	0.01	0.15	0.00
Private households	0.18***	0.01	0.05	0.00
<i>Household:</i>				
At least one child 0–5	0.40***	0.01	0.30	0.01
At least one child 6–15	0.47***	0.01	0.31	0.01
Household size	4.40***	0.04	3.59	0.04
<i>Location:</i>				
Urban area	0.70	0.00	0.69	0.00
Traditional rural area	0.25***	0.00	0.21	0.00

(Continued).

Table 1 Continued.

<i>Averages or proportions</i>	<i>Women</i>	<i>SE</i>	<i>Men</i>	<i>SE</i>
Commercial farming area	0.03***	0.00	0.05	0.00
Mining areas	0.02***	0.00	0.05	0.00

Notes: The data have been weighted to represent population statistics. The sample is all African adults (ages 16–64) with wage or self-employment. Although not described here, dummies variables for the nine provinces in South Africa are also included in the regressions to account for regional labor market differences. ***, **, * denote significance at the 1, 5, and 10 percent levels, respectively.

Source: Quarterly Labour Force Survey (2017: 1).

Most African women and men are in wage employment, with the share significantly higher among women (86 percent) than men (82 percent). There are no significant differences by gender in length of tenure in current employment. However, Table 1 describes significant variation in the industry of employment, suggesting persistent gender segregation in paid work. For example, women’s employment is concentrated in three sectors – wholesale and retail trade, community services, and private households (largely working as paid domestic workers) – accounting for the employment of 71 percent of African working-age women. The employment of African men, in contrast, is more dispersed across the traditionally “male-dominated” sectors of mining, construction, and transport, as well as manufacturing and agriculture.

There is considerable variation in educational attainment in South Africa, but the gender gap in earnings is not reflected in men’s educational advantage: women are significantly less likely than men not to have completed schooling; there is no difference in the share of women and men holding at most a grade 12 or “matric” education (the final year of schooling in South Africa); and women are significantly more likely than men to report post-secondary education (21 percent compared to 15 percent of African women and men, respectively).

As is well-documented in several studies, marriage rates among African women and men in South Africa are very low (Posel and Casale 2013; Posel and Hall 2021), and the modal marital status among African working-age women with employment is to have never married and not be cohabiting with a partner. Women are significantly more likely than men to have never married, or to be previously married (and now widowed or divorced). However, women are also substantially more likely than men to live with at least one child younger than 6 years of age (40 percent of African women compared to 30 percent of African men), and at least one child between the ages of 6 and 15 years (47 percent of African women compared to 31 percent of African men). In the context of higher rates of marriage and cohabitation among men, these differences in co-residence with children

highlight how the link between marriage and childbearing has been dislodged in South Africa (Garenne et al. 2001; Hunter 2010), and they help to explain why African women live in significantly larger households than African men (Posel, Casale, and Grapsa 2016).

RESULTS: SORTED EFFECTS AND THE CLASSIFICATION ANALYSIS

The average and sorted partial effects of the earnings gap between African women and men in South Africa are summarized in Figure 1, which also plots the 90 percent confidence intervals of the estimates.¹⁰ The left frame shows an average partial effect (APE) of -0.257 when the monthly gender wage gap is analyzed.¹¹ However, the sorted partial effects (SPEs) range from -0.729 to 0.094 , indicating substantial heterogeneity in the gender wage gap.¹²

This wide range of SPEs brings to light an important contribution of the method; namely, that it shows how the average partial effect typically reported in studies on the gender earnings gap masks considerable variation in the disadvantage experienced by women relative to observably comparable men. More specifically, in our sample there is a group of disadvantaged African women for whom the gender earnings gap is almost three times the average gap (-0.729 versus -0.257), while there are African women whose monthly earnings are indeed higher than those

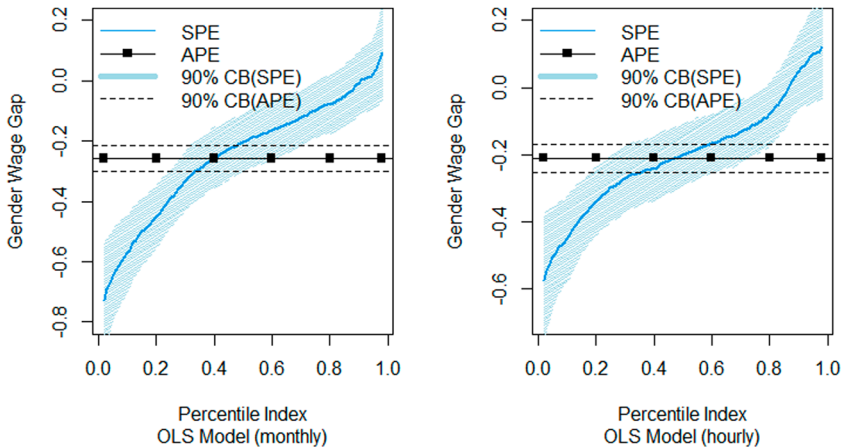


Figure 1 Sorted effects of the gender wage gap for women with wage or self-employment.

Notes: APE $\sim$ average partial effect; SPE $\sim$ sorted partial effect; CB $\sim$ confidence band.

Source: Quarterly Labour Force Survey (2017: 1).

of observably similar African men. About 40 percent of African women experience a gender earnings gap that is larger than the average gap (shown by the point on the graph where the APE and SPE curves intersect), while roughly 10 percent of African women experience an earnings advantage relative to observably comparable African men (the part of the curve where the SPE is above zero).

As expected, given that women work fewer hours on average than men, the APE for the gender earnings gap is somewhat lower at -0.211 when hourly earnings are used as the dependent variable (right frame of Figure 1). Nonetheless, there remains substantial heterogeneity in the gender earnings gap, with the SPEs ranging from -0.577 to 0.122 . When analyzing hourly earnings, the sorted effects method uncovers that almost half of African women experience a gap that is larger than the average gap, with around 13 percent experiencing a positive gap or an earnings advantage relative to observably similar men.

