Worker Flows in South Africa From 2008 to 2019:

Evidence From Matched Quarterly Labour Force Survey

by

Sammy Pan

Thesis Supervisor: Associate Professor Andrew Kerr

A dissertation submitted in partial fulfilment

of the requirements for the degree of

Masters of Commerce

specialising in

Applied Economics

University of Cape Town

July 2022

Acknowledgements

To consider the completion of a Master's thesis as a personal success may be overblown, but if it is some sort of success, then I cannot have achieved it on my own without the support of my supervisor, family members, and friends. This is not my success, but also your success.

First and foremost, I would like to extend my sincerest gratitude to my supervisor, Associate Professor Andrew Kerr. One cannot have asked for a better mentor and teacher. Compassionate but rigorous, kind but impartial, patient but prudent, you have motivated me to become a better academic and researcher. Without your guidance, your detailed feedback on my sometimes terrible drafts, and your sound advice, this thesis could not have been completed.

I would like to thank my parents. While they still do not really know what it is that I am studying, they have supported me in every decision I have made for myself and in every career path that I have taken. I am so fortunate to be around them – wherever they are is my home. Being the first person in my family to have completed a Master's degree, I hope I can make them proud.

Lastly, I would like to thank my best friend Bianca. You have accompanied me throughout this entire process, from when I chose the thesis topic until its completion. You cheer me up when I am down, you inspire me when I feel deflated, and you make me feel proud of myself even when I achieve little. Thank you.

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.

Signed by candidate

Sammy Pan

Table of Content

List of Tables

List of Figures

Abstract

This study examines the extent, cyclicality, and heterogeneity of worker flows for the working-age population of South Africa from 2008 to 2019. Worker flows are measured using quarterly panel data at the individual level, which is constructed by matching individual records between consecutive quarters of the cross-sectional data of the Quarterly Labour Force Survey (QLFS) from 2008q1 to 2019q2. This study expends a significant amount of effort in proposing and implementing a method for evaluating the quality of different matching procedures for the QLFS, which has never been accomplished systematically in previous literature.

Using the matched QLFS panel, we find that the average quarterly worker flow rate was 46.26% over the period 2008 – 2019, which is large when compared with other developing countries. This is surprising in a labour market considered by many as rigid. Gross worker flows are procyclical. But its procyclicality is driven by the procyclical worker flows in the informal sector. In contrast, worker flows in the formal sector are acyclical. There is substantial heterogeneity in worker flows across different firms, jobs, industries, and sub-population groups. A series of descriptive analyses, complemented by a regression analysis of separation probabilities, reveal that worker flows are much larger for some groups (e.g., those who are younger or have lower levels of education), firms (e.g., those that are smaller in size or private), jobs (e.g., those with fixed-term contract relative to permanent contract), or industries (e.g., construction, agriculture, and domestic services) when compared to the rest. Therefore, although substantial gross worker flows in South Africa imply that the labour market as a whole is not as rigid as most people think, certain sub-sections of the labour market are definitely more rigid than others.

Despite a myriad of data challenges and obstacles, we believe that this study has managed to capture the true extent and pattern of worker flows in South Africa. Measurement errors may exist in our estimates, but a robustness check shows a high degree of consistency between our worker flow estimates and those in the past literature, implying that measurement errors are moderate.

1. Introduction

The labour market of South Africa seems static. Its rate of unemployment is persistently high. Labour market institutions such as trade unions, minimum wages, and employment protection regulations are often said to stifle the dynamics of the labour market by stemming the growth and vitality of formal-sector wage employment (Bhorat, Kanbur & Mayet, 2012; Fields, 2014; Magruder, 2012; Piek, von Fintel & Kirsten, 2020). Even South Africa's small informal sector appears to lack the vigour that is often observed in a typical informal economy (Heintz & Posel, 2008). Unable to find jobs, many of the unemployed stop searching for jobs and become discouraged (Verick, 2012). In contrast to their plight, demands for better remuneration from trade unions often occupy news headlines.

In light of the apparent rigidity of the labour market portrayed above, how substantial are worker flows? Worker flows refer to separations and hires of individuals, whether they be transitions between employment and non-employment, or changes in employers. Surprisingly, past literature on labour market dynamics and worker flows suggests that, overall, South African labours are considerably mobile (Cichello, Leibbrandt & Woolard, 2014; Essers, 2017; Kerr, 2018). Behind the seemingly stagnant headline indicators such as the unemployment rate and despite various labour market institutions such as trade unions and employment protection regulations, millions of South Africans move into and out of employment or switch jobs on a quarterly, half-yearly, or annual basis.

A better understanding of worker flows in South Africa is important. First, it enhances our understanding of South Africa's unemployment. The unemployed are often thought of as a stagnant pool of unfortunate individuals without jobs, but the presence of worker flows suggests that unemployment is dynamic as opposed to static. At any moment, some individuals enter the unemployment pool while others exit it. It is important to shed light on this process, which helps shift the direction of the discourse on the nature of South Africa's unemployment so that it is grounded in labour market theories and models that are more realistic. Second, worker flows reveal the efficiency of the South African labour market in terms of matching the job-hunters with job vacancies. Research on unemployment in the context of South Africa has a strong emphasis on the structural mismatch between the human capital of the unemployed and the skills required by employers (e.g., Banerjee et al., 2008), but there is limited knowledge on the matching efficiency of the labour market. Worker flows illuminate how much labour market frictions contribute to South Africa's extremely high unemployment levels, which carries important implications for policymakers hoping to reduce unemployment. Lastly, it is important to study, in its own right, how frequently people move into, out of, and between jobs. These transitions affect the welfare of the individuals involved.

To contribute to our understanding of worker flows in South Africa, this paper examines the extent, cyclicality, and the heterogeneous pattern of worker flows using quarterly individual-level panel data which spans over a 10-year period from 2008 to 2019. This panel data was constructed by matching individuals in consecutive cross-sectional data from the Quarterly Labour Force Survey (QLFS).

This study makes several contributions to the existing literature on worker flows and labour market dynamics in South Africa. First, a majority of the existing South African literature on labour market dynamics lacks appropriate guiding theoretical frameworks. As such, different studies devise different methods for measuring labour market dynamics, and these measures are often incomparable with each other, as well as incompatible with empirical evidence from other countries. As a consequence, our knowledge on the labour market dynamics of South Africa is fragmented and there is a disconnect between the South African literature and a body of related literature outside of South Africa. To address this, we utilise the matching model as the theoretical framework of the study, which guides our efforts in defining and measuring worker flows. In doing so, we not only connect the South African literature on labour market dynamics to a large and vibrant body of theoretical and empirical literature on worker flows in the West, but we also build a solid theoretical and methodological foundation for future efforts in research on labour market dynamics.

As one of the only previous studies on worker flows in South Africa, Kerr (2018) examined annual worker flows for formal-sector employees over a three-year period from 2011 to 2014. However, due to limited data coverage, the study cannot reveal the overall extent of worker flows in South Africa. Covering a relatively short period of time, it also cannot show the long-term trend and cyclical variation of worker flows. Furthermore, due to a lack of relevant information in the data, it cannot conduct a desirable analysis of the heterogeneity of worker flows by worker, firm, and job characteristics. All of these limitations are addressed in this paper. We utilise panel data based on more than 40 QLFS cross-sectional datasets from 2008q1 to 2019q2. Because the QLFS is representative of the South African population, we are able to derive quarterly worker flows for all employed individuals, whether they be workers in the formal sector or informal sector. Given the long time span covered by the data, we can also examine the temporal variation and cyclicality of worker flows. Furthermore, there is a wealth of information collected on individuals in the QLFS data, which enables us to conduct a thorough heterogeneity analysis of worker flows.

This paper also makes important contributions to the technical task of creating panel data from surveys with a rotating panel design, such as the QLFS. A panel dataset of individuals who are matched between consecutive quarters of the QLFS is extremely valuable for those wishing to conduct microeconomic research on labour market issues because it allows one to control for time-invariant individual characteristics in analyses. But users of such panel data inevitably face two challenges. First, several different procedures for data matching have been conceived in past literature but it is unclear which "matching procedure" generates panel data of the best quality. Second, there has been no systematic investigation into the relationship between "matching procedures" and "matching quality". This paper addresses these two challenges by presenting a consolidated and comprehensive summary of data matching procedures used in previous literature, and by proposing and implementing a method for evaluating the quality of panel data created from these matching procedures.

The rest of the paper is organised as follows. Section 2 details the theoretical perspective of the matching model and the relevant definitions and concepts of worker flows. Section 3 presents a review of the existing South African and international literature on worker flows. Section 4 describes the datasets used in this study, the QLFS and the Post-Apartheid Labour Market Series (PALMS). To construct a set of quarterly panel samples of individuals for worker flows analyses, we match individuals across consecutive quarters of the QLFS data from 2008q1 to 2019q2. To ensure that the optimal matching procedure is adopted for this endeavour, in Section 5, we conduct a thorough investigation into different QLFS matching algorithms and the effect of matching algorithms on the matching quality of the resulting panel. Based on this investigation, an optimal matching procedure is chosen for our panel sample creation – the sample creation process and the process by which individual weights are adjusted for panel attrition are described in Section 6. With our panel sample, worker flows are measured, and how this is done is outlined in Section 7. Finally, findings on worker flows are presented: Section 8 examines the extent of worker flows, Section 9 investigates the cyclicality of worker flows, and Section 10 studies the heterogeneity of worker flows by individual, firm, job, and industry characteristics using both descriptive and regression methods. Section 11 concludes.

2. Theoretical Background

2.1. The matching model of the labour market

The inadequacy of the competitive model of the labour market is exposed when one tries to explain labour mobility and unemployment. At the equilibrium of a competitive labour market, the prevailing wage rate is equal to the marginal product of labour. And because a worker's marginal product is a function of one's productivity, which does not increase when he/she switches to a different employer, there is simply no incentive for a worker to switch jobs. Even if there is an exogenous shock to the labour demand or supply, the only transition experienced by a representative worker is between economic inactivity and employment. Because a competitive labour market adjusts instantaneously to economic shocks under the assumption of perfect information, no one will become unemployed in the process of labour market adjustment (Borjas, 2013:146). Needless to say, what has been described is rarely what one observes in reality, where the unemployment level may be high and persistent, and the unemployed may stay jobless for an extended period of time. Market-based economies are also characterised by substantial labour mobility across firms and labour market status – these transitions are motivated by reasons other than economic shocks, such as job dissatisfaction, job shopping, dismissal, or retraining.

In light of the inconsistency between the theory and the real world, it is crucial to adjust the theory to better reflect reality. The theoretical framework that is capable of explaining the co-existence of unemployment and labour mobility is the matching model, best exemplified by Mortensen and Pissarides (1994). In the simplified version, it assumes that the labour market possesses intrinsic frictions – caused by factors such as the geographical separation between those looking for jobs and those wanting to hire, and the imperfect information regarding the characteristics of the employer, the vacancy, and the job-hunter. These frictions impose costs on recruitment and job-searching processes. When firms with vacancies and job seekers finally "meet", a job match is not automatically formed (which is when the worker is hired) but, instead, they will engage in a process of bargaining. During the bargaining process, the imbalance and the distribution of bargaining power will affect the distribution of the surplus value of the job match as well as the wage for the worker. But, due to imperfect information in the labour market, workers do not necessarily know which job is best for him/her and employers might also be ignorant of the true productivity of the worker until it is revealed on the job over time. As a result, a recent hire might prove to be a mismatch, in which case the worker quits or is fired, even in the absence of economic shocks. A separated worker might spend some time before she finds the next suitable job vacancy, and more time could elapse until the next successful job match is formed. Those who are in-between employment become unemployed, in which case unemployment is frictional because it is the result of matching inefficiency and imperfect information. Unemployment is also dynamic since individuals are constantly moving in and out of the unemployment pool, the size of which changes depending on the rate at which workers are separated and hired.

Hires and separations are collectively called worker flows. The number of hires in an economy over a fixed period is described by a matching function. This function takes as inputs the number of vacant jobs, the number of job seekers, and the matching efficiency in the market (Cahuc, Carcillo & Zylberberg, 2014:583). Separations could be initiated by a firm or a worker. Firms might fire unproductive employees, or retrench workers in response to reduced labour demand. But a worker could also voluntarily quit a low-paying job to take up another job that promises a higher wage or better work environment (Donovan, Lu & Schoellman, 2020:18-24).

In the presence of continuous worker flows, the notion of a static labour market equilibrium is no longer relevant. Instead, the most natural concept for the labour market equilibrium is a dynamic one, which is reached when worker flows balance out so that the unemployment rate reaches a steady state. Furthermore, the unemployment rate is intricately related to and influenced by the rate at which vacant jobs are created and destroyed, the rate of hiring and separation, and the matching efficiency in the market.

2.2. The nature of unemployment in South Africa

To contrast the perspective of the competitive model against that of the matching model, we review and analyse two strands of scholarly discourse on the nature of unemployment in the context of South Africa.

2.2.1. The perspective of a competitive model

South Africa has a disconcerting unemployment problem. The official unemployment rate from the last quarter of 2019 stood at a staggering 29,1% (Statistics South Africa [StatsSA], 2019). The official unemployment rate has never dipped below 20% in the 20 years from 1994 to 2014, (Mlatsheni & Leibbrandt, 2014:236). If we include the discouraged workers, unemployment rates have remained above 30% for most of the period between 1995 and 2015 (Festus et al., 2016:584).

These figures raise the question of how to best understand the nature of South Africa's unemployment. One might use a competitive labour market framework, under which large open unemployment indicates a disequilibrium in the labour market. Theoretically, real wages should adjust downwards until the market clears. In reality, however, real wages in the formal sector of South Africa declined little over time while the majority of wage adjustments took place in the informal sector and amongst the self-employed (Kingdon & Knight, 2009:301-302; Kerr, 2021), implying that the formal sector labour market does not clear in a competitive sense. This has led to the belief that there are labour market "rigidities", which restrict formal sector firms from responding efficiently to changes in the underlying labour demand or supply (Kingdon & Knight, 2009; Go, et al., 2010; Fedderke, 2012). As is evident through this line of logic, such beliefs are rooted in the competitive theoretical framework of the labour market.

While the exact definition of "rigidity" differs across different studies, two types of rigidity can be distinguished. The first is the rigidity of firing and hiring, which refers to South African employers' inability to change the optimal size and composition of their workplaces (Go et al, 2010:1483; Bertrand & Crepon, 2020). The firing/hiring rigidity is often said to be a consequence of South Africa's stringent employment protection regulation (Bhorat, Naidoo & Yu, 2014:6) or, at least, a perception thereof (Benjamin, 2014; Bertrand & Crepon, 2020).

The second type of rigidity is wage rigidity, which is manifested in the fact that formal-sector wages appear unresponsive to changes in the underlying economic environment (Fedderke, 2012; von Fintel, 2016). Empirical evidence suggests that wage rigidity in South Africa is maintained by its various labour market institutions, such as bargaining councils (Magruder, 2012), trade unions (Kerr & Wittenberg, 2021), and minimum wage laws (Bhorat, Kanbur & Mayet, 2012; Piek, von Fintel & Kirsten, 2020).

2.2.2. Shifting to the perspective of a matching model

As Fields (2000) articulated, one of the barriers to solving the problem of unemployment in South Africa is the lack of appropriate labour market models that could adequately explain the unique features of its labour market and link together a complex web of agents and institutions. He further argued that the competitive model is not the right one, highlighting the fact that it is incompatible with the large extent of wage rigidities observed in the formal sector of South Africa.

Subsequently, Wittenberg (2002) asserted that a fruitful approach to labour market research involves incorporating knowledge from the literature of matching models. He demonstrated this by showing that certain features of the South African labour market, such as the high unemployment rate of African youths, can be better accounted for by acknowledging that the labour market contains friction and imperfect information.

Indeed, there appear to be substantial frictions that obstruct the efficient matching between job seekers and firms, and this renders the application of a matching model feasible for the context of South Africa. To begin with, both wage rigidities and the rigidities of firing and hiring, as defined before, make up a substantial part of the labour market frictions.

But other forms of friction also exist. First, there is the spatial mismatch between business centres where job opportunities are abundant and residential areas – a problem that becomes much more pronounced for the lower-income residential areas, such as townships, where the unemployment rate is higher. This has created massive informational barriers between labour-seeing firms and job-seeking workers, thus delaying the process of recruitment and obstructing job searching (Banerjee et al., 2008:734). While spatial segregation affects all countries to varying degrees, it is a much more deep-rooted challenge for South Africa. The urban landscapes today remain a legacy of the segregationist policies implemented before and under apartheid, which were intended to separate the black, coloured and Indian population away from jobs in urban areas (von Fintel, 2016; van der Merwe & Krygsman, 2020).

The problem of spatial mismatch would be ameliorated if there is a reliable and affordable network of public transportation. But this is not the case for most South Africans, and especially not so for those who are economically deprived (Kerr, 2017). Poor interconnectivity of public transport and the geographical dislocation of the low-income residential area imply that low-income job seekers could only look for jobs within a small geographical range (van der Merwe & Krygsman, 2020).

Secondly, labour market frictions also result from information asymmetry between firms and jobseekers with regard to workers' productivity. Ample evidence shows that educational attainment is only partially capable of signalling productivity in the South African labour market. While matric and higher education credentials signal higher productivity, lower levels of education do not increase one's employability (Wittenberg, 2002; Duff and Fryer, 2005). To overcome this informational barrier, firms have increasingly relied on job referrals (Schoer, Rankin & Roberts, 2014) and past work experience (Ingle and Mlatsheni, 2017) as screening devices, but this has severely disadvantaged unemployed youths since they possess little social connection or past experiences (Mlatsheni & Leibbrandt, 2014:238; Wittenberg, 2002). Furthermore, having no or little prior work experience, some workseekers might be ignorant of their true productivity (Carranza et al., 2020), which prevents them from choosing the most effective search channels according to their skills.

The above-mentioned labour market friction poses an immense challenge to the orthodox view of unemployment in South Arica, which sees it as a static and disequilibrium phenomenon in an otherwise competitive labour market. According to the matching theory, labour market frictions render searching and matching inefficient, and as a result, unemployment arises as a dynamic equilibrium phenomenon.

In light of labour market frictions, some South African scholars have incorporated the matching model theories into their toolkit for analysing and re-examining various issues concerning unemployment[1]. This body of literature has generated fresh insights on unemployment, illustrating that, by incorporating the knowledge from the literature on matching models, there lies a fruitful research avenue for labour market-related issues.

However, despite these efforts, our understanding of worker flows in South Africa remains limited. Besides Kerr (2018), there is a dearth of literature on the topic, despite worker flows being a crucial component of the matching model framework and are fundamental to our understanding of, amongst others, wage formation, dynamics of labour demand and supply, and matching efficiency.

We intend to fill this knowledge gap. With individual-level panel data constructed from a series of Quarterly Labour Force Survey (QLFS) from StatsSA, which spans over 10 years, we study South Africa's worker flows. By examining their magnitude, heterogeneous patterns, and cyclicality, we contribute to the ongoing efforts of incorporating matching theories into unemployment research in South Africa.

But a better knowledge of worker flows in South Africa is important in its own right. How frequently workers move into, out of and between jobs carry significant welfare implication for the workers involved. By studying worker flows, we also shed light on the extent to which employment protection

---

[1] For instance, Schoer, Rankin & Roberts (2014) studied how workers and firms self-select into different channels of recruitment or searching, and how this contributes to inefficient matching. They used this to explain why the rate of unemployment is high among African youths. Abel, Burger & Piraino (2020) studied how information asymmetry regarding job seekers' true productivity decreases job matching efficiency and match quality. They relate their findings to the observation that South African employers are increasingly reliant on social connections and referrals for new hires. Carranza et al., (2020) extended the previous study but showed that information asymmetry affects both firms and workers, i.e., both parties make sub-optimal decisions regarding recruiting or searching. Van der Merwe & Krygsman (2020) explore how geographical mismatch between low-income work-finders and job raises matching frictions. This was then used to account for the pattern of employment duration and worker turnover in South Africa.

regulation in South Africa constrains firms from adjusting their workforce. This problem is frequently found at the centre of public debates and media attention and is therefore important to address.

2.3. Concepts and definitions of worker flows

To connect the theoretical discussion above to the empirical discussions that will follow below, we now define key concepts in the worker flow literature. Conventionally, worker flows are measured from the perspective of the firm with firm-level panel data (e.g., Bellmann, Gerner & Upward, 2018) or matched employee-employer records (e.g., Burgess, Lane & Stevens, 2000; Huber & Smeral, 2006). But this paper is concerned with worker flows measured from the perspective of the workers using individual-level panel data. Some scholars have attempted the latter (e.g., Bell & Smith, 2002; Fallick & Fleischman, 2004; OECD, 2009; Schoiswohl & Wuger, 2016) and their studies provide a rich reservoir of knowledge that is fundamental to the ensuing discussion.

The main difference between the worker-level and firm-level worker flows is in the data used for measuring the two concepts. The firm-level panel data – based either on firm surveys as seen in Bellmann, Gerner and Upward (2018) or on employer-employee matched data as seen in Burgess, Lane and Stevens(2000) – enables researchers to compare the employment information of the same firm across time. From this, it is possible to determine the number of hires and separations in each firm. These hires and separations take place from the perspective of the firm. On the other hand, worker-level panel data – based mostly on individual-level surveys – enables us to compare labour market information for the same person across time. From this, we determine who is hired or separated. These hires and separations occur from the perspective of the workers.

Besides data differences, worker-level worker flows and firm-level worker flows essentially measure the same thing – the extent to which hires and separations take place in the labour market. After all, every hire or separation in a firm corresponds to a hired or a separated person, and vice versa. However, one should exercise caution in making direct comparisons between the two measures, because firm-level and worker-level panel datasets are rarely similar in data coverage, target population, time periods, or panel frequency (Bassanini & Garnero, 2013). In the following section, we first introduce the concepts and definitions for firm-level worker flows, which are more conventionally measured in the literature. This is followed by an exposition of worker-level worker flow concepts and definitions.

2.3.1. Firm-level worker flows

Presented in equation (1) is an accounting identity that captures the employment dynamics of a representative firm i between time $t-1$ and t. E_{it} is the level of employment at the firm i at time t,

E_{it-1} denotes its employment level at time $t-1$, and the net change in employment, $E_{i,t} - E_{i,t-1}$, is simply hires (H_{it}) minus separations (S_{it}) between $t-1$ and t. When we sum up the hires and separations across all firms, we obtain gross hires (H_t) and gross separation (S_t), as per equation (2).

$$1)\ \Delta \mathrm{E}_{\mathrm{it}} = E_{it} - E_{it-1} = H_{it} - S_{it}$$

$$2)\ H_t = \sum_i H_{it} \,; S_t = \sum_i S_{it}$$

Gross worker flows (WF_t) from the perspective of the firms are the sum of gross hires and separations, as shown in equation (3). The gross worker flow rates (WFR_t) is calculated by dividing the WF_t by the average employment stock across all firms between $t-1$ and t (Davis & Haltiwanger, 1999:2718), as shown in equation (4). Using the average employment stock as a denominator for WFR yields several desirable properties (Davis, Faberman & Haltiwanger, 2006:5): firstly, the WRF will be bounded at -200 and +200 percent. Secondly, it handles regression to the mean: the intertemporal movement of the gross worker flow rates will be smoothed despite fluctuations in employment stocks from one period to the next.

$$3)\ WF_t = H_t + S_t$$

$$4)\ N_t = \frac{\sum_i E_{it-1+} \sum_i E_{it}}{2}; \ WFR_t = \frac{WF_t}{N_t}$$

2.3.2. Worker-level worker flows

Similar to firm-level worker flows, worker-level worker flows refer to the movement of workers into and out of firms. Therefore, in the definitions provided below, we only consider wage employees, and not self-employed, employers, or non-wage employees.

