Monopsony and Measurement Error in the South African Labour Market

By

Nicholas Forster

FRSNIC006

Dissertation Supervisor: Associate Professor Andrew Kerr

A dissertation submitted in partial fulfilment of

the requirements for the degree of

Master of Commerce

specialising in

Applied Economics

University of Cape Town

December 2023

Acknowledgments

There are a few people I would like to thank who contributed significantly to my completion of this dissertation. Without their support, writing this dissertation would not have been possible.

First, I am extremely grateful for the support of my supervisor, Associate Professor Andrew Kerr. He is a supervisor who genuinely cares for his students and wants them to do as well as they can possibly do. He provided me with extremely valuable feedback and guidance and did so in a way that always left me encouraged and excited to implement it. His gentle but clear correction gave me such confidence that I could get this done, and get it done well. Thank you, Andrew!

Second, I am grateful to those responsible for the Louis Wiener Scholarship, who have generously provided me with funding for this Masters.

Third, I must thank my closest friends. My five digs' mates provided me with endless support, love, banter and fun which I could not have done without. Living with them made writing this dissertation easier. I am eternally grateful to God for having such solid brothers in Christ upon whom I can depend. Jeff, Josh, Ben, Hamish, and Matt – I could not have asked for better friends. I must also thank Emma, my wife to be, whose love, encouragement and grace kept me joyful in Christ throughout this process. I thank you and love you Em.

Fourth, I am so grateful to my parents. They have raised, supported, and loved me to the moon and back, and this dissertation would not have even nearly been possible without them. Their generosity in providing me with an incredible education and their pushing of me to ensure I do the best I can is a big part of the reason this dissertation exists. Mom and Dad, I love you and thank you!

Last, I thank God for his grace, mercy, and love. He has been so faithful to give me the ability and the strength to complete this dissertation. I submit and surrender this dissertation entirely to his glory, for he is worthy of it all. *"Whatever you do, work heartily, as for the Lord and not for men, knowing that from the Lord you will receive the inheritance as your reward"* – Colossians 3:23-24 ESV.

Plagiarism Declaration

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

I have used the UCT-Harvard convention for citation and referencing. Each contribution to and quotation in this thesis from the works of other people has been attributed and has been cited and referenced. Any section taken from an internet source has been referenced to that source.

This thesis is my own work and is in my own words (except where I have attributed it to others).

I have not allowed and will not allow anyone to copy my work with the intention of passing it off as his or her own work.

I acknowledge that copying someone else's assignment or essay, or part of it, is wrong, and declare that this is my own work.

Word Count: 23 235

Signed by candidate

Nicholas Forster

Table of Contents

List of Tables

List of Figures

Abstract

In this paper, I analyse the extent to which monopsony power is present in the South African labour market by estimating the wage elasticity of labour supply to the firm, following Manning's (2003) method. I also consider the extent to which the method of identifying job-to-job separations and the use of poorly imputed earnings data by StatsSA changes results. I use panel data from the Labour Force Survey and Quarterly Labour Force Survey in South Africa, although only for waves in the QLFS in which I have earnings data with no imputations by StatsSA.

I find a low elasticity of labour supply to the firm in South Africa of between 0,68 and 0,83, which suggests that there is substantial monopsony power in South Africa. These estimates are far off infinity, which suggests that using perfect competition to model the South African labour market is not a realistic assumption. I find little difference in this elasticity to the firm across gender and race overall, but some differences across gender within race and across race within gender. I find that the labour supply of more educated individuals is more elastic to the firm than those with less education and that within higher educated groups, men are supplied more elastically than women. This suggests that education lessens vulnerability to monopsony power, and more so for men. Lastly, I find that results are very sensitive to the method of identifying separations, as well as to the use of the imputed StatsSA earnings data. In both cases, the elasticities estimated are less than half the above estimates.

Despite measurement error being a concern, the low estimates of the elasticity to the firm are unlikely to be biased downwards to the extent that the South African labour market more closely resembles perfect competition. Thus, those designing policy for the South African labour market should do so from an assumption that the market is imperfectly competitive. Furthermore, these estimates are not far off estimates using administrative data, which suggests that survey data can inform on monopsony power in a labour market when carefully analysed.

1. Introduction

In 2003, Alan Manning released his book, *Monopsony in Motion*. His work led to a number of authors using his method to estimate the wage elasticity of labour supply to the firm. The basic idea was that this elasticity was not infinite, as the model of perfect competition would predict. In fact, it was not anywhere near infinity, leading Manning (2003) to conclude that modelling the labour market as imperfect was a much better assumption.

Perhaps the success of Manning's (2003) book can be put down to its very simple intuition. Crucial to this intuition is a question. When a firm marginally drops the wage it pays, why does not everyone leave the firm? This is, after all, what the perfectly competitive model predicts. However, this is not the case in reality as many people would likely stay in the same job. The lesson to draw from this simple thought experiment is that firms have some power over their employees. If a firm can drop wages, and people do not leave, perhaps the firm has the power to drop the wage some more.

Although much research has been conducted on this topic, most of it focused on developed countries. In this paper, I estimate the wage elasticity of labour supply to the firm in South Africa, a developing country, using the Labour Force Survey [LFS] and Quarterly Labour Force Survey [QLFS] from StatsSA. In doing so, I gain some understanding of the extent of monopsony power in South Africa. As discussed below, South Africa's segregated past and existing high levels of unemployment suggest that monopsony power is likely extensive. As far as I am aware, only one author has done this up to now. Bassier (2023) estimates the wage elasticity of labour supply to the firm using administrative tax data as his main avenue of analysis and using labour force survey data in his appendix. However, this labour force survey contained poorly imputed earnings data, and thus in this paper, I used unimputed earnings data to estimate the elasticity.

As a secondary focus, I investigate how this estimate may be sensitive to several factors. The first of those is the method by which I identify job-to-job changes. As I will explain, the identification of job-to-job separations is crucial in estimating the elasticity of labour supply to the firm. However, identifying separations to another job is difficult when using household survey data and there is likely to be measurement error. I thus investigate how sensitive results are to changes in how I identify separations. Second, I briefly consider the sensitivity of the results to mismeasurement in earnings. I investigate whether there is mismeasurement in the publicly released and imputed StatsSA earnings data in the Quarterly Labour Force Survey by comparing it to non-public and unimputed data. I also compare the estimates of the elasticity of labour supply to the firm obtained using these different earnings variables.

This sensitivity analysis is central to the relevance of this paper. Although administrative data is often used in monopsony work, it is often more scarcely available than survey data and does not take into account the informal sector. This is especially important in developing countries. Therefore, I seek to understand in this paper the extent to which and the assumptions under which survey data is useful in monopsony work.

My findings can be summarised as follows. I find very low elasticities of labour supply to the firm of between 0,68 and 0,83. Although measurement error is a concern, these estimates are far off positive infinity, and not far off estimates found using administrative data. This suggests that the South African labour market is more similar to an imperfectly competitive labour market model than a perfectly competitive model. I find little difference in the male and female elasticities, although I do find that Coloured women and African men are more elastically supplied to the firm than Coloured men. I find that the those with greater education are more elastically supplied to the firm, and that education seems to benefit men more in this regard. Lastly, I find the results to be sensitive to how I identify separations, as well as to the use of the imputed StatsSA data. In both cases, the elasticities decrease substantially. Thus, although survey data can produce estimates not far off those found using administrative data, careful consideration of measurement error is necessary.

This dissertation is structured as follows. In the next section, I review the broad monopsony literature while including a discussion on why monopsony is relevant in South Africa. In Section 3, I explain the theoretical model that underpins my analysis, as derived by Manning (2003). I then explain the data sources I use, including a discussion on how I make use of panel data using the labour force surveys from StatsSA. In Section 5 I discuss the methodology, followed by a discussion of how I identify job separations in the presence of measurement error in Section 6. Section 7 includes some descriptive statistics, while in Section 8 I present and discuss my regression analysis and the resulting estimates of the elasticity of labour supply to the firm. In Section 9 I consider the robustness of the results, while in Section 10 I explore how poor wage imputation on StatsSA's part affects the results.

2. Literature Review

In this section, I summarise the current literature on monopsony. I first introduce the term and discuss some of the basic assumptions and implications of its use in studying the labour market. Second, I discuss approaches to monopsony that estimate the wage elasticity of labour supply to the firm, followed lastly by a discussion on why monopsony is relevant in South Africa.

2.1. Introduction to Monopsony

Monopsony in the labour market is about the power employers have over their employees. The literature is broad and includes a wide variety of approaches and emphases that range from applications of monopsony to minimum wage laws and inequality to the concentration of labour markets and their resulting wages (Manning, 2020). Despite this variety, the literature all relates to Manning's (2003) central question- if a firm drops its wages by one cent, why does everyone not leave? Manning's (2003) point is that firms have some power over their employees not to pay them their marginal revenue product, and thus the labour market cannot be perfectly competitive.

The term 'monopsony' originates with Robinson (1933) as cited in Manning (2003), who used it in discussing imperfections in the labour market. The imperfection, according to Robinson (1933) as cited in Manning (2003), is that firms may need to pay a higher wage to attract more workers. If firms need to pay a higher wage to attract more workers, then this implies an upward-sloping labour supply to the firm curve, rather than a horizontal curve as a perfectly competitive labour market model would imply. Furthermore, it implies a positive elasticity of labour supply to the firm that is not infinite as perfect competition would suggest. This example by Robinson (1933) is analogous to Manning's (2003) point that if a firm drops its wage by 1 cent, not everyone will leave. Thus, to answer his own question of why not everyone leaves, Manning (2003) summarises Robinson (1933) by pointing to the ignorance of other job opportunities, heterogeneous preferences for the type of firm one wants to work at, and the frictions and costs involved in having to search for a new job – otherwise known as immobility. An upward sloping curve of labour supply to the firm implies that firms have power over their employees. If a firm can drop its wages below its employees' marginal revenue products and not lose all its employees, then the firm has some power over their employees not to pay them what perfect competition would suggest is a 'market wage'.

This approach to studying the labour market is relevant today. Manning (2020) argues that the idea that employers have power over their employees has become an attractive idea to those seeking to explain high levels on inequality and the decline of the labour share of national income. Indeed, inequality has increased in wealthier countries over the past 40 years (Fuchs-Schündeln, Krueger & Sommer, 2010; Cingano, 2014; Guvenen, Pistaferri & Violante, 2022) and although there is some disagreement over whether it has increased or not in South Africa (Bhorat et al., 2020; Bassier & Woolard, 2021; Kerr, 2023), it is nevertheless very high. Differences in wages across firms have been shown to drive inequality (Card, Heining & Kline, 2013), and differences in monopsony power across different groups or labour markets (i.e. across gender or industry) could also feasibly drive inequality (Manning, 2020). Bassier (2023) shows that differences in firms explain a larger share of wage variation in South Africa than in the rest of the world, which suggests that firm monopsony power could be an important driver of wage inequality in South Africa.

There is also substantial evidence globally of a decrease in labour's share of national income (Rodriguez & Jayadev, 2013). A firm being able to pay a worker a lower share of their productivity necessarily implies a lower labour share of national income, as firm owners capture more of the profits from their employees' productivity (Manning, 2020).

Monopsony also has several interesting applications to important issues in the labour market. One example is the minimum wage. Under perfect competition, the imposition of any minimum wage is predicted to cause job losses. However, a growing body of literature has challenged this prediction with empirical evidence showing little or no impact on employment (Doucouliagos & Stanley, 2009; Broecke, Forti & Vandeweyer, 2017). Monopsony can account for a very small or absent impact of the minimum wage on employment as well as a definite impact on employment if the minimum wage is set too high, as shown theoretically in Ehrenberg, Smith & Hallock (2021).

Before proceeding, an important clarification regarding the term 'monopsony power' should be made. Central to Manning's (2003) approach is the assumption that employers set wages. However, Manning (2003) argues that one example of where this assumption may not be realistic is in the presence of unions, bargaining councils and minimum wages which prevents firms from exercising their monopsony power. The intuition is as follows. The primary measure of monopsony power is the elasticity of labour supply to the firm. The inverse of this is the rate of exploitation, which is how much greater in percentage terms a worker's marginal revenue product is compared to their wage. It is thus a measure of the markdown a firm could theoretically impose on their workers' marginal revenue product. However, firms may not mark down workers' wages by this much. As Manning (2003) argues, unions, bargaining councils, and minimum wages restrict the exercise of monopsony power by firms by preventing very low wages. Thus, a more accurate term for what much of the monopsony literature explores is what Manning (2003) has termed 'potential monopsony power', as the extent to which firms exercise this power to reduce their employees' wages is not being studied.

Some literature has however studied the direct effect of firms' monopsony power on wages. This literature studies the relationship between labour market concentration (in employer terms) and average wages in those labour markets (Rinz, 2018; Azar et al., 2022; Benmelech et al., 2022). The link to monopsony here is that if labour markets are more concentrated, then the ability of individuals to leave their jobs and find another employer is limited, thus giving the employer some monopsony power to reduce their wages. All three of Rinz (2018), Azar et al. (2022) and Benmelech et al. (2022) found negative relationships between labour market concentration and average wages in the USA, suggesting that firms do impose their monopsony power to a certain extent.

Having introduced the term and briefly discussed its relevance, I move in the next subsection to discussing the literature focusing on the wage elasticity of labour supply to the firm. This is a key indicator of the extent of monopsony power in the labour market, as it gives an indication of how steep the labour supply curve to the firm is. It is also the primary parameter of interest in this paper.

2.2. The Wage Elasticity of Labour Supply to the Firm

A key contribution of Manning (2003) was to demonstrate a way in which the elasticity of labour supply to the firm could be estimated. Manning (2003), using a model from Burdett & Mortensen (1998), showed that the elasticity of labour supply to the firm could be derived from a combination of the elasticity of separations from and recruits to a firm.[1] Thus, much of what followed in the literature were attempts to estimate the elasticity of separations from, and

[1] This model is explained in more detail in Section 3.

recruits to, a firm. These elasticities are the percentage change in the number of recruits to and quits from a firm in response to a 1% increase in the wage offered by the firm.

As will be discussed in more detail below, the type of data available to a researcher restricts which of these two elasticities they are able to estimate. Therefore, Manning (2003) shows under relatively reasonable assumptions that the elasticity of recruitment is equal to the negative of the elasticity of separations. This enables the estimation of the elasticity of labour supply to the firm with only one of the separation or recruitment elasticities. This subsection of literature can thus be divided into which approach is used. Within these two approaches, the literature can further be divided into what type of data is used, whether a household survey, a matched employer-employee dataset, or data from an experiment. Despite a wide variety of datasets and approaches, all papers have one thing in common – the elasticity of labour supply to the firm is low and far from infinite.

2.2.1. The Elasticity of Separations from the Firm

2.2.1.1. Household Survey Data

The earlier papers in the literature use household survey data to study the elasticity of separations from the firm, and thus the wage elasticity of labour supply to the firm. I begin with Manning (2003), who uses data from 4 sources - all panel datasets from nationally representative surveys. From the USA, the author uses the Panel Study of Income Dynamics [PSID] and the National Longitudinal Study of Youth [NLSY]. From the UK, the author uses the Labour Force Survey [LFS (UK)] and the British Household Panel Survey [BHPS]. All surveys have a yearly interval except the LFS (UK) which has a quarterly interval. Of these, the PSID, BHPS and LFS (UK) are the most comparable to my work because the NLSY limits its sample to younger individuals.

Manning (2003) uses two methods to estimate the elasticity of labour supply to the firm. In the first, he regresses a dummy variable indicating whether an individual separated or not on the log wage and other controls. He estimates this using a hazard model, and then simply multiplies the elasticity by -2 to get the elasticity of labour supply to the firm, using his relatively simple proof that the elasticity of labour supply to the firm is equal to the elasticity of recruitment less the elasticity of separation, which themselves are equal in absolute values.[2] I refer to this as the

[2] See Section 3

simple approach. In the second approach, Manning (2003) runs two different regressions, one for separations to employment, and the other for separations to non-employment. A sum of the two, weighted by the share of separations to employment, is used to calculate the elasticity of labour supply to the firm.[3] I refer to this as the complex method.

The simple elasticity to the firm is 1,95 in the PSID, 1,44 in the BHPS and 1 in the LFS (UK). Manning (2003) finds complex elasticities of 1,38, 0,75, and 0,75 for the PSID, NLSY, BHPS, and LFS (UK) respectively. Therefore, using the complex method results in a smaller estimate of the wage elasticity of labour supply to the firm. Manning (2003) does not report the elasticity of labour supply to the firm across gender. However, he finds very little difference across gender in terms of the elasticity of separations, which implies little difference in the simple elasticity of labour supply to the firm. Manning (2003) finds yearly separation rates of 21% in the PSID, 19% in the BHPS and a quarterly rate of 5,8% in the LFS (UK).

Booth & Katic (2011) also use household survey data, from the Household, Income and Labour Dynamics in Australia [HILDA] survey. Using a logit, instead of a hazard model like Manning (2003), they find an elasticity of labour supply to the firm of 0,72 when using Manning's (2003) simple approach. Therefore, when using a similar approach, they find smaller elasticities than in Manning (2003). Booth & Katic (2011) also pursue a more complex approach, although it differs slightly from Manning's (2003) complex approach.[4] They find an elasticity of 0,71 when using their complex approach and when including controls for time and industry, the elasticity increases to 0,79. They also find that the supply of men to the firm is more elastic compared to women and that the separation rate is 14% in Australia, which is lower than what was found in the UK and US by Manning (2003).

2.2.1.2. Endogeneity Concerns

Using household survey data to study the relationship between separations and the wage has a few issues. Thus, at this point it is worth discussing the endogeneity of the wage variable, and how the literature has sought to deal with this. Ransom & Sims (2010) note that the concern is that workers have unobservable characteristics that influence both their wage and their

[3] Manning's (2003) complex approach is a little more complex than explained here. See the methodology section for a full explanation.

[4] Booth & Katic (2011) assume that the elasticity of recruits from non-employment is equal to the elasticity of separations to non-employment. Manning (2003) does not, but rather follows the approach explained in the methodology.

likelihood of separation. For example, assume someone places high value on finding fulfilment in their job. This motivates them to work diligently and excellently. This, in turn, leads to higher wages. However, later in a job, if they begin to feel unfulfilled in their role, they then become more likely to quit their job in search of greater fulfilment. However, one cannot observe this desire for fulfilment in a household survey. Therefore, not accounting for it would result in an underestimation of the true negative impact of wages on the likelihood of separation.

One way in which Manning (2003) and Booth & Katic (2011) seek to ensure the exogeneity of wages is through adequate controls that control for the average level of wages. This is because workers evaluate their current wages in light of wages paid by other firms in the same market (Manning, 2003). For example, assume average wages over the entire labour market increased. It is unlikely that separation behaviour would change, as the individual's relative wage would not change. Thus, a set of covariates usually included in a wage regression should be included. Failing to do so will likely bias the coefficient, with the direction of this bias depending on the individual covariate excluded. Controls generally include gender, education, race, marital status, number of children, regional controls, health status, experience and time period (Manning, 2003; Booth & Katic, 2011).

One covariate in particular that has been discussed extensively is tenure, as including or excluding tenure are both justifiable options (Manning, 2003). On the one hand, wages may influence tenure as firms with higher wages likely have employees with longer tenures (Manning, 2003). Thus, controlling for tenure would be akin to controlling for a variable through which wages affect the likelihood of separation, resulting in the coefficient on wages being downwardly biased. In this case, it is better to exclude tenure. On the other hand, firms might have formalized pay-scale structures based on seniority. In this case, individuals are incentivized to stay at the firm longer and thus tenure would directly affect the separation likelihood (Manning, 2003). Excluding it then biases the coefficient on log wage upward. Both arguments have some merit and specifications that include and exclude tenure are often both reported. For example, Booth & Katic (2011) show that when including tenure, their estimate of the elasticity to the firm falls substantially to around 0,4. However, both Booth & Katic (2011) and Manning (2003) tend to focus on the estimates they obtained excluding tenure. This indicates a preference for tenure as a factor through which wage affects the elasticity of separation. I, therefore, will focus on estimates where I exclude tenure.

Booth & Katic (2011) also attempt to ensure the exogeneity of the wage variable using a fixed-effects logit model. They aim to solve for endogeneity by controlling for the fixed and unobserved variables. To pursue this approach workers must be present in at least two pairs of two waves (i.e. four waves). Thus a pair of adjacent waves would be 2001-2002 and 2002-2003. When doing so, the elasticity of labour supply to the firm decreases to 0,4. However, when including controls for time and industry, the elasticity increases to 1,1.

A further endogeneity concern mentioned by neither Manning (2003) nor Booth & Katic (2011) is measurement error in the separations variable, particularly with regards to separations to another job. Manning (2003) and Booth & Katic (2011) both identify separations using a tenure variable. If tenure is less than the time elapsed since the previous interview, then the individual has moved to a new job. However, if tenure is mismeasured, then one may misidentify separations. Accurately identifying separations to employment is crucial as measurement error in the dependent variable when using a non-linear regression model causes bias in the coefficients (Hausman, Abrevaya & Scott-Morton, 1998). Two non-monopsony papers that consider measurement error in tenure and its impact on identifying job changes are Brown & Light (1992) and Bergin (2015). Although discussed in more detail in Section 6, Brown & Light (1992) show that there is substantial measurement error in tenure in the PSID, which is the dataset used by Manning (2003). Bergin (2015) shows substantial measurement error in tenure in a household survey in Ireland. Brown & Light (1992) also show that this measurement error in the dependent variable can bias coefficients downwards in a separations regression. Therefore, I discuss this in more detail with regard to my data in Section 6.

2.2.1.3. Matched Employer-Employee Data

One way in which both the endogeneity of the wage variable and the measurement error in the dependent variable can be addressed is through the use of matched employer-employee data. As firms and employees are matched in this data, a richer set of firm controls is often available to the researcher to control for any confounding effects on the wage. Secondly, because a firm identifier exists, separations to employment can be identified with much less measurement error than in household survey data. The papers below use this type of data. Almost all the estimates remain very small and far off infinity, although slightly larger than in Manning (2003) and Booth & Katic (2011).

Hirsch & Jan (2015) use a German social security dataset that merges firm and worker information. They then estimate the elasticity of labour supply to the firm for natives and immigrants using the more complex approach suggested by Manning (2003). Using a fixed effects approach, like Booth & Katic (2011), they find elasticities of 1,14 and 1,36 for immigrants and natives respectively. These estimates are generally larger than the non-fixed effects approaches in Manning (2003) and Booth & Katic (2011), which shows that unobserved characteristics in employees and firms may affect both separations and wages.

Some approaches differ from this standard approach, however, such as Barth & Dale-Olsen (2009) who use matched employer-employee data from Norway. They estimate a firm-level wage premium for various types of workers by regressing wages on a host of characteristics including firm dummies. They then regress the average separation rate on this firm-specific premium which gives an estimate of the elasticity of separation. The authors calculate this across worker groups that differ in human capital and gender (Barth & Dale-Olsen, 2009). They find elasticities of labour supply to the firm of 1,49 and 1,14 for lesser-educated men and women respectively. For higher-educated men and women, they find elasticities of 1,18 and 1,09 respectively. Men thus have a larger elasticity than women, as in Booth & Katic (2011). Once again, these estimates are slightly larger than the estimates in Manning (2003) and Booth & Katic (2011). They are similar to Hirsch & Jan (2015).

Ransom & Sims (2010) also use matched employer-employee data, but in Missouri in the USA. However, they also use a slightly different approach to estimate the elasticity of separations for public school teachers. They argue they can observe exogenous shifts in wages because they use the public-school salary schedules for all sub-districts within Missouri (Ransom & Sims, 2010). These salary schedules instrument for the true wage, and the authors argue that the schedules are good instruments because they are correlated with the actual wage teachers are paid, but uncorrelated with any of the unobservable teacher characteristics. When using this instrumental variable approach, Ransom & Sims (2010) find a labour supply elasticity of around 3,57. Thus, when an explicit identification approach is used, the coefficient elasticity is larger than the estimates in Manning (2003) and Booth & Katic (2011) are even larger than the other matched employer-employee estimates discussed thus far. Although this is likely in part due to their IV approach, as if the instrumental variable is indeed exogenous, the true effect of wages on separations may be larger. As in the example in Section 2.2.1.2., an instrument such as a wage schedule would be unrelated to the employee's preference for fulfilment, thus isolating a larger effect of the wage on separation. The larger coefficients may however also be because they restrict their analysis to a single-occupation labour market. This may also explain their low yearly separation rate of 9,8%.

A slightly different approach was used by Ransom & Oaxaca (2010). They estimate the elasticity to the firm using data from a single firm with multiple establishments. Their analysis is thus restricted to one firm within the retail grocery industry. The estimation approach is similar to Manning's (2003) approach, but they argue that wages are exogenous because they are fixed by the outcome of a bargaining process, and thus the firm has no direct control over the wages. Ransom & Oaxaca (2011) find estimates of around 2,8 to 3 for men, and 2,1 to 2,4 for women, all when excluding tenure. When including tenure, the male range drops to between 2,4 and 2,7, and women between 1,5 and 1,8. They thus also find that men have a larger elasticity to the firm. They find a separation rate of 16%.

The literature on developing countries is sparse. One exception is Vick (2017). The author uses a matched employer-employee dataset for the formal sector in Brazil and uses a method very similar to Booth & Katic's (2011) complex approach to estimate the elasticity. Depending on the controls used, Vick (2017) finds elasticities to the firm of between 1 and 2. The author also finds that men have a statistically significantly larger elasticity than women.

A recent and innovative approach is found in Bassier, Dube & Naidu (2021) and Bassier (2023). First, the authors attempt to isolate the firm component of wages by using an approach by Abowd, Kramarz & Margolis (1999) [AKM]. They regress wages on a dummy variable indicating whether an individual is employed at a particular firm in a particular period. This regression also includes worker and time fixed effects. By doing this, they decompose wages into a firm and worker component and then only use the firm component of wages as produced by the AKM approach in their analysis. They then propose an innovative solution to the problem of unobserved worker heterogeneity. They use workers with similar employment histories who begin work at firms paying different wages. These workers had similar recent wages, tenures and employers. The authors track their subsequent separation behaviour and use the difference in wages and behaviour to estimate an elasticity to the firm. They call this the 'movers' approach.

Bassier, Dube & Naidu (2021) use matched employer-employee data from Oregon in the USA. They find an elasticity of labour supply to the firm of 4,2. This is substantially larger than the estimates found in Manning (2003). Nevertheless, it is still an elasticity far from positive infinity. They find that there are lower labour supply elasticities in lower-wage markets, which suggests that lower-wage employees in lower-wage markets faces higher levels of potential monopsony power.

Bassier (2023) is a particularly important paper because it is the only paper that estimates the elasticity of labour supply to the firm in South Africa. The author uses administrative tax data that includes 8 million employees, matched to their firms between 2011 and 2016. Descriptively, Bassier (2023) finds an average yearly separation rate over the period of around 37%, which is high. In addition to the 'movers' approach used in Bassier, Dube & Naidu (2021), the author uses two others. The first is Manning’s (2003) complex approach. The second is a first-difference regression that instruments changes in wage with changes in value-added per worker at the firm level. This approach exploits within-firm variations in value-added per worker over time, which likely influences wages. Unobserved heterogeneity in firms is therefore controlled for, while also using a plausibly exogenous instrument for wages. The 'movers' approach is the author's preferred approach. Manning's (2003) complex approach yields an elasticity of 0,86. When including industry-by-region market fixed effects it decreases to 0,76. When using the first-differenced approach, the elasticity is 0,74, and when using the ‘movers’ approach, the elasticity is 1,6. Thus, once again, when an explicit identification approach is used to deal with the endogeneity of wages, the elasticity is larger. Bassier (2023) does use Manning’s (2023) simple approach with the QLFS between 2010 and 2015, finding elasticities to the firm of 0,32 and 0,3 in the formal and informal sectors respectively. Bassier (2023) cites measurement error and fewer controls as the primary reason for these lower elasticities. Poorly imputed earnings data is also likely a reason for these low estimates, which I discuss later in the dissertation.

Differences between Bassier, Dube & Naidu (2021) and Bassier (2023) also suggest that the elasticity to the firm in South Africa is smaller than in the USA. Although Bassier, Dube & Naidu (2021) only estimates the elasticity in a single US state, the fact that South Africa's is much lower in Bassier (2023) suggests that the USA has a larger elasticity of labour supply to the firm than South Africa. Lastly, Bassier (2023) provides suggestive evidence that high unemployment may be driving high monopsony power, as regions in South Africa with high unemployment have lower elasticities.

The evidence from the literature discussed above suggests that the elasticity of labour supply to the firm is low and far from infinity. Further, it appears that when using matched employee-employer data and implementing an explicit identification strategy, the estimates of the elasticity of labour supply to the firm are slightly larger than when using household survey data, but nevertheless far off infinity. All of the above estimates, however, were estimated using the elasticity of separations. In the next subsection, I discuss the estimates found using the elasticity of recruitment to the firm.

2.2.2. The Elasticity of Recruitment to the Firm

Most of the literature discussed in this subsection uses experimental data because the experimenters are able to vary the wage offered to potential employees, and thus ensure the exogeneity of the wage and observe all of the wage offers. The approaches to estimating the recruitment elasticity in the papers below are much more complex than the approach to estimating the separations elasticity used in this paper and thus a full explanation of how they estimate their recruitment elasticities is beyond the scope of this paper. They do however all make use of the relationship Manning (2003) derives between the elasticity of recruits and the elasticity of labour supply.

