# JOB MOBILITY AND WAGE GROWTH IN SOUTH AFRICAN MANUFACTURING

Ayanda Hlatshwayo

2019

# Abstract

We estimate the relative contribution of worker and firm fixed effects (i.e. the unobserved time invariant worker and firm characteristics) to wages levels and wage growth in the manufacturing sector using the SARS-NT tax administrative panel dataset between 2011 to 2016. We find that worker fixed effects dominate firm fixed effects for both wage levels and wage growth. Low income employees' have higher firm fixed effects and lower worker effects compared to high income employees. This shows us that the company you work for is more important for low wage workers wage level. Whereas high wage workers can obtain a high wage regardless of the company they are employed in. However, for wage growth we find that the worker fixed effect explains a similar proportion of the variation in wages across low and high wage workers. We also find that the importance of firm fixed effects in explaining the variation in wage levels increases with firm size.

We find a negative correlation between the worker and firm fixed effects in both the wage levels and growth indicating negative assortative matching in the manufacturing sector i.e. high wage workers seem to sort into low wage firms which is the same finding as most of the previous literature. However, our firm size correlations reveal negative assortative matching for small firms (less than 500 employees) and positive assortative matching for medium and large firms (500 or more employees). Therefore, the correlation across the sample, which is similar to the findings in the literature, hides the differences we find when we disaggregate by firm size.



JEL Codes: J31, J63, C23.

Keywords: Job mobility; linked employer-employee data; worker and firm heterogeneity; wages.

# 1 Introduction

Job mobility and wage growth have been studied for several decades[1]. The availability and use of linked employer-employee panel datasets has led to an increased interest in understanding the share of individual and firm level unobserved time-invariant characteristics in explaining wages. However, the literature has mainly focused on wage levels and the large degree of dispersion in wages. This paper closely follows Sorensen and Vejlin (2011) using the South African linked employer-employee database complied from tax administrative data for the manufacturing sector between 2011 and 2016. We take a closer look at wages in formal South African manufacturing firms analysing both wage levels and wage growth with a focus on the relative importance of observable characteristics and unobservable[2] individual/worker and firm fixed effects. Thus, contributing to the literature on unobserved worker and firm heterogeneity in both wage levels and wage growth.

The Abowd et al. (1999) seminal paper on high wage workers and high wage firms was the first paper to introduce the analysis of both observed and unobserved heterogeneity, as well as the extent of sorting between workers and firms. Their methodology has been adopted by a number of studies using data from various countries which include Goux and Maurin, (1999); Barth and Dale-Olsen (2003), Andrews et al. (2008), Grutter and Lalive, (2004); and Hyslop and Mare (2009). Bhorat et al. (2017)[3] were the first to use the South African linked employer-employee database to decompose wage differentials by worker and firm characteristics using the Abowd et al. (1999) method and found results consistent with French and Austrian labour markets that approximately 61 per cent of the variation in wages was due to individual effects while 13 per cent was due to firm-level effects. Our wage level estimates are restricted to only manufacturing sector jobs and we find that worker fixed effects explain 53 per cent of the variation in wages and firm fixed effects 26 per cent with the remainder is accounted for by the observable characteristics and residual[4]. This is aligned with the results of most papers in the literature which find that worker fixed effects are more important than firm fixed effects in the explanation of the variation in wage levels even though the estimates in the literature are based on the entire economy. Some examples include Sorensen and Vejlin (2011) who used Danish data

---

[1] Bartel and Borjas (1978) and Mincer (1986)

[2] As early as Topel and Ward (1988) we understood that individuals may differ for unobservable reasons/charateristics in their decision to change jobs. However, estimating these effects has become possible with linked employer-employee data.

[3] The Bhorat et. al. paper merges in the CIT data which reduced the number of observations and allows for analysis across all sectors. Whereas this paper only uses the IRP5 data giving us more employees, movers and connected firms even between branches (if the PAYE numbers are different to the main company). Further, we keep all the jobs held by an individual in manufacturing firms.

[4] Generally, the relative shares of the worker fixed effects, firm fixed effects, observable characteristics and residual sum to 1, however, the covariance’s can become negative making it difficult to interpret the numbers as shares (Cornelissen, 2008).

and found that worker fixed effects explained 58 per cent of the variation in wage levels and 14 per cent for firm fixed effects. Card et al. (2013) using German data found that worker fixed effects explain between 50 and 60 per cent and around 20 per cent for firm effects. As well as Jinkins and Morin (2018) using Danish data found that 78 per cent wage heterogeneity is explained by the worker fixed effects and 12 per cent by the firm fixed effects. Bassier (2019) also used the South African linked employer-employee panel between 2011 and 2014 and found that South African firms have a large role in determining wages relative to the international literature on firm wage premia and worker effects. He finds that firm effects explain 25 per cent of the total variance in log wages (similar to our estimate although he used all the sectors in the economy), 60 per cent of the average gender wage gap, and 40 per cent of the average gap between middle and lower income decile workers.

Most of the literature has focused on wage levels. However, Sorensen and Vejlin (2011) and Jinkins and Morin (2018) investigate using the two-way fixed effect model for wage growth instead of levels using Danish data. Sorensen and Vejlin (2011) estimate that worker fixed effects only explained 7 to 12 per cent of the variation in wage growth and 4 to 10 per cent for firm fixed effects. While, Jinkins and Morin (2018) assume that the worker effect is differenced away and find that firm fixed effects explain 21 per cent of the variance in wage growth for job to job movers. Our estimates are higher for the full sample[5] with worker fixed effects explaining 42 per cent of the wage growth heterogeneity and 7 per cent for firm fixed effects. m Our paper also looks at worker and firm fixed effects across gender and firm size. Jinkins and Morin (2018) also extend Sorensen and Vejlin (2011) by including match effects. They find that 66 per cent of the variance in wage growth of job to job movers stems from the variance in of the change in match quality (match effect) and firm fixed effects 16 per cent.

The literature has also focused on looking into whether or not there is a positive correlation between worker and firm fixed effects (i.e. positive assortative matching). The expected result being that high wage workers sort into high wage firms. However, numerous studies (Goux and Maurin, 1999; Barth and Dale-Olsen, 2003; Andrews et al., 2008; and Grutter and Lalive, 2004) have found evidence of a small negative correlation between firm and worker effects implying that high wage workers tend to sort into low wage firms. Our paper also finds a negative correlation between worker and firm fixed effects in both the wage levels and wage growth analysis. This introduced a puzzle in the literature which has been explained by Abowd et al. (2004) and Andrews et al. (2008) as limited mobility bias. The bias is larger the fewer movers you have in the sample. However, Andrews et al. (2008) also finds that while the bias can be considerable it is not large enough to remove the negative correlation entirely.

[5] All manufacturing jobs

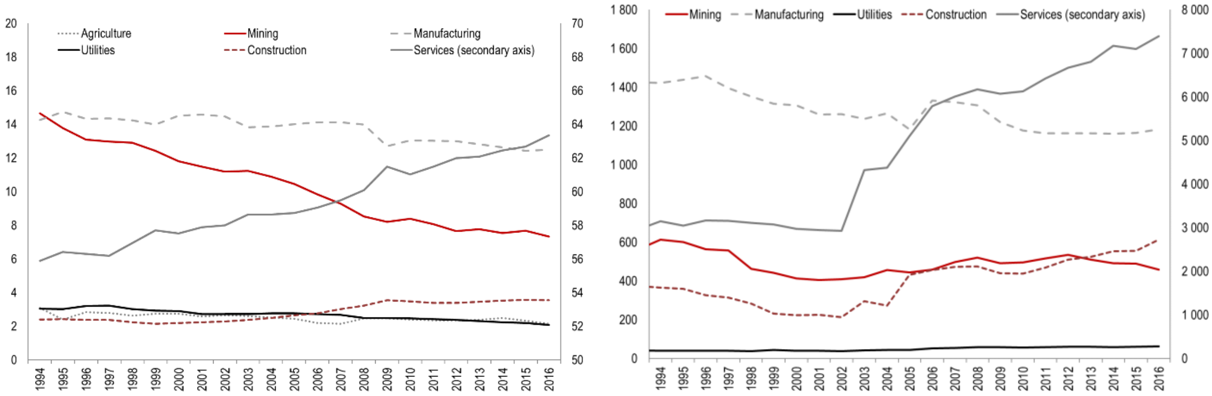
Cornelißen and Hübler (2011) provide a different explanation for the negative correlation. They find that among large German firms (more than 1 000 employees), low wage firms tend to be stable firms and high wage workers are more stable workers which increases their incentive to choose stable low wage firms. However, this does not hold for small firms. This paper finds negative assortative matching for small (less than 500 employees) and positive assortative matching for medium and large firms (500 or more employees). The negative assortative matching could reflect the increasing productivity in the manufacturing sector at the expense of relatively low wage jobs.

The paper is structured as follows. Section 2 provides the background for the manufacturing sector. Section 3 explains the data used and the descriptive statistics. Section 4 presents the methodology applied. Section 5 details the empirical results and Section 6 concludes.

## 2 Background – Manufacturing trends and structural change

The South African manufacturing sector has undergone significant change over the last three decades. It is important to understand the evolution of the sector to provide the context for interpreting our results. The sectors contribution to GDP has gradually declined from 14.4 per cent in 1994 to 12.5 per cent in 2016 and employment decreased from 1.42 million in 1994 to 1.18 million in 2016. This is in part due to the structural change in the economy which is becoming more services orientated shift away from the primary and secondary sectors (see Figure 1).

