The Evolution of Female and Male Earnings Inequality in post-Apartheid South Africa

Janina Hundenborn\textsuperscript{1,3} Ingrid Woolard\textsuperscript{2,3} Murray Leibbrandt\textsuperscript{1,3}

\textsuperscript{1}University of Cape Town
\textsuperscript{2}University of Stellenbosch
\textsuperscript{3}Southern African Labour and Development Unit

WIDER Development Conference
12 September 2019
1. Background
   Motivation

2. Literature
   South African Labour Market
   Micro-Simulations

3. Data
   Issues
   Summary Statistics

4. Methodology
   Microsimulation Approach

5. Counterfactual Micro-Simulations
   Estimating Effects

6. Conclusion
   Results
Motivation

Despite affirmative action laws and other policy interventions implemented since the first democratic government in 1994, the inherited inequalities have worsened in the post-apartheid era. This increase is driven by inequality from the labour market (Leibbrandt et al., 2010; Hundenborn et al., 2016).
Motivation

Despite affirmative action laws and other policy interventions implemented since the first democratic government in 1994, the inherited inequalities have worsened in the post-apartheid era. This increase is driven by inequality from the labour market (Leibbrandt et al., 2010; Hundenborn et al., 2016).

When analyzing the increase in inequality, there are two important issues to account for:

1. Due to the high levels of inequality, analysis has to go beyond the mean.

2. Imperative to account for the high levels of unemployment prevalent in South Africa.
Motivation

Despite affirmative action laws and other policy interventions implemented since the first democratic government in 1994, the inherited inequalities have worsened in the post-apartheid era. This increase is driven by inequality from the labour market (Leibbrandt et al., 2010; Hundenborn et al., 2016).

When analyzing the increase in inequality, there are two important issues to account for:

1. Due to the high levels of inequality, analysis has to go beyond the mean.

2. Imperative to account for the high levels of unemployment prevalent in South Africa.

Therefore, our research applies advanced micro-simulations to the earnings distribution which model the distortion effects of structural unemployment.
Wittenberg (2016a,b) offer a detailed analysis of earnings inequality and of measurement issues underlying PALMS data set. Findings show compression of incomes below the mean while top of the distribution moved away from the median.


Wittenberg and Ntuli (2013) investigate the changes in the labour force participation of black women using regression analysis.

Kwenda & Ntuli (2018) use RIFs to show that gender pay gap in SA is larger in the private sector.

Van der Westhuizen et al. (2007) relate to the work of Bhorat and Leibbrandt (1999) as well as Casale (2004). Their findings manifest the increase in female unemployment in the first decade after apartheid.
Literature on Micro-Simulations

- Oaxaca-Blinder decomposition (Oaxaca, 1973 and Blinder, 1973); ‘price effect’ and ‘endowment effect’.
- Growing literature that is looking for methods to decompose differences across entire distributions rather than across means (including Dinardo et al., 1996; Lemieux, 2002; Firpo et al., 2009; Fortin et al., 2011).
- Bourguignon et al. (2008) moved ‘Beyond Oaxaca-Blinder’, allowing -among other things- to decompose the ‘price effect’ and ‘endowment effect’ not just at the mean but rather across the entire distribution.
- Garlick (2016) finds that changes in the distribution of education increased inequality in total labour earnings.
Stacked cross sectional dataset containing micro-data from 54 household surveys conducted by Statistics South Africa between 1994 and 2017, as well as the 1993 PSLSD.

Detailed comparable information on demographics and labour market participation across years.

Finn and Leibbrandt (2018) find a significant shift in earnings inequality after 2012 which is possibly caused by a change in the methodology of imputations performed by Stats SA.

Therefore, years chosen for this study range from 1993 to the second quarter of 2012.
Distribution of Earnings between 1993 and 2012

- Increase in Gini coefficient for both men and women, larger for men.
- Median incomes increased for women but not for men.
- Mean incomes increased significantly for women.
- Top of the earnings distribution moving away from the median (Wittenberg, 2016b).