Another useful feature of the sorting method by Chernozhukov, Fernandez-Val, and Luo (2018) is that it reveals *which* women are most disadvantaged and *which* women are least disadvantaged when compared to observably similar men, across the range of covariates included in the model. This is done here by comparing the characteristics of the 20 percent of women most affected or disadvantaged by the earnings gap (those with the lowest, or most negative SPE values) to the characteristics of the 20 percent of women least affected or disadvantaged by the earnings gap (those with the highest, or largest positive SPE values). The results of this classification analysis for the monthly wage gap are shown in Table 2. A negative number, in the column that shows the difference between the two groups of women (column 5), indicates that the characteristic is more common among the most disadvantaged women, while a positive number indicates that the characteristic is more common among those women least disadvantaged relative to men.¹³

The classification analysis on monthly earnings suggests the typical human capital variables are important in distinguishing between those least and most affected by the gender gap. For instance, 41 percent of women in the least affected group have not completed secondary education (that is, Grade 12) compared to 66 percent of those in the most affected group, a significant difference of -25 percentage points (as shown in column 5). In contrast, 32 percent of women in the least affected group have a post-secondary education (either a degree or diploma), compared to only 9 percent of women in the most affected group, a statistically significant difference of 23 percentage points. However, there is no significant difference in the percentage of women in the least and most affected groups with a completed Grade 12. These results are consistent with other literature for South Africa which finds that Grade 12 (or the matric certificate) offers little protection against weak employment outcomes,

SORTING THE GENDER EARNINGS GAP

Table 2 Classification analysis for employed women (least and most affected 20 percent) – monthly earnings gap

	20% least affected		20% most affected		Difference (least-most affected)		
	Mean	SE	Mean	SE	Estimate	SE	P-val
<i>Individual characteristics:</i>							
Age	37.71	1.64	41.87	1.10	− 4.16	2.63	0.06*
Age squared	1531.9	138.4	1856.3	95.73	− 324.40	221.97	0.07*
Less than grade 12	0.41	0.08	0.66	0.05	− 0.25	0.11	0.02**
Grade 12	0.27	0.07	0.26	0.04	0.01	0.10	0.45
Grade 12 + post-secondary	0.32	0.09	0.09	0.03	0.23	0.12	0.02**
Tenure < 5 yrs	0.45	0.07	0.72	0.04	− 0.27	0.10	0.00***
Tenure ≥ 5 yrs, < 10 yrs	0.28	0.06	0.15	0.03	0.12	0.08	0.06*
Tenure ≥ 10 yrs, < 20 yrs	0.14	0.05	0.11	0.03	0.04	0.08	0.32
Tenure ≥ 20 yrs	0.13	0.05	0.03	0.02	0.10	0.06	0.06*
Married	0.27	0.06	0.51	0.05	− 0.24	0.11	0.01***
Widowed	0.06	0.05	0.09	0.04	− 0.03	0.09	0.37
Divorced	0.02	0.02	0.05	0.03	− 0.03	0.05	0.23
Never married	0.65	0.07	0.35	0.05	0.30	0.11	0.00***
<i>Household characteristics:</i>							
At least one child 0–5	0.37	0.07	0.45	0.05	− 0.08	0.12	0.24
At least one child 6–15	0.20	0.07	0.64	0.05	− 0.44	0.11	0.00***
Household size	3.50	0.34	5.09	0.31	− 1.59	0.62	0.01***
<i>Location variables:</i>							
Urban area	0.92	0.05	0.44	0.05	0.48	0.09	0.00***
Traditional	0.07	0.04	0.44	0.05	− 0.37	0.09	0.00***
Commercial farming area	0.01	0.01	0.04	0.01	− 0.03	0.02	0.09*
Mining areas	0.00	0.01	0.08	0.02	− 0.07	0.02	0.00***
Western Cape	0.17	0.03	0.00	0.00	0.17	0.03	0.00***
Eastern Cape	0.11	0.05	0.09	0.03	0.02	0.08	0.41
Northern Cape	0.00	0.01	0.01	0.01	− 0.01	0.01	0.30
Free State	0.07	0.04	0.04	0.02	0.03	0.05	0.29
KwaZulu-Natal	0.08	0.05	0.21	0.04	− 0.13	0.08	0.06*
Northwest	0.00	0.00	0.22	0.04	− 0.22	0.04	0.00***
Gauteng	0.49	0.08	0.15	0.04	0.35	0.11	0.00***
Mpumalanga	0.07	0.03	0.11	0.03	− 0.04	0.06	0.27
Limpopo	0.01	0.02	0.18	0.04	− 0.17	0.06	0.00***
<i>Employment characteristics:</i>							
Self-employed	0.00	0.01	0.52	0.06	− 0.52	0.07	0.00***
Agriculture	0.07	0.03	0.01	0.01	0.06	0.03	0.04**
Mining	0.02	0.01	0.00	0.00	0.02	0.02	0.15
Manufacturing	0.06	0.04	0.06	0.02	− 0.01	0.06	0.44
Utilities	0.01	0.01	0.00	0.00	0.01	0.01	0.18

(Continued).

Table 2 Continued.

	20% least affected		20% most affected		Difference (least-most affected)		
	Mean	SE	Mean	SE	Estimate	SE	P-vals
Construction	0.00	0.00	0.13	0.02	− 0.13	0.02	0.00***
Wholesale and retail trade	0.14	0.06	0.37	0.05	− 0.23	0.10	0.01***
Transport	0.04	0.03	0.00	0.00	0.04	0.03	0.08*
Finance	0.30	0.06	0.02	0.01	0.28	0.07	0.00***
Community services	0.08	0.06	0.37	0.07	− 0.28	0.12	0.01***
Private households	0.27	0.09	0.03	0.03	0.24	0.12	0.02*

Notes: The sample is African adults (ages 16–64) with wage and self-employment. The “ca” command in the Sorted Effects package is used to produce the estimates in this table. The analysis is weighted, and SEs are based on 500 bootstrap replications. ***, **, * denote significance at the 1, 5, and 10 percent levels, respectively.

Source: Quarterly Labour Force Survey (2017: 1).

and that tertiary education is key to labor market success in the post-apartheid period (Posel, Casale, and Vermaak 2014; Mosomi 2018; Roberts and Schoer 2021).

The results suggest age confers disadvantage in terms of the gender gap, a common finding for many countries, given interruptions to women’s experience-earnings profile in the child-rearing years (Sigle-Rushton and Waldfogel 2007; Mosomi 2019a, 2019b). However, after controlling for age, longer tenure in the current job is protective of women’s earnings. More specifically, a considerably smaller share of women in the least affected group has only been in their current job for less than five years (45 percent) compared to those in the most affected group (72 percent) – a – 27 percentage point difference. In contrast, a larger share of women in the least affected group has been in their current job for five or more years compared to those in the most affected group (with a positive difference of around 10 percentage points, depending on length of tenure).