Dal Bó, Finan & Rossi (2013) is one such paper. They estimate the elasticity of recruitment using data from an experiment in Mexico where different wages were randomly offered for the same job to individuals. These offers were distributed across different job-posting platforms in different locations. They estimate an elasticity of labour supply to the firm of 2,15.

Another example is Caldwell & Oehlsen (2018). These authors used an experiment where Uber drivers were randomly offered higher wages. By studying drivers who also use other ride-share apps, the authors could study the extent to which they shift hours to Uber as the Uber wage increases. They find an elasticity of labour supply to the firm of around 2,7 for men and 4 for women. Their finding of a higher female elasticity of labour supply to the firm is in contrast with much of the above literature.

A different experiment was conducted by Dube, Jacobs, Naidu & Suri (2020). They conducted experiments on an online task request website to estimate the elasticity of labour supply to the firm. Amazon Mechanical Turk (MTurk) is an online micro-task platform that allows people to advertise small tasks for pay. Registered workers on the platform can then choose tasks to complete for the advertised wage. The platform is useful for experimenters as they can randomly vary the wage for the same task and then use then time it takes for the task to be accepted as a measure of the probability of acceptance. They find an extremely small elasticity of labour supply of between 0,05 and 0,14 (Dube et al., 2020). This they deem surprising, as MTurk is a very large and diverse online labour market.

2.3. Summary of the Elasticity to the Firm Literature

The above review has given insight into the literature concerning the elasticity of labour supply to the firm. However, the literature is certainly larger than has been discussed above. A fuller treatment can be found in Sokolova & Sorensen (2021). This meta-study includes 1320 estimates of the elasticity of labour supply to the firm from 53 papers. Of all the estimates they survey that estimate the elasticity directly, the median is 1,4 while the median is 1,73 in papers that estimate it via the elasticity of separations. The authors find some other interesting results in their meta-analysis. They find evidence of estimates published in top journals being higher. However, this result disappears when conditioned on the presence of an identification strategy. Therefore, approaches without an explicit identification strategy will likely produce lower estimates. The low estimates produced by Manning (2003) and Booth & Katic (2011) are evidence of this. It is therefore also likely to be the case in this dissertation.

The above review has also revealed a few other important aspects of monopsony analysis. First, men tend to have a higher elasticity of labour supply to the firm, which is what Sokolova & Sorensen (2021) find. This suggests that they face less potential monopsony power than women. Second, the elasticity in South Africa seems to be quite low, as shown by Bassier (2023). Thus, I expect to find elasticities that are smaller than those found in Manning (2003) and Booth & Katic (2011). Lastly, Bassier (2023) also reveals that separation rates are likely to be quite high in South Africa.

2.4. Why continue with the simple approach?

It is necessary to consider whether using only the approach of Manning (2003) and Booth & Katic (2011) is worth it, since the literature has moved on from Manning's (2003) approach and from the use of survey data primarily to deal with the endogeneity of the wage variable. Using Manning's (2003) approach with survey data results in smaller estimates than the rest of the literature and so it is likely to be the case in analysis using South African survey data. I persist with this approach for three reasons.

First, no one, to my knowledge, has pursued Manning's (2003) approach using the relatively large pool of South African survey data. Bassier (2023) does make use of Manning's (2003) framework but uses administrative data in which firms are identified. Additionally, much of the work done so far in the entire monopsony literature has focused on developed countries. In fact, out of the 1320 estimates of the elasticity in the meta-study by Sokolova & Sorensen (2021), only 136 of them are from developing countries. Thus, studying South Africa makes the literature more representative.

Second, estimating the elasticity of separations using survey data is not an entirely outdated approach. Langella & Manning (2021) use a similarly simple approach to estimating the elasticity of separations and in doing so they make some important inferences about the labour market. Thus, using an estimate of the elasticity of substitution obtained using survey data is still a relevant approach.

Third, Manning's (2003) primary argument still holds. That is that although the estimate of the elasticity of labour supply to the firm may be downward biased, the magnitude of the bias would have to be huge to suggest that the labour market is perfectly competitive. As such, I may find downward-biased estimates of the elasticity of labour supply in South Africa. But if their magnitude is sufficiently small then it can be argued that the South African labour market is not perfectly competitive. Understanding the evidence for monopsony in South Africa is an important endeavour, to which I now turn my attention. In the following subsection, I discuss the relevance of studying monopsony in South Africa.

2.5. The Relevance of Monopsony in South Africa

To consider why studying monopsony in South Africa is important, it is helpful to go back to the causes of monopsony as suggested by Robinson (1933) and adopted by Manning (2003). These are heterogeneous preferences, ignorance, and immobility.

Immobility is a particularly important friction given South Africa's history. The segregation and oppression of black people caused by Apartheid left a significant mark on the South African labour market. A simple look at labour market outcomes using the Post-Apartheid Labour Market Series [PALMS] by Kerr, Lam & Wittenberg (2019) reveals this. Consider the average unemployment rate of 25% one finds if averaging over the 25 years or so in PALMS. For Africans, this is 29%, Coloureds 22%, Indians and Asians 12% and Whites 6%. There are thus clear differences in labour market outcomes across race. This makes South Africa an interesting country in which to study monopsony in that one can explore to what extent these inequalities are present in potential monopsony power.

Some previous research discussed the importance of a monopsonistic labour market in apartheid South Africa. The seminal work on this is Francis Wilson's 1972 book, *Labour in the South African Gold Mines 1911-1969*. Lucas (1985) also discussed monopsony in gold mine labour in apartheid South Africa. As discussed in both Wilson (1972) and Lucas (1985), the hiring of African workers for gold mines was managed by a single agency during Apartheid. This agency was therefore a pure monopsonist, being the only purchaser of labour in the market. This pure monopsonist implemented a maximum wage in conjunction with the Chamber of Mines. Furthermore, up until the 1970s, it was a criminal offence for workers to quit their jobs before their contract was up (Lucas, 1985). Both Wilson (1972) and Lucas (1985) show that these monopsonistic conditions had a detrimental impact on these workers. Wilson (1972) showed that real wages did not increase for these mine workers for 60 years despite concurrent increases in the price of gold, while Lucas (1985) presented evidence for an upward-sloping labour supply curve of black workers to the single monopsonist. Lucas (1985) estimated a labour supply function using time series data by regressing the log of aggregate employment numbers for African mine workers on average wages over the course of 30 years. For foreign and domestic black workers, the elasticity to the firm was 0,14, and for South African black workers alone, the elasticity to the firm was 0,5.

Although this research was done on the labour market for mine workers, it is likely indicative of the entire African labour market during Apartheid. Eidelberg (1997) gives a helpful summary of the contents of some early Apartheid legislature. Amongst other legislation, the Group Areas Act of 1950 and the Native Laws Amendment Act of 1952 placed severe restrictions on the ability of African workers to live and work in so-called 'White areas'. The Group Areas Act gave the centralised state control over the prevention of African land ownership in urban areas and these acts enabled the state to enforce influx control for urban areas. Africans could only stay within a 'white' area for more than 72 hours if they could prove they were employed by a white firm owner. The Natives Abolition of Passes and Coordination of Documents Acts of 1952 created a centralised documentation system that Africans were required to use to gain access to 'White areas'.

Consider these restrictions through the lens of monopsony. This legislation prevented the freedom of movement of Africans in urban areas, thus logically infringing on the job mobility of Africans. If an African individual could not freely move to find new or better employment, then they were less likely to leave their job. Therefore, an employer could pay them a lower wage.

There is evidence that the Apartheid legislature still impacts the labour market today, which suggests that the causes of monopsony are still important characteristics of the labour market today. For example, work on spatial inequality and poverty in South Africa has shown that areas which were once Apartheid homelands are still worse off than other areas today (David et. al, 2018). Thus, the geographical immobility caused by Apartheid still exists today, thus likely creating frictions in the labour market for those who live there.

A more direct link to monopsony can be made through the literature on transport costs. High transport costs make it costly to travel to look for new work, thus preventing workers from quitting their jobs, thereby giving their employers monopsony power. Kerr (2017) showed that transport costs are very high in South African cities compared to OECD countries, with transport costs acting as an effective tax of around 30% on commuters' earnings. Shah & Sturzenegger (2022) embed transport costs into a simple search and matching model for the South Africa labour market and argue that these transport costs negatively impact labour market outcomes for South Africans.

As discussed above, there is also recent evidence of the importance of monopsony in South Africa, as shown by Bassier (2023). Depending on the method used, the wage elasticity of labour supply to the firm is estimated to be between 0,86 and 1,6. These are low compared to the rest of the world, and far off infinity, which suggests that there is substantial monopsony power in South Africa.

The main points from this discussion are three-fold. First, South Africa has a history of monopsonistic conditions. Second, South Africa has high frictions and persistent inequalities in the labour market. Third, there is some recent evidence of monopsony power in South Africa. My conclusion is therefore that South Africa likely exhibits significant potential monopsony power. Thus, adding to the currently sparse empirical evidence for this is important.

Before proceeding, it is worth mentioning again that what I study in this paper is 'potential monopsony power'. In reality, employers may not exercise their full potential monopsony power because of the prevalence of unions, bargaining councils and minimum wages. South Africa has a relatively high prevalence of unions and bargaining councils as well as a national minimum wage implemented in 2019, which replaced sectorally determined minimum wages. Bhorat & Stanwix (2018) estimate that in 2014, 38,8% of workers in South Africa were either in a union or covered by some sort of bargaining agreement. Furthermore, evidence in South Africa suggests that bargaining councils not only increase the wages of those in covered firms but also in closely related firms not covered by bargaining council agreements (Bassier, 2022). Therefore, the full extent of the potential monopsony power I expect to find is likely not exercised because unions, bargaining councils and minimum wages are prevalent in South Africa.

Having established that monopsony is a study of the extent to which firms may have power over their employees to pay them less than their marginal revenue product, and that evidence both globally and in South Africa suggests that this power is substantial, I explain in the next section Manning's (2003) theoretical model underpinning a study of monopsony in the labour market.

3. Theory

3.1. The Wage Elasticity of Supply to the Firm

The theory informing the study of monopsony in the labour market comes from Manning's (2003) book, *Monopsony in Motion.* Manning (2003) presents several theoretical models that enable the study of monopsony. The important parameter in these models is the wage elasticity of labour supply to an individual firm. The theory enables the use of simple econometric techniques to estimate the extent of monopsony power in the labour market. I discuss three of these models here. I first briefly discuss the simple static model, followed by the simple dynamic model and then then the Burdett-Mortensen [BM] model.

3.2. The Simple Static Model

I first discuss the simple static model, which introduces to the wage elasticity of labour supply to the firm. This simple static model considers a profit-maximising firm with the profit function:

$$\pi = Y(N) - w(N).N \tag{1}$$

Y refers to production, N to labour, and w to wages. If we consider the wage elasticity of labour supply to the firm to be as follows:

$$\varepsilon = wN'(w)/N(w) \tag{2}$$

then we can rearrange the first-order condition of (1) to the following:

$$\frac{Y'-w}{w} = \frac{1}{\varepsilon_{Nw}} = \varepsilon \tag{3}$$

Equation (3) represents what is sometimes referred to as the rate of exploitation. This is the difference between the marginal revenue product of a worker and the wage they are paid and Equation (3) shows that it is equal to the inverse of the wage elasticity of labour supply to the firm. Under perfect competition, workers are paid their marginal revenue product and thus the elasticity of labour supply to the firm is infinite. However, under monopsonistic conditions, workers are paid below their marginal revenue productivity. Thus, the elasticity is finite and low. This therefore makes the estimation of the wage elasticity of supply to the firm an important endeavour.

3.3. The Simple Dynamic Model

Manning (2003) then shows in a dynamic partial equilibrium model how ε_{Nw} can be derived. Consider Equation (4), which represents the number of workers at a firm in the current period.

$$N_t = [1 - s(w_t)]N_{t-1} + R(w_t) \tag{4}$$

N_t and N_{t-1} represent the number of workers in the firm in the current and previous periods. $s(w_t)$ represents the separation rate of employees from the previous period. $R(w_t)$ represents the number of recruits to the firm in the current period. The wage is assumed to be negatively related to the separation rate and positively related to the recruitment level. Under steady-state conditions, employment in the current and previous periods are equal. Under such conditions, the following condition holds as separations are equal to recruits:

$$s(w)N(w) \ = \ R(w) \tag{5}$$

The above condition can be rearranged to let labour supply equal the recruitment level over the separation rate. If we then differentiate this rearranged condition, multiply through by w/N, and substitute in the steady state condition for N, we can solve for the wage elasticity of supply to the firm, found in Equation (6) below.

$$\varepsilon_{Nw} = \frac{w}{R}\frac{dR}{dw} - \frac{w}{s}\frac{ds}{dw} = \varepsilon_{Rw} - \varepsilon_{sw} \tag{6}$$

The wage elasticity of labour supply to the firm is thus equal to the elasticity of recruitment less the elasticity of separation. The recruitment and separation elasticities are thus crucial in estimating the elasticity of labour supply to the firm. However, the above model is only a partial equilibrium model focusing on a single firm. I derive the general equilibrium model from Manning (2003) below. This is a version of the model derived by Burdett & Mortensen (1998).

3.4. The Burdett-Mortensen Model

Manning (2003) derives a general equilibrium model to explore monopsony power in the labour market, building on the model developed by Burdett & Mortensen (1998). Key to this model is the inclusion of interactions with other firms in the market.

The assumptions of the model are as follows, as discussed in Manning (2003). There exist M_w workers who are equally productive and attach a value b to leisure. There are M_f employers, who are infinitesimally small relative to the market. They exhibit constant returns to scale and contain workers with a productivity of p. These employers set wages, with a single wage paid in a firm. The cdf of wages across firms is given by $F(w)$. All workers, regardless of their employment status, receive job offers at a rate of λ. These are drawn randomly from $F(w)$. There exists an exogenous job destruction rate of δ.

One assumption worth drawing attention to is the assumption that firms set wages. How wages are set is a topic of debate in labour economics and Manning (2011) discusses this in some detail. Manning's (2011) conclusion is that it is not unreasonable to assume that firms set wages, which I agree with. It may be true that some negotiation is present. However, for many workers, it is likely that the wage is simply advertised by the employer and workers then decide whether to apply for the job or not.

Two assumptions about worker behaviour are of crucial importance in this model. The first is the assumption that workers move firms if they receive a higher wage offer than they are currently earning. Second is that unemployed workers take a job if the wage offered exceeds their reservation wage. This is assumed equal to the value of leisure. The assumption that worker movement depends on their relative wage is important, as it enables one to derive a relationship between the separation and recruitment elasticities. This will be discussed in more detail in the next subsection.

One last characteristic of the model is worth discussing before proceeding. Manning (2003) shows how the model produces a continuous cdf of wages across firms. This is attractive in one sense and not in another. It is attractive as it produces wage dispersion despite homogenous firms and workers. This is a characteristic of the labour market in reality. However, the continuous nature of the cdf is less realistic. This is because wage bunching is an observed labour market phenomenon (Manning, 2003).

Having discussed the assumptions, I now explain Manning's (2003) version of the BM model. Chapter 2 of Manning (2003) gives the full treatment. The separation rate is modelled as follows:

$$s(w; F) = \delta + \lambda[1 - F(w)] \tag{7}$$

The job offer rate is multiplied by the proportion of firms with a higher wage than the current firm. This is then added to the job destruction rate. Deriving the recruitment rate is a little more complex and requires some prior derivations. Firstly, the non-employment rate is represented as follows:

$$u = \frac{\delta}{\delta + \lambda} \tag{8}$$

Intuitively, if the job destruction rate is high relative to the job offer rate, unemployment is high. A derivation of the distribution of wages across workers, $G(w)$, is also necessary (not to be confused with the distribution of wages across firms, $F(w)$). This is derived by exploiting the equilibrium condition of separations equal to recruits in a firm paying less than w. $G(w)$ can thus be shown to be as follows:

$$G(w; F) = \frac{\delta F(w)}{\delta + \lambda[1 - F(w)]} \tag{9}$$

One will notice that $F(w)$ stochastically dominates $G(w; F)$. Therefore, in this model, workers are concentrated in higher-paying firms. This is a well-established labour market fact.

From the above, one can calculate the flow of recruits to the firm. Manning (2003) shows this can be represented by Equation (10) below.

$$R(w; F) = \frac{\delta\lambda}{M[\delta + \lambda(1 - F(w))]} \tag{10}$$

M in the above equation represents the ratio of firms to workers in the market. The steady-state condition, $sN = R$, can then be used to obtain the labour supply to the firm:

$$N(w;F) = \frac{\delta\lambda}{M[\delta+\lambda(1-F(w))]^2} \quad (11)$$

This labour supply to the firm is increasing in the wage. The assumptions that the separation rate and recruitment flow are negatively and positively related to the wage cause this.

From here, the equilibrium labour supply and profits can be solved for. This allows one to solve for the wage distribution of the firms, and thus the expected wage of a worker. Equation (12) below shows the expected wage of the worker:

$$E(w) = \frac{\delta}{\delta+\lambda} b + \frac{\lambda}{\lambda+\delta} p \quad (12)$$

The expected wage is a weighted average of the worker's productivity, p, and the worker's reservation wage (set equal to the price of leisure), b. The weight on the marginal product can further be expressed as follows:

$$\frac{\frac{\lambda}{\delta}}{1+\frac{\lambda}{\delta}} \quad (13)$$

Thus, the importance of productivity in setting wages is increasing in the ratio of the job offer rate to the job destruction rate. Thus, if job offers arrive quickly, the ratio is large and the worker's wage is closer to their productivity than their reservation wage. This indicates less monopsony power. Thus, this ratio is a good measure of the extent of monopsony power in the market.

However, this ratio is rather difficult to calculate. Thus, Manning (2003) develops a simple back-of-the-envelope measure of monopsony power. This is the share of recruits from non-employment, which is a little easier to estimate than directly estimating the ratio in equation (13). Manning (2003) shows that recruitment from non-employment is decreasing in the above ratio. Therefore, estimating the share of recruits from non-employment enables the researcher to gain an idea of the extent of monopsony power in the market. In this model the higher this proportion, the more monopsony power is present.

The key characteristic of the Manning (2003) version of the BM model is that is implies an upward sloping labour supply curve to the firm. Thus, the model characterises an imperfect labour market where the wage elasticity of labour supply to the firm is not infinite, which indicates that there is monopsony power in the labour market. I have also shown in this section that the wage elasticity of labour supply to the firm can be estimated using the elasticity of separations from and recruits to a firm. Estimating these elasticities requires data in which either separations or recruits can be observed alongside the wage of the individual. I thus turn in the next section to discussing the data I use.

4. Data Description

There is a wealth of household survey data in South Africa. However, up to now this data has not been used to study monopsony. This section describes which datasets I use and how I am able to track individuals over time. I make use of two household surveys conducted by Statistics South Africa [StatsSA] that focus on the labour force, the Labour Force Survey [LFS] and the Quarterly Labour Force Survey [QLFS]. Another well know household survey in South Africa that tracks individuals over time is the National Income Dynamics Study [NIDS]. However, I do not use NIDS in this paper because the time between surveys is too large (2 to 3 years). I thus use the LFS and QLFS, which have smaller intervals between waves.

4.1. LFS

The first panel I use is from the Labour Force Surveys (LFS) conducted between September 2001 and September 2007. The LFS was a bi-annual household survey conducted by StatsSA. It was a household survey focusing on the labour market characteristics of its respondents (StatsSA, 2001). A rotating panel design was used whereby 20% of the dwellings in the sample were rotated out every survey. StatsSA conducted the surveys in February/March and September every year between 2000 and 2007. I use the LFS data as found in the Post-Apartheid Labour Market Series [PALMS], which is a dataset in which Kerr, Lam & Wittenberg (2019) harmonised all South African labour force data.

However, there is a slight difference in the data I use between the surveys from September 2001 to March 2004 and the surveys from September 2004 to September 2007. In 2001:2-2004:1 data, I use panel identifiers created by StatsSA, which they created using individual's names to more accurately track individuals over time.[5] These panel identifiers are not publicly available, and are the same identifiers are used by Banerjee et. al. (2008).[6]

In 2004:2-2007:2, I use the publicly available data, which does not contain these panel identifiers. In this data, a household number and person number are provided, which together serve as a personal identifier. However, there is no guarantee that the identifier represents the same individual as surveyed in the previous wave. The persons could have, for example, moved

[5] Note that the number after the colon indicates which of the two surveys in a single year I am referring to. I.E. 2003:1 refers to the LFS conducted in March 2003. For the QLFS, this number refers to the quarter.
[6] This dataset was provided to me by my supervisor, Andrew Kerr.

out of the dwelling or the person number assigned within the household could have changed. Therefore, to use the Labour Force Surveys as a panel, the construction of an individual panel identifier is necessary. The household identifier and other immutable characteristics can be used to do this. There are differing approaches in the literature to doing this, but an extensive discussion on the merits of different approaches is beyond the scope of this paper. However, the reader can find such a discussion in Pan (2022).

Pan (2022) identifies 4 previous approaches to creating a panel using labour force surveys in South Africa. The author conducts a thorough evaluation of each of the approaches. The author settles on a slight alteration of the approach pursued by Ranchhod & Dinkelman (2008) [RD].

I implement Pan's (2022) altered RD approach as follows.[7] I matched people across waves who have the same household ID, self-declared race, and gender. Age differences of up to 1 year are allowed. I remove any matches where individuals decrease in their education by more than 1 year. I also remove any matches where people go from being married or divorced in one wave, to never having been married in the next. Pan (2022) notes one important cost of using the RD approach. In this approach, one cannot distinguish between individuals in the same household with the same race and gender and only differ in age by 1 year. Thus, removing these people from the data is necessary. I follow Pan (2022) in adding an additional criterion to the RD approach. I remove matches of individuals who change educational categories by more than 1.[8] For example, I remove an individual who moves from primary education to tertiary education across one wave.

The 2001:2-2004:1 panel contains 224 425 observations, when restricted to the working-age population. There are 74 606 individuals in this panel. Around 42% of the individuals are present in two waves, 31% in three and the rest in more than three. The 2004:2-2007:2 panel contains 231 692 observations, when restricted to the working age population. There are 90 534 individuals in this panel. On average across all waves, only 32% of each individual wave's cross-section could be matched in the next wave, although it should be noted that included in the unmatched proportion of each cross-section are individuals in the 20% of households rotated out of the survey in the next wave.

---

[7] Significant thanks must go to Sammy Pan for providing me with his Stata code.
[8] These categories are 1) No education, 2) Some primary education, 3) Some secondary education, 4) Matric, 5) A degree or some tertiary education.

4.2. QLFS

The next dataset I use is the Quarterly Labour Force Survey as found in PALMS. This was StatsSA's replacement for the Labour Force Survey and began in 2008. The QLFS also employs a rotating panel design, with 25% of the dwellings being rotated in and out each quarter. I once again construct a panel using the matching algorithm adapted from Pan (2022). However, I do not use all the available QLFS data between 2008 and 2023. Rather, I use data only from the four quarters in 2011, 2018 and 2019 because in these years I have unimputed earnings data in each quarter.[9] This is important as there are significant issues with the publicly released earnings data in the QLFS (Kerr & Witternberg, 2021; Kerr, 2023). Kerr & Wittenberg (2021) and Kerr (2023) show that imputations made by StatsSA for those who responded in brackets or with don't knows and refusals are very poor. Therefore, I only use the unimputed data in my main analysis.

The final QLFS panel for the working age population contains 399 172 observations, with 143 066 individuals. About 32% of the individuals are present in 2 waves, 34% in 3 and 34% in 4. On average across all waves, only 54% of each individual wave's cross-section could be matched in the next wave, although it should be noted that included in the unmatched proportion of each cross-section are those individuals in the 25% of households rotated out of the survey in the next wave.

4.3. Weights for Missing Data and Attrition

Some issues of missing data and attrition should be discussed here, as well how I use weights in the data to account for these issues. I discuss the weights used to account for unit non-response, then item non-response in earnings, and then finally how these non-response weights are adjusted for attrition.

As with any survey, unit non-response is an issue. I use two different weights for the two groups of datasets to adjust for this. In the LFS Panel between 2001 and 2007, I use the cross-entropy weights developed in the creation of PALMS. However, in the QLFS panel from 2011 and 2018/19, I use the StatsSA weights. I do not use the cross-entropy weights available in PALMS as they have not been created for quarters 3 and 4 of 2019. Thus, I use the regular StatsSA weights to ensure I include the whole of 2019.

[9] This data was provided to me by my dissertation supervisor, Andrew Kerr, and is not publicly available.

I use an additional weight when analysing wages because some individuals respond to the earnings question in a bracket. This is a type of item non-response, as although we have some information on the individual's wage, we do not have the actual wage amount. If one uses the earnings variable in the surveys without any adjustment, those responding in brackets are ignored. To solve this issue, the bracket weights developed in PALMS are used which were first suggested by Wittenberg (2008). These bracket weights weight up those who responded with an amount to represent those who responded in brackets within the same earnings category. Although useful, this does not solve the issue of complete refusals or non-responses. I do not impute for earnings myself, as this is likely to add measurement error to my primary independent variable.

I adjust all of the above weights to account for attrition. Attrition is a missing data issue where an individual is not observed in certain time periods (Wooldridge, 2010). This results in an unbalanced panel and if attrition is non-random this may cause bias in my coefficients (Wooldridge, 2010). Using a matching algorithm to create the panel could exacerbate the attrition issue. Consider the point made by Pan (2022). Assume that the matching algorithm rejects a true match between two waves because the variables used in the algorithm are mismeasured. This would result in non-random attrition which renders any statistics inferred as biased. I use an inverse probability weighting approach to solve this issue (Wooldridge, 2010).

In order to construct these weights I regress a dummy variable indicating a successful match in the next period on observable characteristics in the current period. Using a probit regression, I then obtain the predicted values from this regression. I then adjust the weights used above by the inverse of these predicted values. I run these regressions for every single wave which has a subsequent adjacent wave. I do so to account for any changes in the relationship between the matching algorithm and observable characteristics across waves.

There is some difference of opinion about what the relevant sample should be. Both Pan (2022) and Ranchhod & Dinkelman (2008) used all individuals aged 15-64. However, Ranchhod & Dinkelmann (2008) only use those households in a particular period that appear in the next wave. They do this so as to not include households that were deliberately rotated out in creating the weights. However, Pan (2022) argues that households that refuse to respond in the following period should not be excluded and I follow his approach.

I report results from the attrition regressions for three pairs of the LFS waves, and three pairs of the QLFS waves in Appendix A. In both panels, there are a number of statistically significant independent variables. This suggests that attrition is non-random. I discuss a few noteworthy results below.

In most cases, older individuals are more likely to match across wave than those in the 15-24 age category. Individuals in dwellings with higher proportions of children are more likely to match across wave, while those in households with a higher proportion of working age men are less likely to match across wave. Although in some waves women are more likely to match across waves, the difference in likelihood is generally small and weakly significant.

4.4. Quality of the Constructed Panels

As discussed above, the QLFS panels and the panel in the LFS between 2004:2 and 2007:2 are constructed by a matching algorithm. Thus, the ability of the matching algorithm to accurately match individuals across wave should be briefly discussed. A comprehensive discussion of several methods and their accuracy in the QLFS is done by Pan (2022). I discuss whether Pan's (2022) method does as well in the LFS as he shows it does in the QLFS. I also compare the accuracy of Pan's (2022) method to the other three methods he discusses. Pan (2022) bases his method on Ranchhod & Dinkelmann's (2008) [RD] strict approach. Their expanded approach simply implements less constraints on who is matched. Anand, Kothari & Kumar (2016) [AKK] use StatsSA's person number variable to identify matches and then impose restrictions on any changes in age, race or gender. Leung, Stampini & Vencatachellum (2014) [LSV] also use this variable, but impose slightly different conditions than AKK. A detailed discussion of these approaches is done by Pan (2022).

Table 5A was taken from Pan (2022), who used a confusion matrix from Tharwat (2021) to classify possible outcomes when using the matching algorithm. Consider two waves, i and j. If the matching algorithm matches the correct person in wave i to a person in wave j, then it is a "True Positive" [TP]. If matched to an incorrect person in wave i, whether a true match exists in wave j or not, it is a "False Positive" [FP]. If the algorithm correctly does not match a person in wave i to anyone in wave j, it is a "True Negative" [TN]. If the algorithm does not match a person in wave i to anyone in wave j, and there is a true match, then it is a "False Negative" [FN].

Table 5A: Confusion Matrix

		Actual Matching Outcome		
		i_j	i_k(k ≠ j)	i_null
Predicted Matching Outcome	i_j	TP	FP	FP
	i_null	FN	FN	TN

Key: TP – True Positive, FP – False Positive, FN – False Negative, TN – True Negative.

Several measures of accuracy can be calculated using these categories, following Pan (2022). First, I calculate a true positive and true negative rate using the formulas below. These are the proportion of true matches or non-matches correctly predicted by the algorithm.

$$TPR = \frac{TP}{TP+FN}; TNR = \frac{TN}{TN+FP} \tag{14}$$

Two more metrics can be calculated. These are the positive and negative predictive values. These are the proportion of all predicted matches and non-matches that are correct. The formulas for these metrics are displayed below.

$$PPV = \frac{TP}{TP+FP}; NPV = \frac{TN}{TN+FN} \tag{15}$$

To evaluate the accuracy of Pan's (2022) matching algorithm in the LFS, I compare a panel created by the algorithm for 2001:2-2004:1 to the official StatsSA panel. This assumes that StatsSA's 2001:2-2004:1 panel is accurate, but I cannot be entirely certain of this. As far as I am aware, StatsSA created the panel after all surveys were conducted by reviewing the names in their records. Thus, the panel should be accurate, but mistakes could have been made, which should be kept in mind. The resulting metrics are reported in Table 5B below. The results from Pan (2022) are also reported here.