*Figure 1: Share of GDP (%) and Formal employment ('000) by economic activity*




Source: Reyes et. al. (2019) using Statistics South Africa (1st panel) and Statistics South Africa historical QLFS series (2nd panel) between 1994 and 2016. Note that long term employment in agriculture is not provided in this long term data series.

Despite the declining contribution to GDP and employment the manufacturing sector is becoming more productive. Using the SARS-NT panel Kreuser and Newman (2018) found that productivity in the manufacturing sector grew on average between 2010 and 2013. However, there is significant heterogeneity in productivity within and between the subsectors. Further, TFP increases with firm size and TFP growth is driven by the larger firms in terms of number of employees.

Gabriel (2016) also found that manufacturing sectors display positive productivity growth but negative employment growth using a new social accounting matrix (SAM). Reyes et al. (2019) corroborates this result using the SARS-NT panel and show that although firms in South Africa (incl. manufacturing) have increased their productivity it has not translated into growing employment. Even manufacturing value added has increased as the expense of labour.

The trend that is emerging is one of a sector that is becoming more productive while shedding jobs i.e. the sector is becoming more capital intensive. Earlier research on the impact of trade on domestic production which led to strong growth in capital-intensive exports and import penetration in ultra-labour-intensive sectors indicated the shift towards more capital-intensive production in South Africa, especially in manufacturing industries (Edwards, 2001).

Fedderke et al. (2018) also used the SARS-NT panel 2010-2012 and found that manufacturing firm exit rate exceeded the entry rate consistent with the rising average concentration of the manufacturing sector. They also found significant entry and exit of small firms, with exit dominating the flow. This is consistent with our finding of higher job exits compared to entry from 2014.

Bhorat and Rooney (2017) use the LFS and QLFS between 2001 and 2014 to analyse employment growth in the manufacturing sector and a shift in the composition of skills in the sector. Managers (skilled) and elementary workers (unskilled) increased while clerks, operators and assemblers (semi-skilled) declined.

All these trends taken together indicate that the South African manufacturing sector becoming more productive, capital intensive and concentrated while shedding jobs which are likely to be semi-skilled and increasing managerial positions.

# 3 Data and descriptive statistics

The South African Revenue Service and National Treasury tax administrative panel dataset (SARS-NT panel) from 2011 to 2016 is used to conduct this research. The panel dataset consists of four data sources the Personal Income Tax employee returns (IRP5 and IT3a), Company Income Tax data (CIT) for registered companies (IT14 and ITR14), Transactions records from traders (Customs) and Value-Added Tax records (VAT). This paper mainly uses the IRP5 component of the data. Individuals are identified by their unique identity numbers.

Firms are identified by a unique Pay As You Earn (PAYE) number and CIT number. A firm can have multiple PAYE numbers which can be matched to a unique CIT number, however, the reverse is not possible. As such, the PAYE number represents multiple branches of the same organisation. In this study we use the PAYE number as the firm identifier. This implies that workers switching jobs between branches in an organisation will be considered movers.

The IRP5 form is submitted by employers and includes information on the total employee tax amount, taxable income, retirement fund income, gender, date of birth, employment start and end date and industry. An employer issues an IRP5 certificate to each employee if remuneration is paid and tax on that remuneration has been deducted. If no tax has been deducted, and the employee remuneration is equal or greater than R2 000 per year, then an IT3a certificate is issued. More importantly, firms with workers that only earn below the R2 000 per year are not taxed and an IRP5 or an IT3a form are not issued, as such, these employees are not captured in the panel. Resultantly, any analysis of wages does not cover the lower end of the wage distribution (Pieterse et al., 2018).

We start the data cleaning process by dropping duplicate certificates, these are likely to be revisions. We exclude labour brokers from the analysis. We then generate total earnings per job by summing gross non-retirement fund income (the salary paid to an individual from which contributions to medical aid and UIF are deducted), non-taxable income (which includes arbitration awards, purchased annuities, travel reimbursements, subsistence allowances, uniform allowances, and other allowances) and gross retirement income (or pension contributions). Earnings are converted to monthly and are deflated using the economy wide consumer price index (CPI). We drop outliers by trimming the top percentile of wages which deleted approximately 5 million observations.

To calculate employment duration for each job we take the absolute value of the difference between the start and end date. We drop observations with employment duration greater than 30000 days and those with missing duration. We keep all the manufacturing sector jobs held by an individual i.e. only jobs which were in the manufacturing sector. This allows us to have more connected firms as firms with the same worker in them can be compared in the fixed effect estimation. A mover/switcher is an individual employed in firm x in time t and moves to another firm in either in time t or t+1. Individuals who enter the job market for the first time are called entrants.

The IRP5 form does not include firm size, thus firm size is calculated as the number of employees reported as employed in a firm in a given year weighted by the number of months employed. We also use the gender and date of birth variables in the analysis. Age is restricted between 19 and 65 years.

The sample is restricted to only include the manufacturing sector[6] which accounts for a 15% share of employment in the panel. The largest sector in the panel is the financial services sector which accounts for 28. 37% of the formal jobs in South Africa, followed by government (16.32%) and then manufacturing.

*Table 1: Distribution of firms by sector*

| Sector | Share of employment |
|---|---|
| Agriculture | 7.88 |
| Mining | 3.67 |
| **Manufacturing** | **15.13** |
| Utilities | 1.01 |
| Construction | 3.11 |
| Trade | 10.93 |
| Transport & communication | 3.55 |
| Tourism | 2.59 |
| Financial services | 28.37 |
| Government | 16.32 |
| Non-government community services | 6.97 |
| Not classified | 0.46 |

Source: SARS-NT panel (own calculations)

The number of workers in the sample ranges from 1.4 million in 2011 to 1.8 million in 2016 with 75 per cent males and 25 per cent females employed in the sector. There are more jobs than workers as some workers hold multiple jobs. Most workers hold one job between 2011 and 2016. The number of firms in the sample gradually increases from 41 714 in 2011 to 44 756 in 2016[7].

[6] This is due to estimation limitations discussed in the methodology section.
[7] This is not an indication of pure firm entry it also captures expansion of branches within an organisation.

Table 2: Key summary statistics

| | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 |
|---|---|---|---|---|---|---|
| **Number of workers** | 1,489,280 | 1,612,539 | 1,730,818 | 1,826,034 | 1,842,098 | 1,827,503 |
| **Number of firms** | 41,714 | 43,016 | 45,927 | 49,594 | 48,822 | 44,756 |
| **Number of jobs** | 1,537,289 | 1,670,940 | 1,796,050 | 1,910,703 | 1,937,138 | 1,896,391 |
| **Number of females** | 502,391 | 543,652 | 586,065 | 622,359 | 631,844 | 631,481 |
| **Number of males** | 986,898 | 1,068,901 | 1,144,770 | 1,203,691 | 1,210,271 | 1,196,041 |
| **Mean real monthly wage** | 14635.23 | 14615.21 | 14729.07 | 14494.37 | 14777.46 | 14881.71 |
| **Median real monthly wage** | 7121.04 | 7214.87 | 7245.57 | 7314.38 | 7428.92 | 7616.86 |
| **Mean worker age** | 37 | 37 | 37 | 37 | 37 | 38 |
| **Median worker age** | 35 | 35 | 35 | 35 | 35 | 35 |
| **Mean number of jobs per worker** | 1.07 | 1.08 | 1.08 | 1.1 | 1.1 | 1.1 |
| **Mean number of workers per firm** | 1466 | 1501 | 1472 | 1503 | 1508 | 1634 |
| **Median number of workers per firm** | 175 | 185 | 190 | 191 | 198 | 217 |

Source: SARS-NT panel (own calculations)

The average and median wage both increase over time, however, the mean wage is double the median wage. There are very large firms that skew the distribution of firms evidenced by the large difference between the median and the mean. The median manufacturing firm in the sample has approximately 200 workers while the mean firm has around 1 000 workers.

The firm size distribution shows that most manucturing firms have between 100 to 499 employees. However, firms with bewtween 1 000 and 49 999 employees disportionaltely have the largest share of employment.

Table 3: Firm size and employment share

| **Firm size** | **Proportion** | **Employment share** | | | | | |
|---|---|---|---|---|---|---|---|
| | 2011-2016 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 |
| 1 - 9 | 5.3 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 |
| 10 - 49 | 21.0 | 0.4 | 0.4 | 0.4 | 0.4 | 0.4 | 0.4 |
| 50 - 99 | 11.9 | 0.6 | 0.6 | 0.6 | 0.6 | 0.6 | 0.5 |
| 100 - 499 | 27.9 | 4.5 | 4.4 | 4.8 | 4.6 | 4.6 | 4.22 |
| 500 - 999 | 9.8 | 4.7 | 4.6 | 4.8 | 4.6 | 4.7 | 4.5 |
| 1000 - 49999 | 24.1 | 89,70 | 90,00 | 89,70 | 89,80 | 89,70 | 90,40 |
| Total | 100 | 100 | 100 | 100 | 100 | 100 | 100 |

Source: SARS-NT panel (own calculations)

The wages in this paper are reported as monthly wages and log monthly wages are used for most of the analysis. Wage inequality is high in South Africa evidenced by the distribution of wages even at the top end of the wage distribution. The income of the $99^{th}$ percentile is almost double that of the $95^{th}$ percentile see Table 4.

