WP/16/137

IMF Working Paper

South Africa: Labor Market Dynamics and Inequality

by Rahul Anand, Siddharth Kothari, Naresh Kumar

I N T E R N A T I O N A L   M O N E T A R Y   F U N D

 WP/16/137

IMF Working Paper

African Department

South Africa: Labor Market Dynamics and Inequality[1]

Prepared by Rahul Anand, Siddharth Kothari, and Naresh Kumar

Authorized for distribution by Laura Papi

July 2016

Abstract

This paper analyzes the determinants of high unemployment in South Africa by studying labor market dynamics using individual level panel data from the Quarterly Labor Force Survey. While prior work experience and gender are found to be important determinants of the job-finding rate, education attainment and race are important determinants of the job-exit rate. Using stock-flow equations, counterfactual exercises are conducted to quantify the role of these different transition rates on unemployment. The paper also explores the contribution of unemployment towards inequality. Reducing unemployment is found to be important for reducing inequality – estimates suggest that a 10 percentage point reduction in unemployment lowers the Gini coefficient by 3 percent. Achieving a similar reduction solely through transfers would require a 40 percent increase in government transfers.

JEL Classification Numbers: J63, J21, D31

Keywords: Uunemployment, South Africa, Labor market transitions, Inequality

Author’s E-Mail Address: ranand@imf.org; skothari@imf.org; nkumar@imf.org

[1] We are grateful to Laura Papi, Axel Schimmelpfennig, Domenico Fanizza, Jose Torres, Yi Wu, Ahmat Jidoud and our colleagues in the African Department for helpful comments and discussions. We benefited from the feedback received from participants at the Economics Society of South Africa Conference (ESSA) (2015) at the University of Cape Town, South Africa.

Contents Page

I. INTRODUCTION

South Africa has made significant strides in economic and social development since its first democratic elections about two decades ago. Growth averaging 3.3 percent since 1994 and social assistance that now reaches more than half of all households have resulted in a 40 percent increase in real per capita GDP and a 10 percentage point drop in the poverty rate.

Yet South Africa's economy faces important structural challenges—the two most important being the high levels of unemployment and inequality (Figure 1). The Gini coefficient at about 65 is one of the highest in the world. The unemployment rate is 25 percent (35 percent including discouraged workers), with the youth (age 15-24) unemployment rate being even higher, despite one of the lowest participation rates in the world. This paper uses household survey data to shed light on the factors driving these two structural challenges. In particular, we analyze the determinants of labor market outcomes in South Africa, and the contribution of unemployment in explaining the high levels of inequality.

To study labor market outcomes, we construct a panel dataset using the Quarterly Labor Force Survey (QLFS). As the panel allows us to follow the same individual across quarters, we use this dataset to analyze the probability of individuals transitioning into and out of unemployment (job-exit and job-finding rates, respectively), and how these transition rates differ by individuals' characteristics: education levels, experience, age, etc. We then try to quantify the effect of different job-finding and job-exit rates on aggregate unemployment using simple stock-flow equations (Shimer 2012; Cortes et al. 2014). These counterfactual exercises allow us to ascertain the macro effects of the differences in transition rates, which we document from individual level regressions.

Our results suggest that experience, age, race, sex, and insider-outsider dynamics have played an important role in determining labor market outcomes and unemployment in South Africa. Focusing on the job-finding rate, we find that prior work-experience plays an important role in determining the employability of individuals with education playing a minor role (potentially due to poor quality). People with prior work-experience have almost 50 percent higher job-finding rate than those without experience, with experience being even more important for young job seekers—compounding the problem of high youth unemployment. The counterfactual exercise suggests that the difference in job-finding rates between the experienced and those without experience can result in an aggregate unemployment rate difference of 11 percentage points. Similarly, long-term unemployment lowers future job finding rates, and women also have lower job finding rates compared to men.

We then turn to the job-exit rate (probability of individual transitioning from employment to unemployment), and find that education is an important determinant of job security—individuals with higher education have significantly lower job-exit rates. All other factors that improve employability of an individual also matter for retaining a job. Young individuals and women not only have low job-finding rates, but also have high job-exit rates, resulting in high unemployment rates among these groups. Blacks also have higher job-exit rates compared to other racial groups.

We also analyze the role of labor market institutions in determining labor market dynamics in South Africa. We focus on the effect of trade-union membership, and find that being a union member increases job-security. However, we also find evidence for the harmful effect of union density on outsiders. In particular, unemployed individuals who have previously worked in highly unionized industries find it more difficult to find a job in the future.

In addition to the results on job-finding and job-exit rates, we analyze the probability of an individual transitioning into formal sector employment. We find that the informal sector acts as a stepping stone to formal employment: individuals who are unemployed have a much lower probability of finding a formal sector job compared to those working in the informal sector.

