Labour Market Inequality in South Africa: A decomposition of changes in earnings from 2001 to 2011

A dissertation submitted to the Economics Department, University of Cape Town, in partial fulfilment of the requirements for the award of a Master's Degree in Applied Economics.

By Scott Hosking

Under supervision of Haroon Bhorat

Plagiarism Declaration

I, Scott Hosking hereby declare that the work on which this dissertation is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university.

1. I know that plagiarism is wrong. Plagiarism is to use another's work and pretend that it is my own.

2. I have used the Harvard referencing guide for citation and referencing. Each contribution to, and quotation in this thesis from the work(s) of other people has been contributed, and has been cited and referenced.

3. This thesis is my own work.

4. I have not allowed, and will not allow, anyone to copy my work.

Signature: Signed by candidate

Date: 2016 – 11 - 10

Contents

List of Figures ... 4

List of Tables ... 4

Acknowledgments ... 5

Abstract ... 6

Section 1: Introduction ... 6

Section 2: Tasks and Technological Change ... 9

Section 3: The South African Economy, Labour Market and Tasks ... 11

South African Literature ... 15

Returns to Education ... 15

Unions ... 17

Minimum Wages ... 18

Section 4: Methodology ... 19

Wage Setting Model and Tasks ... 19

Decomposition Framework ... 21

Oaxaca-Blinder Decomposition ... 22

Quantile Decomposition ... 23

Specification ... 25

Inference ... 26

Section 5: Data ... 27

Wage Data ... 27

Measuring Tasks ... 28

Descriptive Statistics ... 29

Reweighting ... 30

Section 6: Analysis ... 30

RIF-Regressions ... 30

Decomposition Results ... 32

Main Composition Effects ... 34

Main Wage Structure Effects ... 34

Secondary results ... 36

Section 7: Conclusion ... 37

Bibliography ... 39

Appendix A: Inference with Multiple Imputation ... 42

Appendix B: RIF-Regression Tables and Graphs ... 43

Appendix C: Data and Reweighting ... 44

List of Figures

Figure 1: Change in earnings by percentile between 2001 and 2011 ... 7

Figure 2:Log of real hourly earnings kernel densities in 2001 and 2011 ... 8

Figure 3: Task Intensity Across the Earnings Distribution ... 15

Figure 4:Distribution of educational attainment by birth cohort category, adult males 25-50 ... 16

Figure 5: RIF Coefficients at each percentile in the earnings distribution ... 32

Figure 6: Selected Wage Structure and Composition Effects ... 35

List of Tables

Table 1: Income Inequality Metrics for Wage Earners in 2001 and 2011 ... 7

Table 2: Between and Within Sector Occupation Shifts ... 13

Table 3: Sector share of occupation employment ... 14

Table 4: Aggregate Decomposition Results ... 33

Table 5: Main Effects ... 34

Table 6: Sector and Discrimination Effects ... 36

Table 7: RIF Regression Coefficients ... 43

Table 8: Sector contributions to GVA over time ... 44

Table 9: Descriptive Statistics ... 45

Table 10: Reweighting Regression ... 47

Table 11: Codes for table 10 ... 50

Table 12: Occupation Task Scores ... 50

Acknowledgments

This author is deeply indebted to Haroon Bhorat for his supervision of this thesis. Likewise, the code made publicly available by Nicole Fortin and the task measurement data by David Dorn were at the heart of the analysis, and without which it is very unlikely that thesis would have been completed.

Abstract

The relatively stable overall wage inequality in South Africa between 2001 and 2011 has hidden two distinct trends. Strong growth above the median for high wage earners has increased inequality at the top of the earnings distribution, whilst similarly, strong growth below the median has decreased inequality at the bottom of the distribution. This paper uses the 'task' approach alongside a Recentred Influence Function decomposition framework to explore the factors associated with this pattern of change. The findings suggest that routine-biased technical change and minimum wage laws enacted over the decade have important roles to play in the changes.

Section 1: Introduction

The extraordinarily high levels of household income inequality in South Africa have been well documented (see Van Der Berg, 2010; Leibbrandt et al., 2010 and Leibbrandt et al., 2012). Despite the introduction of an extensive grant support system, the South African labour market still accounts for the overwhelming majority of household income. It thus has an integral role to play in influencing the country's income inequality levels. (Leibbrandt et al., 2012). Trends in labour market earnings are thus central to any discussion concernining South Africa's income inequality, as any change in the overall distribution of labour earnings will have material effects on country's household income distribution. A recent paper by Wittenberg (2014) shows that wage changes in the past 15 years have resulted in a systematic shift in the employee earnings distribution. While there has been relatively little change in the overall wage inequality, as measured by the Gini coefficient, there has been a radical shift in the nature of that inequality. Table 1 uses the PALMS dataset employed by Wittenberg (2014) to reproduce some of the paper's results over the 2001 to 2011 period.[1]

The period saw the wage Gini grow by 5.06%. A closer look at quantile ratios however shows that this overall increase hides two large and countervailing effects. The 9.28% increase in the p90/p50 earnings ratio reflects strong growth at the top of the distribution relative to the middle, as would be expected with an increase in the Gini. At the same time, however the 9.52% in the p10/p50 reflects lower inequality at the bottom of the distribution. This can be seen in the fact that there is virtually no change in the P90/p10 ratio. Graphically, these dynamics can be clearly seen in Figure 1. The graph shows the difference between the log of real hourly earnings in 2001 and 2011 at each percentile of the distribution. What the data suggests is that there has been steady relative growth at the bottom of the distribution, and to a lesser degree at the top of the distribution. At the same time growth between the 30th and 60th percentiles has stagnated. This collapse in the relative earnings of the middle distribution is representative of a fundamental shift in the earnings profile of the South African labour market. This is well captured by comparing kernel density plots in 2001 and 2011. Figure 3 shows that, while there has been a compression of the distribution between the 10th and 50th percentiles in 2011 relative to 2001, there has been a fanning out of the distribution between the 50th and 90th percentiles.

[1] The data is introduced and described fully in Section 5

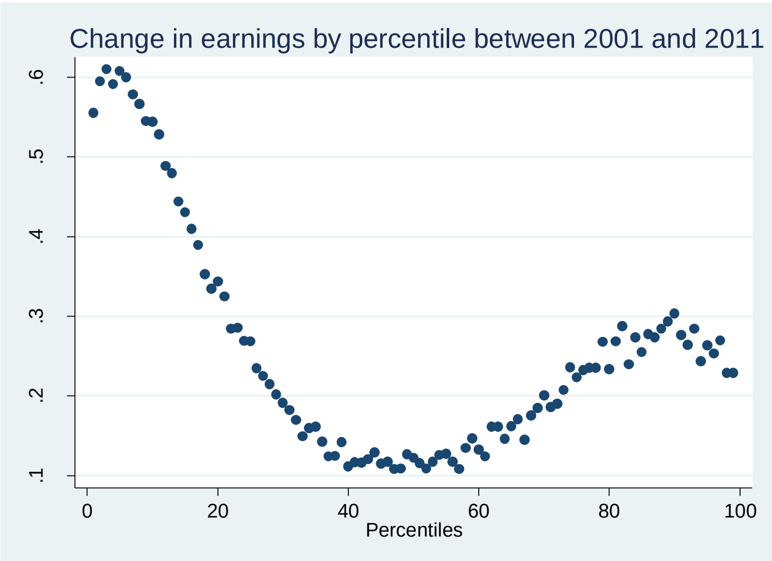
Table 1: Income Inequality Metrics for Wage Earners in 2001 and 2011

Inequality Measures	**2001**	**2011**	**Percentage Change**
90/10	12.498	12.460	-0.3
	(0.419)	(0.153)	
90/50	3.389	3.703	9.28
	(0.066)	(0.056)	
10/50	0.271	0.297	9.52
	(0.009)	(0.004)	
Gini	0.498	0.523	5.06
	(0.006)	(0.003)	
Median	1700.74	2139.18	25.78
	(16.30)	(32.35)	
Mean	2642.68	3727.27	41.04
	(15.58)	(25.15)	

Sample reflect wage earners between the ages of 15 and 65. Errors calculation is explained in Section 4.1.

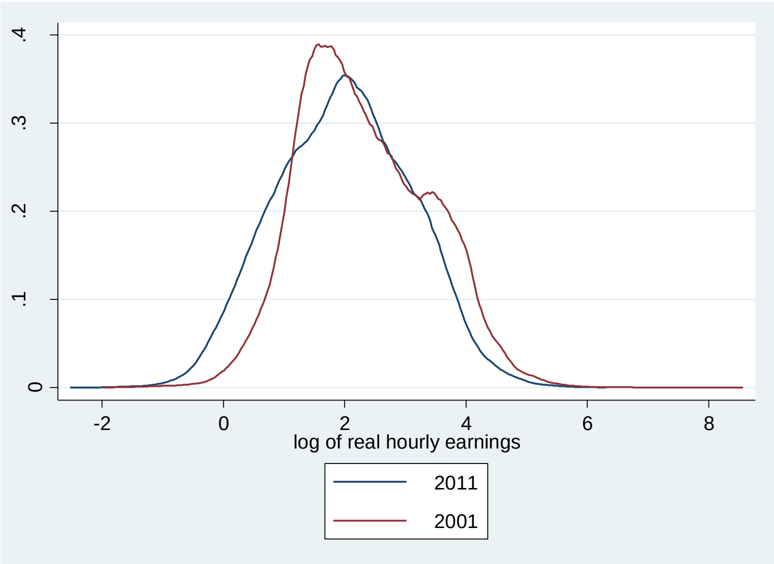
Source: PALMS dataset, Authors calculations.

Figure 1: Change in earnings by percentile between 2001 and 2011



Source 1: Author's calculations using PALM's data

Figure 2:Log of real hourly earnings kernel densities in 2001 and 2011



Source 2: PALMS Data, Authors calculations

What are the forces driving these wage trends? A range of studies that focused on just the mean have found that the large - and growing – tertiary education premium has played a significant role in driving inequality in the country (Branson et al. 2013; Keswell & Poswell 2004; Hertz 2003). Likewise, it is probably the case that institutions have played a role, with a range of studies having shown that being part of a union is associated with a wage premium at the mean, and several recent ones providing evidence that minimum wages have had a significant effect on the average incomes of workers in the impacted sectors (Banerjee et al. 2008; Bhorat, Goga, et al. 2012; Bhorat et al. 2013). Furthermore, several authours have put forward evidence of skills-biased biased technological change in the country, suggesting that this could also be an important factor shaping he wage distribution(Bhorat et al. 2013). The studies, while indicative, do not provide insight into the role each of the factors play at different points in the distribution. The conditional quantile regression, which is a popular tool in labour economics, does allow for a closer look at the relative effects at different points. A weakness it suffers, however, is that it returns the effect of X on a conditional quantile of Y given X, rather than the unconditional effect. Typically, we care about the latter. For example, we care about the effect of unions on the 30th percentile of the *wage distribution*, rather than the effect of distributions on the university graduate that is performing poorly enough to be at the 30th percentile (Autor 2012). A recent paper by Firpo et al., (2007) introduces an unconditional quantile regression that gets around the problem. This allows for the estimation of the unconditional effects of covariates on the wage distribution, as well as for the application of detailed quantile decompositions.

An important use that Firpo et al.,(2013) put the approach to in a follow up paper is estimating and comparing the effects of technology on the wage distribution using tasks. The approach is centred on the idea that different types of occupations are alternatively complemented or substituted for by technology, depending on what type of tasks they are associated with. The value in this more nuanced approach is that it accommodates technology having non-monotonic effects through the earnings distribution. Indeed a significant body of literature in the USA has argued, based on the approach, that technological change has served to lower wages in the middle of the distribution relative to the top and bottom (Acemoglu & Autor 2011; Autor et al. 2008). We follow Firpo et al.,(2013) in using the approach for a similar exploration using the South African data. In doing so, we are seeking to test whether the same forces are at play shaping the wage distribution in South Africa, a developing country, as those occurring in a developed country. Overall, the paper finds important roles for technology, education and the minimum wage in changes in the distribution. The remainder of the paper will proceed as follows. Section 2 introduces the international task literature. Section 3, using tools from the task literature, makes the case for technological change playing an important role in shaping wages in South Africa before reviewing the South African wage formation literature. Section 4 introduces the Firpo et al., (2013) methodology, and links it to the wage determinants introduced in the literature review. Section 5 introduces the data, Section 6 presents the results and, finally, Section 7 concludes.

Section 2: Tasks and Technological Change

The dialogue on technological change in the international literature has, for the most part, viewed it as skills biased. This refers to technology complementing high-skilled workers, thereby increasing labour market demand for them. The increased demand is then responsible for the earnings premium growth associated with college workers (Katz & Autor 1999). This, however, is not the story in its entirety. A recent branch of literature has developed a more nuanced approach to technological change. In a seminal contribution, Autor, Levy and Murnane (2003) (henceforth ALM) develop an innovative approach to argue that technology has had non-monotonic effects across the wage distribution. Rather than focusing on the endowments of workers, their education and experience levels, they look at the tasks to which the workers apply those endowments to perform. The argument centres around the rapid advances in the capability of computer technology, and the related price decreases in computer capital. These have had important implications for firms' decisions regarding their labour-capital input mix.

ALM divide tasks into three distinct categories. Routine tasks are well-defined, repetitive and simple to codify. This, in turn, means they are easily substituted by capital, with the result that labours' contributions to these types of tasks diminishes over time. These types of tasks are related to the administrative, clerical and production occupations associated with the middle class in developed countries. Abstract tasks, on the other hand, are activities that require problem-solving, intuition and creativity, and are associated with workers that have high levels of education and analytical ability. These tasks are information intensive – this being the case workers are more productive when the price of accessing, organizing and storing information decreases, with the result that technological change is complementary to these abstract oriented workers. Finally, manual non-routine tasks are those activities that require situational adaptability, personal skills and visual and language recognition. While these activities – be they driving a car, pruning a hedge or interacting with a difficult customer – are beyond the ability of all but the most advanced computers, they are easy for a relatively low skilled worker to perform.

ALM use the USA Dictionary of Occupational Titles (DOT) to develop measures of the intensity of routine, manual and abstract tasks associated with an occupation. They then go on to build and test a model that examines how the workplace demand for the labour input of tasks changes in response to the decreasing price of technology. In doing so, they show that industries that were originally intensive in routine tasks made relatively larger investments in capital as the price computer fell, substituting the labour with technology and increasing the demand for abstract skills. Given that high skill, college graduates typically have a comparative advantage in abstract tasks, this process provides a causal explanation for the role of technology in driving the increased demand for high-skill workers that had been observed in the literature.

Goos & Manning (2007) build on ALM's routinization hypotheses by showing a drop in the relative demand for occupations that are intense in routine tasks in the United Kingdom. These occupations sit in the middle of the distribution, and the result is what Goos and Manning (2007) describe as job polarization – a relative increase in high skill abstract oriented occupations that typically sit at the top of the earnings distribution and an increase in occupations focused on non-routine manual tasks typically found near the bottom of the distribution. They link this shift to changes in inequality in the UK between 1976 and 1995, estimate that job polarisation accounted for 54% of the increase in the 90-50 ratio, and 33 % of the 50-10. Autor et al., (2008) observe similar job polarization in the USA, and at the same time increasing wages for occupations at the bottom and top of the distribution relative to the middle. Like Goos and Manning (2007) they ascribe these changes to the routinization hypothesis developed by ALM, arguing that the polarized wage growth was being driven by the polarized demand shifts.

This hypothesis has since been explored and developed in a number of papers. Firpo, Fortin and Lemieux (2013) use a decomposition approach to examine changes in the wage distribution and find that, even after controlling for institutions, discrimination and education levels, there is strong evidence that routine biased technical change has played a role in shaping the USA wage distribution. In particular, over the 1990's, the period that the USA saw polarized wage growth like that observed in Figure 1, Firpo, Fortin and Lemieux (2013) attribute 40% of the increase in the 90-50 gap, and 21% of the decrease in the 50-10 gap to technological change. Acemoglu and Autor (2011) show that workers with a comparative advantage in routine tasks have had decreasing wages over the 1980's, 1990's and 2000's, while workers with a comparative advantage in abstract and manual tasks have experienced increasing wages[2]. Autor and Dorn (2013) show that most employment growth at the bottom of the distribution can be attributed to low-skill service occupations – some examples being food service workers, security guards, custodians and hairdressers. Internationally there has also been support for the routine biased technical change hypothesis. Goos, Manning and Salomons (2014) for example show that 16 European countries are going through employment polarization and present evidence that routine biased technical growth has played a central role in driving the observed employment shifts.

[2] The papers identify offshoring as another factor that could potentially be driving the wage polarization in developed economies. There is little evidence that it plays a role in South Africa however, and will not be included in the analysis.

At the heart of the routine biased technical literature is a focus on the tasks associated with a worker's occupation. Acemoglu and Autor (2011) consolidate much of the task literature and highlight the value of examining the labour market per the tasks that workers perform as opposed to limiting it to their endowments. They present a model in which a worker with a given level of endowments can perform a range of different tasks, and change the tasks that they perform in response to changes in the labour market. Workers will choose to allocate their human capital to a specific task based on their comparative advantage, the prevailing 'prices' of tasks and the wages for different types of skills. Technology can influence the productivity of all workers, certain types of workers (thereby changing their comparative advantage) or by completely displacing the need for local workers through capital substitution or offshoring. Wages of different skill workers are determined by their relative supplies and the types of tasks they are performing.