Table 4: Monthly real wage percentiles

| **Percentile** | **Monthly wage** |
|---|---|
| 25 | R3 782.11 |
| 50 | R7 331.89 |
| 75 | R16 069.04 |
| 90 | R34 241.95 |
| 95 | R50 843.16 |
| 99 | R100 024.10 |

Source: SARS-NT panel (own calculations)

Roughly 60 per cent of the workers stay in the same job and 40 per cent get a new job between 2012[8] and 2016. Same job refers to a unique job was held for two consecutive years. The number of jobs held for a minimum of two periods has increased between 2012 and 2016. New job refers to a job that is held by an individual in year t and was not held in year t-1. New jobs increased from 2012 until 2014 and have gradually declined in 2015 and 2016. Job entry and exit[9] are done at the individual level so it captures an individual that is entering the panel for the first time or leaving the panel[10]. The number of individuals entering the manufacturing sector has gradually declined over time and the number of individuals exiting has increased. There is more exit than entry in the sample from 2014.

Table 5: Same job, new job, entry and exit

| | 2012 | 2013 | 2014 | 2015 | 2016 |
|---|---|---|---|---|---|
| Same job | 1,183,066 | 1,286,133 | 1,390,939 | 1,426,445 | 1,445,616 |
| New job | 487,874 | 509,917 | 519,764 | 510,693 | 450,775 |
| Entry | 417,146 | 430,364 | 417,750 | 378,225 | 367,484 |
| Exit | 325,432 | 354,242 | 435,425 | 585,103 | - |

Source: SARS-NT panel (own calculations)

Most workers in the sample stay in the same job. The transition matrix for 2014 shows us that 96.39 per cent of the workers that were in the same job in 2013 were still in the same job in 2014 and only 3.61 per cent were in a new job in 2014. Among workers that had a new job in 2013, 83.21 per cent were still in that job in 2014 and 16.79 per cent were in another new job in 2014. These proportions are relatively the similar across all the years in the panel.

Table 6: Job level transitions

<table>
<tr><td></td><td></td><td colspan="2">2014</td></tr>
<tr><td></td><td></td><td>Same job</td><td>New job</td></tr>
<tr><td rowspan="2">2013</td><td>Same job</td><td>96.39</td><td>3.61</td></tr>
<tr><td>New Job</td><td>83.21</td><td>16.79</td></tr>
</table>

[8] 2011 is the first year of the panel as such all the workers are entrants in that year with new jobs.
[9] There is no exit in 2016 because it is the last year of the panel
[10] It is possible to leave a job but not the panel if you had multiple jobs.

Source: SARS-NT panel (own calculations)

We divide the sample into wage quantiles creating four categories where the first quantile represents individuals with low wages and the fourth quantile individuals with high wages. We find that most of the new jobs are among the low wage workers and most of the individuals staying the same jobs are high medium and high wage workers. When looking at individuals who had a new job in the previous year and a new job in the following year within the manufacturing sector we find that higher wage workers make up a larger proportion of these churners. This could be due to low income churners mostly exiting the sector which we discuss later in the paper (see figure 3 below). As such, more high wage churners remain in the manufacturing allowing us to track their movement within the sector.

Table 7: Wage quantiles in new job, same job and churners

| | Low wage | Low medium wage | High medium wage | High wage |
|---|---|---|---|---|
| Same job | 18.46 | 24.69 | 29.41 | 29.41 |
| New job | 35.96 | 25.53 | 20.91 | 17.60 |
| Churners | 21.13 | 24.45 | 26.20 | 28.22 |

Source: SARS-NT panel (own calculations)

Off the workers in low wages in 2013, 82.86% of them remained in low wages and 14.95% transitioned to low medium wages. This transition matrix shows us that workers are likely to stay in the same wage quantile which shows much less mobility compared to the findings of Vermaak (2011) and Finn et al. (2012) whose studies sampled mainly the extreme low end of the wage distribution and informal employment which are discussed below.

Table 8: Wage quantile transitions

| | | 2014 | | | |
|---|---|---|---|---|---|
| | | Low wage | Low medium wage | High medium wage | High wage |
| 2013 | Low wage | 82.86 | 14.95 | 1.56 | 0.62 |
| | Low medium wage | 9.30 | 77.04 | 12.82 | 0.84 |
| | High medium wage | 1.80 | 7.10 | 82.53 | 8.57 |
| | High wage | 0.96 | 0.9 | 5.02 | 93.11 |

Source: SARS-NT panel (own calculations)

Prior to the availability of the SARS-NT panel, the NIDS, KIDS and LFS panels were used for analysing income or wage mobility. Cichello et al. (2003) analysed earnings dynamics among Africans in KwaZulu-Natal between 1993 and 1998 using the KwaZulu-Natal Income Dynamics Study. They found that working aged Africans in KZN experienced large gains in earnings between 1993 and 1998. Further, they found that while obtaining formal employment was an important pathway to growth in earnings, they also found that most of those who ‘got ahead’ did so by remaining in the same sector.

Vermaak (2010) used six waves of the South Africa’s labour force panel (2001 to 2004) to assess low wage mobility and found that low wage workers who maintain employment are more likely to experience upward than downward earnings mobility. The aggregate transition matrices in their paper show that the probability of an individual earning less than R800 remaining in the same earnings category is less than 50% and the probability of people earning more than R800 remaining in that earnings category is 91.45% (see table 9). This reflects a high level of mobility in low wages (i.e. less than R800). The greater than R800 category has a very high range and does not adequately capture the earnings mobility of the workers in it.

Table 9: Aggregate transition patterns between earnings categories for all employed individuals

| | | t | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | <R150 | R150-R299 | R300-R499 | R500-R799 | R800+ | Total | % of t-1 |
| | <R150 | **32.64** | 28.56 | 17.28 | 10.48 | 11.05 | 100 | 2.78 |
| | R150-R299 | 10.81 | **43.33** | 22.84 | 10.76 | 12.26 | 100 | 7.00 |
| t-1 | R300-R499 | 4.84 | 14.53 | **42.70** | 21.15 | 16.79 | 100 | 9.96 |
| | R500-R799 | 2.30 | 5.93 | 15.41 | **46.57** | 29.80 | 100 | 11.56 |
| | At least R800 | 0.38 | 1.16 | 2.28 | 4.73 | **91.45** | 100 | 68.70 |

Source: Vermaak (2010) using LFS panel, Sep 2001 to March 2004.

Finn et al. (2012) also analysed income mobility in South African, however, they used the first two waves of the National Income Dynamics Survey (NIDS). They found that individuals earning less than R515 and those earning more than R1898 displayed the least mobility and those in the middle had the most mobility (see table 10).

Table 10: Transitions across earnings across the first two NIDS waves

| | Wave 2 | | | | |
|---|---|---|---|---|---|
| Wave 1 | | | | | |
| | <515 | 515-948 | 949-1898 | >1898 | |
| <515 | 70% | 20 | 7 | 3 | 100% |
| 515-948 | 41 | 31% | 21 | 7 | 100% |
| 949-1898 | 20 | 22 | 35% | 22 | 100% |
| >1898 | 5 | 5 | 14 | 76% | 100% |

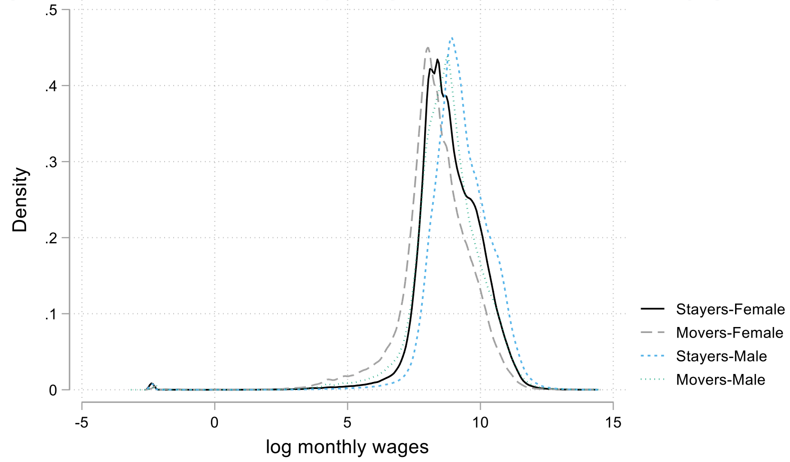
Source: Finn et al. (2012) using NIDS wave 1 and 2

The South African studies above were done on individual level surveys which oversample the lower end of the wage distribution and the informal sector. Whereas in this paper we cover formal manufacturing sector for individuals earning more than R2000 per year. Resultantly, we find less income mobility across compared to these studies.

We then looked at the kdensity of monthly wages for males and females who have the same job (stayers) and those with new jobs (movers) in figure 2 above. For both males and females, the movers’ distribution is slightly to the left of the stayers, indicating higher wages of stayers.

We also see that there is a higher density of females in the lower wage segment who are movers. Further, the male movers and stayers distributions are slightly to the left of the female movers and stayers distributions, showing that male workers earn slightly more than females. We already know that most of the workers changing jobs are low wage workers (see Table 7 above). However, are these low wage workers merely changing jobs or are they also changing industries?