Next we examine the role of unemployment in explaining the high levels of inequality in South Africa, using data from the third wave of the National Income Dynamics Study (NIDS) conducted in 2012. Our analysis suggests that reducing unemployment by 10 percentage points would lead to a fall in the Gini coefficient from 0.665 to 0.645. This may appear small, but to achieve a similar reduction in Gini solely through fiscal transfers would require an increase in transfers by about 40 percent (equivalent to 3.3 percent of GDP in 2012 or about 11.1 percent of government expenditure). Therefore, without progress on reducing unemployment, reduction in inequality may be difficult to achieve through fiscal transfers alone.

Overall, the results suggest that while improving educational outcomes remains crucial to reducing unemployment, providing work experience (formal or informal) holds the key to lowering unemployment in the short run, especially for the most disadvantaged groups – youth, women, and blacks. Targeted policy interventions, such as the Youth Tax Incentive Scheme, may help improve the employability of young South Africans. Simultaneously, labor market reforms to increase the influence of outsiders may also help in creating jobs.[2]

There is a large literature which tries to explain the high unemployment rate in South Africa. As documented in Kingdon and Knight (2009), the abolition of apartheid was followed by a large increase in the supply of labor, while demand for labor stagnated. Many reasons have been suggested for labor demand lagging supply. Lack of quality education and a mismatch between demand and supply of skills been argued to be an important driver of unemployment. Skill biased technical change and decline in mining and agriculture has held back demand for low skilled workers. The influx of women and blacks into the labor force, following the abolition of apartheid, has on the other hand greatly increased the supply of low skilled workers, thus contributing to an increase in unemployment (Banerjee et al. 2008). However, Levinsohn (2008) argues that while improving education outcomes is important in the long run, it cannot solve the problem of unemployment in the short term. Instead, he calls

[2] Given the nature of the study, the paper cannot cover all reform areas that could give rise to an increase in labor demand. For example, product market reform, trade liberalization, reduced policy uncertainty, and improvements in the business environment may be important for boosting labor demand, but this study does not cover these reform areas.

for a targeted wage subsidy for young inexperienced workers to encourage firms to hire individuals who have recently entered the labor force.

Rigid labor markets have propped up real wages, preventing labor markets from equilibrating. Banerjee et al. (2008) find a substantial union premium for wages, suggesting that trade unions are propping up wages for insiders. Magruder (2012) finds that industries with centralized bargaining councils have about 8-13 percent lower employment, with losses concentrated in small firms. Other important factors that might be contributing to subdued labor demand and high levels of unemployment include: large costs of commuting to work due to poor public transportation and long distances between business centers and residential areas (Ngarachu et al. (forthcoming)); high crime rates which discourage the setting up of informal firms (Banerjee et al. 2008); and product market restrictions and supply side bottlenecks.[3]

A key contribution of this paper is to create a quarterly panel from the QLFS in order to study the determinants of labor market outcomes. There are a few recent papers which are most closely related to this paper as they use the same dataset to study particular transition rates. Leung et al. (2014), following the same individual over time, examine the probability of a person being employed (without distinguishing between finding and exit rates when doing their regression analysis) and show that human capital can act as a buffer to external shocks. Essers (2014) focuses solely on the probability of an individual exiting from formal sector employment and how this probability has changed during the Global Financial Crisis (GFC). Verick (2012) focuses on the increase in the share of discouraged workers during the recession, but does not conduct regression analysis for transition rates. In comparison to these studies, this paper conducts a more comprehensive analysis of transitions in and out of employment, and into formal sector jobs. We also quantify the possible impact of different job finding and exit rates on aggregate unemployment.

The paper is organized as follows. In Section II, we present the data and discuss the methodology used in the paper. In Section III, we discuss the main results on unemployment. Section IV discusses the main results on inequality, and Section V concludes the paper.

II. Data and Methodology

A. QLFS—Constructing the Panel

We use data from the Quarterly Labor Force Survey (QLFS) of South Africa from 2008Q1 to 2014Q3.[4] The survey, conducted by Statistics South Africa (Stats SA), interviews approximately 30,000 households every quarter, and asks detailed questions regarding the labor force status of individuals as well as their demographic characteristics.

---

[3] Anand et al. (2016) show that such bottlenecks have reduced the response of firms to aggregate shocks such as an exchange rate depreciation.

[4] All the waves of the QLFS are publically available thanks to StatsSA.

The survey is designed as a rotating panel, with 25 percent of the housing units rotated out of the sample every quarter. Once a housing unit is selected to be a part of the sample, it is surveyed for four consecutive quarters. However, the panel dimension is at the level of the housing unit and not the household i.e. if one household moves out of the housing unit (or address) and another moves in then the new household will be surveyed.