Some of the most important features distinguishing the least and most affected women have to do with family life and household composition, characterized in our model by the variables representing marital status, living with children, and household size. For example, being married is associated with substantial disadvantage, with only 27 percent of women in the least affected group being married (or cohabiting with a partner), compared to 51 percent of those in the most affected group, a significant difference of – 24 percentage points. In contrast, women least affected by the gap are far more likely to have never married compared to those most affected, with a 30-point difference between the two groups. Living with at least one child between 6 and 15 years is associated with an

even greater disadvantage, with only 20 percent of women in the least affected group living with a child this age compared to 64 percent in the most affected group, a percentage point difference of -44 . Even after controlling for children in the household, household size is still associated with a larger gender earnings gap, suggesting that living with other people in the household creates additional care responsibilities for women.

As has been documented in other work for South Africa, *where* a woman lives profoundly affects her life prospects, with African women in rural areas among the most vulnerable groups when measured across a range of objective and subjective indicators (Casale and Posel 2010a). Our analysis similarly finds that only 7 percent of women in the least affected group live in traditional rural areas, compared to 44 percent in the most affected group (a significant difference of -37 percentage points). In sharp contrast, 92 percent of women in the least affected group live in urban areas, compared to 44 percent in the most affected group (a significant difference of 48 percentage points). In a similar vein, living in the wealthier provinces of Gauteng and the Western Cape, which include the two largest metropolitan areas (Johannesburg and Cape Town), is associated with being least affected (with positive differences of 35 and 17 percentage points respectively compared to the most affected women), while living in the less well-off provinces of Northwest, KwaZulu-Natal, and Limpopo is associated with being most affected (differences of -22 , -17 , and -13 percentage points, respectively, compared to the least affected women).

In terms of employment characteristics, being in self-employment relative to wage employment confers the greatest disadvantage in terms of the size of the gender earnings gap. Almost no women in the least affected group are in self-employment, compared to 52 percent of those in the most affected group. While this might be capturing discrimination on the part of customers and suppliers, it probably also reflects the gender difference in the nature of the self-employment. Women's self-employment is more likely to be informal and piecemeal in nature while men's self-employment is more likely to be registered and in larger businesses that employ others (Casale, Posel, and Mosomi 2021).

The sectoral breakdown indicates that women in the most affected group are also substantially more likely to work in the sectors of wholesale and retail trade and community services, which tend to be dominated by women, than women in the least affected group, resulting in considerable percentage point differences of -23 and -28 points respectively. In contrast, women in the least affected group are much more likely to be working in finance relative to the most affected group (a difference of 28 percentage points). The analysis also shows that women in the least affected group are more likely to be working in private households (almost all of whom are domestic workers) compared to women in the most affected

group, although as we explain below, this is driven entirely by differences in the hours worked in this sector.

The summary statistics and graphs mapping heterogeneity in the gender earnings gap showed that part of the gap derives from women working fewer hours than men on average. We therefore repeat the classification analysis in Table 3 controlling for hours worked by using the log of hourly earnings as the outcome variable. While the results for many of the characteristics are similar to those in the previous analysis for the log of monthly earnings, there are some noteworthy differences.

The first is that the tenure variables lose significance and shrink in size. In other words, once hours worked are controlled for, tenure is no longer an important feature distinguishing those women least and most affected by the gender earnings gap. This is because length of tenure is correlated with working full time or working longer hours on the job, and the same family-related factors that likely affect the number of uninterrupted years a woman (or man) remains in a job may also affect the hours they are able to work in that job.

The second set of key differences between the results in Tables 2 and 3 similarly reflect the well-documented inequality in family-related responsibilities in South Africa (Casale, Posel, and Mosomi 2021). After controlling for hours worked, the benefit associated with being single shrinks (from 30 to 24 percentage points), as does the disadvantage associated with living with at least one child ages 6–15 years (from -44 to -30 percentage points) and living in larger households (from -1.59 to -0.97). This suggests that the care responsibilities associated with living with others, and especially living with children, prevent women from working longer hours or taking on jobs that demand long working hours. Nonetheless, even after accounting for variation in hours worked, large and significant differences between the least and most affected subgroups of women remain for the family-related variables.

Finally, the differences between the least and most affected groups become somewhat more muted for the location and sector variables after controlling for hours worked, as the kinds of jobs performed in different locations and sectors obviously vary, and correspondingly so would the hours worked in those jobs. However, the direction of the association and the significance levels are largely unchanged. The one exception is for the private household sector, for which the difference between least and most affected groups flips from 24 to -16 percentage points. In other words, after controlling for hours worked, being employed by a private household is associated with disadvantage rather than advantage. Almost all workers in this category are women domestic workers, with the small percentage of men working in private households generally employed as gardeners. This result is interesting because gardening is largely an informal occupation not regulated by South African labor law. Instead, in 2002, there were

SORTING THE GENDER EARNINGS GAP

Table 3 Classification analysis for employed women (least and most affected 20 percent) – hourly earnings gap