We first break down the working-age population into two groups – those who are in wage employment (E^w) as opposed to those who are not. If a person j was employed at a wage-paying job at $t-1$ but was no longer so in the next period t, she experienced a wage employment outflow ($E_{jt}^{outflow}$). Conversely, if she was not in a wage employment at $t-1$ but was in a wage employment in the next period t, then she experienced a wage employment inflow (E_{jt}^{inflow}). If she was continuously employed at a wage-paying job in both $t-1$ and t, but changed her employer, then it is an employer-to-employer flow ($EEflow_{jt}$). Following this, gross hires between $t-1$ and t (H_t) equal to the sum of wage employment inflows and EE flows across the working-age population, and gross separations between $t-1$ and t (S_t) equal to the sum of wage employment outflows and EE flows across the working-age population – as per equation (5) and (6).

$$5)\ H_t = \sum_j E_{jt}^{inflow} + \sum_j EEflow_{jt}$$

$$6)\ S_t = \sum_j E_{jt}^{outflow} + \sum_j EEflow_{jt}$$

Therefore, WF_t from the perspective of the workers between $t-1$ to t equals the sum of H_t and S_t, as was defined above in equation (3). Following the convention, to obtain gross worker flow rate (WFR_t), we divide WF_t by the average stocks of wage employment between $t-1$ and t, as per equation (7).

$$7)\ \mathrm{N_t} = \frac{Wage\ Employment\ stock_{t-1} + Wage\ Employment\ stock_t}{2};\ WFR_t = \frac{WF_t}{N_t}$$

2.3.3. Reallocation method and turnover method

Two methods are available when it comes to measuring worker flows (Huber & Smeral, 2006). The "*reallocation method*" specifies that we compare a worker's labour status across two discrete points in time, which thus helps us determine if a hire or separation has taken place. In contrast, the "*turnover method*" assumes that time is continuous, which means that we should count all hires and separations that have transpired throughout the time interval.

The key difference between the two concepts lies in the treatment of short-term spells. Consider the following example: suppose someone works for company A at $t-1$. At some point between $t-1$ and t, she was fired from A, leading to a brief period of unemployment. However, before time t, she was hired by company B. In this example, the reallocation method stipulates that we compare her employment status between the two points in time and ignore her brief period of unemployment. Accordingly, we would count her labour market transition as an EE flow (between employer A and B). In contrast, the turnover method stipulates that we acknowledge the short-lived spell of unemployment that she experienced between $t-1$ and t. Accordingly, we would count both a separation (from company A) and a hire (into company B).

Huber and Smeral (2006) finds that short-term spells are a substantial part of gross worker flows. As such, ignoring short-term spells could lead to the underestimation of gross worker flows. The extent of the underestimation, however, depends on the time period between $t-1$ and t: the longer this period is, the more pronounced the underestimation becomes. Therefore, the decision between the reallocation and turnover method often depends on the data at hand and the importance of short-term spells for the research question.

3. Literature Review

Having discussed the relevant theoretical background for this study and the motivation to investigate worker flows in the context of South Africa, this section examines the existing empirical literature on worker flows.

3.1. The extent of worker flows in South Africa

To our knowledge, Kerr (2018) remains the only paper that has directly examined the extent of worker flows in South Africa[2]. Using the personal income tax data from the South African Revenue Service (SARS), the author studied firm-level worker flows in the formal sector between 2011 and 2014. His finding suggests that the average yearly gross worker flow rate in the formal sector of South Africa is 53%, that is, more than one in two of the job matches either breaks up or forms each year. This study underscores the pervasiveness of labour movement in South Africa's labour market.

That there are substantial movements in the labour market is a recurrent finding for a body of literature that examines the dynamics of the South African labour market. A substantial proportion of the working-age population find/lose jobs, adopt/give up job-searching, or enter/exit the labour market, whether it is every quarter (Essers, 2017), every six months (Ranchhod & Dinkelman, 2008), or every two years (Posel, Casale & Vermaak, 2014). Flows between firms may also be large. Cichello, Leibbrandt and Woolard (2014) studied labour movements across different employment types (regular wage, casual or self-employment), employment sectors (primary, secondary, tertiary or household), and occupations (professional/managerial, semi-skilled and elementary). Their findings suggest that, over a two-year period, around 11,57% of the employed changed employment type, 22,41% changed industry, and 30,29% changed occupation[3]. We interpret this as evidence of large EE flows, despite the legitimate concerns that some of these flows do not represent flows (e.g., a person who changed occupations could have done so within the same firm). As valid as these concerns are, we believe that

---

[2] Although Donovan, Lu and Schoellman (2020) also directly examined worker flows in South Africa, they did so in a cross-country context and only presented the estimated size of worker flows of South Africa in scatterplots alongside other countries.

[3] These figures on based on Tables 6, 9, and 12 in Cichello, Leibbrandt and Woolard (2014:73-75)

such findings remain important for a preliminary understanding of the magnitude of EE flows in South Africa, especially when there lacks alternative evidence.

While these studies have utilised worker-level panel datasets to shed light on labour market dynamics in South Africa, none of them examined worker flows. This is despite the fact that it is possible to measure worker flows using worker-level panel datasets. As a result, empirical knowledge on the size of worker flows in South Africa remains scant. This study intends to fill this gap in the literature.

3.2. Worker flows and employment protection stringency

Looking beyond South Africa, there is a large body of literature on worker flows. The earliest studies were concentrated in the U.S. A more recent study by Davis and Haltiwanger (2014:3-4) puts the U.S. quarterly firm-level worker flow rate at 24.1% in 2010. Subsequently, there have been studies that compare worker flows across different economies. A recurring finding from these studies is that worker flows vary significantly across different economies, even within economic blocs that are relatively homogenous in terms of the level of development and close in geographical proximity. OECD (2009) used panel data based on labour force surveys to measure worker flow rates for some OECD member countries[4] between 2000 and 2005. They found that the average worker flow rate in these countries is around 33% yearly and that there is substantial cross-country variation: Turkey, the U.K., Spain and the U.S. had the highest rates, all exceeding 45%. While Greece, Austria and Italy had the lowest rates, all lower than 30%. (OECD, 2009:129).

Cross-country variation in worker flows is ascribed to, amongst others, the stringency of labour market regulation, especially in terms of employment protection (Bassanini & Garnero, 2013). OECD (2009:119) observed that worker flows are larger in countries where there are lax dismissal laws, such as in the U.S.. But worker flows are smaller in countries where the protection of employment is more extensive, such as Germany and France. Similarly, Bellmann, Gerner and Upward (2018) commented that regulatory inflexibility in the labour market in Europe is the reason why worker flows are smaller in most European countries than in the U.S.

Empirical studies have subsequently verified the causal link between the stringency of employment protection and the reduction in worker flow rates. Bassanini and Garnero (2013) studied the impact of employment protection law on hiring and separation rates in 24 countries, all OECD members, using

---

[4] Including the U.S., U.K., and most European countries, but excluding Korea and Japan.

an annual worker-level panel from 1995 to 2007. The stringency of the employment protection law is proxied on OECD Indicators of Employment Protection which helped them easily compare employment security and the level of firm adaptability across countries. Utilising a difference-in-difference approach, they showed that the restrictiveness of the employment protection of a country had a significantly negative impact on that country's aggregate rates of hiring and separations (Bassanini & Garnero, 2013:31).

Employment in developing countries is much less protected, so one might speculate that worker flows are substantially higher in these countries. Although knowledge on the size of worker flows in developing countries remains limited, the existing knowledge appears to confirm this speculation. Donovan, Lu and Schoellman (2020) examined measures of worker mobility in countries that range vastly in terms of economic development and geographical location. They found that worker flows in the poorer countries are much larger than in the developed countries. Worker flows in some of the poorest countries in their sample can be up to 2 to 3 times larger than those in the richest countries.

Further evidence can be found in a small body of studies that utilised matched worker-firm data in developing countries. Kaplan, Gonzalez and Robertson (2007) and Corseuil et al. (2014) measured worker flow rates at the firm level in Mexico and Brazil using, respectively, Mexican social security administration data and Brazilian administrative records of employment. The average annual worker flow rate in the private formal sector of Mexico was found to be 71.3% from 1985 to 2001. In Brazil, it is 165.1% for youths (15-24 years old) and 84.1% for adults (25+ years old) in the formal sector between 1996 and 2010. These figures are significant even in comparison with worker flow estimates from the U.S., where employment protection is generally considered to be lax (Bassanini & Garnero, 2013).

Where does South Africa fit in this debate? While South Africa is not as economically developed as most OECD countries, its labour market regulation is widely considered to be stringent. Benjamin, Bhorat and Cheadle (2010) studied the "Employing Workers" indices compiled by the World Bank[5], which are numerical indicators showing how strict labour regulations are with regard to firing, hiring, working hours, and the employment relationship. Indices are comparable across countries, and a higher

[5] It must be noted that, while "Employing Workers" indices were widely reported in the media and used in empirical research, Benjamin, Bhorat and Cheadle (2010) pointed out their inadequacies. Firstly, these indices capture only the stringency of main clauses of labour regulation while ignoring subclauses; secondly, they are silent on the implementing capacity of labour market institutions; and thirdly, they ignore how courts and relevant institutions interpret the regulation, which may be very different from the actual regulation.

index value indicates increased labour regulation stringency. The authors illustrated that South Africa's indices for hiring and firing difficulty are noticeably above the world average and the world median (Benjamin, Bhorat & Cheadle, 2010:76), suggesting that the legal provisions constrain South African firms' autonomy in hiring and firing.

There is also evidence that South African businesses overestimate the level of stringency implied by the local labour market regulations (Bertrand & Crepon, 2020). This might exacerbate the effects of labour regulation stringency on the labour market. Evidence from the randomised control trials suggests that, because South African employers believe labour laws are more pro-worker than they really are, they hire fewer workers and are more hesitant in firing existing workers (Bertrand & Crepon, 2020).

Whether it is because of the strict labour regulatory environment, a distorted perception of the labour regulatory environment, or both, one might expect worker flows to be low in South Africa, at least in comparison with countries that are known for having lax labour regulations such as the U.S. But existing evidence from Kerr (2018) defies our expectations since it was shown that, with a yearly gross worker flow rate averaging 53%, the labour market in South Africa is much more fluid than in most OECD countries, as fluid as in the U.S.[6], and only less fluid than in the developing countries for which we have reliable estimates of worker flows for, i.e., Mexico, Turkey and Brazil.

Why are worker flows unusually large in South Africa? There are three possible explanations. The first is that Kerr (2018) overestimated worker flows. This explanation posits that the true worker flows are much smaller and more commensurate with South Africa's legislative context. The second explanation is that South African employment protection is not as stringent as is commonly believed, contrary to the popular belief that the employment protection laws are stricter in South Africa than in "virtually anywhere else in the world" (Benjamin, 2014:256).

Finally, the third explanation denies neither Kerr's finding nor the widespread belief regarding South Africa's labour market regulation stringency. Instead, it posits that certain industries, sectors, groups of people, or job contracts enjoy substantially more employment protection. In contrast, there exist relatively "unregulated" sections in the labour market where the reach of labour regulation is limited (e.g., informal sector) or the amount of employment protection guaranteed by the labour regulation is

---

[6] Based on 2011 U.S. worker flow rates in Cahuc, Carcillo and Zylberberg (2014:569)

partial (e.g., casually-contracted workers compared to workers on permanent contracts). As a result of this, risk-averse employers might be incentivised to hire more workers into the “unregulated” section, thus pushing up the aggregate level of worker flows (Martin & Scarpetta, 2012:99). This paper intends to adjudicate amongst these different explanations for the size of the worker flows in South Africa.

3.3. Heterogeneous pattern and cyclicality of worker flows

Measured at an aggregate level, gross worker flows could conceal a substantial amount of heterogeneity at the individual, firm, and industry level. It is important to examine the heterogeneous patterns of worker flows in South Africa for two reasons. First, it sheds light on whether employment protection is selectively enjoyed by some sub-sections of the labour market, which helps us adjudicate on the contending theories proposed above for explaining the large worker flow rates observed for South Africa. For instance, if casual workers experience larger worker flows than those on permanent contracts, then this might indicate differing levels of employment protection for the two groups.

With this said, variation of worker flows does not always indicate the presence of selective employment protection. Other factors may be at play. For instance, as entrants to the labour market, youths might want to improve their job match quality or move up the job ladder. They could achieve this by changing jobs more frequently. But as they become older and become more settled in their career, they change jobs less frequently. As a result, the fact that worker flows are higher amongst younger workers than amongst the elderly can be adequately explained by the life-cycle dynamics, without needing to consider varying degrees of employment protection.

A second reason for examining the heterogenous pattern of worker flows is because the varied experiences of different groups in the labour market are worth studying in their own right. Researchers have devoted a significant amount of attention to better understand the extent and nature of inequality in the context of South Africa, and have generally shown that inequality in South Africa prevails along a multitude of socio-economic dimensions, such as income, wealth, employment and health. By highlighting the diverse experiences of different population groups with regard to worker flows, we shed light on this new layer of inequality for workers in South Africa.

We now review the heterogeneous pattern of worker flows from empirical studies. Worker flows differ vastly across industries, with workers in goods-making industries (e.g., manufacturing, construction, electricity, gas and water supply) experiencing lower worker flows, and service industries workers experiencing higher worker flows (e.g. retail, wholesale, hospitality, tourism and restaurant) (Lane, Stevens & Burgess, 1996; Burgess, Lane & Stevens, 2000:481-482; Davis, Faberman & Haltiwanger, 2006:7; OECD, 2009:124-127). These patterns exist due to differences in industry-specific labour training costs, labour skill requirements, and variations in hiring and screening conventions (Lane, Isaac & Stevens, 1996).

Within industries, worker flows display substantial variations across firms. They decrease as firms become larger and older (Lane, Stevens & Burgess, 1996; Huber & Smeral, 2006:1692) – a pattern that is likely a result of the larger and older firms having managers that are more capable of screening workers. Another reason for this pattern is that workers might perceive new and smaller firms as being insufficient in terms of providing career advancement opportunities, and therefore leave these firms more frequently for opportunities in larger and older firms (Lane, Stevens & Burgess, 1996).

Individual characteristics are also determinants of one's experience of worker flows. OECD (2009:146-150) found that worker flows are larger for the younger population, females, and people at the extremes of the skill distribution (poorly educated or highly educated).

In the South African context, Kerr (2018) examined industry- and firm-level heterogeneity in worker flows. His results suggest that, while worker flows appear lower than average in goods-making industries (such as "mining and quarrying", "textile" and "vehicles, parts and accessories"), exceptions exist in industries such as "other manufacturing" and "construction", where the worker flow rates exceed the average of worker flow rates across all industries. Similarly, although worker flows in some service sectors are higher than average, they are low in other sectors – notably, worker flow rates in "educational services" and "wholesale" are lower or on par with the average worker flow rate across all industries. These findings indicate that the institutional and societal characteristics unique to South Africa may play a role in affecting the distribution of worker flows across industries.

Industry-level heterogeneity of worker flows might also be influenced by variations in firm sizes and earnings across different industries. Kerr (2018) found worker flows decrease with firm size and earnings, consistent with international studies. So, it is plausible that a part of the cross-industry variation of worker flow rates is attributable to variations in firm sizes and earnings by industry. This highlights the fact that, in the heterogeneity analysis, the correlations between covariant characteristics make isolating the effect of any one characteristic on worker flows difficult. To ameliorate this issue, we could use parametric methods that control for other covariant variables so as to isolate the impact of one specific characteristic of interest (e.g., Zizzamia & Ranchhod, 2019).

In terms of individual characteristics, Ker (2018) found worker flows to be larger for those at the lower end of the wage distribution. However, the data used in the study notably contained scant information on workers' individual characteristics, and thus a full examination of the variation of worker flows by individual characteristics was not possible. With that said, there exists an abundant body of labour dynamics literature in South Africa that points to the important roles of race, gender, age, education, and trade union membership in influencing dynamic patterns of labour market movement. How these characteristics affect worker flows will be examined in this paper.

Lastly, worker flows vary with business cycles. Davis and Haltiwanger (1998:97-104) calculated the correlation coefficient between net employment growth and gross worker flows in the U.S. from 1930 to 1980. They found the correlation coefficient to be 0.31, indicating weak procyclicality of worker flows. Cahuc, Carcillo and Zylberberg attribute the weak procyclicality of aggregate worker flows to a combination of procyclical hiring and acyclical separations (Cahuc, Carcillo & Zylberberg, 2014:572-573).

3.4. Efficiency and welfare implication of worker flows

Haltiwanger et. al. (2004:192) remark that there exists a trade-off as a labour market becomes more fluid, that is, as the magnitude of worker flows increases. On the one hand, increased worker flows reflect increased allocative efficiency. On the other hand, large worker flows have the potential to disrupt the job security of workers, thereby undermining the welfare of those holding precarious jobs. While it is beyond the scope of this paper to explore the full implications of worker flows, we summarise some arguments from both sides of the debate.

3.4.1. Links between worker flows and efficiency

In a review of the theoretical link between worker flows and efficiency, Donovan, Lu and Schoellman (2020) summarised that the conventional thinking in the literature is that the poor functioning of the labour market leads to the misallocation of workers across regions, firms, and jobs, which result in lower labour productivity. A fluid labour market is more capable of adapting to economic and technological shocks and it, therefore, facilitates the efficient reallocation of productive resources. This enhances aggregate productivity.

But there is evidence against this conventional wisdom. When Donovan, Lu and Schoellman (2020:14) compared measures of labour market dynamics across countries at different levels of development (measured by GDP per capita), they found that the extent of labour market flows decreased with the level of a country's economic development, which is "at odds with the conventional wisdom".

Furthermore, at a disaggregated level, a lower level of worker flows may itself be a result of productivity improvement. Davis and Haltiwanger, (2014:10-11) used the example of U.S. chain supermarkets such as Walmart to illustrate this point. They noted that, as the retail sector become dominated by national chain stores like Walmart, worker flow rates in the retail sector subsequently fell. However, this reduced labour market fluidity is concomitant with an increase in the retail sector productivity.

Large worker flows could also undermine economic efficiency. To illustrate this, Martin and Scarpetta (2012:99) offered a critical view of the deregulation of temporary work contracts in Europe in recent times. They noted that the partial deregulation of the labour market pushed firms to replace permanent work contracts with temporary ones. Although worker flows increased as a result, the deregulation has also distorted the optimal allocation of employment, reduced workers' commitment to the firm, and reduced firms' incentive to invest in workers – all of these had a long-term negative impact on economic productivity. This view is particularly pertinent for South Africa in light of, on the one hand, what seems to be large worker flows (Kerr, 2018) and, on the other hand, the increasing popularity of labour brokers and the rising share of temporary employment (Bhorat & Lilenstein, 2016).

3.4.2. Links between worker flows and welfare

Notwithstanding the efficiency implication of worker flows, there is little doubt that worker flows have a direct bearing on workers' welfare. If increased worker flows result in efficiency gain, and this gain is distributed equitably across all workers, then the predominant effect of labour market fluidity will be benign. In reality, however, efficiency gain may be distributed unequally, with some people benefitting and others losing.

One group of such "losers" are the displaced workers. In a fluid labour market, job security is constantly influenced by fluctuations in aggregate demand. When workers are retrenched due to a fall in demand, they become displaced. Davis and Haltiwanger (2014:10-11) found involuntary separations as such negatively affect displaced workers in terms of "health, marital stability, emotional well-being". Work displacement also negatively affects the children of the workers in terms of educational achievement and cognitive development.

Employment protection is often proposed as a way to protect the employed population by reducing excessive fluidity in the labour market. However, while they protect the employed, they are likely to hurt the unemployed. Reduced labour market fluidity leads to longer spells of unemployment. This depreciates the market-relevant human capital of the unemployed. Consequently, the unemployed are more likely to travel a "destructive" path where long-term erosion of human capital reinforces their marginalisation (Davis & Haltiwanger, 2014).

Where there are "losers", there are "winners". If workers quit their jobs due to skill mismatch or preferences and if they subsequently find better jobs elsewhere, then their welfare will be increased. For instance, Bassanini et al. (2010) showed that, in several European countries, people who changed jobs from 1995 to 2001 managed to achieve on average a 6% wage premium in the new job. Certainly, not all job switchers experience an improvement in their welfare. People could move from a higher-paying job to a lower-paying job. Furthermore, in some cases, the job-specific human capital is destroyed as they move from one job to another, which negatively affects their future career prospects. In other cases, job movers are forced to relocate geographically, which negatively impacts their and their family's well-being (OECD, 2009:118).

Overall, there is a trade-off between efficiency and welfare when it comes to worker flows, but this relationship is more complicated than is often caricatured. For this reason, interpretations of findings on worker flows in the context of South Africa should be done in such a way that considers both sides of the debate – that is, by looking at both workers' welfare and economic efficiency.

4. Data Description

This study examines worker-level worker flows in the context of South Africa by utilising both QLFS data and the Post-Apartheid Labour Market Series (PALMS). We describe these two datasets below.

4.1. Quarterly Labour Force Survey

Introduced to replace Labour Force Survey (LFS), the QLFS has been conducted on a quarterly basis by StatsSA since 2008 and collects labour market information on individuals aged 15 years or older living in South Africa. QLFS cross-sectional samples are nationally and provincially representative, and they are the official data source of South Africa's headline labour indicators such as the unemployment rate. Its questionnaire covers a broad range of topics, including respondents' socio-economic background, job-searching strategies for the unemployed, coping strategies for the economically inactive, and work arrangements for the employed.

Released as quarterly cross-sectional datasets, the QLFS is designed as a rotating panel of dwelling units (DU). A dwelling unit is a piece of ground containing one or several housing units where households live. We make the explicit differentiation between a "household", which refers to the people who live under the same roof and share resources, and a "housing unit", which refers to a building structure where a household resides. Around 30 000 dwellings were surveyed in QLFS every quarter from 2008 to 2015. This number increased to 33 000 from 2015 onwards following a re-design of the master sample.

In accordance with the rotating panel design, all dwellings were assigned to one of the four rotation groups. Every quarter, 1/4 of the dwellings enumerated in the previous survey are rotated out of the sampling frame and are replaced by new dwellings. Thus, 3/4 of the dwellings from the previous quarter are not rotated out, and they are revisited by fieldworkers. Therefore, between every two consecutive QLFS cross-sectional samples, three-quarters of the sampled dwellings units should overlap. As long as there are some households that stayed in the same dwelling across the two quarters and some individuals have not moved out of these households, we are able to construct a quarterly panel of individuals from the households in the overlapping dwellings.

As is indicated, QLFS does not track households or individuals over time, but, instead, it tracks dwellings. Therefore, we are unable to track an individual from one quarter to the next if 1) her entire household moves out of the dwelling 2) she moves out of the household 3) there is household or individual non-response 4) she passes away. Furthermore, when the master sample is refreshed, such as seen in 2015, all dwellings from the old master sample are completely rotated out to make room for the new ones, and none of the individuals from the old master sample can be tracked to the next quarter. This implies that no individuals in the 2014q4 QLFS sample could be tracked to the 2015q1 sample.

4.2. Post-Apartheid Labour Market Series

In addition to using the original data files from StatsSA, we also utilise the Post-Apartheid Labour Market Series (PALMS) dataset (see Kerr, Wittenberg & Lam, 2019). PALMS is a stacked cross-sectional dataset which, in its most recent version (version 3.3), comprises 69 household surveys conducted between 1994 and 2019 in South Africa. This includes a series of QLFS ranging from 2008q1 to 2019q2, which is the time range studied in this paper.

I utilise the PALMS dataset instead of replicating the work of appending the QLFS series because considerable efforts have been made to ensure that the constituent survey data in PALMS is harmonised and consistent across time (Kerr & Wittenberg, 2019). One such feature that we take advantage of is the cross-entropy (CE) weights in PALMS, which we use as the basis for the adjustment of weights for panel attrition. The CE weights are benchmarked to consistent demographic and geographic trends that are projected for the population totals. From a time-consistency perspective, this is superior to the original weight variables which are benchmarked to time-inconsistent population aggregates that exhibit jumps in values from time to time (Branson & Wittenberg, 2014:20). Basing the panel weights on CE weights, as opposed to the original weights, ensures that the time variation in the patterns of labour market dynamics reveals useful facts about the underlying labour market, rather than reflecting the inconsistency of weights in various surveys.

I also exclude from this study the more recent rounds of QLFS that are yet to be incorporated into PALMS (i.e. QLFS 2019q3 or later). Harmonising the new rounds of QLFS and creating new CE weights falls beyond the scope of this paper.

5. Matching Algorithms

The rotating sample design of the QLFS, described above, enables us to match individuals across consecutive QLFS cross-sections, thus giving us a set of quarterly individual-level panel data that can be used to examine worker flows in the context of South Africa. This section deals with the matching algorithm used in the data-matching procedure.