As we are interested in studying labor market dynamics of individuals, we need to construct a panel dataset at the individual level rather than at the housing unit level. To do so, we implement a matching algorithm similar to the one used in Ranchhod and Dinkelman (2007).[5] We first match housing units and individuals across quarters using the household and individual identifiers provided with the dataset. Then to ensure that we do not have false matches (due to households moving out for example) we check to see if matched individuals have the same sex and race across surveys and that their age does not differ by more than one year. Of the total of 2,355,379 observations, 1,318,338 were matched to the next quarter, implying a match rate of 56 percent.[6]

While Stats SA does not generally release an official individual level panel for the QLFS, they did release a one-off individual level panel between 2013 Q3 and 2013 Q4. We can use this official panel (which was based on non-publically available data, including names of individuals) to evaluate the performance of our matching algorithm. In particular, we can compare the results of our matching algorithm for these two quarters to the official individual level panel. The results of this comparison are presented in Table 1. Out of the total observations in 2013 Q3, both Stats SA and our algorithm matched 62.70 percent and failed to match 33.98 percent. Therefore, for 96.68 percent of the sample, our algorithm gave the same outcome as the official panel. As the main aim of this paper is to analyze labor market transitions, we would be especially concerned if our matching algorithm gave us a large number of false matches as this would give rise to more transitions than is actually the case. Therefore, it is reassuring to see that our algorithm only generated 0.36 percent false matches.

One inherent limitation of the data is that attrition and household moves can result in the matched sample not being representative of the population as a whole. While it is impossible to exactly quantify the effect of such a bias, Table 2 shows that the matched sample seems quite similar to the full sample, at least in terms of some common observables. In particular, the share of the population in different age, sex, race, and education groups is quite similar for the full sample and for the sub-sample that we were able to match to the next quarter.

---

[5] Similar individual level panel datasets using the QLFS were constructed by Verick (2012) and Leung et al. (2014).

[6] Note that given the rotating panel structure of the survey, theoretically the match rate should not have been more than 75 percent to begin with. Furthermore, given the attrition in the survey (a housing unit was interviewed in one quarter but dropped out in the next quarter for some reason) we expect the match rate to be below 75 percent.

In all our analysis of labor market dynamics, we restrict the sample to individuals between the age of 15 and 64, i.e. working age population.

B. Methodology—Labor Market Dynamics

We are interested in analyzing how labor market transition probabilities depend on individual characteristics. For example, we want to test whether the probability of an unemployed getting employed (or an employed becoming unemployed) depends on their education status and past experience.[7] We do this by using different dummy variables corresponding to different labor market transition in a simple probit regression model.

In particular, we consider the following three left hand side variable:

1) **Job-finding rate:** The LHS variable is a dummy which takes value 1 if an individual was unemployed in quarter *t* but became employed in quarter *t+1*, and 0 if an individual was unemployed in quarter *t* and was *not* employed in quarter *t+1*.

$$f_{i,t} = \begin{cases} 1 & unemployed\ in\ t\ but\ employed\ in\ t+1 \\ 0 & unemployed\ in\ t\ and\ not\ employed\ in\ t+1 \end{cases}$$

This variable on the LHS of the probit regression will allow us to measure how the probability of an individual transitioning out of unemployment and into employment depends on individual level characteristics.

2) **Job-exit rate:** The LHS variable is a dummy which takes value 1 if an individual was employed in quarter *t* but became unemployed in quarter *t+1*, and 0 if an individual was employed in quarter *t* and was *not* unemployed in quarter *t+1*.

$$d_{i,t} = \begin{cases} 1 & employed\ in\ t\ but\ unemployed\ in\ t+1 \\ 0 & employed\ in\ t\ and\ not\ unemployed\ in\ t+1 \end{cases}$$

This variable on the LHS of the probit regression will allow us to measure how the probability of an individual transitioning out of employment and into unemployment depends on individual level characteristics.

3) **Probability of transitioning to formal employment:** The LHS variable is a dummy which takes value 1 if an individual was unemployed or working in the informal sector in quarter *t* but was employed in the formal sector in *t+1*, and 0 if an individual was unemployed or working in the informal sector in quarter *t* and was *not* employed in the formal sector in *t+1*.

$$i_{i,t} = \begin{cases} 1 & unemployed/informal\ sector\ in\ t\ but\ formal\ sector\ \ in\ t+1 \\ 0 & unemployed/informal\ sector\ in\ t\ and\ not\ formal\ sector\ \ in\ t+1 \end{cases}$$

This variable will allow us to measures the probability of an individual to transition into a formal sector job. We are especially interested in testing whether an individual who is working in the informal sector is more likely to transition into a formal sector job as compared to an unemployed individual.

[7] See Kerr et al. (2014) who look at job creation and destruction in South Africa using firm level data.

For all the three variables described above, we run probit regressions of the form:

$$y_{i,t} = \Phi\left(\beta X_{i,t}\right)$$

Where $y_{i,t}$ can be $f_{i,t}, d_{i,t}, or\ i_{i,t}$ depending on the transition probability being considered, $X_{i,t}$ represents individual level characteristics like age, sex, race etc. The individual level characteristics are included as categorical variables i.e. for race we include a dummy for each race category (excluding one category to avoid multicollinearity). Therefore, the coefficients (β) measure the difference between the particular group's transition rate and that of the excluded group.