Modelling and empirically examining at the task (and therefore occupational) level has two major value adds. The first of these is that it allows for demand shocks to have different effects across different types of workers. One obvious benefit of this from the discussion above is that it presents a tool to perform a more nuanced analysis of how technological change has impacted the wage distribution. The approach has also been applied to other related inquiries; in the USA offshoring for example has emerged has having a significant effect on the wage distribution, like technological change certain types of tasks are more susceptible to being offshored than others, with the result that the phenomenon has been extensively examined using the approach. (Oldenski 2014; Acemoglu & Autor 2011). It is to be expected that, as the relatively new literature grows, it will be put to further applications. The other value add is that the literature presents is that it allows for the modelling of how shifts in the occupational structure affect inequality. Thus, for example, Goos et al., (2007) were able to show that growth in inequality in the UK was tied to shifts in the structure of occupations in the country. Both of these aspects are useful in an enquiry into changes in the South African wage distribution.

Section 3: The South African Economy, Labour Market and Tasks

South Africa's economy is in some ways more like developed economies than its middle-income comparators. As has been well documented, the past four decades have seen the economy's primary and manufacturing sectors shrink at the expense of domestically oriented services (e.g., Rodrik, 2008; Fedderke, 2012; Bhorat et al., 2014). The primary sectors, mining and quarrying and agriculture, farming and fishery accounted for 22% of gross value add (GVA) of the economy in 1980, but fell to just 11.5% in 2011. Between 2001 and 2011, our period of focus, mining's GVA fell by 3.5 percentage points, despite the high prevailing commodity prices[3]. Manufacturing tells a similar if less drastic tale. The sector decreased its GVA from 17.9% in 1980 to 15.8% in 2001. Thereafter, despite adequate performance up to 2007, the sector was negatively impacted by the

[3] The full names of the sectors are: agriculture, hunting, forestry and fishing; mining and quarrying; manufacturing; electricity, gas and water supply; construction; wholesale and retail trade; transport, storage and communication; financial intermediation, insurance, real estate and business services; community, social and personal services; and private households, exterritorial organizations, representatives of foreign governments and other activities not adequately defined. For the remainder of the paper, for brevity's sake, we refer to them in shortened form.

recession -its contribution falling to 13.97% in 2009 and only recovering to 14.34% in 2011. The sectors that have seen growth are all domestic-facing services. The finance and business services sector has seen explosive growth, increasing its share from 13.36% in 1980 to 18.2% in 2001, and then to 21.2% in 2011. Transport and trade have also grown, albeit not as dramatically. Transport increased its share from 6.02% to 8.42% between 1980 and 2001, while trade grew from 13.29% to 14.41% over the first period, increasing slightly between 2001 and 2011 to 14.97% of gross value added[4]. The reasons behind this structural shift are complex, and well beyond the scope of the paper. Rather, we are interested in how the changing structure of the economy influenced the demand for different types of workers and thereby influenced wages. If shifts in the occupational structure can largely be attributed to the between sector movements described above, it is less likely that other demand shocks played a role in the labour market. Conversely, if shifts in the occupational structure can be attributed to within-sector shifts, then other demand shocks – technological change being one them - have a role to play in explaining changes in the labour market (Katz & Murphy, 1992).[5]

We follow Acemoglu and Autor (2011) in using a shift-share decomposition to establish the extent to which changes in occupation shares are attributable to between sector and within sector shocks. The decomposition has the form:

$$\Delta E_{jt} = \sum_k \Delta E_{kt}\gamma_{jk} + \sum_j \Delta \gamma_{jkt}E_k, \quad \textbf{(1)}$$

where ΔE_{jt} is the change in occupation j's share of employment over period t. The first term accounts for between sector changes, with ΔE_{kt} equal to the change in the share of employment for sector k over period t, and γ_{jk} equal to the average share of employment in sector k attributable to occupation j. The second term represents within-sector shifts. The first part of the term, $\Delta\gamma_{jkt}$, is equal to the change in occupation j's share of employment of sector k. The second is the average overall share of employment of sector k in the two-time periods.

Occupations used in the shift share are grouped per the tasks with which they are associated. This is achieved by mapping the routine, manual and abstract occupational task measures developed by ALM onto the South African data and grouping occupations by their intensity in the various tasks.[6] A slightly more nuanced view of the relationship between technical change and tasks can be achieved by grouping across five task oriented categories. The categories are as follows: occupations intensive in abstract tasks, occupations intensive in routine-cognitive tasks, occupations intensive in routine-manual tasks, occupations that score very low on routine tasks and occupations dominated by manual tasks[7]. Based on the theory presented in the literature

---

[4] These changes are reported in Table 9 in Appendix C

[5] This section draws loosely on Acemoglu and Autor (2011) and Autor et al.,(2008).

[6] The measures have been used extensively in the literature. Section 5 provides details on how they were merged into the South African data.

[7] The groupings largely follow the single-digit SASCO codes. Abstract occupations consist of legislators, senior officials and managers, professionals, and technicians and associate professionals; routine-cognitive occupations consist of clerks; routine-manual consists of craft and related trades workers and plant and machinery operators and assemblers. The one exception to the routine-manual group is that taxi drivers and light vehicle drivers, which score very low on routine tasks,

review, one would expect to see technical change associated with decreasing relative shares of the two sets of occupations intensive in routine tasks and increases in the relative shares of the group of occupations intensive in abstract tasks. The effect of technical change on non-routine/ manual tasks – elementary workers and services, sales and drivers – would result, based on the theory, in neutral or positive changes on the relative share of workers[8]. A final point is that, when examining theses changes through the lens of the shift-share analysis, any shifts driven by technical change would be associated with within-sector shifts rather than between-sector shifts.

Table 3 shows the results for the shift-share decomposition using single-digit industry codes and the five occupation groupings introduced above. The results align with the technological change hypothesis in that both the routine intensive occupation groups see negative within-sector growth while abstract and non-routine manual occupations see positive within-sector growth. A closer examination of $\Delta\gamma_{jkt}$ shows that the within-sector contraction for clerks was driven by the financial and business sector, while manufacturing, construction and mining drove the within-sector contraction of operators and craftsmen. The between-sector shifts described above also influence occupational shares. Table 3 reflects the share of each of the five occupation groups that each sector accounts for in 2001, and the change over the period. It shows that service occupations, abstract occupations and clerks are relatively concentrated in the domestic facing non-tradable sectors that saw growth over the period, while elementary workers and operators are concentrated in mining, agriculture and manufacturing. Unsurprisingly, given the growth patterns described above, the former group saw positive between-sector driven growth in their employment shares, while the latter saw negative.

Table 2: Between and Within Sector Occupation Shifts

Occupation	**2001 Share**	**2011 Share**	**Overall Change (ΔE_{jt})**	**Within ($\sum_j \Delta \gamma_{jkt} E_k$)**	**Between ($\sum_k \Delta E_{kt} \gamma_{jk}$)**
Clerks (Routine-Cognitive)	11.61	12.52	0.90	-0.57	1.47
Machinery Operators and Crafts (Routine-Manual)	22.78	19.25	-3.52	-2.22	-1.31
Services, Sales and Drivers (Non-Routine)	14.30	15.20	0.90	-0.39	1.29
Legislators, Professionals and Technicians (Abstract)	20.10	23.48	3.38	0.86	2.52
Elementary Workers (Manual)	31.22	29.56	-1.66	2.32	-3.98

are merged with services and sales workers to form a non-routine/ flexible group (ideally this group would be characterised by scoring highly on interactiveness and flexibility); elementary occupations are made up of skilled agricultural and fishery and elementary workers.

[8] Where technical change is not posited to directly complement or substitute for non-routine manual workers, assuming that routine tasks complement non-routine tasks, the effect of cheaper routine task inputs as a result of technical change could result in a greater demand for non-routine manual tasks.

Values represent percentage points.

The within-sector shifts point to dynamics other than the industrial structure of the economy driving labour demand in the country. The nature of the shifts suggests several distinct processes. On the one hand, the within-sector contraction of routine intense occupations points to capital-labour substitution, driven by decreases in the relative cost of capital as technology improves. Alternatively, high skilled workers made more productive by advancing technology could be subsuming tasks formerly produced by routine task intensive occupations. The within sector expansion of demand for abstract oriented occupations, on the other hand, aligns with the classic conception of skills-biased technical change, where technology improvements increase the productivity of occupations that are intense in information, thereby increasing the demand for them.

Table 3: Sector share of occupation employment

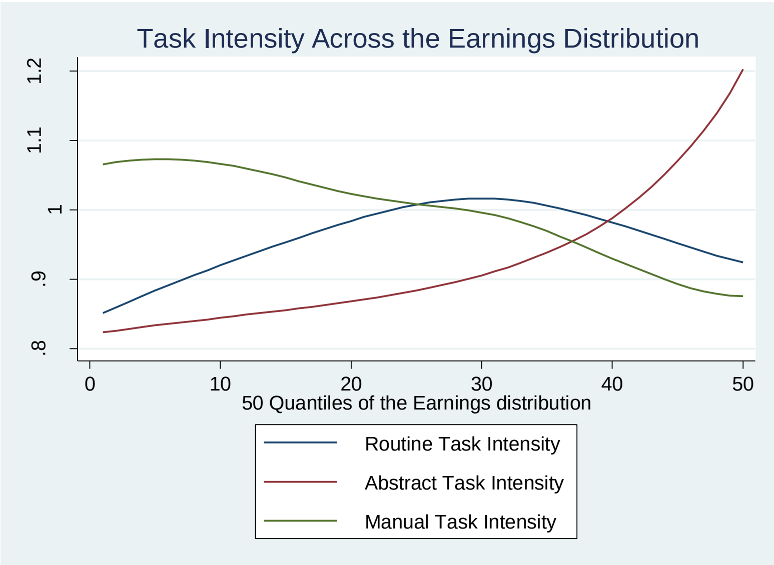
	Clerks		**Machinery Operators and Crafts**		**Services, Sales and Drivers**		**Legislators and Professionals**		**Elementary Workers**	
	2001	Change	2001	Change	2001	Change	2001	Change	2001	Change
Agriculture	1.09	-0.81	2.36	-1.28	7.51	-4.51	0.65	0.15	8.99	-3.79
Mining	2.65	-1.49	15.76	-7.79	8.65	-6.69	1.96	-0.28	6.17	-3.11
Manufacturing	13.51	-3.94	35.99	-1.29	4.80	-1.20	12.66	-2.53	15.94	-2.22
Utilities	1.22	-0.65	2.23	-0.60	0.39	-0.20	1.09	-0.24	1.06	-0.36
Construction	1.20	0.92	16.69	4.31	0.57	0.37	1.03	1.96	5.47	1.61
Trade	24.62	4.79	12.30	1.92	36.76	0.54	9.76	1.92	15.52	2.39
Transport	9.26	-0.70	10.00	2.24	2.18	-0.52	4.86	0.19	5.31	0.41
Finance	21.89	-2.12	1.28	1.66	14.78	6.98	17.01	2.93	9.48	3.77
Services	24.55	4.00	3.40	0.81	23.78	5.03	50.98	-4.09	20.13	3.06
Domestic	0.00	0.00	0.00	0.02	0.57	0.21	0.01	-0.01	11.94	-1.78

Change represents change in sector share from 2001 to 2011. Values represent percentage points.

The link between the shifts described above and the polarized wage growth is observed in Figure 3. It shows that South Africa mirrors the international literature in that routine task intense occupations are concentrated in the middle of the earnings distribution while abstract tasks are concentrated at the top and manual at the bottom.[9] This implies that, if the drop in employment share for routine intense occupations is associated with a decrease in the relative return for these types of jobs, this will contribute to lower wage growth in the middle of the distribution compared to the top and bottom. Similarly, the growth in the share of abstract occupations could have contributed to the wage growth observed in the upper part of the distribution.

[9] The earnings distribution is divided into 50 quantiles and the average task intensity associated with each of those quantiles calculated. These are then plotted against earnings distribution quantiles using a locally weighted smoothing regression bandwidth 0.5.

Figure 3: Task Intensity Across the Earnings Distribution



Refers to 2001 earnings distribution. Author's calculations using PALMS data.

The decreases in the relative share of employment of routine oriented workers presents itself as a likely explanation of the polarized wage growth that South Africa has experienced, but it is by no means the only one. Before continuing to present a wage setting model that incorporates tasks, a review of the South African literature on wage determination, with the aim of highlighting other factors that could have influenced wage growth over the period is presented below.

South African Literature

Returns to Education

The skills biased labour demand described above has been met with a rapid expansion in the supply of labour. Between 1995 and 2008 the working age population grew from 23 million to 29 million people. At the same time, labour market participation increased from 49% to 55% (Leibbrandt et al., 2010)[10]. This has resulted in an extra 5 million people entering the labour market. Most of this is due to an increase in the participation of individuals under the age of 30, and in a drastic rise in the participation of African women in the labour force. This surge in the supply of unskilled labour, along with the demand dynamics described above, played a significant role in the rapid rise of the unemployment rate in the late 1990's and early 2000's (Burger & Woolard 2005).

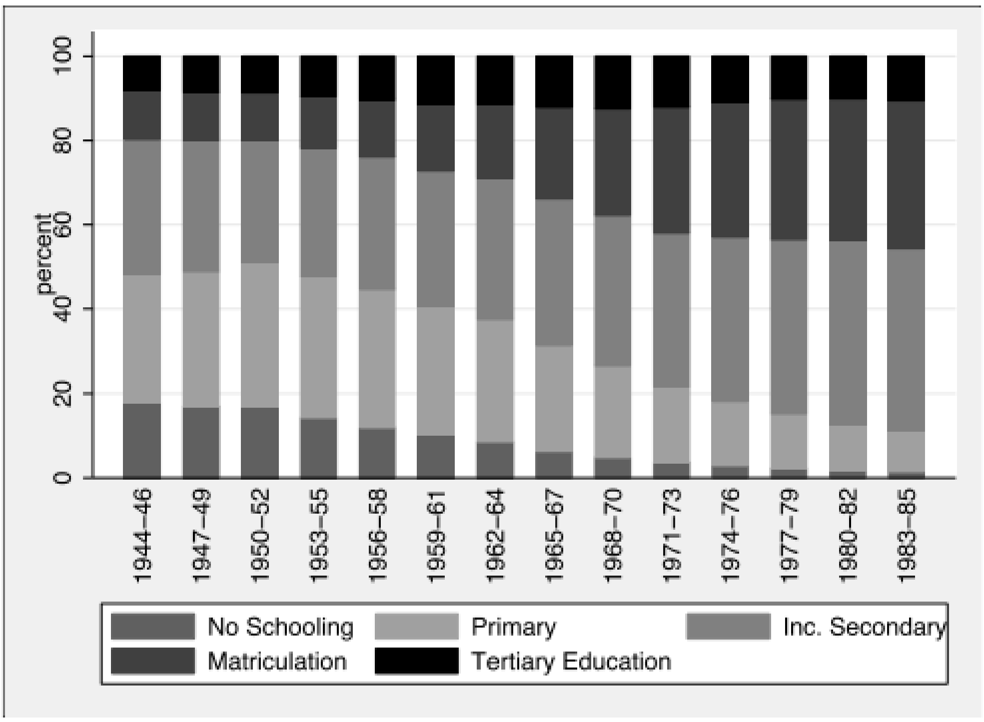
Labour market outcomes are intrinsically linked to the stock of human capital of its participants. Two important elements in this regard in the South African context are the upgrading of the levels of education achieved across cohorts, and the strong variation in the quality of that education. Branson et al., (2013) use labour force surveys,

[10] Participation rates have dropped from a high of 58.4% in 2005, possibly because of the long term nature of unemployment many participants face.

spanning span from 1994 to 2010, to examine the relative levels of education achieved by each cohort. The study tracks and compares outcomes for individuals born between 1944 and 1985. Among Africans, unsurprisingly, there is a strong improvement in educational attainment between generations; only 10% of Africans born 1944 to 1946 had 11 or more year's education. Of those born from 1980 to 1982 however, 50% of African's had reached a similar level. While data did not permit an enquiry into more recent cohorts, it is likely that the trend of higher attainment will continue given the broad improvements that have been achieved in access to education in the democratic era. Figure 4 shows the distribution of educational levels obtained by each cohort. Evidently, the proportion of each cohort entering tertiary education has remained relatively unchanged, while there has been a significant increase in the quantity of matriculants and individuals with some secondary education.

The premiums associated with tertiary education have been well established in the South Africa literature (Keswell & Poswell 2004). Branson et al., (2013) show that this premium gets progressively higher as one moves to more recent cohorts, and is particularly high for the latest. This outcome fits well within a supply demand framework where there has been skill biased labour demand and a stable supply of tertiary graduates. The absolute return to a completed matric declined significantly across newer cohorts, although interestingly, the relative return when compared to grade 10 and 11 completion remained positive and stable across time. Two potential explanations for this development are the increased supply of both matriculants and secondary school drop-outs, and the decreasing quality of secondary education (Branson et al. 2013).