Although previous studies offer several matching algorithms that can be used on the QLFS data, there has been very little effort by researchers to investigate the implication of different matching algorithms on the matching quality of the resulting panel. As such, there is substantial ambiguity on how accurately the matched panel data tracks the same individuals over time, and whether the matching accuracy is influenced by the choice of matching algorithm. These questions are important to us, particularly because it is difficult to claim credibility in our findings unless we are certain that the most optimal matching algorithm is followed in the making of our panel samples and that it accurately tracks the same individuals over time. Therefore, in this section, we provide a detailed description and a thorough investigation into matching algorithms for the QLFS data. A method for evaluating the quality of different matching algorithms is proposed, which is then used to select an optimal matching strategy used in this study.

A brief outline of this section is helpful. At the outset, we provide an overview of different terminologies that appear in the rest of the section. Following this, we provide a summary of the different matching algorithms used by researchers in the past. In order to quantitatively evaluate the matching quality of matching algorithms, an evaluation method is proposed. This evaluation method is then undertaken to help us evaluate the matching quality of four different types of matching algorithms. Finally, based on the evaluation results, we choose an optimal matching algorithm.

5.1. Terminologies

A total of 46 rounds of QLFS cross-sections are utilised for this paper, starting with QLFS 2008q1 and ending with QLFS 2019q2. Following the naming convention in PALMS, cross-sections are also called "waves" – QLFS 2008q1 is wave 23, QLFS 2008q2 is wave 24, and so on.

If we use letter "t" to denote a specific quarter, wave $t-1$ denotes the QLFS cross-section of the previous quarter, while wave $t+1$ denotes the next quarter. Therefore, the panel creation process is essentially a process whereby we match individual records that represent the same person between wave $t-1$ and wave t. A matched panel based on wave $t-1$ and wave t is called "panel t", where "t" is its "panel number". For instance, panel 23 contains matched records from QLFS wave 22 and wave 23.

We will provide a review of several matching algorithms used in previous studies. To differentiate amongst the matching algorithms, we name algorithms by the abbreviated surnames of the associated authors. For instance, the matching algorithm conceived by Leung, Stampini and Vencatachellum (2014) is named "LSV" algorithm. We differentiate panels produced from different algorithms in an analogous fashion – e.g., a LSV panel is based on LSV algorithm.

Finally, we differentiate between "official panels" and "matched panels". Official QLFS panels are produced by StatsSA. Like matched panels, official panels track individuals over time. An important distinction between the two is that the official panels may be qualitatively different from the matched panels because StatsSA has access to confidential information about the individuals (e.g., their names), which could potentially be used to improve the official panel's matching quality.

5.2. Description of QLFS matching algorithms

5.2.1. The general approach

It is not a straightforward process to link the same individuals across consecutive cross-sections of the QLFS. To see this, one first needs to understand how identifiers are recorded in the data. Individuals in a specific QLFS cross-section are uniquely identified by two variables: a household identifier (UQNR) and an individual line number within the household (PERSONNR). When a dwelling is enumerated for the first time in the QLFS, all households in the dwelling are approached by the fieldworker for the survey. Each successfully interviewed household is given a unique UQNR. Therefore, UQNR uniquely identifies households within each cross-section. However, it is important to note that UQNR is technically a geographic indicator: each UQNR refers to the location where a household resides (e.g., house, shack, etc.). During the survey interview, the fieldworker gives a unique line number to each household member: coded as PERSONNR, it simply represents the order in which a household member is interviewed (e.g., if a household member is the first person to be interviewed within that household, then his/her PERSONNR is 1). Therefore, PERSONNR uniquely identifies individuals within each household. As such, every individual in a cross-sectional QLFS dataset has a unique UNQR-PERSONNR combination.

However, the UNQR-PERSONNR combination cannot be used to uniquely identify the same individuals over time. That is, PERSONNR i in UQNR j in wave $t-1$ does not necessarily represent the same person as PERSONNR i in UQNR j in wave t. On the one hand, an entire household could move in or out of a residential location denoted by UQNR j: in this case, the family living in UQNR j in wave $t-1$ is different from the family living in the same UQNR j in wave t. On the other hand, even if the household stays in the same residential location denoted by UQNR j over time, the household members could be given different line number in different waves – in this case, even if the family staying in UQNR j in wave $t-1$ is the same family living in UQNR j in wave t, PERSONNR i in wave $t-1$ could be a different person from PERSONNR i in wave t. Thus, we cannot track individuals between consecutive quarters of the QLFS data by merging unique combinations of UNQR-PERSONNR across quarters.

Lacking an intertemporally consistent individual identifier, researchers track individuals across waves by merging their UQNR and observable demographic characteristics, such as age, gender, self-declared race. For instance, if there is a 25-year-old Indian woman living in the same location denoted by a specific UQNR number across two consecutive waves of the QLFS, then one could reasonably assume that these two records represent the same person.

5.2.2. Matching algorithms in previous studies

Several matching algorithms have been proposed in the past literature, which are summarised below. By highlighting the similarities and differences between different algorithms, we also integrate and consolidate the knowledge base on the topic at hand, which has, thus far, been disconnected and fragmented.

The first two algorithms come from Ranchhod and Dinkelman (2008) who made individual-level panels by matching consecutive LFS cross-sections over 2001 – 2003. The LFS has the same rotating panel design and has similar questionnaire designs as the QLFS, but it is conducted every six months. Thus, matching algorithms developed for LFS can be used on the QLFS data. These authors proposed two algorithms with slightly different matching criteria. The first algorithm (which we call RD-expanded) matches individual records who have the same UQNR, gender, and self-declared race between wave $t-1$ and wave t. It allows for a cross-wave age difference of up to one year. The second algorithm (which we call RD-strict) imposed additional criteria. In addition to the criteria in RD-expanded, RD-strict also excludes matched records who, from t-1 to t, reported a decrease in the ability to speak/write, a decrease in educational attainment, a decrease in age, or if the matched person changed from being "once married" at t-1 to "never married" at t. Note that information on one's ability to read/write is only collected in LFS but not in QLFS.

Two additional features of these two algorithms are worth noting. Firstly, Ranchhod and Dinkelman (2008) excluded records from the matched panel if they reported having moved into a new dwelling unit between $t-1$ and t. Whether or not someone recently moved into a dwelling unit was information collected only in LFS but not in QLFS. Secondly, before matching is performed, Ranchhod and Dinkelman (2008:4) deleted any Individual records from the same household in a particular wave with the same race, gender and age differing by at most one year. They did this because the deleted records cannot be reliably matched. To see why, consider the following example. Suppose there are two 25-year-old African women living in a house, and in the next wave there are also two 25-year-old African women living in the same house. Since we only differentiate records based on age, gender and race, we simply cannot tell apart the records that belong to the first woman vis-á-vis those belonging to the second.

Anand, Kothari and Kumar (2016) (hereafter AKK algorithm) proposed an algorithm for matching consecutive QLFS cross-sections. The AKK algorithm merges individual records who, between $t-1$ and t, have identical UQNR, PERSONNR, gender, and self-declared race. It also allows for the cross-wave age difference of up to one year. As discussed above, PERSONNR is assigned to a household member within every household but it cannot be used to identify the same person across time. By imposing the PERSONNR restriction, the AKK algorithm will likely exclude people who should have been included in the panel but had different PERSONNR between t-1 and t.

Lastly, the algorithm used in Leung, Stampini and Vencatachellum (2014) (hereafter the LSV algorithm) merges observations who have, between $t-1$ and t, identical UQNR, PERSONNR, gender, race and education in years. The criteria for the above-mentioned algorithms are tabulated and presented below in Table 1.

Table 1: Summary of matching criteria for different algorithms

<table>
<tr><th></th><th>RD-expanded</th><th>RD-strict</th><th>AKK</th><th>LSV</th></tr>
<tr><td>Pre-requisite for matching</td><td colspan="2">Individual records from the same household in a particular wave with the same race, gender and age differing by at most one year are deleted before matching starts</td><td colspan="2">None</td></tr>
<tr><td>UQNR</td><td colspan="4">Identical match between t-1 and t</td></tr>
<tr><td>PERSONNR</td><td colspan="2">None</td><td colspan="2">Identical match between t-1 and t</td></tr>
<tr><td>Gender</td><td colspan="4" rowspan="2">Identical match between t-1 and t</td></tr>
<tr><td>Self-declared race</td></tr>
<tr><td>Age</td><td>$|age_{t-1} - age_t| \leq 1$</td><td>$age_{t-1} = age_t, or$ $age_{t-1} = age_t - 1$</td><td>$|age_{t-1} - age_t| \leq 1$</td><td>None</td></tr>
<tr><td>Education in years</td><td rowspan="3">None</td><td>$Educ_{t-1} \leq Educ_t$</td><td>None</td><td>$Educ_{t-1} = Educ_t$</td></tr>
<tr><td>Ability to read/write</td><td>Ability in reading or writing not allowed to decrease from t-1 to t</td><td colspan="2" rowspan="3">None</td></tr>
<tr><td>Marital status</td><td>Marital status not allowed to change from “once married” at t-1 to “never married” at t</td></tr>
<tr><td>Other criteria</td><td colspan="2">Any matched records who reported moving into a new dwelling unit between t-1 and t are deleted</td></tr>
<tr><td>First used in</td><td colspan="2">Labour Force Survey</td><td colspan="2">Quarterly Labour Force Survey</td></tr>
</table>

Notes: The matching criteria presented in the table refer to what the following authors would have done to match LFS/QLFS cross-sections between t-1 and t; ***Sources:*** Ranchhod and Dinkelman (2008); Leung, Stampini and Vencatachellum (2014); Anand, Kothari and Kumar (2016).

5.3. An evaluation method for QLFS matching algorithms

The matching quality of the QLFS panel (i.e., how accurately it tracks the same individual over time) is likely to vary among matching algorithms. However, the relationship between matching quality and matching algorithm remains ambiguous, and, to our knowledge, has not been investigated systematically. To address this lacuna in the literature, we present an evaluation approach that enables us to quantify and compare the matching quality of different algorithms.

5.3.1. The guiding principle

We were given access to a series of official QLFS panels[7]. Made by StatsSA for internal use, the official panels contain matched individuals who were present in the QLFS cross-sections across consecutive quarters from 2010q1 (wave 31) to 2013q4 (wave 46). Since there is a total of 16 QLFS cross-sections between 2010q1 and 2013q4, there are 15 official panels. Following the naming convention, these panels are referred to as "official panel 32" (matched between wave 31 and 32), "official panel 33" (matched between wave 32 and 33), and so on.

With this data on hand, the guiding principle for evaluating the quality of a matching algorithm is as follows. The matching quality of an algorithm is evaluated by comparing the matched panels against the official panels. Under the assumption that the official panels are superior in matching quality, the magnitude of the discrepancy between the matched panel and the official panel becomes an indicator of the underperformance of the matching algorithm.

This method relies on the assumption that the official panels are superior in matching quality. We believe this is the case. In the making of the official QLFS panels, StatsSA matches individuals on names and surnames across waves, in addition to using observable demographic variables (StatsSA, 2018:98). Names and surnames are confidential information, to which only StatsSA has access, and matching individuals based on names as well as demographic characteristics is clearly more accurate than just relying on demographic characteristics. The assumption that official panels are superior in quality is also made by Anand, Kothari and Kumar (2016:7).

However, if official panels contain many false matches or simply fail to match many people who should have been included in the panel, then the assumption of its superiority may no longer be valid. There is no way to test this, but we acknowledge such a possibility. If the "superiority" assumption about the official panels is invalid, then our approach will also be invalid. In that case, comparisons between the matched panels and the official panels reveal nothing about the matching quality of the underlying matching algorithm. Instead, comparisons as such will simply reflect how closely the matching algorithm can replicate official panels.

5.3.2. Performance metrics

The extent to which a matched panel is similar to/different from an official panel can be quantified using "performance metrics", which are indicators of the matching quality for a matching algorithm.

---

[7] Thanks to Associate Professor Andrew Kerr from the University of Cape Town.

Suppose there is an individual record called i in wave $t-1$. The matching algorithm either matches it to another record called j in wave t (i_j match), or considers it a non-match (i_null match). These are the matching outcomes generated by the matching algorithm, which we call the "predicted matching outcomes". For the same individual record i in wave $t-1$, StatsSA will match it to record j at t, to record k at t (where $k \neq j$), or consider it to be a non-match. This information will be available in the official panel. These matching outcomes are assumed to be the "true matching outcomes", since we assume the official panels to be superior in quality.

Then, for every person i in wave $t-1$, we can compare her "predicted matching outcome" against the corresponding "true matching outcome". Drawing on the literature of confusion matrix (Tharwat, 2018), we can classify the predicted matching outcome into four distinct categories, shown below in Table 2. If i is matched to the same record j in both matched panel and official panel, it is considered a true match – which we call "true positive" (TP). If i is matched by the algorithm to j, but is matched to different record k in the official panel, then it is a false match – "false positive" (FP). If i is not matched to any record by the algorithm, but it is matched to a record in the official panel, then we have a case of false non-match, or "false negative" (FN). If i is neither matched by the algorithm, nor matched in the official panel, we have a true non-match – "true negative" (TN).

Table 2: Confusion matrix for matching outcomes

	i_j	i_k $(k \neq j)$	i_null
i_j	TP	FP	FP
i_null	FN	FN	TN

Notes: Row = Predicted Matching Outcome
Column = True Matching Outcome

With these four elements (TP, FP, FN, TN), we can calculate some performance metrics for an underlying algorithm[8] (Tharwat, 2020). The first metric is the true positive rate (TPR), calculated as proportion of true matches correctly predicted by an algorithm. The second metric is the true negative rate (TNR), calculated as the proportion of true non-matches correctly predicted by an algorithm. The formula for TPR and TNR are presented in equation (8) below. Collectively, TPR and TNR indicate

[8] All performance metrics presented below all have well-known distributions and statistical properties which fall beyond the scope of this paper. However, this may be a fruitful avenue for future research.

how an algorithm performs from two different angles – how well it captures true matches and how well it captures true non-matches.

$$8)\ TPR = \frac{TP}{TP + FN};\ TNR = \frac{TN}{TN + FP}$$

We present two more metrics. Firstly, positive predictive value (PPV) is defined as the proportion of matches predicted by the algorithm that are correct. Secondly, negative predictive value (NPV) is defined as the proportion of non-matches predicted by the algorithm that are correct. PPV and NPV are defined more precisely in equation (9). Collectively, PPV and NPV indicate the predictive accuracy of an algorithm.

$$9)\ PPV = \frac{TP}{FP + TP};NPV = \frac{TN}{FN + TN}$$

According to Madrian and Lefgren (1999), if an algorithm performs better in terms of both TPR and TNR, then it is more capable of capturing both true matches and true non-matches. Therefore, it should be superior. Similarly, if an algorithm generates better PPV and NPV, it is more accurate in terms of predicting both matches and non-matches. Then by the same logic, it is also superior. These rules help us rank the performance of different algorithms, thus enabling us to evaluate and compare, in a systematic way, how different algorithms affect the matching quality of the resulting matched panels.

5.4. Evaluating QLFS matching algorithms

Using the evaluation method outlined above, we can investigate the quality of the above-mentioned matching algorithms used in previous literature – AKK algorithm, RD-expanded algorithm, RD-strict algorithm, and LSV algorithm. Such an exercise will aid us in choosing an optimal algorithm, which will be used for this study.

First of all, we generate four sets of QLFS matched panels from these four matching algorithms. Individual records are matched across consecutive QLFS cross-sections from 2008q1 to 2019q2. Consistent with the age restriction in the official panel, the matched individuals in a matched panel t

are aged 15-64 inclusive in both wave $t-1$ and wave t[9]. A total of 44 quarter-to-quarter matched panels are generated per algorithm. Records cannot be matched between QLFS 2014q4 and 2015q1, and this is because the QLFS master sample was redesigned at the beginning of 2015.

When implementing RD-expanded and RD-strict algorithms on QLFS data, we were not able to replicate all matching criteria. In particular, we could not ascertain if a matched person recently moved into the dwelling unit between $t-1$ and t, and we could not inspect the time-inconsistency in one's ability to read/write across waves. There is insufficient information in the QLFS data for conducting these checks.

5.4.1. Sample size comparison

The size of the panel sample generated by each algorithm reveals the first informative insight into the differences among the algorithms. AKK algorithm is able to identify an average of 30170 individuals in a quarter-to-quarter QLFS panel. This is followed by RD-expanded, which identifies an average of 29639 individuals in each panel. The average sample size of the RD-strict panel is smaller at 25881. The LSV panels are the smallest in size: about 23966 individuals are identified in an average panel. To put these figures into perspective, there are around 50013 working-age individuals in an average QLFS cross-section from 2008q1 to 2019q2. And from 2010q1 to 2013q4, there are on average 32607 individuals in a quarter-to-quarter official panel dataset.

There is a precipitous decline in the sample size of all matched panels from the year 2015. This is caused by a large reduction in the QLFS sample size from 2015. This is because, after the 2015 redesign of the QLFS master sample, a larger share of the sampled clusters of QLFS came from Gauteng, where the rate of urbanisation is higher and the individual response rate is, as a result, lower (Kerr and Wittenberg, 2019:9).

5.4.2. Match rate comparison

To calculate the match rate of an algorithm for wave t, we divide the number of matched individuals between t and $t+1$ by the total number of working-age individuals in wave t. Match rate for wave t

[9] In the official panel, individuals who are matched between t-1 and t are aged 15-64 at $t-1$, but their age ranged from 15 to 67 at t. For consistency, we restrict our attention to matched records in the official data whose age is between 15-64 in both $t-1$ and t. Doing so means ignoring 1320 matched individuals who are aged 64 in $t-1$ but are more than 64 year old in t in the official panel. But this omission will have negligible effect in our analyses below as the deleted sample represents less than 0,01% of the sample size.

is therefore the proportion of all working-age individuals in wave t that are successively matched by an algorithm to the next wave. For instance, if the AKK match rate for QLFS 2009q1 is 52.92%, then 52.92% of all working age individuals in QLFS 2009q1 were matched by the AKK algorithm to the next quarter.

We calculated the match rates by algorithm and quarter from 2008q1 to 2019q1, and we presented this information in Figure 1 below. For comparison purposes, we also presented the match rates for the official panels by quarter from 2010q1 to 2013q3. Also included in figure 1 is a "max rate", which is the percentage of working age individuals in the wave t who live in houses that were successfully interviewed in wave $t + 1$[10]. "Max rate" represents the maximum portion of working-age individuals at wave t who could be matched to wave $t + 1$, assuming that 1) the same household lived in the interviewed house between t and $t + 1$; 2) no one moved out of the household between t and $t + 1$; and 3) there are no deaths and no individual non-response[11].

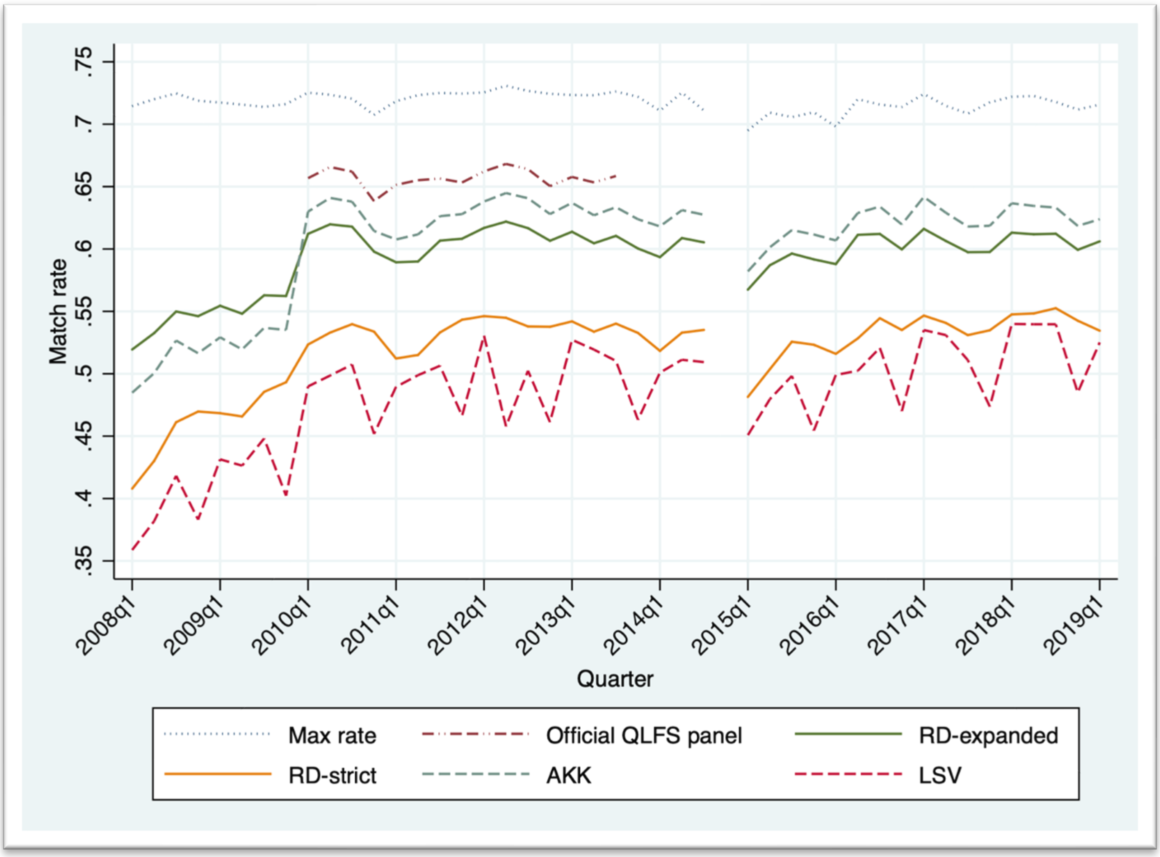
Figure 1 shows that the official QLFS panels have the highest match rates (with an average of 65,69%) whereas the lowest match rates are associated with the LSV algorithm (with an average of 48,20%). Of the four algorithms we study, the RD-expanded and AKK contend for the highest match rate, but, on average, the AKK match rate (60.57%) is slightly bigger than that of RD-expanded (59,39%). We also observe that, the more strict an algorithm is, the lower its match rate. For instance, the RD-strict match rate is much lower than that of RD-expanded, because the former has stricter matching criteria.

Since the "max rate" in Figure 1 represents the theoretically maximum match rate, a comparison against it reveals the proportion of cross-sectional units that the algorithm fails to match. We find that even the official match rates, which are the highest amongst all matched panels, are on average 6,09 percentage points lower than the "max rate". A large portion of this gap may be explained by individual non-response, migration or death, as it is not inconceivable for these factors to affect 6% of the working-age population every quarter.

[10] Whether a house is successfully interviewed in wave $t + 1$ is determined by examining if its UQNR appears in both t and $t + 1$.

[11] Theoretically, the "max rate" should be approximately 75%. Because 3/4 of all dwelling units are revisited every quarter, roughly 3/4 of all houses at wave t should also be revisited in wave t+1. The reason why the actual max rate is below 75%, as is seen in the graph below, is because some houses in wave t refused or failed to respond at wave t+1 even though they were enumerated.

Figure 1: Match rate by algorithm, 2008q1 – 2019q2



Notes: The break in the line between 2014 and 2015 is due to non-matching as a result of the master sample revision. ***Sources***: Author's calculations from QLFS, PALMS, and the official QLFS panel data.

With this said, non-response, migration and mortality cannot adequately explain the large drop in the match rates for some of the algorithms, such as LSV and RD-strict. Their match rate plots lie far below the "max rate", differing by as much as 20-40 percentage points. This is indicative of the algorithms' failure to track people who have appeared across consecutive waves.

The zigzagged pattern displayed by the LSV is seasonal – the plot dips in the last quarter of most years but it rebounds in the first quarter of a new year. The school calendar seems to be the reason behind this pattern. School attendees who are aged 15 or older might report a one-year increment in their educational attainment in the first quarter of a new year, and since LSV requires identical education of matched records between consecutive quarters, these "students" will be excluded from the panels made from matching the q4 sample of QLFS and its q1 sample of the next year.