The individual level characteristics $\left(X_{i,t}\right)$ that we include in the regressions vary depending on the transition probability being considered. Four sets of control variables are included in all regression. These are: (i) education level categorized into four groups—less than primary, primary but less than secondary, secondary but less than university, and university or more; (ii) race categorized into four groups—Black, Colored, Indian/Asian, and White; (iii) age categorized into five groups—15 to 24, 25 to 34; 35 to 44; 45 to 54, and 55 to 64; and (iv) gender—male and female.

For the job-finding rate regressions, in addition to the usual demographic variables, we also include two additional relevant variables:

1) Dummy for long-term unemployed: variable takes value 1 for those who have been unemployed for less than a year and 0 otherwise.[8] Including this variable in the regression allows us to test the scarring effect of long-term unemployment (whether being unemployed for a long period reduces employment prospects).
2) Dummy for experience: variable takes value 1 if the individual has never been employed before and 0 if she has been employed before. Including this variable in the regression allows us to test the importance of prior experience in determining the future employability of an individual.

Similarly, for the job-exit rate regression, we include the following two variables in addition to the usual demographic variables:

1) Sector of employment: variable takes on three values depending on whether the individual is employed in the formal sector, informal sector, or in a private household.[9] This allows us to test whether informal sector employees are more likely to exit from employment.
2) Contract duration: variable takes four values depending on whether an individual is on a limited term contract, permanent contract, unspecified duration contract, and finally if contract duration is not applicable for the job (self-employed for example). This controls for the fact that temporary workers with limited duration contracts are more likely to exit employment.

[8] This is based on a question in the survey asking individuals to report their duration of unemployment.
[9] Another possibility would be to run separate regressions for the job-exit rate of formal and informal sector workers. However, the informal sector comprises only 25 percent of the workforce. Therefore, we do not run separate regressions but control for the sector of employment by including it as an explanatory variable in our regression.

As a robustness check, we construct two versions of each of the LHS variables described above. Our baseline results use the broad definition of unemployment (which includes discouraged individuals who are not actively looking for work but would be willing to work if offered a job) when constructing the LHS variables, but we do robustness checks where we use the narrow definition.

All regressions include time dummies and dummies for the province in which the individual resides. This is especially important as our sample period includes the Global Financial Crisis (GFC) when the job-finding rate or destruction rate might have changed substantially. By including time dummies, we ensure that the coefficients on the individual level characteristics are identified from within time-period variation only.

We use survey weights in all regressions and cluster standard errors at the household level unless otherwise mentioned.

III. Analysis of Unemployment

A. Micro-Regression

Result I. Previous experience is an important determinant of job-finding rates, while education has almost no effect.

Table 3 reports results of the probit regression—the marginal effects, and not the raw coefficients—where the left hand side variable is a dummy that measures the transition probability from unemployment to employment i.e. the LHS variable is the dummy $f_{i,t}$ described above. The rows report the difference between the job-finding rate of the group in question compared to the excluded group (mentioned in the first column of the table).[10] Appendix II reports the levels of the job-finding rate from the same regressions.

As evident from column 1, the job-finding rate does not differ substantially across different education groups and race. We might have expected more educated individuals to transition more quickly from unemployment to employment; however, this is not the case. This may be due to several reasons. Individuals with high education could have higher reservation wages and therefore wait longer before accepting a job. Also, as the unemployment rate among more educated individuals is lower (due to lower job-exit rates), the group of educated unemployed may be a selected sample of the universe of educated individuals.[11] Finally, it

[10] Kerr et al. (2014) study job creation and destruction using a firm level dataset. Our results are not strictly comparable to theirs as a transition from one job to another without passing through unemployment would be recorded as job creation and destruction in the firm level data, but does not show up as a transition in our individual level panel.

[11] In particular, this result should not be interpreted as saying that the finding rate of a randomly selected educated individual is the same as that of a randomly selected person with low education level. This is because, our regressions condition on being unemployed and the group of educated unemployed are arguably a selected sample of the universe of educated individuals. However, from a macro perspective the result is still informative

(continued…)

may be suggesting that because of low quality, educational attainment is not a reliable signal of productivity, and hence employability.[12]

The level of experience on the other hand is a very important determinant of the job-finding rate—an individual with no prior work experience has a lower job-finding probability than someone with experience (after controlling for age, education etc). In terms of magnitude, estimated coefficients suggest that the job finding rate of individuals with no experience is about 46 percent lower than that of individuals with prior work experience.[13] This supports the view that employers use experience as an important screening device when hiring workers (potentially because schooling is not a good screening device due to poor quality of education).

Result II. Long-term unemployment reduces future employability. Moreover, the youth and women have lower probability of finding jobs.

Similarly, those who are unemployed for short periods have a finding probability which is 0.072 (45 percent) more on average than the long-term unemployed. Age and gender are also important determinant of the job-finding rate with older individuals and men more likely to transition into employment compared to 15 to 24 year olds and women.

Result III. Previous experience matters more for the youth.

Column 2 in Table 3 repeats the same probit regression but now includes an interaction of age with the experience variable. This allows us to test whether experience within the same age group affect the employability of workers. The excluded group is 15 to 24 year olds who have worked before. As can be seen, 15 to 24 year olds with no experience have a much lower job-finding probability compared to those who have worked before. Furthermore, even 25 to 34 year olds with no experience have a much lower job-finding probability compared to 15 to 24 year olds with experience.