All matching algorithms perform worse in the LFS than in the QLFS, across all metrics except the negative predicted value. Pan's (2022) matching algorithm only matches 52% of the true matches from wave to wave. In fact, the method which correctly predicts the largest proportion of the true matches is the RD-expanded approach. This only predicts 71% of the true matches. This is a concern. If our panel quality is much lower in the 2004:2 to 2007:2 panel, then this could drive differences between the datasets in our results. However, this can be checked by using Pan's (2022) method with the 2001:2-2004:1 panel.

Table 5B: Matching Algorithm Accuracy Metrics

Metric	TPR	TNR	PPV	NPV
LFS – Pan (2022)'s alt-RD	52%	87%	72%	75%
RD Expanded	71%	82%	70%	83%
RD Strict	53%	87%	72%	75%
AKK	63%	88%	76%	80%
LSV	38%	89%	68%	70%
QLFS – Pan (2022)'s alt-RD	81%	99%	99%	73%

Source: Own calculations from StatsSA. RD – Ranchhod & Dinkelmann (2008), AKK -Anand, Kothari & Kumar (2016) , LSV – Leung, Stampini & Vencatachellum (2014). TPR – True Positive Rate, TNR – True Negative Rate, PPV – Positive Predictive Value, NPV – Negative Predicted Value. All proportions calulcated with the official StatsSA panel of the LFS between September 2001 and March 2004.

It also seems that two other algorithms perform better than Pan's (2022) approach. The RD expanded approach returns better TPR and NPV metrics than Pan's (2022) method, and the AKK method outperforms Pan's (2022) method in all metrics. However, I still choose to use Pan's (2022) method. I do not use the RD-expanded approach because it performs worse than Pan's (2022) approach according to the PPV. Although the difference is only 2 percentage points, this represents a difference of around 12 000 in the number of false positives. False positives will bias results. I do not use the AKK approach because it is based on StatsSA's person number variable. As Pan (2022) shows, the quality of this variable has been prone to change, and thus I cannot be sure that this variable does as well in the 2004:2 to 2007:2 data as it does in the 2001:2 to 2004:1 data. As well as this, using the same method in the LFS and the QLFS guards against any changes that might be driven by the change in method. Therefore, despite its average performance in the LFS, I use Pan's (2022) approach in both the LFS and QLFS.

Now that I have described in detail the panel data I use in this paper, I turn to describing how I will estimate the elasticity of labour supply to the firm using this panel data.

5. Methodology

So far, I have explained from the literature that studying monopsony is concerned with understanding the imperfections of the labour market. Estimating the wage elasticity of labour supply to the firm is one way of quantifying these imperfections. Having described the data, it is now necessary to explain how I go about estimating the wage elasticity of labour supply to the firm. I follow Manning (2003) closely in doing so.

5.1. Specifications

Manning (2003) has two methods of estimating the wage elasticity of labour supply to the firm. I refer to these as the simple and complex approach and use both in my analysis. The simple approach begins by regressing a binary variable indicating whether an employee has separated or not on the log wage and other controls, as modelled in (16) below.

$$Pr(s_{it} = 1) = \mathrm{G}(\beta_0 + \beta_1 lnw_{it-1} + \lambda'\boldsymbol{X}_{it-1} + \mathrm{u}_{it-1}) \tag{16}$$

s_{it} is a dummy variable indicating separation to both employment and non-employment, as compared to those remaining in employment with the same employer. G represents the logit function. lnw_{it-1} is the natural log of the hourly wage in the previous period, while $\boldsymbol{X}_{it-1}$ is a vector of personal and household characteristics included as controls. In the main analysis, specifications with three different sets of covariates are used. The first includes no other independent variables other than the log real wage. The second is a set of controls akin to what is used in Manning (2003), namely age, age squared, years of education, race, gender, marital status, wave and province. A third set of controls includes industry, occupation and a dummy indicating whether the individual works in the public sector or not. Given the preference for excluding tenure explained in the literature review and the measurement error in tenure discussed above, I do not include it in any specifications.

Although Manning (2003) uses hazard models to estimate the separations regressions above, I use logit models as in Booth & Katic (2011). I use logit models largely because hazard models are beyond the scope of my expertise, and because they are still commonly used in the literature, such as in Booth & Katic (2011). β_1 is obtained via the calculation of the average partial effects and then is divided by the mean of s_{it} to obtain the wage elasticity of separation.

In the simple approach, Manning (2003) proves under relatively strict conditions that the elasticity of labour supply to the firm is equal to twice the negative of the wage elasticity of separations from the firm.[10]

$$\varepsilon_{Nw} = -2\varepsilon_{sw} \tag{17}$$

In the more complex approach, Manning (2003) relaxes an important assumption that was used in deriving (17), that separations to and recruits from non-employment are independent of the wage. Thus, they do not affect the elasticity of the separations and recruits. However, this might not be the case. Separations to and recruits from non-employment could be related to the wage. For example, consider an individual who experiences significant disutility from working. Ignoring their wage, they may desire to not work at all. However, they may still not quit to non-employment if their wage is large enough to render them better off and so relaxing this assumption is therefore more realistic. Thus, Manning (2003) extends the model and assumes that separations to and recruits from non-employment are dependent on the wage. The implication of assuming this is that the wage elasticity of labour supply to the firm is made up of a weighted share of the various elasticities of recruitment from and separation to employment and non-employment. Equation (18) shows this.

$$\varepsilon_{Nw} = \theta_R \varepsilon^{e}_{Rw} + (1 - \theta_R)\varepsilon^{n}_{Rw} - \varepsilon^{e}_{sw} - (1 - \theta_s)\varepsilon^{n}_{sw} \tag{18}$$

Manning (2003) then shows that a suitably weighted average of the elasticity of recruitment from employment equals the negative of a suitably weighted elasticity of separation to employment.[11] However, Manning (2003) states this is not possible for the separation and recruitment elasticities to and from non-employment. The reason for this is that one firm's separation to non-employment doesn't directly correspond to a hire from non-employment. As such, the elasticities are not related. However, Manning (2003) does derive an estimate of the elasticity of recruitment from non-employment via the elasticity of recruitment from employment, which is presented in equation (19) below (eqn. 4.23 in Manning (2003)):[12]

[10] The derivations for this condition can be found on page 97 of Manning (2003).
[11] See page 99 of Manning (2003).
[12] See page 100 of Manning (2003).

$$\varepsilon_{Rw}^{n}(w) = \varepsilon_{Rw}^{e}(w) - \frac{w\theta_R'(w)}{\theta_R(w)[1-\theta_R(w)]} \tag{19}$$

That leaves one needing to estimate the parameters shown in (20). (20) is simply (19) substituted into (18), but with $-\varepsilon_{sw}^{e}$ substituted for ε_{Rw}^{e} and γ substituted for $\frac{w\theta_R'(w)}{\theta_R(w)[1-\theta_R(w)]}$. θ_e represents the share of all separations to employment. In equilibrium, this equals the share of hires from employment. This is why the share of hires from employment is not included.

$$\varepsilon_{Nw} = \theta_e(-\varepsilon_{sw}^{e}) + (1-\theta_e)(-\varepsilon_{sw}^{e} - \gamma) - \theta_e(\varepsilon_{sw}^{e}) - (1-\theta_e)(\varepsilon_{sw}^{ne}) \tag{20}$$

The following equations detail how the necessary parameters in (20) are estimated.

$$Pr(s_{it}^{e} = 1) = G(\beta_0 + \beta_1 lnw_{it-1} + \lambda' \boldsymbol{X}_{it-1} + \mathrm{u}_{it-1}) \tag{21}$$

$$Pr(s_{it}^{ne} = 1) \; = \; G(\beta_0 + \beta_1 lnw_{it-1} + \lambda' \boldsymbol{X}_{it-1} + \mathrm{u}_{it-1}) \tag{22}$$

s_{it}^{e} is a dummy indicating separations to employment, as detailed above, and s_{it}^{ne} is a dummy variable indicating separations to non-employment. The above regressions are estimated as logit models via maximum likelihood estimation. ε_{sw}^{e} and ε_{sw}^{ne} are obtained by dividing the average partial effect of β_1 in (21) and (22) by the means of the dependent variables. γ is obtained from (23) below, in which a dummy variable indicating whether an employee was hired from employment or not is regressed on the log wage.[13] The rest of the variables are exactly the same as in the specifications including separations to employment and non-employment.

$$Pr(nh_{it}^{e} = 1) = G(\beta_0 + \gamma lnw_{it} + \lambda_{it}' \boldsymbol{X}_{it} + u_{it}) \tag{23}$$

5.2. Use of Weights

As discussed in the data description, a combination of attrition and bracket weights are used in addition to the PALMS created cross-entropy weights found in the LFS and the regular StatsSA weights found in the QLFS. I will refer to these as the base weights. However, some specifics

[13] Manning (2003) shows on page 100 and in Appendix 4A that γ does indeed equal the furthest right hand side term in (34).

as to when which exact weights are used is necessary. Firstly, when running the regressions used in the main specifications above, the bracket weights adjusted for attrition are used. Consider a three-period sample, where the first period is denoted t. To adjust for attrition in period $t + 2$, one runs a probit regression in period $t + 1$ with a dummy indicating whether an individual is present in the following period. Then, using the fitted probability of being observed in the following period, one weights up the period $t + 2$ weight. The approach differs slightly in this paper. In our separation regressions, the separation variable is constructed in the current period based on whether the individual has separated from the previous period or not and thus the log real wage and all controls are lags (i.e. observed in the previous period). As such, using the weight from the previous period is the correct course of action. However, if I were to follow the usual attrition weighting approach, this weight would be adjusted for attrition from the period before that. In other words, if I ran a separation regression in $t + 2$ using a lagged weight (i.e. the $t + 1$ weight), the weight would be adjusting for attrition from t to $t + 1$. But in my separation regression, I want to adjust for attrition from $t + 1$ to $t + 2$. Thus, I weight up the base weights in the same period in which I ran the attrition regression, so that when I use the lagged weight in our separation regressions, I am adjusting for attrition in the correct wave.

There is one case in which this approach is altered to the normal attrition weighting approach. When regressing the binary variable indicating a new hire from employment on the log real wage in the current period, say $t + 1$, we do want to account for attrition from t to $t + 1$. Using the approach used in the separation regressions would result in us adjusting for attrition from $t + 1$ to $t + 2$. Thus, in these regressions, we use the attrition weights to weight up the base weights in the following period. In all non-earnings analysis, like calculating the mean of the dependent variable, I use the base weights adjusted for attrition, as opposed to using the bracket weights as well.

This section has detailed how I can use the relationship between separations and wages to estimate the elasticity of labour supply to the firm. Thus, how I identify separations is of crucial importance to an unbiased estimation of the elasticity of labour supply to the firm. In the next section, I explain why this is the case, and what options I have for identifying separations when measurement error is likely an issue.

6. Identifying Separations in the presence of Measurement Error

Identifying separations and new hires accurately is crucial when investigating the elasticity of labour supply to the firm. This section discusses possible methods for identifying separations and hires using survey data. I begin with the tenure approach used by Manning (2003) and Booth & Katic (2011). As there are some measurement error issues with tenure, I discuss an alternate approach used by Pan (2022) to identify job flows in the South African survey data context. This section focuses on identifying separations to and hires from employment. This is because separations to and hires from non-employment are easily identifiable using an employment status variable.

6.1. The Tenure Approach

6.1.1. Explanation

Manning (2003) uses tenure to identify hires and separations. If an individual's tenure is less than the time elapsed since the previous wave, then the individual is assumed to be in a new job. Booth & Katic (2011) also use this approach. Individuals are not directly asked about their tenure in the LFS or the QLFS. However, they are asked for the month and year they started working for their current employer. Therefore, one can construct a tenure variable using the survey and job start dates.

In the QLFS I have the exact survey date for 2018 and 2019 and so can accurately construct tenure length in months. For the LFS and the 2011 QLFS, however, I only know the month in which the survey was conducted. This is only an issue if interviews in the same wave were conducted in different months, as then the interval between interviews could differ across individuals. It appears that interviews in the LFS were conducted in a single month because the September 2001 LFS release states that the survey was "conducted in September 2001" (StatsSA, 2001:ii). Therefore, the monthly interval remains the same across all individuals. However, in the QLFS, StatsSA conducts interviews over the course of 3 months. I know this for two reasons. First, the 2011 Quarter 1 release states that the survey was "conducted in January-March 2011" (StatsSA, 2011). Second, I have the survey date variable in the 2018 and 2019 QLFS data and this shows that interviews were conducted evenly across the 3 months in the wave. This thus presents an issue in the 2011 QLFS. Without further information, the monthly intervals between interviews could differ across individuals and I would not know. However, the 2018 and 2019 data reveal that 99% of individuals were interviewed 3 months apart. Therefore, it is reasonable to assume that almost all individuals in the 2011 QLFS were interviewed exactly 3 months apart.

Under this assumption, the separation variable is constructed in the following way. The sample is all individuals employed in the previous wave. I observe a separation to non-employment if the individual is now in non-employment. I observe a separation to employment if the individual's tenure is less than the time passed since the previous wave. In the LFS, this is 6 months except in the second LFS of 2002, where it is 7 months. This is because the first LFS of 2002 was conducted in February and not in March (StatsSA, 2002). In the 2011 QLFS, it is either 3 months or 6 months, depending on whether I use a 1 or 2-wave interval. In the 2018 and 2019 QLFS, it is the difference in months between the survey dates observed, which for 99% of individuals is 3 months or 6 months depending on which interval I use.

6.1.2. Measurement Error

The accuracy of the tenure variable is important, as it enables the creation of the separation variable, which is my primary dependent variable. Measurement error is traditionally not an issue when it is in the dependent variable. However, it can be an issue when it is a binary response variable as shown by Hausman, Abrevaya & Scott-Morton (1998). These authors show that measurement error in a discrete-choice dependent variable causes downward bias in the regression coefficients when using a non-linear estimation approach such as a probit or logit.

There are a few ways measurement error in tenure can be analysed. Brown & Light (1992) provide a helpful analysis of measurement error in tenure and its impact on identifying job changes. They analyse the American PSID, which is one of the data sources used by Manning (2003). They analyse the internal consistency of the tenure variable to understand the extent to which tenure is likely mismeasured and how results differ when using different approaches to identify changes in jobs.

Brown & Light (1992) first identify separate jobs in a very similar manner to the method I discussed above and that I use below. If tenure is less than the time elapsed since the last interval, a new job is recorded. Then, within each job, they calculate the proportion of job start dates across wave that match. They then report the proportion of all jobs present in more than one wave that have a consistent start date across each wave. This serves as a measure of internal consistency in tenure. They also include measures of internal consistency that allow for 3 and 6-month differences in the job start date within a job and across waves. In the PSID between 1976 and 1985, only 7% of jobs had a consistent job start date across wave (Brown & Light, 1992). 35% of jobs had start dates within 3 months of each other, while 46% of jobs had start dates within 6 months of each other (Brown & Light, 1992). Therefore, tenure in the PSID between 1976 and 1985 was highly inconsistent.

The Brown & Light (1992) measure of internal consistency is imperfect because it relies on a mismeasured tenure variable to identify different jobs. Therefore, it may be incorrectly differentiating jobs. Thus, if mismeasurement in tenure is causing an incorrect differentiation of jobs, this measure of internal consistency will not pick it up and thus likely underestimates true total measurement error. Nevertheless, because the measure still informs on the internal consistency of tenure, it gives an idea of how badly tenure is mismeasured.

Brown & Light (1992) do not discuss what this measurement error does to the separation rate. Unfortunately, it is unclear and depends on whether tenure is over or under-estimated. Consider the following examples. If tenure is generally overestimated, then I am less likely to see tenure values of less than 6 months. Therefore, I should identify fewer separations if identifying separations over a 6-month period. On the other hand, if tenure is generally underestimated, then I should see too many tenures below 6 months. I would therefore identify too many separations. However, tenure may not be biased in one direction, and thus it is difficult to know how it will impact the separation rate.

Brown & Light (1992) do however discuss how this might impact regression analysis. The authors' findings also show that the way they use tenure to identify job changes impacts the coefficient on wage in a regression of job separations on wage. Although all their proposed methods use tenure to identify job changes, the conditions for how they identify jobs are varied across method. Their recommended approach is the one I discuss above and use below. However, they show that their other methods, which are less stringent in their conditions, result in lower coefficients on wage in a separations regression, meaning that how I identify separations likely matters for the primary coefficient of interest.

Some other simple methods can identify the extent to which tenure is mismeasured. Firstly, any tenure values less than 0 are mismeasured. Secondly, any tenure values that exceed an individual's age minus the minimum working age are likely mismeasured. For example, a 20-year-old individual is unlikely to have been working for 10 years. Thirdly, someone who was in non-employment in the previous wave should have a tenure less than the wave interval. If their tenure exceeds the wave interval, it is likely mismeasured. I can then calculate the proportion of all tenure observations with the above characteristics. I conduct these analyses in addition to the Brown & Light (1992) analysis below.

6.1.3. Analysis

Table 6A gives the proportion of jobs with consistent job start dates across each of the individual datasets. A 6-month interval is used for both the LFS and QLFS. The table shows that job start dates are much more consistent in the QLFS than they are in the LFS. 73% of the jobs identified in the QLFS have the same job start date, whereas only 18% of jobs in the LFS have a consistent start date across wave. This is a remarkable difference in internal consistency. The consistency in the QLFS is very high, even when compared to the literature. Brown & Light (1992) find an internal consistency of 7% in the PSID, while Bergin (2015) finds an internal consistency of 20% in Ireland. The constructed LFS panel used from 2004:2 to 2007:2 is slightly more consistent than the official StatsSA panel from 2001:2 to 2004:1. If I use the matching algorithm for the 2001:2-2004:1 panel instead of the official StatsSA panel, the exact consistency increases slightly to 18%.

Table 6A: Proportion of Jobs with Consistent Start Dates

	LFS 0104	LFS0407	LFS Avg	QLFS 11	QLFS 18/19	QLFS Avg
Exact	15%	21%	18%	74%	73%	73%
Within 3 Months	25%	25%	29%	79%	79%	79%
Within 6 Months	31%	39%	35%	82%	82%	79%
Jobs		40630			36964	

Source: Own calculations from StatsSA. LFS – Labour Force Survey. QLFS - Quarterly Labour Force Survey. New jobs defined by tenure being less than the time elapsed between waves. 6 Month intervals used in both surveys. Period Labels: 0104: 2001:2-2004:1. 0407: 2004:2-2007:2. 11: 2011. 1819: 2018-2019.

There are a few possible reasons for the differences in consistency in job start dates. Firstly, because I am only using a 6-month interval in the QLFS despite it being a quarterly survey, I observe a maximum of two periods per job. This is because dwellings only remain in the QLFS for four quarters. In the LFS I observe as many as 6 because only 20% of dwellings in the LFS are rotated out every 6 months. The more waves, the less likely it is that the job start date will be consistent across all of them. However, when using a quarterly interval in the QLFS and observing jobs across as many as 4 quarters, the internal consistency in the QLFS remains very high. This reason therefore seems unlikely.

Secondly, as discussed above, Brown & Light's (1992) measure is imperfect as it uses a mismeasured tenure variable to identify different jobs. Therefore, if the QLFS identifies more jobs per wave and fewer waves per job than it should, I may be overinflating the relative internal consistency of the QLFS. For example, assume that a job that lasts 4 waves is erroneously separated into two jobs in the QLFS. Let us also assume that the job start dates correspond within each of the erroneous job-wave pairs. However, assume that they do not correspond across all 4 waves. Then, using the Brown & Light (1992) measure inflates the internal consistency of tenure because it considers the two jobs-wave pairs as internally consistent. However, consistency in the QLFS remains high whether a 3 or 6-month interval is used. In the example above, a 6-month interval would exhibit a lower consistency, because job start dates would not match across waves with a 6-month interval. Therefore, I do not think this is the reason.

Thirdly, it may simply be down to a smaller time interval in the QLFS. A smaller interval may enable individuals to better remember how they responded to the previous survey. This would increase consistency. I cannot, however, confirm this is the case.

Lastly, it could be due to differences in panel quality. For example, if two different people are matched to each other across wave, then they are very likely to have different tenures. As discussed in the previous section, the panel quality in the QLFS is higher than in the 2004:2-2007:2 LFS. However, the official StatsSA panel between 2001:2 and 2004:1 is slightly less consistent than the 2004:2-2007:2 matched panel. Further, if I use the matching algorithm in the 2001:2-2004:1 data, the panel consistency increases. Therefore, if panel quality is the issue, this would suggest that the official StatsSA panel in 2001:2-2004:1 is of poor quality. If it is not in fact of poor quality, then it is likely not differences in panel quality driving the differences in consistency between the LFS and QLFS.

It is therefore unclear why the QLFS is so much more internally consistent. Although I have no way of telling, perhaps StatsSA implemented some sort of imputation based on responses in other waves. As I cannot tell, I simply assume the QLFS is more internally consistent than the LFS.

The way in which the internal inconsistency of the tenure variable affects separations likely depends on the direction in which tenure is biased. To investigate this, I calculate the difference between a likely mismeasured tenure value and its likely true value. I can identify these true values by identifying jobs in which the majority of waves have the same job start date. Then, by looking at whether the supposedly mismeasured job start date is before or after the actual date, I gain an indication of the direction in which tenure is biased. As Figure 1A and 1B show below, there is no clear direction in which the job start dates are mismeasured, meaning that there is no definite direction in which tenure is biased. Therefore, there is little I can say on whether I believe that the number of separations will be under or over-estimated.

Figure 1A: Difference between Mismeasured and Mode Tenure by Job: LFS

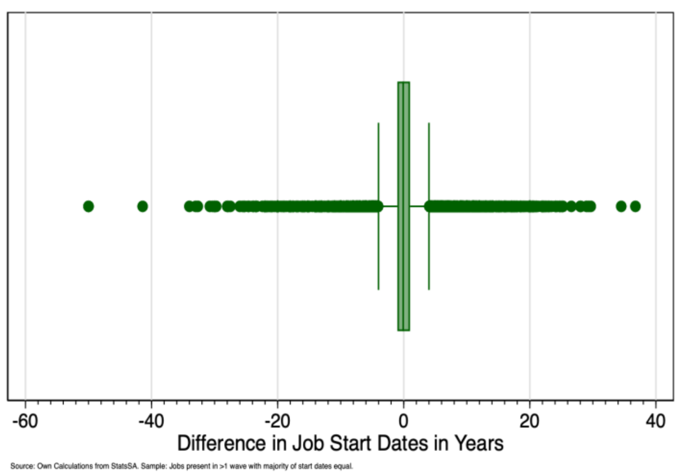
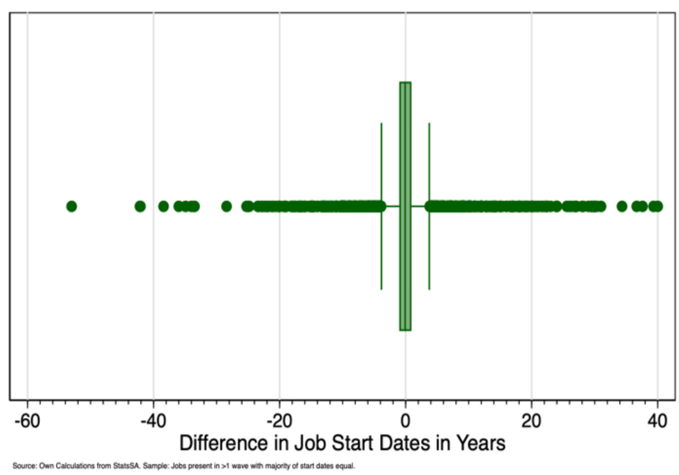


Figure 1B: Difference between Mismeasured and Mode Tenure by Job: QLFS



In Table 6B below, I consider the rest of the ways in which mismeasurement may be identified. Slightly more observations in the LFS have a tenure less than 0 and more than age minus 15. Of all new hires from non-employment, tenure exceeds 6 months for 57% and 54% of the observations in the LFS. This suggests substantial mismeasurement of tenure. In the QLFS, the same proportions are 42% and 33% respectively and are thus smaller but still substantial. For my analysis in the rest of this paper, I exclude those observations with tenure less than 0. I also exclude those whose tenure exceeds the amount of time someone could have been of working age. For those who were previously unemployed and now exhibit a tenure more than the wave interval, I recode tenure to be less than the wave interval.

This section has shown that measurement error is a concern when using tenure to identify hires and separations, more so in the LFS than the QLFS. However, because the direction of this mismeasurement is unclear, I cannot be sure whether the differences in tenure inconsistency will increase or decrease separations. The results of Hausman, Abrevaya & Scott-Morton (1998) and Brown & Light (1992) imply that this mismeasurement in tenure will likely bias the coefficients on wage in the separations regressions downwards. An alternative method of identifying hires and separations is therefore explored next.

Table 6B: Other Indicators of Mismeasurement

	LFS 01/04	LFS 04/07	QLFS 11	QLFS 18/19
Proportion of observations with tenure < 0	0,2%	0,2%	0,0%	0,0%
Proportion of observations with tenure > age minus 15.	1,4%	1,1%	0,4%	0,4%
Proportion of new hires from non-employment with tenure > 6m	57%	54%	42%	33%

Source: Own Calculations from StatsSA. LFS – Labour Force Survey. QLFS - Quarterly Labour Force Survey. 6m – 6 months. Period Labels: 0104: 2001:2-2004:1. 0407: 2004:2-2007:2. 11: 2011. 1819: 2018-2019.

6.2. The Firm Characteristics Approach

6.2.1. Explanation

Pan (2022) used changes in industry, public/private sector status and formality to identify job flows in addition to tenure changes. The intuition behind this method is that firms are unlikely to change their industry, sector or formality. Therefore, any changes observed in these variables for an individual must imply a change in the job and thus a separation. This method alone fails to identify job-to-job changes where the industry, sector and formality all remain the same. Pan (2022) therefore uses it in addition to the tenure approach described above.[14] Replicating Pan's (2022) approach is possible in the QLFS, but not exactly possible in the LFS. Pan (2022) uses a combination of unemployment contributions and registration for income tax to infer the formality of a worker. However, the LFS does not contain a tax registration variable for all waves. Although the LFS does ask individuals directly about their formality status, the changes between this variable in the LFS and a constructed variable ala Pan (2022) in the QLFS are very large. I therefore choose not to use it. This is not a significant issue, as identifying changes via formality is not definitely necessary in my research. If an individual has stayed in the same job, but the firm has changed formality, then I would not want to identify that as a separation. Therefore, I only use industry and sector to identify changes in firm characteristics.

[14] Pan's (2023) use of the tenure approach is more restrictive than mine because he does not assume that 3 months have elapsed between interviews in the QLFS. He only identifies separations where the job start date occurs within the quarter that a person is interviewed in.

6.2.2. Analysis

It is important to understand to what extent the firm characteristics approach alters the tenure approach. The firm characteristics approach should only be used in addition to the tenure approach because it cannot identify job-to-job changes where firm characteristics have not changed. I therefore analyse the implications of this approach by comparing it to the tenure approach. This is done by analysing the consistency of these two variables within the jobs identified by the tenure approach. Table 6C below reports the consistency of firm characteristics within the jobs identified by the tenure approach discussed above. I also include consistency in occupation and UIF status, as these should both remain relatively consistent across wave and within a job.

Table 6C: Consistency of Firm Characteristics within Job

	LFS 0104	LFS 0407	LFS Avg	QLFS 11	QLFS 18/19	QLFS Avg
Industry	78%	77%	78%	92%	92%	92%
Sector	92%	93%	92%	97%	97%	97%
UIF	64%	70%	67%	90%	89%	89%
Occupation	60%	62%	61%	85%	87%	86%

Source: Own calculations using StatsSA. LFS – Labour Force Survey. QLFS - Quarterly Labour Force Survey. New jobs defined by tenure being less than the time elapsed between waves. 6 Month intervals used in both surveys. Period Labels: 0104: 2001:2-2004:1. 0407: 2004:2-2007:2. 11: 2011. 1819: 2018-2019.

All firm characteristics are more consistent in the QLFS than in the LFS. In the LFS, only 78% of jobs had the same industry across waves, whereas this was 92% in the QLFS. This indicates that identifying separations by changes in industry will substantially increase the number of separations, more in the LFS than the QLFS. Unsurprisingly, consistency is much higher when the sector variable is analysed and the difference in consistency between the two surveys is also lower. UIF and occupation in the LFS are quite inconsistent, while in the QLFS they are slightly less consistent than industry and sector.