5.4.3. AKK match rates vs RD-expanded match rates

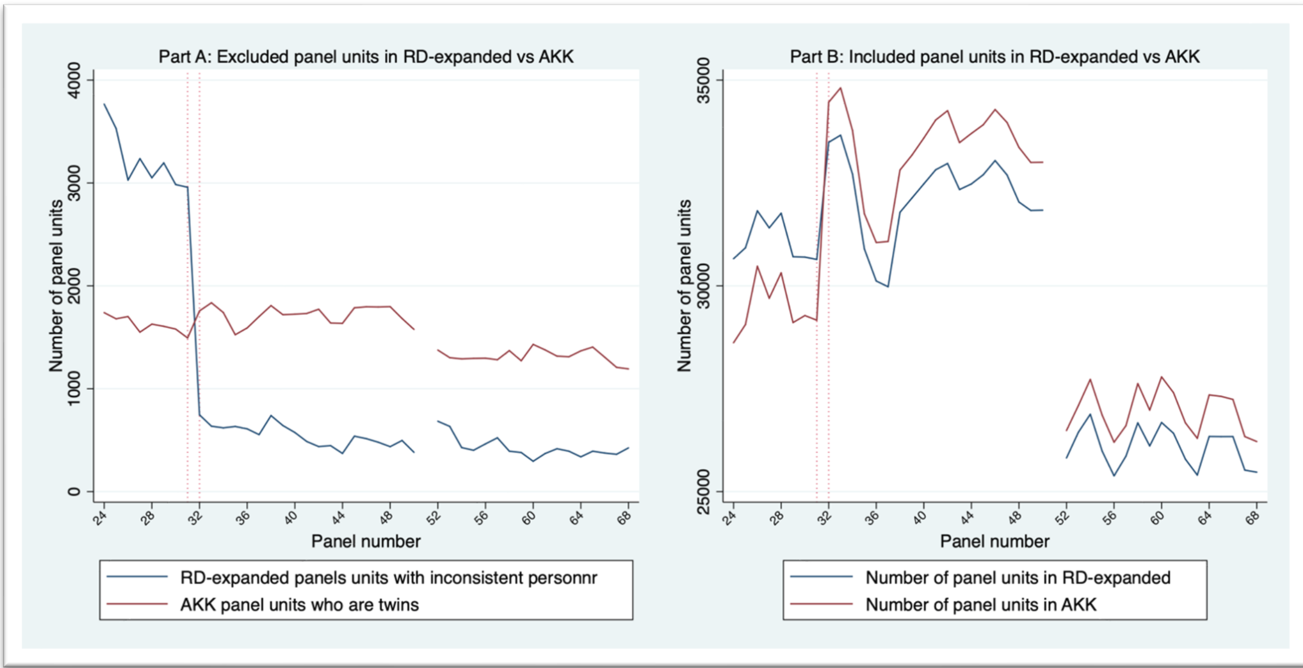
Figure 1 also shows that the RD-expanded match rate is consistently above that of AKK prior to 2010q1. But after 2010q1, this relationship was reversed and the match rate of AKK panels became larger and consistently so for the next 9 years. This is peculiar because the AKK algorithm is identical to the RD-expanded algorithm in terms of matching criteria on gender, age and self-declared race, but AKK furthermore requires identical matches on PERSONNR across consecutive waves. Therefore, AKK is technically more restrictive than RD-expanded, which means that we expect the AKK match rate to be lower.

What happened in 2010 that led to the higher match rate observed for the AKK algorithm? To answer this question, we must consider two sides of the story. On the one hand, some RD-expanded panel individuals will be excluded from AKK panels for having different PERSONNR across consecutive quarters. But on the other hand, some AKK panel individuals will also be excluded from the RD-expanded matched panels. Before the RD-expanded algorithm is implemented, Ranchhod and Dinkelman (2008) deleted any cross-sectional records who live in the same house, have the same race, gender, and have ages that differ by less or equal to a year. For convenience, we refer to these deleted records as "twins"[12]. While "twins" are excluded from the RD-expanded panels, they are not excluded from the AKK panels. This is because PERSONNR, which AKK utilises as a part of its algorithm, can help differentiate "twin" records across panels. Consequently, the presence of "twins" will reduce the match rate for RD-expanded relative to AKK.

Therefore, one possible explanation for the match rate of RD-expanded to drop below that of AKK is that, from 2010, there are more "twins" than there are panel individuals with different PERSONNR across consecutive waves. We find this explanation to be valid and present our evidence in Figure 2 below. Part A of Figure 2 plots two lines: one indicates the number of AKK panel individuals who are part of "twin" records, and the other line plots the number of RD-expanded panel individuals who have time-inconsistent PERSONNR. The former represents AKK panel individuals who would have been deleted by the RD-expanded algorithm, while the latter represents the RD-expanded panel individuals who would have been deleted by the AKK algorithm. On the right, Part B of Figure 2 plots the sample size of the two panels by panel number.

[12] Since they share common demographic traits with other household members. Of course, they might not be biological twins despite sharing age, race, and household. And there may be more than two household members who are all considered "twins", e.g., three 25-year-old Indian women living in the same household in a wave.

Figure 2: Sample size comparison between AKK and RD-expanded panels



Notes: (1) The two vertical dotted lines represents panel 31 (matched between 2009q3 and 2009q4) and panel 32 (matched between 2009q4 and 2010q1) (2) The break in the line between 2014 and 2015 is due to non-matching as a result of the master sample revision. ***Sources***: Author's calculations from QLFS and PALMS data.

As we see in part A of Figure 2 (where panel 32 is matched between QLFS 2009q4 and 2010q1): before the year 2010, AKK rejects **more** panel units with inconsistent PERSONNR than RD-expanded rejects "twins", but after 2010, RD-expanded rejects **more** "twins" than AKK rejects panel units with inconsistent PERSONNR. As a result of this, the AKK match rates exceeded that of RD-expanded from 2010 onwards.

Furthermore, the drop in instances of inconsistent PERSONNR is not caused by the reduced sample size of RD-expanded panels. In fact, both panels increased in size drastically following 2010 as shown by the steep rise of two plots from panel 31 to 32 in Part B of Figure 3. Moreover, the sample size of AKK panel data expanded much more than that of RD-expanded panel data as a result of the reduced prevalence of time-inconsistent PERSONNR.

Our findings suggest that StatsSA became better at maintaining PERSONNR from 2010. Furthermore, the quality of PERSONNR improved in the long run, because it remained well-maintained from 2010 until 2019. Unfortunately, we cannot provide a satisfactory account of what exactly led to the improvement. We compared the pre-2010 questionnaire with the post-2010 questionnaire but found no significant changes in the way PERSONNR is assigned to household members. A reasonable speculation is that a revision was undertaken by StatsSA to improve the intertemporal consistency of

PERSONNR at the end of 2009 as a part of a revamp of the QLFS study, which ultimately led up to, amongst other things, the first publication of the Labour Market Dynamics Report of 2009.

5.4.4. Performance metrics comparison

We now evaluate the matching quality of various matching algorithms by measuring and comparing their performance metrics. Comparisons are made between matched panels and official panels from 2010q1 to 2013q4. This enables us to calculate the four performance metrics: TPR, TNR, PPV and NPV. The mean values of performance metrics by algorithm are presented in Table 3 below. Recall that an algorithm is superior if it has higher TPR and TNR, or if it has higher PPV and NPV.

The first noteworthy observation from Table 3 is that the AKK algorithm generates some of the best metrics. It has the highest TPR and NPV, and reasonably high TNR and PPV. This was not expected from an algorithm that requires identical matches on PERSONNR, which we thought to be poorly maintained over time. The fact that PERSONNR is not as badly maintained as we thought, as shown above, might explain why its performance exceeded our expectations.

Table 3: Mean values of performance metrics by algorithm, 2010q1-2013q4

Algorithm	*Mean TPR*	*Mean TNR*	*Mean PPV*	*Mean NPV*
AKK	95,16%	98,84%	99,27%	91,62%
RD-expanded	92,17%	99,10%	99,45%	86,95%
RD-strict	81,03%	99,42%	99,60%	73,28%
LSV	74,18%	98,20%	98,55%	66,78%

Sources: Author's calculations from the QLFS, PALMS, and the official QLFS panel.

An examination of PPV values shows that all matched panels are predominately comprised of correctly matched individuals. On average, 99,60% of all individuals in the RD-strict panels represent true matches. This is followed by RD-expanded and AKK. Although LSV has the lowest average PPV, it is still quite high at 98,55%. The high PPV values observed across all panels imply that all matched panels contain a negligible proportion of false matches. This is reassuring because false matches bias estimation results.

LSV appears inferior to all other algorithms. Not only does it generate the lowest average values of TPR and TNR, but it also has the lowest PPV and NPV mean values. That is, LSV is worse at capturing and predicting true matches as well as true non-matches. However, a naïve comparison of mean values of performance may not be statistically rigorous. As noted in Fawcett (2006:869), watertight evidence for an algorithm's superiority/inferiority requires us to not only examine the values of the metrics but also to take account of the measures of variance for these metrics. Nevertheless, by directly comparing the mean metrics, we shed light on the poor performance of the LSV algorithm.

Setting LSV apart, it is much more challenging to claim inferiority/superiority for the remaining algorithms. This is because there appears to be an inverse relationship between TPR and TNR and, similarly, an inverse relationship between PPV and NPV. Consequently, in the remaining algorithms, no single algorithm dominates another in both TPR and TNR or in both PPV and NPV.

The trade-off between TPR and TNR, and between PPV and NPV is presented visually in Figure 3. Part A of this graph presents the mean value of TPR on the y-axis and the mean value of NPV on the x-axis. While part B of the graph presents the mean value of PPV on the y axis and mean value of NPV on the x-axis. Setting LSV apart, we can clearly see that, as the TPR of an algorithm improves, its TNR worsens. Similarly, as it becomes superior in PPV, its NPV worsens.

Figure 3: Graphical representation of the trade-offs between performance metrics

Sources: Author's calculations from the QLFS, PALMS, and the official QFLS panel.

These trade-offs are caused by the effects of varying the restrictiveness of an algorithm (Madrian & Lefgren, 1999). If an algorithm is strict, it will reject more false matches. However, at the same time, an excessively restrictive matching algorithm leaves little room for measurement errors which inevitably exist in survey data, and therefore lead to more rejection of true matches. For instance, by requiring a constant age for all matched individuals between $t-1$ and t, an algorithm not only rejects false matches who report contradictory age between $t-1$ and t, but also rejects true matches who mistakenly recalls wrong ages. Therefore, as an algorithm becomes more restrictive, the increased rejection of false matches increases its TNR and PPV, while the increased rejection of true matches will decrease its TPR and NPV. This is clearly illustrated by the fact that RD-strict (the most restrictive algorithm of all) has the highest TNR/PPV and the lowest TPR/NPV when compared to RD-expanded and AKK algorithms.

5.5. Choosing an optimal QLFS matching algorithm

Given the results above, how do we best judge which algorithm works optimally for our study? An objective evaluation of the metrics reveals the apparent inferiority of the LSV algorithm. Not only does it have a lower match rate, but it has also generated lower values of TPR, TNR, PPV and NPV when compared to other algorithms (although we cannot comment on their statistical significance). In addition to this, the quality of the LSV panel is not consistent – school-attending youths are included in some panels but are systematically excluded from others. Given these findings, we are persuaded to reject the LSV algorithm from our options.

Amongst the remaining algorithms, it is much harder to pick the "better" algorithm based solely on performance metrics. When an algorithm performs better in some aspects, it tends to worsen in others. This trade-off compels us to rely on our subjective judgements in addition to inspecting metric values. Following a subjective evaluation of the PERSONNR variable in the QLFS data, we have decided not to use the AKK algorithm. This is despite the AKK algorithm generating some superior metrics. We find it difficult to justify including PERSONNR in the QLFS matching algorithm because, after all, PERSONNR is not designed to be intertemporally consistent (StatsSA, 2006) and its quality may well deteriorate at some point in the future just like it had improved suddenly in 2010.

We are therefore left with RD-expanded and RD-strict algorithms, both originating in Ranchhod and Dinkelman (2008). Between these two, we are inclined to choose the RD-strict algorithm. Although the RD-expanded panel data is larger in sample size, it contains individuals who have missing data on education and marital status, as well as people who have unbelievable inter-quarter changes in education and marital status. While many of them represent true matches, these records are unhelpful for our data analysis which relies on the correct measurement of the aforementioned variables – as we hope to measure worker flows for people with different education, age and marital status. For this reason, we consider the RD-strict algorithm to be an optimal algorithm for our study.

6. Data Preparation

In the section above, we devoted a substantial amount of effort in, firstly, reviewing and comparing different matching algorithms for QLFS and LFS used by researchers in previous studies and, secondly, quantifying and contrasting the matching quality of these algorithms. Thanks to these efforts, we are confident in our choice of RD-strict algorithm, originating in Ranchhod and Dinkelman (2008), as the foundation of the matching algorithm in the making of our panel sample. We now proceed to the next stage of this study – data preparation. This involves constructing the panel sample using the RD-strict algorithm and adjusting the sample weights for panel attrition.

6.1. Panel sample construction

Guided by the findings from the previous section, we use the RD-strict matching algorithm to match consecutive quarters of QLFS samples from 2008q1 to 2019q2. To reiterate, individuals are matched from one quarter to the next using UQNR, gender, race and age, with additional consistency checks on years of education and marital status to minimise false matches[13]. Because we are interested in the worker flows for the entire working age population of South Africa, individuals in our sample are aged 15-64 inclusive in both quarters of the quarter-to-quarter panel.

To further calibrate our panel sample to suit the purpose of this study, we imposed additional matching criteria ensuring that there are no inconsistent increases in educational attainment across waves. We deleted matched individuals who experienced more than one increment in their educational category. Five education categories are defined: in ascending order, they are "no education" (0 years of education), "primary education" (1-7 years of education), "incomplete secondary education" (8-11 years of education), "matric" (12 years of education) and "tertiary education" (> 12 years of education).

Following these steps, the average match rate of the panel sample is 51,66%, ranging from 40,54% in panel 24 (linking 2008q1-2008q2) to 54,96% in panel 66 (linking 2018q3-2018q4). On average, there

[13] We do not check if an individual's ability to write or read decreases from one quarter to the next, nor do we check if someone recently moved into a dwelling across consecutive quarters. Because the information necessary for conducting these checks is not available in the QLFS.

are 25743 individuals in a panel, ranging from 21768 in panel 52 (linking 2015q1-2015q2) to 29046 in panel 46 (linking 2013q3-2013q4).

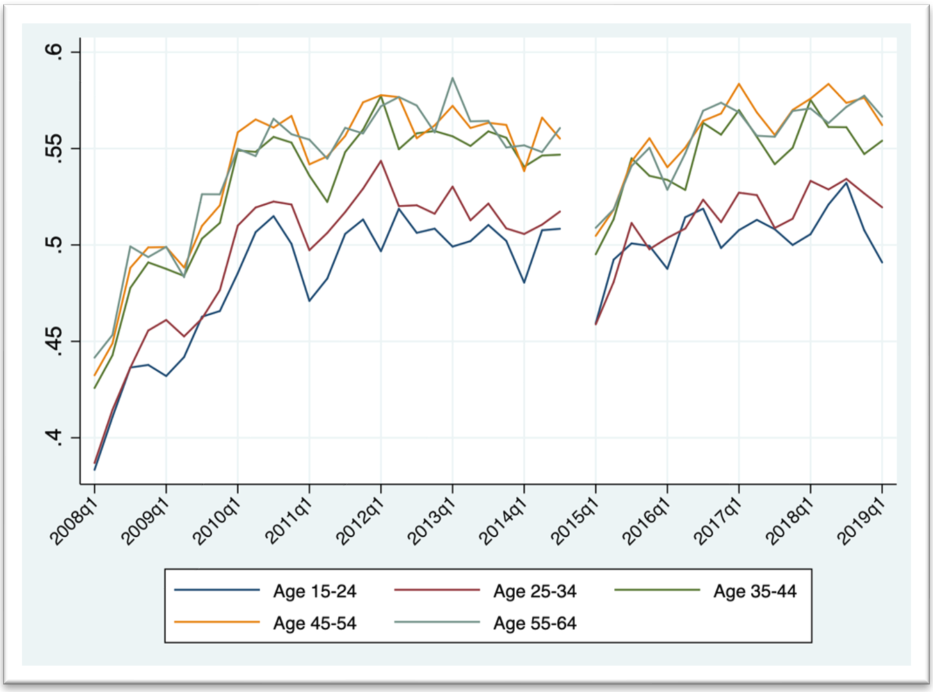
6.2. Non-random panel attrition

Like any other longitudinal data, attrition is of concern for our panel sample. In our case, attrition occurs when someone, who could have been matched between two consecutive waves, was not matched. This could occur for several reasons: household non-response, individual non-response, migration, or death. But algorithm-specific factors could also play a role. Certain variables were used by the matching algorithm as inputs. If these variables are miscoded or were mistakenly recalled, then the algorithm could incorrectly reject records representing the same individual. Therefore, the extent of attrition in the matched panels is exacerbated by measurement and recall errors with regard to the demographic variables used as inputs of the matching algorithm.

Attrition is an issue because it impedes our ability to infer useful statistics about the relevant population. In this paper, we are interested in studying the extent of worker flows for the working-age population of South Africa. If the attrited individuals are a non-random subset of the population, and if the estimates of worker flow are derived from the matched individuals using unadjusted weights, our estimates will not be representative of the population – they will be biased. (Lohr, 2019:331-332).

Following Schoiswohl and Wuger (2016:12), we can examine the non-random pattern of attrition by comparing the match rates for population subgroups with different observable characteristics. For instance, we could contrast the match rate for men against that of women. If attrition is completely random, it should affect both genders equally. Therefore, men and women should have similar match rates. On the other hand, if match rates are visibly different by gender, then attrition must be non-random. This could happen either because of a direct correlation between attrition and gender (e.g., men are more inclined than women to move away for work, leading to a higher attrition rate), or due to an indirect correlation through a third factor which we may or may not observe (e.g., men might have more disposable income than women, and people with higher disposable income are more likely to be absent from home, and therefore the higher non-response rate amongst men leads to a higher attrition rate). Figure 4a-4f below contrasts the quarterly match rates by age, gender, race, marital status, education and employment status.

Figure 4a: Match rates by population subgroups (age groups), 2018q1 – 2019q2



Notes: (1) y axis represents match rates and x axis represents quarter (2) The break in the line between 2014 and 2015 is due to non-matching as a result of the master sample revision. ***Sources***: Author's calculations from the QLFS and PALMS data.

Figure 4b: Match rates by population subgroups (gender), 2018q1 – 2019q2

Notes & Sources: Same as above

Figure 4c: Match rates by population subgroups (racial groups), 2018q1 – 2019q2

Notes & Sources: Same as above.

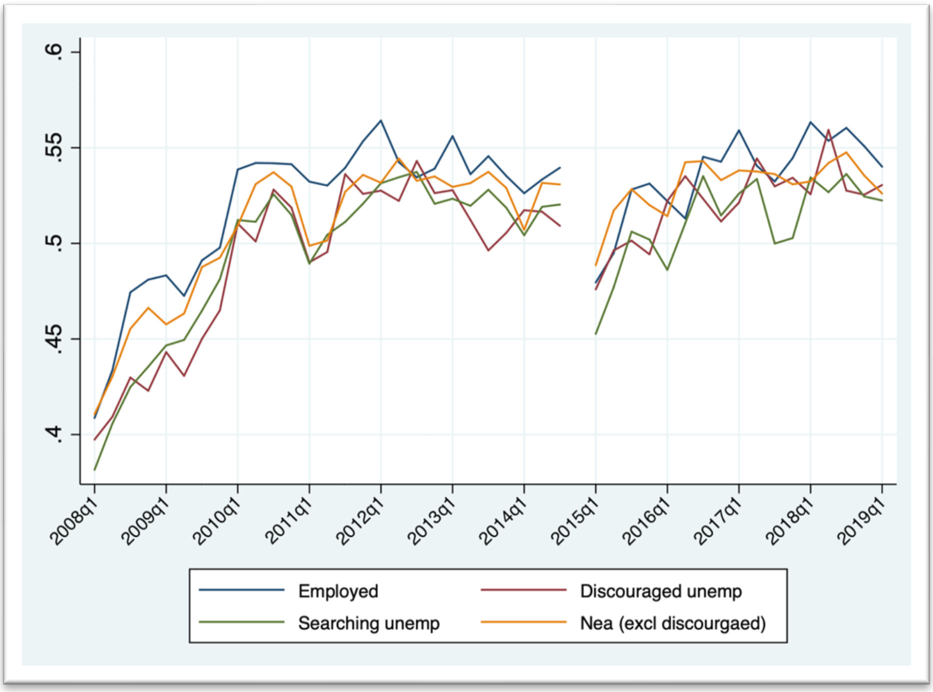
Figure 4d: Match rates by population subgroups (Marital status), 2018q1 – 2019q2

Notes & Sources: Same as above.

Figure 4e: Match rates by population subgroups (education), 2018q1 – 2019q2

Notes & Sources: Same as above.

Figure 4f: Match rates by population subgroups (employment status), 2018q1 – 2019q2



Notes & Sources: Same as above.

A visual inspection of these graphs reveals a non-random pattern of panel attrition. The algorithm is more likely to match older cohorts, female, Indian or white, people who are or were married, widows or widowers, and those with no formal education. In terms of labour market status, of people aged 15 to 64, the matching algorithm appears more capable of matching people who are employed. Most of these patterns are what one would expect. Match rates are higher for people who are less likely to move away, such as the older, the employed, and those who are married.

The differing match rates by observable characteristics, as displayed in Figure 4a-4f, suggest that the panel attrition in the sample is non-random. If we completely ignore the individuals whom we failed to match, and estimate worker flows using the matched subsample and the unadjusted weights, our estimators will not be representative of the working-age population of South Africa. Therefore, it is necessary to correct weights for attrition bias.

6.3. Adjusting cross-sectional weights for attrition

To the extent that attrition is only correlated with one's observable characteristics, we can correct attrition-induced biases by adjusting the cross-sectional weights for the matched subsample with inverse probability weighting (IPW) techniques. In every panel t, we take a matched individual's cross-sectional weight at wave $t-1$ and multiply it with the inverse of probability of her being matched between wave $t-1$ and t. The resultant weight, which we call panel weight, is applied to the matched subsample for deriving estimates that are representative of the population. The resultant estimates are unbiased as long as the matching probabilities are accurately obtained (Lohr, 2019:340).

We estimate the conditional probability of being matched between $t-1$ and t using probit models. The dependent variable is defined for every individual aged 15-64 at $t-1$. It is equal to 1 if she was successfully matched between $t-1$ to t, but it is 0 otherwise. The following independent variables are chosen due to their plausible correlation with the probability of matching: age intervals, gender, self-declared race, marital status, education, labour market status, province, geographic location (urban or rural), and household structure variables (household size and household composition). Also included as an explanatory variable is a dummy showing if someone is responding to the survey in person. Proxy interviews are administered on behalf of someone who is absent from the interview, but they are more likely to record the wrong information for the person concerned. This leads to potential measurement errors which could reduce one's matching probability.

We suspect that the effects of the chosen independent variables on matching probabilities vary across time. Therefore, we run a separate probit regression for each panel. As a result, a total of 44 regressions were run.

When we run the probit regression to predict the matching probabilities between $t-1$ at t, we set the regression sample to the full cross-sectional sample at wave $t-1$. In other words, a matching probability between $t-1$ and t was predicted for every individual present in the QLFS cross-section at wave $t-1$, whether or not she resided in a dwelling that would be re-interviewed in the next wave t. This is different from Ranchhod and Dinkelman (2008), as their IPW regression sample comprised of individuals at wave $t-1$ who lived in dwelling units that would be successfully surveyed in the next wave t. Our method represents a superior weight adjustment methodology. The sample restrictions in the IPW regressions in Ranchhod and Dinkelman (2008) would have excluded those who were successfully surveyed in wave $t-1$ but lived in households that refused to answer the survey in wave t despite being re-interviewed. Because household non-response is not random, excluding people from the non-responding households leads to selection bias. Consequently, their IPW regressions predicted the matching probabilities conditional on both independent variables **and** wave t household non-response, instead of **just** matching probabilities conditional on independent variables. It is the latter conditional probability that we need in order to render the panel sample representative of the population and, to obtain that, the IPW regressions should be run on a full cross-sectional sample.