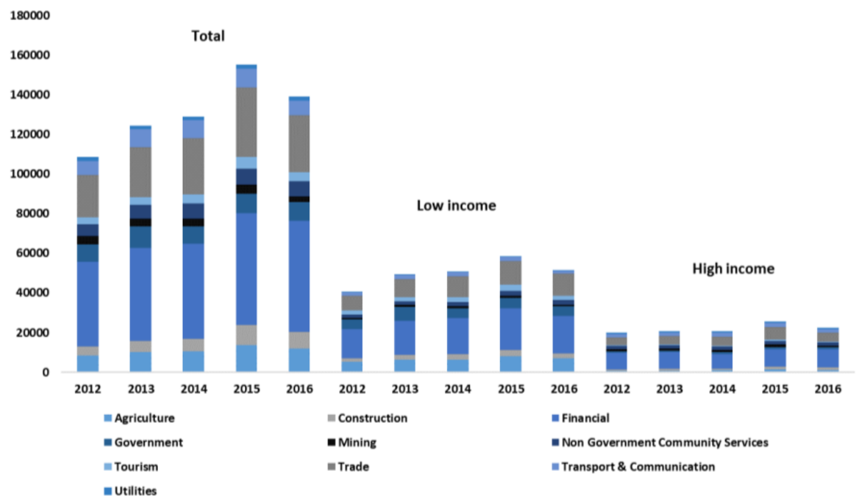
Figure 2: Individual monthly wage movers vs stayers by gender




Source: SARS-NT panel (own calculations)

Beyond changing jobs there are a number of workers that change industries and move out of the manufacturing sector. Most of these individuals move to the financial services and trade sectors. Figure 3 also shows that most of the workers leaving the sector are low wage workers (i.e. the bottom quantile). Thus, we see that the manufacturing sector is mostly losing low income workers to other sectors. This could be a potential explanation of why we find that most workers changing jobs continuously are high wage workers. Most of the low income churners could be leaving the sector as the industry sheds low income jobs.

Figure 3: Individuals moving from the manufacturing sector to other industries




Source: SARS-NT panel (own calculations)

# 3 Methodology

The paper follows the two-way fixed effects estimator of Abowd et al (1999). Closely following Sorensen and Vejlin (2011) and Cornelißen and Hübler (2011). The aim for the analysis is to estimate the determinants of wage levels and wage growth along with unobserved individual and firm effects (which maybe be correlated with the observables). We then look at the correlation of the estimated components to see their relative importance in the determination of wages and wage growth.

We estimate the following fixed effect model:

$$Y_{it} = \beta' x_{it} + \theta_i + \psi_{j(i,t)} + \varepsilon_{it} \tag{1}$$

Where:

- i denotes a work (i=1….N) at time t (t=1….T) in firm j (j=1.....J);
- $Y_{it}$ is the log wage/ $\Delta Y_{it}$ log change in wage (wage growth)
- $x_{it}$ is a vector with K observable characteristics i.e. age, age squared, age cubed, gender and time effects (a full set of year dummy variables).
- $\beta$ is the co-efficient that captures the effects of observed time-varying worker and firm characteristics (including time effects);
- $\theta_i$ is the worker fixed effects;
- $\psi_{j(i,t)}$ is the firm fixed effects; and
- $\varepsilon_{it}$ is the idiosyncratic error term.

Our main interest is the estimated unobservable worker $\theta_i$ and firm $\psi_j$ fixed effects. The underlying assumption is that error term orthogonal to all regressors and to the worker and fixed effects (Abowd, 1999)

$$E[\varepsilon_{it} | x_{it}, i, j(i,t)] = 0. \tag{2}$$

This implies strict exogeneity which means that workers' mobility decisions are independent of $\varepsilon_{it}$. This rules out endogenous mobility which is an important consideration, however, in this paper we are interested in measuring the relative role of firm and worker heterogeneity in wages and wage growth.

In matrix notation we have

$$Y = \boldsymbol{X}\beta + \mathbf{D}\theta + \boldsymbol{F}\psi + \varepsilon \tag{3}$$

where X is a N* x K matrix of observable covariates, D is a N* x N matrix of worker dummy variables for the worker effect, F is a N* x J matrix of firm dummy variables for the firm effect, Y (wages) and $\varepsilon$ (residual/error) are N* x 1 vectors, and N*=NT. To estimate equation 3 we need to compute N worker effects and J firm effects, N is usually in millions and J in thousands making the estimation unfeasible with standard estimation methods (Sorensen and Vejlin, 2011). We use the grouping algorithm (conjugate gradient) from Abowd et al. (2002) which is applied in the felsdvreg command by Conelissen (2008) based on the fixed effects model and least squared dummy variable model (FEiLSDVj). The felsdvreg command is used to estimate equation 3. The time that it takes to run this command and memory required dictated our sample selection. Thus, we could not look at multiple sectors and chose to restrict the sample to only the manufacturing sector.

When estimating individual and firm fixed effects it should be noted that you will get a worker effect for all workers but not all firms. Only firms connected by worker mobility will have an estimate i.e. only connected groups of firms which have had workers move to another firm can have fixed effects estimated. For example, there are many firms across the country that will never be connected because workers in certain areas will never move to jobs in other parts of the country. As mentioned above Abowd et al. (2002) provide an algorithm to determine these groups. In each group one firm is used as the reference firm and does not get a fixed effect estimate calculated for it. For the analysis the sample is divided by male and female and then males and females are further divided into the first and fourth quantile (as a proxy for low income and high income workers), and lastly by firm size.

In the full sample (i.e. all manufacturing firms over all the years in the panel) there are 56 693 unique firms and 3 229 064 individuals. Most of the workers in the panel only appear once and approximately 20% are seen 6 times in the sample and 4% are seen more 7 times or more (see table 11). Workers that appear less than 3 times in the sample account for 46.58%.

Table 11: Number of observations per worker

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 or more |
|---|---|---|---|---|---|---|---|
| Frequency | 912,095 | 591,946 | 386,956 | 306,175 | 272,059 | 632,728 | 127,104 |
| % | 28.25 | 18.33 | 11.98 | 9.48 | 8.43 | 19.59 | 3.94 |

Source: SARS-NT panel (own calculations)

There is very little movement in our sample with only 15.8% of the workers having more than 2 employers in the sample. This is much lower than the movement in French firms (50%) from Abowd et al. (1999) and Sorensen and Vejlin (2011) for Danish firms (66%). However, our panel only has 6 years of observations and Abowd et al. (1999) had 10 years and Sorensen and Vejlin (2011) had 26 years. Further, our results only apply to the manufacturing sector.

Table 12: Number of firms a worker is employed in

| | 1 | 2 | 3 | 4 or more |
|---|---|---|---|---|
| Frequency | 2 720 030 | 412 725 | 73 973 | 22 335 |
| % | 84.24 | 12.78 | 2.29 | 0.69 |

Source: SARS-NT panel (own calculations)

What is important to note is that the estimation and analysis are only conducted on the largest connected group of firms is used. It is possible to look at the correlations across all the groups but this requires certain assumptions to be made for the interpretation to be meaningful. As such, we only look at the largest connected group of firms in each sub sample for the analysis and the sample sizes are given in table 13 and 14 below. A summary of the total number of workers, firms, fixed effects estimated and groups is provided in the Table 1 and 2 in the Annexure for every sub sample.

The largest connected group in the wage level analysis in the sample contains approximately 3.2 million workers in 44 162 firms which are connected by 508 449 movers. The sample for wage growth is smaller with around 2.1 million workers in 32 975 firms connected by 203 109 movers. We will refer to these groups as the full sample.

Table 13: Sample sizes using the largest connected group

| | Wage levels | | | Wage growth | | |
|---|---|---|---|---|---|---|
| | Number of workers | Number of firms | Number of movers | Number of workers | Number of firms | Number of movers |
| Total | 3 175 896 | 44 162 | 508 449 | 2 090 977 | 32 975 | 203 109 |
| Female | 1 076 228 | 27 821 | 160 017 | 636 424 | 15 938 | 58 844 |
| High Income (q4) | 215 646 | 9 832 | 32 336 | 161 648 | 5 646 | 16 224 |
| Low income (q1) | 422 862 | 6 539 | 31 935 | 135,468 | 1 781 | 4 705 |
| Male | 2 035 063 | 40 029 | 346 632 | 1 350 526 | 28 547 | 141 063 |
| High Income (q4) | 431 883 | 14 729 | 65 979 | 346 001 | 9 293 | 35 940 |
| Low income (q1) | 825 341 | 20 248 | 79 508 | 282 956 | 6 413 | 13 378 |
| Firm size | | | | | | |
| 10 – 49 | 729 904 | 21 326 | 54 524 | 225 414 | 8 460 | 12 557 |
| 50 - 99 | 550 890 | 5 714 | 26 146 | 200 898 | 2 875 | 5 981 |
| 100 - 499 | 1 149 271 | 3 921 | 90 753 | 709 513 | 3 568 | 32 265 |
| 500 - 999 | 451 057 | 474 | 18 313 | 273 314 | 417 | 6 977 |
| 1000 - 49999 | 882 214 | 248 | 46 354 | 583 511 | 239 | 20 902 |

Source: SARS-NT panel (own calculations)

# 4 Results

The aim of our analysis is to estimate the worker and firm fixed effects and to see the relative importance of each component's contribution to the explanation of the variance in the dependent variables. We do this for both wage growth and wage levels to compare the models. Thus the variance decomposition is given as follows:

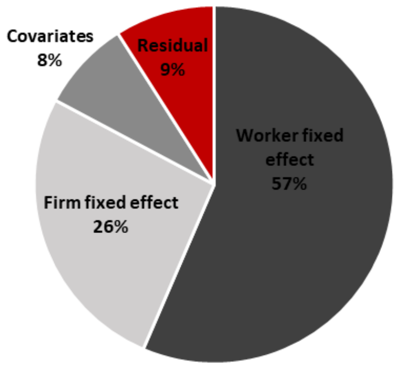
$$Var(Y_{it}) = Cov\big(Y_{it}, \beta' x_{it} + \theta_i + \psi_{j(i,t)} + \varepsilon_{it}\big) = Cov(Y_{it}, \beta' x_{it}) + Cov(Y_{it}, \theta_i) + Cov(Y_{it}, \psi_{j(i,t)}) + Cov(Y_{it}, \varepsilon_{it}) \quad (4)$$

We then divide through by the variance of the wages or the change in wages

$$\frac{Cov(Y_{it}, \beta' x_{it})}{Var(Y_{it})} + \frac{Cov(Y_{it}, \theta_i)}{Var(Y_{it})} + \frac{Cov(Y_{it}, \psi_{j(i,t)})}{Var(Y_{it})} + \frac{Cov(Y_{it}, \varepsilon_{it})}{Var(Y_{it})} = 1 \quad (5)$$

Abowd et al. (1999) and Sorensen and Vejlin (2011) make a decomposition similar to this one and determine that worker effects are more important than firm effects in explaining the variation in wages and wage growth. The estimation results and their variance decomposition are shown on Table 3 and 4 in the annexure. Starting with the full sample on the wage levels regressions which is the standard Abowd (1999) model we find that worker fixed effects dominate the firm fixed effects and explain 57 per cent of the variance in wages while firm fixed effects explain 26 per cent of the variation in wages, the observable characteristics/covariates 8 per cent and the residual 9 per cent. The worker effect estimates are similar to Sorensen and Vejlin (2011) who found 58 per cent, however, firm effects only explained 14 per cent of the variation in Danish wages and 26 per cent for South African manufacturing firms fixed effects. Jinkins and Morin (2018) estimated higher worker fixed effects of 78 per cent and firm effects of 12 per cent for Danish wages. Although the worker fixed effects dominate the firm fixed effects which is similar to the findings in the literature, the firm fixed effects in manufacturing companies in South Africa are higher than those in the literature indicating that these firms play are larger role in explaining the variation in wage levels. It should be noted that the studies in literature estimate these effects across all sectors whereas this paper only looks at manufacturing.

Figure 4: Wage level variance decomposition for the full sample

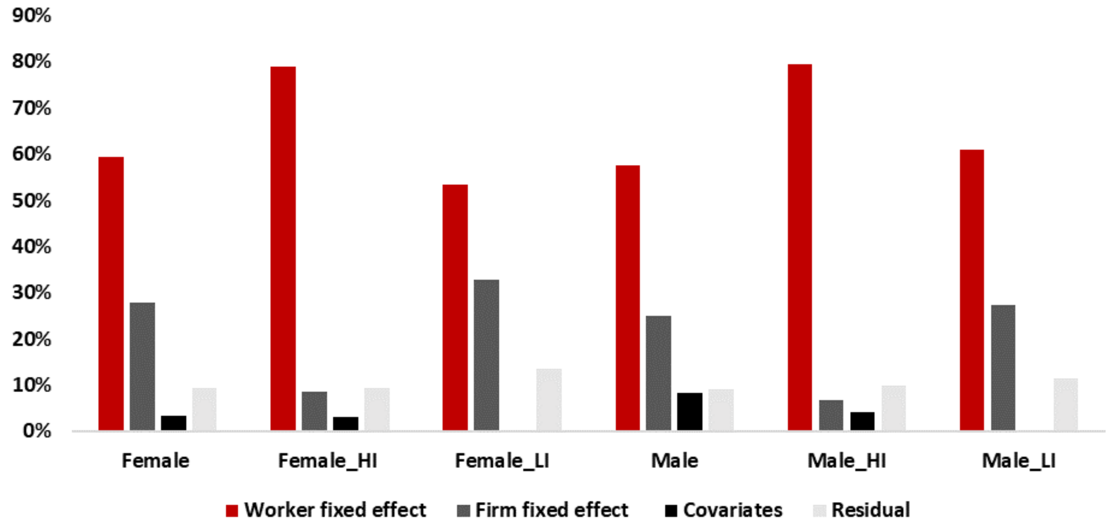



Source: SARS-NT panel (own calculations)

The results for the full sample are very similar to those for males and females which reflects that gender is not the main source of heterogeneity. However, both worker and firm effects explain slightly more of the variation in female wages. When we divide the gender sub samples by wage/income quartiles and examine the first and last quartiles as low income (LI) and high income (HI) respectively, we find a lot more heterogeneity. Education level would have been a more appropriate indicator to use, however, education is not reported in the data. Thus, we use the income quantile a worker falls under to compare workers in the absence of education levels. We find that for low income employees' worker fixed effects have a lower explanatory power compared to high income workers effects[11] and their firm fixed effects have a higher explanatory power compared to high income employees. This essentially means a larger proportion of the low income employees' wages are explained by the firm they are employed in and their personal attributes determine less of their wages compared to high income employees. For low income females their worker effects explain 53 per cent of the variation in their wages while it explains 79 per cent for high income females. For low income males their firm effects explain 27 per cent of the variation in their wages and only 7 per cent for high income males. These trends are similar for both males and females and are presented in figure 5 below.

This finding makes sense intuitively and is what we would expect. Low income workers have lower bargaining power and are more likely to take the wage offered to them by a firm. However, high income workers are likely to have a higher skill level giving them more leverage to negotiate a higher wage regardless of the firm.

Figure 5: Wage level variance decomposition across gender and income quantiles



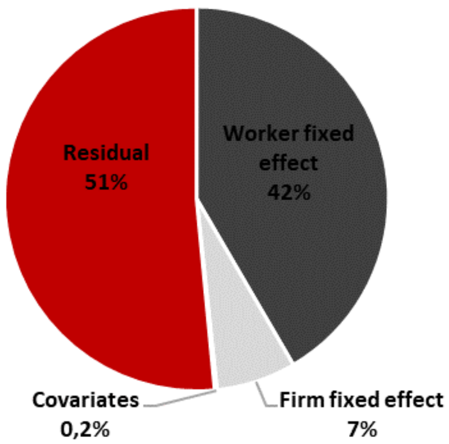

Source: SARS-NT panel (own calculations)

We now look at the wage growth estimations and variance decomposition. Notably, the error term explains a lot more of the variation in wage growth indicating that we know less about

[11] A lot of individual level characteristics are not available in the data and characteristics like education which are unlikely to change once an individual enters the labour market will fall into the fixed effect estimates along with race, worker ability and motivation.

what explains the variation in wage growth. Sorensen and Vejlin (2011) find that the residual explains 85 per cent of the variation in wage growth and Jinkins and Morin (2018) find a residual of 72 per cent. We find that the residual only explains 51 per cent of the variation in wages which is lower than previous studies. What is interesting is that our worker effects explain 42 per cent of the variation in wages which is much higher than the Sorensen and Vejlin (2011) estimate of 9 per cent. Jinkins and Morin (2018) difference away the worker fixed effect. Our firm fixed effects explain 7 per cent of the variation in wages compared to 4 per cent by Sorensen and Vejlin (2011) and 24 per cent by Jinkins and Morin (2018). This indicates that in the manufacturing sector the worker effect explains a much higher proportion of wage growth resulting in a lower residual.

Figure 6: Wage growth variance decomposition for the full sample

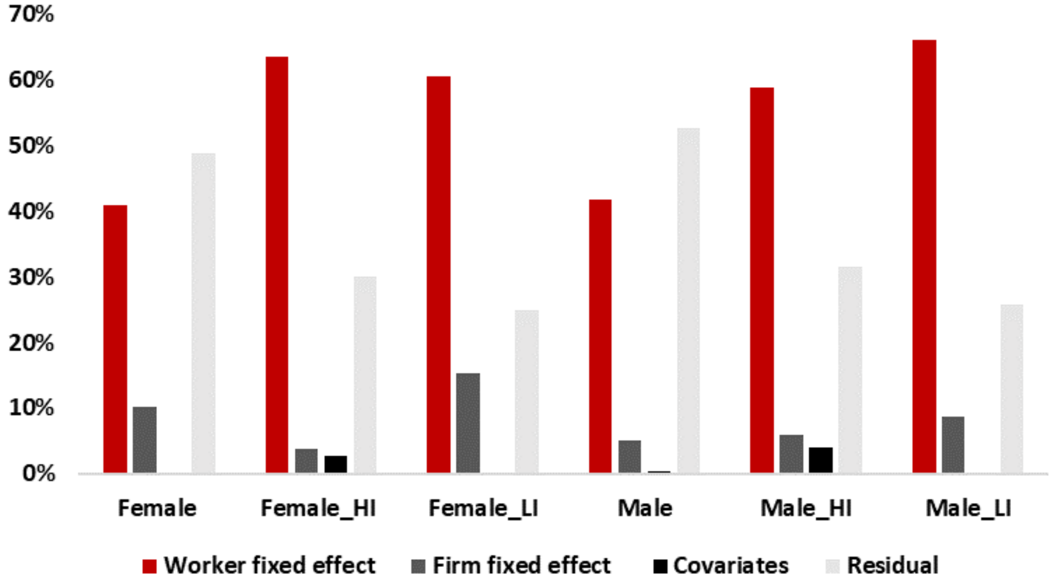



Source: SARS-NT panel (own calculations)