However, industry and sector may also be mismeasured. For example, of all 22% of jobs with inconsistent industry in the LFS, it is unlikely that all of those have mismeasured tenure. I can investigate this by comparing the within-job consistency of tenure and industry. Tables 6E and 6D report a cross-tabulation of job start date consistency and industry consistency within jobs identified by the tenure approach. Cells in green show the totals. Cells in yellow show the column percentage, which is the proportion of those within each category of tenure consistency who have consistent industry or not. Cells in orange show the row percentage, which is the proportion of those within each industry consistency category who have consistent tenure or not. Lastly, cells in pink show the cell percentage, which reports the proportion of all observations in each tenure-industry consistency category. In the case of the above example, 12% of jobs present in more than one wave with inconsistent industry have a consistent tenure, which suggests that industry and not tenure may be mismeasured.

Table 6D: Matching Tabulation for the jobs present in > 1 wave in the LFS

Match Type	Tenure Match					
	Category	No		Yes		Total
Industry Match	No	**7868**	**88%**	**1120**	**12%**	**8988**
		24%	**19%**	**16%**	**3%**	**22%**
	Yes	**25562**	**81%**	**6080**	**19%**	**31642**
		76%	**63%**	**84%**	**15%**	**78%**
	Total	**33430**	**82%**	**7200**	**18%**	**40630**

Source: Own calculations from StatsSA. Green cells show the total, orange cells the row percentage, yellow cells the column percentage, and pink cells the cell percentage. Sample: all jobs present in more than one wave. New jobs defined by tenure being less than the time elapsed between waves. 6 Month intervals used in both surveys.

Table 6E: Matching Tabulation for the jobs present in > 1 wave in the QLFS

Match Type	Tenure Match					
	Category	No		Yes		Total
Industry Match	No	**1220**	**39%**	**1879**	**61%**	**3099**
		12%	**3%**	**7%**	**5%**	**8%**
	Yes	**8644**	**26%**	**25221**	**74%**	**33865**
		88%	**23%**	**93%**	**68%**	**92%**
	Total	**9864**	**27%**	**27100**	**73%**	**36964**

Source: Own calculations from StatsSA. Green cells show the total, orange cells the row percentage, yellow cells the column percentage, and pink cells the cell percentage. Sample: all jobs present in more than one wave. New jobs defined by tenure being less than the time elapsed between waves. 6 Month intervals used in both surveys.

The stark contrast in the internal consistency of the QLFS and the LFS is revealed by the cell percentages in pink. While only 15% of jobs in more than one wave have a matching tenure and industry in the LFS, 68% do so in the QLFS.

Thus, using the employer characteristics approach to identify separations will likely increase the separation rate substantially because industry is relatively inconsistent within jobs identified by the tenure approach. This increase should be larger in the LFS, as industry is more inconsistent within the jobs identified by tenure in the LFS as compared to the QLFS. Industry is also inconsistent in some jobs where tenure is consistent, which suggests that industry could also be mismeasured. Therefore, in the next subsection, I calculate the worker flows produced by each method, and compare to the literature to understand which I believe to be more accurate.

6.3. Comparison of Worker Flows across Methods

I begin with calculating the separation rates using both methods, and report these in Table 6F below. A 6-month interval in both the LFS and QLFS is used below.

Table 6F: Separation Rates by Dataset and Method

	LFS 0104	LFS 0407	QLFS 11	QLFS 1819
Tenure	19%	19%	10%	11%
Industry	32%	34%	16%	16%
Sector	20%	18%	10%	10%
All 3	38%	40%	20%	21%

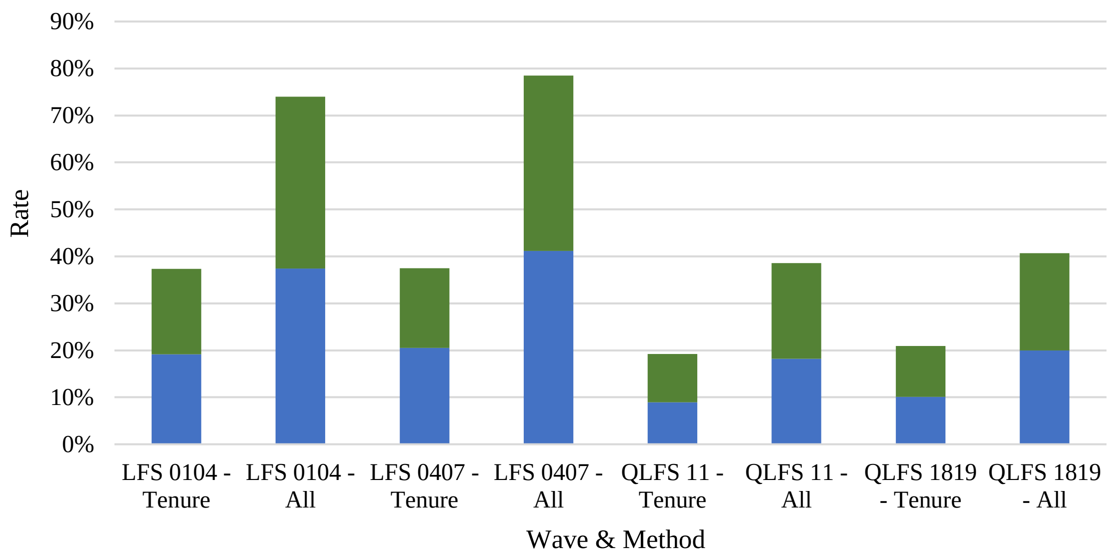
Source: Own Calculations from StatsSA. LFS – Labour Force Survey. QLFS – Quarterly Labour Force Survey. 6-month intervals are used for both the LFS and QLFS. Survey weights used. Period Labels: 0104: 2001:2-2004:1. 0407: 2004:2-2007:2. 11: 2011. 1819: 2018-2019.

The separation rate in the LFS is higher in the QLFS across all methods. In all datasets, the separation rate in the LFS is around double the separation rate in the QLFS. The use of tenure and sector together doubles the separation rate. The separation rate in the LFS when using tenure, industry and sector appears very high. It seems unlikely that 38% of those employed 6 months ago are either in a different job or in non-employment 6 months later. For example, consider that in Bassier (2023), the yearly separation rate using administrative tax data is 37%. If the LFS, which is using a 6-month interval, is exhibiting a higher separation rate than yearly data, this is likely overestimated. Although Bassier (2023) uses later data, it seems unlikely that the separation rate could change so drastically within that time.

These differences are also evident in the average worker flow rates reported in Figure 2 below. I calculate this by dividing the sum of separations and hires in a period by the average total employment over the previous and current periods. Average worker flow rates are between 70% and 80% in the LFS, which seem very high. In the QLFS, they remain around 40%. When only using tenure, the worker flow rates of around 20% in the QLFS seem very low. External evidence reveals this to be the case. Kerr (2018) finds an average worker flow rate of 53%

using tax data with a yearly interval, whereas we use a half-yearly interval.[15] Therefore, when using tenure alone, the LFS estimates appear more reasonable if not a little small, and the QLFS estimates appear too small. When using the tenure, industry and sector approach, the LFS estimates appear too large, and the QLFS estimates appear more reasonable.

Figure 2: Worker Flow Rates by Dataset and Method



Source: Own calculations from StatsSA. Survey weights used.

In the discussion on measurement error above, it appeared as if tenure was more mismeasured in the LFS than in the QLFS. Yet here, when using tenure alone, the QLFS estimates of worker flows do not seem to square with external evidence in the form of the tax data in Kerr (2018).

6.4. Summary and Solution

The issues discussed above make it difficult to decide how to proceed. The issues with each of the approaches are as follows. If I use tenure alone to identify separations, I risk ignoring that tenure is very inconsistent in the LFS. I also produce a total number of separations and hires in the QLFS that appears too low when compared to external evidence. If I use tenure and industry and sector, I produce a total number of separations in the LFS that appears too large. In the

[15] It should be noted that although Kerr (2018) does not report a separation rate, his worker flow rates imply a separation rate of around 26,5%, if I assume that the separation rate is half the worker flow rate. This is much lower than the separation rate found by Bassier (2023) using similar data. Exploring the reason for the differences in these two should be an avenue for future research.

QLFS, I likely introduce more measurement error to a separation variable based on a relatively internally consistent tenure variable.

I decide to use the tenure-alone approach for the following reasons. First, Pan's (2022) justification for using changes in employer characteristics is that the tenure approach is imprecise. This is because Pan (2022) could not observe the exact survey dates. In this paper, for 2018 and 2019, I have the survey dates. They indicate that using exactly a quarter or 6-month interval in the QLFS is reasonable. Therefore, the lack of survey dates in the 2011 QLFS is not a significant issue. In the LFS, all surveys were conducted in a single month, and therefore the lack of survey dates is also not a significant issue. Imprecision in tenure due to varying survey dates is thus not a major issue.

Second, the literature so far has overwhelmingly used the tenure approach and there are few examples in the literature that use changes in employer characteristics to identify changes in jobs, and thus I do not find strong enough reason to depart from the tenure approach.

Third, the LFS worker flows are extremely high for a survey with a 6-month interval when using tenure, industry and sector. They are especially high compared to the numbers found in Kerr (2018), who uses a yearly interval. Using a yearly interval in the LFS would increase the average worker flow rate even more, to a point at which it would be implausibly large. There is therefore likely measurement error in the industry variables too. I, therefore, believe that using industry and sector will add too much noise to the separations variable. The QLFS numbers using tenure alone appear a little low. However, this is more easily explainable than the overestimation in the LFS. For example, I do not identify changes in informality, which Kerr (2018) implicitly does and I use a 6-month interval, as opposed to a yearly interval in Kerr (2018). Lastly, the apparent high internal consistency of the QLFS gives me some confidence that measurement error in tenure may not be that bad. I therefore proceed with some descriptive analysis using the tenure-alone approach.

7. Descriptive Statistics

7.1. Key Variables

Table 7A below presents descriptive statistics for important variables across the two datasets. Although the QLFS is a quarterly survey, a 6-month interval is used below to provide clear comparability with the LFS. The separation rate decreases from 19% in the LFS to 11% in the QLFS. As discussed above, it is unclear whether the greater inconsistency in tenure in the LFS is driving this or not. Both the separation rates to employment and non-employment are larger in the LFS. Tenure is slightly larger in the LFS, which is surprising given that the separation rate is higher. Most other variables are similar across dataset, except for the average real hourly wage. This increases from R41,80 an hour to R55,27. This is a 30% increase in the average real wage, which although substantial is unsurprising because an upward trend in real wages has been shown by Wittenberg (2017).

Table 7A: Descriptive Statistics for Important Variables across Dataset

Variables	LFS - 6 Month	QLFS - 6 Month
Separation All	0,19	0,11
Separation to Employment	0,06	0,04
Separation to Non-Employment	0,14	0,07
Female	0,44	0,46
African	0,66	0,74
Coloured	0,14	0,12
Indian/Asian	0,04	0,03
White	0,17	0,11
Age	38,18	38,53
Years of Education	9,63	11,03
Married	0,59	0,51
Union Member	0,34	0,31
Publicly Employed	0,24	0,22
Tenure	7,75	6,91
Real Hourly Wage	41,80	55,27

Source: Own Calculations from StatsSA. Q/LFS – Quarterly/Labour Force Survey. Interval of 6 months used for both the LFS and QLFS. Survey weights used.

7.2. Separation Rates across Key Variables

Table 7B below reports how separation rates differ across important variables. In most categories, both the separation rates to employment and non-employment are larger in the LFS. The separation rates are very similar between men and women and across race, Africans and Coloureds have the highest separation rates, followed by Indians /Asians, and then Whites. The largest differences across race are observed for separations to non-employment. There is also a large increase in the separation rate to non-employment for the 55-64 age category in the LFS, likely due to retirement. Across education, only those with a tertiary education have a lower separation rate in the LFS, while in the QLFS, those with a matric do as well.

Table 7B: Separation Rates across Key Variables

		Separation Rate					
		All		To Employment		To Non-Employment	
Variable	Panel	LFS	QLFS	LFS	QLFS	LFS	QLFS
Gender	Male	0,18	0,11	0,06	0,04	0,12	0,07
	Female	0,20***	0,11	0,05***	0,04**	0,16***	0,07
Race	African	0,21	0,12	0,06	0,05	0,16	0,08
	Coloured	0,19***	0,13	0,07***	0,05	0,13***	0,08
	Indian/Asian	0,17***	0,05***	0,06	0,02***	0,13***	0,03***
	White	0,12***	0,03***	0,05***	0,02***	0,08***	0,01***
Age Category	15-24	0,40	0,23	0,15	0,08	0,30	0,16
	25-34	0,21***	0,13***	0,07***	0,05***	0,15***	0,08***
	35-44	0,15***	0,09***	0,05***	0,04***	0,10***	0,05***
	45-54	0,12***	0,07***	0,03***	0,03***	0,10***	0,05***
	55-64	0,19***	0,08***	0,03***	0,02***	0,17***	0,06***
Education Category	None	0,21	0,16	0,06	0,07	0,17	0,10
	Some Primary	0,23***	0,16	0,07	0,07	0,17	0,11
	Some Secondary	0,24**	0,15	0,07*	0,06	0,18	0,10
	Matric	0,18***	0,10***	0,06	0,03***	0,13***	0,06***
	Some Tertiary	0,08***	0,03***	0,04***	0,02***	0,05***	0,02***
Real Hourly Wage	1-3	0,39	0,26	0,12	0,10	0,31	0,18
	4-14	0,28***	0,19**	0,08***	0,07	0,21***	0,12**
	15-35	0,19***	0,13***	0,06***	0,06*	0,14***	0,08***
	36-92	0,09***	0,06***	0,03***	0,03***	0,06***	0,03***
	93-313	0,07***	0,03***	0,03***	0,01***	0,04***	0,01***
	313 & above	0,03***	0,03***	0,00***	0,02**	0,02***	0,02***

Source: Own calculations from StatsSA. Q/LFS – Quarterly/Labour Force Survey. An interval of 6 months is used in both the LFS and the QLFS. Survey weights used. Significance stars indicate significant difference from first category. Significance Levels: *** < 0,01. ** < 0,05. * < 0,1. Real Hourly Wage categories chosen from the brackets used by StatsSA in the QLFS.

Lastly, the separation rate decreases as the real monthly earnings increases. Those who earn between R4 and R14 an hour have a separation rate of 28% and 21% in the LFS and QLFS respectively. The separation rates for those who earn between R93 and R313 and hour are 7% and 3% in the LFS and QLFS respectively. I therefore expect to find a negative relationship between separations and the real hourly wage in the regressions below.

I can also describe the relationship between wages and separations by plotting Kernel Density estimates for wages across the separation variable. This can be found in Figures 3A and 3B below. In both the LFS and the QLFS, the distribution of log real wages for those who separate is to the left of the distribution for those who do not. Therefore, those at lower wage levels appear more likely to separate.

Figure 3A: Wage Distribution across Separation: LFS

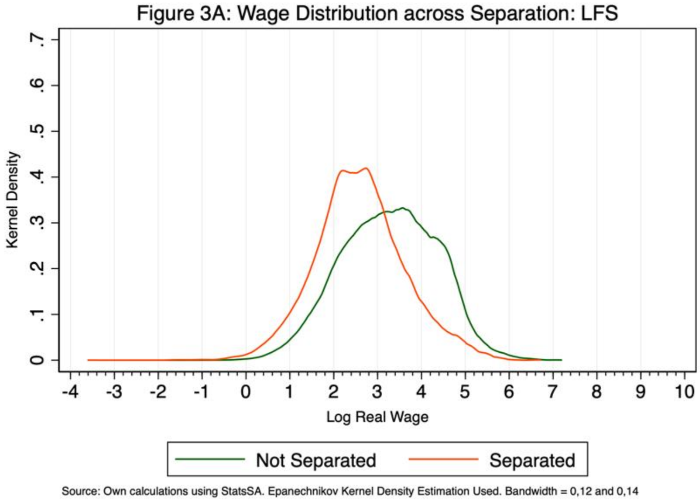
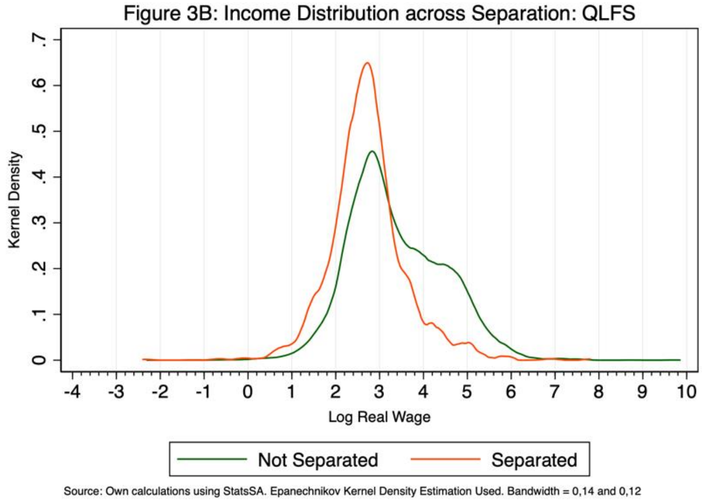


Figure 3B: Income Distribution across Separation: QLFS

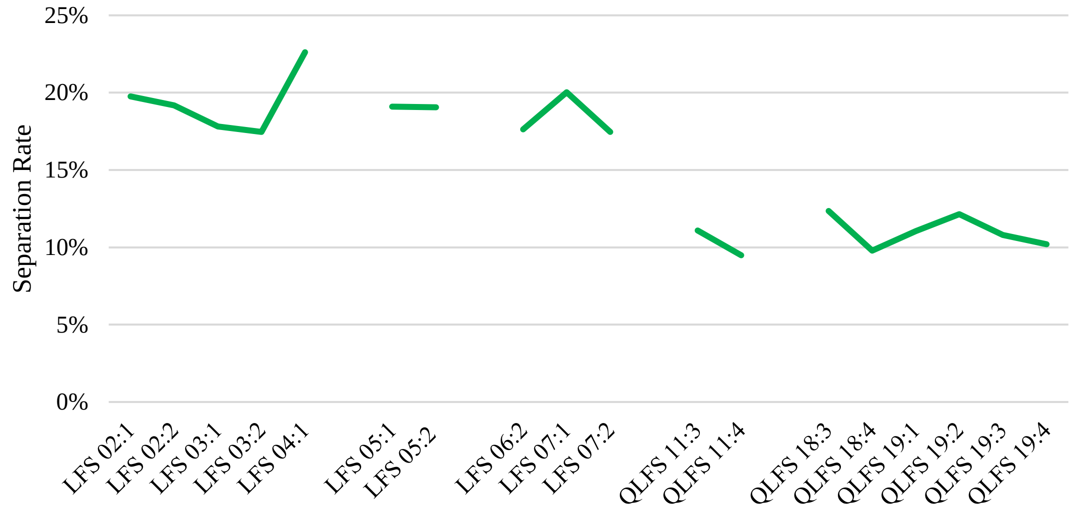


I do, however, also see some differences in the distribution of income across the two surveys. In the QLFS, we see a sharper spike in the distribution around the log wage of 2,5 to 3. However, in the LFS, this distribution is a little more spread out with less sharp spikes. The LFS also has a longer tail to the left, indicating more small wage values. The QLFS, on the other hand, has a longer tail to the right, indicating more large wage values. There are especially long tales for those who separate in the LFS, and those who don't separate in the QLFS.

7.3. Separation Rates and Worker Flows across Wave

Figure 4 below displays how the separation rates vary across wave, with the gaps denoting where waves are not consecutive. Several observations are worth mentioning. First, the separation rates do not differ much within each survey and across sub-dataset. There is a large drop in the separation rate between the LFS and QLFS.

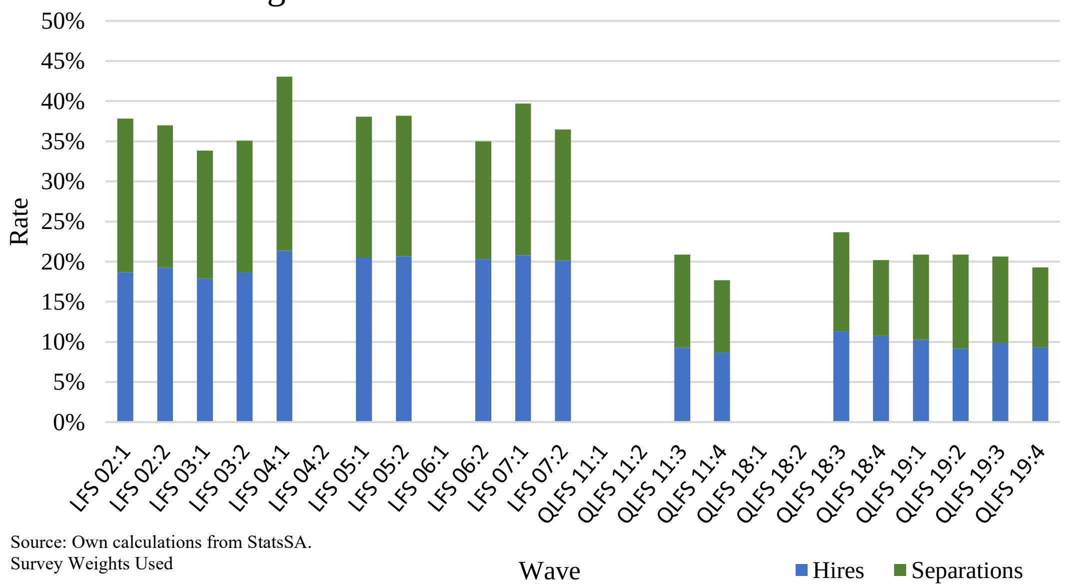
Figure 4: Separation Rate across Wave



Source: Own calculations from StatsSA. Survey Weights Used

As discussed in the identifying separations section above, I can also analyse worker flows across waves. Figure 5 below reports the average worker flow rate across each wave, again showing the worker flow rate is much higher in the LFS than in the QLFS. It seems unlikely that such a large drop in the worker flow rate has occurred, and thus some of this difference is likely due to measurement error.

Figure 5: Worker Flow Rate across Wave



Source: Own calculations from StatsSA. Survey Weights Used

7.4. Hires from Non-employment

As discussed in the theory section above, the proportion of hires from non-employment is a simple 'back of the envelope' measure for the level of monopsony. The theory suggests that the higher the proportion of hires from unemployment, the more monopsonistic a labour market is. Therefore, I investigate how this measure varies across some important variables.

These proportions are reported in Table 7C below, with the significance of a t-test to the base category included. The proportion of hires from non-employment is higher in the LFS at 0,72 than in the QLFS at 0,61. In almost all categories the proportion of hires from non-employment is higher in the LFS than in the QLFS. This suggests that the level of monopsony in South Africa has decreased over time. It is unlikely that this difference is driven by poor panel quality in the 2004-2007 panel due to the matching algorithm, as both the 2001:2-2004:1 and 2004:2-2007:2 panel yield very similar estimates. However, poor panel quality in both the official StatsSA 2001:2-2004:1 panel and the matched 2004:2-2007:2 panel could drive this difference.

Females, Africans and younger individuals exhibit higher proportions than men, Whites and those in the middle-aged categories. The higher female proportion squares with the literature when it suggests women face higher monopsony power than men. The historical disadvantage faced by Africans suggests that they would be subject to more monopsony power than Whites, which this table also shows. It is also unsurprising that this measure suggests young people may be more subject to monopsony power than middle-aged individuals. This is because they are new to the workforce with little experience or on-the-job skills.

A lower proportion of those with some tertiary education are hired from unemployment in the LFS. There are no other statistically significant differences across education in either the LFS or QLFS. That more skilled individuals face less monopsony power is unsurprising. Lastly, in the LFS, a lower proportion of high-wage earners are hired from non-employment as compared to the base category. This has implications for the logit regression of the hired from employment dummy on wage. It suggests that log wage will have a positive coefficient. However, in the QLFS, none of the larger wage categories are significantly smaller than the base category. In fact, there is no clear pattern in how the proportion changes across income category.

Table 7C: Proportion of New Recruits from Non-Employment

		LFS		QLFS	
Variable		Proportion	Significance	Proportion	Significance
Overall		0,72		0,61	
Gender	Male	0,68		0,59	
	Female	0,77	***	0,63	**
Race	African	0,75		0,62	
	Coloured	0,67	***	0,55	**
	Indian/Asian	0,67	**	0,49	
	White	0,62	***	0,50	**
Age Category	15-24	0,80		0,73	
	25-34	0,71	***	0,60	***
	35-44	0,65	***	0,58	***
	45-54	0,73	***	0,55	***
	55-64	0,78		0,54	***
Education Category	None	0,73		0,59	
	Some Primary	0,70		0,58	
	Some Secondary	0,74		0,61	
	Matric	0,75		0,62	
	Some Tertiary	0,61	***	0,57	
Real Hourly Wage	1-3	0,77		0,65	
	4-14	0,74		0,64	
	15-35	0,73	*	0,60	
	36-92	0,60	***	0,53	
	93-313	0,44	***	0,50	
	313 & above	0,37	**	0,80	

Source: Own calculations from StatsSA. Significance Levels: *** < 0.01. ** < 0.05. * < 0.1. Significance indicates the statistical significance of the difference between the relevant category and the first (base) category. Q/LFS – Quarterly/Labour Force Survey. 6-month interval used for both surveys. Survey weights used.

7.5. Summary Statistics from 3-Month & 1-Year Intervals

Table 7D below reports a summarised version of above statistics using a 3-month interval in the QLFS and a 1-year interval in the LFS. This enables better comparison to the literature. Manning (2003) finds that the yearly separation rates are 21% in the PSID (US), and 19% in the BHPS (UK). Table 8D shows that South Africa had a higher yearly separation rate of 25% in the early 2000s. In the UK LFS, which is a quarterly survey, the separation rate is 6%, whereas in the QLFS it is 7%. There is thus evidence that the separation rate is generally higher in South Africa than in the UK and USA. This differs to Kerr (2018) who finds that worker flows in South Africa were lower than in the USA. The yearly separation rate in the LFS is also substantially higher than the 14% found by Booth & Katic (2011) in Australia. Compared to previous South African evidence, the yearly separation rate of 25% is smaller than the 37% found by Bassier (2023). These are calculated over different time periods and thus are not directly comparable, but it does suggest that using survey data underestimates the separation rate.

Table 7D: Descriptive Statistics by Alternate Intervals

Variables	LFS - 1 Year	QLFS - 3 Month
Separation All	0,25	0,07
Separation to Employment	0,10	0,02
Separation to Non-Employment	0,16	0,06
Female	0,45	0,46
African	0,65	0,73
Coloured	0,14	0,12
Indian/Asian	0,04	0,03
White	0,17	0,12
Age	38,08	38,53
Years of Education	9,64	11,00
Married	0,58	0,50
Union Member	0,34	0,31
Publicly Employed	0,24	0,21
Tenure	7,75	6,88
Real Hourly Wage	41,53	54,54
Proportion of Hires from Non-employment	0,66	0,78

Source: Own calculations from StatsSA. Q/LFS – Quarterly/Labour Force Survey. Survey weights used.

The other difference to Manning (2003) and Booth & Katic (2011) is that the separation rate to non-employment in South Africa is higher than the separation rate to employment. In the PSID and BHPS, the separation rate to employment is 12%, and to non-employment it is 7% and 6% respectively. In South Africa, the separation rate to employment in the yearly LFS is 10%, and 17% to non-employment. In the UK LFS, the separation rate to employment is 3,2% and to non-employment is 2,5%. In the QLFS, it is 2% and 6% respectively. These results are unsurprising. South Africa's very high unemployment rate leads us to expect a higher separation rate to non-employment.

For hires from non-employment, the proportion is 71% in the PSID and 64% in the BHPS. The proportion in the LFS is 66% which is quite similar to the values in Manning (2003). This suggests little difference in the extent of monopsony. The proportion of hires from non-employment is 49% in the LFS (UK). In the South African QLFS, it is 78%, which suggests a higher level of monopsony in South Africa as compared to the UK.

7.6. Summary

The descriptive analysis suggests the following conclusions. First, the separation rate is higher in the LFS than in the QLFS. It is unclear whether this is due to differences in the mismeasurement of tenure in the two datasets. Second, South African separation rates appear larger than in the developed countries studied by Manning (2003) and Booth & Katic (2011). Third, although separation rates do not differ across gender, they do across race. African individuals have the highest separation rates, followed by Coloured, Indian and Asian, and then White individuals. Fourth, the separation rate decreases as the wage category increases. This leads me to expect a negative coefficient in the separations regressions below. Lastly, the proportion of hires from non-employment is higher in South Africa than in the UK. It is also higher for women, Africans and those without a tertiary education. This suggests a higher level of monopsony faced by these groups. I proceed with the estimation of the elasticity of labour supply to the firm in the next section.

8. Results

8.1. Main Regression Results

In the previous sections I explained the process by which I investigate the extent of monopsony power in South Africa. After reviewing the literature, I explained the theoretical model and thus how the elasticity of labour supply to the firm indicates the extent of potential monopsony power. Then, I discussed how the elasticity of labour supply to the firm can be estimated via the elasticity of separations. I then discussed how these separations can be identified using tenure in survey data, and in the preceding section, I presented some descriptive statistics and preliminary evidence on the extent of monopsony in South Africa. In this section, I estimate the elasticity of labour supply to the firm and discuss what implications this has for the extent of potential monopsony power in South Africa.

Table 8A below reports the main results. I compare the results across the two datasets and across the three sets of controls. A six-month interval between periods is used to make comparisons between the LFS and QLFS easier. This is the interval between waves in the LFS but represents a two-wave interval in the QLFS. I only report the average partial effects of the primary variable of interest, log real wage. Full regression results are reported in Appendix 2. The first seven rows show the results obtained from the simplified Manning (2003) approach. The rest of the table reports results obtained from using the more complex Manning (2003) approach.