Figure 4:Distribution of educational attainment by birth cohort category, adult males 25-50



Source 3: Branson et al.,(2013)

The rise in returns to tertiary schooling relative to secondary education has garnered significant attention in recent literature. Lam et al., (2012) show that while overall inequality remained relatively stable between 1997 and 2007, this stability concealed two strong counteracting forces. On the one hand, the overall increase in education levels had an inequality-reducing effect while on the other, higher returns to education had an inequality-increasing effect. Leibbrandt & Levinsohn (2011) use semi-parametric decomposition techniques to decompose changes in household income in South Africa. They find that if endowments of household members had been rewarded in the same manner in 2008 as they had been in 1993, there would have been a pervasive increase real household income. In reality, however, there have only been increases at the very top and very bottom of the distribution. While the technique does not allow for the exploration of changes in returns to specific endowments, the authors speculate that the skills twist documented above would explain the dampening of endowment effects.

In Section 1, Table 1 illustrates that earnings have grown faster for workers at the top of the wage distribution than for workers in the middle of the wage distribution. This growth in the p50/p90 ratio resonates strongly with the narrative of skills-biased demand and a stable supply of tertiary graduates. Moreover, it also strongly aligns with task literature discussed above, where high skill workers specializing in abstract tasks are complemented by technology, which concurrently acts as a substitute for middle-skill workers.

This, however, is not the entire story of changes in the wage distribution in South Africa. The evidence in Section 1 also shows that wages have also grown more quickly for individuals at the 10th percentile than the 50th percentile. In the international task literature, this increase is ascribed to the increased demand for low-skill service occupations as businesses benefit from cheaper "routine task input". It is unlikely however that this is the story in South Africa. The bottom of the earnings distribution is dominated by elementary occupations, which saw their share of employment drop over the decade. Moreover, in a relatively non-rigid system, one would expect the massive increase in the supply of low-skilled workers to depress wages at the bottom of the distribution. The sharp rise observed implies that institutional forces have some role in wage setting in the lower part of the distribution. Practically two separate institutions have garnered significant attention in the international and local literature – unions and minimum wages.

Unions

An analysis of union membership in South Africa shows several significant trends between 2001 and 2010. In the private sector, overall union membership has been dropping. In 2001 30.6% of private sector workers were part of a union but by 2010, this number had dropped to 26.3%. This is a result of the shrinking contribution of mining, the most heavily unionized sector, to employment, and declining union membership in agriculture, manufacturing, construction and finance. Conversely, Public sector unionisation climbed from 70.1% in 2001 to 74.6% in 2010. A comparison of these unionisation rates across OECD countries show that South Africa is not unusually highly unionized, average unionisation rates in South Africa in 2005 were 37.5%, while the OECD average was 30%. Brazil in comparison was 71% (Bhorat et al. 2014).

The bargaining system in South Africa occurs at two levels. Bargaining councils can be established by one or more registered unions, and one or more registered employer organisations in a specific sector and area. Issues to be negotiated are left to the discretion of the two parties, and participation is voluntary, with wage formation typically a central area of concern for the respective parties. Agreements made at bargaining council level can extend to non-parties within the relevant sector and area. A second point at which bargaining occurs is at the plant level, by unions representing employees at the specific company. Thus, unionized employee wages could potentially be party to influence from both bargaining councils and plant level bargaining.

There is an extensive literature on the premium that unions command in South Africa's labour market (eg.Azam & Rospabe 2007;Banerjee et al., 2008). Variety in the way authors have treated for union endogeneity, or treated for it at all, as well as the extent to which authors include the influence of bargaining councils in their specifications, has led to a wide range of results. Azam and Rospabe (2007) for example find a premium of 100.46%, Butcher & Rouse (2001) include bargaining councils in their specification and find that union members that fall under a bargaining council earn a premium of around 30%, though this premium declines to around 10% for workers that just fall under a bargaining council agreement. Similarly, Bhorat et al. (2012), who also control for whether an employee falls under a bargaining council agreement and introduce employee benefit variables, find a premium of just 7% for union members outside the bargaining council system, or 22% within it. While the evidence is not conclusive, it appears that the union wage premium is definitely below 25%, and perhaps as low as 7% (Bhorat et al., 2012). While the upper bound is high, it is not excessive in relation to comparator countries (Bhorat et al. 2014).

Minimum Wages

The declining value of the real minimum wage was shown to have significant impacts on the p10/p50 inequality ratio during the 1980's in the USA (DiNardo et al. 1996). In South Africa, as minimum wage laws were introduced in a number of sectors between 2001 and 2011, it is possible that these laws had a similar yet reverse impact. In a series of papers, Bhorat et al., (2013, 2014) estimate the impact of the newly introduced minimum wage laws on earnings, hours worked and employment. The 2012 paper focuses on just the Agricultural sector, while the second paper analyses changes in Retail, Domestic workers, Forestry, Taxi workers, and Private Security sectors. Together these six sectors make up 76% of the sample of minimum wage workers in the September 2007 Labour Force Survey (Bhorat et al.,2013). Of relevance in this instance is the effect of the laws on real earnings of workers in the respective sectors. Bhorat et al.,(2014) investigate this relationship in the Agriculture sector using difference-in-differences estimation and find that real earnings rose significantly in the post-law period. In the 2013 paper, the authors use the same technique to examine the non-agricultural sectors. They find that in four of the five sectors, the introduction of the laws were associated with positive increases in real hourly earnings in the post-law period.[11] One potential limitation in both studies is the LFS dataset's failure to account for non-wage remuneration, such as housing, food etc. Thus it is possible that the effects identified are reflecting employers 'trading' non-wage remuneration for wage remuneration in response to the minimum wage laws.

[11] The Retail, Domestic, Taxi and Security sectors

The literature review highlights a set of factors that have been shown to play important roles in determining wages in South Africa. On the one hand, we can expect changing returns to education, driven by supply and demand dynamics, to have a role in driving upper distribution inequality. Minimum wages and unions on the other hand are more likely to have a role to play in the bottom half of the distribution. Similarly, on the evidence from the task literature review task analysis above, one might expect the routine task variables to have an impact in the middle of the distribution. Section 4 below presents a framework that allows for the estimation and comparison of the effects that each of these aspects have had in driving changes in real earnings.

Section 4: Methodology

In presenting a unified analysis of these forces, the paper follows closely Firpo et al.,(2013). The authors develop a wage setting model that links task prices to the skill returns and at the same time incorporates institutional and discriminatory factors. They use it to justify a set of covariates in an exhaustive decomposition of changes in the wage structure in the USA from the 1970's to 2000's. The methodology has four distinct subsections. The first introduces the wage setting model. The second explains the decomposition framework used to explore the model's implications. The third subsection describes the covariates included in the analysis. Finally, the fourth subsection briefly explains the multiple imputation and bootstrap approach used for inference.

Wage Setting Model and Tasks

FFL posit that an occupation j involves producing an occupation-specific task Y_j, which is, in turn, an input in the firm's production function. The occupation-specific task, Y_j, consists of a range of different activities, with the mix and intensity of those activities depending on the type of occupation in question. Workers are characterized by a k-dimension set of skills $S_i = [S_{i1}, S_{i2}, S_{i3},\ldots S_{iK}]$, where some of these skills are observed – education and experience for example – and others, like people skills, are not. Production of Y_j produced by individual i is assumed to linearly depend on skill such that:

$$Y_{ij} = \sum_{k=1}^{K} \alpha_{jk} S_{ik}, \quad \mathbf{(2)}$$

where α_{jk} represents the productivity of skill k in occupation j. Firms produce goods and services through the combination of the occupation-specific tasks per some firm-specific production function. In a competitive wage setting environment, workers are paid according to the value of the tasks (inputs) they produce. Thus, if the market price of task Y_j at time t equals p_{jt}, an individual i producing Y_j would be paid $p_{jt}Y_{ij}$. Including time specific shocks and occupation-specific compensating differentials, δ_t and c_j, respectively, as well as a vector of institutional and discrimination factors Z_{it} gives the wage setting equation:

$$w_{ijt} = \delta_t + c_j + Z_{it}\varphi_t + p_{jt}Y_{ij} \equiv \delta_t + c_j + Z_{it}\varphi_t + p_{jt}\sum_{k=1}^{K} \alpha_{jk} S_{ik}. \quad \mathbf{(3)}$$

In this framework demand shocks influence wages purely through changes in the market price of tasks, p_{jt}. It is important to note here that the productivity schedules α_{jk} are fixed over time. This implies that technological change does not influence the productivity of a skill in performing a given task. While this may be true in some instances, it certainly doesn't hold across the board. It is likely that, especially for higher skilled workers, advances in technology allow them to take on a greater number of tasks. This being the case, changes in wages could then be attributed to an increase in task prices p_{jt} when in fact they are the result of greater productivity of a given skill allowing a worker to produce more Y_{ij}. In equation (3) p_{jt} and α_{jk} enter multiplicatively making them impossible to untangle empirically, and we follow FFL by treating this product as an occupation-specific return to skill. Thus, while the remainder of the paper will talk about the price of tasks, the effect could also be related to shifting productivity schedules associated with technological change.

Linking the occupational changes described in section 3 to changes in the returns to tasks.

A drop in the demand for the labour provision of an occupation-specific task input Y_j, driven either by a firm's decision to substitute labour for capital or as it is subsumed by a more productive type of worker, would lead to a decrease in the price of that task p_{jt}. FFL include the effect of changing task prices through the linear specification:

$$p_{jt} = \pi_{0t} + \sum_{h=1}^{H} \pi_{ht} T_{jh} + \mu_{jt}. \quad \textbf{(4)}$$

In this specification the T_{jh} is a task content measure, and changes in the price of the occupation-specific input task Y_j are a result of changes in π_{0t} and π_{ht}[12]. Changes in π_{ht} reflect changes in the returns to the subset of tasks (that have been measured) that constitute Y_j, while π_{0t} captures changes in skill prices that are common to all occupations[13]. Substituting (3) into (2) returns:

$$w_{ijt} = \delta_t + c_j + Z_{it}\varphi_t + \left[\pi_{0t} + \sum_{h=1}^{3} \pi_{ht} T_{jh} + \mu_{jt}\right] \sum_{k=1}^{K} \alpha_{jk} S_{ik}. \quad \textbf{(5)}$$

Equation 4 represents a complex relationship between task prices and skills. A valuable insight however that FFL offer is that changes in π_{ht} would impact both between-occupation inequality and within-occupation inequality. If for example, the average level of skills differed across occupations (think for example different levels of education across different jobs) then a change in task prices would influence the dispersion of wages between them. Similarly, because some of the skills used to produce the task are unobserved, changes in the price of that task would impact within-occupation wage inequality even though education and experience are controlled for. FFL (2009), in a separate paper, develop a framework that can be used to decompose wage changes at each quantile of the distribution into components linked to changes in the returns to covariates and changes in the distributions of those covariates. FFL use the decomposition to empirically explore the implications of the wage setting equation described above. Of interest are the distribution wide effects of changes in the price of tasks (π_{ht}), and of changes in the overall allocation of occupations in the economy, compared to the price and quantity

---

[12] In a departure from FFL I use a single task content measure for reasons discussed in section 4.3.

[13] Differing returns to education due to the changes in the relative supplies of workers for example is captured by π_{0t}

effects associated with other common explanations in the literature: changes in education levels, institutional forces and discrimination[14].

Before proceeding with an introduction of the decomposition framework, there are several limitations to the FFL specification that should be noted. Acemoglu and Autor (2011) develop and solve a model that links tasks to skills, and where workers are endogenously allocated to occupations, but in doing so make some restrictive assumptions. FFL, in a different approach, chose the more flexible specification described above. In so doing however, they make it difficult to derive and solve a model that fully describes how workers choose occupations and how supply and demand affect wages in general equilibrium, and therefore choose to work in a partial equilibrium framework. This has important implications. General equilibrium effects that would arise from workers leaving occupations that have been negatively affected by the shock, thereby increasing the labour supply to other occupations are not considered. In theory, this would mean a depression of wages in the occupations receiving the workers, with the result that an empirical analysis using the FFL model would then understate the effects of occupation-specific demand shocks to changes in the wage distribution. A similar, but reverse situation would occur when an occupation benefits from a demand shock. A second related concern is that, in an empirical examination, there is significant room for self-selection bias. If, for example, a subset of workers select into an occupation based on some unobserved ability in response to changes in task prices, and this affected the overall composition of skills in that occupation, an estimate of the effect of task prices on the wage of that occupation would be biased. FFL use a paper by Cortes (2012) that analyses changes in task prices using panel data to argue that, in a similar manner to the general equilibrium effects, the selection effects understate the impact of occupation-specific demand shocks on the wage distribution. This serial understating means, they argue, that if the task content of occupations helps account for some of the change in the wage distribution in the decomposition exercise, one can be sure that the occupation specific demand shocks have a role to play in shaping the wage distribution. The data, unfortunately, are not available to undertake a similar comparative analysis in the South African case.

Decomposition Framework

The decomposition used by FFL combines two separate approaches. The first of these is in effect a generalized version of the popular Oaxaca-Blinder (OB) decomposition (Oaxaca 1973; Blinder 1973). The second is the semiparametric reweighting technique first introduced by Dinardo et al.,(1996). The section will introduce the original OB framework and then briefly show how it is generalized to quantiles. From there it will explain the value of incorporating the semi-parametric reweighting technique into the approach and show how it is achieved.

---

[14] In Appendix B of the 2013 paper, FFL formally develop a connection between changes in task prices π_{ht}, and the wage structure effect – elaborated on below - associated with T_j.

Oaxaca-Blinder Decomposition

The OB framework is used to decompose differences in the mean between two mutually exclusive groups into two distinct effects. The first of these is attributable to differences in the distributions of the observed characteristics of each group, and the second attributable to differences in the relationship between the variable of comparison and the features of that group[15]. For the remainder of the paper these two effects will be referred to as the composition and wage structure effects respectively. An example commonly used is a decomposition of the average wage difference between men and women. Part of that difference is due to differences in the mean levels of education and experience between the two groups, the composition effect, and part of the difference is because men and women are rewarded differently for equal levels of education and experience – the wage structure effect. It is possible to show this algebraically. Let an outcome variable Y be linearly related to the covariates, X the error term *v* such that:

$$Y_g = X\beta_g + v_g \quad g = A, B. \quad (\mathbf{6})$$

Where A and B are two mutually exclusive groups, X is the vector of covariates ($X_i = [X_{i1}, \ldots X_{iK}]$) and β_g is a vector of returns to those covariates. It is assumed that $E(V_g|X) = E(V_g) = 0$. Letting D=1 be an indicator of group B membership and taking expectations over X, the overall difference in average outcomes between group B and A can be written as:

$$\Delta_O^{\mu} = E[Y_B|D = 1] - E[Y_A|D = 0]. \quad (\mathbf{7})$$

Some algebraic manipulation, and the addition and subtraction of the counterfactual wage that group B workers would have earned under the wage structure of group A, $E[X|D=1]\beta_A$, allow the expression to be written as:

$$\Delta_O^{\mu} = E[X|D = 1](\beta_B - \beta_A) + (E[X|D = 1] - E[X|D = 0])\beta_A. \quad (\mathbf{8})$$

The first expression in this equation reflects the wage structure effect - the difference between the returns to covariates for group A and group B. The second reflects the composition effect – the difference in the distribution of covariates for group A and group B. Equation 3 can be estimated using a regression framework as follows:

$$\hat{\Delta}_O^{\mu} = (\hat{\beta}_{B0} - \hat{\beta}_{A0}) + \sum_{k=1}^{K} \bar{X}_{Bk} (\hat{\beta}_{Bk} - \hat{\beta}_{Ak}) + \sum_{k=1}^{K} (\bar{X}_{Bk} - \bar{X}_{Ak})\hat{\beta}_{Ak} \quad k = 1, \ldots, K \ , \ (\mathbf{9})$$

where the Beta-hats are estimated intercept and slope coefficients of regression models for groups A and B. Importantly, the additive linearity of the approach means that it is easy to identify the impact of each covariate X_k on the composition and wage structure effects.

Aside from the additive linearity and independence of error assumptions implied above there are two critical assumptions associated with the OB decomposition that carry through to the RIF-decomposition framework.

[15] These effects are referred to differently throughout the literature.

Simple counterfactual treatment

This assumption holds that a counterfactual wage distribution for workers in 2001 can be developed by plugging the covariate distribution of 2001 workers into the wage structure of workers in 2011 and vice versa. This assumption means that we are working in a partial equilibrium framework – we are ignoring the potential effects that changing the distribution of covariates would have on the returns associated with those covariates. This is one of the weakest points in any OB type decomposition (Autor, 2012).

Overlapping support

This assumption holds that the inputs in the wage setting functions in each group are consistent – an example would be that there are no inputs in 2011 that play a significant role in determining wages that played no role in 2001.