The results for the full sample are similar to those for the male and female sub samples. However, when we divide the gender sub samples by income quantiles we find that worker fixed effects increases and explains on average around 60 per cent of wage growth and the residual falls to around 30 per cent. What is interesting is that the worker effect explains a similar amount of variation in wages regardless of whether you are a high or low income male or female. Notably, for low income males and females their worker effect explains a higher proportion of the variation in wage growth compared to the wage level. This could indicate that low income workers are likely to accept the wage offered (resulting a lower worker effect in the wage level estimation) but once they're in the firm, they gain more bargaining power to negotiate a higher wage which is reflected in the higher worker effect in the wage growth estimation.  Malindi (2016) using nationally representative household panel data from Statistics South Africa, found that found that black workers had much larger wage growth from an additional year of firm tenure than they did from an additional year of labour market experience. The opposite was true for white workers. Their results provide evidence in favour of greater ex ante uncertainty around the expected productivity of black workers and the quality worker-firm matches these workers enter into as the key mechanism behind the relatively larger wage returns firm tenure for black workers because black workers face a greater penalty in the wage returns as a result of the wedge between potential and actual experience. This corroborates our finding of low wage workers having to accept the wage offered by a firm when they enter the firm and the increased ability to earn more when they have experience working in the firm. This is assuming that black workers are a proxy for our low wage quantile in the absence of race in our dataset.

In terms of the firm fixed effects low income females have the highest proportion of variation explained by firm effects at 15 percent while all the sub samples are less than 8 per cent. The observable characteristics/covariates are generally low across the sample.

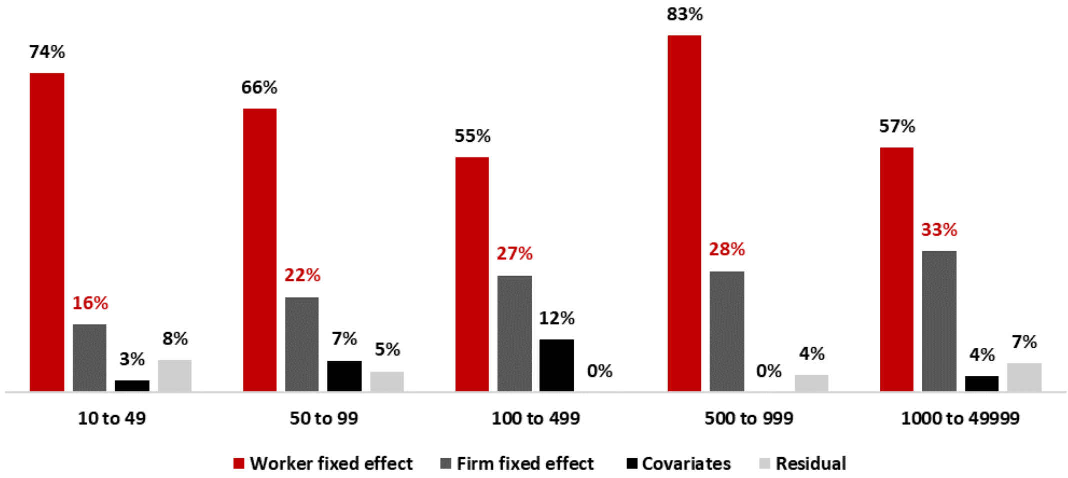
Figure 7: Wage growth variance decomposition across gender and income quantiles




Source: SARS-NT panel (own calculations)

We also divided the sample by firm size using the number of employees in the firm (see table 4 in the annexure). Firms with 1 to 9 employees had too many small connected groups to be able to compute worker and firm fixed effects. As such employees employed in these firms are excluded from this analysis. The wage level regressions also show that worker fixed effects dominate firm fixed effects. However, they reveal an interesting pattern where firm fixed effects estimates increase with firm size (see figure 8). So the larger the firm the more important the firm fixed effect becomes in explaining the variation in wages.

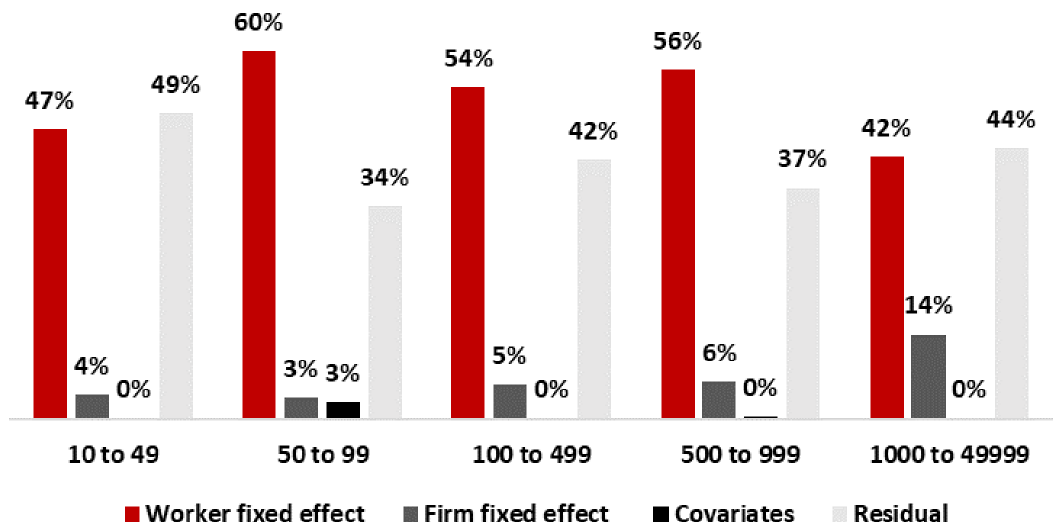
Figure 8: Wage level variance decomposition across firm size




Source: SARS-NT panel (own calculations)

The worker fixed effect for wage growth explains less of the variation in wages compared to wage levels using firm size. Firm fixed effects explain very little of the variation in wage growth ranging from 3 to 6 per cent for all firm sizes with the exception of firms employing 1 000 to 49 999 workers where they explain 14 per cent –these firms have the largest share of employment. Again the residual is much higher in the wage growth estimation compared to the wage levels estimation. However, our residuals are lower than both Sorensen and Vejlin (2011) and Jinkins and Morin (2018).

Figure 9: Wage growth variance decomposition across firm size




Source: SARS-NT panel (own calculations)

We also analyse the correlation of the worker and firm fixed effects on the both wage levels and growth see table 14 and 15. The expected result being that high wage workers sort into high wage firms i.e. positive assortative matching (Abowd et al., 1999). However, like most of the previous literature we find a negative correlation between the firm and work effects, across the full sample and all the gender wage level sub samples. A negative correlation means high income workers are on average sorting into low wage firms. Thus, on an aggregate level we find negative assortative matching. Notably, there is a stronger negative correlation in the wage growth estimations and the correlation is negative across the full sample and all sub samples of wage growth.

Table 14: Worker and firm effects correlation across gender and income quantiles

| | | **Firm effects** | |
|---|---|---|---|
| | | **Wage levels** | **Wage growth** |
| **Worker effects** | Total | -0.0165 | -0.6056 |
| | Female | -0.0809 | -0.6869 |
| | Male | -0.0518 | -0.6587 |
| | Female_LI | -0.2603 | -0.5712 |
| | Female_HI | -0.2439 | -0.7480 |
| | Male_LI | -0.2563 | -0.5566 |
| | Male_HI | -0.3025 | -0.7127 |

Note: (1) Worker fixed effects - $\theta_i$, Firm fixed effects - $\psi_{j(i,t)}$, (2) The correlations are estimated from equation 3.

There are several studies that have also found evidence of a small negative correlation between firm and worker effects in wage levels, introducing a puzzle into the literature which has been explained by Abowd et al. (2004) and Andrews et al. (2008). They suggest that the estimation of the worker and firm fixed effects are carried out with error, thus it is possible that the estimated correlation is biased downward because an over estimation of worker effects can lead to an under estimation of firm effects. As such, the bias is bigger when the data has a fewer movers termed 'limited mobility bias'. However, Andrews et al. (2008) also finds that while the bias can be considerable it is not large enough to remove the negative correlation entirely. While this result could in part be due to limited mobility bias, we can assume based on the Andrews et al. (2008) finding that the bias would not be enough to change the negative sign. Cornelißen and Hübler (2011) provide a different explanation for the negative correlation. They find that among large German firms (more than 1 000 employees), low wage firms tend to be stable firms and high wage workers are more stable workers which increases their incentive to choose stable low wage firms. However, this does not hold for small firms where low wage firms are unstable and high wage firms are stable. As such, it remains puzzling why high-wage workers are matched with small low wage firms.