The key result from the specification of choice, named 'Medium Controls' [Med Controls] in the table, is that an increase in the average wage by 1% is associated with an average decrease in the likelihood of separation by 0,4% in the LFS and 0,41% in the QLFS. All coefficients are significant at the 1% level in these regressions. To estimate the wage elasticity of labour supply to the firm, I multiply the separation elasticity by -2. An increase in the wage by 1% is associated with an increase in the labour supply to the firm of 0,8% in the LFS and 0,83% in the QLFS. These elasticities are far off infinity, thus giving strong evidence of an imperfectly competitive labour market. Further, using the simple approach, it seems that the level of monopsony power in South Africa has not changed over time. Without any controls, the elasticities increase, although by much more in the QLFS than the LFS. However, when I add controls for occupation, industry, and public sector employment, both coefficients decrease by a similar amount. The change in coefficient shows that these variables are related to both separations and wages. In other words, wages and separations differ across occupation, industry and sector.

Table 8A: Regression Analysis and Elasticity Calculations - 6 Month Interval

Type	Indicator	LFS			QLFS		
		No Controls	Med Controls	Med+ Controls	No Controls	Med Controls	Med+ Controls
All Separations	Coefficient	-0,08	-0,08	-0,06	-0,06	-0,05	-0,04
	Std. Err.	0,00	0,00	0,00	0,00	0,00	0,00
	Significance	***	***	***	***	***	***
	Rate	0,19	0,19	0,19	0,11	0,11	0,11
	Elasticity	-0,43	-0,40	-0,33	-0,58	-0,41	-0,35
Labour Supply	Simple Elasticity	0,85	0,80	0,66	1,16	0,83	0,70
Separations to Employment	Coefficient	-0,02	-0,03	-0,02	-0,02	-0,02	-0,01
	Std. Err.	0,00	0,00	0,00	0,00	0,00	0,00
	Significance	***	***	***	***	***	***
	Rate	0,06	0,06	0,06	0,04	0,04	0,04
	Elasticity	-0,40	-0,45	-0,36	-0,55	-0,41	-0,34
Separations to Non-Employment	Coefficient	-0,06	-0,06	-0,05	-0,04	-0,03	-0,03
	Std. Err.	0,00	0,00	0,00	0,00	0,00	0,00
	Significance	***	***	***	***	***	***
	Rate	0,14	0,14	0,14	0,07	0,07	0,07
	Elasticity	-0,46	-0,41	-0,34	-0,62	-0,44	-0,37
Hires from Employment	Coefficient	0,30	0,26	0,27	0,13	0,08	0,04
	Std. Err.	0,05	0,05	0,05	0,09	0,09	0,09
	Significance	***	***	***			
Proportion of Separations	To Employment	0,26	0,26	0,26	0,36	0,36	0,36
Proportion of Separations	To Non-Employment	0,74	0,74	0,74	0,64	0,64	0,64
Labour Supply	Complex Elasticity	0,62	0,68	0,51	1,07	0,79	0,68
Observations		46811	46811	46702	20729	20729	20705

Source: Own Calculations from StatsSA. Significance Levels: *** < 0,01. ** < 0,05. * < 0,1. Q/LFS – Quarterly/Labour Force Survey. Med controls: age, age squared, years of education, race, gender, marital status, province, wave. Med+ controls: Med controls plus occupation, industry and public sector. Hires from employment regression has a dependent variable = 1 if hired from employment and = 0 if hired from non-employment. Survey weights used. Full regression results reported in Appendix B.

In the second part of Table 8A I report the results obtained using Manning's (2003) more complex approach, where the elasticity of separations to employment and non-employment are calculated separately. I focus on the main specification (med controls) in discussing the results, as these controls are most similar to those used in Manning (2003). A 1% average increase in the real wage is associated with a 0,45% and 0,41% average decrease in the likelihood of separation to employment in the LFS and QLFS respectively. For separations to non-employment, the elasticity is -0,41 in the LFS and -0,44 and in the QLFS. Thus, the results are very similar across dataset and separations to employment and non-employment.

The final wage elasticities of labour supply to the firm using Manning's (2003) more complex approach are also similar across dataset, although they differ a little more than when using the complex approach. An increase in the real wage of 1% is associated with an average increase in the labour supply to the firm of 0,68% in the LFS, and 0,79% in the QLFS. The main reason for this difference is that the coefficient on log real wage in the hires from employment regression is larger in the LFS than in the QLFS. This coefficient is equal to 0,26 in the LFS, but only equal to 0,08 in the QLFS.

The last part of Table 8A shows that the proportion of separations to employment is higher in the LFS compared to the QLFS. This means that more weight is placed on the elasticity of separations to employment rather than non-employment when calculating the wage elasticity of labour supply to the firm.

I can use the elasticities to calculate the implied rates of exploitation, which is the inverse of the elasticity of labour supply to the firm. The rate is how much, in percentage terms, a worker's marginal revenue product is greater than their wage. The complex elasticities imply a potential exploitation rate of 147% in the LFS and 127% in the QLFS. This means that if the marginal revenue productivity of workers in real terms was R50 an hour, then the maximum a firm could mark down their worker's wage to is R20,24 in the LFS and R22,07 in the QLFS. Thus, the levels of monopsony power are very similar across dataset.

Before moving on to the next subsection, it is worth noting that the relative similarity of the results in the LFS and QLFS is surprising. In the previous sections, I documented differences in the quality of the panels, the average level of real wages, the quality of the tenure variable and the level of separations. Yet, the largest difference in the elasticities between the LFS and QLFS is only 0,11.

8.2. Results across Gender

I now turn to an analysis of the elasticity estimates across gender. Table 8B below reports the elasticity estimates obtained using the medium controls specification. I also report the differences in the two estimates by dataset. I use interaction terms in the regressions in order to test for differences in the coefficients.

Ai & Nortion (2003) warn against using interaction terms in non-linear regressions and estimating average partial effects. The authors argue that including an interaction term in a non-linear regression and interpreting the average partial effect of this interaction term is incorrect. However, Karaca-Mandic, Norton & Dowd (2012) show how this can be estimated accurately using Stata. I could also estimate the regressions separately. However, this does not allow me to test for significance across category. Further, Alisson (1999) argues that the comparability of coefficients across separate regressions is not always possible. Therefore, I calculate these elasticities by using the average partial effect of the log real wage on separation across gender, using the approach specified by Karaca-Mandic, Norton & Dowd (2012).

The male and female elasticities are quite similar to each other in most dataset-method combinations. The male complex elasticities are marginally larger in both the LFS and the QLFS, but this is only by 0,12 in both datasets. Given the wide variety of parameters involved in estimating the complex elasticities, it is difficult to test for a significant difference. However, of all the average partial effects of separations on the wage, only the average partial effects of separations to non-employment in the LFS are significantly different across gender. Therefore, I conclude that there is not enough evidence in the LFS and QLFS to suggest that monopsony power differs between men and women on the whole in South Africa.

That the estimates are similar across gender is surprising for two reasons. First, several papers including Booth & Katic (2011) and Ransom & Oaxaca (2010) find lower female elasticities. Second, research in South Africa shows that women experience worse labour market outcomes than men (Casale & Posel, 2011; Bhorat & Goga, 2013; Mosomi, 2019). If their labour market outcomes differ, one would expect that monopsony power differs, but it does not. However, the results are similar to Manning (2003), who, although he does not explicitly report the wage elasticities of the labour supply of men and women to the firm, reports the elasticity of separations for both genders. Manning (2003) finds very little difference in the elasticities of separations of men and women.

Table 8B: Regression Analysis and Elasticity Calculations across Gender - 6 Month Interval

Type	Indicator	LFS			QLFS		
		Male	Female	Difference	Male	Female	Difference
All	Coefficient	-0,073	-0,081	0,008	-0,045	-0,045	0,000
Separations	Std. Err.	0,00	0,00	1,54	0,005	0,005	0,047
	Significance	***	***		***	***	
	Rate	0,179	0,204	***	0,110	0,108	
	Elasticity	-0,408	-0,395		-0,408	-0,419	
Labour Supply	Simple Elasticity	0,82	0,79		0,82	0,84	
Separations	Coefficient	-0,028	-0,024	-0,004	-0,019	-0,016	-0,003
to	Std. Err.	0,003	0,003	-1,033	0,004	0,003	-0,632
Employment	Significance	***	***		***	***	
	Rate	0,062	0,053	***	0,044	0,039	**
	Elasticity	-0,445	-0,448		-0,421	-0,402	
Separations	Coefficient	-0,052	-0,064	0,012	-0,029	-0,033	0,004
to	Std. Err.	0,003	0,004	2,772	0,003	0,004	0,791
Non-Employment	Significance	***	***	***	***	***	
	Rate	0,125	0,160	***	0,069	0,071	
	Elasticity	-0,422	-0,402		-0,420	-0,457	
Hires	Coefficient	0,212	0,314	-0,010	0,002	0,186	0,184
from	Std. Err.	0,072	4,251	0,119	0,124	1,640	0,165
Employment	Significance	***	***				
Proportion of Separations	To Employment	0,305	0,219		0,375	0,339	
	To Non-employment	0,695	0,781		0,625	0,661	
Labour Supply	Manning (2003) Elasticity	0,73	0,61		0,84	0,72	
Observations		24653	24653	22158		10074	10655

Source: Own calculations from StatsSA. Significance Levels: *** < 0,01. ** < 0,05. * < 0,1. Q/LFS – Quarterly/Labour Force Survey. Controls: age, age squared, years of education, race, marital status, province, wave. Hires from employment regression has a dependent variable = 1 if hired from employment and = 0 if hired from non-employment. Survey weights used. Full regression results reported in Appendix C.

8.3. Results across Race

Table 8C below reports the results across race. I only report the simple estimates for differences across race. I do this because some of the sample sizes are very small when running the hires from employment regression in the complex method. Many of the differences between races in the average partial effects below are statistically insignificant, although the separation rates are generally statistically different from each other. Thus, the results should be taken as suggestive evidence unless otherwise stated.

Table 8C: Regression Coefficients and Elasticities by Race

Type	Indicator	LFS				QLFS			
		African	Coloured	Indian /Asian	White	African	Coloured	Indian /Asian	White
All Separations	Coefficient	-0,08	-0,06	-0,10	-0,05	-0,05	-0,04	0,00	-0,04
	Std. Err.	0,00	0,01	0,02	0,02	0,00	0,02	0,03	0,02
	Significance	***	***	***	***	***	**		**
Significance	Compared to African		***						
	Rate	0,21	0,19	0,17	0,12	0,12	0,13	0,05	0,03
Significance	Compared to African		***	***	***			***	***
	Elasticity	-0,38	-0,32	-0,58	-0,42	-0,38	-0,32	-0,09	-1,24
Labour Supply	Simple Elasticity	0,76	0,64	1,17	0,83	0,76	0,65	0,18	2,47
		31797	10309	1376	3298	17209	2306	256	820

Source: Own calculations from StatsSA. Significance Levels: *** < 0,01. ** < 0,05. * < 0,1. Q/LFS – Quarterly/Labour Force Survey. Controls: age, age squared, years of education, gender, marital status, province, wave. Survey weights used. Full results reported in Appendix D (including complex method).

Although Whites may be more elastically supplied to the firm than Africans and Coloureds, it is difficult to confirm this statistically, as the sample sizes are small. The largest elasticity to the firm is for the supply of Indian and Asians in the LFS, but this elasticity to the firm is then the lowest in the QLFS, possibly due to the very small sample size observed in the QLFS. The African and Coloured elasticities are quite similar to each other in both waves, although the African elasticity is slightly larger, with the average partial effects and separation rates statistically significantly different from each other in the LFS.

Having reported gender and race separately, I can also analyse how the elasticities differ across race within gender. The elasticities are reported in Table 8D below. In both the LFS and the

QLFS, Coloured women are supplied more elastically to the firm than Coloured men, with the average partial effect only significantly different in the LFS. This suggests Coloured men face more potential monopsony power than women. This is a surprising result for the same reasons discussed in the gender section. We would expect women to face more potential monopsony power, and thus it is interesting that Coloured men face more in South Africa. The difference in the Coloured male and African male elasticities is also likely significant in the LFS, as the average partial effects are significantly different. None of the other differences are significantly different in the average partial effects. For Indians and Asians, a negative female elasticity is found in the QLFS, likely due to a small sample size.

Table 8D: Difference in Elasticities across Gender and Race

Dataset	LFS		QLFS	
Race/Gender	Male	Female	Male	Female
African	0,75	0,76	0,76	0,77
Coloured	0,46	0,83	0,43	0,89
Indian/Asian	1,11	1,14	1,25	-0,64
White	1,05	0,58	2,53	2,93

Source: Own Calculations from StatsSA. Q/LFS – Quarterly/Labour Force Survey. Controls: age, age squared, years of education, gender, marital status, province, wave. Survey weights used. Regression coefficients and significance reported in Appendix E.

8.4. Results across other Covariates.

I also analyse the simple elasticities across some of the other covariates. Table 8E below reports the results across education category and gender.

Table 8E: Elasticity across Gender and Education Category

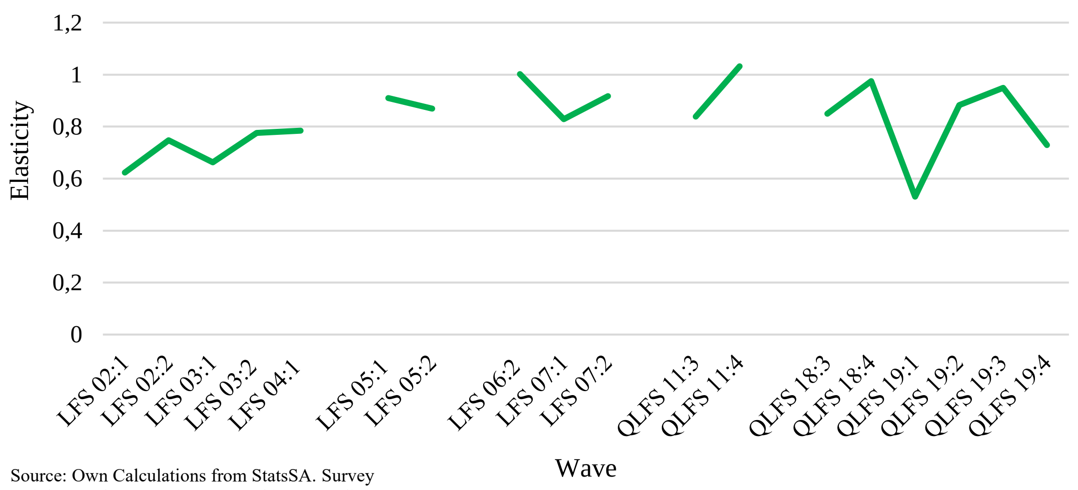
Dataset	LFS			QLFS		
Education/Gender	Male	Female	Overall	Male	Female	Overall
No Education	0,28	0,60	0,49	0,17	0,69	0,44
Some Primary	0,63	0,48	0,57	0,43	0,32	0,41
Some Secondary	0,72	0,55	0,64	0,35	0,41	0,37
Matric	0,96	0,82	0,89	1,63	1,15	1,40
Some Tertiary / Degree	1,81	1,61	1,65	4,45	1,52	2,58

Source: Own calculations from StatsSA. Q/LFS – Quarterly/Labour Force Survey. Controls: age, age squared, years of education, gender, marital status, province, wave. Survey weights used. Regression coefficients and significance reported in Appendix E.

The elasticity of labour supply to the firm increases with education, with the average partial effects for those with a matric or tertiary education significantly larger than some of the lower categories for both men and women in the LFS and just men in the QLFS. Therefore, those with more education are less vulnerable to potential monopsony power. The elasticities are particularly high for the supply of men with tertiary education. This implies that attaining a tertiary education protects one from potential monopsony power, and more so for men. In fact, the only educational categories in which men have a statistically significantly larger average partial effect than women in the LFS are those who have a matric or tertiary education. There is thus some evidence that it is only with higher levels of education that women face more monopsony power as compared to men.

Lastly, Figure 6 below reports how the elasticity of labour supply to the firm differs across wave. The elasticity has remained relatively consistent, with all values within 0,3 of each other. However, one small change is important to note. The elasticity of labour supply to the firm is lower in the 2001:2-2004:1 LFS panel compared to the 2004:2-2007:2 panel. The average elasticity up to the first half of 2004 is 0,72, while in the second half of the LFS, it is 0,91. Therefore, the close similarity of the LFS and QLFS elasticity estimates above are in fact due to a slight increase in these estimates in latter half of the LFS.

Figure 6: Elasticity of Labour Supply to the Firm across Wave



Source: Own Calculations from StatsSA. Survey weights used. Regression reported in Appendix E.

8.5. Summary

In summary, the estimated elasticity of labour supply to the firm is low at between 0,68 and 0,83. The results therefore imply a South African labour market that is far from perfectly competitive. Results do not differ substantially across the LFS and the QLFS when using either method, although the complex elasticities are marginally larger in QLFS as compared to the LFS. Results are also very similar across gender when using both methods, which is surprising given the evidence in the literature. Across race, a low Coloured male elasticity was found, and some suggestive evidence that the White elasticity to the firm is largest. The wage elasticity of labour supply decreases with education, and the elasticity to the firm is particularly high for the supply of tertiary educated individuals to the firm. Lastly, greater education appears to benefit men more than women in terms of the monopsony power they face. In the following section I conduct a series of robustness checks to understand whether the above results are sensitive to the wave intervals used and the method of identifying separations.

9. Robustness

9.1. Different Wave Intervals

The first robustness check examines whether results differ when the wave interval changes. This is done for two reasons. First, it will reveal whether differing wave intervals impacts the elasticity of labour supply to the firm. Second, it will enable greater comparability with the literature. I do not expect the coefficients to differ substantially when changing the wave interval, but checking this to ensure that the estimates of the elasticity are as comparable to the literature is necessary.

Table 9A below reports the results from a 1-year interval in the LFS and a quarterly interval in the QLFS. The results do not differ substantially when the wave interval is changed. The yearly LFS coefficients of 0,74 and 0,7 are very similar to their 6-month counterparts of 0,8 and 0,68. The quarterly QLFS coefficients of 0,82 and 0,77 are also very similar to their 6-month counterparts of 0,83 and 0,79.

Table 9B below reports these elasticities alongside some estimates from other key papers. Bassier (2023), using the complex method, finds an elasticity of 0,86 in South Africa, which is similar to my estimates. This is somewhat surprising as I believe that measurement error in the dependent variable is worse in the LFS and QLFS than in the tax data that Bassier (2023) uses. Bassier's (2023) is slightly larger than our yearly LFS estimate here. A larger elasticity of separations to employment drives this. There are several reasons Bassier's (2023) simple estimate may be slightly larger. First, Bassier (2023) analyses the years between 2011 and 2016. This is later than our LFS data. Second, as mentioned above, Bassier (2023) uses administrative tax data, and thus likely suffers from less measurement error than in our data. Third, Bassier (2023) has less rich a set of individual controls, which may push his estimate upwards because unobserved heterogeneity might be influencing both the worker's wage and separation behaviour. Indeed, the yearly estimate from the LFS is 0,8 without controls, which is a little closer to Bassier's (2023) estimate. Bassier's (2023) estimate when using the 'movers' approached discussed above is 1,6, which is around double the LFS and QLFS estimates. In summary, when the methods are similar, the estimates in the survey data and the tax data are not too different. But when an explicit identification approach is used as in Bassier (2023), the estimates in the survey data are much smaller. Thus, using survey data may not be as much of an issue as compared to not having an explicit identification approach. My estimates are also larger than Bassier's (2023) estimates using the QLFS, which suggests that my use of unimputed earnings data might be making a difference, although this is investigated in Section 10.

Table 9A: Regression Analysis and Elasticity Calculations

Type	Indicator	LFS - 1 Year			QLFS - 3 Months		
		No Controls	Med Controls	Med+ Controls	No Controls	Med Controls	Med+ Controls
All Separations	Coefficient	-0,10	-0,09	-0,08	-0,04	-0,03	-0,02
	Std. Err.	0,00	0,00	0,01	0,00	0,00	0,00
	Significance	***	***	***	***	***	***
	Rate	0,25	0,25	0,25	0,07	0,07	0,07
	Elasticity	-0,40	-0,37	-0,32	-0,60	-0,41	-0,34
Labour Supply	Simple Elasticity	0,80	0,74	0,63	1,19	0,82	0,67
Separations to Employment	Coefficient	-0,04	-0,04	-0,04	-0,01	-0,01	0,00
	Std. Err.	0,00	0,00	0,01	0,00	0,00	0,00
	Significance	***	***	***	***	***	***
	Rate	0,10	0,10	0,10	0,02	0,02	0,02
	Elasticity	-0,40	-0,42	-0,35	-0,61	-0,42	-0,27
Separations to Non-Employment	Coefficient	-0,07	-0,06	-0,06	-0,03	-0,02	-0,02
	Std. Err.	0,00	0,00	0,00	0,00	0,00	0,00
	Significance	***	***	***	***	***	***
	Rate	0,16	0,16	0,16	0,06	0,06	0,06
	Elasticity	-0,44	-0,39	-0,34	-0,60	-0,41	-0,36
Hires from Employment	Coefficient	0,24	0,17	0,15	0,10	0,09	0,11
	Std. Err.	0,05	0,05	0,06	0,06	0,07	0,07
	Significance	***	***	***			
Proportion of Separations	To Employment	0,34	0,34	0,34	0,23	0,23	0,23
Proportion of Separations	To Unemployment	0,66	0,66	0,66	0,77	0,77	0,77
Labour Supply	Complex Elasticity	0,67	0,70	0,60	1,14	0,77	0,52
Observations		25754	25754	25700	44216	44216	44180

Source: Own Calculations from StatsSA. Significance Levels: *** < 0,01. ** < 0,05. * < 0,1. Q/LFS – Quarterly/Labour Force Survey. Med controls: age, age squared, years of education, race, gender, marital status, province, wave. Med+ controls: Med controls plus occupation, industry and public sector. Hires from employment regression has a dependent variable = 1 if hired from employment and = 0 if hired from non-employment. Survey weights used. Full results reported in Appendix F.

Manning's (2003) simple yearly estimates are 1,44 and 1,95. Therefore South Africa has a lower elasticity of labour supply to the firm than in the US or UK when the simple method is used. Booth & Katic (2011) find a simple estimate of 0,72 in Australia. This is surprisingly similar given that Australia is a developed country. We would expect Australia to have a more similar elasticity to those found in Manning (2003). The complex estimates are a little more similar to Manning's results. Although the LFS estimate of 0,68 is still smaller than the PSID's value of 1,38 and the BHPS's value of 0,75, it is closer than for the simple estimates. The results from the 3-month interval in the QLFS can be compared to the results from the UK LFS in Manning (2003). In Manning (2003) the LFS (UK), the simple elasticity is 1 and the complex is 0,75. Therefore my simple estimate is smaller than the UK estimate, but our complex estimate is very similar.

Table 9B: Comparison of Estimates to the Literature

Measure	Category	PSID	BHPS	SA Tax Bassier (2023)	SA LFS	UK LFS	SA QLFS
Time Interval		1 Yr	1 Yr	1 Yr	1 Yr	1 Qtr	1 Qtr
Elasticity of Separations	All	-0,97	-0,72	-0,31	-0,37	-0,50	-0,41
Elasticity of Labour Supply to the Firm	Simple	**1,95**	**1,44**	**0,62**	**0,74**	**1,00**	**0,82**
Elasticity of Separations	To Employment	-0,87	-0,63	-0,51	-0,42	-0,53	-0,42
	To Non-Employment	-0,89	-0,63	-0,29	-0,39	-0,58	-0,41
Hires from Employment	Logit Coefficient	0,95	1,38	[a]	0,17	0,75	0,09
Share of Separations	To Employment	0,62	0,63	[b]	0,34	0,56	0,23
	To Non-Employment	0,38	0,37	[c]	0,66	0,44	0,77
Elasticity of Labour Supply to the Firm	Complex	**1,38**	**0,75**	**0,86**	**0,70**	**0,75**	**0,78**

Source: Own calculations from StatsSA, Manning (2003), Bassier (2023). [a]Bassier (2023) uses a slight variation of Eqn. to calculate the elasticity to the firm. Thus for this parameter, he calculates an elasticity instead of using a logit coefficient. His elasticity is equal to 0,041. If I calculate an elasticity for the LFS coefficient, it is 0,11. [b]Bassier (2023) uses the share of recruits from employment, which Manning (2023) says in equilibrium should be equal to the share of separations to employment. Bassier (2023) finds this to be 0,45.[c]Accordingly, the share of recruits from non-employment implied by Bassier (2023) is 0,55.

Manning (2003) also reports most of the components of the elasticity formula. Therefore, I can investigate what drives the differences between my results and Manning (2003). The larger simple elasticity of labour supply to the firm in the UK is driven by a larger elasticity of separations in all datasets in Manning (2003). My complex estimates are more similar to his due to the logit coefficient on wage in the hires from employment regression, which is much larger in the datasets used by Manning (2003) and so reduces the simple elasticity by more than in the LFS and QLFS. Table 9B also shows that the proportion of separations to employment is very different in South Africa than in the US and UK. In the yearly data, the proportion of separations to employment is around 63% in the US and UK, but only 34% in the LFS. When using quarterly data, it is 56% in the UK and 23% in South Africa.

9.2. Identifying Separations using Employer Characteristics

As discussed in section 7, the results may be sensitive to the way in which I identify separations to employment. The other way I could identify separations is through observing changes in the industry and the public/private sector of the individual. This is done in addition to identification using tenure. I did not use this method in the main analysis above, as I believe that it likely introduces substantial measurement error into the dependent variable. As per Hausman et. al. (1998), this would bias the regression coefficients downwards. I investigate this in Table 9D.

All coefficients decrease substantially when using employer characteristics to identify separations. In both the simple and complex methods, the elasticities more than halve. These are very low estimates of the wage elasticity of labour supply to the firm compared to those found by Manning (2003), Booth & Katic (2011) and Bassier (2023). If true, they imply a very high level of potential monopsony power in South Africa.

The estimates of the elasticity of the labour supply to the firm when using employer characteristics to identify separations are very low, and thus likely a result of substantial measurement error in the dependent variable. Thus, using the tenure-alone method is likely a better approach. This also implies that Pan's (2022) estimates of worker flows may be upwardly biased.

9.3. Robustness Summary

Some important conclusions have followed from the robustness analysis. First, changes in the wave interval have not changed the results. Second, I find a similar but slightly smaller complex elasticity than South Africa than Bassier (2023) when using a similar method, but a much smaller elasticity compared to his 'movers' approach. Third, when using the simple approach, I find a lower wage elasticity of labour supply to the firm in South Africa than in the UK and the US but a similar elasticity to Australia. I find more similar complex elasticities to the UK. This similarity is driven by large differences in the logit coefficient on log wage in the hired from employment regression. Fourth, I find much lower elasticities when using the firm characteristics approach to identify separations, which is likely due to substantial measurement error in the dependent variable.

Table 9D: Regression Analysis and Elasticity Calculations - 6 Month Interval

Type	Indicator	LFS	QLFS
		Med Controls	Med Controls
All	Coefficient	-0,06	-0,04
Separations	Std. Err.	0,00	0,00
	Significance	***	***
	Rate	0,39	0,21
	Elasticity	-0,16	-0,17
Labour Supply	Simple Elasticity	0,31	0,35
Separations	Coefficient	-0,02	-0,01
to	Std. Err.	0,00	0,00
Employment	Significance	***	***
	Rate	0,29	0,15
	Elasticity	-0,08	-0,09
Separations	Coefficient	-0,05	-0,03
to	Std. Err.	0,00	0,00
Non-Employment	Significance	***	***
	Rate	0,13	0,07
	Elasticity	-0,40	-0,44
Hires	Coefficient	0,59	0,37
from	Std. Err.	0,03	0,08
Employment	Significance	***	***
Proportion of Separations	To Employment	0,65	0,67
Proportion of Separations	To Unemployment	0,35	0,33
Labour Supply	Complex Elasticity	0,06	0,17
Observations		48567	20810

Source: Own Calculations from StatsSA. Significance Levels: *** < 0,01. ** < 0,05. * < 0,1. Q/LFS – Quarterly/Labour Force Survey. Med controls: age, age squared, years of education, race, gender, marital status, province, wave. Hires from employment regression has a dependent variable = 1 if hired from employment and = 0 if hired from non-employment. Survey weights used. Full regression results reported in Appendix G.