Columns 3 and 4 of Table 3 repeat the same regressions as column 1 and 2 but using the narrow definition of unemployed as opposed to the broad definition. The results are qualitatively similar.[14]

---

as it suggests that the lower level of unemployment among the higher educated individuals is not being driven by the higher finding rate among the educated unemployed (which is the relevant sample for estimating the finding-rate which determines unemployment within each education group), but rather by their lower job-exit rates (as discussed in Table 4 below).

[12] Given the fact that different racial groups might be attending very different types of educational institutions, we also tried looking at the interaction of race and education. There was no significant difference in job finding rate for different education levels within a given race either.

[13] In particular, the job-finding probability of those who have worked before is 0.131, while those who have never worked before is 0.071 (Table 1 in Appendix II) which translates into a 46 percent lower job-finding rate for the inexperienced workers.

[14] StatsSA (2013) which used the panel dimension between 2013Q3 and 2013Q4 also documented transition rates by various demographic characteristics and found that people who had work experience were almost three times more likely to find a job compared with those who had never worked before (StatsSA, 2013). However,

(continued…)

Result IV. Higher education is associated with lower job-exit rate.

Table 4 reports results for the probit regression where the left hand side variable is a dummy that measures the transition probability from employment to unemployment i.e. the LHS variable is the dummy $d_{i,t}$ described above. The table again reports the marginal effects from the probit (and not the raw coefficients). In particular, the rows report the difference between the job-exit rate of the group in question compared to the excluded group (mentioned in the first column of the table). Appendix II reports the levels of the job-exit rate from the same regressions.

Unlike for the job-finding rate, the level of education is an important determinant of the job-exit rate. Individuals with university education have much more stable jobs on average—the job-exit rate of individuals with university education is about 55 percent less than that of individuals with less than primary education.[15]

Result V. The youth are more likely to be separated from a job, while race also matters.

Race and age are also important determinants of job security. Whites and Indian/Asians have a much lower job-exit rate compared to Blacks. On average, Whites have a job-exit rate which is 0.032 less than that of Blacks (almost 60 percent lower). Young individuals have extremely high job-exit rates compared to older people—55 to 64 year olds have a job-exit rate which is 0.06 less than 15 to 24 year olds (75 percent lower).[16] Finally, the sector of employment also determines the job-exit rate with formal sector workers less likely to be separated from their job compared to informal sector workers.[17]

Column 3 of Table 4 repeats the same regressions as column 1 but using the narrow definition of unemployed as opposed to the broad definition. The results are qualitatively similar.[18]

---

our results have several advantages over that of StatsSA (2013). These include: (i) the multivariate probit setting allows us to look at marginal effects after controlling for various other factors while StatsSA (2013) looked at one variable at a time; (ii) StatsSA constructed the panel across two quarters only while we construct a panel from 2008Q1 to 2014Q3, thus giving us a much larger sample.

[15] The job-exit probability of individuals with less than primary education is 0.055, while that of university graduates is 0.025 (Table 2 in Appendix II).

[16] Again we also looked at the interaction of race and education. The job-exit rate was lower for more highly educated individuals within a racial group, and was also lower for whites of the same education level compared to Blacks.

[17] We also control for the contract duration of an employed individual i.e. whether the individual is a temporary or permanent worker. As expected, temporary workers are more likely to transition to unemployment compared to permanent workers.

[18] Essers (2104) looked at the related probability of an individual continuing in formal sector employment (rather than the overall job-exit rate which is the focus of this paper) and finds that mid-aged and more highly educated workers were more likely to remain in regular wage employment.

Result VI. A job in the informal sector improves the probability of getting a formal sector job.

Table 5 reports results for the probit regression where the left hand side variable is a dummy that measures the probability of transitioning into formal sector employment i.e. the LHS variable is the dummy $i_{i,t}$ described above. The table again reports the marginal effects from the probit (and not the raw coefficients). In particular, the rows report the difference between the transition rate of the group in question compared to the excluded group (mentioned in the first column of the table). Appendix II reports the levels of the transition rate from the same regressions.

The key hypothesis that we want to test here is whether the informal sector acts as a stepping-stone into formal sector employment. Therefore, we want to test whether people employed in the informal sector are more likely to transition into formal sector employment as compared to unemployed individuals. To do this, we include a categorical variable which takes three values: 1 for unemployed individuals who have worked before (excluded group in the regression), 2 for unemployed individuals who have no work experience, and 3 for individuals who are currently employed in the informal sector.

As reported in Table 5 column 2, unemployed individuals who have never worked before have a lower probability of transitioning into formal sector employment compared to unemployed individuals who have worked before. Furthermore, people who are working in the informal sector are more likely to transition into formal sector employment than the unemployed. Estimated coefficients suggest that workers in the informal sector have a 55 percent higher probability of transitioning into formal sector employment than the unemployed.[19] Furthermore, people with higher levels of education are more likely to transition into formal sector employment as are Whites compared to the other racial groups.