Regression results are not shown in full due to space constraints. For illustration purposes, however, we report in Appendix A the average partial effects and the statistical significance of some independent variables from the first and the last probit regression – the first regression predicts the conditional probability of being matched between QLFS 2008q1 to 2008q2, while the last regression predicts the conditional matching probability between QLFS 2019q1 and 2019q2.

Both regressions suggest that, even after controlling for relevant observable characteristics, the older cohorts, those who have no education, live in smaller households, and were never married are significantly more likely to be matched to the next quarter than the younger cohort, those who have matric education, live in bigger households and have been divorced. Aligning with our expectation, whether someone is responding in person significantly affects her chance of being matched – in 2008q1 or 2019q1, a self-respondent is 2,34 or 1,24 percentage points more likely to be matched to the next quarter relative to someone who did not respond in person. People who are employed in the household sector were significantly less likely to be matched across quarters than the economically inactive.

Probit regression results are used to predict the conditional matching probability for the matched individuals, the inverse of which is multiplied by their cross-sectional weights to generate the panel weights. Following StatsSA (2018:100) and Essers (2017:197), for each panel, if a panel weight value is lower than the 1st percentile value, we set it to the 1st percentile value; if a panel weight value is larger than the 99th percentile value, we set it to the 99th percentile value.

It is important to note that, while panel weights correct biases induced by non-random attritions based on observable characteristics, it does not address the issue of non-random attrition if it is correlated with unobserved characteristics. Perhaps most worrisome is the unobserved worker flows between $t-1$ and $t.$ As is explained in Ranchhod and Dinkelman (2008:7), if we assume that people who experienced worker flows between $t-1$ and t are more likely to alter their residential locations, thereby less likely to be included in our panel, and if we have not corrected for this using the IPW technique, then our estimated worker flows will likely *underestim*ate the true extent of worker flows for the working age population.

7. Measurement of worker flows

With a set of nationally representative panel data, we are one step closer to examining worker flows in the context of South Africa. In this section, the procedure undertaken to accurately measure worker flows is outlined. In developing this procedure, we are closely guided by the widely used definitions and concepts of worker flows discussed above in Section 2.3.

One concept that needs clarification from the outset is the treatment of short-term spells of employment/non-employment that both start and end within the period between $t-1$ and t in our QLFS panel data. To recall: these short-term spells are counted as part of worker flows if we use the turnover method, but they are excluded if we use the reallocation method. Failure to short-term spells leads to substantial underestimation of worker flows if there exist many such short-term spells. In our panel data, there is a short period between every two consecutive waves – approximately three months. As such, we believe there are not many short-term spells, and the omission of them would not result in a substantial underestimation of worker flows. For this reason, we adopt the reallocation method.

The concept of "workers" also needs some clarification. In this paper, workers refer to "wage workers" – this definition excludes the self-employed, unpaid workers in household businesses, and employers. This definition is not only consistent with the theoretical framework of the matching model where worker flows refer to the reallocation of wage-earning labour across wage-paying firms, but it is also the definition adopted in prior literature (e.g., see Donovan, Lu and Schoellman 2020).

7.1. Measurement methodology

7.1.1. Gross worker flows and gross worker flow rates

Between every two consecutive quarters $t-1$ and t in the data, the quarterly gross worker flow (WF_t) is equal to the sum of hires (H_t) and separations (S_t) of wage workers. We estimate H_t as the weighted sum of working-age population who were newly hired into a wage job at time t, which includes those who were unemployed or economically inactive at $t-1$ but were working at a wage-paying job at t, non-wage workers at $t-1$ who became wage worker at t, and those who were continuously employed in the wage sector but switched employers between $t-1$ and t. Similarly, S_t is estimated as the weighted sum of wage workers at $t-1$ who had become, at time t, unemployed, economical inactive, employed in a non-wage job, or switched to new wage job. Then, following the convention, we estimate the quarterly gross worker flow rate (WFR_t) by dividing WF_t by the average weighted wage employment stock between $t-1$ and t (N_t). In mathematical formats, we can express WF_t and WFR_t as follows:

$$WF_t = H_t + S_t; WFR_t = \frac{WF_t}{N_t}$$

7.1.2. Worker flows and worker flow rates by population sub-group

We also estimate worker flows and worker flow rates for various population sub-groups defined by education, race, gender, etc. For instance, to estimate worker flows for men between $t-1$ and t, we calculate the weighted number of men who were hired (H_t^{male}), the weighted number of men who were separated (S_t^{male}), and the average weighted number of men who were in wage employment (N_t^{male}) between $t-1$ and t. Accordingly, worker flows for men (WF_t^{male}) and the worker flow rates for males (WFR_t^{male}) are estimated as follows:

$$WF_t^{male} = H_t^{male} + S_t^{male}; WFR_t^{male} = \frac{WF_t^{male}}{N_t^{male}}$$

Worker flows and worker flow rates by other population sub-groups are estimated in a similar way.

7.1.3. Worker flows and worker flow rates by firm characteristics

Similarly, we can measure worker flows and worker flow rates by firm characteristics. For instance, to measure the extent of worker flows for workers in the formal sector between $t-1$ and t, we estimate the weighted number of working-age population who were hired into formal sector (H_t^{formal}), the weighted number of working-age population who were separated from the formal sector (S_t^{formal}), and the average stock of the formal wage employment (N_t^{formal}). Then, formal-sector worker flows (WF_t^{formal}) and formal-sector worker flow rates (WFR_t^{formal}) are calculated as follows:

$$WF_t^{formal} = H_t^{formal} + S_t^{formal}; WFR_t^{formal} = \frac{WF_t^{formal}}{N_t^{formal}}$$

Worker flows and worker flow rates by other firm characteristics are estimated in a similar way.

7.2. Estimating employer-to-employer flows

The QLFS contains detailed questions that enable us to determine the employment status for all working-age individuals, and whether someone is a wage worker. This information enables us to identify transitions between unemployment/economic inactivity and wage employment, as well as between wage sector employment and non-wage sector employment. However, to actualise the measurement of worker flows, we also need to identify employer-to-employer flows (EE flows). Three different methods are utilised to estimate EE flows in this paper – the tenure method, the method whereby inter-quarter changes in employers are observed, and the method whereby inter-quarter changes in sector formality are observed. These methods are described below and the motivations for adopting all three methods are provided.

7.2.1. EE flows using the tenure method

In previous literature, EE flows are measured using information on the job tenure of workers. Job tenure is the length of time over which an employed person has been with the current employer. In the QLFS data, job tenure could be accurately estimated if we knew the time interval between the commencement date of someone's current wage job and the date on which this person is interviewed by the survey fieldworker (StatsSA, 2018:56). Then, assuming that two consecutive QLFS interviews on a matched individual are conducted exactly 3 months apart, for someone employed in the wage sector in both $t-1$ and t, having a tenure shorter than 3 months in wave t would imply that she experienced an EE flow (i.e., changed her job) at some point between $t-1$ and t.

Nevertheless, the data limitation of the QLFS and its survey design prevent us from carrying out the aforementioned "tenure method". First, the period between two consecutive interviews for the matched individual in QLFS may not be 3 months. The QLFS data is collected evenly throughout the entire quarter, with approximately 10000 dwelling units interviewed in each of the three months of the quarter (StatsSA, 2008:11). Therefore, the time interval between two consecutive QLFS surveys for a matched individual ranges from 1 month to 5 months. Second, it is impossible to know the month of the interview for a surveyed individual in the QLFS data. While such information is collected by Statistics SA, it is not released to the public. As such, it is impossible to compare the starting date of someone's job with the date of his/her survey interview. Therefore, job tenure cannot be accurately determined for individuals in our data.

While it is impossible to use the conventional "tenure method", a "stricter" version of the "tenure method" can still be adopted to help us determine the occurrence of some EE flows in the QLFS data. Despite not knowing the exact month of the survey interview for an individual, there is information on the commencement month of someone's current wage job. If a wage worker reports that her current job commenced in the current quarter, and she was also employed in the wage sector in the previous quarter, then she must have experienced an EE flow. This is because, although we do not know the date of her previous survey interview, we are sure that the interview occurred before the current quarter. Therefore, we are certain that the commencement month for her most recent job must be subsequent to the month of her last interview [14], which implies that she changed her job after the last interview – thus, she must have experienced an EE flow between the last interview at $t-1$ and the current interview at t. This is the first method by which EE flows are calculated in this paper.

7.2.2. EE flows as inter-quarter changes in employer characteristics

In addition to EE flows derived based on the "stricter" version of the "tenure method", we could also infer EE flows by studying the inter-quarter changes in employer characteristics. Some changes in employer characteristics indicate changes in employers. QLFS data allows us to ascertain if someone works for a public or private institution and the job's industry grouping. To the extent that continuing firms rarely change their ownership type and broad industry groupings (we use the one-digit industry code, $jobindcode$, in PALMS), these employer-level changes should imply changes in employers, hence EE flows. This is the second type of EE flows calculated for this study.

7.2.3. EE flows as inter-quarter changes in sector formality

Finally, we count the flows between the formal sector and informal sector as part of the EE flows. This is the third type of EE flows calculated in this study. We consider sector formality changes of continuously employed wage workers as EE flows for two different reasons. The first is a technical reason. Inter-quarter changes in sector formality are common in the sample data. Even if they are not counted as EE flows in the estimation of gross worker flows, we must still count them as hires or separations in the estimation of formal-sector or informal-sector worker flows. For instance, an inter-quarter change from the informal sector to the formal sector of an employed individual represents a separation in the informal sector and a hire in the formal sector. But, if inter-quarter changes in sector formality are only accounted for in the estimation of formal-sector and informal-sector worker flows but not in the estimation of gross worker flows, the gross worker flows estimate will end up being smaller than both informal-sector and formal-sector worker flows estimates. This cannot be true because, in theory, gross worker flows are equal to the weighted average value of formal-sector and

[14] For instance, suppose someone is continuously employed in a wage job in both 2018q1 and 2018q2, but reported in 2018q2 that her job commenced in the current quarter – i.e., in April, May, or June. Although we do not know when her data was collected in the first quarter of 2018, we are sure that it was collected before April. Therefore, the commencement date for her most recent job must be subsequent to the date of her last interview, hence an EE flow.

informal-sector worker flows. The only way to fix this is to measure inter-quarter changes in sector formality as EE flows in the estimation of gross worker flows.

Although this is not explicitly stated, formality changes of firms are counted as job flows in Kerr, Wittenberg & Arrow (2014), and they are considered to result in worker flows in Kerr (2018). The former study utilises firm-level panel data based on the Quarterly Employment Statistics survey from StatsSA, which excludes all unregistered enterprises. Thus, if an enterprise in the survey cancels its registration status in the next period to become a part of the informal sector, it would not be included in the survey in the next period – the authors would have considered this as an "enterprise death", hence "job destruction". The latter study utilises annual panel data based on employer-employee matches derived from the SARS database for income tax certificates. Income tax certificates are only issued by registered, thus formal-sector, firms. Therefore, if an enterprise becomes informal and stops issuing income tax certificates to its employees, none of its employees would appear in the next wave of the SARS panel data. The author would have counted its employees as being "separated". As seen, formality changes of a firm is implicitly considered to lead to both job flows and worker flows in the two most pertinent previous studies on worker flows in the context of South Africa. For this reason, we believe it is reasonable to include changes in sector formality for wage workers as part of the EE flows in this paper.

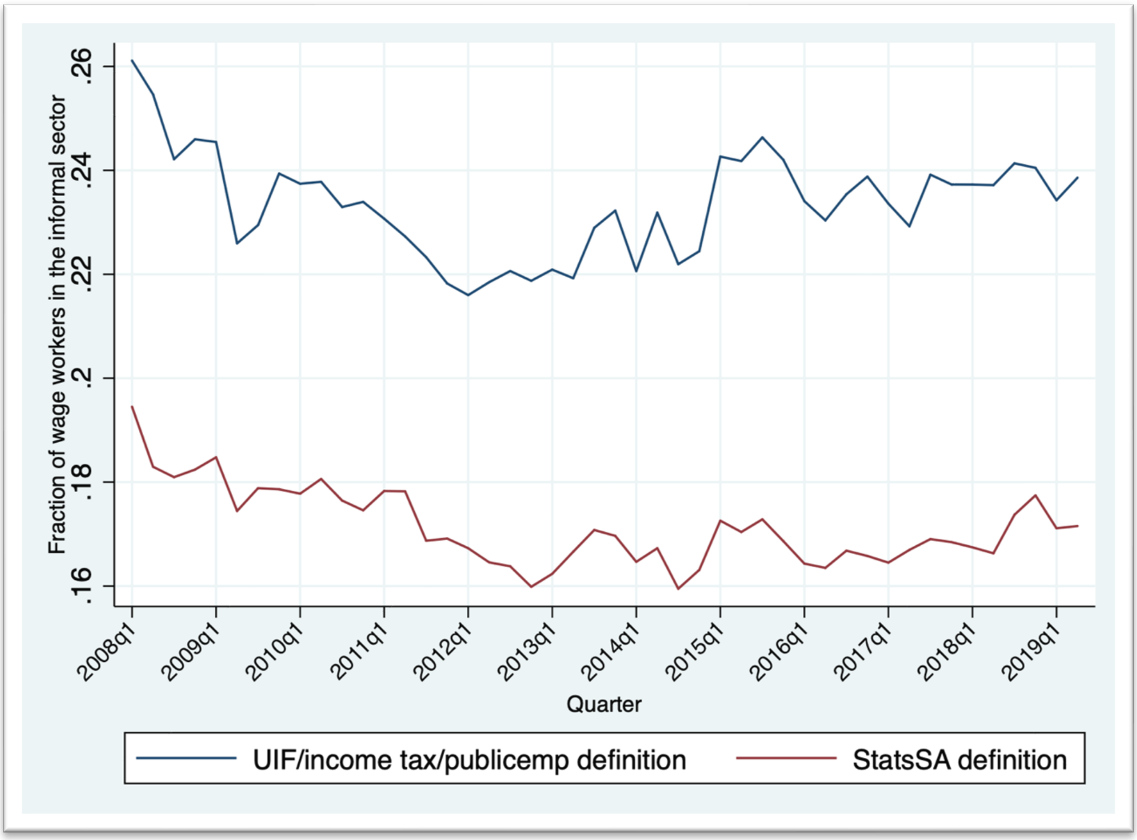
According to the definition from StatsSA, wage workers belong to the informal sector if they work for establishments that employ <5 employees **and** if income tax is not deducted from their wages (StatsSA, 2008:2). But, as Kerr and Wittenberg (2019:3) noted, this definition may be flawed. Suppose someone is *registered* for income tax and also works for enterprises that employ fewer than 5 people. Their tax registration status implies that the firm is in the formal sector, but if they earn below the tax threshold (which is way above the median income in South Africa), they would be erroneously classified as part of the informal sector. Given its inconsistency, we do not use the StatsSA sector informality variable.

Following the advice of Kerr and Wittenberg (2019:2), we identify sector informality using a variable on whether a worker's firm deducted Unemployment Insurance Fund (UIF) contributions from a worker's earnings. In particular, anyone who pays UIF contributions is part of the formal sector. But we impose two additional checks on the informality variable based on UIF because we believe that one's reported UIF status could suffer from recall or measurement errors. Firstly, we count someone as part of the formal sector if she reported paying income tax, irrespective of her UIF contribution status. Secondly, if someone reports not paying UIF contribution, nor income tax, but is employed in the public sector, then she is also included in the formal sector. This is because public servants in South

Africa are excluded from contributing to UIF, but they are, undoubtedly, part of the formal sector. Only when none of the three aforementioned conditions is valid do we consider someone part of the informal sector. If the information on these three conditions is missing, then a wage worker's sector formality is set to missing. However, such cases do not occur and, therefore, all wage workers in the sample have non-missing values for sector formality.

Using the new definition, the average share of wage employment belonging to the informal sector from 2008q1 to 2019q2 is estimated at 23,35%, which is 6,21% points larger than if we used the sector informality variable derived by StatsSA (see Figure 5 for the time series plots of the informal sector size by definition). The increase in the estimated size of the informal sector under the new informality variable is reasonable, in our opinion, particularly because we do not blanketly categorise all enterprises with >=5 employees as part of the formal sector. Furthermore, the trends underlying the two graphs closely follow each other, which is an assurance that the new informality variable adequately captures the temporal trend associated with the StatsSA informality variable.

Figure 5: Informal sector size by sector formality definition, 2008q1 – 2019q2



Notes: Estimates are weighted using ceweight1. ***Sources*****:** Author's calculations from QLFS and PALMS data.

To summarise the crucial points for this section. First, we include as part of the EE flows anyone who is continuously employed in a wage job but reports to have obtained her current job in the current survey quarter. Second, EE flows also include people who remained in wage employment but reported an inter-quarter change in firm ownership and broad industry groupings. Third, inter-quarter changes in wage employees' sector formality are also considered EE flows. Importantly, the variable of sector formality is based on UIF registration, income tax deduction, and employer type (private or public).

7.3. Measurement issues

The credibility of our results hinges on there being minimal measurement errors in the data, especially with respect to the variables mentioned in the foregoing discussion, which are instrumental to the measurement of worker flows. Therefore, the legitimate concern that measurement errors in the data could lead to spurious worker flows is something worth tackling in future research. Although this issue is not directly addressed in the present research, which only aims at laying a solid groundwork for worker flow analysis in South Africa, we have provided sufficient robustness checks on our worker flow estimates based on a comparison of our results with those in Kerr (2018).

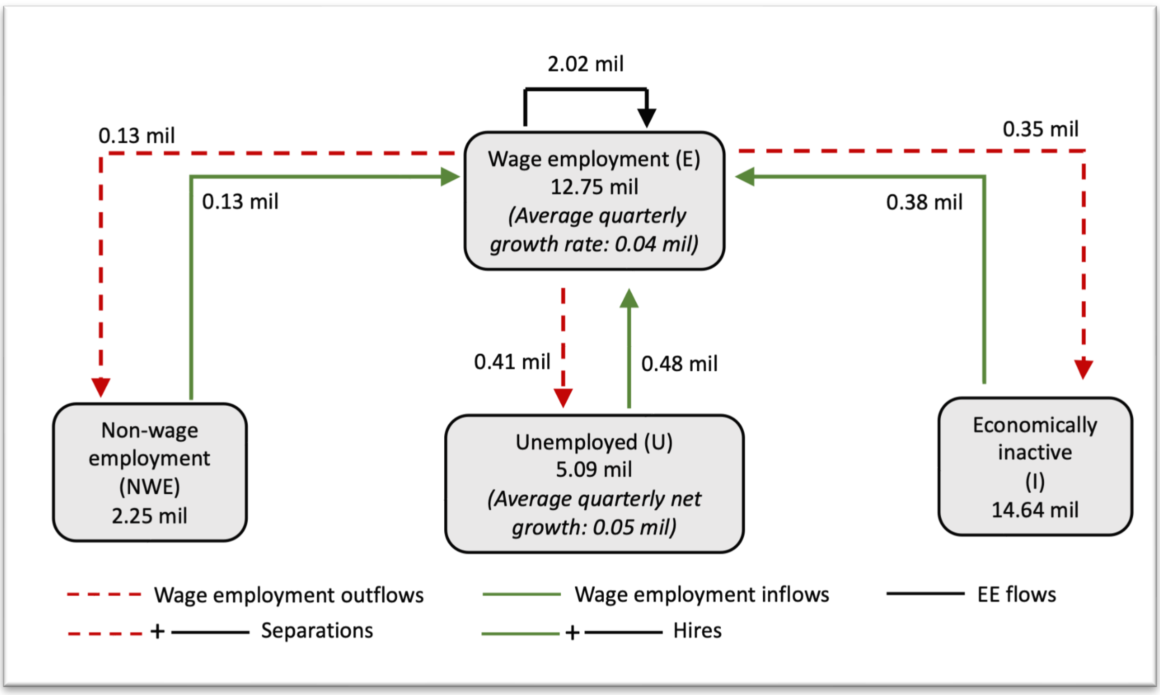
8. Extent of worker flows

8.1. Worker flows and stocks

The existing South African literature finds that substantial flows hide underneath labour market stocks, but few scholars have focused on worker flows – the sum of hires and separations. In light of this, we fill the research gap by demonstrating the sharp contrast between the vibrancy of worker flows in South Africa and the relative dormancy of the labour market as portrayed by the slow growth of wage employment stocks.

We do this by presenting Figure 6 below, which shows the average stock size of different labour market states (employment, non-wage employment, unemployment, and economic inactivity) as well as the size of quarterly flows between different states in terms of millions of persons from 2008q1 to 2019q2. The red dashed lines represent wage employment outflows, and the green solid lines represent wage employment inflows. The EE flows, which count both as outflows and inflows of wage employment, are represented as a black solid line.

Figure 6: Average quarterly worker flows and stocks, 2008q1 – 2019q2



Notes: All estimates are weighted using panel weights. ***Sources***: Author's calculations from QLFS and PALMS data.

As is shown in this flow chart, the average quarterly net increase in wage employment stock was merely 40000. The anaemic growth of wage employment stock, however, belies substantial worker flows. On average, 2,91 million[15] wage-employed workers were separated. Simultaneously, the group of individuals who are hired in an average quarter is of comparable size, 3.01 million[16]. Combining hires and separations, as many as 4 million unique individuals were hired and separated in an average quarter, and the average quarterly gross worker flows amounted to 6 million. These figures are considerable in comparison with the quarterly net growth of wage employment stocks.

A majority of the separations (69%) and hires (67%) are EE flows, which means that they switched employers over a quarter or managed to find a new wage job within a quarter. That more than half of the separations or hires are EE flows might concern those who suspect that the measurement errors in the data have inflated the estimated size of EE flows. While this concern is legitimate, we are reassured to some degree in light of the fact that EE flows often exceed more than half of annual hires or separations across all OECD countries (Cahuc, Carcillo & Zylberberg, 2014). Therefore, our results are consistent with the EE flow patterns found in more advanced economies.

8.2. Temporal trends and variability in worker flows and stocks

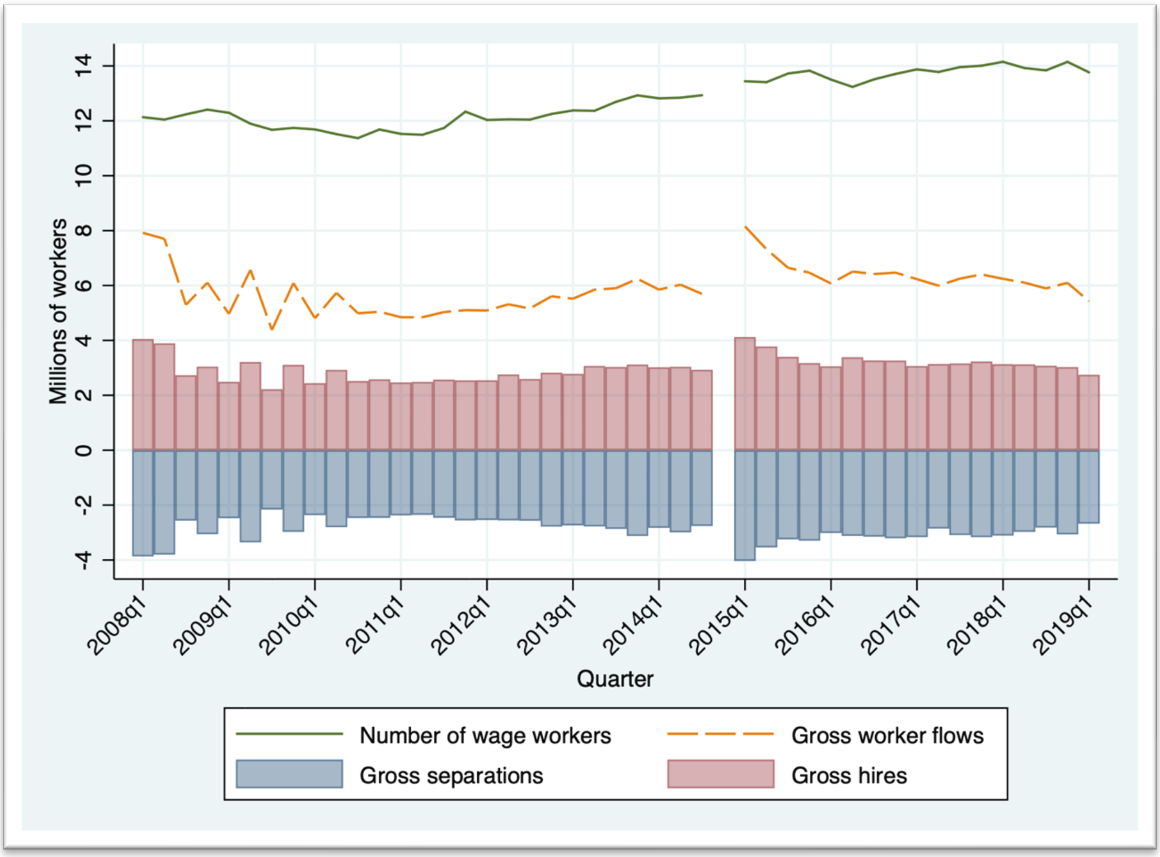
One major drawback of Figure 6 is that it takes the average across a prolonged period spanning over 10 years. This could conceal any temporal variabilities in gross worker flows that may be of interest to researchers. We address this issue in Figure 7, where we plot the time series for the size of wage employment stock, and the magnitude of gross hires, separations, and worker flows. Gross separations are plotted below the x-axis as negative values as a visual contrast with gross hires.