10. Measurement Error in the Publicly Available Earnings Data

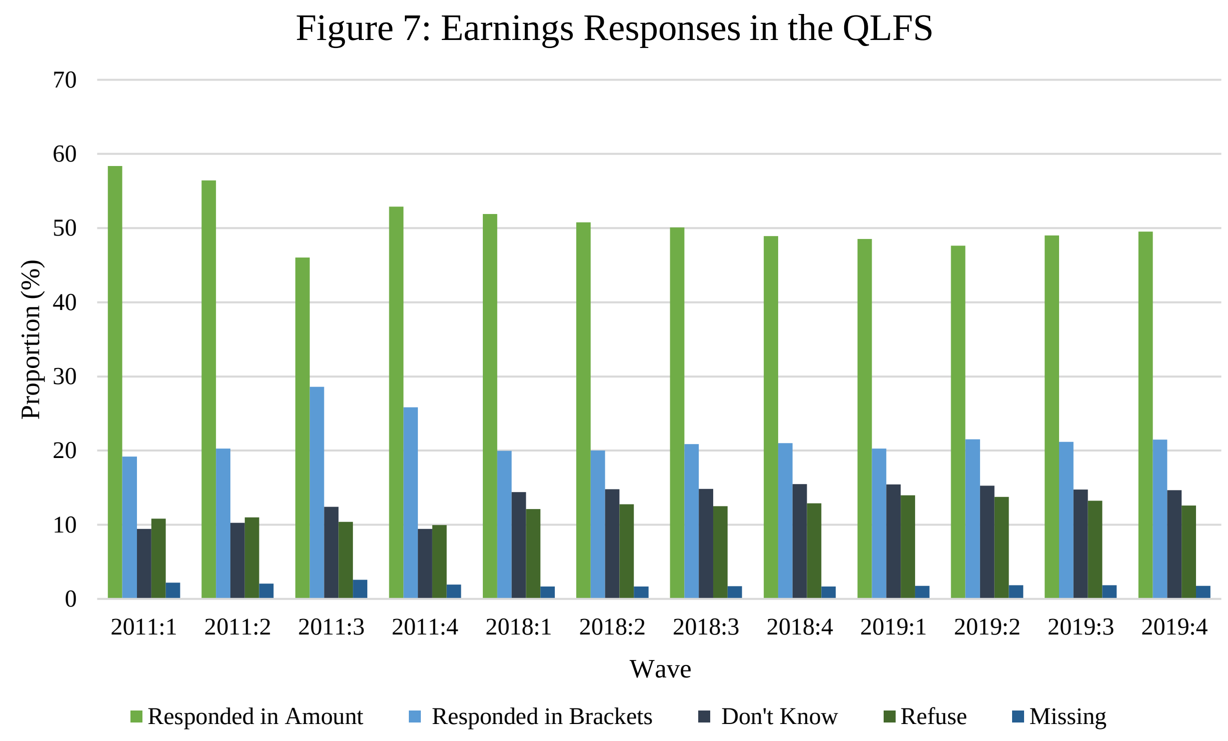
10.1. Introduction

The final issue worth discussing in this paper is measurement error in the key independent variable, real wages. In the above analysis, I only use unimputed earnings data from 2011, 2018 and 2019 to estimate the wage elasticity of labour supply to the firm in the QLFS. This is because the publicly available StatsSA earnings data in the QLFS is regarded as unreliable (Kerr & Wittenberg, 2021; Kerr, 2023), because the imputation undertaken by StatsSA appears to have been poorly done and is not publicly documented (Kerr & Wittenberg, 2021; Kerr, 2023). Kerr & Wittenberg (2021) show how measurement error in earnings attenuated the union wage premium in the QLFS data. I therefore believe it important to investigate how different my results would have been if I had used the publicly available data. This is an important issue because quality wage data is clearly crucial in understanding the condition of the South African economy.

I first conduct some descriptive analysis comparing the imputed and unimputed data to gauge who StatsSA imputed for and how badly they might have done so. Then, I analyse the effects of any potential measurement error on my monopsony analysis. Measurement error in the independent variable generally causes attenuation bias in the coefficients (Wooldridge, 2019). Therefore, I can observe whether this prediction holds by running the same regression as above with the publicly available data.

10.2. Descriptive Analysis

Before looking at the quality of StatsSA's imputation, it is helpful to understand how people respond to the earnings question and in what proportions. This is only possible with the non-public data. This is reported in Figure 7 below. The proportion of individuals who respond with an amount ranges between 47% and 59%. It is larger in 2011 and decreases in 2018 and 2019. The proportion responding in brackets remains relatively consistent at around 20%, except for two spikes in the second half of 2011. The proportion of those responding 'Don't know' increases from 10% in 2011 to 14,7% in 2018/19. Refusals also increase from 10,7% to 13,1% but remain slightly below 'Don't Know'. A category not mentioned in either Kerr & Wittenberg (2021) or Kerr (2023) are those completely missing. There is no category for these individuals in the StatsSA questionnaires either, and thus it is unclear how these completely missing observations come about. They decrease slightly from 2,1% In 2011 to 1,7% in 2018 and 2019.



Source: Own calculations from StatsSA.

Very large differences between 2011 and 2018/19 emerge when considering imputation. Table 10A below reports the proportion of non-responses imputed for by StatsSA. In 2011, StatsSA imputed for almost all those who responded either in brackets, with a "don't know" or with a refusal. 71% of those with a completely missing answer had their incomes imputed.

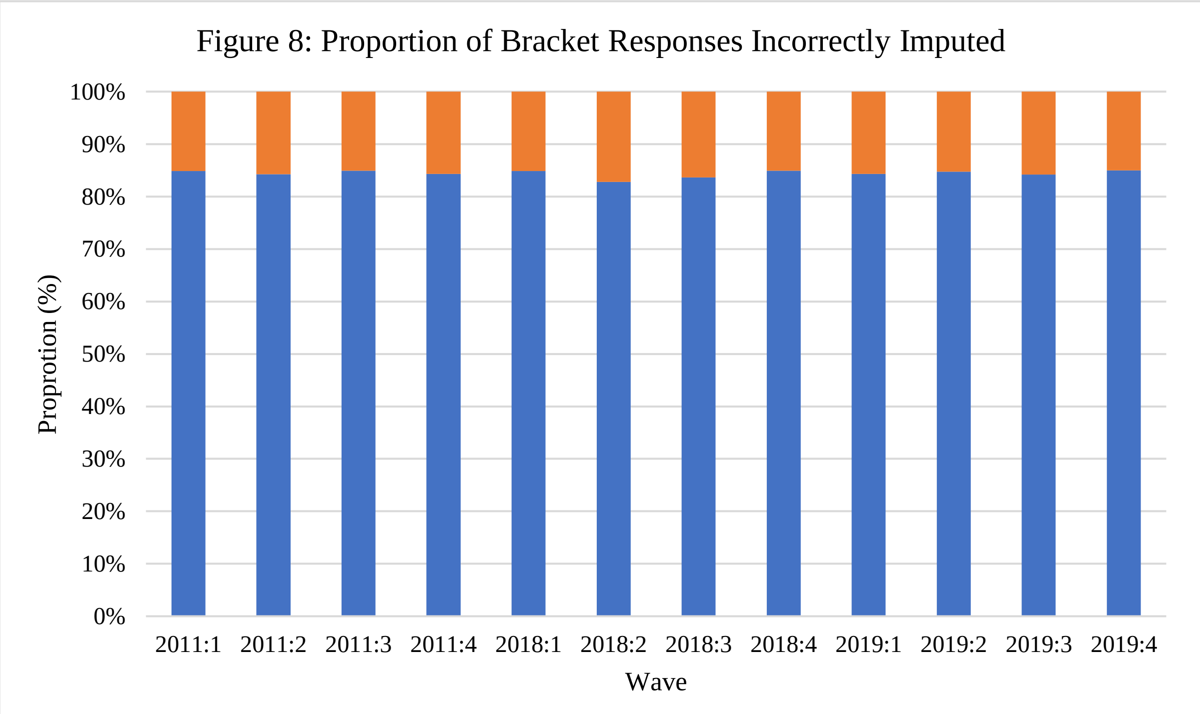
Table 10A: Proportion Imputed by StatsSA

Category	QLFS 2011	QLFS 2018/19
Response in brackets	100%	56%
Don't Know	99%	55%
Refuse	99%	45%
Missing	71%	35%

Source: Own Calculations using unimputed StatsSA data.

In the 2018/19 QLFS, only just over half of those responding with brackets or "don't know" are imputed. Under half of the refusals are imputed, and only 35% of the completely missing responses are imputed for. There is no clear pattern in the data of who they choose to impute for. Comparing the public and non-public data, all those who responded in brackets but who are not imputed for have a 'don't know' or a refusal. It thus seems that StatsSA changed the answers of some individuals in the publicly released data. These results square with what Kerr (2023) finds in the 2012 QLFS.

I can also provide some insight into the quality of the imputations. I do this by comparing the imputed income with the bracket response for those who responded in brackets. I use a similar approach to Kerr (2023), who finds a similar result in the 2012 QLFS. This is reported in Figure 8 below and the results are very concerning. For those who responded in brackets, around 85% of the imputed amounts by StatsSA fell outside of the bracket in which they responded. There is therefore strong evidence of inconsistent and poor imputation across the QLFS. Poor imputation results in individuals being assigned incomes that are not accurate. This results in substantial measurement error. I, therefore, expect the estimates of the elasticity of labour supply to be much lower when using this data.



Source: Own calculations from StatsSA

10.3. Regression Analysis

Table 10B below reports the regression results using the mismeasured earnings data. I report the estimates for both a 3- and 6-month interval. The estimated elasticities are much smaller, and are closer to and in fact even smaller than Bassier's (2023) survey data estimates. For the 6-month interval, the simple elasticity using the unimputed data is 3,6 times larger than when using the imputed data. For the complex elasticity, it is 6,5 times larger. Therefore, there is some strong evidence to suggest that using the publicly released StatsSA earnings data leads to downwardly biased coefficients. Poor earnings data severely inhibits the ability of researchers to conduct accurate analysis of the labour market. Thus, StatsSA should release all unimputed data since the beginning of the QLFS to enable researchers to conduct more accurate analysis.

Table 10B: Mismeasured Wages Regression

		6 Months		3 Months	
Type	Indicator	Imputed	Unimputed	Imputed	Unimputed
			Med Controls		Med Controls
All	Coefficient	-0,01	-0,05	-0,01	-0,03
Separations	Std. Err.	0,00	0,00	0,00	0,00
	Significance	***	***	***	***
	Rate	0,11	0,11	0,07	0,07
	Elasticity	-0,11	-0,41	-0,10	-0,41
Labour Supply	Simple Elasticity	0,23	0,83	0,20	0,82
Separations	Coefficient	0,00	-0,02	0,00	-0,01
to	Std. Err.	0,00	0,00	0,00	0,00
Employment	Significance	*	***	***	***
	Rate	0,04	0,04	0,02	0,02
	Elasticity	-0,05	-0,41	-0,09	-0,42
Separations	Coefficient	-0,01	-0,03	-0,01	-0,02
to	Std. Err.	0,00	0,00	0,00	0,00
Non-Employment	Significance	***	***	***	***
	Rate	0,07	0,07	0,06	0,06
	Elasticity	-0,15	-0,44	-0,19	-0,41
Hires	Coefficient	0,08	0,08	0,07	0,09
from	Std. Err.	0,05	0,09	0,04	0,07
Employment	Significance				
Proportion of Separations	To Employment	0,36	0,36	0,23	0,23
Proportion of Separations	To Unemployment	0,64	0,64	0,77	0,77
Labour Supply	Complex Elasticity	0,12	0,79	0,20	0,77
Observations		34126	20729	73249	44216

Source: Own Calculations from StatsSA. Significance Levels: *** < 0,01. ** < 0,05. * < 0,1. Q/LFS – Quarterly/Labour Force Survey. Med controls: age, age squared, years of education, race, gender, marital status, province, wave. Hires from employment regression has a dependent variable = 1 if hired from employment and = 0 if hired from non-employment. Survey weights used. Full regressions results reported in Appendix H.

11. Conclusion

In this paper, I have used household survey data in South Africa to estimate the extent of potential monopsony power in the labour market. To do this, I estimated the elasticity of labour supply to the firm in a manner proposed by Manning (2003). I produced four estimates using two slightly different methods over two different periods, finding that the elasticity of labour supply to the firm in South Africa is low, with all estimates below 1. The estimates ranged between 0,68 and 0,83 which imply rates of exploitation of between 120% and 147%. In other words, firms have significant power in South Africa to pay workers a wage far below their marginal revenue product.

Few definite differences in the elasticity to the firm were found across gender and race as a whole. However, Coloured women and African men were found to be supplied more elastically to the firm than Coloured men. Although the White elasticity was large, a small sample size prevented me from statistically inferring that it is larger than the African and Coloured elasticities. The most interesting cross-group results were found for differences in education. Those with a matric or tertiary education were found to be supplied more elastically to the firm than those with less education. Thus, education appears to be an important avenue through which workers are protected against firms' monopsony power. Men appeared to benefit from these differences more, with evidence of a significant gender gap in monopsony power for those with more education, and none for those with little.

I do recognise that measurement error is a significant concern. Thus, my estimates are likely downwardly biased. However, they are unlikely to be biased enough such that the labour market is actually perfectly competitive. Bassier's (2023) estimate of 0,86 using a similar method and administrative tax data show that my estimates may not be substantially biased. However, when Bassier (2023) uses an explicit identification approach, he finds an elasticity of 1,6, and thus I acknowledge that there are limitations to Manning's (2003) method.

A secondary aim of this paper was to understand how sensitive this analysis is to several factors. First, I showed that how I identified job-to-job flows has significant implications for the estimate of the elasticity found. Although not discussed at all by Manning (2003) and Booth & Katic (2011), tenure is a may be a highly mismeasured variable in survey data, as shown by Brown & Light (1992) and Bergin (2015). I showed that tenure seems to be much more mismeasured in the LFS than in the QLFS, although the QLFS showed surprisingly high levels of internal consistency. To provide an alternative, I investigated a method used by Pan (2022), which used changes in employer characteristics in addition to changes in tenure. This method produced extremely high job flow numbers in the LFS and very low estimates of the elasticity of labour supply to the firm. I suspect this is due to significant mismeasurement of job flows using this method.

Lastly, I showed that results are sensitive to measurement error in earnings. By comparing StatsSA's publicly released imputed earnings and the unimputed data, I provided evidence of very poor imputation on StatsSA's part. I also showed that the elasticities decrease substantially when using this data.

My results have important implications for policymaking in South Africa. As discussed in the literature review, one application of monopsony is the minimum wage. Under a perfectly competitive framework, all minimum wage introductions cause employment losses. However, under an imperfectly competitive framework, a minimum wage of the right level will not decrease employment, and could, in fact, increase it (Ehrenberg, Smith & Hallock, 2021). Thus, as South Africa is an imperfectly competitive market where both unemployment and poverty are high, a minimum wage, such as was implemented in 2019, could improve working conditions without necessarily reducing employment. This increases the need to pursue empirical research into whether the 2019 minimum wage is set at the right level.

My results have further relevance with respect to the use of survey data to estimate the elasticity of labour supply to the firm. Although much of the recent literature has used administrative data to estimate this elasticity due to measurement error in survey data, survey data is often more accessible to researchers and has the added benefit of including the informal sector. The fact that Bassier (2023) found an elasticity to the firm of 0,86 using a similar method with administrative data suggests that with some careful assumptions and an understanding of survey data's imperfections, a result not far of the arguably superior administrative data can be found. Thus, where administrative data may be sparse, especially in developing countries, survey data can be used to some extent to inform on the level of monopsony power in a labour market.

The most important result from this paper is that the wage elasticity of labour supply to the firm is low in South Africa, which implies that there is substantial monopsony power in the South African labour market. Although it may be downwardly biased, the magnitude of the bias would have to be substantial for there to be little monopsony power in South Africa. South Africa does have a high prevalence of unions and bargaining councils and a national minimum wage which all likely serve to restrain this power, but it is nevertheless very low. As such, when analysing the South African labour market and considering how to solve the significant challenges it presents, researchers and policymakers should do so with an imperfectly competitive model of the labour market in mind.

12. Reference List

Abowd, J.M., Kramarz, F. & Margolis, D.N. 1999. High Wage Workers and High Wage Firms. *Econometrica*. 67(2):251–333. DOI: 10.1111/1468-0262.00020.

Ai, C. & Norton, E.C. 2003. Interaction terms in logit and probit models. *Economics Letters*. 80(1):123–129. DOI: 10.1016/S0165-1765(03)00032-6.

Allison, P.D. 1999. Comparing Logit and Probit Coefficients Across Groups. *Sociological Methods & Research*. 28(2):186–208. DOI: 10.1177/0049124199028002003.

Anand, R., Kothari, S. & Kumar, N. 2016. *South Africa: Labor Market Dynamics and Inequality*. IMF Working Papers 16(137):1. DOI: 10.5089/9781498358934.001.

Azar, J., Marinescu, I. & Steinbaum, M. 2022. Labor Market Concentration. *Journal of Human Resources*. 57(Special Issue 1): S167–S199. DOI: 10.3368/jhr.monopsony.1218-9914R1.

Banerjee, A., Galiani, S., Levinsohn, J., McLaren, Z. & Woolard, I. 2008. Why has unemployment risen in the New South Africa?. *Economics of Transition*. 16(4):715–740. DOI: 10.1111/j.1468-0351.2008.00340.x.

Barth, E. & Dale-Olsen, H. 2009. Monopsonistic discrimination, worker turnover, and the gender wage gap. *Labour Economics*. 16(5):589–597. DOI: 10.1016/j.labeco.2009.02.004.

Bassier, I., Dube, A. & Naidu, S. 2021. Monopsony in Movers: The Elasticity of Labor Supply to Firm Wage Policies. *Journal of Human Resources*. 56(2):1–52. DOI: 10.3368/jhr.monopsony.0319-10111R1.

Bassier, I. & Woolard, I. 2021. Exclusive Growth? Rapidly Increasing Top Incomes Amid Low National Growth in South Africa. *South African Journal of Economics*. 89(2):246–273. DOI: 10.1111/saje.12274.

Bassier, I. 2022. *Collective bargaining and spillovers in local labour markets*. Centre for Economic Performance Discussion Paper No. 1895. London: Centre for Economic

Performance, London School of Economics and Political Science. Available: http://eprints.lse.ac.uk/118057/1/Bassier_Bargaining_council_spillovers_published.pdf. [2023, November 28].

Bassier, I. 2023. Firms and inequality when unemployment is high. *Journal of Development Economics*. 161:103029. DOI: 10.1016/j.jdeveco.2022.103029.

Benmelech, E., Bergman, N.K. & Kim, H. 2022. Strong Employers and Weak Employees How Does Employer Concentration Affect Wages? *Journal of Human Resources*. 57(Special Issue 1):S201–S250. DOI: 10.3368/jhr.monopsony.0119-10007R1.

Bergin, A. 2015. Employer Changes and Wage Changes: Estimation with Measurement Error in a Binary Variable. *LABOUR*. 29(2):194–223. DOI: 10.1111/labr.12051.

Bhorat, H. & Goga, S. 2013. The Gender Wage Gap in Post-Apartheid South Africa: A Re-examination. *Journal of African Economies*. 22(5):827–848. DOI: 10.1093/jae/ejt008.

Bhorat, H. & Stanwix, B. 2018. Wage Setting and Labor Regulatory Challenges in a Middle-Income Country Setting: The Case of South Africa. Background note in *An Incomplete Transition: Overcoming the Legacy of Exclusion in South Africa.* World Bank Group. Available at: https://documents1.worldbank.org/curated/en/661661529320791504/pdf/127303-Wage-Setting-and-Labor-Regulatory-Challenges-in-a-Middle-Income-Country-Setting-The-Case-of-South-Africa.pdf. [Accessed 30 October 2023].

Bhorat, H., Lilenstein, K., Oosthuizen, M. & Thornton, A. 2020. *Wage polarization in a high-inequality emerging economy: The case of South Africa*. WIDER Working Paper 2020/55. Helsinki: UNU-WIDER. DOI: 10.35188/UNU-WIDER/2020/812-2.

Booth, A.L. & Katic, P. 2011. Estimating the wage elasticity of labour supply to a firm: What evidence is there for monopsony? *Economic Record*. 87(278):359–369. DOI: 10.1111/j.1475-4932.2011.00728.x.

Broecke, S., Forti, A. & Vandeweyer, M. 2017. The effect of minimum wages on employment in emerging economies: a survey and meta-analysis. *Oxford Development Studies*. 45(3):366–391. DOI: 10.1080/13600818.2017.1279134.

Brown, J.N. and Light, A. 1992. Interpreting Panel Data on Job Tenure. *Journal of Labor Economics*. 10(3):219–257. DOI: 10.1086/298286.

Burdett, K. & Mortensen, D.T. 1998. Wage Differentials, Employer Size, and Unemployment. *International Economic Review*. 39(2):257. DOI: 10.2307/2527292.

Caldwell, S. & Oehlsen, E. 2018. *Monopsony and the Gender Wage Gap: Experimental Evidence from the Gig Economy*. Massachusetts Institute of Technology Working Paper. Cambridge. Available: https://www.haas.berkeley.edu/wp-content/uploads/Sydnee-Caldwell-Monopsony-and-the-Gender-Wage-Gap.pdf [2023, March 15].

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. DOI: 10.1093/qje/qjt006.

Casale, D. & Posel, D. 2011. Unions and the Gender Wage Gap in South Africa. *Journal of African Economies*. 20(1):27–59. DOI: 10.1093/jae/ejq029.

Cingano, F. 2014. *Trends in Income Inequality and its Impact of Economic Growth.* OECD Social, Employment and Migration Working Papers 163. Paris: OECD. DOI: 10.1787/5jxrjncwxv6j-en

Dal Bó, E., Finan, F. & Rossi, M.A. 2013. Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service. *The Quarterly Journal of Economics*. 128(3):1169–1218. DOI: 10.1093/qje/qjt008.

David, A., Guilbert, N., Hamaguchi, N., Higashi, Y., Hino, H., Leibbrandt, M. & Shifa, M. 2018. *Spatial poverty and inequality in South Africa: A municipality level analysis*. SALDRU Working Paper 221. Cape Town: SALDRU, UCT.

Doucouliagos, H. & Stanley, T.D. 2009. Publication Selection Bias in Minimum-Wage Research? A Meta-Regression Analysis. *British Journal of Industrial Relations*. 47(2):406–428. DOI: 10.1111/j.1467-8543.2009.00723.x.

Dube, A., Jacobs, J., Naidu, S. & Suri, S. 2020. Monopsony in Online Labor Markets. *American Economic Review: Insights*. 2(1):33–46. DOI: 10.1257/aeri.20180150.

Ehrenberg, R.G., Smith, R.S. & Hallock, K.F. 2021. *Modern Labor Economics*. 14th ed. New York: Routledge. DOI: 10.4324/9780429327209.

Eidelberg Phil. 1997. South African Apartheid: The Homeland-Township Nexus, 1948–1986. *South African Historical Journal*. 36(1):88–112. DOI: 10.1080/02582479708671270.

Finn, A. and Ranchhod, V., 2017. *Short-run differences between static and dynamic measures of earnings inequality in South Africa.* REDI3x3 Working Paper 35. Cape Town: The Research Project on Employment, Income Distribution and Inclusive Growth.

Fuchs-Schündeln, N., Krueger, D. & Sommer, M. 2010. Inequality trends for Germany in the last two decades: A tale of two countries. *Review of Economic Dynamics*. 13(1):103–132. DOI: 10.1016/j.red.2009.09.004.

Guvenen, F., Pistaferri, L. & Violante, G.L. 2022. Global trends in income inequality and income dynamics: New insights from GRID. *Quantitative Economics*. 13(4):1321–1360. DOI: 10.3982/QE2260.

Hausman, J.A., Abrevaya, J. & Scott-Morton, F.M. 1998. Misclassification of the dependent variable in a discrete-response setting. *Journal of Econometrics*. 87(2):239–269. DOI: 10.1016/S0304-4076(98)00015-3.

Hirsch, B. & Jahn, E.J. 2015. Is There Monopsonistic Discrimination against Immigrants?. *ILR Review*. 68(3):501–528. DOI: 10.1177/0019793915572124.

Karaca-Mandic, P., Norton, E.C. & Dowd, B. 2012. Interaction Terms in Nonlinear Models. *Health Services Research*. 47(1pt1):255–274. DOI: 10.1111/j.1475-6773.2011.01314.x.

Kerr, A. 2017. Tax(i)ing the Poor? Commuting Costs in South African Cities. *South African Journal of Economics*. 85(3):321–340. DOI: 10.1111/saje.12161.

Kerr, A. 2018. Job Flows, Worker Flows and Churning in South Africa. *South African Journal of Economics*. 86(S1):141–166. DOI: 10.1111/saje.12168.

Kerr, A. Lam, D. and M. Wittenberg. 2019. *Post-Apartheid Labour Market Series 1993-2019* [dataset]. Version 3.3.1. Cape Town: DataFirst [producer and distributor], DOI: 10.25828/gtr1-8r20.

Kerr, A. & Wittenberg, M. 2021. Union wage premia and wage inequality in South Africa. *Economic Modelling*. 97:255–271. DOI: 10.1016/j.econmod.2020.12.005.

Kerr, A. 2023. *Measuring earnings inequality in South Africa using household survey and administrative tax microdata.* Unpublished Working Paper.

Langella, M. & Manning, A. 2021. Marshall Lecture 2020: The Measure of Monopsony. *Journal of the European Economic Association*. 19(6):2929–2957. DOI: 10.1093/jeea/jvab039.

Leung, R., Stampini, M. & Vencatachellum, D. 2014. Does Human Capital Protect Workers against Exogenous Shocks? Evidence from Panel Data on South Africa during the 2008-2009 Crisis. *South African Journal of Economics*. 82(1):99–116. DOI: 10.1111/saje.12029.

Lucas, R.E. 1985. Mines and Migrants in South Africa. *The American Economic Review*. 75(5):1094–1108. Available: https://www.jstor.org/stable/1818648 [2023, October 30].

Manning, A. 2003. *Monopsony in Motion*. Princeton University Press. DOI: 10.1515/9781400850679.

Manning, A. 2011. 'Imperfect Competition in the Labor Market', in D. Card and O. Ashenfelter (eds) *Handbook Labor of Economics*. 1st edn. Saint Louis: Elsevier Science & Technology, pp. 973–1041. DOI: 10.1016/S0169-7218(11)02409-9.

Manning, A. 2020. Monopsony in Labor Markets: A Review. *ILR Review*. 74(1):3–26. DOI: 10.1177/0019793920922499.

Mosomi, J. 2019. An empirical analysis of trends in female labour force participation and the gender wage gap in South Africa. *Agenda*. 33(4):29–43. DOI: 10.1080/10130950.2019.1656090.

Organisation for Economic Co-operation and Development. 2011. *Divided We Stand*. OECD. DOI: 10.1787/9789264119536-en.

Pan, S. 2022. *Worker Flows in South Africa from 2008 to 2019: Evidence from Matched Quarterly Labour Force Survey*. Unpublished M.Com Dissertation. University of Cape Town.

Ranchhod, V. & Dinkelman, T. (2008) *Labour Market Transitions in South Africa: What can we learn from matched Labour Force Survey data?*. Southern Africa Labour and Development Research Unit Working Paper Number 14. Cape Town: SALDRU, University of Cape Town.

Ransom, M.R. & Oaxaca, R.L. 2010. New Market Power Models and Sex Differences in Pay. *Journal of Labor Economics*. 28(2). DOI: 10.1086/651245

Ransom, M.R. & Sims, D.P. 2010. Estimating the Firm's Labor Supply Curve in a "New Monopsony" Framework: Schoolteachers in Missouri. *Journal of Labor Economics*. 28(2). DOI: 10.1086/649904

Rinz, K. 2018. *Labor Market Concentration, Earnings Inequality, and Earnings Mobility Labor Market Concentration, Earnings Inequality, and Earnings Mobility*. CARRA Working Paper Series 2018–10. Washington: Centre for Administrative Records Research and Applications, US Census Bureau. Available: https://irs.princeton.edu/sites/g/files/toruqf276/files/concentration.pdf [2023, January 26].

Robinson, J. 1933. *The Economics of Imperfect Competition*. London: Palgrave Macmillan UK. DOI: 10.1007/978-1-349-15320-6.

Rodriguez, F. & Jayadev, A. 2013. The Declining Labor Share of Income. *Journal of Globalization and Development*. 3(2):1–18. DOI: 10.1515/jgd-2012-0028.

Shah, K. & Sturzenegger, F. 2022. *Search, Transport Costs, and Labor Markets in South Africa*. CID Research Fellow and Graduate Student Working Paper Series 2022.142. Cambridge, MA: Harvard University. Available: https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37374949 [2023, 27 November].

Sokolova, A. & Sorensen, T. 2021. Monopsony in Labor Markets: A Meta-Analysis. *ILR Review*. 74(1):27–55. DOI: 10.1177/0019793920965562.

Statistics South Africa (2001) 'Labour Force Survey, September 2001', Statistical Release. Pretoria. Available at: https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/126/download/1498 (Accessed: 5 October 2023).

Statistics South Africa (2002) 'Labour Force Survey, February 2002', Statistical Release. Pretoria. Available at: https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/150/download/1951 (Accessed: 5 October 2023).

Statistics South Africa (2011) 'Quarterly Labour Force Survey, Quarter 1 2011', Statistical Release. Pretoria. Available at: https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/215/download/3119 (Accessed: 5 October 2023).

Tharwat, A. 2021. Classification assessment methods. *Applied Computing and Informatics*. 17(1):168–192. DOI: 10.1016/j.aci.2018.08.003.

Vick, B. 2017. Measuring links between labor monopsony and the gender pay gap in Brazil. *IZA Journal of Development and Migration*. 7(1):10. DOI: 10.1186/s40176-017-0099-x.

Wilson, F. 1972. *Labour in the South African gold mines 1911-1969*. Cambridge: University Press.

Wittenberg, M., 2008. *Nonparametric estimation when income is reported in bands and at points.* Cape Town: Economic Research Southern Africa Working Paper no. 94.