	<i>20% least affected</i>		<i>20% most affected</i>		<i>Difference (least-most affected)</i>		
	<i>Mean</i>	<i>SE</i>	<i>Mean</i>	<i>SE</i>	<i>Estimate</i>	<i>SE</i>	<i>P-vals</i>
<i>Individual characteristics:</i>							
Age	36.75	1.74	41.25	1.45	− 4.50	3.17	0.08*
Age squared	1450.3	144.4	1799.9	124.7	− 349.6	274.1	0.10*
Less than grade 12	0.42	0.07	0.62	0.06	− 0.20	0.12	0.05**
Grade 12	0.30	0.07	0.26	0.05	0.05	0.11	0.33
Grade 12 + post-secondary	0.27	0.07	0.12	0.04	0.15	0.11	0.09*
Tenure < 5 yrs	0.56	0.06	0.67	0.06	− 0.11	0.12	0.18
Tenure ≥ 5 yrs, < 10 yrs	0.25	0.05	0.15	0.04	0.10	0.09	0.13
Tenure ≥ 10 yrs, < 20 yrs	0.10	0.04	0.13	0.04	− 0.03	0.08	0.36
Tenure ≥ 20 yrs	0.08	0.05	0.05	0.03	0.04	0.07	0.32
Married	0.32	0.06	0.55	0.06	− 0.23	0.12	0.03**
Widowed	0.06	0.05	0.05	0.05	0.01	0.10	0.47
Divorced	0.03	0.02	0.04	0.03	− 0.01	0.05	0.39
Never married	0.59	0.07	0.35	0.06	0.24	0.12	0.03**
<i>Household characteristics:</i>							
At least one child 0–5	0.41	0.07	0.42	0.06	− 0.01	0.13	0.48
At least one child 6–15	0.33	0.06	0.63	0.07	− 0.30	0.13	0.01***
Household size	3.90	0.33	4.88	0.35	− 0.97	0.69	0.08*
<i>Location variables:</i>							
Urban area	0.80	0.05	0.50	0.07	0.30	0.11	0.00***
Traditional	0.15	0.04	0.39	0.06	− 0.24	0.10	0.01***
Commercial farming area	0.04	0.01	0.04	0.02	0.00	0.03	0.50
Mining areas	0.01	0.01	0.07	0.02	− 0.06	0.03	0.02**
Western Cape	0.16	0.03	0.00	0.00	0.16	0.03	0.00***
Eastern Cape	0.10	0.04	0.05	0.03	0.05	0.08	0.28
Northern Cape	0.00	0.00	0.02	0.01	− 0.01	0.02	0.19
Free State	0.07	0.04	0.03	0.02	0.04	0.06	0.22
KwaZulu-Natal	0.14	0.05	0.14	0.05	0.00	0.10	0.49
Northwest	0.00	0.00	0.27	0.04	− 0.27	0.04	0.00***
Gauteng	0.40	0.07	0.16	0.06	0.25	0.11	0.01***
Mpumalanga	0.11	0.04	0.07	0.04	0.03	0.07	0.32
Limpopo	0.02	0.02	0.26	0.06	− 0.24	0.07	0.00***
<i>Employment characteristics:</i>							
Self-employed	0.00	0.01	0.52	0.09	− 0.51	0.10	0.00***
Agriculture	0.13	0.03	0.00	0.00	0.13	0.03	0.00***
Mining	0.02	0.01	0.00	0.00	0.02	0.01	0.07*
Manufacturing	0.05	0.04	0.07	0.02	− 0.02	0.06	0.34

(Continued).

Table 3 Continued.

	20% least affected		20% most affected		Difference (least-most affected)		
	Mean	SE	Mean	SE	Estimate	SE	P-vals
Utilities	0.02	0.01	0.00	0.00	0.02	0.01	0.06*
Construction	0.00	0.01	0.05	0.03	− 0.05	0.03	0.06*
Wholesale and retail trade	0.16	0.07	0.38	0.08	− 0.22	0.12	0.03**
Transport	0.09	0.03	0.00	0.00	0.09	0.03	0.00***
Finance	0.44	0.06	0.01	0.01	0.43	0.07	0.00***
Community services	0.06	0.07	0.30	0.08	− 0.24	0.14	0.04**
Private households	0.02	0.03	0.18	0.08	− 0.16	0.11	0.07*

Notes: The sample is African adults (ages 16–64) with wage and self-employment. ***, **, * denote significance at the 1, 5, and 10 percent levels, respectively.

Source: Quarterly Labour Force Survey (2017: 1).

legislative amendments to the labor laws aimed at formalizing domestic worker contracts and mandating minimum hourly wages in the sector. This finding suggests that while legislative protections in occupations and sectors that are traditionally low paid and dominated by women and often open to exploitation are crucial, monitoring of compliance is equally important (Bhorat, Kanbur, and Mayet 2013).

As a robustness check, we conducted the analysis without what might be considered “bad controls,” namely variables that may be endogenous to earnings or themselves considered outcomes variable, such as the family-related variables and sector/industry of employment. Appendix Figure A1 and Table A1 present the results for the more limited “human capital specification.” As we would expect, the average partial effect is larger and the range of the SPEs somewhat narrower than the results based on the regressions with the fuller set of controls. Nonetheless, we still find substantial heterogeneity in the earnings gap. The APE when using monthly earnings is -0.370 , while the SPEs range from -0.795 to -0.039 . For hourly earnings, the APE is -0.243 while the range of the SPEs is -0.653 to -0.028 . Furthermore, the classification analysis produces results that are largely consistent with those obtained when the fuller specification was used, although as we might expect some of the differences are more pronounced given fewer controls in the regression.¹⁴

As a final exercise, we also performed the classification analysis for men, comparing the characteristics of men most and least affected or advantaged by the gender wage gap. The results (available from the authors on request) show that many of the characteristics associated with disadvantage among women are associated with advantage among men, although for some of the effects, the size and significance differ. Perhaps the most striking

findings concern marital status. As in the case of women, marital status is a key characteristic contributing to heterogeneity among men, but for men, marriage is associated with substantial advantage: 58 percent of men who are most advantaged by the gender wage gap are married, while 39 percent of the least advantaged are married, a difference of 19 percentage points. Similarly, while living with a child (ages 6–15) is associated with greater disadvantage for women, it is associated with greater advantage for men: 44 percent of men in the most advantaged group live with a child this age compared to only 13 percent of the least advantaged men, a difference of 31 percentage points.

DISCUSSION

In this study, we used a sorting method, developed in Chernozhukov, Fernandez-Val, and Luo (2018), to uncover considerable heterogeneity in the gender earnings gap in South Africa. The extent of this heterogeneity is not surprising given that South Africa has such high rates of inequality, a main source of which lies in the labor market. In addition, the sorting method allowed us to establish which women are most and which women least affected by gender earnings inequality, across the range of covariates included in the model.

The analysis identified marital status as an important source of difference among women (as well as among men). African women who are married experience a significantly greater earnings disadvantage than other African women. For men, the association is reversed: compared to men who are not married, married men are considerably more likely to be advantaged by the gender wage gap. The different associations between marriage and the gender earnings gap are likely to be partly explained by different selection effects. Research has documented that in South Africa, higher-earning African men are significantly more likely to marry, perhaps because they are better able to overcome the income hurdles (including bridewealth payments) to marriage (Casale and Posel 2010b). In contrast, less productive women, or women with less attachment to employment, may be selected into marriage (Blau and Kahn 2017). However, marriage may also contribute directly to earnings differences, providing benefits to men but reducing women's earnings potential partly because of their non-market responsibilities and because intramarital power differences may reduce women's autonomy in choosing the nature of their employment (Ahituv and Lerman 2007).