Firstly, we observe that the South African labour market exhibits cyclicality. From 2008 to 2011, the size of wage employment stock either stagnated or trended downwards. This adverse labour market condition might reflect the economic downturns in South Africa in the aftermath of 2008 global financial. Since 2011, the picture turned slightly more positive as there has been a sustained increase in the wage employment stock. However, this lasted until 2017, after which year the growth of wage employment stock evened out. We examine the cyclicality of worker flows in more depth later.

[15] This is the sum of all wage employment outflows and EE flows: 0.13 + 0.41 + 0.35 + 2.02
[16] This is the sum of all wage employment inflows and EE flows: 0.13 + 0.48 + 0.38 + 2.02

The temporal variability of gross worker flows is somewhat muted. Except for the beginning of 2008 and the year 2015, gross worker flows fluctuated around 6 million over the entire sampled period. Gross worker flows peaked at 8 million in the first quarter of 2008 and 2015. While the surge of gross worker flows in 2008 could imply an equilibrium readjustment process in the South African labour market as it transitioned from being more fluid before the global financial crisis to being less fluid thereafter, the sudden rise of gross worker flows in 2015 could simply be due to data incomparability resulted from changes in QLFS master sample. Recall that the QLFS master sample prior to 2015 was different from that after 2015. This could result in a temporal discontinuity in the value of variables derived from the dataset before and after 2015 (Kerr and Wittenberg, 2019).

Figure 7: Quarterly worker flows vs stocks, 2008q1 – 2019q2



Notes: All estimates are weighted using panel weights. **Sources:** Author's calculations from QLFS and PALMS data.

Lastly, we note that gross separations tend to match gross hires in size in all quarters: they look almost symmetrical across the x-axis in the graph. This not only reflects a well-known stylised fact about the strong and positive correlation between gross hirings and separations (OECD, 2009), but it also suggests that the South African labour market may be settling into a dynamic equilibrium where two sides of the worker flows balance each other, leaving labour market headline indicators stationary. Indeed, the data suggests that, between 2008q1 and 2019q2, the quarterly official unemployment rate moved no more than 3% point above or below its ten-year average value of 25,4%, and the quarterly wage employment absorption rate moved no more than 2% point above or below its average value of 36,8%. The co-existence of a large and persistent unemployment rate and the balanced as well as substantial work flows negates the widely adopted narrative that interprets South African unemployment as a disequilibrium phenomenon in an otherwise perfectly competitive labour market, as Wittenberg (2002) noted.

8.3. The extent of worker flow rates and robustness check

Thus far, we have only presented gross worker flows in absolute terms. To gauge the extent of gross worker flows relative to the stock of wage employment and to aid cross-study comparisons of the extent of worker flows, we also calculated gross worker flow rates (WFR). We find that the average quarterly gross WFR in South Africa is 46,26%, as is shown in the last cell of Table 4 below. In other words, almost half of the employee-employer matches either form or break up every quarter over 2008 – 2019.

Table 4: Average quarterly worker flow rates by sector formality and time periods

	Formal sector	*Informal sector*	*All sectors*
Before master sample change (2008 – 2014)	36.99%	78.83%	46.18%
After master sample change (2015 – 2019)	37.87%	75.46%	46.39%
All times (2008 – 2019)	37.33%	77.53%	46.26%

Notes: (1) Author's own definition for sector formality is used – see a detailed explanation in section 6.1. (2) All estimates were weighted with panel weights. **Sources:** Author's calculations from QLFS and PALMS data.

Because this is the first time that QLFS has been used to study worker flows for South Africa, we want to determine if our estimates are reasonable. To do this, we compare our results with Kerr (2018), who measured gross WFR for the formal sector of South Africa using a methodology similar to ours. Because workers represented in the QLFS sample are from both formal and informal sectors, we have estimated WFR by sector formality.

Furthermore, as was suggested in an earlier discussion, the change of the QLFS master sample in the year 2015 might have led to an "artificial" increase in gross worker flows. Therefore, to test the robustness of the WFR estimates, we divided the sampled period into two sub-intervals – one before 2015 and the other one after 2015 – and estimated WFR for each time interval separately.

These procedures resulted in a series of WFR estimates, relevant to different time periods and sector formality. These estimates are presented in Table 4. From the table, we can see that the WFR estimates are quite robust to changes in the master sample and time period. When we switch between QLFS samples before 2015 and those after 2015, the WFR estimates differed by less than 1% point amongst the formal sector workers and differed by only 3% points amongst those in informal sectors. In light of this, for simplicity, we will only utilise the WFR estimates based on the entire sampled period without making the explicit distinction between the pre-2015 sample and the post-2015 sample in ensuing discussions.

We note that the average quarterly formal-sector WFR is 37.3% from 2008 to 2019. In comparison, Kerr (2018) reported that the annual WFR for formal sector workers is on average 53% from 2011 to 2014. Unfortunately, a direct comparison of the two estimates is difficult for three reasons. First, our WFR estimates are quarterly whereas those reported by Kerr (2018) are yearly. Second, because worker flows are measured using the reallocation method, which omits short-term spells within the measurement period, we cannot simply multiply our quarterly WFR estimate to make them comparable with yearly estimates (Huber & Smeral, 2006). Lastly, because the sampling design of QLFS only allows researchers to match surveyed individuals over a maximum period of three-quarters of a year, it is impossible to construct yearly QLFS panel data.

Bearing in mind these limitations, we construct two additional sets of data: half-yearly and "three-quarterly" QLFS panels. These additional datasets will enable us to estimate the half-yearly and "three-quarterly" WFR. Then, by studying the increasing trend of WFR estimates with increased time intervals between the consecutive waves of the QLFS panel data, we can tell if WFR estimates based on the QLFS data are comparable with the yearly estimates in Kerr (2018).

Following the exact same procedures used in the construction of the quarterly QLFS data – which includes matching working-age individuals based on the algorithm outlined in Section 5.1, creating appropriate panel weights for the matched sample using the IPW technique specified in Section 5.3, and measuring worker flows using the method detailed in Section 6 – we constructed two additional sets of QLFS panel data: the first consists of working-age individuals matched across every two-quarters of the QLFS from 2008q1 to 2019q2, and the second consists of working-age individuals matched across every three-quarters of the QLFS over the same period.

Using these datasets, we estimated the average WFR for workers in the formal sector over the entire sample period as well as for the period from 2011 to 2014 – the latter is the period over which WFR estimates are made in Kerr (2018). We present in Table 5 the average WFR estimate based on quarterly, half-yearly, and "three-quarterly" QLFS panel data alongside the estimates in Kerr (2018).

Table 5: Robustness check for the WFR estimates based on QLFS data

	Quarterly	*Half-yearly*	*Every three quarters*	*Yearly*
Average formal-sector WFR over ***2008 – 2019*** *based on* ***QLFS***	37,33%	48,82%	55,27%	N/A
Average formal-sector WFR over ***2011 – 2014*** *based on* ***QLFS***	35,62%	46,58 %	53,46%	N/A
Average formal-sector WFR over ***2011 – 2014*** *based on* ***SARS income tax data*** *(Kerr 2018)*	N/A	N/A	N/A	53%

Sources: Estimates in the first two rows are based on author's calculations from QLFS and PALMS data

Over the period 2008 – 2019, the average WFR for formal-sector wage workers based on QLFS data increases gradually from 37,33% every quarter to 48,82% every half a year, and then to 55,27% every three quarters. These numbers will be slightly lower if we focus on the period between 2011 to 2014. However, judging from these trends, it is reasonable to infer that the yearly WFR estimate based on yearly QLFS panel data – had it been possible to create such data – should be around 60%. Although it might be slightly bigger than 53%[17], the yearly estimate for formal-sector WFR obtained by Kerr (2018), the difference does not appear large despite the large data differences between this study and Kerr (2018). Therefore, we believe that this is strong evidence that the WFR estimates based on the QLFS data are credible and that this study has managed to capture the true extent of gross worker flows in South Africa despite a myriad of data challenges and obstacles.

8.4. Cross-country comparisons of worker flow rates

Cross-country variations in gross WFR is often ascribed to, amongst others, variations in the stringency of worker protection (Bassanini & Garnero, 2013), and it is widely believed by South African business groups and economic-political commentators that the worker protection in South Africa restricts firm's ability to hire and fire. Therefore, one should expect gross WFR to be much lower in South Africa compared to other countries with more relaxed regulations for worker protections. We put this to test

[17] And this is what one would expect. For instance, the SARS income tax data may only consist of a subset of the formal-sector workers captured in the QLFS data, and it may also include more individuals with high earnings than the QLFS would. Factors such as these could result in lower WFR being estimated in the SARS income tax data in comparison with the QLFS data.

by comparing worker flows, a proxy for labour market fluidity (Nakamura et al., 2020), between South Africa and other countries.

Table 6 lists several gross WFR estimates from a diverse range of countries. Naturally, we focus on studies that used data similar to ours (data based on matched rotating panel survey of individuals), but we have also included studies utilising panel data based on matched employee-employer records. All studies included in the table have adopted the reallocation counting method, which we also used.

Table 6: Selected estimates for WFR outside of South Africa

Country	*Year*	*Quarterly WFR*	*Data*	*Representativeness*	*Source*
U.K.	1996 - 2010	12,40% (based on figure 1)	Individual-level survey data	Nationally representative	Gomes, 2012
Australia	1996 - 2002	18,79%	Matched employee-employer data	All registered firms except for those in agriculture and public services	Huber & Smeral, 2006
U.S.	2010	24.1%	Firm-level survey data	Non-farm private sector	Davis & Haltiwanger, 2014
U.S.	1985 - 1994	32,20%	Matched employee-employer data	Non-manufacturing sector in the State of Maryland, U.S.	Burgess, Lane & Stevens, 2000
Austria	2011-2014	41,2% (based on figure 6)	Individual-level survey data	Nationally representative	Schoiswohl & Wuger, 2016
Ethiopia	2011 - 2018	38%* (half-yearly)	Matched employee-employer data	Private firms	Shiferaw & Soderbom, 2021
OECD average	2000 - 2005	33%* (yearly)	Individual-level survey data	Nationally representative	OECD, 2009

We proceed with cross-country comparisons with caution, because estimates from different studies are not directly comparable due to differences in the underlying data and time periods. With this caveat in mind, we note that, with an average quarterly gross WFR of 46,26%, the South African labour market appears much more fluid than most developed countries. Average quarterly gross WFR is larger in South Africa than it is in the U.K. (12,4%), Australia (18,79%), and Austria (41,2%). Even though we do not have average quarterly WFR estimates for OECD countries, the fact that the yearly gross WFR (33%) in OECD is smaller than South Africa's quarterly gross WFR is sufficient evidence that labour market is more fluid in South Africa than an average OECD country. Furthermore, even though U.S. is generally considered to be a country with lax worker protection laws, and the most fluid labour market within the developed countries (Bellmann, Gerner & Upward, 2018), its quarterly gross WFR trails behind that of South Africa.

How does South Africa compare to a developing country in sub-Saharan Africa like Ethiopia? Shiferaw and Soderbom (2021) estimated half-yearly WFR for formal-sector private firms in Ethiopia, and found it to be 38% on average. This is close to our formal-sector WFR estimate for South Africa, which is 37,3%. But, because our WFR estimates are quarterly, we conclude that the formal sector of the South Africa is still more fluid than that of Ethiopia, despite the latter having much less restrictive worker protections (Shiferaw & Soderbom, 2021:8).

The sheer size of the quarterly gross WFR in South Africa is hardly compatible with the narrative that firms across the board have difficulty in hiring and firing. It is possible that certain sub-sections of the labour market enjoy better worker protection than the rest, and firms in these sub-sections are much more restricted in their human resources practices. Indeed, table 4 shows that formal sector workers, who are better protected by the labour regulation, enjoy a lower WFR than those in the informal sector. We shall shed more light on this in the upcoming heterogeneity analyses.

However, with the evidence presented thus far, it appears that the South African labour market is more fluid than what is commonly believed. This conclusion is consistent with Kerr (2018:152), who noted that worker flows “are a pervasive part of the way the labour market operates” in South Africa.

9. Cyclicality analysis

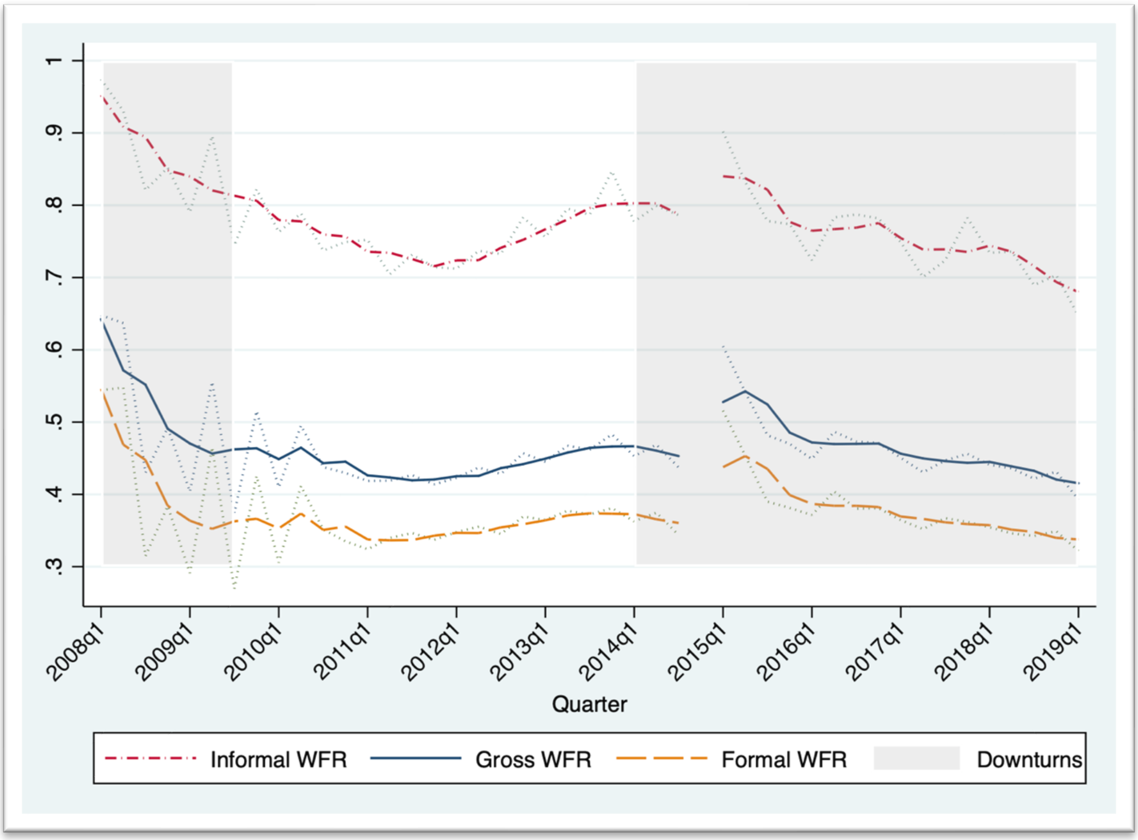
The South African Reserve Bank (SARB) tracks the cyclical fluctuation of the aggregate economy of South Africa around its long-term growth trend by regularly monitoring the movement of three composite business cycle indicators – leading indicators, coincident indicators, and lagging indicators – which are compiled from a diverse range of economic and financial variables. It then publishes, in the SARB Quarterly Bulletin report, a set of reference turning points that represent the beginning/end of an upward phase – when the aggregate economy grows on par with or faster than its long-term growth trend – and a downward phase – when the growth of the aggregate economy dips below its long-term trend. A complete chronology of these reference points is available from the year 1946. (Venter, 2011)

According to the SARB Quarterly Bulletin of June 2021, an economic downturn started in November 2007 as a result of the global financial crisis and lasted until September 2009. This was followed by an economic boom that endured from September 2009 to November 2013 and a subsequent economic slowdown that is ongoing at the time of writing (SARB, 2021). Therefore, stretching over ten years from 2008 to 2019, our data covers a complete business cycle. This affords us a unique opportunity to examine the cyclicality of worker flows, which are important to understand for both policymakers and trade unions in South Africa but have hitherto received little academic attention. To the extent that a larger extent of worker flows indicates increased labour market fluidity, which speeds up labour reallocation and economic efficiency, an examination of past time trends and cyclicality of WFR enables policymakers to better monitor fluctuations of labour market efficiency. On the other hand, the ensuing discussions should also be pertinent to trade unions and worker associations which are more concerned with improving the employment security for their members, because large WFR would imply more precarious employment.

9.1. Temporal and cyclical trends of worker flow rates

Figure 8 displays the temporal trends for quarterly WFR by sector formality. Following Gomes (2012), to remove seasonality, we smoothed the time series using four-quarter moving averages with uniform weights. This means that a smoothed WFR estimate at time t is calculated as the simple average of WFR estimates at $t-2$, $t-1$, t, and $t+1$. We experimented with shifting the "smoothing window" forwards and backwards by one or two periods to test if doing so would visibly alter the temporal patterns of the smoothed series. It does not. The smoothed series are highlighted in the graph and are juxtaposed on top of the unsmoothed time series. Finally, we used the grey share to indicate when the South African aggregate economy is in a downward phase.

Figure 8: Time trends of quarterly WFR, 2008 – 2019



Notes: (1) Time series are smoothed using 4-quarter moving averages (2 lagged, 1 current, and 1 forward observations) with uniform weights to remove seasonality. (2) Smoothed time series are juxtaposed on the line plots for the actual observations (3) All estimates are weighted using panel weights; **Sources**: Author's calculations from QLFS and PALMS. Recession dates were provided by SARB in its 2021 June Quarterly Bulletin report.

An inspection of the graph yields a few notable findings. First, the informal sector exhibits a visibly higher degree of time variability in WFR than the formal sector. This could indicate that the informal sectors react more sharply to underlying economic conditions, which is consistent with Kingdon and Knight (2009) who observed that informal sector wages are more responsive to variations in unemployment rate than wages set in the formal sector.

Second, WFR appears procyclical irrespectively of whether we focus on the formal sector, informal sector, or the gross WFR. This is indicated by a clear, downward trend of WFR during the economic recession in 2008 and 2009, an upward trend from 2011 to 2014 during the economic booms, and another downward trend during the economic downturn that is ongoing.

9.2. Measuring the cyclicality of worker flow rates

A visual inspection of the time trends above indicated that worker flow rates are procyclical, but the exact extent of the cyclicality remains to be investigated, which we do below. Guided by Gomes (2012), we measure the cyclicality of a variable, say Z, by running an ordinary least square regression, where the dependent variable is the log transformation of Z and the independent variables are the unemployment rate, quarter dummies, and a linear annual time trend. Accordingly, the cyclicality of Z is indicated by the estimated regression coefficient on the unemployment rate – if it is insignificant, then Z is acyclical; If it is significantly positive, then Z is counter-cyclical; If it is significantly negative, then Z is procyclical.

The strength of the regression method over a non-parametric method, in which case we would calculate the correlation coefficient between the variable of interest and a cyclical economic indicator such as the unemployment rate, is that the regression method not only reveals the nature of cyclicality for the variable of interest but also allows us to examine the statistical significance of its cyclicality, as well as allowing us to control for confounding time or seasonal trends.

To be consistent with the foregoing investigation, we distinguish between the formal sector and the informal sector. Furthermore, because worker flow rates are numerically smaller than 1, which would result in negative values if it was log-transformed, we use $ln\ (WFR + 1)$ instead of $ln\ (WFR)$ as the dependent variable in the relevant regression[18]. Following this, we run three regressions and report our results below. The cyclicality for gross WFR, formal WFR, and informal WFR are indicated in Table 7 as the size and the statistical significance of the estimated regression coefficient on the unemployment rate ($\hat{\beta}_{Urate}$) in the regression where the dependent variable is the $ln\,(1 + variable\ of\ interest)$. To conserve space, we do not report regression coefficients on other independent variables.

Our results suggest that, after controlling for seasonality and the time trend, only informal-sector WFR is procyclical. Formal-sector WFR is acyclical. Therefore, it is the labour reallocation process in the informal sector that drives up the procyclicality of gross WFR. The fact that worker flows in the formal sector appear irresponsive to cyclical trends echoes Verick (2012), who speculated that the hiring and

[18] This is not what Gomes (2012) does because he uses gross worker flow values, which are measured in terms of the actual number of individuals.

firing costs of the formal sector employment prevented the formal-sector employers from adjusting to the business cycles.

Table 7: Cyclical variation of worker flow rates using econometric approach

Full sample		*Formal sector*		*Informal sector*	
Variable of interest	$\hat{\beta}_{Urate}$	*Variable of interest*	$\hat{\beta}_{Urate}$	*Variable of interest*	$\hat{\beta}_{Urate}$
Gross WFR	-0.0161*	Formal WFR	-0.0143	Informal WFR	-0.0198**

Notes: (1) The natural logarithm of 1 + worker flow rate between $t - 1$ and t is run on the unemployment rate at $t - 1$, along with other independent variables: quarter dummies and an annual time trend. (2) Each of the regression has a total of 42 observations. (2) Standard errors are reported in parentheses:*** p<0.01,** p<0.05,* p<0.1. ***Sources***: Author's calculations from QLFS and PALMS data.

Because WFR is procyclical in the informal sector but acyclical in the formal sector, we suspect that the gross WFR is only weakly procyclical. We can confirm this by calculating the correlation coefficient between gross WFR and the unemployment rate, following Nakamura et al. (2020). This yields negative 0.31, which is closer to zero than it is to one, indicating the weakly procyclicality of gross WFR. This is consistent with the existing international literature (Davis and Haltiwanger, 1998).

10. Heterogeneity analysis

Thus far, we have predominantly focused on South African worker flows at an aggregate level, which we have shown to be substantial in size in comparison with other advanced economies. However, aggregate estimates of worker flows could conceal considerable heterogeneities at a more disaggregated level – be it industry, firm, or individual. Uncovering these heterogeneities is crucial for a more nuanced understanding of the labour market fluidity in South Africa because its high levels of gross worker flows do not necessarily imply that all workers have the same experience across industries and firms, or across population groups with different individual characteristics such as age, education, and population groups. For these reasons, we now proceed to investigate the heterogenous patterns of worker flows.

10.1. Worker flows by industry

We start our analysis by investigating variations in worker flows across industries. Studies have shown that the distribution of worker flows across different industries is similar across countries (OECD, 2009) – worker flows tend to be lower in goods-making industries and higher in service industries (Lane, Stevens & Burgess, 1996; Burgess, Lane & Stevens, 2000; Davis, Faberman & Haltiwanger, 2006).

We test whether this is true for South Africa by analysing industry-level variations in worker flows using the QLFS panel data. The industry variable in QLFS is based on the self-reporting of the employee. To differentiate across industries, we use the one-digit industry code in the data, which divides the South African economy into ten industries. Although this classification is too coarse for us to appreciate the nuances across smaller economic sub-sectors, we have decided to use it because a more refined industry classification would lead to the problem of small sample sizes here exits, rendering the industry-level estimates unreliable.