**Table 1: Distributional Statistics of Real Earnings**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Female</td>
<td>Male</td>
<td>Female</td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td>Median</td>
<td>1 822.44</td>
<td>996.73</td>
<td>1 728.51</td>
<td>1 364.26</td>
<td>1 412.43</td>
<td>753.30</td>
</tr>
<tr>
<td>Mean</td>
<td>3 387.86</td>
<td>1 919.08</td>
<td>3 089.60</td>
<td>1 965.42</td>
<td>2 450.14</td>
<td>1 642.20</td>
</tr>
<tr>
<td>Gini</td>
<td>0.591</td>
<td>0.583</td>
<td>0.563</td>
<td>0.511</td>
<td>0.550</td>
<td>0.583</td>
</tr>
</tbody>
</table>

Source: Authors’ calculations based on weighted PALMS Data (1993 - 2012)

All earnings values are presented in 2000 Rand.
Descriptive Statistics

- Overall participation has increased for women but decreased for men.
- Male participation is consistently higher than female participation.
- Employment of males is significantly higher than that of females.
- Africans are consistently less likely to be employed, and African women are the most marginalized (Ntuli and Wittenberg, 2013).
- Average education levels increased to almost 10 years for men and women.
- Small increase of women in higher skilled occupations (management, professionals).
- Still very large share of women in elementary occupations including domestic workers.
Simulating changes in earnings distribution

Borrowing from Bourguignon et al. (2008) and González-Rozada and Menendez (2006), the distribution of labour income depends on the participation rate $P_t$, the employment rate $E_t$, the occupational structure $O_t$, observed characteristics $X_t$, the returns to individual characteristics $R_t$ and unobservable components $\epsilon_t$:

$$Y_{Lt} = f(P_t, E_t, O_t, X_t, R_t, \epsilon_t)$$ (1)

It is possible to estimate any inequality measure over this distribution

$$\vartheta_t = \Phi(f(P_t, E_t, O_t, X_t, R_t, \epsilon_t))$$

To assess the effect of changes in the participation rate for example, we estimate the difference of $\vartheta - \hat{\vartheta}$ where

$$\hat{\vartheta} = \Phi(f(P_{t+1}, E_t, O_t, X_t, R_t, \epsilon_t))$$ (2)
In order to calculate the reduced-form equation (1), a set of structural equations have to be estimated:

1. Individual Labour Market Outcomes are calculated through a bivariate probit model using the Heckman selection correction.
2. Individual Occupational Allocation is estimated using an ordered logit model.
3. Individual Wages can be estimated with a log-linear Mincer model.

The coefficients estimated through these steps are used to simulate counterfactuals by changing one element at a time and comparing the differences of the distributions.
Table 2: Counterfactual Estimations