Women's responsibility for the provision of care would explain why African women who live with children also experience significantly more disadvantage in the labor market compared to other African women. This effect is large and significant in the analysis of the monthly gender earnings gap, but more muted (although still sizeable and significant)

for the hourly gender earnings gap analysis, signaling that an important source of disadvantage derives from women reducing their working hours. While the existence of a motherhood penalty in the labor market is well-documented in the feminist literature (Sigle-Rushton and Waldfogel 2007; Grimshaw and Rubery 2015), this finding carries additional weight in the South African context given the highly gendered nature of care. Low rates of marriage and cohabitation among Africans, coupled with high rates of extramarital childbearing, have resulted in women being far more likely than men to live with children (Posel and Hall 2021). It is not surprising then that most African children are cared for physically and financially by women (Hatch and Posel 2018). However, even where men and women are married or are living with a child, women spend more time on both care work and general household chores than men (Casale, Posel, and Mosomi 2021). Our analysis underlines the substantial disadvantage that responsibilities associated with both marriage and children confer on African women, and by implication, the children whom they support.

Many studies of the gender earnings gap in South Africa have focused on the wage employed, perhaps because information on the earnings of the self-employed is difficult to capture (Posel and Muller 2008; Mosomi 2019a, 2019b). Notwithstanding these concerns, our study showed that women in self-employment experience more gender earnings inequality than women in wage employment. This finding likely derives partly from the different kinds of businesses men and women are involved in (and the different constraints that they face), as well as perhaps reflecting the protective effects of labor legislation (including minimum wage legislation) governing wage employment.

The sorting method revealed also how spatial geographies, created during apartheid, still contribute considerably to inequality in South Africa. Compared to African women living in urban areas, those in tribal (or traditional) rural areas, where there are typically fewer employment opportunities and where patriarchal norms are perhaps more entrenched, are significantly more likely to be negatively affected by the gender wage gap. Analogously, a significant share of women who experience an earnings advantage relative to comparable men live in the wealthier province of Gauteng (which includes Johannesburg, the largest metropolitan area in the country), where earnings opportunities are far greater and more varied and gender norms more progressive.

There are several specific characteristics of the South African context, including its apartheid past, very high levels of inequality, low marriage rates, and the nature of household formation. Nonetheless, the ubiquity of the gender division of labor means that many of the insights uncovered in the South African labor market are likely to be relevant in other countries, and particularly the contribution of family-related variables to gender earnings inequality.

In addition, the sorting method applied in this study has broader appeal for the analysis of heterogeneity. By mapping the complete range of the partial effects for all observations in the sample, the method uncovers the extent of heterogeneity as well as the characteristics of people who experience the largest and smallest degrees of inequality. It does this in a way that is visually intuitive and easy to report, which may be particularly appealing for policy discussions. Further, in the case of gender earnings inequality, it is arguably more useful to focus on women (and their characteristics) in the tails of the distribution than on the characteristics associated with the average earnings gap when reflecting on policy implications. While South Africa may be an outlier in terms of the degree of earnings inequality, this method may be particularly useful in other countries with relatively high levels of inequality, not just in earnings but also in wealth and other dimensions of wellbeing.

A limitation of the sorted effects method, common to many other approaches that use observational data to analyze (gender) wage gaps, is that it is still based on observable characteristics measured in available micro-datasets. The method is therefore no more useful than other approaches in uncovering how unobserved factors (such as personality) contribute to earnings inequality. This shortcoming could be addressed partly through improved data collection, but perhaps it would be better illuminated through other experimental approaches. Nonetheless, given the frequency and availability of standard micro-data, such as labor force surveys, across both developed and developing regions, the sorted effects method, in our view, remains a valuable approach to mapping the extent and correlates of heterogeneity.

Dorrit Posel

*School of Economics and Finance, University of the Witwatersrand Johannesburg
Johannesburg, Gauteng, South Africa
email: dorrit.posel@wits.ac.za*

Dambala Gelo

*School of Economics and Finance, University of the Witwatersrand Johannesburg
Johannesburg, Gauteng, South Africa
email: dambala.kutela@wits.ac.za*

Daniela Casale

*School of Economics and Finance, University of the Witwatersrand Johannesburg
Johannesburg, Gauteng, South Africa
email: daniela.casale@wits.ac.za*

Adeola Oyenubi

*School of Economics and Finance, University of the Witwatersrand Johannesburg
Johannesburg, Gauteng, South Africa
email: adeola.oyenubi@wits.ac.za*

NOTES ON CONTRIBUTORS

Dorrit (Dori) Posel is a distinguished professor and holds the Helen Suzman Chair in the School of Economics and Finance at the University of the Witwatersrand. Much of her research over the past thirty years has been within the field of feminist economics, and she has published widely on issues related to marriage and family formation, labor migration, labor force participation and the gender division of labor, the economics and demographics of language use, and measures of well-being.

Dambala Gelo is Associate Professor in the School of Economics and Finance at the University of the Witwatersrand, South Africa. His research concerns applied microeconomics of development and environment with a focus on institutional and organizational questions, human capital, technology adoption, poverty trap, and energy transition.

Daniela Casale is Professor in the School of Economics and Finance at the University of the Witwatersrand, South Africa. Daniela is a development economist who has been involved in research in the field of feminist economics for over twenty years. Other broad areas of study include labor, well-being, health, and education. She has published widely across a range of economics, development, and public health journals.

Adeola Oyenu is Associate Professor in the School of Economics and Finance, University of the Witwatersrand, South Africa. Adeola's research interest is in the areas of health and development economics with a specific focus data driven approaches like Machine Learning.

ACKNOWLEDGMENTS

The authors thank Jacqueline Mosomi for all her input with the PALMS data, and the reviewers and the associate editor for the very helpful comments received.

NOTES

- ¹ Apartheid was a system of institutionalized racism or segregation that privileged “whites” (or people of European descent) over the other race groups in South Africa. It was formalized in 1948 after the National Party took power and ended in 1994 with the appointment of the first democratic African National Congress government.
- ² Four racial categories, which were designated during the Apartheid period, remain commonly identified in all household surveys and the Population Census conducted by the official statistical agency (Statistics South Africa): “Africans,” also referred to as black South Africans, who constitute just over 80 percent of the South African population; “Coloureds” (mixed race), “Indians/Asians,” and “Whites.”
- ³ This concern was addressed in subsequent household and labor force surveys in South Africa.