A final issue that carries through to RIF decompositions is associated with the use of base groups and categorical variables. When using a categorical variable it is normal practice to use a dummy variable for each category, and to omit one category as the base group. This approach, however, poses a difficulty in attempting to identify the detailed decomposition effects for the wage structure. It is impossible to find how much of the difference in the intercepts between group A and group B ($\hat{\beta}_{B0} - \hat{\beta}_{A0}$) is due to differences in the base category, and how much of it is due to true 'unexplained' difference between the groups that would otherwise have been captured in the intercepts. The size of the effect 'hidden' in the intercepts will differ depending on the base group chosen, thereby affecting the overall contribution of the categorical variable to the wage structure. There is no uniformly accepted approach to dealing with the problem, and we follow FFL in choosing categories that are comparable to the international literature.

Quantile Decomposition

An important use of OB type decompositions has been to explore the reasons for changing inequality over time, with the two groups in the decomposition consisting of the two years being compared. In this instance, a decomposition of the mean provides little insight - distributional statistics such as variance, Gini coefficients and quantiles are more relevant. It is tempting to decompose quantiles in a similar manner to the mean using conditional quantile regressions and the OB methodology described above. Unfortunately, this is not an option; referring to equations 6 and 7, when applying the OB decomposition we are interested in the difference between the unconditional means of Y_A and Y_B. An essential attribute of the OLS regression is that it can be used to assess the impact of a change in the mean value of X on the *unconditional* mean of Y. This is easily shown; taking expectations over X of $E(Y_i|X_i)=X_i'\beta$ gives $E_x[E(Y_i|X_i)] = E_x(X_i'\beta)$, through the law of iterated expectations this simplifies to $E(Y)=E(X_i')\beta$. As is the case at the mean, to apply the OB decomposition at the quantile level, the effect of β_τ on the unconditional quantile Q_τ needs to be recovered. While the conditional quantile regression is a commonly used tool in labour economics, the law of iterated expectations does not apply to it. This means that, while $Q_\tau(Y_i|X_i)=X_i'\beta_\tau$ can be interpreted as returning the effect of X on the τ^{th} conditional quantile of Y given X, one cannot take expectations $E_x[Q_\tau(Y_i|X_i)] = E(X_i)\beta_\tau$ to return the effects of β_τ on the unconditional quantile Q_τ needed to perform the OB decomposition.

Firpo et al.,(2009) propose a quantile regression technique that gets around this problem. The approach makes use of a common statistical tool, the influence function, which measures the influence of a single observation on a distributional statistic. In the case of a quantile, an influence function (IF) is equal to:

$$IF(Y, Q_\tau) = \frac{\tau - \mathbf{1}\{Y \leq Q_\tau\}}{f_y(Q_\tau)}. \quad \mathbf{(10)}$$

Where Q_τ is the population τ-quantile of the unconditional distribution of Y, $\mathbf{1}\{y \leq Q_\tau\}$ is a binary indicator function as to whether the outcome variable Y is less than or equal to the quantile Q_τ and $f_y(Q_\tau)$ is the density of the marginal distribution of Y. Adding back the original statistic to the IF gives the recentred influence function (RIF):

$$RIF(Y; Q_\tau) = Q_\tau + \frac{\tau - \mathbf{1}\{Y \leq Q_\tau\}}{f_y(Q_\tau)}. \quad \mathbf{(11)}$$

Firpo et al.,(2009,2013) show that by treating the $RIF(y; Q_\tau)$ as an outcome variable in a linear regression on X, and because the law of iterated expectations applies to RIF's, it is possible to estimate the unconditional partial effect of X on Q_τ. This in turn allows for an OB type decomposition to be undertaken across the wage distribution. Equation 11 is analogous to equation 7, with $\Delta_O^{Q_\tau}$ referring to the overall difference between group A and B at Q_τ, and the coefficients $\gamma_A^{Q_\tau}$ and $\gamma_B^{Q_\tau}$ coming from RIF-regressions at Q_τ for groups A and B respectively.

$$\Delta_O^{Q_\tau} = E[X|D=1]^T\left(\gamma_B^{Q_\tau} - \gamma_A^{Q_\tau}\right) + \left(E[X|D=1] - E[X|D=0]\right)^T \gamma_A^{Q_\tau} \quad \mathbf{(12)}$$

The first expression in the equation represents the wage structure effect and the second represents the composition effect. As is the case with the OB mean decomposition, the linear specification of the approach allows for the easy recovery of the individual contribution of each covariate. By letting group A be workers in 2001 and group B be workers in 2011, this approach gives a computationally simple way to decompose changes in the wages throughout the distribution over time. While it would be possible to run the decomposition as currently specified, there is potential for bias. FFL point out that the linear specification used in the RIF-regression is local, and will not hold for larger changes in the covariates. The result is that the decomposition terms could be biased, as changes in γ^{Q_τ} over time reflect changes in the composition of covariates rather than changes in the wage structure. They propose a solution that, under the ignorability assumption, guarantees each term will only reflect differences in the wage or composition effect.

The idea is first to use reweighting to create the counterfactual distribution AC, where workers in group A are given characteristics of group B, but maintain their wage structure. Thereafter, OB decompositions will be applied using this counterfactual distribution. A decomposition of differences between groups AC and B returns 'pure' wage structure effects, while, similarly, a decomposition between AC and A returns 'pure' composition effects. FFL use the same reweighting function developed by DiNardo et al., (1996)[16]:

$$\psi(X) = \frac{\Pr(D=1|X)\,/P(D=1)}{\Pr(D=0|X)\,/\Pr(D=0)}. \quad (\mathbf{13})$$

The counterfactual mean is equal to $\overline{X_{AC}} = \sum_{i \in A} \psi(X_i).X_i \ \rightarrow \overline{X_B}$ and a regression of $\widehat{RIF}(Y_A, Q_\tau)$ on the reweighted sample returns the coefficient $\gamma_{AC}^{Q_\tau}$. If the linear specification used in the RIF-regression accurately reflects the relationship between Q_τ and X, then $plim(\hat{\gamma}_{AC}^{Q_\tau}) = plim(\hat{\gamma}_{A}^{Q_\tau}) = \gamma_{A}^{Q_\tau}$. Taking this into account, an RIF OB decomposition between group A and the counterfactual group AC would return:

$$\widehat{\Delta^{Q_\tau} =} \ (\overline{X_{AC}} - \overline{X_A})\,\hat{\gamma}_A^{Q_\tau} + \overline{X_{AC}}(\hat{\gamma}_{AC}^{Q_\tau} - \hat{\gamma}_A^{Q_\tau}). \quad (\mathbf{14})$$

The first term in equation 13 reflects the pure composition effect. The second term measures the specification error. The specification error represents the difference between the composition effect measured using the reweighting, and the composition effect estimated by the RIF-framework. A significant specification error would reflect the fact that the linear specification is not accurately capturing the true relationship between Q_τ and X. A second, similar decomposition between groups B and AC returns:

$$\Delta^{Q_\tau} = \ \overline{X_B}(\hat{\gamma}_B^{Q_\tau} - \hat{\gamma}_{AC}^{Q_\tau}) + (\overline{X_B} - \overline{X_{AC}})\,\hat{\gamma}_{AC}^{Q_\tau}. \quad (\mathbf{15})$$

In this case, the first term of equation 15 reflects the wage structure effect. The second term indicates how accurately the reweighting function was estimated – if it was done so perfectly the second term would be zero.

Specification

FFL use two task measure aimed at capturing the effects of technological change on the wage distribution, the first of these captures the information content of jobs while the second reflects the overall automization of jobs.[17] In a departure from FFL we use task measures based on those developed by ALM. Routine tasks, or the ability of tasks to be automated, are captured by the Routine Task Intensity (RTI) measure developed by Autor & Dorn (2013). The measure is, in effect, the ratio of routine tasks to abstract and manual tasks in an occupation as measured by the ALM measures. The information orientation of tasks is captured by the abstract task measure developed by ALM. The reason for this is, simply, availability of data. While South African data are not rich enough to create adequate task measures, it is a relatively simple endeavour to map the publicly available Autor and Dorn (2013) measures to the South African data. Further details on the measures are available in Section 5. We also include interactions of RTI with manufacturing, construction and finance. The choice of these sectors relates to the shift-share analysis in section three. Manufacturing and construction saw decreases in the share

[16] In practice a probit regression on the pooled data, with a binary indicating whether an observation is in group A or B on the LHS, and with a rich specification of interactions on the RHS is used to estimate Pr(D=d|X).

[17] FFL also include three measures aimed at capturing the effects offshoring, they have little relevance in the South African context however.

of operators and craftsmen in employment in the sector, while finance and business services saw a decline in clerks. [18]

Turning to other covariates included in the decomposition.

As is the case in FFL, education and potential experience are included by apportioning individuals into groups depending on their highest level of education achieved or years of potential experience.[19] This approach is particularly useful in the South African context, where the non-linearity of returns to education has been well documented (e.g.,. Lam et al. 2012). Concerning the impact of institutions, we include a union coverage, minimum wage and sector level dummies. The sector level dummies account for (some) bargaining council effects, as well as any changes in sector level wage premia. Marriage, gender and race are also controlled for.

It is important to note here that the independence assumption is almost sure to be violated for at least a few of these covariates, and that RIF-regressions do not, at present, have a simple framework allowing for a correction of this endogeneity. Nevertheless, as was alluded to in the methodology, with a weaker assumption it is still possible to obtain results that are interesting. The ignorability assumption from the evaluation literature holds that covariates and errors do not need to be independent if the conditional distributions between errors and covariates are consistent across groups. In making this assumption we acknowledge that some point estimates of the coefficients $\hat{\gamma}^{Q_\tau}$ may be biased, but, as long as the ignorability assumption holds, can be confident that we are accurately estimating wage structure and composition effects (Firpo et al. 2013).

Inference

There are two important considerations when dealing with inference. The first of these is that standard errors need to reflect the fact that the technique described above involves two stages of estimation, the reweighting logit and RIF-regressions. The easiest way to achieve this is to bootstrap the entire procedure, and we follow FFL in doing so. The second is that we use multiply imputed earnings data included in the PALM's dataset in the analysis (Kerr & Wittenberg, 2013). To accommodate for the variance introduced by using imputed data, we run the decompositions and calculate bootstrap errors (using 100 repetitions) for each of the ten multiply imputed datasets provided, and then combine the point and error estimates according to the rules established by Rubin (1987). These are further described in Appendix A.[20]

[18] Mining also saw a decrease in the relative share of operators. The sector is highly unionised, and has a strong bargaining council presence, and its inclusion as an interaction clouded results.

[19] Age minus years of education.

[20] The analysis and inference makes use of the stata code generously made publically available by Professor Nicole Fortin on her public profile.

Section 5: Data

Wage Data

The analysis uses 2001 to 2011 data from PALMS v2.1. The dataset consists of stacked biannual Labour Force Surveys from 2001 to 2007, and the Quarterly Labour Force Surveys from 2008 to 2011. The surveys have been commonly used in the literature and are considered one of the more reliable sources of South African labour market data available. A feature unique to the PALMS dataset is that it includes a cross entropy derived weight, courtesy of Branson & Wittenberg (2014), which gives consistent trends across time – something that is missing with the survey weights. These cross-entropy weights are used throughout the analysis. A second valuable feature of the dataset is the way it treats income. The dataset includes ten sets of multiply imputed real earnings variables that take into account bracket-responses, missing values and outliers (Kerr & Wittenberg 2013). The analysis makes use of these imputed earnings variables and accommodates them as described in section 4.4. Aside from earnings, variables of interest that the dataset contains are age, years of education, union membership status, occupation, industry, hours worked last week, marriage and race.

The sample is limited to wage earners between the ages of 15 and 65 out of the military. The decision to focus on the employed is grounded on the observation by Wittenberg (2014) that systematic shifts in the distribution of income have occurred only for employees, not the self-employed. Wage earners make up 86% of labour market participants in 2011 and account for 80% of labour earnings. Also, following Wittenberg (2014), we drop those observations with a reported value of zero for the earnings variable. Including zero-incomes make an appreciable difference to the income distribution. However, the considerable variance in the number of these types of earners suggest that they are providing inadequate data. Hourly wages are calculated by dividing the "monthly real earnings" variable by 4, and then by the "number of hours worked last week" variable. In the decomposition analysis, following the literature, the "hourly wage" variable is logged. Potential experience is simply the age of the respondent minus the number of years completed education. In the interests of a larger sample, the analysis for 2001 is undertaken on a pooled sample of both LFS surveys for the year. Similarly, analysis for 2011 is undertaken on a pooled sample of the last two waves of the QLFS surveys.[21]

Other than the task, hourly earnings and potential experience variables, the minimum-wage dummy is the only variable manually constructed. The purpose of the minimum-wage variable is to identify whether the sector-specific minimum-wage laws introduced over the period had positive effects on the real earnings of the occupations to which they applied. Following Bhorat et al., (2013) the workers to whom the wage laws apply in the 2011 sample, and will apply to in the 2001 sample, are identified using an overlap of the South African Standard Classification of Occupations (SASCO) and Standard Industrial Classification (SIC) codes available in PALMS. From there, in an acknowledgement of the bluntness of SASCO, and following FFL's specification of the minimum wage, we further limit the dummy to those workers that have a nominal income of less than or equal

[21] The number of observations in the 2001 and 2011 pooled samples are 42952 and 34834 respectively.

to the highest prescribed minimum wage for their sector in the associated period.[22] Because all the minimum wage laws were enacted between 2001 and 2011, in the case of the first period this meant deflating each of the occupation-specific minimum wages from their respective inception dates to 2001 prices. The aim of this specification is to identify those workers that the minimum wage laws were enacted to support – those earning less than or equal to the prescribed minimum wage in a designated sector/ occupation – to examine whether the group is associated with an increase in real earnings over time.

Measuring Tasks

Following the precedent set by Goos et al., (2014) we map the publicly available task measurement variables developed by Autor & Dorn (2013) onto the South African data[23]. The variables assign scores to each occupation for the intensity of routine, manual and abstract tasks associated with it. Therefore, bringing the measures to the South African data required a link between USA census data occupations and SASCO. This link was developed by first mapping from USA Census data to ISCO-88, the International Standard Classification of Occupation's, via publicly available crosswalks.[24] SASCO is closely related to ISCO-88. Differences come, primarily, from the inclusion in SASCO of catchall "not elsewhere counted" occupations and South African specific occupations. We amalgamate these codes with similar occupations that map to ISCO-88 and merge the task measures to the 2001 distribution of occupations.[25] A final step is to aggregate the task measure up to the three-digit level, weighted by the 2001 distribution of occupations. This step is to limit the measurement error introduced by the mapping across different occupation classification systems. It is worth noting that a strong assumption in the mapping is that equivalent occupations in the USA and South African data perform the same tasks. This in turn implies that that technology use in production is equivalent across countries. While this may be the case in certain sectors it will also likely fall down in certain areas.

Autor and Dorn (2013) closely follow ALM in using DOT to develop their task measures. DOT evaluates US occupations on 44 dimensions, some examples being training times, physical demands and required worker aptitudes, temperaments and interests. A full discussion of the task measure construction, as well as their strengths and weaknesses, is available in ALM (2003). Saliently, however, each of the task measurements is developed using a combination of DOT measurements. In the case of abstract tasks these variables measure the direction, control and planning, and quantitative reasoning requirements associated with an occupation. Routine tasks use variables measuring the finger dexterity and the extent to which an occupation is related to

[22] The minimum wages for each sector and occupation are available at: http://www.labour.gov.za/DOL/legislation/sectoral-determinations/sectoral-determination.
[23] The measures are publicly available at http://www.ddorn.net/data.htm.
[24] Crosswalks are available from the National Crosswalk Service Center, a USA agency that provides technical support and information on occupations in the country. See http://www.xwalkcenter.org/index.php/downloads under xwalks.
[25] Overall 28 occupations were amalgamated, 23 of these occupations accounted for 0.05% or less of the combined 2001 and 2011 sample. Of the remaining occupations, teaching associate professionals "not elsewhere counted" were merged with primary education teaching associate professionals, spaza shop owners were merged with kiosk salespersons, shebeen operators were merged with bar tenders, minibus taxi drivers were merged with taxi cab drivers and landlords were merged with hotel stewards.

set limits, tolerances or standards. Finally, the manual measurement is based on a variable measuring eye-hand-foot coordination.

The RTI measure used by Autor and Dorn (2013) is equal to the log of routine task variable minus the sum of the logs of abstract and manual task variables. The idea is that those occupations that score highly on routine tasks relative to the other types of tasks are more likely to be substituted for by technology. The reason for using the ratio rather than just the raw measure is a weakness in the DOT variables. Several high skill science oriented occupations for example score very highly on routine tasks in the measures. The RTI measure, by taking the ratio of tasks, overcomes this problem. We follow Goos et al.,(2014) in standardizing the RTI further. It is important to note here that the results are sensitive to the choice of standardization. Equation 8 shows that the size of X-bar influences the size of the wage structure effect. Similarly, the magnitude of the beta-hats is affected by th standard deviation. Rather than make an arbitrary choice on these factors, as one would do normalizing the data, we standardize the data to lie between one and zero. The rationale behind this is that it puts the measures on a comparable scale to the other variables in the data, which are all dummies. The Abstract task measure is the DOT measure, logged then standardized. Despite its inclusion in the RTI measure the correlation between the two is low, at just 0.22. The DOT measures do not perfectly capture the tasks associated with an occupation, particularly when one considers that occupational tasks vary both across time and within occupations themselves. This weakness is likely compounded by measurement errors introduced via the mapping from USA to South African data. Despite this, the measures still provide a relatively accurate and intuitive characterization of occupations. Table 12 in Appendix C provides a full list of each occupation and its associated measure.