Table 8 presents the estimated WFR by industry classification, averaged over the period 2008q1 – 2019q2. What is clear is that there exists substantial variation in WFR across industries. Average WFR range from 22,63% in the “Mining and quarrying” to 79,04% in the “Construction”.

Table 8: Average quarterly worker flow rates by industry, 2008q1-2019q2

Industry classification	*WFR average*
Agriculture, forestry and fishing	60,64%
Mining and quarrying	22,63%
Manufacturing	45,42%
Utilities	52,60%
Construction	79,04%
Trade	49,11%
Transport	50,25%
Finance	43,55%
Services	33,67%
Domestic Services	51,01%
Gross WFR	46,26%

Note*:* All estimates are weighted using panel weights. ***Sources*****:** Author's calculations from QLFS and PALMS

The results appear at odds with the international literature, which finds that goods-making industries have smaller WFR than service industries. In contrast, our data show that, apart from "Mining and quarrying", most goods-making industries have relatively high WFR. This includes "Agriculture, forestry and fishing" (60,64%), "Utilities" (52,60%), and "Construction" (79,04%). The average WFR in "Manufacturing" (45,42%) is on par with the average gross WFR across all industries. On the other hand, South African workers in service industries experience relatively low levels of WFR, with the notable exceptions of workers in "Domestic services" and "Transport" which have high average WFR (both are above 50%). The average WFR in "Services" (33,67%) and "Finances" (43,55%) are both below the average gross WFR across all industries.

While our results are incompatible with evidence from developed countries such as the U.S., they are comparable with Kerr (2018), which is another South African study on worker flows. Similar to our results, Kerr (2018:155) finds that some of the lowest values of WFR are observed in the mining sectors while the construction and agricultural sectors have some of the highest WFR values, and certain service sectors, such as "Public Administration" and "Education", have very low WFR values.

The only exception is that he uncovers a relatively high WFR for the finance sector and a low WFR for the utilities sector, whereas Table 8 shows that the opposite is true. This difference may be caused by differences in the data utilised, as Kerr (2018:156) cautions that the industry codes in the South African income tax data do not fit the standard industrial classification system used by StatsSA and, furthermore, they are not as reliable as the QLFS industry codes.

10.2. Worker flows by firm and job characteristics

In addition to its industry-level heterogeneity, WFR could also vary across workers who are employed in different types of firms and jobs. Kerr (2018) examined the variation of WFR by firm size and found it to be lower in larger firms. He also attempted to differentiate between publicly-run firms and private firms in order to ascertain the WFR differences between the two sectors, but could not do so properly due to data constraints. Furthermore, there was no job-level information on employees in the SARS income tax data, which prevented the author from investigating the variability of WFR by job characteristics.

We build on Kerr (2018), as well as extend it, by analysing the variation of WFR across workers whose employment differs by the firm and job characteristics. In particular, we focus on firm size, the public/private sector, and two job characteristics – job contract type and duration. Table 9 presents the average estimated WFR by firm and job characteristics from 2008q1 to 2019q2.

Table 9: Average quarterly worker flow rates by firm/job characteristics, 2008q1-2019q2

Firm and job characteristics	*WFR average*
Firm size	
Small firms (0-9 employees)	57,23%
Medium firms (10-49 employees)	44,86%
Large firms (>50 employees)	37,08%
Public vs private	
Public sector	29,56%
Private sector	50,46%
Contract type	
Written contract	40,70%
Verbal contract	68,64%
Contract duration	
Permanent contract	31,96%
Fixed-term contract	77,40%
Gross WFR	46,26%

Notes: All estimates are weighted using panel weights. ***Sources:*** Author's calculations from QLFS and PALMS

As found in Kerr (2018), table 9 confirms that WFR declines with firm size. WFR is only 37% for wage workers who self-report to work for firms with 50 or more employees, but rises up to 57,23% for those reportedly working for firms with fewer than 10 employees.

Wage workers in the public sector have a drastically different experience in terms of worker flows in comparison with those in the private sector: public-sector WFR is as low as 30% whereas it is much higher, at 50%, in the private sector. Due to data limitations, Kerr (2018) could only identify public-sector employees by identifying people working in the "public administration" industry. But doing so would exclude any public-sector workers who are outside of the "public administration" industry but are employed by firms that are owned by the state, such as public-run hospitals, schools, or utility companies. We are able to address this data issue using a variable from QLFS data that is based on a direct question asking the respondent whether she/he works for the government, state-owned business, or private enterprises.

As a result of South Africa's strict worker protection regulations (Bhorat, Naidoo & Yu, 2014), or at least a perception thereof (Benjamin, 2014; Bertrand & Crepon, 2020), firms that hope to maintain their flexibility in terms of hiring and firing might opt for employment contracts of limited duration or verbal contracts, since these types of work arrangements are less protected by labour market regulation than permanent and written contracts. Therefore, unless all contract types are equally protected by the labour regulation, gross WFR will be "pushed up" by the higher WFR in the unregulated section of the labour market (Martin & Scarpetta, 2012). To what extent this is true can be examined by studying how WFR varies across workers who have different contract types and durations. We find that WFR is much lower amongst wage workers with written contracts (40,70%) than those with verbal contracts (68,64%), and much lower for permanent workers (31,96%) than those working for a limited duration (77,40%). Therefore, WFR for wage workers with verbal or fixed-term contracts is much higher than their counterparts holding written or permanent contracts. This appears to confirm Martin and Scarpetta's (2012) argument – employees that are less protected by the worker protection regulation will bear the main brunt of employment adjustment.

To address the concern that workers with "unprotected" contracts are more likely to be in the informal sector[19], we estimated the WFR by contract type and contract duration for formal-sector workers only. Limiting the sample to formal-sector workers, however, does not change the tentative conclusion made above. Formal-sector workers with written contracts continue to have a much lower average WFR (36,04%) than those with verbal contracts (71,15%). Similarly, WFR for formal-sector permanent workers is merely 29,77%, compared to the much higher WFR for those holding contracts of a limited term (63,29%).

[19] And this is indeed the case: the share of informal-sector workers amongst verbal contract holders is an overwhelming 85,18%, compared to a mere 6,74% for written contract holders. A higher percentage of workers with fixed-term contracts are in the informal sector (25,65%), compared to a much lower percentage amongst permanent contract holders (4,86%).

10.3. Worker flows by individual characteristics

Industry, firm, and job characteristics are important determinants of the extent of worker flows that a worker experiences, but his/her experience also varies by individual characteristics. We now explore a set of individual characteristics and their correlation with WFR. A greater WFR for a population group indicates more precarious employment for that group, conditional on all else being held constant. Although we cannot hold all other factors constant in the correlation analysis below, our results can still present some important preliminary insights into the volatility of employment for different groups of workers.

Table 10 presents the estimated average WFR for different groups of workers, as defined by different population groups, gender, age, education, and union membership. These WFR estimates are relevant for the period 2008q1 – 2019q2 except for union membership, which is not available in the data prior to 2010q3. Therefore, the average WFR by union membership is only relevant to the period 2010q3 – 2019q2.

Although differences in workers' earnings are shown to drive variations in WFR (Kerr, 2018), they are not examined here. This is because a majority of the earning information in the QLFS panel data is either unavailable or unreliable. Earning information was not collected in the QLFS surveys in 2008 and 2009 (Kerr & Wittenberg, 2019). From 2010 onwards, although earnings had become available in QLFS, its data quality was severely undermined by the earning imputations performed by StatsSA – it is impossible to distinguish the unimputed earning from the rest, nor is the imputation methodology made public by StatsSA. Furthermore, the imputation scheme appeared to have changed in 2012 (Kerr & Wittenberg, 2019). As a result, earning series based on QLFS data from 2010 onwards are not comparable across time and produced unbelievable changes in earning inequalities (Kerr & Wittenberg, 2021).

Also left out of Table 10 are geographical locations, even though they are possibly important drivers of variations in WFR since they are shown to affect one's probability of entry and exit of employment (Zizzamia & Ranchhod, 2019). The reason for their omission is that our panel data cannot track individuals who change living locations across quarters, which means that it cannot capture worker flows that involved changes in geographical locations. Because people who changed geographical locations are likely to be systematically different from those who stay in the same location, their worker flow patterns are also likely to differ. Therefore, if we used our data to estimate WFR by geographical locations, the resulting estimates will suffer from sample selection biases.

Table 10: Average quarterly worker flow rates by individual characteristics, 2008q1-2019q2

Individual characteristics	*WFR average*
Population groups	
African/Black	48,52%
Coloured	46,19%
Indian/Asian	38,23%
White	35,60%
Gender	
Male	47,86%
Female	44,36%
Age groups	
Age: 15-24	69,53%
Age: 25-34	49,98%
Age: 35-44	41,30%
Age: 45-54	38,50%
Age: 55-64	40,63%
Education	
None educ	57,58%
Some primary educ	56,19%
Incomplete secondary	54,85%
Matric	44,93%
Tertiary educ	29,03%
Union membership	
Union member	23,85%
Not union members	54,39%
Gross WFR	46,26%

Note: (1) All estimates are weighted using panel weights (2) WFR estimates by union membership is only relevant to the period 2010q3 – 2019q2 because this information was not collected prior to 2010q3. **Sources**: Author's calculations from QLFS and PALMS.

We first contrast worker flow experiences by racial groups. We find that Black South Africans are far more likely to be hired or separated in a given quarter, whereas white South Africans are least likely amongst all racial groups to experience quarterly worker flows. This is reflected in the variation of WFR amongst racial groups, which ranges from 35,6% for whites to 48,52% for black South Africans. The WFR for coloured South Africans is 46,19%, meaning that the extent of labour mobility for them is more comparable to that of black South Africans. In contrast, the WFR for Indians/Asians is 38,23%, which is closer to that of whites.

In OECD countries, worker reallocation occurs at a faster pace for the younger, females, and people at the extreme ends of the educational attainment (poorly educated or highly educated) (OECD, 2009:146-150). We examine whether this is true for South Africa.

When we estimate the average WFR by age, gender, and educational attainment, we find that, firstly, men have a slightly higher WFR (47,86%) than women (44,36%). Secondly, the average quarterly WFR for those aged 15-24 (69,53%) and 25-34 (49,98%) are higher than those who are older. But, once surpassing the age of 35, one's age does not monotonically decrease his/her probability of being hired or separated. In fact, WFR for those aged 55-64 (40,63%) is higher than the generation younger them who are aged 45-54 (38,50%). The slight rise in WFR from the second oldest age group to the oldest age group might reflect the rise in the number of retirees when people enter their sixties. Therefore, the findings regarding the variation of WFR by gender and age are slightly different from the evidence from OECD countries.

Lastly, we find that there is a monotonic relationship between WFR and education levels. WFR is higher amongst the less educated, and it is higher for the better educated. Interestingly, WFR declines only slightly from "no education" (57,58%) to "some primary education" (56,19%), and from "some primary education" to "incomplete secondary education" (54,85%). The first significant drop in WFR occurs from "incompletely secondary education" to "matric" (44,93%). The second and much more sizeable drop in WFR occurs from "matric" to "tertiary educated" (29,03%). The finding that education decreases WFR in a monotonic way is also inconsistent with evidence from the OECD countries where the average WFR is higher for both poorly educated and highly educated.

The varied pattern of WFR by education observed in the South African context may be explained in consideration of the fact that only matric and higher education credentials could signal higher productivity in the South African labour market while lower levels of education do not (Wittenberg, 2002; Duff and Fryer, 2005). Due to the poor signalling ability of lower levels of education, the real value of a job match involving a worker with a lower level of education is not revealed to the employer until after that worker has worked for some time. As a result, employers cannot screen these workers in an efficient way, which might lead them to adopt a strategy that involves hiring poorly educated workers in larger numbers while also simultaneously firing them whenever it is revealed that they are unproductive. This increases the WFR for those with low levels of education. In contrast, it is much easier to screen workers with higher levels of education. Therefore, employers can simply refuse to hire those whom they deem unproductive before forming a job match with the said person. This reduces the worker flow rates for the highly educated population.

The average WFR for unionised workers is extremely low at 23,85%. In contrast, the average WFR for non-unionised workers is 54,39%, almost twice as high as that for unionised workers. Despite this sharp contrast, it is difficult to assert, based on this finding, that union membership leads to greater difficulty in hiring or firing for firms, even though this is a narrative repeatedly portrayed in the media. This is because being unionised is correlated with other factors that will influence a worker's mobility. For instance, Kerr & Wittenberg (2021) noted that unionised members are increasingly concentrated among the high-income earners and those working in the public sector. Because both high earnings and public sectors are associated with reduced WFR, it is impossible to know from a simple correlation analysis that it is the trade union that caused the lower WFR observed amongst union members.

10.4. Reflection on the "labour rigidity" debate

We started this section with the hypothesis that, although gross worker flows are substantial in South Africa, people may have vastly different experiences in terms of labour mobility for being in different industries, firms, jobs, and population groups. Thus far, we have tentatively affirmed this premise with correlation analyses demonstrating that WFR does vary substantially at more disaggregated levels. At the industry level, wage workers in the "mining and quarrying", "finance", and "service" sectors appear to experience much lower WFR than their counterparts in "construction", "agriculture", and "utilities". At the firm level, employment at larger or public firms is more stable than that in smaller or private firms. At the job level, wage employees working on verbal or fixed-term contracts change jobs more frequently than those whose employment is backed by written or permanent contracts. At the individual level, workers who are black, coloured, male, young, or those who do not have high levels of education or union membership are much more susceptible to worker flows than the rest of the population.

At this point, it is relevant to revisit the "labour rigidity" debate. Puzzled by the question of why worker flows are so large in South Africa based on estimates provided by Kerr (2018), we provided three explanations. The first is that Kerr (2018) overestimated the extent of worker flows in South Africa. However, using an alternative set of data constructed from QLFS surveys that span over ten years, we found worker flows to be sizeable in both formal and informal sectors of South Africa. This shows that the estimate provided by Kerr (2018) is plausible.

The second explanation contends that the considerable magnitude of worker flows must imply that South African employment protection is not as stringent as is commonly believed. However, the heterogeneity analysis conducted above demonstrated that certain sections of the labour market exhibit considerable “rigidity”, as indicated by extremely low rates of worker flows, while the rest of the labour market bears the main brunt of employment adjustment

Therefore, the evidence presented compels us to conclude that the third explanation is the better solution to the debate. This explanation denies neither Kerr’s finding nor the widespread belief regarding South Africa’s labour market regulation stringency but, instead, it posits that rigidity differs across different sections of the South African labour market.

10.5. Regression analysis of the probability of separation

Although we have shown that the experience of worker flows varies across different sub-sections of the labour market, one of the challenges in our analysis is that the different characteristics investigated thus far are not independent of each other. For instance, although workers in “mining and quarrying” have smaller WFR, they are also more likely to be unionised, work for larger and formal firms, and have written contracts. Because all these characteristics are simultaneously correlated with lower levels of worker flows, it is impossible for our analysis to show whether the lower levels of worker flows in “Mining and quarrying” is a result of industry-specific factors, or simply due to the type of firms, jobs, and individuals that are more likely to be found in “Mining and quarrying”.

This drawback dampens the usefulness of our analysis. On the one hand, for policymakers who wish to address labour market rigidity, our analysis is of limited use in terms of showing where exactly in the labour market are hirings or firings more difficult. On the other hand, for those who want to better understand the drivers of precarious employment for vulnerable population groups, our analysis cannot adjudicate between the following two viewpoints: the first is that some people have precarious employment due to labour market discrimination, and the second is that they have precarious employment because they are more likely to enter precarious employment as a result of their individual characteristics.

To address some, but not all, of the aforementioned concerns, we estimate a probit model in a regression analysis to demonstrate the isolated effects of various characteristics on wage workers’ probability of *separation*. Ideally, we would want to examine both sides of the labour reallocation process – hires and separations. However, if “hiring” was our dependent variable in the regression, we would need to run a regression on a sample of individuals who initially do not have wage employment. As a result, we cannot control for various characteristics at the industry, firm, or job level. This would render our regression analysis much less informative. One way to resolve this issue is to vary the outcome variable, and estimate several models with each model predicting the probability of being hired into one specific industry, firm type, job type, as well as all the combinations of the different types of industries, firms, and jobs. While this is theoretically feasible, we have decided to forego this exercise due to its computational complexity. Instead, we have decided to focus on the "separation side" of the worker flows. The worker-level analysis of separation probability has been widely used in a large body of literature, either to examine the wage elasticity of firm-level labour supply (e.g. Ransom & Oaxaca, 2010; Booth & Katic, 2011) or to examine the underlying drivers of worker churning[20] (e.g., Shiferaw & Soderbom, 2021). Their works provide a rich reservoir of knowledge from which we can draw.

Similar to Shiferaw & Soderbom (2021), who analysed the probability of separation on a sample of Ethiopian formal-sector workers, our probit model controls for industry, firm, and worker characteristics as per the equation presented below. In this model, i indicates individuals and t indicates wave. The dependent variable S_{it} is a dummy that equals 1 if someone was separated from a wage job from $t-1$ to t, and zero otherwise (this variable is only defined for those who are wage workers at $t-1$). As is made clear in the equation, we use one's characteristics at $t-1$ to predict his/her probability of separation between $t-1$ and t.

$$S_{it} = \mu + \left(\sum_{k=1}^{9} \alpha_k industry_{i,t-1,k}\right) + \left(\sum_{k=1}^{6} \beta_k firm_{i,t-1,k}\right) + \left(\sum_{k=1}^{j} \gamma_k worker_{i,t-1,k}\right) + \varepsilon_{it}$$

Nine industry dummies ($\sum_{k=1}^{9} industry_k$) are included. These are "Agriculture", "Mining and quarrying", "Manufacturing", "Utilities", "Construction", "Transport", "Finance", "Services", and "Domestic Services". The missing category, "Trade", is the base industry. Six firm/job level characteristics ($\sum_{k=1}^{6} firm_k$) are included as independent variables, indicating if the worker is in the formal sector, is in the public sector, has a written contract, has a permanent contract, works for a medium-sized firm with 10-49 employees, or works for a large firm with more than 50 employees (the small-sized firm is the base category). We include the following worker characteristics ($\sum_{k=1}^{j} \gamma_k worker_k$): race, age group, gender, levels of education, and union membership. Lastly, we also control for quarter dummies and included a linear annual time trend.

[20] Worker churning is a concept related to worker flows. In a given firm, worker churning between $t-1$ and t is the excess worker flows over and above the net change in its employment.

We denote the residual error as ε_{it}. Clearly, given the parsimonious model specification, it is impossible for us to have controlled for all relevant observed/unobserved factors that both affect one's separation probability and are correlated with the independent variables. Therefore, it is implausible to presume that the residual errors have a zero conditional mean. This implies that our regression analysis cannot shed light on any causal relationships. For this reason, It is important to bear in mind that our regression analysis is no more than a more refined descriptive analysis, which is only insightful when it is compared against the results of correlation analyses conducted in earlier sections.

We run the probit model predicting separation probabilities on a pooled sample of 216184 individuals from the QLFS cross-sections over 2010q3 – 2019q1. As mentioned, union membership is not collected in the QLFS data prior to 2010q3, which forced us to forego observations prior to 2010q3[21]. In the estimation of the standard errors of the regression coefficients, we take into consideration of clustering effect at the level of the primary sampling unit[22]. The regression results are presented in Table 11 below, where the average partial effects (APE) of the independent variables and the associated statistical significance are reported.

We focus first on industries. Overall, if we rank industries by the reported APE on separation probability, and compare it to the industry ranking by average WFR calculated earlier, we find that these two rankings are not completely aligned, nor are they completely different. This suggests that a part of the variation in industry-level WFR can be explained by relevant firm/job/worker characteristics within industries, and therefore cannot be wholly attributed to industry-specific factors. For instance, wage workers in "Mining and quarrying", "Services", and "Domestic services" are less likely to be separated than those in other industries. This is indicated by the fact that only these industries have negative APE on separation probability relative to "Trade". This is only partially aligned with our expectation based on the average WFR reported for each industry – while average

[21] To test the robustness of the regression results, we also run the probit model without the union membership variable on the pooled sample comprising individuals from 2008q1 to 2019q1. We found that doing so does not change the sign or the statistical significance of the regression coefficient on most variables. To conserve space, we do not report these results in the paper.

[22] This is different from Zizzamia and Ranchhod (2019) who cluster the standard errors at the individual level based on the fact that the pooled sample of panel data includes repeated observations of the same individuals. To test the robustness of our results, we re-run the regression using a time-consistent individual identifier as the clustering variable in the regression. We find that doing so does not change the statistical significance of the regression coefficient or its level of significance, except for the numerical variable for the year of the survey: "Survey year" is statistically significant at a 10% level if the regression clusters at the individual level, but it loses its statistical significance if it is clustered at the level of the primary sampling unit.

WFR is the lowest in “Mining and quarrying” and “Services”, workers in “Domestic Services” have a higher average WFR than those in “Trade”.

Table 11: Average marginal effects on the probability of separating from wage employment between t-1 and t, 2010q3-2019q1

Dependent variable = Separation between t-1 and t	*Average marginal effect*
Industry characteristics	
Agriculture (Base: Trade)	0.00647
Mining and quarrying	-0.0474***
Manufacturing	0.0194***
Utilities	0.0820***
Construction	0.0735***
Transport	0.0608***
Finance	0.00799**
Services	-0.0125***
Domestic Services	-0.0325***
Firm and job characteristics	
Formal sector	-0.166***
Medium firms (10-49) (Base: small firms with 0-9 employees)	-0.0213***
Large firms (>50)	-0.0286***
Public firms	-0.0132***
Written contract	0.0395***
Permanent contract	-0.0990***
Worker characteristics	
Coloured (Base: Black)	-0.00476
Indian/Asian	-0.00320
White	0.00518
Female	-0.00156
Age: 25-34 (Base: Age 15-24)	-0.0285***
Age: 35-44	-0.0401***
Age: 45-54	-0.0492***
Age: 55-64	-0.0321***
Some primary educ (Base: No educ)	-0.00971
Incomplete secondary educ	-0.0282***
Matric	-0.0472***
Tertiary educ	-0.0741***
Union member	-0.0418***
Time control	
Quarter 2 (Base: quarter 1)	-0.00121
Quarter 3	-0.00741***
Quarter 4	-0.00319

Survey year	0.000701
Observations	216,184

Notes: (1) Standard errors are clustered at the cluster level (2) Panel weights are applied (3) Statistical significance of the average partial effects is reported: *** $p<0.01$, ** $p<0.05$, * $p<0.1$ (4) The quarter fixed effects and a linear trend are included in the regression but not reported in the table; **Sources**: Author's calculations using the QLFS and PALMS data.

In terms of firm/job characteristics, we find that, after controlling for the worker's industry and individual characteristics, one's separation probability is reduced if she/he belongs to the formal or public sector, works for a larger firm, or has a permanent contract. These results are aligned with our expectations based on prior correlation analyses of WFR. However, having a written contract increases a wage worker's probability of being separated by 4% in comparison to those with verbal contract, ceteris paribus. This is counter-intuitive and at odds with the earlier findings where we showed that WFR is lower amongst workers with verbal contracts than those with written contracts. To investigate this issue further, we break down separation into two sub-components – job-loss (flows from wage employment to unemployment/economic inactivity/non-wage employment) and EE flows. We then run two probit regressions, where the first regression predicts the probability of job loss and the second predicts the probability of EE flows[23], using the same independent variables and specifications for standard errors as in the separation regression. Our results show that a written contract reduces a wage worker's probability of job loss by 1,89%, but it increases his/her probability of EE flows by 6,27%, ceteris paribus. Both effects are statistically significant at a 1% level. Therefore, when viewed together, written contract increases a worker's probability of separation, but it is mostly driven by workers on written contracts who switch jobs within three months rather than by workers exiting employment – why this is the case is a pertinent issue for future studies.