- ⁴ Note that alternatively we can represent the linear combination of the parameters and the covariates as follows: $\gamma x = x_i' \gamma = \sum_{j=1}^p \gamma_j x_{ij}$ and $\mu x = x_i' \mu = \sum_{j=1}^p \mu_j x_{ij}$, where j indicates the column of the covariate matrix, X .
- ⁵ A negative $\Lambda(x)$ is interpreted as a gender earnings penalty experienced by woman i with characteristics x_i , while a positive $\Lambda(x)$ is interpreted as a gender earnings premium experienced by woman i with characteristics x_i .
- ⁶ As explained earlier, this is different from existing literature using a decomposition approach (such as Oaxaca-Blinder) which commonly estimates the average partial effect as a counterfactual distribution that allows for decomposing the gap into the effects of coefficients, covariates, and residuals. In a similar vein, literature that uses unconditional quantile regressions analyzes heterogeneity but only at specific points along the outcome distribution (Melly 2005). In contrast to these approaches, the SPE method (i) focuses on the heterogeneity in the explained part among subpopulations defined by covariate values and (ii) uses this heterogeneity as a basis for classifying observational units into most or least affected groups and summarizing their characteristics. The focus on the extremes is arguably better for informing policy targeting.
- ⁷ The PALMS version of the QLFS data is created by a team at DataFirst at the University of Cape Town, who collate the labor market data from the official labor force surveys conducted by Statistics South Africa and package them in a comparable and reliable form for public use, along with suitable weights (Kerr, Lam, and Wittenberg 2019).
- ⁸ The last wave of the National Income Dynamics Study (NIDS) for South Africa, which was also conducted in 2017, has the advantage over the QLFS of interviewing each adult member of the household, rather than asking only one member of the household to respond on behalf of others, as is the case in the QLFS. However, information on the earnings of the self-employed is largely missing in NIDS and as a longitudinal survey, national representativeness is compromised by attrition across the waves.
- ⁹ The Rand-US Dollar exchange rate was R17.62 on February 6, 2023.
- ¹⁰ We use the “spe” command in the R Sorted Effects package (Chen et al. 2019) to produce the estimates in Figure 1. The 90 percent bootstrap uniform confidence bands are shaded in blue and are based on 500 bootstrap replications (this follows algorithm 2.1 in Chernozhukov, Fernandez-Val, and Luo 2018: 1920). The analysis was performed using survey weights as indicated earlier.
- ¹¹ The APE is marginally larger than the unconditioned average gap, reported in Table 1, signaling that women have characteristics which would favor higher earnings.
- ¹² Although the samples and specifications are not directly comparable, we find considerably more variation in the SPEs than do Chernozhukov et al. (2018) using data from the US.
- ¹³ We use the 20 percent cut-off to analyze the tails of the distribution as this provides us with a substantial enough sample size to compare the most and least affected groups, although other papers applying the method have also used 5 percent or 10 percent. As a robustness check, we re-estimated the classification analysis comparing the 10 percent most affected women to the 10 percent least affected women. We found our results to be largely consistent with the 20 percent cut-off analysis, although standard errors are larger because samples are smaller. As would be expected when using more extreme groups, the size of the differences in characteristics between the two groups is often larger than when using the 20 percent samples. To give one example, when the 10 percent sample is used, the difference in the probability of being married between the least and most affected groups of women is -29 percentage points (SE of 0.14) compared to -25 percentage points (SE of 0.11) when the 20 percent sample is used (see Table 2), although in both cases the difference is statistically significant.

- ¹⁴ Although not shown here due to space constraints, we also estimated the analysis on the sample of wage employed only, given concerns that the earnings data for the self-employed might be recorded in the survey with greater error than the earnings for the wage employed. Again, our results suggest substantial heterogeneity in the gender gap, and the classification analysis produces a consistent story (both when comparing the results for monthly and hourly earnings gaps, and when comparing those with and without a full set of controls as done above).

REFERENCES

- Ahituv, Avner and Robert I. Lerman. 2007. "How Do Marital Status, Work Effort and Wage Rates Interact?" *Demography* 44(3): 623–47.
- Allanson, Paul, Jonathan P. Atkins, and Timothy Hinks. 2000. "A Multilateral Decomposition of Racial Wage Differentials in the 1994 South African Labour Market." *Journal of Development Studies* 37(1): 93–120.
- Angrist, Joshua D. and Jörn-Steffen Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.
- Bhorat, Haroon and Sumayya Goga. 2013. "The Gender Wage Gap in Post-Apartheid South Africa: A Re-Examination." *Journal of African Economies* 22(2): 1–22.
- Bhorat, Haroon, Ravi Kanbur, and Natasha Mayet. 2013. "The Impact of Sectoral Minimum Wage Laws on Employment, Wages, and Hours of Work in South Africa." *IZA Journal of Labor & Development* 2: Article 1.
- Blau, Francine D. and Lawrence M. Kahn. 2017. "The Gender Wage Gap: Extent, Trends and Explanations." *Journal of Economic Literature* 55(3): 789–865.
- Blau, Francine D. and Anne E. Winkler. 2018. *The Economics of Women, Men and Work*. 8th ed. New York: Oxford University Press.
- Bliner, Alan S. 1973. "Wage Discrimination: Reduced Form and Structural Estimates." *The Journal of Human Resources* 8(4): 436–55.
- Casale, Daniela. 2004. "What has the Feminisation of the Labour Market 'Bought' Women in South Africa? Trends in Labour Force Participation, Employment and Earnings, 1995–2001." *Journal of Interdisciplinary Economics* 15(3–4): 251–75.
- Casale, Daniela and Dorrit Posel. 2010a. "Investigating the Well-Being of Rural Women in South Africa." *Agenda (Durban, South Africa)* 84: 44–50.
- . 2010b. "The Male Marital Earnings Premium in the Context of Bride Wealth Payments: Evidence from South Africa." *Economic Development and Cultural Change* 58(2): 211–30.
- . 2011. "Unions and the Gender Wage Gap in South Africa." *Journal of African Economies* 20(1): 27–59.
- Casale, Daniela, Dorrit Posel, and Jacqueline Mosomi. 2021. "Gender and the Economy in South Africa." In *The Oxford Handbook of the South African Economy*, edited by Fiona Tregenna, Imraan Valodia, and Arkebe Oqubay, 735–56. Oxford: Oxford University Press.
- Chen, Shuowen, Victor Chernozhukov, Ivan Fernandez-Val, and Ye Luo. 2019. "Sorted Effects: Sorted Causal Effects in R." arXiv: 1909.00836 [econ.EM].
- Chernozhukov, Victor, Ivan Fernandez-Val, and Ye Luo. 2018. "The Sorted Effects Method: Discovering Heterogeneous Effects Beyond Their Averages." *Econometrica* 86(6): 1911–38.
- Cock, Jacklyn. 1980. *Maids and Madams: A Study in the Politics of Exploitation*. Johannesburg: Ravan Press.