Descriptive Statistics

Table 9 in Appendix C reports descriptive statistics for the variables of interest. Other than the task measures, the covariates are dummy variables, with the result that the means multiplied by a hundred reflect percentages of the population. The changes in education over the period correspond to the situation described by Branson et al.,(2013). The portion of workers that have some high school, or have completed high school, rose by 2 percentage points and 6.8 percentage points respectively. At the same time, the number of workers with very low levels of education dropped significantly, with the no schooling, some primary and completed primary schooling categories falling by 4, 7.3 and 2 percentage points respectively. The portion of workers with tertiary education over the period remained relatively stable, with an increase of 4 percentage points for individuals with less than three years' tertiary education, a 1.3 percentage point increase in the portion of workers with a bachelor's degree and a 0.1 percentage point decrease in postgraduate degree workers.

The potential experience categories are also fairly stationary over the period, with the most significant movements being increases of 1.7, 0.6 and 0.7 percentage points in the participation of individuals that have between 10 – 15, 15 - 20 and 20 - 25 years' potential experience respectively. This likely reflects the increased numbers of young, unskilled workers in the labour market described by Leibbrandt et al. ,(2010).

Turning to institutions; there are more workers in the minimum wage sectors captured in 2011 than 2001. This is attributable to the rapid growth of service, trade and finance employment, which account for a significant portion of minimum wage workers. In line with the evidence put forward by Bhorat et al.,(2012) of dropping unionisation rates in the private sector, the relative number of union members has decreased by 3.6 percentage points. The discrimination variables provide some interesting insights into the changing nature of the labour market. The base group is non-married, black men. This indicates a dramatic 7.2 percentage point decrease in the share of married labour market participants. Encouragingly, and also in line with Leibbrandt et al.,(2010), there has been an increase in female participation in the labour market. Finally, the sector employment shares reflect the changing nature of the South African labour market. In line with the trends described in Section 3, the low and semi-skilled intensive sectors -mining, agriculture and manufacturing - decrease their shares of employment while growth is seen in services, finance, trade and construction.

Reweighting

The decomposition analysis described in section 4 consists of two steps. The first of these is a reweighting of the data to create a counterfactual distribution where workers in 2001 are given the endowments of workers in 2011 but maintain the same wage structure. The second step uses the counterfactual in conjunction with RIF-regressions to calculate detailed composition and wage structure effects. In practice, the reweighting is achieved by multiplying the survey weights associated with observations in the 2001 sample by equation 13. The calculation of equation 13 requires the estimation of $Pr(D=1)$ and $\Pr(D=1|X)$. The first of these terms is simply the ratio of observations in 2011 and the combined sample of 2001 and 2011. The second term is calculated using a probit regression, with the left hand side featuring a dummy indicating whether an observation is part of the 2001 or 2011 sample, and the right hand side consisting of a rich set of covariates. As well as the education, union, race, gender and marriage variables included in the decomposition, the probit specification includes one-digit occupations and a range of interactions. These consist of interactions of experience and education, union-status and education, union-status and experience, and one-digit occupations and education. The regression results are reported in Appendix C. The purpose of the regression however is the manufacturing of an accurate counterfactual, and the efficacy of the specification doing so can be measured via the reweighting errors introduced in equation 14 and presented in section 6.

Section 6: Analysis

RIF-Regressions

Before analysing the decomposition, some interesting insights are provided via the RIF-regression covariates used in the decompositions. Figure 4 shows a set of graphs comparing the returns to covariates at different points in the earnings distribution in 2001 and 2011. The results are from regressions at every 5th percentile, with a base group of non-married, black men that have 15 to 20 years' experience, some high school and work in the finance sector. This base group is also used in the decomposition.

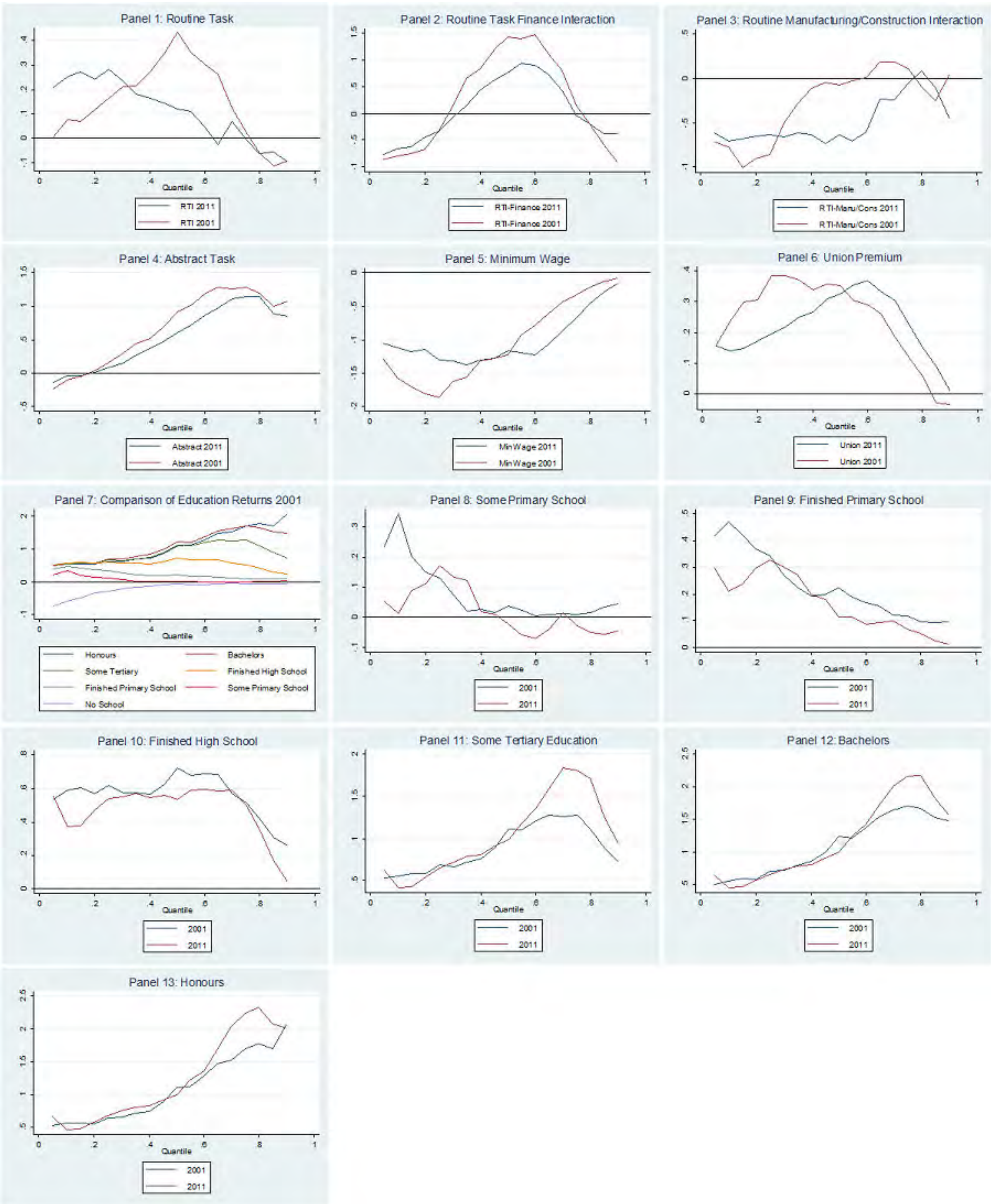
The first three panels relate to technological change. The first panel shows the beta-coefficients associated with RTI in 2001 and 2011. Panels 2 and 3 go on to display the sum of the RTI coefficients and the coefficients of the interaction terms. Panel 1 shows that in 2001 occupations intense in routine tasks had distinctly higher returns at the median relative to the 10th and 90th percentiles. Table 7 shows that the results are significant at the 1%

level at the 50th percentile, and at the 10% level at the 10th percentile. In 2011 the returns to the covariates shift drastically. Returns are high and significant at the 1% level at bottom of the distribution, and are both small and lose significance in the middle of the distribution. This pattern of change is consistent with routine biased technical change – occupations that were formerly associated with high returns in the middle of the distribution go on to command a premium at the bottom as their relative ranking in the earnings distribution decreases. Returns to routine tasks for workers in the finance sector are concave and higher in 2001 than 2011. They are significant at the 1% level at the 10th and 50th percentiles in 2001 and 2011. This pattern of returns is likely driven by office clerks, of which the sector employs a large share. The lower middle of the distribution returns in 2011 are consistent with routine biased technological change and, along with the declining share of clerks employed in the sector, suggest that technology is substituting for workers. The returns to routine tasks for workers in the manufacturing and construction sectors also show evidence of routine-biased technical change. In 2001 the returns are concave but negative though most of the distribution. In 2001 the returns are negative and stable between the 10th and 60th percentiles before climbing slightly at the top of the distribution. The results are significant at the 1% level at the 10th and 50th percentiles in 2011, and at the 10th percentile in 2001. Panel 4 shows returns to abstract tasks. As expected, in both cases returns improve monotonically from the 10th through to the 80th percentiles, before dipping slightly at the top of the distribution. What is somewhat counterintuitive given the theory however, is the lower returns to abstract tasks in 2011 compared to 2001. The result is likely driven by the inclusion of both tertiary education and abstract tasks in the specification.

Panel 5 shows the returns to minimum wage dummy. There is a marked improvement in returns between the 10th and 40th percentiles in 2011 compared to 2001. It is possible that the result is driven, at least in part, by high ability individuals self-selecting into the minimum wage sector. Regardless, the result does provide evidence that, despite the low compliance noted by Bhorat et al. (2010), the minimum wage laws enacted over the period have played a significant role in increasing incomes for individuals at the bottom of the distribution. Panel 6 compares the returns to unions across the distribution in 2001 and 2011 - the results mirror much of the international literature with higher returns at the bottom of the distribution that steadily decrease. This pattern of returns, which is inequality-enhancing for the lower half of the distribution, and inequality-decreasing for the upper half, explain the composition effects associated with de-unionization.

Turning to education. Panels 7-12 compare returns to education in the two periods. The change in tertiary education shows a similar pattern in all three categories; 2011 has lower returns between the 10th and 60th percentiles, it then climbs higher between the 60th and 80th before dropping dramatically. The higher 2011 returns to tertiary education are expected and align with the literature that has argued that returns to post-secondary schooling have played an inequality-enhancing role in the country (e.g., Branson et al., 2013). However, the sharp drop in returns above the 80th percentile is somewhat surprising. This shape of returns is similar to the change in the earnings distribution described in Figure 2, where percentage earnings increases for high-income individuals peak between the 80th and 90th percentiles before showing a slight decrease. One other possible explanation for the result is data. This is the first study to use the QLFS data in comparing the evolution of the returns to education over time. The dip in the change of earnings at the top of the distribution, as well as the dip in returns to education, could both be indicative of undersampling of the top percentiles of the earnings distribution in the QLFS relative to the LFS.

Figure 5: RIF Coefficients at each percentile in the earnings distribution



Decomposition Results

Table 4 below gives the aggregate results for the OB decomposition. The results show similar evidence to that presented by Lam et al.,(2012). Composition effects have been inequality reducing across the board. Between the 90th and 50th percentiles these gains, however, have been overwhelmed by a strong inequality-enhancing wage structure effect. A fresh insight that the quantile decomposition offers is that wage structure effects have also had a strong inequality-reducing effect between the 10th and 50th percentiles. This is well illustrated in the first panel of Figure 3. The overall changes predicted by the model closely follow the growth in earnings described in Figure 1. Wage structure effects are responsible for the U-shaped wage growth observed, while composition effects remain fairly consistent throughout the distribution, peaking at the 20th percentile and slowly decreasing through to the 90th percentile. At all three points wage structure effects dominate composition effects. At the p90/p10, the wage structure effect accounts for about 80% of the change. Likewise, the p90/p50 the wage structure effect accounts for more than a 100% of the change and is negated by the composition effect. Finally, at the p50/p10 ratio, the wage structure effect amounts to close to a 100% of the change. The reweighting function performs well in that the errors are small and insignificant. Similarly, the small and insignificant specification errors mean that the composition effects estimated from the linear specification are very close to those that would have been estimated from a semi-parametric decomposition using just the reweighting function.

Table 4: Aggregate Decomposition Results

	p90/p10	***p90/p50***	***p50/p10***
Overall Change	-25.799***	16.777***	-42.576***
	(2.275)	(2.002)	(1.459)
Composition Effects	-4.462**	-3.298**	-1.164
	(1.525)	(1.192)	(0.908)
Wage Structure Effects	-22.147***	17.241***	-39.387***
	(2.686)	(2.464)	(1.764)
Reweighting Error	0.476	1.506	-1.030
	(1.503)	(1.198)	(0.687)
Specification Error	0.334	1.329	-0.995
	(2.602)	(2.296)	(1.799)

The values represent log wage differentials multiplied by 100. Errors in brackets are calculated from the bootstrapping and imputations described in section 4.4. Asterisks indicate significance at the 10%(*), 5%(**) and 1%(***) levels with 9 degrees' freedom. The formulas for each of the effects are as follows.

Composition Effect: $(\overline{X_{AC}} - \overline{X_A})\,\hat{\gamma}_A^{Q_\tau}$

Wage Structure Effect: $\overline{X_B}(\hat{\gamma}_B^{Q_\tau} - \hat{\gamma}_{AC}^{Q_\tau})$

Reweighting Error: $(\overline{X_B} - \overline{X_{AC}})\,\hat{\gamma}_{AC}^{Q_\tau}$

Specification Error: $\overline{X_{AC}}(\hat{\gamma}_{AC}^{Q_\tau} - \hat{\gamma}_A^{Q_\tau})$.

The detailed wage structure and composition effects allow for a closer look at the factors driving the aggregate effects. Table 5 reports the detailed effects of the minimum wage, routine task, abstract task, education, union and experience - the main factors identified in the literature. These effects for these covariates are also graphed in Figure 5. The sum of the effects of the routine technology variables gives the technology effect. Similarly, the sum of the effects for finished high school and below gives primary school effects while the sum of the three tertiary education categories gives the tertiary effects.