<table>
<thead>
<tr>
<th>Counterfactual</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_{lt} = f(P_{t+1}, E_t, O_t, X_t, R_t, \varepsilon_t)$</td>
<td>Individuals’ participation choice in period $t$ is simulated as if the choice was made in period $t + l$. The counterfactual participation sample takes size $N^*_P$. Allocation into participation is semi-random.</td>
</tr>
<tr>
<td>$Y_{lt} = f(P_{t+1}, E_{t+l}, O_t, X_t, R_t, \varepsilon_t)$</td>
<td>Additional to the above, the employment outcome in period $t$ is simulated as if in period $t + l$. The counterfactual employment sample takes size $N^*_E$. Allocation into employment is semi-random.</td>
</tr>
<tr>
<td>$Y_{lt} = f(P_{t+1}, E_{t+l}, O_{t+l}, X_t, R_t, \varepsilon_t)$</td>
<td>Additional to the above, individuals are allocated into counterfactual occupations. No randomization.</td>
</tr>
<tr>
<td>$Y_{lt} = f(P_{t+1}, E_{t+l}, O_{t+l}, X_{t+l}, R_t, \varepsilon_t)$</td>
<td>Additional to the above, socio-demographic characteristics are transformed in a rank-preserving exercise. The population weight is re-scaled.</td>
</tr>
<tr>
<td>$Y_{lt} = f(P_{t+1}, E_{t+l}, O_{t+l}, X_{t+l}, R_{t+l}, \varepsilon_t)$</td>
<td>Additional to the above, returns to characteristics from period $t + l$ are applied to the counterfactual employment sample.</td>
</tr>
<tr>
<td>$Y_{lt} = f(P_{t+1}, E_{t+l}, O_{t+l}, X_t, R_t, (\hat{\sigma}^{t+l}/\hat{\sigma}^t)\varepsilon_t)$</td>
<td>Additional to the above, the distribution of the residuals is re-scaled.</td>
</tr>
</tbody>
</table>
A logical first step is to simulate a counterfactual labour force by applying the coefficients that determine participation in $t + l$ to individuals’ characteristics in period $t$.

- Estimate a linear prediction index $S^*_p$ of individuals in time $t$ where
  $$S^*_p = \beta_{t+l}^P X_t.$$  
- Randomizing the allocation into participation by drawing a random number $\xi_1$ from a standard uniform distribution $U(0, 1)$.

$$P^*_p = \frac{\exp(\beta_{t+l}^P X_t^P + \xi_1)}{1 + \exp(\beta_{t+l}^P X_t^P + \xi_1)}.$$  

The distribution of earnings is estimated for the individuals identified in the counterfactual participation sample $N^*_P = N_t \times P_{t+l}$. 

\[ (3) \]
Counterfactual Estimations - Women

Figure 1: Decomposition of Average Effects for Women 1993 - 2012
Counterfactual Estimations - Men

Figure 2: Decomposition of Average Effects for Men 1993 - 2012

Janina Hundenborn
Evolution of earnings inequality
UNU-WIDER Conference
Results

- Changes in the participation of women accounts for very little.
- The changes in employment rates as well as in the expected returns to employment increased inequality significantly.
- The endowment effect also had an increasing effect on earnings inequality; particularly for women.
- The price effect was negative for both men and women but significantly larger for men after effect of changes in characteristics has already been accounted for.
- Interaction effects are larger for women then for men. Decreasing inequality for both.
High levels of earnings inequality in South Africa between 1993 and 2012.

Unreasonable shift in inequality measured after 2012, therefore no more recent data could be included (Wittenberg, 20116a,b; Finn and Leibbrandt, 2018).

Micro-simulations new tool for in-depth examination of trends in the evolution of earnings inequality.

Improve access to employment given the discrepancy between participation and employment, particularly for women.

Need for addressing inequality of education.
Thank you for your attention.
1. Background
   Motivation

2. Literature
   South African Labour Market
   Micro-Simulations

3. Data
   Issues
   Summary Statistics

4. Methodology
   Microsimulation Approach

5. Counterfactual Micro-Simulations
   Estimating Effects

6. Conclusion
   Results
### Table 3: Participation Rates of Working Age Population

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>0.843</td>
<td>0.798</td>
<td>0.852</td>
<td>0.850</td>
<td>0.794</td>
<td>0.804</td>
</tr>
<tr>
<td>Female</td>
<td>0.592</td>
<td>0.575</td>
<td>0.732</td>
<td>0.731</td>
<td>0.627</td>
<td>0.647</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>By Population group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>African</td>
<td>0.827</td>
<td>0.772</td>
<td>0.843</td>
<td>0.845</td>
<td>0.780</td>
<td>0.792</td>
</tr>
<tr>
<td>Coloured</td>
<td>0.867</td>
<td>0.853</td>
<td>0.875</td>
<td>0.854</td>
<td>0.831</td>
<td>0.839</td>
</tr>
<tr>
<td>Indian/Asian</td>
<td>0.891</td>
<td>0.884</td>
<td>0.885</td>
<td>0.855</td>
<td>0.824</td>
<td>0.825</td>
</tr>
<tr>
<td>White</td>
<td>0.894</td>
<td>0.877</td>
<td>0.873</td>
<td>0.873</td>
<td>0.852</td>
<td>0.858</td>
</tr>
</tbody>
</table>