- DiNardo, John, Nicole M. Fortin, and Thomas Lemieux. 1996. "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach." *Econometrica* 64(5): 1001-44.
- Garenne, Michel, Stephen Tollman, Kathleen Kahn, Terri Collins, and Shirley Ngwenya. 2001. "Understanding Marital and Premarital Fertility in Rural South Africa." *Journal of Southern African Studies* 27(2): 277-90.
- Grimshaw, Damian and Jill Rubery. 2015. "The Motherhood Pay Gap: A Review of the Issues, Theory and International Evidence." Conditions of Work and Employment Series, No. 57. Geneva: International Labor Office.
- Grün, Carola. 2004. "Direct and Indirect Gender Discrimination in the South African Labour Market." *International Journal of Manpower* 25(3/4): 321-42.
- Hatch, Michelle and Dorrit Posel. 2018. "Who Cares for Children? A Quantitative Study of Childcare in South Africa." *Development Southern Africa* 35(2): 267-82.
- Hinks, Timothy. 2002. "Gender Wage Differentials and Discrimination in the New South Africa." *Applied Economics* 34(16): 2043-52.
- Hunter, Mark. 2010. *Love in the Time of AIDS*. Pietermaritzburg: University of KwaZulu-Natal Press.
- Kerr, Andrew, David Lam, and Martin Wittenberg. 2019. "Post-Apartheid Labour Market Series 1993-2019 [dataset]." Version 3.3. DataFirst, Cape Town.
- Kroczyk, Martin and Philipp Kugler. 2022. "Heterogeneous Effects of Monetary and Non-monetary Job Characteristics on Job Attractiveness in Nursing." SSRN 4233437. Social Science Research Network.
- Leibbrandt, Murray, Ingrid Woolard, Arden Finn, and Jonathan Argent. 2010. "Trends in the South African Income Distribution and Poverty Since the Fall of Apartheid." OECD Social, Employment and Migration Working Papers No. 101, Organisation for Economic Co-operation and Development.
- Maconachie, Moira. 1989. "Looking for Patterns of Women's Employment and Educational Achievements in the 1985 Census." *Agenda (Durban, South Africa)* 5: 80-92.
- Melly, Blaise. 2005. "Decomposition of Differences in Distribution Using Quantile Regression." *Labour Economics* 12(4): 577-90.
- Mincer, Jacob A. 1974. *Schooling, Experience, and Earnings*. Cambridge, MA: NBER.
- Mosomi, Jacqueline. 2018. "Distributional Changes in the Gender Wage Gap in the Post-Apartheid South African Labour Market." PhD thesis, University of Cape Town.
- . 2019a. "An Empirical Analysis of Trends in Female Labour Force Participation and the Gender Wage Gap in South Africa." *Agenda (Durban, South Africa)* 33(4): 29-43.
- . 2019b. "Distributional Changes in the Gender Wage Gap in the Post-Apartheid South African Labour Market." UNU-WIDER Working Paper 2019/17, UNU-WIDER.
- Muller, Colette. 2009. "Trends in the Gender Wage Gap and Gender Discrimination Among Part-Time and Fulltime Workers in Post-Apartheid South Africa." Working Paper 124. Durban, South Africa: University of KwaZulu-Natal.
- Ntuli, Miracle. 2007. "Exploring Gender Wage Discrimination in South Africa, 1995-2004: A Quantile Regression Approach." IPC Working Paper Series No. 56. Gerald R. Ford School of Public Policy, University of Michigan.
- Oaxaca, Ronald. 1973. "Male-Female Wage Differentials in Urban Labor Markets." *International Economic Review* 14(3): 693-709.
- Paul, Mark, Khaing Zaw, and William Darity. 2022. "Returns in the Labor Market: A Nuanced View of Penalties at the Intersection of Race and Gender in the US." *Feminist Economics* 28(2): 1-31.
- Posel, Dorrit and Daniela Casale. 2013. "The Relationship Between sex Ratios and Marriage Rates in South Africa." *Applied Economics* 45(3): 663-76.
- . 2019. "Gender and the Economy in Post-Apartheid South Africa: Changes and Challenges." *Agenda (durban, South Africa)* 33(4): 1-8.