We also divided our sample by firm size although firms with less than 10 employees are excluded. Focusing on the wage level correlations we find a negative correlation between firm and worker fixed effects in firms less than 500 employees. However, we find that firms with 500 to 999 and 1 000 to 49 999 display positive assortative matching (see Table 15). For simplicity, I will refer to firms employing less than 500 employees as small, firms employing between 500 and 999 employees as medium and firms with more than 1 000 employees as large. Thus, among medium and large South African manufacturing firms high wage workers sort into high wage firms. Unlike Cornelißen and Hübler (2011) we cannot speak to the job stability in these firms given that we do not estimate job duration functions. However, we also find that among small firms high wage workers sort into low wage firms. Given the aggregate trends in the South African manufacturing sector where productivity is increasing at the cost of labour and particularly low wage workers, we could be picking up highly paid owners/managers working in small firms. We find a negative correlation across all firm sizes for wage growth. Thus on aggregate you find a negative correlation which is consistent with the literature but this hides the differences by firm size which shows that firms employing more than 500 workers in South Africa have positive assortative matching.

Table 15: Worker and firm effects correlation across firm size

| | | Firm effects | |
|---|---|---|---|
| | | **Wage levels** | **Wage growth** |
| **Worker effects** | 10 to 49 | -0.4367 | -0.3513 |
| | 50 to 99 | -0.3549 | -0.8090 |
| | 100 to 499 | -0.0014 | -0.5516 |
| | 500 to 999 | 0.1714 | -0.7706 |
| | 1000 to 49999 | 0.0932 | -0.2577 |

Note: (1) Worker fixed effects - $\theta_i$, Firm fixed effects - $\psi_{j(i,t)}$, (2) The correlations are estimated from equation 3.

# 5 Conclusion

The manufacturing sector is gradually evolving over time. There is more job exit than entry in sector and the sector is shedding mostly low wage quantile jobs. While there is a fair amount of movement in the sector workers who stay in the same job tend to have higher wages on average. When we conduct a variance decomposition on wage levels and wage growth we find that a workers individual unobserved characteristics are consistently more important than firm fixed effects in explaining the variation in wages. However, the relative size of the worker and firm fixed effects and their correlation allow us to learn a little bit more about manufacturing wages. We find that low income workers have lower worker fixed effects and higher firm fixed effects compare to high income workers which could reflect that they have lower bargaining power and are more likely to take the wage offered to them by a firm. High income workers are likely to have a higher skill level giving them more leverage to negotiate a higher wage regardless of the firm. However, when we turn to wage growth we find that the worker effects explain a similar proportion of the change in the wage regardless of the income level. This adds another layer to our understanding where we initially found that that low income workers are likely to accept the wage offered (resulting a lower worker effect in the wage level estimation) but once they're in the firm, they gain more bargaining power to negotiate a higher wage which is reflected in the higher worker effect in the wage growth estimation. Disaggregating by firm size also shows an interesting dynamic at the wage level where the firm effect increases with firm size. Thus, the larger the firm an individual works in the more important the firm fixed effect becomes in explaining the variation in wage levels.

Further, we find a negative correlation between firm and worker fixed effects across the full sample. Thus, on an aggregate level we find negative assortative matching so high wage workers sort into low wage firms which is a similar finding with the rest of the literature. However, when we disaggregate by firm size find that only small firms with less than 500 employees have negative assortative matching whereas medium (500 to 999) and large (1 000 to 49 999) have positive assortative matching. Given the general aggregate trends in the South African manufacturing sector where productivity is increasing at the cost of labour and particularly low wage workers, we could be picking up highly paid owners/managers working in small firms. As such, on aggregate you find a negative correlation which is consistent with the literature but this hides the differences by firm size which shows that firms employing more than 500 workers in South Africa have positive assortative matching.

## References

Abowd, J. M., Kramarz, F., & Margolis, D. N. (1999). High Wage Workers and High Wage Firms. *Econometrica*, 67 (2), 251–334.

Abowd, J., Creecy, R. & Kramarz, F. (2002), Computing person and firm effects using linked longitudinal employer-employee data, Technical Paper 2002-06, U.S. Census Bureau.

Abowd, J., Kramarz, F., Lengermann, P. & Perez-Duarte, S. (2004), Are good workers employed by good firms? A test of a simple assortative matching model for France and the United States, mimeo.

Andrews, M. J., Gill, L., Schank, T., & Upward, R. (2008). High Wage Workers and Low Wage Firms: Negative Assortive Matching or Limited Mobility Bias. *Journal of the Royal Statistical Society. Series A (Statsitics in Society)*, 171(3), 673-697.

Aterido, R., Hlatshwayo, A., Pieterse, D., & Steenkamp, A. (2019). Firm Dynamics, Job Outcomes, and Productivity: South African Formal Businesses, 2010–14. Policy Research Working Paper 8788, World Bank Group.

Barthel, A. P., & Borjas, G. J. (1978). Wage Growth and Job Turnover: An Empirical Analysis. NBER Working Paper 285, National Bureau of Economic Research.

Barth, E., Bryson, A., Davis, J. C., & Freeman, R. (2014). It's Where You Work: Increases in Earnings Dispersion across Establishments and Individuals in the U.S. NBER Working Papers 20447, National Bureau of Economic Research.

Bassier, I. (2019). The Wage-setting Power of Firms: Rent-sharing and Monopsony in South Africa. Southern Africa –Towards Inclusive Economic Development (SA-TIED) Working Paper No. 59.

Bhorat, H., Oosthuizen, M., Lilenstein, K., & Steenkamp, F. (2017). Firm-level Determinants of Earnings in the Formal Sector of the South African Labour Market. UNU-WIDER Working Paper 2017/25.

Bhorat, H., & Rooney, C. (2017). State of Manufacturing in South Africa. Development Policy Research Unit Working Paper 201702. DPRU, University of Cape Town.

Card, D., Heining, J. & Kline, P. (2013). Workplace Heterogeneity and the Rise of West German Wage Inequality. *The Quarterly Journal of Economics*, 128 (3), 967–1015.

Cichello, P.L., Fields, G.S., & Leibbrandt, M. (2003). Earnings and Employment Dynamics for Africans in Post-apartheid South Africa: A Panel Study of KwaZulu-Natal. Development Policy Research Unit Working Paper 03/77.

Cornelißen, T. & Hübler, O. (2007). Unobserved Individual and Firm Heterogeneity in Wage and Tenure Functions: Evidence from German Linked Employer-Employee Data. IZA Discussion Paper No. 2741.

Cornelissen, T. (2008). The Stata Command Felsdvreg to Fit a Linear Model with Two High-dimensional Fixed Effects. *Stata Journal*, 8(2), 170.

Edwards, L. (2001). Trade and the Structure of South African Production, 1984–97. *Development Southern Africa*, 18(4), 471–91.

Fedderke, J., Obikili, N., Viegi, N. (2018). Mark-up's and Concentration in South African Manufacturing Sectors: An Analysis with Administrative Data. *South African Journal of Economics*, 86(S1), 120–140.

Finn, A., Leibbrandt, M and Levinsohn, J. (2012). Income Mobility in South Africa: Evidence from the First Two Waves of the National Income Dynamics Study. Cape Town: SALDRU, University of Cape Town. SALDRU Working Paper Number 82/ NIDS Discussion Paper 2012/5.

Gabriel, S. (2016). Productivity Backcasting: Economy-wide Trends, 1993 – 2013. Presentation at UNU/WIDER Conference: Growth and development policy – new data, new approaches, and new evidence, Pretoria, South Africa.

Goux, D. & Maurin, E. (1999). Persistence of inter-industry Wage Differentials: A Re-examination using matched worker-firm panel data. *Journal of Labor Economics*,17, 492–533.

Gruetter, M. & Lalive, R. (2004), The importance of firms in wage determination, Discussion Paper No. 1367, IZA.

Hyslop, D. R., & Maré, D. C. (2009). Job Mobility and Wage Dynamics. Wellington: Statistics New Zealand.

Kreuser, C. F. and Newman, C. (2018). Total Factor Productivity in South African Manufacturing Firms. *South African Journal of Economics*, 86(S1), 40-78.

Malindi, K. (2016). Wage Effects of on-the-job Training for South African Workers. CSAE 2017 conference paper 1142.

Mincer, J. (1986). Wage Changes in Job Changes. NBER Working Paper 1907, National Bureau of Economic Research.

Pieterse, D., Gavin, E. & C. Kreuser, F. (2018). Introduction to the South African Revenue Service and National Treasury firm-level panel, Firm level analysis using administrative record data. *South African Journal of Economics.* 86(S1), 6-39.

Sorensen, K. L. & Vejlin, R. (2011). Worker and Firm Heterogeneity in Wage Growth: An AKM Approach. *Labour*, 25(4), 485-507.

Topel, R.H. & Ward, M. P. (1988). Job Mobility and the Careers of Young Men. NBER Working Paper 2649, National Bureau of Economic Research.