Column 4 of Table 5 repeats the regression but uses the narrow definition of unemployment. The results are qualitatively the same.

Result VII. During the crisis, unemployment increased due to fall in job-finding rate.

All the regressions for job-finding rate and job-exit rate discussed previously included time dummies although these were not reported in Tables 3 and 4. However, these time dummies are themselves of interest, especially because our sample period includes the GFC which saw unemployment shoot up in South Africa. The coefficients on the time dummies can inform us about the relative contribution of job-exit versus job-finding rates in the rise of unemployment in this period.

Figure 2 plots the job-finding rate and job-destruction rate relative to 2008Q1 i.e. it plots the coefficients on the time dummies from the regressions reported in Table 3 column 1 and

[19] The probability that an unemployed individual (with prior work experience) finds a formal sector job is 0.072 while that for an informal sector worker is 0.112 (Table 3 in Appendix II).

Table 4 column 1 respectively. Perhaps somewhat surprisingly, the job-exit rate remained more or less unchanged during this period, although the job-finding rate declined substantially. Therefore, the increase in unemployment was not because of a large increase in job separations, but rather because it was very hard for the unemployed to become employed again. This result is similar to the finding in the US literature that most of the fluctuation in unemployment over the business cycle is due to fluctuation in job-finding rate, with job-separation rate being acyclical (Hall 2005; Shimer 2005).

Result VIII. There is some evidence of insider-outside dynamics at play, but results need to be interpreted with caution.

South African labor markets are characterized by high levels of trade union membership and collective bargaining agreements. This can give rise to insider-outside dynamics with trade-union members having job-security and above market clearing wages while outsiders find it more difficult to find jobs. We find some descriptive evidence in favor of this hypothesis.

The QLFS started asking employed individuals whether they are trade union members or not beginning in 2010 Q3. To see whether trade union membership affects job security, we include an individual's trade union membership status as an explanatory variable for the job-exit regressions ($d_{i,t}$ on the LHS). The results for this are reported in column 2 of table 4.[20] As can be seen, not being a trade union member reduces job security by increasing the job-exit probability by around 58 percent.[21] Therefore, trade union membership is beneficial for insiders and provides job security, but at the same time outsiders face more uncertainty.

To further explore the effect of trade unions on outsiders, we test whether an unemployed individual who was previously employed in a high trade union density sector finds it more difficult to transition to employment. If skills are industry specific and unemployed individuals are more likely to look for a job in their previous industry of work, then lower job-finding rates for individuals who previously worked in high trade union density industries would be consistent with the hypothesis of trade unions benefiting insiders at the expense of outsiders. We compute the trade union density of each industry at the 1-digit level.[22] We then include the trade union density of an unemployed individual's previous industry as an additional explanatory variable in the job-finding regressions (where $f_{i,t}$ is the RHS variable).[23] We find that the coefficient on this new variable is -0.231 and significantly different from zero at the 1 percent level of significance.[24] This implies that if one has

[20] Note that the sample size for this regression is smaller than for the other regressions in the table as we need to restrict the sample to after 2010Q3.
[21] The job-exit probability of individuals with union membership is 0.031, while that of non-members is 0.049 (Table 2 in Appendix II).
[22] There are 10 industries at the 1-digit level.
[23] We can do this because unemployed individuals in the QLFS report their previous industry of work.
[24] Note that -0.231 is the raw coefficient on the trade union density variable and not the marginal effect. This is because the trade union density variable is not a categorical variable but a continuous one and therefore we cannot compute marginal effects compared to an excluded group. Furthermore, for this regression, we cluster standard errors at the industry-time level to account for the fact that the trade union density variable varies only at the industry level.

previously worked in an industry with high trade union density (but have subsequently become unemployed), it reduces the probability of getting employed in the future.

To further explore this sectoral heterogeneity and get a sense of which industries are driving this result, we plot the job-finding rate of unemployed from different previous industries against the trade union density of that industry in Figure 3.[25] As can be seen, individuals who had worked in high trade union density industries like mining and "electricity, gas and water", have a much lower job-finding rate as compared to people who have previously worked in industries like agriculture, construction or trade.

While this result is interesting, it is important to interpret it with caution. In particular, we cannot claim causality from high trade union density in previous industry to low job-finding rates as industries might differ substantially across a number of other relevant dimensions (specialization for example). However, the strong and clear negative correlation observed in Figure 3 does lend support to the hypothesis that trade unions benefit insiders at the expense of outsiders, including people previously employed by sectors with high trade union density.

B. Aggregate Counterfactuals

So far we have documented how the job-finding rate and job-exit rate differ at the micro level based on individual characteristics. We now do some partial equilibrium calculations to quantify the macro effects of different job-finding and exit rates on the unemployment rate. As an example, we showed in result I, that the inexperienced have a much lower job-finding rate compared to the experienced. However, what is the effect of this lower job-finding rate on aggregate unemployment? To get a good quantitative answer to this question would probably require a structural model which specifies the reason for the difference in job-finding rates between the experienced and inexperienced. Here we take a reduced form approach where we use steady state versions of simple stock-flow equations of labor market transitions to quantify the effects of different job-finding and exit rates on aggregate unemployment.