Source: Authors’ calculations based on weighted PALMS Data (1993 - 2012)

### Table 4: Employment Rates Conditional on Participation

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>0.756</td>
<td>0.678</td>
<td>0.673</td>
<td>0.716</td>
<td>0.710</td>
<td>0.710</td>
</tr>
<tr>
<td>Female</td>
<td>0.683</td>
<td>0.532</td>
<td>0.528</td>
<td>0.644</td>
<td>0.644</td>
<td>0.645</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>By Population group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>African</td>
<td>0.687</td>
<td>0.614</td>
<td>0.609</td>
<td>0.667</td>
<td>0.666</td>
<td>0.666</td>
</tr>
<tr>
<td>Coloured</td>
<td>0.838</td>
<td>0.771</td>
<td>0.771</td>
<td>0.769</td>
<td>0.735</td>
<td>0.735</td>
</tr>
<tr>
<td>Indian/Asian</td>
<td>0.925</td>
<td>0.842</td>
<td>0.853</td>
<td>0.905</td>
<td>0.869</td>
<td>0.869</td>
</tr>
<tr>
<td>White</td>
<td>0.969</td>
<td>0.938</td>
<td>0.944</td>
<td>0.941</td>
<td>0.944</td>
<td>0.944</td>
</tr>
</tbody>
</table>

Source: Authors’ calculations based on weighted PALMS Data (1993 - 2012)
Using the participation effect as an example, for the randomization process we...

- draw random numbers from a uniform distribution $U(0, 1)$ for each member of the sample in period $t$
- multiply this number by the standard deviation of the participation scores in $t + 1$
- compute a rescaled probability of participation $P^*_P$ which includes both the deterministic elements of observed characteristics and estimated returns to characteristics and the random component
- sort $P^*_P$ and select $N^*_P = N_t \times P_{t+1}$ for the counterfactual sample
- repeat this step with a new random number draw for each round of the estimation
Since the probability of selection is dependent on the probability of participation, those with a higher participation probability are more likely to be in the counterfactual sample $N_P^*$ but through the repeated drawing, some individuals with high probabilities of participation may not be selected, and some individuals with low probabilities of participation may be selected. This aims to simulate labour market distortions that may prevent some individuals from achieving their optimal participation choice. The estimation can be repeated for any number of times, we choose 1000 repetitions and calculate confidence intervals to interpret the results.
Bhorat, H., and Leibbrandt, M.  
Correlates of vulnerability in the South African labour market.  

Blinder, A. S.  
Wage discrimination: Reduced form and structural estimates.  

Bourguignon, F., Ferreira, F. H., and Leite, P. G.  
Beyond Oaxaca-Blinder: Accounting for differences in household income distributions.  

Casale, D.  

DiNardo, J., Fortin, N. M., and Lemieux, T.
*Econometrica 64*, 5 (September 1996), 1001–1044.

**Finn, A., and Leibbrandt, M.**
The evolution and determination of earnings inequality in post-apartheid South Africa. 

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

**Fortin, N., Lemieux, T., and Firpo, S.**
Decomposition methods in economics. 

**Garlick, J.**
Changes in the inequality of employment earnings in South Africa. 


LEMIEUX, T.

NTULI, M., AND WITTENBERG, M.

OAXACA, R.

VAN DER WESTHUIZEN, C., GOGA, S., AND OOSTHUIZEN, M.

WITTENBERG, M.
WITTENBERG, M.