- Posel, Dorrit, Daniela Casale, and Erofilo Grapsa. 2016. "Re-estimating Gender Differences in Income in South Africa: The Implications of Equivalence Scales." *Development Southern Africa* 33(4): 425–41.
- Posel, Dorrit, Daniela Casale, and Claire Vermaak. 2014. "Job Search and the Measure of Unemployment." *South African Journal of Economics* 82(1): 66–80.
- Posel, Dorrit and Katharine Hall. 2021. "The Economics of Households in South Africa." In *The Oxford Handbook of the South African Economy*, edited by Fiona Tregenna, Imraan Valodia, and Arkebe Oqubay, 800–22. Oxford: Oxford University Press.
- Posel, Dorrit and Colette Muller. 2008. "Is There Evidence of a Wage Penalty to Female Part-Time Employment in South Africa?" *South African Journal of Economics* 76(3): 466–79.
- Pullabhotla, Hemant K. and Mateus Souza. 2022. "Air Pollution from Agricultural Fires Increases Hypertension Risk." *Journal of Environmental Economics and Management* 115: 102723.
- Roberts, Gareth and Volker Schoer. 2021. "Gender-Based Segregation in Education, Jobs and Earnings in South Africa." *World Development Perspectives* 23: 100348.
- Rospabé, Sandrine. 2001. "An Empirical Evaluation of Gender Discrimination in Employment, Occupation Attainment and Wage in South Africa in the Late 1990s." DPRU/FES Workshop on Labour Markets and Poverty in South Africa, Johannesburg, November.
- Shepherd, Debra. 2008. "Post-Apartheid Trends in Gender Discrimination in South Africa: Analysis through Decomposition Techniques." Stellenbosch Economic Working Papers: 06/08, University of Stellenbosch, South Africa.
- Sigle-Rushton, Wendy and Jane Waldfogel. 2007. "Motherhood and Women's Earnings in Anglo-American, Continental European, and Nordic Countries." *Feminist Economics* 13(2): 55–91.
- Spaull, Nic and Nwabisa Makaluza. 2019. "Girls Do Better: The Pro-Female Gender Gap in Learning Outcomes in South Africa 1995–2018." *Agenda (Durban, South Africa)* 33(4): 11–28.
- Tijdens, Kea G. 2002. "Gender Roles and Labor Use Strategies: Women's Part-Time Work in the European Union." *Feminist Economics* 8(1): 71–99.
- Van Broekhuizen, Hendrik and Nic Spaull. 2017. "The 'Martha Effect': The Compounding Female Advantage in South African Higher Education." Stellenbosch Economic Working Paper 14. Stellenbosch: Stellenbosch University.
- Walker, Cheryl, ed. 1990. *Women and Gender in Southern Africa to 1945*. Cape Town: David Philip Publishers.
- Wittenberg, Martin. 2017. "Wages and Wage Inequality in South Africa 1994–2011: Part 2 – Inequality Measurement and Trends." *South African Journal of Economics* 85(2): 298–318.
- World Bank. 2022. *Inequality in Southern Africa: An Assessment of the South African Customs Union*. Washington, DC: International Bank for Reconstruction and Development/World Bank.

SORTING THE GENDER EARNINGS GAP

APPENDIX

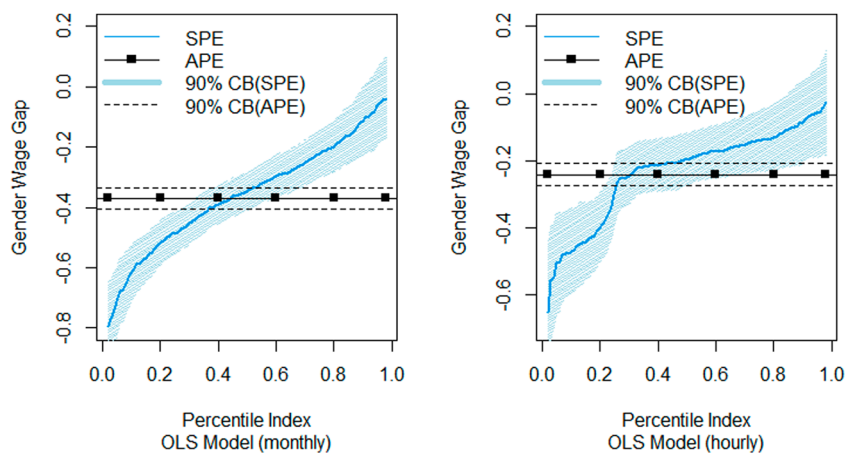


Figure A1 Sorted effects of the gender wage gap for women with wage or self-employment, using a limited set of controls.

Notes: APE ~ average partial effect; SPE ~ sorted partial effect; CB ~ confidence band.

Source: Quarterly Labour Force Survey (2017: 1).

Table A1 Classification analysis for employed women (least and most affected 20 percent) – monthly and hourly earnings gap, using a limited set of controls

	<i>Monthly earnings gap</i>			<i>Hourly earnings gap</i>		
	<i>Difference (least-most affected)</i>			<i>Difference (least-most affected)</i>		
	<i>Estimate</i>	<i>SE</i>	<i>P-val</i> s	<i>Estimate</i>	<i>SE</i>	<i>P-val</i> s
<i>Individual characteristics:</i>						
Age	− 12.10	2.77	0.00***	− 9.27	3.48	0.00***
Age squared	− 961.95	225.70	0.00***	− 722.1	281.7	0.01***
Less than grade 12	− 0.67	0.10	0.00***	− 0.47	0.16	0.00***
Grade 12	0.06	0.13	0.31	− 0.04	0.15	0.39
Grade 12 + post-secondary	0.61	0.13	0.00***	0.52	0.19	0.00***
Tenure < 5 yrs	− 0.21	0.12	0.04**	− 0.09	0.15	0.26
Tenure ≥ 5 yrs, < 10 yrs	0.13	0.09	0.08*	0.16	0.13	0.11
Tenure ≥ 10 yrs, < 20 yrs	0.02	0.08	0.43	− 0.08	0.10	0.22
Tenure ≥ 20 yrs	0.06	0.07	0.18	0.01	0.10	0.45
<i>Location variables:</i>						
Urban area	0.53	0.11	0.00***	0.35	0.12	0.00***
Traditional	− 0.43	0.10	0.00***	− 0.35	0.12	0.00***
Commercial farming area	0.00	0.03	0.50	0.08	0.04	0.02**
Mining areas	− 0.10	0.02	0.00***	− 0.08	0.02	0.00***
Western Cape	0.17	0.04	0.00***	0.20	0.03	0.00***
Eastern Cape	0.04	0.08	0.30	0.03	0.11	0.39
Northern Cape	− 0.01	0.02	0.29	− 0.02	0.02	0.22
Free State	0.06	0.06	0.16	0.07	0.08	0.20
KwaZulu-Natal	− 0.14	0.11	0.10*	− 0.03	0.10	0.37
Northwest	− 0.26	0.04	0.00***	− 0.29	0.03	0.00***
Gauteng	0.39	0.12	0.00***	0.28	0.15	0.03**
Mpumalanga	− 0.05	0.07	0.27	0.04	0.08	0.32
Limpopo	− 0.22	0.07	0.00***	− 0.27	0.09	0.00***
<i>Employment characteristics:</i>						
Self-employed	− 0.44	0.08	0.00***	− 0.56	0.10	0.00

Notes: The sample is African adults (ages 16–64) with wage and self-employment. ***, **, * denote significance at the 1, 5, and 10 percent levels, respectively.

Source: Quarterly Labour Force Survey (2017: 1).