The methodology we use is based on Cortes et al. (2015) and Shimer (2012). Let $e_t, u_t, and\ n_t$ to be the share of individuals between the age of 15 and 64 who are employed, unemployed, and out of the labor force respectively at time *t*. Define $\rho_{xy,t}$ to be the transition rate from state *x* (employed, unemployed, or out of labor force) to state *y* at time *t* i.e. it is the share of individuals in state *x* who transition to state *y* between time *t* and *t+1*. Therefore the job-finding rate and job-exit rate analyzed above correspond to $\rho_{ue,t}$ and $\rho_{eu,t}$ respectively.

[25] The job-finding rates by previous industry are computed by running the same regressions as reported in Table 3, but by also including the previous industry of the unemployed as an additional RHS variable. We then use this probit regression to compute the job-finding rate of each industry. The advantage of using the probit rather than simply computing the transition probability directly for each previous industry is that this controls for differences in composition of workforce across industries i.e. if one industry has more skilled or older workers than others, then the difference in job-finding rate due to skill or age will not be attributed to the industry when we run the probit.

The evolution of $e_t, u_t, and\ n_t$ over time can be expressed as a set of three equations in matrix notation:

$$\begin{pmatrix} e_{t+1} \\ u_{t+1} \\ n_{t+1} \end{pmatrix} = \begin{pmatrix} \rho_{ee,t} & \rho_{ue,t} & \rho_{ne,t} \\ \rho_{eu,t} & \rho_{uu,t} & \rho_{nu,t} \\ \rho_{en,t} & \rho_{un,t} & \rho_{nn,t} \end{pmatrix} \begin{pmatrix} e_t \\ u_t \\ n_t \end{pmatrix} \qquad (1)$$

Now consider a steady state where the transition rates do not change over time. Using the first two equations above along with the fact that $n = 1 - e - u$, $\rho_{ee} = 1 - \rho_{eu} - \rho_{en}$, $\rho_{uu} = 1 - \rho_{ue} - \rho_{un}$, and $\rho_{nn} = 1 - \rho_{ne} - \rho_{nu}$ we can solve for the steady state values of e and u (where variables without the sub-script *t* represent steady state values). These are given by (see appendix for details[26]):

$$\begin{pmatrix} \rho_{eu} + \rho_{en} + \rho_{ne} & \rho_{ne} - \rho_{ue} \\ \rho_{nu} - \rho_{eu} & \rho_{ue} + \rho_{un} + \rho_{nu} \end{pmatrix} \begin{pmatrix} e \\ u \end{pmatrix} = \begin{pmatrix} \rho_{ne} \\ \rho_{nu} \end{pmatrix}$$

Therefore,

$$\begin{pmatrix} e \\ u \end{pmatrix} = \begin{pmatrix} \rho_{eu} + \rho_{en} + \rho_{ne} & \rho_{ne} - \rho_{ue} \\ \rho_{nu} - \rho_{eu} & \rho_{ue} + \rho_{un} + \rho_{nu} \end{pmatrix}^{\mathbf{-1}} \begin{pmatrix} \rho_{ne} \\ \rho_{nu} \end{pmatrix} \qquad (2)$$

Table 6 gives the average value of $\rho's$ for our entire sample. Substituting into the above equation we get the steady state value of unemployment to be 32.2 percent.[27] This is very close to the average unemployment rate for our pooled sample (i.e. pooling across quarters) which is 31.8 percent.

We can now use equation (2) to conduct counterfactual exercises. In particular, we can use the equations to ascertain what the unemployment rate would be under different assumptions for the job-finding rate and job-exit rate. To continue the example discussed above, the job-finding rate for those who have been employed before is 0.131, while the average for the entire sample is 0.106. What would the counterfactual unemployment rate be if all individuals had the job-finding rate of experienced individuals? This counterfactual unemployment rate is given by

[27] Note that u is not the unemployment rate but rather the share of individuals between the age of 15 and 64 who are unemployed. The unemployment rate is calculated as $\frac{u}{u+e}$.

$$\begin{pmatrix} e^c \\ u^c \end{pmatrix} = \begin{pmatrix} \rho_{eu} + \rho_{en} + \rho_{ne} & \rho_{ne} - \boldsymbol{\rho_{ue}^c} \\ \rho_{nu} - \rho_{eu} & \boldsymbol{\rho_{ue}^c} + \rho_{un} + \rho_{nu} \end{pmatrix}^{-1} \begin{pmatrix} \rho_{ne} \\ \rho_{nu} \end{pmatrix}$$

where we have replaced the job-finding rate with the counterfactual job-finding rate ρ_{ue}^c. Table 7 reports the result for this case. If all individuals had the finding rate of experienced individuals then the unemployment rate would be 28 percent in steady state which is a full 4 percentage points less than average unemployment rate in the sample. On the other hand, if all individuals had a job-finding rate of the inexperienced, then the unemployment rate would have been 39 percent. Therefore, the difference in job-finding rate between the experienced and the inexperienced results in an aggregate unemployment rate difference of 11 percentage points.