Once we have controlled for industry, firm and job characteristics of a worker's employment, as well as his/her education, age, and union membership, we find that one's racial groups and gender no longer have any significant impacts on his/her separation probability. This suggests that what we observed before as the variation in WFR across different racial groups and genders may simply be due to the fact that different racial groups/gender tend to take up different types of jobs in different industries or have different education, age, or union membership profiles. This finding is surprising because both race and gender are shown by past studies to have large and statistically significant impacts on one's

[23] The regression predicting the probability of job loss is run on the same sample of individuals as in the regression predicting the probability of separation, whereas the regression predicting the probability of EE flows is run on a sample of individuals who are employed in both $t-1$ and t.

employment status (whether someone is employed, unemployed, or economically inactive) and one's level of earnings (even when a multitude of control variables are included).

Other results regarding individual characteristics are aligned with our expectations. The higher one's educational level, the less likely she/he will be separated from employment. This effect is most pronounced at the highest level – tertiary educated workers are 7% less likely to be separated than workers without any formal education, ceteris paribus. Separation probability also decreases with age, but only until one reaches the second oldest age group, or just before they reach the age of retirement. As a result, those aged 35-44 and 45-54 both appear to be less likely to separate from work than the oldest age group (55 – 64). Union members are 4,2% less likely to separate from work than non-unionised members if we hold constant the other relevant factors.

Overall, some regression results are at odds with the correlation analyses conducted before. In particular, the variation in WFR across workers in different industries appears to be driven by both industry-specific factors and within-industry heterogeneities in terms of workers. Secondly, WFR variation across population groups defined by race and gender may be entirely due to racial/gender differences in employment characteristics and other individual characteristics such as education and age. With this said, many regression results are also consistent with the pattern of WFR discovered before, especially those regarding the effect of firm and job characteristics, age, education, and trade union membership.

11. Conclusion

Utilising a large set of quarterly panel data tracking individuals between consecutive waves of the QLFS from 2008q1 to 2019q2, we provided a detailed examination of worker flows in South Africa over the ten-year period. This is the primary aim of this paper. Similar to Kerr (2018), we find substantial worker flows. The average quarterly worker flow rate between 2008 and 2019 was 46,26%. It is unlikely that such a large estimate is the result of some methodological or data-related errors, since our robustness checks show that our formal-sector worker flows estimates are very similar to those in Kerr (2018) over a comparable period. A cross-country comparison finds that South African worker flows are large, even compared to other developing countries and countries in Africa. This is surprising in a labour market considered by many as rigid, but it is consistent with Kerr (2018).

In terms of the cyclicality of worker flows, we find that worker flows are pro-cyclical in the informal sector, but they are acyclical in the formal sector. As a result, gross worker flows are weakly procyclical. This means that the matching efficiency of the labour market improves during expansions and weakens during downturns, but this is mainly due to the cyclical variations in hires and separations in the informal sector rather than those in the formal sector. The fact that worker flows in the formal sector appear irresponsive to cyclical trends echoes Verick (2012), who argued that rigidities of the formal sector prevented the formal-sector employers from adjusting to the business cycles.

Lastly, we find there to be substantial heterogeneity in worker flows by individual, firm, job, and industry characteristics. A series of descriptive analyses show that worker flows are much larger for some individuals (e.g., those who are black, younger, without trade union membership, or have little education), firms (e.g., those that are smaller in size or private), jobs (e.g., jobs with fixed-term or verbal contract), or industries (e.g., construction, agriculture, and domestic services) when compared to the rest. However, after adopting a regression analytical approach where interrelated confounding factors are held constant, we find that variation in worker flows across industries were driven not only by industry-specific factors but also within-industry heterogeneities of workers, and that some heterogeneous patterns of worker flows across different population groups (e.g., race and gender) may be caused by their differences in employment characteristics, education, and age. With this said, both descriptive and regression results reveal substantial heterogeneity of worker flows with regard to industry, firm characteristics, job characteristics, age, education, and trade union membership. This means that, although substantial gross worker flows of South Africa imply that the South African labour market as a whole is not as rigid as most people think, some sub-sections of the labour market are more rigid than others.

The secondary aim of this paper is to draw attention to the lack of systematic study in South Africa on different data matching algorithms for surveys with a rotating panel design such as the QLFS and their disparate matching quality. As such, we first described four distinct matching algorithms used on the QLFS or its predecessor LFS from three previous studies: Ranchhod and Dinkelman (2008), Anand, Kothari and Kumar (2016), and Leung, Stampini and Vencatachellum (2014). Then, utilising a set of official QLFS panel data from Stats SA which was matched on confidential information of individuals such as their names and street address, we proposed and implemented a method for evaluating the matching quality of these four matching algorithms. This investigation reveals that the algorithm used by Leung, Stampini and Vencatachellum (2014) appears inferior in matching quality to all other algorithms. Setting this algorithm apart, however, it is much more difficult to discriminate amongst the remaining algorithms in terms of matching quality – this is because, as one changes the strictness of the matching criteria for a matching algorithm, there will be a trade-off between its ability to reject false matches and its ability to accept true matches. This is consistent with the conclusions in Madrian and Lefgren (1999), who studied matching procedures for Current Population Survey from the United States.

Our investigation did generate one surprising finding. After comparing the matching algorithm in Ranchhod and Dinkelman (2008) with that in Anand, Kothari and Kumar (2016), we found strong evidence that Stats SA became better at maintaining the intertemporal consistency of PERSONNR from 2010 onwards. This implies that, from QLFS 2010q1 onwards, one can produce individual-level panel data of remarkable matching quality simply by merging individuals' household identifier (UQNR) and individual line numbers within the household (PERSONNR) between consecutive waves of the QLFS. Unfortunately, it is unclear what led to the sudden improvement in the quality of PERSONNR.

On the one hand, this paper extends Kerr's (2018) study on worker flows by using data that contains much richer information and covers a much longer period of time, and we replicated some of the key findings and stylised facts from the international literature on the cyclicality and heterogeneity of worker flows. On the other hand, we laid the methodological groundwork for future research on labour market dynamics in South Africa and solved some significant technical challenges that might be encountered by future researchers hoping to construct their own panel data based on matched surveys with a rotating panel design.

While our findings on worker flows are an important addition to the existing literature on labour market dynamics in South Africa, we must also acknowledge the limitations of this study. A major limitation of the study lies in our decision to measure employer-to-employer flows using an unconventional methodology based on a combination of tenure changes, inter-quarter changes in employer characteristics, and inter-quarter changes in employer sector formality. While the necessity to use this unconventional methodology is explained in detail, it could be a major source of measurement errors in our worker flows estimates. With this said, one finds assurance in the fact that our formal-sector worker flows estimates are very similar to those in Kerr (2018), indicating that the measurement errors on worker flows in this paper might only be moderate. However, future studies should either improve upon our methodology for measuring employer-to-employer flows, or inspect the accuracy of the employer-to-employer flows estimated in this paper using external data sources or information.

Appendices

Appendix A

Selected probit regression output for matching probabilities

	Average partial effect	
Dependent variable: *Individual is matched between t-1 and t* *(=1 if someone is successfully matched, 0 otherwise)*	*(1)* *2008q1-2008q2*	*(2)* *2019q1-2019q2*
Age:		
Age 25-34 (Base category: Age 15-24)	0.00459	0.0264***
Age 35-44	0.0451***	0.0596***
Age 45-54	0.0652***	0.0767***
Age 55-64	0.0694***	0.0481***
Gender:		
Female (Base category = male)	0.00955*	0.00871
Population group:		
Base category = African black		
Colored	0.0357***	-0.0232**
Indian	0.0346**	0.00421
White	0.0391***	-0.00123
Marital status:		
Married (Base category = never married)	-0.0237***	-0.00406
Widowed/widower	-0.0428***	-0.0231
Divorced	-0.124***	-0.0523***
Education status:		
None educ (Base category = matric education)	0.0655***	0.0521***
Some primary educ	-0.0196***	0.0105
Incomplete secondary educ	5.11e-05	0.0167***
Tertiary educ	-0.00530	-0.0182**
Employment status:		
(Base category = economically inactive excluding discouraged workers)		
Employed in formal sector but not in private household	0.00267	0.0217***
Employed in informal sector and not in private household	-0.0145*	-0.00684
Employed in household sector	-0.0507***	-0.0670***
Discouraged unemployment	-0.0113	0.00511
Searching unemployment	-0.0192***	0.00496
Geo-type:		
Rural (Base category = urban)	-0.00640	0.0185***

Response in person:		
(Base category = not responding in person)		
Responding in person	0.0234***	0.0124**
Response in person not specified	-0.0789	0.239***
Household size	-0.00901***	-0.00787***
Control for province	Yes	Yes
Household composition:		
(Base category = %female pensioner (55> yrs) in the household)		
%female kid (<15 yrs) in the household	0.155***	-0.0163
%male kid (<15 yrs) in the household	0.134***	0.0148
%female prime aged (16-55 yrs) in the household	0.0205	-0.0848***
%male prime aged (16-55 yrs) in the household	-0.0226	-0.118***
%male pensioner (55> yrs) in the household	-0.000800	-0.0363
Number of observations	59025	42024

Notes: (1) Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1 (2) The first probit regression is run on the QLFS 2008q1 cross-sectional individuals aged 15-64, and the second probit regression is run on QLFS 2019q1 cross-sectional individuals aged 15-64. (3) Only the size and the statistical significance of the average partial effect of each independent variables are reported. Province dummies are included in the regression but are not reported in this table. (4) Regressions are not weighted ***Sources***: Author's calculations from QLFS and PALMS data.

Bibliography

Abel, M., Burger, R. & Piraino, P. 2020. The value of reference letters: Experimental Evidence from South Africa. *American Economic Journal: Applied Economics*. 12(3):40-71.

Anand, R., Kothari, S. & Kumar, N. 2016. South Africa: Labor market dynamics and inequality. *International Monetary Fund Working Paper*. 16(137). DOI:10.5089/9781498358934.001.

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.

Bassanini, A. & Garnero, A. 2013. Dismissal protection and worker flows in OECD countries: Evidence from cross-country/cross-industry data. *Labour Economics*. 21:25-41.

Bassanini, A., Garnero, A., Marianna, P. & Martin, S. 2010. Institutional determinants of worker flows: A cross-country/cross-industry approach. *OECD Social, Employment and Migration Working Papers*. 107. DOI:10.1787/5kmbqvstc09x-en.

Bellmann, L., Gerner, H.D. & Upward, R. 2018. Job and worker turnover in German establishments. *The Manchester School*. 86(4):417-445.

Benjamin, P. 2014. Labour Law. In The *Oxford Companion to the Economics of South Africa*. H. Bhorat, A. Hirsch, R. Kanbur, M. Ncube, Eds. Oxford: Oxford University Press. 250-258.

Benjamin, P., Bhorat, H. & Cheadle, H. 2010. The cost of "doing business" and labour regulation: The case of South Africa. *International Labour Review*. 149(1):73-91.

Bertrand, M. & Crepon, B. 2020. Teaching Labor Laws: Evidence from a randomized control trial in South Africa. *IZA Discussion Paper.* 13513. *DOI:*10.2139/ssrn.3660259.

Bhorat, H. & Lilenstein, A. 2016. From regulatory amendments to employment shifts: The temporary employment services sector in South Africa. (Unpublished). Available: https://apso.org.za/images/documents/research/From-Regulatory-Amendments-to-Employment-Shifts-TES-Update-2018-.pdf [2022, July 14].

Bhorat, H., Kanbur, R. & Mayet, N. 2012. The Impact of Sectoral Minimum Wage Laws on Employment, Wages and Hours of Work in South Africa. *DPRU Working Paper*. 12(155). Available: http://www.dpru.uct.ac.za/sites/default/files/image_tool/images/36/DPRU%20WP12-154.pdf [2022, Jul 14].

Bhorat, H., Naidoo, K. & Yu, D. 2014. Trade unions in an emerging economy: The case of South Africa, *WIDER Working Paper* 2014(55). DOI:10.35188/UNU-WIDER/2014/776-9.

Booth, A.L. and 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.

Borjas, G. 2013. *Labor Economics*. Rev. 6th ed. New York: McGraw-Hill.

Branson, N. & Wittenberg, M. 2014. Reweighting South African National Household Survey Data to Create a Consistent Series Over Time: A Cross-Entropy Estimation Approach. *South African Journal of Economics*. 82(1):19-38.

Burgess, S., Lane, J. & Stevens, D. 2000. Job flows, worker flows, and churning. *Journal of labor economics*. 18(3):473-502.

Cahuc, P., Carcillo, S. & Zylberberg, A. 2014. *Labor Economics*. Cambridge, Massachusetts: The MIT Press.

Carranza, E., Garlick, R., Orkin, K. & Rankin, N. 2020. Job Search and Hiring With Two-Sided Limited Information About Workseekers' Skills. *Upjohn Institute Working Paper*. 20(328). DOI:10.2139/ssrn.3634690.

Cichello, P., Leibbrandt, M. & Woolard, I. 2014. Winners and losers: South African labour-market dynamics between 2008 and 2010. *Development Southern Africa.* 31(1):65-84.

Corseuil, C.H., Foguel, M., Gonzaga, G. & Ribeiro, E.P. 2014. Youth turnover in Brazil: Job and worker flows and an evaluation of a youth-targeted training program. *National University of La Plata Working Paper*. 155. Available: http://hdl.handle.net/10419/127666 [2021, Aug 30].

Davis, S. J. & Haltiwanger, J. 1998. Measuring gross worker and job flows. In *Labor Statistics Measurement Issues*. J. Haltiwanger, M. E. Manser & R. Topel, Eds. Chicago: University of Chicago Press. 77-122.

Davis, S. J. & Haltiwanger, J. 1999. Gross job flows. In *Handbook of Labor Economics: Volume 3B.* O. Ashenfelter & D. Card, Eds. North Holland: Elsevier. 2711–2805.

Davis, S. J. & Haltiwanger, J. 2014. Labor market fluidity and economic performance. *National Bureau of Economic Research Working Paper.* 20479. Available: http://www.nber.org/papers/w20479 [2021, Aug 28].

Davis, S. J., Faberman, R. J. & Haltiwanger, J. 2006. The flow approach to labor markets: New data sources and micro-macro links. *Journal of Economic perspectives*. 20(3):3-26.

Donovan, K., Lu, J., Schoellman, T. 2020. Labor market dynamics and development. *Yale University Economic Growth Center Discussion Paper*. 1071. Available: http://hdl.handle.net/10419/227653 [2021, Aug 28].

Duff, P. & Fryer, D. 2005. Market failure, human capital, and job search dynamics in South Africa: The case of Duncan village. *DPRU Working Paper*. 5(98). Available: http://hdl.handle.net/11427/7365 [2022, Feb 1]. Cape Town: Development Policy Research Unit, University of Cape Town.

Essers, D. 2017. South African labour market transitions since the global financial and economic crisis: Evidence from two longitudinal datasets. *Journal of African Economies*. 26(2):92-222.

Fawcett, T. 2006. An introduction to ROC analysis. *Pattern recognition letters*. 27(8):861-874.

Fedderke, J. 2012. The Cost of Rigidity: the case of the South African labor market. *Comparative Economic Studies*. 54(4):809-842.

Festus, L., Kasongo, A., Moses, M. & Yu, D. 2016. The South African labour market, 1995–2015. *Development Southern Africa*. 33(5):579-599.

Fields, G. S. 2014. Segmented labour markets in South Africa. In The *Oxford Companion to the Economics of South Africa*. H. Bhorat, A. Hirsch, R. Kanbur, M. Ncube, Eds. Oxford: Oxford University Press. 244-249.

Fields, G.S. 2000. The Employment Problem in South Africa. In *Trade & Industry Monitor Volume 16*. Available: https://hdl.handle.net/1813/74996 [2022, Feb 1].

Go, F., Kearney , M., Korman, V., Robinson, S. & Thierfelder, K. 2010. Wage subsidy and labour market flexibility in South Africa. *The Journal of Development Studies*. 46(9):1481-1502.

Gomes, P. 2012. Labour market flows: Facts from the United Kingdom. *Labour Economics*. 19(2):165-175.

Haltiwanger, J., Kugler, A., Kugler, M., Micco, A. & Pages, C. 2004. Effects of tariffs and real exchange rates on job reallocation: evidence from Latin America. *The Journal of Policy Reform.* 7(4):191-208.

Heintz, J. and Posel, D. 2008. Revisiting informal employment and segmentation in the South African labour market. *South African Journal of Economics*. 76(1): 26-44.

Huber, P. & Smeral, K. 2006. Measuring worker flows. *Applied Economics*. 38(14):1689-1695.

Ingle, K. & Mlatsheni, C. 2017. The extent of churn in the South African youth labour market: Evidence from NIDS 2008-2015. *SALDRU Working Paper Series*. 201(1). Available: http://hdl.handle.net/11090/884 [2022, Feb 1].

Kaplan, D.S., Martínez, G. & Robertson, R. 2007. Mexican employment dynamics: Evidence from matched firm-worker data. *World Bank Policy Research Working Paper*. 4433. Available: https://openknowledge.worldbank.org/bitstream/handle/10986/7599/wps4433.pdf?sequence=1 [2021, Aug 30].

Kerr, A. & Wittenberg, M. 2019. Earnings and employment microdata in South Africa. *World Institute for Development Economics Research Working Paper*. 47. DOI:10.35188/UNU-WIDER/2019/681-4.

Kerr, A. & Wittenberg, M. 2021. Union wage premia and wage inequality in South Africa. *Economic Modelling*. 97:255-271.

Kerr, A. 2018. Job flows, worker flows and churning in South Africa. *South African Journal of Economics*. 86:141-166.

Kerr, A. 2021. Measuring earnings inequality in South Africa using household survey and administrative tax microdata. *WIDER Working Paper*. 82(2021). DOI: 10.35188/UNU-WIDER/2021/020-7.

Kerr, A., Lam, D. & Wittenberg, M. Post-Apartheid Labour Market Series 1993-2019 [dataset]. Version 3.3. Cape Town: DataFirst [producer and distributor], 2019. DOI:10.25828/gtr1-8r20.

Kerr, A., Wittenberg, M. and Arrow, J. 2014. Job Creation and Destruction in South Africa. *South African Journal of Economics*. 82(1):1-18.

Kingdon, G. & Knight, J. 2009. Unemployment: South Africa's Achilles' heel. In *South African economic policy under democracy*. J. Aron, B. Kahn, G. Kingdon, Eds. New York: Oxford University Press Inc. 300-326.

Lane, J.I., Isaac, A.G. & Stevens, D.W. 1996. Firm heterogeneity and worker turnover. *Review of Industrial Organization*. 11(3):275-291.

Lane, J.I., Stevens, D. & Burgess, S. 1996. Worker and job flows. *Economics Letters*. 51(1):109-113.

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.

Lohr, S. 2019. *Sampling: Design and Analysis*. Rev. 2nd ed. Florida: Taylor & Francis Group.

Madrian, C. & Lefgren, J. 1999. An approach to longitudinally matching Current Population Survey (CPS) respondents. *National Bureau of Economic Research Technical Working Paper.* 247. DOI:10.3386/t0247.

Magruder, J.R. 2012. High unemployment yet few small firms: The role of centralized bargaining in South Africa. *American Economic Journal: Applied Economics*. 4(3):138-66.

Martin, J.P. & Scarpetta, S. 2012. Setting it right: Employment protection, labour reallocation and productivity. *De Economist*. 160(2):89-116.

Mlatsheni, C. & Leibbrandt, M. 2014. Unemployment in South Africa. In The *Oxford Companion to the Economics of South Africa*. H. Bhorat, A. Hirsch, R. Kanbur, M. Ncube, Eds. Oxford: Oxford University Press. 236-243.

Mortensen, D.T. & Pissarides, C.A. 1994. Job creation and job destruction in the theory of unemployment. *The Review of Economic Studies*. 61(3):397-415.

Nakamura, A., Nakamura, E., Phong, K. & Steinsson, J. 2020. Worker reallocation over the business cycle: Evidence from Canada. Available: https://eml.berkeley.edu/~jsteinsson/papers/cwf.pdf [2021, June 11].

Organisation for Economic Co-operation and Development. 2009. How Do Industry, Firm and Worker Characteristics Shape Job and Worker Flows? In *OECD Employment Outlook 2009: Tackling the Jobs Crisis*. S. Scarpetta, Eds. Paris: OECD Publishing. 117-163.

Piek, M., von Fintel, D. & Kirsten, J. 2020. Separating employment effects into job destruction and job creation. *WIDER Working Paper.* 2020(5). DOI:10.35188/UNU-WIDER/2020/808-5.

Posel, D., Casale, D. & Vermaak, C. 2014. Job Search and the Measurement of Unemployment in South Africa. *South African Journal of Economics.* 82(1):66-80.

Ranchhod, V. & Dinkelman, T. 2008 Labour Market Transitions in South Africa: What can we learn from matched Labour Force Survey data? *SALDRU Working Paper Series.* 14. Available: http://hdl.handle.net/11090/40 [2021, June 21].

Ransom, M.R. & Oaxaca, R.L. 2010. New market power models and sex differences in pay. *Journal of Labor Economics*. 28(2):267-289.

Schoer, V., Rankin, N. & Roberts, G. 2014. Accessing the first job in a slack labour market: Job matching in South Africa. *Journal of International Development*. 26(1):1-22.

Schoiswohl, F. & Wuger, M. 2016. A Representative Matched Cross-section Survey for Austria – Measuring Worker Flow Dynamics with the Austrian Labour Force Survey. *Department of Economics Working Paper Series.* 223. Available: https://epub.wu.ac.at/id/eprint/4935 [2021, June 11].

Shiferaw, A. & Soderbom, M. 2021. Worker Turnover and Job Reallocation: Evidence from Matched Employer-Employee Data. *Economic Development and Cultural Change*. DOI:10.1086/718686.

South African Reserve Bank. 2021. *Statistical Tables: General Economic Indicators.* In *South African Reserve Bank Quarterly Bulletin June 2021*. Pretoria: South African Reserve Bank. Available: https://www.resbank.co.za/content/dam/sarb/publications/quarterly-bulletins/quarterly-bulletin-publications/2021/back-up-folder-june-/12Statistical%20tables%20General%20economic%20indicators.pdf [2022, Feb 1].

Statistics South Africa. 2006. *The South African Labour Force Panel Study: Methodology document*. Pretoria: Statistics South Africa. Available: https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/483/download/6406 [2021, June 16].

Statistics South Africa. 2018. *Labour market dynamics in South African, 2018.* Pretoria: Department of Statistics South Africa. Available: http://www.statssa.gov.za/publications/Report-02-11-02/Report-02-11-022018.pdf [2021, Aug 23].

Statistics South Africa. 2019. *Statistical release: Quarterly Labour Force Survey Quarter 4:2019.* Pretoria: Department of Statistics South Africa. Available: https://www.statssa.gov.za/publications/P0211/P02114thQuarter2019.pdf [2021, Aug 23].

Tharwat, A. 2020. Classification assessment methods. *Applied Computing and Informatics*. 17(1):168-192.

van der Merwe, J. & Krygsman, S. 2020. The relationship between transport accessibility and employment duration. *WIDER Working Paper.* 2020(56). DOI:10.35188/UNU-WIDER/2020/813-9.

Venter, J.C. 2011. Business cycles in South Africa during the period 2007 to 2009. In *South African Reserve Bank Quarterly Bulletin June 2011*. Pretoria: South African Reserve Bank. Available: https://www.resbank.co.za/content/dam/sarb/publications/quarterly-bulletins/articles-and-notes/2011/4678/Article---Business-cycles-in-South-Africa-during-the-period-2007-to-2009.pdf [2022, Feb 16].

Verick, S. 2012. Giving up job search during a recession: The impact of the global financial crisis on the South African labour market. *Journal of African Economies*. 21(3):373-408.

von Fintel, D. 2016. Wage flexibility in a high unemployment regime: Spatial heterogeneity and the size of local labour markets. *Stellenbosch Economic Working Papers*. 9. Available: https://ideas.repec.org/p/sza/wpaper/wpapers265.html [2022, Feb 1].

Wittenberg, M. 2002. Job search in South Africa: A nonparametric analysis. *South African Journal of Economics*. 70(8):1163-1196.

Zizzamia, R. & Ranchhod, V. 2019. Measuring employment volatility in South Africa using NIDS: 2008-2017. *SALDRU Working Paper Series*. 246. Available: http://www.opensaldru.uct.ac.za/handle/11090/961 [2021, June 11].