Table 5: Main Effects

	Composition			**Wage Structure**		
	p90/p10	p90/p50	p50/p10	p90/p10	p90/p50	p50/p10
Minimum Wage	-2.425***	-2.165***	-0.261*	-8.086***	-1.665	-6.422**
	(0.452)	(0.420)	(0.123)	*(2.107)*	*(1.494)*	*(2.390)*
Technology	0.178	-1.744***	1.922***	-1.665	16.480**	-18.145***
	(0.341)	(0.420)	(0.346)	*(7.534)*	*(7.169)*	*(5.489)*
Abstract Task	2.582***	1.001***	1.580***	-10.585	4.343	-14.928**
	(0.449)	(0.233)	(0.279)	*(8.587)*	*(8.776)*	*(5.114)*
Primary Education	-4.209***	-3.114***	-1.095*	-11.862*	0.549	-12.411*
	(0.502)	(0.404)	(0.550)	*(5.985)*	*(4.336)*	*(6.264)*
Tertiary Education	6.944***	2.542**	4.402***	3.056	7.004**	-3.948
	(1.365)	(0.906)	(0.660)	*(2.865)*	*(2.706)*	*(2.234)*
Union	1.106***	2.207***	-1.101***	3.623*	1.517	2.107
	(0.218)	(0.320)	(0.188)	*(1.971)*	*(1.837)*	*(1.192)*
Experience	-0.087	-0.101	0.014	1.004	-6.989	7.994*
	(0.157)	(0.141)	(0.086)	*(5.224)*	*(4.752)*	*(4.277)*

The values represent log wage differentials multiplied by 100. Errors in brackets are calculated from the bootstrapping and imputations described in section 4.4. Asterisks indicate significance at the 10%(*), 5%(**) and 1%(***) levels with 9 degrees' freedom

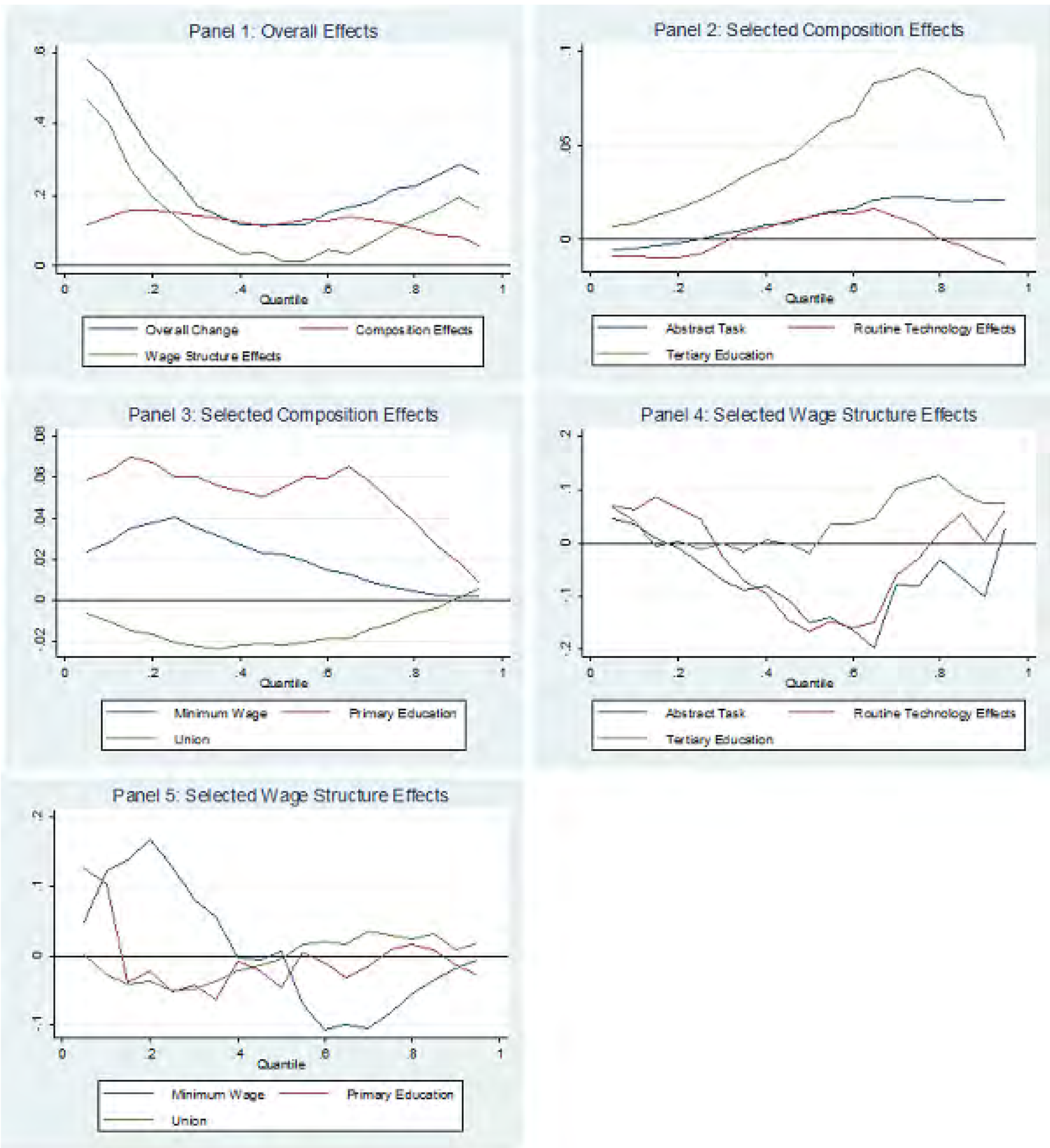
Main Composition Effects

Composition effects are dominated by education. The increased number of tertiary graduates, and the large wage premiums they receive, has driven a strong inequality-enhancing composition effect. Primary education, as expected, has significant inequality-weakening effects. These effects are overwhelmed by tertiary education composition effects at the p90/p10 ratio, but not at the p90/p50. Routine technology effects, because they are the sum of the interactions and routine task variable are difficult to interpret clearly. The slightly stronger effect at the median relative to the top and bottom of the distribution likely reflects the fact that between-sector forces have resulted in a growth in the number of clerks in the economy. Similarly, the inequality enhancing abstract task composition effects reflect the trends described in Table 3, where between and within sector forces are increasing the number of information intense occupations. The union composition effects reflect the de-unionization presented in the descriptive statistics. The effect is inequality enhancing at the top of the distribution, and inequality-lowering at the bottom. This effect is analogous to that described by FFL and aligns with the non-monotonic nature of union effects.

Main Wage Structure Effects

Wage structure effects should be interpreted with caution. As was highlighted above the size of the effects is subject to the choice of base for the categorical variable. Keeping this in mind, several interesting insights arise from the effects. The minimum wage has large and significant inequality reducing effects at the p90/p10 and p50/p10 ratios. Consistent with lower returns in the middle of the distribution, technology has large and significant inequality-enhancing impact at the p90/p50 ratio and an inequality-reducing one at the p50/p10 ratio. Returns to abstract tasks are insignificant except at the p50/p10 ratio, where they show a large inequality-reducing effect. Figure 5 shows that wage structure effects for abstract tasks have a similar u-shape to the routine tasks. Where this type of change was expected for routine tasks, it is not for abstract tasks. As was pointed out, the RIF results in Figure 4 shows that while abstract tasks are monotonically increasing in returns as one moves up the wage distribution, the returns in the middle of the distribution are significantly lower in 2011 than 2001. This result is not easily interpretable, and a deeper exploration of whether it reflects an anomaly in the data or an underlying trend would be an interesting area for future research. Tertiary education wage structure effects support the findings by Lam et al., (2012) of growing returns to tertiary education. They reflect a large and significant inequality-enhancing effect at the p90/p50. In line with the RIF results, unions play a slight inequality-enhancing role at the p90/p10 ratio. The effect, however, is small relative to the minimum wage and tertiary education.

Figure 6: Selected Wage Structure and Composition Effects



Secondary results

Table 6 returns results for the discrimination and sector variables. The base group for race is blacks – thus the significant and inequality decreasing race composition effects reflect the fact that a lower share of whites are participating in the labour market. Interpreting the wage structure effects associated with sectors is tricky, and we do not attempt to do so at length. The sectors that do show significant wage structure effects however all have plausible stories. Agriculture and domestic services both have large inequality reducing effects at the p90/p10 and p50/p10 ratios. Both sectors are intensive in minimum wage workers, and it is likely that the sector effects represent those workers benefitting from minimum wages but excluded by the criteria set in Section 5.1. The p90/p50 wage structure effect associated with mining are likely the result of the mining strikes that occurred over the late 2000's.[26] Community services show large inequality increasing effects at the top of the distribution. The sector serves as a relatively accurate proxy for the public sector, and thus, the effects likely represent increasing public sector wages.

Table 6: Sector and Discrimination Effects

	Composition			***Wage Structure***		
	p90/p10	p90/p50	p50/p10	p90/p10	p90/p50	p50/p10
Marriage	-0.048	0.089	-0.137*	0.016	2.950	-2.934
	(0.042)	(0.052)	(0.072)	(2.458)	(2.298)	(1.893)
Race	-2.066***	-0.825***	-1.241***	-1.711	-2.703	0.993
	(0.435)	(0.212)	(0.260)	(1.741)	(1.788)	(0.613)
Gender	0.061	-0.052	0.112	-3.083	-1.564	-1.519
	(0.088)	(0.072)	(0.073)	(2.379)	(2.237)	(1.707)
Agriculture	-1.964***	0.541**	-2.505***	-4.197**	3.831***	-8.029***
	(0.363)	(0.215)	(0.377)	(1.369)	(1.113)	(0.955)
Mining	0.351	1.259***	-0.908***	1.603	-0.511	2.114***
	(0.289)	(0.325)	(0.217)	(0.961)	(0.888)	(0.493)
Manufacture	0.055	-1.045**	1.101**	4.272	4.315	-0.043
	(0.256)	(0.384)	(0.347)	(4.514)	(4.177)	(2.222)
Utilities	-0.044	0.026	-0.071	0.049	0.038	0.012
	(0.063)	(0.042)	(0.095)	(0.235)	(0.228)	(0.128)
Construction	-0.233*	0.119	-0.352**	0.054	0.776	-0.722
	(0.126)	(0.089)	(0.146)	(1.834)	(1.634)	(1.093)
Trade	-0.305	-2.901***	2.596***	3.231	4.922	-1.690
	(0.552)	(0.621)	(0.444)	(4.507)	(4.034)	(2.440)
Transport	-0.034	0.049	-0.083	1.463	2.627*	-1.164
	(0.109)	(0.151)	(0.249)	(1.423)	(1.305)	(0.865)
Service	0.017	-0.116	0.132	13.206**	16.529**	-3.323
	(0.112)	(0.449)	(0.526)	(5.758)	(5.298)	(3.209)
Domestic	-4.340***	0.931	-5.271***	-5.728**	5.399**	-11.127***

[26] We experimented with a routine task mining interaction early on, which showed that there had been a pervasive rightward shift of the returns to being a routine worker in the mining sector. This is consistent with a rapid rise in wages for machine operators working in mines over the period.

	(0.670)	(0.510)	(0.589)	(2.481)	(2.123)	(1.740)
Constant				-6.808	-40.606	33.798
				(24.975)	*(23.824)*	*(17.293)*

The values represent log wage differentials multiplied by 100. Errors in brackets are calculated from the bootstrapping and imputations described in section 4.4. Asterisks indicate significance at the 10%(*), 5%(**) and 1%(***) levels with 9 degrees' freedom

Section 7: Conclusion

Wages changes in South Africa between 2001 and 2011 were distinctly polarised, with high relative growth at the top and bottom of the distribution. The paper has used a partial equilibrium decomposition framework to explore factors associated with those changes. The decomposition splits changes into two aggregate portions, composition effects, which are related to shifting endowments of workers in the economy, and wage structure effects, which are related to changing prices associated with those endowments. The results suggest that composition effects have had a uniformly positive effect on wage growth, while wage structure effects are responsible for the polarized nature of the growth. Looking more closely at the factors driving the wage structure effect, it appears that technology, minimum wages and shifting returns to tertiary education had important effects at various points in the distribution. Technological change has had non-monotonic effects of the wage distribution. The decomposition shows that decreased returns to routine tasks have had a large and significant effect in contributing to the hollowing out of the distribution. These effects are driven by the finance, manufacturing and construction sectors; all of which, have seen marked decreases in their shares of routine oriented workers. This same technological change has probably, alongside a shift in the economy towards human capital intensive service sectors, contributed towards the inequality enhancing increase in the tertiary education premium observed. Finally, the decomposition confirms that the minimum wage has played an important role in driving relative wage growth at the bottom of the distribution.

While it is important to note that the study does suffer from limitations – the partial equilibrium framework and self-selection issues highlighted in the text are both concerns – there are still several important discussions that it introduces. The evidence presented suggests that South Africa is exposed to the same inequality enhancing trends that have been highlighted in developed countries. Policy then needs to focus on empowering workers to interact in the new labour market.

Bibliography

Acemoglu, D. & Autor, D., 2011. *Skills, Tasks and Technologies: Implications for Employment and Earnings*, Elsevier Inc. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0169721811024105.

Autor, D.H., 2012. Lecture 6: Wage Density Decompositions. *MIT Graduate Labor Economics 14.662 Spring 2012*, p.31. Available at: http://economics.mit.edu/files/7714.

Autor, D.H. & Dorn, D., 2013. The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market. *American Economic Review*, 103(5), pp.1553–1597.

Autor, D.H., Katz, L.F. & Kearney, M.S., 2008. Trends in U.S. Wage Inequality: Revising the Revisionists. *Review of Economics and Statistics*, 90(2), pp.300–323.

Azam, J.-P. & Rospabe, S., 2007. Trade Unions v . Statistical Discrimination : by. *Journal of Development Economics*, 84(1), pp.417–444. Available at: https://core.ac.uk/download/files/153/6376054.pdf.

Banerjee, A. et al., 2008. Why has unemployment risen in the New South Africa? *Economics of Transition*, 16(4), pp.715–740.

Van Der Berg, S., 2010. Current poverty and income distribution in the context of South African history. *Working Paper of the Department of Economics and the Bureau for economic research at the University of Stellenbosch*, pp.1–23. Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1727599.

Bhorat, H., Goga, S. & Van Der Westhuizen, C., 2012. Institutional wage effects: Revisiting union and bargaining council wage premia in South Africa. *South African Journal of Economics*, 80(3), pp.400–414.

Bhorat, H., Kanbur, R. & Mayet, N., 2012. Minimum Wage Enforcement in South Africa. *International Labour Review*, 151(3), pp.277–287.

Bhorat, H., Kanbur, R. & Mayet, N., 2013. The impact of sectoral minimum wage laws in South Africa. *IZA Journal of Labor & Development*, 2(1), pp.1–27. Available at: http://www.izajold.com/content/2/1/1.

Bhorat, H., Naidoo, K. & Yu, D., 2014. Trade Unions in an Emerging Economy. *DPRU Working Paper 201402*, (July), pp.1–27.

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

Branson, N. et al., 2013. Changes in education, employment and earnings in South Africa – A cohort analysis. *SALDRU Working Paper 105*, (105).

Branson, N. & Wittenberg, M., 2014. Reweighting South African National Household Survey Data to Create a Consistent Series Over Time: A Cross-Entropy Estimation Approach. *South African Journal of Economics*, 82(1), pp.19–38. Available at: http://doi.wiley.com/10.1111/saje.12017.

Burger, R. & Woolard, I., 2005. The state of the labour market in South Africa after the first decade of democracy. *Journal of Vocational Education & Training*, 57(4), pp.453–476.

Butcher, K.F. & Rouse, C.E., 2001. Industrial Councils in South Africa. *Industrial and Labour Relations Review*, 54(2), pp.349–370.

Cortes, G.M., 2012. Where Have the Routine Workers Gone ? A Using Panel Data Study of Polarization Where Have the Routine Workers Gone. *Economics Discussion Paper Series 1224*, (October).

DiNardo, J., Fortin, N. & Lemieux, T., 1996. Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach. *Econometrica*, 64, pp.1001–1044. Available at: http://www.jstor.org/stable/2171954.

Fedderke, J., 2012. The Cost of Rigidity: The Case of the South African Labor Market. *Comparative Economic Studies*, 54, pp.809–842.

Firpo, S., Fortin, N.M. & Lemieux, T., 2013. Occupational Tasks and Changes in the Wage Structure. *Unpublished manuscript.*, (February). Available at: http://econ.sites.olt.ubc.ca/files/2014/02/FFL_Occupations_Revised.pdf.

Goos, M. & Manning, A., 2007. Lousy and Lovely Jobs: The Rising Polarization of Work in Britain. *The Review of Economics and Statistics*, 89(1), pp.118–133.

Goos, M., Manning, A. & Salomons, A., 2014. Explaining Job Polarization: Routine-Biased Technological Change and Offshoring. *American Economic Review*, 104(8), pp.2509–26.

Hertz, T., 2003. Upward Bias in the Estimated Returns to Education : Evidence from South Africa. *American Economic Review*, 93(4), pp.1354–1368.

Katz, L. & Autor, D., 1999. Changes in the Wage Structure and Earnings Inequality. *Handbook of Labor Economics*, 3, pp.1463–1555.

Kerr, A. & Wittenberg, M., 2013. Multiply Imputed Labour Income Data. *PALMS Dataset*, 1(August), pp.1–3.

Keswell, M. & Poswell, L., 2004. Returns to Education in South Africa: A retrospective sensitivity analysis of the available evidence. *South African Journal of Economics*, 72(4), pp.834–860.

Lam, D. et al., 2012. Education and Inequality : The South African Case. *SALDRU Working Paper*, (75), p.28.

Leibbrandt, M. et al., 2010. Employment and Inequality Outcomes in South Africa. *SALDRU*, (September 2015), pp.1–54. Available at: http://www.oecd.org/employment/employmentpoliciesanddata/45282868.pdf.

Leibbrandt, M., Finn, A. & Woolard, I., 2012. Describing and decomposing post-apartheid income inequality in South Africa. *Development Southern Africa*, 29(1), pp.19–34.

Leibbrandt, M. & Levinsohn, J., 2011. Fifteen Years On: Household Incomes in South Africa. *the National Bureau of Economic Research*, NBER Worki(January 2011). Available at: http://www.nber.org/papers/w16661.

Murphy, K. & Katz, L.F., 1992. Changes in Relative Wages , 1963-1987 : Supply and Demand Factors Author (s): Lawrence F . Katz and Kevin M . Murphy Published by : Oxford University Press. *The Quarterly Journal* *of Economics*, 107(1), pp.35–78.

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

Oldenski, L., 2014. Offshoring and the Polarization of the U . S . Labor. *Industrial and Labour Relations Review*, 67(December), pp.1–38.

Rodrik, D., 2008. Understanding South Africa's Economic Puzzles. *Economics of Transition*, 16(4), pp.769–797.

Rubin, D., 1987. *Multiple Imputation for Nonresponse in Surveys* I. John Wiley & Sons, ed., Hoboken nj.

Wittenberg, M., 2014. *Analysis of employment , real wage , and productivity trends in South Africa since 1994*, Available at: http://www.ilo.org/wcmsp5/groups/public/---ed_protect/---protrav/---travail/documents/publication/wcms_237808.pdf.

Yuan, Y., 2010. Multiple imputation for missing data: Concepts and new development. *SAS Institute Inc, Rockville, MD*, pp.1–13. Available at: http://www.math.montana.edu/~jimrc/classes/stat506/notes/multipleimputation-SAS.pdf.