Wittenberg, M. 2016. *Trends in Earnings and Earnings Inequality in South Africa: 1993–2014*. Southern Africa Labour and Development Research Unit.

Wittenberg, M. 2017. Wages and Wage Inequality in South Africa 1994-2011: Part 1 – Wage Measurement and Trends. *South Africa Journal of Economics.* 85(2): 279-297. DOI: 10.1111/saje.12148.

Wooldridge, J. M. 2010. *Econometric Analysis of Cross Section and Panel Data*. Second Edition. 1st edition. Cambridge: MIT Press.

Wooldridge, J. M. 2019. *Introductory Econometrics: A Modern Approach*. 7th edition. Mason, OH: Cengage.

13. Appendices

13.1. Appendix A: Attrition Regressions

Table A1: Average Partial Effects - Attrition Logit

VARIABLES	(1) LFS '01 Q2	(2) LFS '04 Q2	(3) LFS '06 Q2	(1) QLFS `11 Q1	(2) QLFS `18 Q1	(3) QLFS `19 Q3
Age 25-34	-0.0166***	-0.0750***	-0.0412***	0.0127*	0.0142*	0.00982
	(0.00574)	(0.00548)	(0.00530)	(0.00683)	(0.00745)	(0.00752)
Age 35-44	0.0345***	-0.0329***	-0.00769	0.0442***	0.0533***	0.0342***
	(0.00685)	(0.00661)	(0.00631)	(0.00789)	(0.00843)	(0.00846)
Age 45-54	0.0841***	-0.00273	0.0302***	0.0594***	0.0671***	0.0684***
	(0.00787)	(0.00756)	(0.00729)	(0.00873)	(0.00927)	(0.00921)
Age 55-64	0.0482***	-6.09e-06	0.0387***	0.0725***	0.0712***	0.0696***
	(0.00909)	(0.00870)	(0.00842)	(0.00989)	(0.0105)	(0.0102)
Female	-0.00202	0.00791*	0.00435	0.00930*	0.0132**	0.00541
	(0.00459)	(0.00442)	(0.00432)	(0.00553)	(0.00605)	(0.00612)
Coloured	0.0492***	0.0786***	0.0763***	0.0850***	-0.00781	-0.0306***
	(0.00810)	(0.00748)	(0.00739)	(0.00904)	(0.0102)	(0.0102)
Indian/Asian	0.136***	0.0692***	0.0791***	-0.00857	0.0366**	-0.0165
	(0.0131)	(0.0138)	(0.0135)	(0.0152)	(0.0167)	(0.0167)
White	-0.0389***	0.0639***	0.0306***	0.0434***	-0.00167	0.00436
	(0.00780)	(0.00812)	(0.00838)	(0.00945)	(0.0109)	(0.0112)
Some Primary	0.0395***	-0.00802	-0.0288***	-0.0573***	-0.0342**	0.0131
	(0.00745)	(0.00718)	(0.00733)	(0.0112)	(0.0152)	(0.0156)
Some Secondary	0.0896***	0.0321***	0.00255	-0.0436***	-0.00871	0.0128
	(0.00757)	(0.00730)	(0.00743)	(0.0110)	(0.0146)	(0.0149)
Matric	0.0959***	0.0705***	0.0297***	-0.0342***	-0.0167	0.0244
	(0.00862)	(0.00835)	(0.00833)	(0.0117)	(0.0151)	(0.0154)
Some Tertiary	0.138***	0.0493***	0.00702	-0.0409***	-0.0279*	-0.0173
	(0.0105)	(0.0106)	(0.0104)	(0.0132)	(0.0162)	(0.0164)
Widow/er	-0.0129	-0.0399***	-0.0164*	-0.0218*	-0.0226*	-0.0391***
	(0.0103)	(0.00929)	(0.00890)	(0.0114)	(0.0137)	(0.0139)
Divorced/Separated	-0.0245**	-0.0616***	-0.0573***	-0.0325**	-0.0672***	-0.0277*
	(0.0115)	(0.0107)	(0.0112)	(0.0138)	(0.0161)	(0.0162)
Never Married	-0.0212***	0.00550	0.0171***	-0.0132**	-0.00876	0.00277
	(0.00550)	(0.00529)	(0.00505)	(0.00618)	(0.00632)	(0.00632)
Employed	-0.0219***	-0.0325***	-0.0355***	0.0143**	0.0147**	0.0114*
	(0.00501)	(0.00471)	(0.00458)	(0.00577)	(0.00627)	(0.00638)
Unemployed	-0.0184***	-0.0436***	-0.0413***	-0.00475	0.00282	0.00795
	(0.00582)	(0.00581)	(0.00555)	(0.00737)	(0.00755)	(0.00741)
Not In Person	0.0112***	-0.0283***	-0.0179***	-0.00623	0.00203	0.00465

	(0.00411)	(0.00394)	(0.00383)	(0.00487)	(0.00532)	(0.00534)
In Person Unspec	0.0151	0.0354	-0.120**	0.151**	0.0762	-0.0773
	(0.0300)	(0.0569)	(0.0535)	(0.0674)	(0.0675)	(0.0578)
HH Size	0.00528***	-0.00411***	-0.00198***	-0.00638***	-0.00931***	-0.00909***
	(0.000747)	(0.000769)	(0.000735)	(0.00101)	(0.00111)	(0.00117)
Eastern Cape	0.0432***	0.0225***	0.111***	0.0270**	-0.0887***	-0.104***
	(0.00911)	(0.00859)	(0.00833)	(0.0107)	(0.0107)	(0.0109)
Northern Cape	0.0493***	-0.0113	0.0233***	-0.0330***	-0.0942***	-0.0694***
	(0.0108)	(0.00897)	(0.00840)	(0.0121)	(0.0144)	(0.0140)
Free State	0.0217**	0.0299***	0.0565***	0.0822***	-0.0921***	-0.0608***
	(0.0101)	(0.00982)	(0.00943)	(0.0114)	(0.0128)	(0.0130)
KZN	-0.0151*	-0.0796***	-0.0110	0.0945***	-0.0676***	-0.0416***
	(0.00912)	(0.00817)	(0.00772)	(0.0105)	(0.0104)	(0.0106)
North West	0.0780***	0.0610***	0.0987***	0.0664***	-0.123***	-0.139***
	(0.00986)	(0.00973)	(0.00939)	(0.0116)	(0.0129)	(0.0134)
Gauteng	-0.0842***	-0.0501***	-0.0281***	0.0639***	-0.118***	-0.111***
	(0.00890)	(0.00892)	(0.00846)	(0.0103)	(0.00976)	(0.00979)
Mpumalanga	0.0257**	0.0423***	0.0143	0.0927***	-0.0687***	-0.0749***
	(0.0102)	(0.0101)	(0.00944)	(0.0114)	(0.0120)	(0.0122)
Limpopo	-0.0253***	0.0841***	0.128***	0.0715***	-0.0996***	-0.0629***
	(0.00977)	(0.00978)	(0.00942)	(0.0112)	(0.0116)	(0.0115)
HH Female Child %	0.109***	0.152***	0.133***	0.0980***	0.100***	0.174***
	(0.0327)	(0.0305)	(0.0287)	(0.0338)	(0.0373)	(0.0356)
HH Male Child %	0.0956***	0.173***	0.0947***	0.0528	0.112***	0.153***
	(0.0327)	(0.0304)	(0.0287)	(0.0337)	(0.0373)	(0.0355)
HH Working Male %	-0.0931***	-0.0575**	-0.0714***	-0.0180	-0.0540	-0.00527
	(0.0310)	(0.0289)	(0.0270)	(0.0313)	(0.0345)	(0.0327)
HH Working Female %	-0.0211	0.0122	-0.0218	0.0117	-0.0371	0.0178
	(0.0308)	(0.0286)	(0.0267)	(0.0313)	(0.0344)	(0.0326)
HH Male Pensioner %	-0.115**	-0.101**	-0.0965**	0.0794	-0.0218	0.114**
	(0.0500)	(0.0467)	(0.0438)	(0.0522)	(0.0571)	(0.0534)
Observations	65,285	67,277	65,998	50,627	42,568	42,080

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Base Categories: Age - 15 to 24, Race - African, Education Category - No education, Marriage Status - Married or living together as husband and wife, Employment category - Not Economically Active, In Person Response - Responded in person, Province - Western Cape, HH proportions - Female Pensioners. Data Source: StatsSA.

13.2. Appendix B: Main Regression Results

Table B1: Main Regressions - No Controls

	LFS – 6 Months				QLFS – 6 Months			
VARIABLES	(1) All Separations	(2) Separations to Employment	(3) Separations to Non-Employment	(4) Hired from Employment	(5) All Separations	(6) Separations to Employment	(7) Separations to Non-Employment	(8) Hired from Employment
Log Real Wage	-0.0809***	-0.0234***	-0.0649***	0.303***	-0.0629***	-0.0233***	-0.0435***	0.131
	(0.00242)	(0.00192)	(0.00206)	(0.0505)	(0.00291)	(0.00205)	(0.00223)	(0.0859)
Constant				-1.703***				-0.815***
				(0.122)				(0.237)
Observations	46,811	39,692	46,811	10,649	20,729	18,956	20,729	2,537

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Columns 1-3 and 5-7 are average partial effects. Columns 4 & 8 are logit coefficients. Columns 1-3 & 5-7 controls are lagged by 6 months. Column 4 & 8 controls are contemporaneous. Survey weights used.

Table B2: Main Regressions - Med Controls

	LFS				QLFS			
VARIABLES	(1) All Separations	(2) Separations to Employment	(3) Separations to Non-Employment	(4) Hired from Employment	(1) All Separations	(2) Separations to Employment	(3) Separations to Non-Employment	(4) Hired from Employment
Log Real Wage	-0.0764***	-0.0260***	-0.0577***	0.260***	-0.0450***	-0.0173***	-0.0306***	0.0832
	(0.00302)	(0.00221)	(0.00266)	(0.0540)	(0.00364)	(0.00279)	(0.00266)	(0.0900)
Age	-0.0251***	-0.00319***	-0.0222***	0.195***	-0.0127***	-0.00158	-0.0115***	0.153***
	(0.00194)	(0.00122)	(0.00169)	(0.0238)	(0.00181)	(0.00133)	(0.00139)	(0.0414)
Age Squared	0.000279***	1.24e-05	0.000262***	-0.00271***	0.000130***	5.90e-06	0.000126***	-0.00171***
	(2.57e-05)	(1.61e-05)	(2.24e-05)	(0.000311)	(2.31e-05)	(1.68e-05)	(1.78e-05)	(0.000534)
Years of Education	-0.000651	-0.000778	-3.57e-05	-0.0338***	-0.00560***	-0.00327***	-0.00274***	-0.0217
	(0.000826)	(0.000561)	(0.000711)	(0.0128)	(0.000952)	(0.000667)	(0.000750)	(0.0242)
Coloured	-0.00262	0.0122**	-0.0162**	0.298**	0.0515***	0.0223**	0.0332***	0.478**
	(0.00867)	(0.00568)	(0.00763)	(0.118)	(0.0141)	(0.0105)	(0.0117)	(0.243)
Indian/Asian	0.0372**	0.0290**	0.0150	0.690***	0.0193	0.0273	-0.0116	1.720***
	(0.0158)	(0.0140)	(0.0136)	(0.224)	(0.0317)	(0.0250)	(0.0222)	(0.197)
White	0.00669	0.0221**	-0.0165	0.611***	-0.0582***	-0.0146	-0.0491***	0.425
	(0.0130)	(0.0108)	(0.0110)	(0.200)	(0.0125)	(0.0111)	(0.00584)	(0.259)
Female	0.0233***	-0.00760**	0.0329***	-0.487***	0.00568	-0.00590	0.0124***	-0.320**
	(0.00521)	(0.00369)	(0.00453)	(0.0813)	(0.00535)	(0.00384)	(0.00425)	(0.126)
Widow/er	0.00568	0.00203	0.00282	0.498***	0.000537	-0.0112	0.00910	-0.187
	(0.0109)	(0.00876)	(0.00908)	(0.176)	(0.0134)	(0.00763)	(0.0113)	(0.379)
Divorced	0.0258	0.0142	0.0161	0.0813	0.00564	-0.00579	0.0119	-0.0360
	(0.0168)	(0.00982)	(0.0152)	(0.193)	(0.0140)	(0.00880)	(0.0118)	(0.381)
Never Married	0.0400***	0.0101**	0.0350***	-0.319***	0.0250***	0.00918**	0.0182***	-0.182
	(0.00653)	(0.00473)	(0.00531)	(0.0968)	(0.00587)	(0.00450)	(0.00434)	(0.146)
Constant				-4.118***				-3.260***
				(0.506)				(0.893)
Observations	46,811	39,692	46,811	10,623	20,729	18,956	20,729	2,537

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Wave and Province Results Excluded. Columns 1-3 & 5-7 are average partial effects. Column 4 & 8 are logit coefficients. Columns 1-3 & 5-7 controls are lagged by 6 months.. Column 4 & 8 controls are contemporaneous. Survey weights used.

Table B3: Main Regressions - Med+ Controls

	LFS				QLFS			
VARIABLES	(1) All Separations	(2) Separations to Employment	(3) Separations to Non-Employment	(4) Hired from Employment	(1) All Separations	(2) Separations to Employment	(3) Separations to Non-Employment	(4) Hired from Employment
Log Real Wage	-0.0624***	-0.0210***	-0.0478***	0.271***	-0.0381***	-0.0145***	-0.0261***	0.0352
	(0.00377)	(0.00306)	(0.00311)	(0.0523)	(0.00411)	(0.00322)	(0.00303)	(0.0894)
Age	-0.0250***	-0.00323***	-0.0222***	0.192***	-0.0133***	-0.00187	-0.0118***	0.148***
	(0.00192)	(0.00123)	(0.00168)	(0.0238)	(0.00180)	(0.00131)	(0.00138)	(0.0417)
Age Squared	0.000282***	1.53e-05	0.000265***	-0.00265***	0.000140***	1.08e-05	0.000132***	-0.00166***
	(2.54e-05)	(1.62e-05)	(2.22e-05)	(0.000310)	(2.28e-05)	(1.65e-05)	(1.75e-05)	(0.000536)
Years of Education	0.000943	-0.000106	0.000936	-0.0201	-0.00313***	-0.00194***	-0.00137*	-0.0256
	(0.000903)	(0.000614)	(0.000777)	(0.0127)	(0.000990)	(0.000702)	(0.000787)	(0.0250)
Coloured	0.00556	0.0155***	-0.0102	0.292**	0.0512***	0.0231**	0.0325***	0.485**
	(0.00863)	(0.00563)	(0.00757)	(0.121)	(0.0142)	(0.0109)	(0.0118)	(0.246)
Indian/Asian	0.0320**	0.0274**	0.0105	0.646***	0.0372	0.0378	-0.00243	1.856***
	(0.0157)	(0.0139)	(0.0134)	(0.230)	(0.0343)	(0.0279)	(0.0251)	(0.280)
White	0.00424	0.0199	-0.0169	0.543**	-0.0626***	-0.0184*	-0.0502***	0.365
	(0.0134)	(0.0122)	(0.0105)	(0.215)	(0.0119)	(0.0101)	(0.00592)	(0.289)
Female	0.0442***	0.00172	0.0476***	-0.449***	0.0284***	0.00443	0.0269***	-0.315**
	(0.00624)	(0.00444)	(0.00553)	(0.0949)	(0.00658)	(0.00482)	(0.00534)	(0.150)
Widow/er	0.00240	0.00127	0.000116	0.497***	-0.00159	-0.0119	0.00725	-0.172
	(0.0108)	(0.00870)	(0.00905)	(0.176)	(0.0131)	(0.00749)	(0.0111)	(0.384)
Divorced	0.0185	0.0127	0.00994	0.118	0.00173	-0.00700	0.00940	-0.150
	(0.0161)	(0.00954)	(0.0144)	(0.196)	(0.0135)	(0.00855)	(0.0115)	(0.398)
Never Married	0.0365***	0.00965**	0.0319***	-0.301***	0.0238***	0.00891**	0.0173***	-0.190
	(0.00654)	(0.00464)	(0.00532)	(0.0940)	(0.00578)	(0.00447)	(0.00425)	(0.146)
Professional	0.00483	-0.00458	0.00987	-0.352	0.0166	0.0138	0.00726	-1.095*
	(0.0343)	(0.0143)	(0.0343)	(0.526)	(0.0389)	(0.0308)	(0.0313)	(0.620)
Technical/Assoc. Prof.	0.00714	0.0188	-0.00914	0.221	-0.0310	-0.00106	-0.0287	-0.524
	(0.0289)	(0.0161)	(0.0271)	(0.446)	(0.0296)	(0.0218)	(0.0243)	(0.524)
Clerks	-0.00550	0.00730	-0.0114	-0.320	-0.0242	-0.0110	-0.0121	-0.392
	(0.0282)	(0.0136)	(0.0270)	(0.435)	(0.0281)	(0.0193)	(0.0243)	(0.472)
Service Workers	0.00822	0.00911	0.00161	-0.337	-0.00355	-0.000513	0.000151	-1.028**
	(0.0281)	(0.0141)	(0.0266)	(0.427)	(0.0279)	(0.0192)	(0.0240)	(0.451)
Skilled Agri	0.0716*	0.0337	0.0474	-0.0921	0.0673	0.0536	0.0255	-1.902***
	(0.0376)	(0.0233)	(0.0334)	(0.510)	(0.0559)	(0.0518)	(0.0410)	(0.730)
Craft/Trade	0.0477*	0.0337**	0.0203	0.0387	0.00926	0.00522	0.00802	-0.645
	(0.0277)	(0.0156)	(0.0249)	(0.423)	(0.0288)	(0.0206)	(0.0241)	(0.472)
Machine Worker	0.0171	0.0206	-0.000379	-0.0624	-0.0322	-0.0156	-0.0152	-1.046**
	(0.0321)	(0.0234)	(0.0254)	(0.416)	(0.0277)	(0.0189)	(0.0238)	(0.532)
Elementary Occ.	0.0357	0.0141	0.0261	-0.311	0.0114	0.00499	0.0110	-0.822*
	(0.0284)	(0.0161)	(0.0255)	(0.420)	(0.0277)	(0.0191)	(0.0238)	(0.447)
Domestic	-0.00166	-0.00250	0.00205	-0.212	-0.0475	-0.0232	-0.0216	-0.898
	(0.0336)	(0.0183)	(0.0302)	(0.478)	(0.0296)	(0.0200)	(0.0259)	(0.566)
Mining	-0.0121	-0.0160*	0.00288	-0.280	-0.0334*	-0.0115	-0.0247*	-0.0967

	(0.0131)	(0.00823)	(0.0112)	(0.251)	(0.0182)	(0.0119)	(0.0130)	(0.570)
Manufacturing	0.0363***	-0.00300	0.0435***	-0.758***	0.0132	0.0118	0.00120	-0.159
	(0.00983)	(0.00695)	(0.00789)	(0.166)	(0.0125)	(0.00953)	(0.00924)	(0.325)
Utilities	0.0987***	-0.00811	0.121***	-0.440	-0.0125	-0.00920	-0.00529	0.157
	(0.0334)	(0.0184)	(0.0342)	(0.485)	(0.0326)	(0.0214)	(0.0268)	(2.036)
Construction	0.184***	0.0604***	0.148***	-0.656***	0.115***	0.0524***	0.0767***	-0.284
	(0.0139)	(0.0105)	(0.0121)	(0.164)	(0.0153)	(0.0122)	(0.0124)	(0.318)
Trade	0.0543***	0.00859	0.0504***	-0.630***	-0.0140	-0.00810	-0.00666	-0.365
	(0.0102)	(0.00783)	(0.00768)	(0.152)	(0.0103)	(0.00709)	(0.00833)	(0.289)
Transport	0.0927***	0.0159	0.0871***	-0.726***	0.0305*	0.0199*	0.0117	0.820**
	(0.0159)	(0.0117)	(0.0135)	(0.213)	(0.0161)	(0.0114)	(0.0126)	(0.414)
Finance	0.0645***	0.0279***	0.0416***	-0.313	-0.00887	-0.00297	-0.00716	0.269
	(0.0128)	(0.0100)	(0.00981)	(0.196)	(0.0112)	(0.00797)	(0.00849)	(0.297)
Services	0.0339***	-0.00240	0.0378***	-0.542***	0.0109	0.00151	0.00953	-0.182
	(0.0111)	(0.00707)	(0.00943)	(0.204)	(0.0116)	(0.00822)	(0.00900)	(0.314)
Domestic Services	0.0765***	0.0308**	0.0530***	-0.481**	0.0552***	0.0415***	0.0181	0.0622
	(0.0175)	(0.0138)	(0.0133)	(0.211)	(0.0161)	(0.0125)	(0.0120)	(0.345)
Public Sector	-0.0599***	-0.0193***	-0.0475***		-0.0380***	-0.0213***	-0.0212***	-0.0879
	(0.00841)	(0.00514)	(0.00727)		(0.00787)	(0.00538)	(0.00620)	(0.234)
Constant				-3.549***				-2.114**
				(0.662)				(1.065)
Observations	46,702	39,605	46,702	10,589	20,705	18,933	20,705	2,534

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Wave and Province Results Excluded. Columns 1-3 & 5-7 are average partial effects. Column 4 & 8 are logit coefficients. Columns 1-3 & 5-7 controls are lagged by 6 months.. Column 4 & 8 controls are contemporaneous. Survey weights used.

13.3. Appendix C: Gender Regression Results

Table C1: All Separations & Separations to Employment

	LFS		QLFS		LFS		QLFS	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Male	Female	Male	Female	Male	Female	Male	Female
VARIABLES	All Separations				Separations to Employment			
Log Real Wage	-0.0730***	-0.0807***	-0.0449***	-0.0452***	-0.0277***	-0.0239***	-0.0186***	-0.0158***
	(0.00397)	(0.00388)	(0.00498)	(0.00454)	(0.00316)	(0.00258)	(0.00403)	(0.00293)
Female	0.0233***	0.0233***	0.00570	0.00570	-0.00760**	-0.00760**	-0.00589	-0.00589
	(0.00521)	(0.00521)	(0.00536)	(0.00536)	(0.00370)	(0.00370)	(0.00385)	(0.00385)
Age	-0.0241***	-0.0264***	-0.0125***	-0.0130***	-0.0033***	-0.00295**	-0.00165	-0.00147
	(0.00181)	(0.00212)	(0.00175)	(0.00191)	(0.00126)	(0.00115)	(0.00138)	(0.00126)
Age squared	0.000268***	0.000294***	0.000127***	0.000133***	1.29e-05	1.14e-05	6.10e-06	5.45e-06
	(2.39e-05)	(2.78e-05)	(2.23e-05)	(2.41e-05)	(1.67e-05)	(1.49e-05)	(1.75e-05)	(1.57e-05)
Years of Education	-0.000622	-0.000681	-0.00549***	-0.00573***	-0.000821	-0.000725	-0.0034***	-0.0031***
	(0.000790)	(0.000867)	(0.000936)	(0.000985)	(0.000590)	(0.000525)	(0.000698)	(0.000655)
Coloured	-0.00250	-0.00275	0.0506***	0.0526***	0.0128**	0.0113**	0.0234**	0.0210**
	(0.00832)	(0.00912)	(0.0140)	(0.0144)	(0.00601)	(0.00528)	(0.0111)	(0.00984)
Indian/Asian	0.0357**	0.0389**	0.0189	0.0197	0.0306**	0.0271**	0.0286	0.0257
	(0.0153)	(0.0166)	(0.0310)	(0.0324)	(0.0147)	(0.0132)	(0.0261)	(0.0238)
White	0.00639	0.00700	-0.0570***	-0.0597***	0.0233**	0.0206**	-0.0154	-0.0137
	(0.0125)	(0.0137)	(0.0122)	(0.0132)	(0.0114)	(0.00997)	(0.0117)	(0.0105)
Widow/er	0.00535	0.00590	0.000592	0.000618	0.00224	0.00197	-0.0117	-0.0104
	(0.0104)	(0.0115)	(0.0131)	(0.0137)	(0.00920)	(0.00809)	(0.00798)	(0.00724)
Divorced	0.0247	0.0272	0.00530	0.00554	0.0150	0.0133	-0.00620	-0.00552
	(0.0161)	(0.0177)	(0.0138)	(0.0143)	(0.0104)	(0.00911)	(0.00929)	(0.00833)
Never Married	0.0384***	0.0421***	0.0245***	0.0255***	0.0106**	0.00939**	0.00963**	0.00860**
	(0.00633)	(0.00689)	(0.00585)	(0.00597)	(0.00502)	(0.00445)	(0.00480)	(0.00420)
Observations	46,811		20,729		39,692		18,956	

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Wave and Province Results Excluded. All controls are lagged by 6 months. Survey weights used.

Table C2: Separations to Non-Employment & Hires from Employment

Dataset	LFS		QLFS		LFS	QLFS
	(1)	(2)	(3)	(4)	(5)	(6)
	Male	Female	Male	Female	All	All
VARIABLES	Separations to Non-Employment				Hires from Employment	
Log Real Wage	-0.0525***	-0.0641***	-0.0289***	-0.0326***	0.212***	0.00235
	(0.00328)	(0.00350)	(0.00323)	(0.00380)	(0.0722)	(0.124)
Female	0.0329***	0.0329***	0.0124***	0.0124***	-0.766***	-0.831*
	(0.00453)	(0.00453)	(0.00425)	(0.00425)	(0.243)	(0.459)
Female * Log Real Wage					0.102	0.184
					(0.0990)	(0.165)
Age	-0.0203***	-0.0246***	-0.0106***	-0.0124***	0.197***	0.155***
	(0.00150)	(0.00196)	(0.00127)	(0.00158)	(0.0237)	(0.0412)
Age Squared	0.000239***	0.000291***	0.000117***	0.000136***	-0.00273***	-0.00174***
	(1.99e-05)	(2.58e-05)	(1.63e-05)	(2.00e-05)	(0.000309)	(0.000533)
Years of Education	-3.19e-05	-3.87e-05	-0.00254***	-0.00296***	-0.0341***	-0.0222
	(0.000647)	(0.000786)	(0.000697)	(0.000821)	(0.0129)	(0.0244)
Coloured	-0.0148**	-0.0180**	0.0309***	0.0359***	0.300**	0.486**
	(0.00693)	(0.00848)	(0.0111)	(0.0126)	(0.118)	(0.243)
Indian/Asian	0.0137	0.0165	-0.0107	-0.0126	0.698***	1.660***
	(0.0125)	(0.0150)	(0.0205)	(0.0240)	(0.224)	(0.212)
White	-0.0150	-0.0183	-0.0453***	-0.0535***	0.610***	0.434*
	(0.0100)	(0.0123)	(0.00539)	(0.00681)	(0.199)	(0.258)
Widow/er	0.00253	0.00312	0.00847	0.00990	0.522***	-0.167
	(0.00822)	(0.0101)	(0.0105)	(0.0123)	(0.175)	(0.381)
Divorced	0.0146	0.0179	0.0109	0.0127	0.0880	-0.0772
	(0.0139)	(0.0170)	(0.0111)	(0.0128)	(0.193)	(0.388)
Never Married	0.0319***	0.0389***	0.0169***	0.0197***	-0.317***	-0.183
	(0.00494)	(0.00582)	(0.00413)	(0.00465)	(0.0966)	(0.145)
Constant					-4.005***	-3.078***
					(0.530)	(0.932)
Observations	46,811		20,729		10,623	2,537

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Wave and Province Results Excluded. Columns 1,2, 3 and 4 are average partial effects. Column 5 and 6 are logit coefficients. Columns 1,2, 3 and 4 controls are lagged by 6 months. Column 5 and 6 controls are contemporaneous. Survey weights used.

13.4. Appendix D: Race Regressions

Table D1: All Separations and Separations to Employment across Race Full Results in the LFS

	(1) African	(2) Coloured	(3) Indian/Asian	(4) White	(5) African	(6) Coloured	(7) Indian/Asian	(8) White
VARIABLES	All Separations				Separations to Employment			
Log Real Wage	-0.0786***	-0.0609***	-0.102***	-0.0517***	-0.0250***	-0.0243***	-0.0459**	-0.0246*
	(0.00311)	(0.00620)	(0.0201)	(0.0185)	(0.00219)	(0.00426)	(0.0213)	(0.0148)
Age	-0.0250***	-0.0252***	-0.0290***	-0.0238***	-0.003***	-0.0037***	-0.00478**	-0.0038***
	(0.00197)	(0.00206)	(0.00246)	(0.00202)	(0.00115)	(0.00137)	(0.00187)	(0.00143)
Age Squared	0.000278***	0.000280***	0.000323***	0.000264***	1.25e-05	1.51e-05	1.97e-05	1.58e-05
	(2.57e-05)	(2.68e-05)	(3.15e-05)	(2.51e-05)	(1.49e-05)	(1.79e-05)	(2.33e-05)	(1.85e-05)
Years of Education	-0.000623	-0.000627	-0.000722	-0.000592	-0.000721	-0.000874	-0.00114	-0.000912
	(0.000823)	(0.000826)	(0.000955)	(0.000789)	(0.000510)	(0.000616)	(0.000839)	(0.000669)
Female	0.0233***	0.0234***	0.0270***	0.0221***	-0.00693**	-0.00841**	-0.0109**	-0.00877**
	(0.00515)	(0.00522)	(0.00610)	(0.00501)	(0.00336)	(0.00414)	(0.00548)	(0.00447)
Widow/er	0.00527	0.00528	0.00620	0.00496	0.00174	0.00212	0.00278	0.00221
	(0.0108)	(0.0108)	(0.0127)	(0.0102)	(0.00806)	(0.00982)	(0.0128)	(0.0103)
Divorced	0.0242	0.0243	0.0282	0.0229	0.0127	0.0155	0.0200	0.0161
	(0.0161)	(0.0163)	(0.0187)	(0.0150)	(0.00909)	(0.0110)	(0.0143)	(0.0118)
Never Married	0.0395***	0.0397***	0.0458***	0.0375***	0.00926**	0.0113**	0.0146**	0.0117*
	(0.00651)	(0.00670)	(0.00788)	(0.00700)	(0.00442)	(0.00544)	(0.00729)	(0.00631)
Observations	46,811				39,692			

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Wave and Province Results Excluded. All controls are lagged by 6 months. Survey weights used.