Vermaak, C. (2010). Low Wage Mobility in the South African Labour Market. DPRU and TIPS Conference paper. http://www.tips.org.za/research-archive/annual-forum-papers/2010/item/2851-low-wage-mobility-in-the-south-african-labour-market

# Annexure

Table 1: Summary of wage level groups

| | Number of workers | Number of firms | Number of groups | Number of movers | Number of estimable effects |
|---|---|---|---|---|---|
| Total | 3 229 063 | 56 693 | 293 | 509 033 | 44 477 |
| Female | 1 147 769 | 49 084 | 1 019 | 161 729 | 28 994 |
| High Income (q4) | 257 439 | 29 795 | 1 073 | 33 864 | 11 137 |
| Low income (q1) | 498 468 | 29 598 | 887 | 33 293 | 7 567 |
| Male | 2 081 360 | 53 358 | 380 | 347 257 | 40 448 |
| High Income (q4) | 476 935 | 34 161 | 1 004 | 67 397 | 15 939 |
| Low income (q1) | 907 785 | 41 688 | 822 | 80 726 | 21 221 |
| Firm size | | | | | |
| 10 – 49 | 845 581 | 26 676 | 342 | 55 746 | 21 757 |
| 50 - 99 | 585 532 | 6 163 | 24 | 26 478 | 5 739 |
| 100 - 499 | 1 149 858 | 3 926 | 2 | 90 753 | 3 290 |
| 500 - 999 | 451 713 | 475 | 2 | 18 313 | 473 |
| 1000 - 49999 | 882 214 | 248 | 1 | 46 354 | 247 |

Note: (1) All the numbers presented incl. workers in firm with no movers (2) Where there is only one group it means all the firms in that category had movers (3) Firms employing 1-9 employees had too many small connected groups preventing estimation.

Table 2: Summary of wage growth groups

| | Number of workers | Number of firms | Number of groups | Number of movers | Number of estimable effects |
|---|---|---|---|---|---|
| Total | 2 218 428 | 52 365 | 800 | 204 730 | 33 886 |
| Female | 762 241 | 45 158 | 1 751 | 61 997 | 18 091 |
| High Income (q4) | 211 786 | 26 463 | 1 095 | 17 838 | 6 989 |
| Low income (q1) | 227 445 | 23 195 | 671 | 5 848 | 2 623 |
| Male | 1 456 211 | 49 219 | 925 | 142 715 | 29 618 |
| High Income (q4) | 402 787 | 30 139 | 1 130 | 37 594 | 10 632 |
| Low income (q1) | 433 844 | 35 113 | 1 357 | 15 626 | 8 161 |
| Firm size | | | | | |
| 10 – 49 | 558 006 | 25 081 | 1 877 | 18 705 | 11 897 |
| 50 - 99 | 362 543 | 5 642 | 311 | 7 845 | 3 373 |
| 100 - 499 | 724 105 | 3 751 | 3 | 32 279 | 3 569 |
| 500 - 999 | 284 463 | 448 | 2 | 6 978 | 417 |
| 1000 - 49999 | 583 571 | 240 | 2 | 20 902 | 238 |

Note: (1) All the numbers presented incl. workers in firm with no movers (2) Firms employing 1-9 employees had too many small connected groups preventing estimation.

Table 3: Variance decomposition across gender and income level

| | Wage levels | | | Wage growth | | |
|---|---|---|---|---|---|---|
| Z | Mean | Std.Dev | Cov(lmwage,Z) / Var(lmwage) | Mean | Std.Dev | Cov(lmwage,Z) / Var(lmwage) |
| Total | | | | | | |
| WE | 4.279074 | 0.9098051 | .5636502 | 3.013183 | .5444191 | .41561039 |
| FE | -.5454053 | .6047237 | .26346799 | -1.563472 | .4756485 | .06711732 |
| xb | 5.202235 | .2964134 | .08187553 | -1.408347 | .0460792 | .0019842 |
| Residual | 2.11e-12 | .3615156 | .09100628 | -6.34e-12 | .3208598 | .51528809 |
| Female | | | | | | |
| WE | 5.055535 | .9769017 | .59464339 | 1.316624 | .4650084 | .40967104 |
| FE | -.4040086 | .6920556 | .27920914 | -.0100501 | .380784 | .10219882 |
| xb | 3.993963 | .1884938 | .0333853 | -1.275018 | .0369401 | -.00012981 |
| Residual | -3.85e-12 | .371478 | .09276217 | 3.54e-12 | .3342256 | .48825995 |
| Male | | | | | | |
| WE | 3.324051 | .9007501 | .57623825 | 3.095621 | .5639434 | .41800133 |
| FE | -.493611 | .5902935 | .24894942 | -1.486188 | .4995839 | .05124081 |
| xb | 6.255011 | .2956061 | .08312834 | -1.56314 | .0493539 | .00365311 |
| Residual | 8.30e-13 | .3510003 | .091684 | -5.64e-12 | .312403 | .52710475 |
| Female_LI | | | | | | |
| WE | -.0450538 | .8600272 | .53255056 | 4.389904 | 1.108121 | .60445467 |
| FE | 9.697247 | .7198288 | .32836809 | -.6854293 | .9433984 | .1529519 |
| xb | -2.441613 | .1026202 | .00304749 | -3.90591 | .1140212 | -.00542565 |
| Residual | 1.43e-12 | .3800516 | .13603387 | 3.39e-11 | .41943 | .24801908 |
| Female_HI | | | | | | |
| WE | 5.892048 | .5244799 | .78874489 | 6.497392 | .3710345 | .63507773 |
| FE | .1018932 | .2409453 | .08640835 | .0564558 | .2189575 | .03679894 |
| xb | 4.137607 | .1177874 | .030962 | -6.441823 | .0665064 | .02671059 |
| Residual | 1.12e-12 | .1692655 | .09388476 | 4.17e-13 | .2116686 | .30141275 |
| Male_LI | | | | | | |
| WE | 8.307413 | .9423038 | .60815352 | 4.983574 | 1.178205 | .66125831 |
| FE | -1.393872 | .761804 | .27193878 | -1.117415 | 1.034876 | .08628576 |
| xb | .8086568 | .0473915 | .00451147 | -3.976487 | .110703 | -.00421671 |
| Residual | -4.55e-12 | .3244387 | .11539624 | 1.75e-11 | .3699545 | .25667264 |
| Male_HI | | | | | | |
| WE | 5.92438 | .5241075 | .79271515 | 9.972391 | .3664355 | .58812171 |
| FE | -.0432501 | .2272101 | .06793129 | -.0782852 | .2305753 | .05854471 |
| xb | 4.629385 | .1347216 | .04150022 | -9.787214 | .0963292 | .03915351 |
| Residual | 1.16e-12 | .1716004 | .09785333 | -8.27e-12 | .2148357 | .31418007 |

Note: (1) Worker fixed effects (WE) - $\theta_i$, Firm fixed effects (FE) - $\psi_{j(i,t)}$, Observable characteristics (xb) – β and the Residual/error term - $\varepsilon_{it}$. (2) The coefficients are estimated from equation 3.

Table 4: Variance decomposition by firm size category

| Z | Wage levels Mean | Wage levels Std.Dev | Wage levels Cov(lmwage,Z) / Var(lmwage) | Wage growth Mean | Wage growth Std.Dev | Wage growth Cov(lmwage,Z) / Var(lmwage) |
|---|---|---|---|---|---|---|
| 10 to 49 | | | | | | |
| WE | 4.547939 | 1.044789 | .73921516 | 1.707722 | .7988567 | .46823497 |
| FE | -.4668932 | .6932983 | .1576731 | -.0562203 | .7480621 | .03890064 |
| xb | 4.728364 | .2284815 | .02723293 | -1.620621 | .0597676 | -.0005256 |
| Residual | -2.49e-12 | .2744737 | .07587882 | -3.74e-12 | .2940672 | .49338998 |
| 50 to 99 | | | | | | |
| WE | 5.704469 | 1.006467 | .65559885 | 4.066088 | 1.112345 | .59507071 |
| FE | -.9693455 | .7235649 | .22139861 | -.6665125 | .9176818 | .03439557 |
| xb | 4.117675 | .2572747 | .07417366 | -3.358832 | .5837494 | .02754981 |
| Residual | 1.33e-12 | .2307965 | .04882887 | .2300463 | .2300463 | .3429839 |
| 100 to 499 | | | | | | |
| WE | 6.496987 | .8649979 | .54545779 | .919368 | .5333128 | .53597393 |
| FE | -.4245569 | .5546859 | .27145848 | .1902268 | .3330597 | .05475753 |
| xb | 2.818789 | .3854466 | .1228265 | -1.067164 | .3149918 | -.0101622 |
| Residual | 4.25e-12 | .2803773 | .06025723 | -7.19e-12 | .265416 | .41943074 |
| 500 to 999 | | | | | | |
| WE | -.3859019 | 1.722545 | .8269908 | 1.759353 | .4957767 | .56316186 |
| FE | .1598072 | .6023965 | .28071506 | .0387914 | .4046301 | .05948129 |
| xb | 9.183968 | 1.375412 | -.15245945 | -1.749268 | .056875 | .00485225 |
| Residual | 7.48e-14 | .2524231 | .04475359 | -1.74e-12 | .2457341 | .3725046 |
| 1000 to 49999 | | | | | | |
| WE | .423326 | 1.063505 | .56639156 | 1.65796 | .370474 | .42443129 |
| FE | -.3981927 | .7759925 | .32708876 | -.0792347 | .2434688 | .13662339 |
| xb | 9.135506 | .3397491 | .03775591 | -1.526642 | .0571369 | .0010153 |
| Residual | 1.08e-12 | .3773023 | .06876377 | 1.39e-11 | .3387803 | .43793002 |

Note: (1) Worker fixed effects (WE) - $\theta_i$, Firm fixed effects (FE) - $\psi_{j(i,t)}$, Observable characteristics (xb) – β and the Residual/error term - $\varepsilon_{it}$. (2) The coefficients are estimated from equation 3.