Appendix A: Inference with Multiple Imputation

This presentation follows Yuan (2010). With m imputations, it is possible to compute m different sets of point and variance estimates for a parameter Q. Letting $\widehat{Q_\iota}$ and $\widehat{U_\iota}$ be the point and variance estimates for the *i*th imputed dataset, where $\iota\epsilon m$, then it is possible to compute variance and point estimates that consider potential errors in the imputation process.

The final point estimate is simply the average each point estimates for the m imputed datasets:

$$\bar{Q} = \frac{1}{m}\sum_{i=1}^{m} \widehat{Q_\iota}$$

In a similar manner, the average of the variance estimate represents the within-imputation variance:

$$\bar{U} = \frac{1}{m}\sum_{i=1}^{m} \widehat{U_\iota}$$

The between-imputation variance is calculated per:

$$B = \frac{1}{m-1}\sum_{i=1}^{m} (\widehat{Q_\iota} - \bar{Q})^2$$

The overall variance estimate associated with Q-bar is equal to:

$$T = \bar{U} + \left(1 + \frac{1}{m}\right) B$$

Inference using T is based on the t-distribution, with $\mathrm{v_m}$ degrees freedom, where:

$$v_m = (m-1)[1 + \frac{\bar{U}}{(1+m^{-1})B}]^2$$

Appendix B: RIF-Regression Tables and Graphs

Table 7: RIF Regression Coefficients

	2011			2001		
	10th Percentile	50th Percentile	90th Percentile	10th Percentile	50th Percentile	90th Percentile
Minimum Wage	-1.091***	-1.188***	-0.168***	-1.369***	-1.203***	-0.115***
	(0.038)	(0.032)	(0.017)	(0.069)	(0.043)	(0.027)
RTI	0.260***	0.096	-0.120	0.160*	0.341***	-0.118
	(0.054)	(0.063)	(0.074)	(0.082)	(0.086)	(0.142)
RTI Finance Interaction	-0.905***	0.674***	-0.234	-0.864***	0.914***	-0.766**
	(0.092)	(0.125)	(0.170)	(0.130)	(0.157)	(0.324)
RTI Manu/Con Interaction	-0.902***	-0.505***	-0.228	-0.759***	-0.192	-0.031
	(0.131)	(0.157)	(0.176)	(0.124)	(0.172)	(0.359)
Abstract Task	-0.058	0.615***	0.857***	-0.086	0.870***	1.030***
	(0.034)	(0.046)	(0.059)	(0.062)	(0.070)	(0.130)
No Primary School	-0.111	0.019	0.026	-0.778***	-0.025	0.037
	(0.072)	(0.049)	(0.019)	(0.099)	(0.040)	(0.038)
Some Primary School	-0.011	0.022	0.007	-0.366***	-0.064*	-0.020
	(0.055)	(0.040)	(0.017)	(0.070)	(0.034)	(0.029)
Finished Primary School	0.190***	0.127***	0.055***	0.119*	0.160***	0.037
	(0.045)	(0.036)	(0.015)	(0.060)	(0.032)	(0.026)
Finished High School	0.348***	0.548***	0.094***	0.265***	0.630***	0.192***
	(0.047)	(0.042)	(0.021)	(0.062)	(0.044)	(0.045)
Some Tertiary Education	0.393***	1.002***	1.037***	0.241***	1.005***	0.690***
	(0.048)	(0.048)	(0.054)	(0.069)	(0.062)	(0.088)
Bachelors	0.422***	1.010***	1.682***	0.239***	1.122***	1.413***
	(0.050)	(0.052)	(0.091)	(0.072)	(0.079)	(0.196)
Honors	0.433***	1.007***	2.157***	0.263***	1.007***	1.942***
	(0.050)	(0.054)	(0.124)	(0.069)	(0.071)	(0.217)
Union Membership	0.147***	0.339***	0.009	0.238***	0.360***	-0.021
	(0.015)	(0.020)	(0.025)	(0.020)	(0.025)	(0.055)
Married	0.023	0.074***	0.063***	0.065**	0.173***	0.105**
	(0.015)	(0.016)	(0.018)	(0.025)	(0.020)	(0.034)
10-15 yrs exp	-0.069	-0.025	-0.048	0.093	-0.145**	-0.079
	(0.044)	(0.047)	(0.072)	(0.084)	(0.059)	(0.072)
20-25 yrs exp	-0.006	-0.020	-0.004	0.121**	-0.015	0.057
	(0.025)	(0.026)	(0.029)	(0.046)	(0.033)	(0.045)
25-30 yrs exp	0.003	-0.006	0.019	0.158***	0.105***	0.212***
	(0.026)	(0.025)	(0.030)	(0.045)	(0.032)	(0.063)
30-35 yrs exp	0.019	0.068**	0.056	0.220***	0.151***	0.243***
	(0.026)	(0.027)	(0.032)	(0.045)	(0.035)	(0.064)
35-40 yrs exp	0.070**	0.098***	0.090**	0.256***	0.237***	0.327***
	(0.026)	(0.027)	(0.034)	(0.043)	(0.039)	(0.081)
40-45 yrs exp	0.023	0.153***	0.151***	0.196***	0.237***	0.403***
	(0.029)	(0.027)	(0.036)	(0.047)	(0.039)	(0.073)
45-50 yrs exp	0.047	0.159***	0.165***	0.269***	0.281***	0.265***
	(0.030)	(0.029)	(0.037)	(0.047)	(0.038)	(0.078)
50 plus yrs exp	0.094**	0.168***	0.111**	0.321***	0.256***	0.283**
	(0.038)	(0.034)	(0.037)	(0.056)	(0.049)	(0.097)
Race	0.031*	0.596***	0.693***	0.045***	0.541***	0.825***

	(0.014)	(0.026)	(0.048)	(0.018)	(0.029)	(0.112)
Gender	-0.128***	-0.154***	-0.180***	-0.154***	-0.146***	-0.137***
	(0.016)	(0.016)	(0.021)	(0.024)	(0.021)	(0.038)
Agriculture	-0.948***	-0.272***	-0.053	-2.117***	0.067	-0.434**
	(0.064)	(0.062)	(0.083)	(0.125)	(0.080)	(0.186)
Mining	-0.306***	0.879***	0.132	-0.248**	0.243**	-0.336
	(0.071)	(0.089)	(0.121)	(0.079)	(0.100)	(0.254)
Manufacture	-0.304***	0.437***	0.084	-0.401***	0.342***	-0.326
	(0.070)	(0.088)	(0.120)	(0.081)	(0.105)	(0.276)
Utilities	-0.660***	0.282**	0.061	-0.588***	0.336**	0.062
	(0.076)	(0.105)	(0.157)	(0.078)	(0.112)	(0.265)
Construction	-0.705***	-0.005	-0.200*	-1.043***	-0.241**	-0.545**
	(0.058)	(0.072)	(0.090)	(0.098)	(0.088)	(0.203)
Trade	-0.185***	0.572***	-0.013	-0.488***	0.364***	-0.497**
	(0.051)	(0.063)	(0.088)	(0.080)	(0.083)	(0.204)
Transport	-0.694***	0.286***	0.004	-0.707***	0.476***	-0.264
	(0.056)	(0.069)	(0.094)	(0.092)	(0.089)	(0.199)
Services	-0.747***	0.243***	-0.032	-0.665***	0.467***	-0.517**
	(0.054)	(0.060)	(0.091)	(0.075)	(0.085)	(0.198)
Domestic Services	-0.664***	0.083	0.169*	-1.585***	0.253**	-0.190
	(0.062)	(0.060)	(0.083)	(0.118)	(0.084)	(0.179)
Constant	1.636***	1.296***	3.107***	1.492***	0.815***	3.032***
	(0.062)	(0.076)	(0.092)	(0.110)	(0.106)	(0.182)

The values represent log wage differentials multiplied by 100. Errors in brackets are calculated from the bootstrapping and imputations described in section 4.4. Asterisks indicate significance at the 10%(*), 5%(**) and 1%(***) levels with 9 degrees' freedom.

Appendix C: Data and Reweighting

Table 8: Sector contributions to GVA over time

Year	**1980**	**2001**	**2011**
Agriculture	3.15	2.82	2.59
Mining	18.74	12.4	8.9
Manufacturing	17.9	15.82	14.37
Utilities	2.31	2.94	2.68
Construction	4.09	2.45	3.73
Trade	13.29	14.41	14.97
Transport	6.02	8.42	9.2
Finance	13.36	18.2	21.2
Services	16.59	16.4	16.45
Domestic	4.56	6.15	5.92

Table 9: Descriptive Statistics

	2001		2011		
Variable	Mean	Standard Deviation	Mean	Standard Deviation	Difference in Means
Log hourly wages*	2.138	1.126	2.374	1.068	0.236
Task Variables					
RTI	0.449	0.156	0.452	0.156	0.003
RTI Finance Interaction	0.045	0.152	0.059	0.165	0.014
RTI Manufacture Interaction	0.079	0.191	0.066	0.1741	-0.011
Education					
No School	0.065	0.247	0.025	0.157	-0.040
Some Primary	0.158	0.364	0.085	0.279	-0.073
Finished Primary	0.071	0.256	0.044	0.205	-0.027
Some High School	0.308	0.461	0.328	0.469	0.020
Finished High School	0.235	0.424	0.303	0.460	0.068
Some Tertiary	0.086	0.280	0.126	0.331	0.040
Bachelors	0.038	0.192	0.051	0.220	0.013
Honours	0.027	0.163	0.026	0.160	-0.001
Institutions					
Minimum Wage Sector	0.237	0.425	0.245	0.430	0.008
Union	0.335	0.472	0.299	0.458	-0.036
Potential Experience					
0 - 10 years	0.027	0.161	0.022	0.146	-0.005
10 - 15 years	0.103	0.304	0.120	0.325	0.017
15 - 20 years	0.154	0.361	0.161	0.367	0.006
20 - 25 years	0.153	0.360	0.160	0.366	0.007
25 - 30 years	0.141	0.348	0.141	0.348	0.000
30 - 35 years	0.127	0.333	0.111	0.314	-0.016
35 - 40 years	0.099	0.299	0.099	0.299	0.000
40 - 45 years	0.080	0.271	0.080	0.271	0.000

45 - 50 years	0.054	0.226	0.053	0.225	-0.001
50 + years	0.050	0.218	0.041	0.198	-0.009
Discrimination					
Married	0.590	0.492	0.517	0.500	-0.072
Race	0.173	0.378	0.143	0.350	-0.030
Gender	0.413	0.492	0.447	0.497	0.034
Sector					
Agriculture	0.089	0.285	0.052	0.222	-0.037
Mining	0.061	0.240	0.031	0.172	-0.031
Manufacturing	0.159	0.366	0.137	0.344	-0.022
Trade	0.155	0.362	0.179	0.383	0.024
Utilities	0.013	0.113	0.007	0.086	-0.006
Services	0.201	0.400	0.232	0.422	0.031
Construction	0.055	0.228	0.071	0.257	0.016
Transport	0.053	0.224	0.057	0.232	0.004
Domestic	0.119	0.323	0.102	0.302	-0.017
Finance	0.095	0.293	0.132	0.339	0.038
No. Observations	41933		34087		
* Calculated from the average of the 10 imputed wage variables					

Table 10: Reweighting Regression

Probit Reweighting Regression	73756
Number of obs	73798
Wald chi2(95)	4973.52
Prob > chi2	0.00
Log pseudolikelihood	-25083442.00
Pseudo R2	0.08

year	Coefficient	Robust Standard Error	z-score	P>|z|	95% Confidence Interval	
Interaction_manager_no	-0.8790	0.4861	-1.81	0.071	-1.832	0.0737
Interaction_manager_sp	-0.6192	0.3243	-1.91	0.056	-1.255	0.0164
Interaction_manager_fp	-0.1882	0.3464	-0.54	0.587	-0.867	0.4908
Interaction_manager_sh	-0.7509	0.1527	-4.92	0.000	-1.05	-0.452
Interaction_professional_fp	-2.0124	0.6504	-3.09	0.002	-3.287	-0.738
Interaction_professional_sh	-0.8242	0.2105	-3.92	0.000	-1.237	-0.412
Interaction_professional_st	-0.4759	0.1941	-2.45	0.014	-0.856	-0.095
Interaction_professional_bch	-1.3844	0.3075	-4.5	0.000	-1.987	-0.782
Interaction_professional_hon	-1.2941	0.2770	-4.67	0.000	-1.837	-0.751
Interaction_technicians_no	-1.0134	0.4214	-2.4	0.016	-1.839	-0.188
Interaction_technicians_sp	-0.8210	0.2109	-3.89	0.000	-1.234	-0.408
Interaction_technicians_fp	-1.2253	0.2387	-5.13	0.000	-1.693	-0.757
Interaction_technicians_sh	-0.8825	0.1348	-6.55	0.000	-1.147	-0.618
Interaction_technicians_st	-0.4872	0.1696	-2.87	0.004	-0.82	-0.155
Interaction_technicians_bch	0.9509	0.3088	3.08	0.002	0.3457	1.5561
Interaction_technicians_hon	0.3388	0.2796	1.21	0.226	-0.209	0.8867
Interaction_clerks_no	-1.1162	0.3496	-3.19	0.001	-1.801	-0.431
Interaction_clerks_sp	-0.9213	0.1982	-4.65	0.000	-1.31	-0.533
Interaction_clerks_fp	-1.2115	0.1893	-6.4	0.000	-1.583	-0.84
Interaction_clerks_sh	-0.8903	0.1326	-6.72	0.000	-1.15	-0.63
Interaction_clerks_st	0.0855	0.1743	0.49	0.624	-0.256	0.4271
Interaction_clerks_bch	0.1101	0.3164	0.35	0.728	-0.51	0.7303
Interaction_clerks_hon	-0.6980	0.3142	-2.22	0.026	-1.314	-0.082
Interaction_sevice_no	-1.0198	0.1795	-5.68	0.000	-1.372	-0.668
Interaction_sevice_sp	-0.8017	0.1471	-5.45	0.000	-1.09	-0.514
Interaction_sevice_fp	-1.3063	0.1519	-8.6	0.000	-1.604	-1.008
Interaction_sevice_sh	-0.8262	0.1307	-6.32	0.000	-1.082	-0.57
Interaction_sevice_st	-0.0278	0.1807	-0.15	0.878	-0.382	0.3263
Interaction_sevice_bch	0.1132	0.3475	0.33	0.745	-0.568	0.7942
Interaction_sevice_hon	-0.6279	0.3718	-1.69	0.091	-1.357	0.1008
Interaction_agriculture_no	-0.2182	0.2008	-1.09	0.277	-0.612	0.1754
Interaction_agriculture_sp	0.3426	0.1653	2.07	0.038	0.0187	0.6665
Interaction_manager_st	-0.0836	0.1791	-0.47	0.641	-0.435	0.2674
Interaction_manager_bch	0.0467	0.3086	0.15	0.880	-0.558	0.6516
Interaction_manager_hon	0.0304	0.2755	0.11	0.912	-0.51	0.5704
Interaction_professional_sp	-0.9651	0.3698	-2.61	0.009	-1.69	-0.24
Interaction_agriculture_fp	0.2587	0.2452	1.06	0.291	-0.222	0.7394