Table D2: All Separations and Separations to Employment across Race Full Results in the QLFS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	African	Coloured	Indian/Asian	White	African	Coloured	Indian/Asian	White
VARIABLES	All Separations				Separations to Employment			
Log Real Wage	-0.0461***	-0.0410**	-0.00492	-0.0387**	-0.0173***	-0.0176	-0.00226	-0.0275
	(0.00347)	(0.0174)	(0.0273)	(0.0183)	(0.00246)	(0.0158)	(0.0193)	(0.0180)
Age	-0.0126***	-0.0171***	-0.0126***	-0.0079***	-0.00154	-0.00222	-0.00204	-0.00145
	(0.00181)	(0.00258)	(0.00315)	(0.00203)	(0.00129)	(0.00186)	(0.00179)	(0.00121)
Age Squared	0.000128***	0.000175***	0.000129***	8.1e-05***	5.90e-06	8.48e-06	7.82e-06	5.56e-06
	(2.29e-05)	(3.20e-05)	(3.49e-05)	(2.21e-05)	(1.62e-05)	(2.33e-05)	(2.15e-05)	(1.49e-05)
Years of Education	-0.00555***	-0.00757***	-0.00557***	-0.0035***	-0.0032***	-0.0046***	-0.00421***	-0.003**
	(0.000931)	(0.00139)	(0.00151)	(0.00108)	(0.000641)	(0.00113)	(0.00158)	(0.00136)
Female	0.00553	0.00754	0.00555	0.00348	-0.00571	-0.00822	-0.00757	-0.00539
	(0.00524)	(0.00713)	(0.00547)	(0.00337)	(0.00367)	(0.00546)	(0.00520)	(0.00410)
Widow/er	0.000662	0.000917	0.000661	0.000409	-0.0107	-0.0156	-0.0143	-0.0101
	(0.0131)	(0.0182)	(0.0131)	(0.00810)	(0.00735)	(0.0108)	(0.0108)	(0.00796)
Divorced	0.00562	0.00778	0.00563	0.00349	-0.00555	-0.00807	-0.00742	-0.00524
	(0.0137)	(0.0189)	(0.0138)	(0.00852)	(0.00847)	(0.0124)	(0.0116)	(0.00833)
Never Married	0.0250***	0.0343***	0.0252***	0.0156***	0.00905**	0.0130**	0.0121*	0.00851
	(0.00573)	(0.00831)	(0.00798)	(0.00526)	(0.00430)	(0.00650)	(0.00714)	(0.00527)
Observations	20729				18956			

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Wave and Province Results Excluded. All controls are lagged by 6 months. Survey weights used.

Table D3: Separations to Non-Employment across Race Full Results in the LFS and QLFS

	LFS				QLFS			
	(1) African	(2) Coloured	(3) Indian/Asian	(4) White	(5) African	(6) Coloured	(7) Indian/Asian	(8) White
VARIABLES	Separations to Non-Employment				Separations to Non-Employment			
Log Real Wage	-0.0607***	-0.0419***	-0.0750***	-0.0349**	-0.0317***	-0.0253***	-0.00412	-0.0156*
	(0.00278)	(0.00532)	(0.0187)	(0.0171)	(0.00272)	(0.00936)	(0.0220)	(0.00941)
Age	-0.0227***	-0.0207***	-0.0253***	-0.0185***	-0.0114***	-0.0160***	-0.00829***	-0.0045***
	(0.00177)	(0.00179)	(0.00249)	(0.00211)	(0.00140)	(0.00228)	(0.00271)	(0.00146)
Age Squared	0.000268***	0.000244***	0.000298***	0.000218***	0.000125***	0.000175***	9.1e-05***	4.9e-05***
	(2.31e-05)	(2.32e-05)	(3.14e-05)	(2.54e-05)	(1.79e-05)	(2.79e-05)	(3.05e-05)	(1.63e-05)
Years of Education	-3.05e-05	-2.78e-05	-3.40e-05	-2.48e-05	-0.00274***	-0.00385***	-0.00200**	-0.00107**
	(0.000731)	(0.000666)	(0.000815)	(0.000595)	(0.000737)	(0.00111)	(0.000809)	(0.000473)
Female	0.0336***	0.0306***	0.0374***	0.0273***	0.0123***	0.0173***	0.00897**	0.00480**
	(0.00461)	(0.00443)	(0.00582)	(0.00462)	(0.00420)	(0.00595)	(0.00420)	(0.00218)
Widow/er	0.00266	0.00240	0.00299	0.00213	0.00907	0.0129	0.00658	0.00348
	(0.00922)	(0.00832)	(0.0104)	(0.00741)	(0.0112)	(0.0158)	(0.00839)	(0.00435)
Divorced	0.0154	0.0139	0.0172	0.0124	0.0116	0.0165	0.00847	0.00448
	(0.0149)	(0.0136)	(0.0167)	(0.0116)	(0.0116)	(0.0165)	(0.00890)	(0.00473)
Never Married	0.0355***	0.0323***	0.0396***	0.0288***	0.0183***	0.0259***	0.0134**	0.00709**
	(0.00536)	(0.00513)	(0.00674)	(0.00554)	(0.00426)	(0.00675)	(0.00520)	(0.00282)
Observations	46,811				20729			

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Wave and Province Results Excluded. All controls are lagged by 6 months. Survey weights used.

Table D5: Hired from Employment Regressions across Race in the LFS and QLFS

	LFS	QLFS
	Hired from Employment	
Log Real Wage	0.223***	0.0597
	(0.0638)	(0.0953)
Coloured	0.232	-0.358
	(0.299)	(0.799)
Indian/Asian	-0.493	-4.259***
	(1.046)	(1.099)
White	-0.873	2.061*
	(0.778)	(1.188)
Coloured*Wage	0.0218	0.282
	(0.109)	(0.255)
Indian/Asian*Wage	0.367	2.137***
	(0.316)	(0.299)
White*Wage	0.391*	-0.412
	(0.218)	(0.279)
Age	0.193***	0.153***
	(0.0237)	(0.0417)
Age Squared	-0.00268***	-0.00172***
	(0.000310)	(0.000538)
Years of Education	-0.0318**	-0.0218
	(0.0126)	(0.0241)
Female	-0.489***	-0.328***
	(0.0818)	(0.126)
Widow/er	0.501***	-0.182
	(0.176)	(0.379)
Divorced	0.0970	0.0202
	(0.192)	(0.384)
Never Married	-0.305***	-0.176
	(0.0960)	(0.146)
Constant	-3.972***	-3.166***
	(0.515)	(0.909)
Observations	10,623	2,537

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Wave and Province Results Excluded. Survey weights used. 6-month interval used for both surveys.

13.5. Appendix E: Interactions Regressions

Table E1: Interactions Regressions in the LFS

Race/Gender	(1) All Separations	Education Category	(2) ... (3) All Separations	(4) Wave	(5) All Separations
African Male	-0.0743***	No Education	-0.0519***	LFS 02:1	-0.0616***
	(0.00403)		(0.00827)		(0.00623)
African Female	-0.0842***	Some Primary	-0.0653***	LFS 02:2	-0.0716***
	(0.00380)		(0.00485)		(0.00530)
Coloured Male	-0.0415***	Some Secondary	-0.0753***	LFS 03:1	-0.0591***
	(0.00779)		(0.00494)		(0.00595)
Coloured Female	-0.0825***	Matric	-0.0782***	LFS 03:2	-0.0677***
	(0.00805)		(0.00600)		(0.00622)
Indian/Asian Male	-0.0845***	Some Tertiary/Degree	-0.0691***	LFS 04:1	-0.0886***
	(0.0246)		(0.0180)		(0.00714)
Indian/Asian Female	-0.120***			LFS 05:1	-0.0868***
	(0.0331)				(0.00782)
White Male	-0.0566**			LFS 05:2	-0.0828***
	(0.0272)				(0.00717)
White Female	-0.0410			LFS 06:2	-0.0884***
	(0.0274)				(0.00839)
				LFS 07:1	-0.0830***
					(0.0126)
				LFS 07:2	-0.0801***
					(0.00995)
Observations	46,811				

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Other controls include age, age squared, years of education, gender, race, province, wave.

Table E2: Interactions Regressions in the QLFS

Race/Gender	(1) All Separations	(2) Education Category	(3) All Separations	(4) Wave	(5) All Separations
African Male	-0.0468***	No Education	-0.0358*	QLFS 2011:1	-0.0465***
	(0.00476)		(0.0184)		(0.00763)
African Female	-0.0453***	Some Primary	-0.0341***	QLFS 2011:2	-0.0490***
	(0.00459)		(0.00876)		(0.00736)
Coloured Male	-0.0275	Some Secondary	-0.0284***	QLFS 2018:1	-0.0525***
	(0.0281)		(0.00685)		(0.00700)
Coloured Female	-0.0558***	Matric	-0.0675***	QLFS 2018:2	-0.0477***
	(0.0183)		(0.00581)		(0.00661)
Indian/Asian Male	-0.0232	Some Tertiary/Degree	-0.0450***	QLFS 2018:3	-0.0293**
	(0.0209)		(0.0122)		(0.0118)
Indian/Asian Female	0.0250			QLFS 2018:4	-0.0536***
	(0.0672)				(0.00992)
White Male	-0.0348*			QLFS 2019:1	-0.0513***
	(0.0186)				(0.00884)
White Female	-0.0516			QLFS 2019:2	-0.0372***
	(0.0394)				(0.00737)
Observations	20,729				

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Other controls include age, age squared, years of education, gender, race, province, wave. Survey weights used.

13.6. Appendix F – Alternate Interval Regressions

Table F1: Alternate Interval Regressions No Controls

	LFS - 1 Year				QLFS - 1 Quarter			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	All Separations	Separations to Employment	Separations to Non-Employment	Hired from Employment	All Separations	Separations to Employment	Separations to Non-Employment	Hired from Employment
Log Real Wage	-0.0988***	-0.0406***	-0.0727***	0.242***	-0.0433***	-0.0110***	-0.0335***	0.0959
	(0.00341)	(0.00286)	(0.00289)	(0.0457)	(0.00166)	(0.000932)	(0.00139)	(0.0596)
Constant				-1.273***				-1.392***
				(0.118)				(0.166)
Observations	25,754	21,335	25,754	7,790	44,216	41,208	44,216	4,242

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Columns 1-3 and 5-7 are average partial effects. Columns 4 & 8 are logit coefficients. Columns 1-3 controls are lagged by 1 Year. Columns 5-7 controls are lagged by 3 months. Columns 4 & 8 controls are contemporaneous. Survey weights used.

Table F2: Alternate Interval Results Med Controls

	LFS - 1 Year				QLFS - 1 Quarter			
VARIABLES	(5) All Separations	(6) Separations to Employment	(7) Separations to Non-Employment	(8) Hired from Employment	(5) All Separations	(6) Separations to Employment	(7) Separations to Non-Employment	(8) Hired from Employment
Log Real Wage	-0.0913***	-0.0421***	-0.0639***	0.173***	-0.0297***	-0.00763***	-0.0230***	0.0920
	(0.00483)	(0.00442)	(0.00389)	(0.0544)	(0.00201)	(0.00123)	(0.00164)	(0.0681)
Age	-0.0364***	-0.00827***	-0.0296***	0.226***	-0.00850***	0.000779	-0.00897***	0.184***
	(0.00259)	(0.00216)	(0.00211)	(0.0280)	(0.000996)	(0.000600)	(0.000851)	(0.0357)
Age Squared	0.0004***	5.17e-05*	0.0004***	-0.003***	8.2e-05***	-1.9e-05**	9.6e-05***	-0.0024***
	(3.33e-05)	(2.80e-05)	(2.70e-05)	(0.000368)	(1.26e-05)	(7.73e-06)	(1.08e-05)	(0.000459)
Years of Education	0.000197	-0.000257	0.000382	-0.0331**	-0.0049***	-0.00185***	-0.00323***	-0.0508***
	(0.00129)	(0.000981)	(0.00110)	(0.0131)	(0.000511)	(0.000309)	(0.000431)	(0.0186)
Coloured	-0.00873	-8.39e-06	-0.0135	0.216	0.0369***	0.0129***	0.0262***	0.340*
	(0.0127)	(0.00998)	(0.0109)	(0.142)	(0.00821)	(0.00495)	(0.00702)	(0.197)
Indian/Asian	-0.0232	-0.00994	-0.0180	0.264	-0.0100	0.00419	-0.0151	0.652*
	(0.0228)	(0.0177)	(0.0191)	(0.281)	(0.0151)	(0.00735)	(0.0117)	(0.377)
White	-0.00918	0.00864	-0.0230	0.526**	-0.0295***	-0.0110***	-0.0206***	-0.443
	(0.0196)	(0.0154)	(0.0171)	(0.209)	(0.00841)	(0.00405)	(0.00734)	(0.380)
Female	0.0110	-0.0104	0.0233***	-0.657***	-0.00216	-0.00779***	0.00556**	-0.605***
	(0.00815)	(0.00674)	(0.00680)	(0.0849)	(0.00295)	(0.00169)	(0.00257)	(0.110)
Widow/er	0.00513	0.0222	-0.00880	0.250	0.00655	-0.00137	0.00712	0.0903
	(0.0166)	(0.0184)	(0.0126)	(0.199)	(0.00726)	(0.00438)	(0.00612)	(0.310)
Divorced	0.00476	0.00602	0.000253	0.155	0.0161	0.0198**	-0.00113	0.675**
	(0.0234)	(0.0148)	(0.0210)	(0.184)	(0.0114)	(0.00987)	(0.00775)	(0.298)
Never Married	0.0411***	0.0193**	0.0300***	-0.326***	0.0164***	0.00224	0.0151***	-0.234*
	(0.00988)	(0.00758)	(0.00833)	(0.102)	(0.00324)	(0.00183)	(0.00278)	(0.124)
Constant				-4.258***				-4.116***
				(0.573)				(0.723)
Observations	25,754	21,335	25,754	7,767	44,216	41,208	44,216	4,242

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Wave and Province Results Excluded. Columns 1-3 & 5-7 are average partial effects. Column 4 & 8 are logit coefficients. Columns 1-3 controls are lagged by 1 year. Columns 5-7 controls are lagged by 3 months.. Column 4 & 8 controls are contemporaneous. Survey weights used.

Table F3: Alternate Interval Results Med+ Controls

	LFS - 1 Year				QLFS - 1 Quarter			
	(5)	(6)	(7)	(8)	(5)	(6)	(7)	(8)
VARIABLES	All Separations	Separations to Employment	Separations to Non-Employment	Hired from Employment	All Separations	Separations to Employment	Separations to Non-Employment	Hired from Employment
Log Real Wage	-0.0781***	-0.0358***	-0.0558***	0.148***	-0.0244***	-0.00478***	-0.0203***	0.109
	(0.00555)	(0.00502)	(0.00459)	(0.0568)	(0.00231)	(0.00147)	(0.00185)	(0.0729)
Age	-0.0357***	-0.00780***	-0.0294***	0.225***	-0.0089***	0.000411	-0.00910***	0.178***
	(0.00259)	(0.00214)	(0.00212)	(0.0277)	(0.000989)	(0.000594)	(0.000846)	(0.0358)
Age Squared	0.000405***	5.01e-05*	0.000358***	-0.00298***	9.0e-05***	-1.34e-05*	9.95e-05***	-0.00223***
	(3.31e-05)	(2.78e-05)	(2.69e-05)	(0.000363)	(1.25e-05)	(7.62e-06)	(1.07e-05)	(0.000459)
Years of Education	0.00154	0.000238	0.00136	-0.0278**	-0.0027***	-0.0008**	-0.002***	-0.0354*
	(0.00137)	(0.00105)	(0.00118)	(0.0139)	(0.000547)	(0.000332)	(0.000464)	(0.0197)
Coloured	0.00338	0.00457	-0.00451	0.219	0.0364***	0.0131***	0.0256***	0.325
	(0.0130)	(0.0102)	(0.0109)	(0.145)	(0.00821)	(0.00496)	(0.00700)	(0.200)
Indian/Asian	-0.0294	-0.0136	-0.0220	0.339	9.00e-05	0.0100	-0.00984	0.610
	(0.0222)	(0.0169)	(0.0188)	(0.286)	(0.0170)	(0.00897)	(0.0131)	(0.399)
White	-0.0159	-0.000409	-0.0216	0.470**	-0.0328***	-0.0121***	-0.0225***	-0.437
	(0.0193)	(0.0144)	(0.0173)	(0.218)	(0.00812)	(0.00380)	(0.00720)	(0.420)
Female	0.0430***	0.00363	0.0480***	-0.617***	0.0160***	-0.00220	0.0188***	-0.558***
	(0.00934)	(0.00766)	(0.00798)	(0.0994)	(0.00369)	(0.00220)	(0.00324)	(0.135)
Widow/er	0.000612	0.0164	-0.0111	0.277	0.00475	-0.00207	0.00587	0.0670
	(0.0161)	(0.0172)	(0.0125)	(0.197)	(0.00704)	(0.00420)	(0.00600)	(0.316)
Divorced	-0.00161	0.00278	-0.00302	0.188	0.0143	0.0184*	-0.00194	0.615**
	(0.0227)	(0.0146)	(0.0206)	(0.188)	(0.0111)	(0.00942)	(0.00760)	(0.306)
Never Married	0.0382***	0.0183**	0.0279***	-0.300***	0.0154***	0.00204	0.0145***	-0.195
	(0.00972)	(0.00752)	(0.00823)	(0.0990)	(0.00318)	(0.00178)	(0.00275)	(0.124)
Professional	-0.0204	0.0294	-0.0600	-0.158	-0.00167	-0.00714	0.00339	-0.308
	(0.0492)	(0.0394)	(0.0439)	(0.415)	(0.0214)	(0.00803)	(0.0201)	(0.680)
Technical/Assoc. Prof.	-0.0451	-0.0152	-0.0390	-0.121	-0.0120	0.00499	-0.0163	0.196
	(0.0360)	(0.0293)	(0.0333)	(0.361)	(0.0177)	(0.00931)	(0.0160)	(0.487)
Clerks	-0.0541	-0.0329	-0.0307	-0.986***	-0.00829	0.00150	-0.00945	0.255
	(0.0365)	(0.0285)	(0.0346)	(0.335)	(0.0167)	(0.00726)	(0.0158)	(0.488)
Service Workers	-0.0236	-0.0201	-0.00789	-0.595*	-0.00530	0.00102	-0.00620	0.00782
	(0.0360)	(0.0282)	(0.0343)	(0.340)	(0.0165)	(0.00716)	(0.0155)	(0.442)
Skilled Agri	0.00537	-0.0218	0.0384	-0.633	0.00687	-0.00346	0.0120	-0.217
	(0.0446)	(0.0350)	(0.0418)	(0.433)	(0.0291)	(0.0112)	(0.0282)	(0.907)
Craft/Trade	0.0275	0.00835	0.0246	-0.595*	0.0233	0.0116	0.0135	0.138

	(0.0372)	(0.0304)	(0.0351)	(0.346)	(0.0172)	(0.00779)	(0.0161)	(0.458)
Machine Worker	-0.0270	-0.0392	0.00635	-0.557	-0.0176	0.000313	-0.0180	-0.291
	(0.0371)	(0.0283)	(0.0355)	(0.342)	(0.0167)	(0.00718)	(0.0158)	(0.470)
Elementary Occ.	-0.0175	-0.0197	0.000364	-0.689**	0.0150	0.00857	0.00832	-0.0325
	(0.0358)	(0.0286)	(0.0336)	(0.333)	(0.0163)	(0.00711)	(0.0154)	(0.439)
Domestic	-0.0913**	-0.0627*	-0.0311	-0.509	-0.0324*	-0.00430	-0.0266*	0.0625
	(0.0425)	(0.0320)	(0.0395)	(0.401)	(0.0170)	(0.00723)	(0.0161)	(0.502)
Mining	-0.0181	-0.0212	-0.00126	-0.249	-0.0268***	-0.0169***	-0.00778	-1.148**
	(0.0200)	(0.0156)	(0.0157)	(0.277)	(0.00841)	(0.00320)	(0.00814)	(0.488)
Manufacturing	0.0565***	0.00362	0.0597***	-0.647***	0.000867	-0.00378	0.00490	-0.479*
	(0.0142)	(0.0114)	(0.0116)	(0.168)	(0.00644)	(0.00366)	(0.00536)	(0.278)
Utilities	0.164***	0.0319	0.169***	-0.453	0.0139	-0.0121	0.0276	-0.379
	(0.0627)	(0.0439)	(0.0644)	(0.561)	(0.0241)	(0.00887)	(0.0236)	(1.082)
Construction	0.205***	0.0848***	0.160***	-0.583***	0.0759***	0.0263***	0.0534***	-0.199
	(0.0218)	(0.0202)	(0.0182)	(0.185)	(0.00856)	(0.00523)	(0.00719)	(0.238)
Trade	0.0600***	0.0104	0.0595***	-0.691***	-0.00986*	-0.00745**	-0.00251	-0.593***
	(0.0139)	(0.0117)	(0.0111)	(0.173)	(0.00547)	(0.00308)	(0.00462)	(0.227)
Transport	0.147***	0.0584**	0.116***	-0.562**	0.0216**	0.00279	0.0198**	0.160
	(0.0255)	(0.0227)	(0.0238)	(0.232)	(0.00952)	(0.00496)	(0.00851)	(0.359)
Finance	0.0741***	0.0273*	0.0543***	-0.435**	-0.00546	-0.00142	-0.00408	-0.183
	(0.0180)	(0.0144)	(0.0142)	(0.219)	(0.00604)	(0.00369)	(0.00490)	(0.240)
Services	0.0287*	-0.0172	0.0475***	-0.579***	0.00404	-0.00524	0.00861	-0.545**
	(0.0173)	(0.0125)	(0.0146)	(0.212)	(0.00623)	(0.00366)	(0.00524)	(0.254)
Domestic Services	0.127***	0.0699**	0.0672***	-0.595***	0.0512***	0.0227***	0.0308***	-0.244
	(0.0292)	(0.0275)	(0.0212)	(0.231)	(0.00959)	(0.00615)	(0.00784)	(0.250)
Public Sector	-0.0493***	-0.0223*	-0.0344***	-0.274*	-0.0316***	-0.0121***	-0.0212***	-0.149
	(0.0148)	(0.0118)	(0.0123)	(0.162)	(0.00425)	(0.00238)	(0.00360)	(0.214)
Constant				-3.092***				-3.962***
				(0.667)				(0.901)
Observations	25,700	21,285	25,700	7,743	44,180	41,177	44,180	4,237

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Wave and Province Results Excluded. Columns 1-3 & 5-7 are average partial effects. Column 4 & 8 are logit coefficients. Columns 1-3 controls are lagged by 1 year. Columns 5-7 controls are lagged by 3 months.. Column 4 & 8 controls are contemporaneous. Survey weights used.

13.7. Appendix G – Separations by Firm Characteristics Results

Table F1: All Separations using Firm Characteristics to Identify Separations

VARIABLES	(1) All Separations	(2) Separation to Employment	(3) Separation to Non-Employment	(4) Hired from Employment	(1) All Separations	(2) Separation to Employment	(3) Separation to Non-Employment	(4) Hired from Employment
Log Real Wage	-0.0601***	-0.0230***	-0.0537***	0.591***	-0.0365***	-0.0129***	-0.0304***	0.368***
	(0.00420)	(0.00453)	(0.00255)	(0.0321)	(0.00466)	(0.00448)	(0.00264)	(0.0828)
Age	-0.0254***	-0.00999***	-0.0215***	0.204***	-0.00917***	0.00193	-0.0111***	0.176***
	(0.00248)	(0.00271)	(0.00164)	(0.0160)	(0.00246)	(0.00244)	(0.00138)	(0.0308)
Age Squared	0.0003***	9.32e-05***	0.0003***	-0.0024***	8.64e-05***	-3.82e-05	0.0001***	-0.0018***
	(3.06e-05)	(3.33e-05)	(2.17e-05)	(0.000201)	(3.11e-05)	(3.06e-05)	(1.78e-05)	(0.000388)
Years of Education	-0.0033***	-0.0038***	-4.50e-05	-0.0198**	-0.0072***	-0.0055***	-0.0026***	0.0241
	(0.00117)	(0.00122)	(0.000685)	(0.00845)	(0.00128)	(0.00121)	(0.000748)	(0.0200)
Coloured	-0.0134	-0.00334	-0.0165**	0.175**	0.0522***	0.0290**	0.0314***	0.179
	(0.0123)	(0.0128)	(0.00741)	(0.0836)	(0.0167)	(0.0148)	(0.0116)	(0.183)
Indian/Asian	0.0656***	0.0594***	0.0132	0.381***	0.0634*	0.0660*	-0.0136	1.468***
	(0.0221)	(0.0228)	(0.0132)	(0.133)	(0.0382)	(0.0359)	(0.0215)	(0.391)
White	0.0818***	0.0905***	-0.0202*	0.452***	0.0218	0.0412**	-0.0491***	0.935***
	(0.0165)	(0.0170)	(0.0104)	(0.133)	(0.0222)	(0.0205)	(0.00580)	(0.264)
Female	-0.0358***	-0.0675***	0.0343***	-0.510***	-0.0120*	-0.0239***	0.0127***	-0.339***
	(0.00706)	(0.00744)	(0.00439)	(0.0525)	(0.00711)	(0.00666)	(0.00424)	(0.0927)
Widow/er	0.00558	0.00333	0.00333	0.130	0.0156	0.00814	0.00889	-0.0571
	(0.0167)	(0.0186)	(0.00881)	(0.112)	(0.0181)	(0.0163)	(0.0112)	(0.253)
Divorced	0.0192	0.0119	0.0163	-0.0572	-0.0195	-0.0283*	0.0119	0.0608
	(0.0190)	(0.0198)	(0.0148)	(0.123)	(0.0191)	(0.0172)	(0.0118)	(0.295)
Never Married	0.0218**	-0.000698	0.0342***	-0.426***	0.0353***	0.0233***	0.0182***	-0.286***
	(0.00986)	(0.0105)	(0.00513)	(0.0609)	(0.00836)	(0.00816)	(0.00432)	(0.107)
Observations	48,567	41,448	48,567	18,585	20,810	19,037	20,810	4,412

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Wave and Province Results Excluded. Columns 1,2 & 3 are average partial effects. Column 4 are logit coefficients. Columns 1,2 & 3 controls are lagged by 6 months. Column 4 controls are contemporaneous. Survey weights used.

13.8. Appendix H – StatsSA Imputed Data Results

Table G1: Separations Regressions with StatsSA Imputed Data (QLFS)

VARIABLES	(1) All Separations	(2) Separations to Employment	(3) Separations to Non-Employment	(4) Hired from Employment
Log Real Wage	-0.0186***	-0.00592***	-0.0137***	0.117*
	(0.00207)	(0.00149)	(0.00152)	(0.0687)
Age	-0.0126***	-0.00282***	-0.0102***	0.146***
	(0.00144)	(0.00104)	(0.00110)	(0.0385)
Age Squared	0.000124***	1.89e-05	0.000108***	-0.00159***
	(1.84e-05)	(1.34e-05)	(1.41e-05)	(0.000498)
Years of Education	-0.00939***	-0.00450***	-0.00538***	-0.00499
	(0.000663)	(0.000470)	(0.000500)	(0.0222)
Coloured	0.0239**	0.0102	0.0152*	0.525**
	(0.0101)	(0.00703)	(0.00827)	(0.223)
Indian/Asian	-0.0236	-0.00160	-0.0256**	1.732***
	(0.0177)	(0.0129)	(0.0123)	(0.328)
White	-0.0570***	-0.0210***	-0.0397***	0.531**
	(0.00797)	(0.00542)	(0.00557)	(0.224)
Female	0.00976**	-0.000665	0.0113***	-0.273**
	(0.00422)	(0.00292)	(0.00336)	(0.117)
Widow/er	0.000415	-0.00683	0.00579	-0.0767
	(0.0106)	(0.00691)	(0.00842)	(0.333)
Divorced	-0.000114	-0.00586	0.00502	0.0780
	(0.0110)	(0.00647)	(0.00930)	(0.370)
Never Married	0.0282***	0.00930***	0.0208***	-0.172
	(0.00473)	(0.00347)	(0.00358)	(0.130)
Constant				-3.661***
				(0.801)
Observations	26,776	24,742	26,776	2,908

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Wave and Province Results Excluded. Columns 1,2 & 3 are average partial effects. Column 4 are logit coefficients. Columns 1,2 & 3 controls are lagged by 6 months. Column 4 controls are contemporaneous. Survey weights used.