Interaction_agriculture_fh	0.8712	0.2594	3.36	0.001	0.3627	1.3797
Interaction_agriculture_st	2.0750	0.5919	3.51	0.000	0.9149	3.2352
Interaction_agriculture_hon	1.8523	0.9188	2.02	0.044	0.0515	3.6531
Interaction_craft_no	-0.9388	0.0973	-9.65	0.000	-1.13	-0.748
Interaction_craft_sp	-0.1934	0.0586	-3.3	0.001	-0.308	-0.079
Interaction_craft_fp	-0.4093	0.0667	-6.14	0.000	-0.54	-0.279
Interaction_craft_fh	0.8154	0.1339	6.09	0.000	0.5529	1.0778
Interaction_craft_st	0.9277	0.1366	6.79	0.000	0.66	1.1954
Interaction_craft_bch	1.5202	0.4317	3.52	0.000	0.674	2.3663
Interaction_craft_hon	-0.0895	0.4044	-0.22	0.825	-0.882	0.7032
Interaction_operator_no	-0.7248	0.0912	-7.95	0.000	-0.903	-0.546
Interaction_operator_sp	-0.3832	0.0584	-6.57	0.000	-0.498	-0.269
Interaction_operator_fp	-0.5021	0.0638	-7.87	0.000	-0.627	-0.377
Interaction_operator_fh	1.1003	0.1347	8.17	0.000	0.8363	1.3643
Interaction_operator_st	1.2957	0.1610	8.05	0.000	0.9801	1.6114
Interaction_operator_bch	1.4218	0.4690	3.03	0.002	0.5026	2.3411
Interaction_elementary_no	-0.6419	0.0509	-12.61	0.000	-0.742	-0.542
Interaction_elementary_sp	-0.2213	0.0410	-5.4	0.000	-0.302	-0.141
Interaction_elementary_fp	-0.4264	0.0435	-9.81	0.000	-0.512	-0.341
Interaction_elementary_fh	0.8525	0.1320	6.46	0.000	0.5938	1.1113
Interaction_elementary_st	0.6847	0.1565	4.38	0.000	0.378	0.9914
Interaction_elementary_bch	1.0683	0.5530	1.93	0.053	-0.016	2.1522
Interaction_elementary_hon	-0.3900	0.6782	-0.58	0.565	-1.719	0.9392
Interaction_manager_no	-0.5989	0.0639	-9.37	0.000	-0.724	-0.474
Interaction_manager_sp	-0.2460	0.0501	-4.91	0.000	-0.344	-0.148
Interaction_manager_fp	-0.3486	0.0575	-6.06	0.000	-0.461	-0.236
Interaction_manager_fh	0.7796	0.1440	5.41	0.000	0.4973	1.062
Interaction_manager_st	0.7698	0.3273	2.35	0.019	0.1283	1.4113
Interaction_manager_bch	0.9227	0.7390	1.25	0.212	-0.526	2.371
manager	0.9837	0.1362	7.22	0.000	0.7167	1.2507
professional	1.8590	0.1479	12.57	0.000	1.5691	2.1489
technicians	0.8149	0.1313	6.21	0.000	0.5575	1.0722
clerks	0.8380	0.1311	6.39	0.000	0.5812	1.0949
sevice	0.9586	0.1316	7.29	0.000	0.7008	1.2165
agriculture	-1.4551	0.1228	-11.85	0.000	-1.696	-1.215
craft	0.0511	0.0389	1.31	0.189	-0.025	0.1273
operator	-0.0922	0.0395	-2.33	0.020	-0.17	-0.015
elementary	0.2792	0.0341	8.19	0.000	0.2123	0.346
Interaction_uni_five	0.0380	0.1217	0.31	0.755	-0.201	0.2766
Interaction_uni_ten	0.0642	0.0530	1.21	0.225	-0.04	0.168
Interaction_uni_twe	-0.0499	0.0436	-1.14	0.253	-0.135	0.0356
Interaction_uni_twf	-0.0907	0.0447	-2.03	0.042	-0.178	-0.003
Interaction_uni_thr	-0.0885	0.0453	-1.95	0.051	-0.177	0.0003
Interaction_uni_thf	0.1578	0.0466	3.39	0.001	0.0664	0.2492
Interaction_uni_for	0.1921	0.0481	3.99	0.000	0.0978	0.2865
Interaction_uni_fof	0.1095	0.0540	2.03	0.042	0.0037	0.2152
Interaction_uni_fift	0.3304	0.0577	5.73	0.000	0.2173	0.4435
Interaction_fif_sp	-0.7465	0.1502	-4.97	0.000	-1.041	-0.452
Interaction_fif_fh	-0.4857	0.0795	-6.11	0.000	-0.642	-0.33
Interaction_fif_st	-0.4379	0.1237	-3.54	0.000	-0.68	-0.195
Interaction_fif_bch	-0.5377	0.2827	-1.9	0.057	-1.092	0.0163

Interaction_fif_hon	-0.5131	0.2587	-1.98	0.047	-1.02	-0.006
Interaction_fof_no	-0.2708	0.0794	-3.41	0.001	-0.426	-0.115
Interaction_fof_sp	-0.2082	0.0524	-3.97	0.000	-0.311	-0.105
Interaction_fof_fh	-0.1318	0.0942	-1.4	0.161	-0.316	0.0527
Interaction_fof_st	0.2671	0.1345	1.99	0.047	0.0035	0.5307
Interaction_fof_bch	-0.0890	0.2965	-0.3	0.764	-0.67	0.4921
Interaction_fof_hon	0.1641	0.2769	0.59	0.554	-0.379	0.7069
Interaction_for_no	-0.2498	0.0918	-2.72	0.007	-0.43	-0.07
Interaction_for_sp	-0.4163	0.0529	-7.87	0.000	-0.52	-0.313
Interaction_for_fh	-0.0534	0.0871	-0.61	0.539	-0.224	0.1172
Interaction_for_st	0.0082	0.1282	0.06	0.949	-0.243	0.2595
Interaction_for_bch	-0.3472	0.2926	-1.19	0.235	-0.921	0.2263
Interaction_for_hon	0.0605	0.2663	0.23	0.820	-0.461	0.5825
Interaction_thf_no	-0.3006	0.1163	-2.58	0.010	-0.529	-0.073
Interaction_thf_sp	-0.4844	0.0562	-8.62	0.000	-0.595	-0.374
Interaction_thf_fh	-0.1370	0.0843	-1.63	0.104	-0.302	0.0282
Interaction_thf_st	-0.0414	0.1239	-0.33	0.738	-0.284	0.2015
Interaction_thf_bch	-0.2809	0.2855	-0.98	0.325	-0.841	0.2787
Interaction_thf_hon	-0.1484	0.2595	-0.57	0.567	-0.657	0.3602
Interaction_thr_no	-0.4008	0.1691	-2.37	0.018	-0.732	-0.069
Interaction_thr_sp	-0.5054	0.0651	-7.77	0.000	-0.633	-0.378
Interaction_thr_fh	-0.2373	0.0813	-2.92	0.003	-0.397	-0.078
Interaction_thr_st	-0.1549	0.1205	-1.29	0.199	-0.391	0.0812
Interaction_thr_bch	-0.2010	0.2839	-0.71	0.479	-0.757	0.3554
Interaction_thr_hon	-0.4489	0.2521	-1.78	0.075	-0.943	0.0452
Interaction_twf_no	-0.4809	0.1939	-2.48	0.013	-0.861	-0.101
Interaction_twf_sp	-0.4282	0.0686	-6.24	0.000	-0.563	-0.294
Interaction_twf_fh	-0.3401	0.0795	-4.28	0.000	-0.496	-0.184
Interaction_twf_st	-0.3051	0.1188	-2.57	0.010	-0.538	-0.072
Interaction_twf_bch	-0.5957	0.2816	-2.12	0.034	-1.148	-0.044
Interaction_twf_hon	-0.4991	0.2478	-2.01	0.044	-0.985	-0.013
Interaction_twe_no	-0.8205	0.4326	-1.9	0.058	-1.668	0.0274
Interaction_twe_sp	-0.5319	0.0804	-6.61	0.000	-0.69	-0.374
Interaction_twe_fh	-0.5255	0.0793	-6.62	0.000	-0.681	-0.37
Interaction_twe_st	-0.3689	0.1185	-3.11	0.002	-0.601	-0.137
Interaction_twe_bch	-0.7609	0.2805	-2.71	0.007	-1.311	-0.211
Interaction_twe_hon	-0.8511	0.2518	-3.38	0.001	-1.345	-0.357
Interaction_ten_fh	-0.6219	0.0896	-6.94	0.000	-0.798	-0.446
Interaction_ten_st	-0.6246	0.1620	-3.86	0.000	-0.942	-0.307
Interaction_ten_bch	-0.9121	0.2992	-3.05	0.002	-1.499	-0.326
Interaction_ten_hon	-0.5132	0.3093	-1.66	0.097	-1.119	0.093
finish_high	-0.2694	0.1463	-1.84	0.065	-0.556	0.0172
unionmem	-0.2135	0.0318	-6.71	0.000	-0.276	-0.151
publicemp	-0.2014	0.0173	-11.67	0.000	-0.235	-0.168
race	-0.4912	0.0197	-24.92	0.000	-0.53	-0.453
gender	0.0490	0.0133	3.68	0.000	0.0229	0.0752
_cons	0.2217	0.0302	7.34	0.000	0.1625	0.2809

Table 11: Codes for table 10

Code	Explanation
no	*No Primary*
sp	*Some Primary*
fp	*Full Primary*
sh	*Some High School*
st	*Some Tertiary*
bch	*Bachelors*
hon	*Honours*
five	*Five years' experience*
ten	*Ten years' experience*
twe	*Twenty years' experience*
twf	*Twenty five years' experience*
thr	*Thirty years' experience*
thf	*Thirty five years' experience*
for	*Forty years' experience*
fof	*Forty five years' experience*
fift	*Fifty years' experience*

Table 12: Occupation Task Scores

Abstract	Routine	Manu1al	Three Digit	Title
0.908290923	0.222086251	0.528397381	111	Legislators
0.908290923	0.222086251	0.528397381	112	Senior government officials
0.908290923	0.222086251	0.528397381	113	Traditional chiefs and heads of villages
0.9808411	0.209110513	0.528397381	114	Senior officials of special-interest organisations
0.908290923	0.222086281	0.44952178	121	Directors and chief executives

0.981398225	0.190402374	0.44952178	122	Production and operations department managers
0.987292588	0.207384378	0.44952178	123	Other department managers
0.968071878	0.207165346	0.535176516	131	General managers
0.948054731	0.791808248	0.739836156	211	Physicists, chemists and related professionals
0.814700007	0.466601282	0.739836156	212	Mathematicians, statisticians and related professionals
0.965021253	0.632646799	0.739836156	213	Computing professionals
0.986008227	0.791866958	0.739836156	214	Architects, engineers and related professionals
0.9678756	0.763958931	0.829145491	221	Life science professionals
0.89930886	0.892113209	0.829145491	222	Health professionals (except nursing)
0.736935854	0.920051455	0.829145491	223	Nursing and midwifery professionals
1	0.098351017	0.834176362	231	College, university and higher education teaching professionals
0.653837323	0.024654256	0.834176362	232	Secondary education teaching professionals
0.766794324	0.193261623	0.834176362	233	Primary and pre-primary education teaching professionals
0.905490935	0.578586042	0.834176362	234	Special education teaching professionals
0.827866733	0.64707464	0.834176362	235	Other teaching professionals
0.983613372	0.625196517	0.261942714	241	Business professionals
0.730475008	0.181404606	0.261942714	242	Legal professionals
0.906828821	0.509105861	0.261942714	243	Archivists, librarians and related information professionals
0.829357922	0.165790096	0.261942714	244	Social science and related professionals
0.647873878	0.24778904	0.261942714	245	Writers and creative or performing artists
0.744779527	0.014446792	0.261942714	246	Religious professionals
0.762925506	0.941696167	0.83489567	311	Physical and engineering science technicians
0.830713511	0.69419992	0.83489567	312	Computer associate professionals
0.609819114	0.798752487	0.83489567	313	Optical and electronic equipment operators
0.754227936	0.613525629	0.83489567	314	Ship and aircraft controllers and technicians
0.516768873	0.070314549	0.83489567	315	Safety and quality inspectors
0.793690681	0.800626278	0.862751901	321	Life science technicians and related associate professionals
0.757791996	0.652863801	0.862751901	322	Modern health associate professionals (except nursing)
0.736935854	0.920051455	0.862751901	323	Nursing and midwifery associate professionals

0.632149756	0.875808716	0.862751901	324	Traditional medicine practitioners and faith healers
0.656660259	0.034451894	0.335184038	331	Primary education teaching associate professionals
0.656660259	0.034451894	0.335184038	332	Pre-primary education teaching associate professionals
0.656660259	0.034451894	0.335184038	333	Special education teaching associate professionals
0.656660259	0.034451894	0.335184038	334	Other teaching associate professionals
0.817722738	0.265995145	0.446696281	341	Finance and sales associate professionals
0.836860657	0.533882856	0.446696281	342	Business services agents and trade brokers
0.689248204	0.875914335	0.446696281	343	Administrative associate professionals
0.914442718	0.784657061	0.446696281	344	Customs, tax and related government associate professionals
0.504203975	0	0.446696281	345	Police inspectors and detectives
0.846199393	0.034214124	0.446696281	346	Social work associate professionals
0.799736321	0.642888725	0.446696281	347	Artistic, entertainment and sports associate professionals
0.744779527	0.014446792	0.446696281	348	Religious associate professionals
0.540925741	0.999152064	0.579299688	411	Secretaries and keyboard-operating clerks
0.525933504	0.841445982	0.579299688	412	Numerical clerks
0.607171297	0.657197595	0.579299688	413	Material-recording and transport clerks
0.138942808	0.483748823	0.579299688	414	Library, mail and related clerks
0.534306109	0.844088733	0.579299688	419	Other office clerks
0.4839423	0.929901242	0	421	Cashiers, tellers and related clerks
0.520453274	0.778603554	0	422	Client information clerks
0.433216959	0.128702596	0.63908875	511	Travel attendants and related workers
0.483670533	0.537969232	0.63908875	512	Housekeeping and restaurant services workers
0.488016278	0.316596597	0.63908875	513	Personal care and related workers
0.536927521	0.882451892	0.63908875	514	Other personal service workers
0.766992152	0.408498853	0.63908875	515	Astrologers, fortune-tellers and related workers
0.747466326	0.306145161	0.63908875	516	Protective services workers
0.560570359	0.324867725	0.485641927	521	Fashion and other models
0.576374531	0.280727178	0.485641927	522	Shop salespersons and demonstrators
0.50110507	0.123199709	0.485641927	523	Stall and market salespersons

0.988618195	0.098577514	0.903482378	611	Market gardeners and crop growers
0.81171906	0.475862294	0.903482378	612	Market-oriented animal producers and related workers
0.988618195	0.098577514	0.903482378	613	Market-oriented crop and animal producers
0.390778244	0.230426371	0.903482378	614	Forestry and related workers
0.276701093	0.3752608	0.903482378	615	Fishery workers, hunters and trappers
0.382541209	0.758947372	0.932926118	711	Miners, shotfirers, stone cutters and carvers
0.51269269	0.905650496	0.932926118	712	Building frame and related trades workers
0.639625371	0.94102931	0.932926118	713	Building finishers and related trades workers
0.439210504	0.928296387	0.932926118	714	Painters, building structure cleaners and related trades workers
0.599591672	0.925061464	0.79603523	721	Metal moulders, welders, sheet-metal workers, structural-metal preparers, and related trades workers
0.700375736	0.902673364	0.79603523	722	Blacksmiths, tool-makers and related trades workers
0.575402737	0.924395859	0.79603523	723	Machinery mechanics and fitters
0.616271496	0.939844251	0.79603523	724	Electrical and electronic equipment mechanics and fitters
0.530919075	1	0.784601569	731	Precision workers in metal and related materials
0.503031373	0.851493001	0.784601569	732	Potters, glass-makers and related trades workers
0.433735371	0.801619709	0.784601569	733	Handicraft workers in wood, textile, leather and related material
0.411804646	0.937965512	0.784601569	734	Printing and related trades workers
0.448698401	0.885620356	0.663674533	741	Food processing and related trades workers
0.544781566	0.87999481	0.663674533	742	Wood treaters, cabinet-makers and related trades workers
0.486788124	0.938564599	0.663674533	743	Textile, garment and related trades workers
0.439952105	0.863500476	0.663674533	744	Pelt, leather and shoemaking trades workers
0.321352482	0.727241874	0.869216084	811	Mining and mineral-processing-plant operators
0.535988867	0.778529286	0.869216084	812	Metal-processing-plant operators
0.308433115	0.768596351	0.869216084	813	Glass, ceramics and related plant-operators
0.169949368	0.760288537	0.869216084	814	Wood-processing-and papermaking-plant operators
0.296295106	0.86158061	0.869216084	815	Chemical-processing-plant operators
0.672335684	0.830305874	0.869216084	816	Power-production and related plant operators
0.941644132	0.791489363	0.869216084	817	Automated-assembly-line and industrial-robot operators
0.930277824	0.789833248	0.77660054	821	Metal-and mineral-products machine operators

0.507222533	0.833759189	0.77660054	822	Chemical-products machine operators
0.798094988	0.814487398	0.77660054	823	Rubber- and plastic-products machine operators
0.152767971	0.761834383	0.77660054	824	Wood-products machine operators
0.409711719	0.863317728	0.77660054	825	Printing-, binding-and paper-products machine operators
0.000962406	0.909055531	0.77660054	826	Textile-, fur-and leather-products machine operators
0.353300512	0.805007935	0.77660054	827	Food and related products machine operators
0.509744227	0.89826566	0.77660054	828	Assemblers
0	0.568016946	0.77660054	829	Other machine operators and assemblers
0.807764053	0.255667567	1	831	Locomotive engine drivers and related workers
0.245327547	0.229224458	1	832	Motor vehicle drivers
0.266666085	0.712948918	1	833	Agricultural and other mobile plant operators
0.733277023	0.644468665	1	834	Ships' deck crews and related workers
0.318549037	0.181913272	0.849359095	911	Street vendors and related workers
0.668770671	0.126058787	0.849359095	912	Shoe cleaning and other street services elementary occupations
0.218727648	0.053779405	0.849359095	913	Domestic and related helpers, cleaners and launderers
0.27118057	0.461632967	0.849359095	914	Building caretakers, window and related cleaners
0.099418461	0.154656902	0.849359095	915	Messengers, porters, doorkeepers and related workers
0.456808031	0.688169837	0.849359095	916	Garbage collectors and related labourers
0.463238478	0.562869728	0.871237755	921	Agricultural, fishery and related labourers
0.35784182	0.784932435	0.763575494	931	Mining and construction labourers
0.193710551	0.710276663	0.763575494	932	Manufacturing labourers
0.643761158	0.758887827	0.763575494	933	Transport labourers and freight handlers

Scores are standardized to lie between 0 and 1.