Similarly, the difference in job-finding rates between the long-term unemployed and the short-term unemployed translates into large differences in aggregate unemployment rates in the counterfactual. If all individuals had the job-finding rate of the short-term unemployed, then the aggregate unemployment rate would be 25 percent in steady state versus 35 percent if everyone had the job-finding rate of long term unemployed.

We can do similar counterfactuals for the job-exit rate. As noted in result IV, the job-exit rate differs significantly by education level. We conduct a counterfactual where we assume everyone has one higher level of education and the corresponding lower job-exit rate. The counterfactual unemployment rate in this case is 30 percent. If we make the extreme assumption that all individuals have the job-exit rate of college graduates, then the counterfactual unemployment rate is 24 percent.

C. Cross-country Comparison of Job Finding and Exit Rates

How does the job-finding and job-exit rate in South Africa compare to that in other countries around the world? We used individual level quarterly panel data to estimate job-finding and job-exit rates in South Africa in the previous two sub-sections. However, such data is not available for a wide range of countries. Elsby et al. (2013) and Shimer (2012) discuss how to estimate finding and exit rates using aggregate data on unemployment duration without the need to follow individuals over time. The ILO has implemented this methodology for a wide range of countries and has made the results available as part of the KLIM dataset.[28] We use this dataset to compare the job-finding and exit rate of South Africa to that of a selected and varied group of countries.

Note that the finding and exit rates from the KLIM dataset are not strictly comparable to those we estimated in the previous two sub-sections due to difference in methodology and also because KLIM reports monthly job-finding and exit rates while we estimate quarterly transition rates. Nevertheless, it is reassuring to see that our estimates are in the same range

[28] Note that in the ILO KILM dataset, what we call job-finding rates are called outflow rate and job-exit rates are called inflow rate.

as in the KLIM dataset. In particular, the ratio of job-finding to exit rate that we estimate is 2.5 (from Table 6), as compared to 3.1 in the KLIM data.

Figure 4 plots the job-finding and exit rates from the KLIM dataset for South Africa and a selection of other countries averaged between 2001 and 2012. As can be seen, the job-finding rate in South Africa is one of the lowest in these countries. Only Romania and Turkey have lower job-finding rates. However, the countries with lower job finding rates also have much lower job-exit rates compared to South Africa. South Africa has the lowest ratio of job-finding to job-exit rate in our sample.

IV. Analysis of Inequality

There are two features of the South African economy that are extreme when compared to its peers, namely the very high unemployment rate and inequality. While the previous analysis has focused exclusively on labor market dynamics and unemployment, we now turn to looking at the potential relation between unemployment and inequality.

In particular, we want to answer the question: How much lower would inequality be if the unemployment rate in South Africa was lower? To answer this question, we do some simple partial equilibrium simulation exercises using data from the third wave of the National Income Dynamics Study (NIDS) conducted in 2012. While the partial equilibrium nature of the exercise is a limitation, we view this as a back-of-the-envelope calculation aimed at assessing the potential relation between unemployment and inequality.

NIDS asks detailed questions regarding labor market outcomes and income levels of individuals surveyed. We use the NIDS instead of the QLFS to look at inequality because of the detailed question on individual income which is available in the NIDS but not in the QLFS.[29] In particular, the NIDS asks individuals to report their post-tax income from various sources (main job, self-employment, casual jobs etc.), as well as income from various government grant schemes (state old age pension, disability grant, child support grant, unemployment insurance etc.). While three waves of the NIDS have been conducted so far, we only use the third wave for our analysis which covered 8,040 households with a total of 32,633 interviewed individuals.

For our baseline results, we use an individual level measure of income inequality. We restrict our sample to individuals in the labor force and assign each individual their personal income reported in NIDS. This is a post-tax measure of income and includes transfers (from government) received by the individual.

We conduct the following simulation to ascertain how much lower income inequality would be if unemployment was lower in South Africa:

[29] The Consumer Expenditure Survey could be another source of income and expenditure data for South Africa. However, individuals in this survey are not asked to report their employment status, therefore we cannot use it to look at the relation between unemployment and inequality.

1) We first simply compute income inequality in the NIDS sample.
2) We then assume that a randomly selected subsample of the unemployed becomes employed.[30]
3) We assign these randomly selected individuals an income level based on quantile regressions for the employed. In particular, we run quantile regressions of log income on age, education, race and sex and assign the predicted income level (if it is higher than their income as unemployed) from this regressions to those randomly selected to become employed.[31]
4) We recomputed inequality based on this simulated income series where the people moved from unemployment to employment have a higher level of income.

The results for this simulation exercise are presented in the right panel of Figure 5. On the x-axis we have the simulated unemployment rate and on the y-axis the simulated Gini coefficient. We use two quantile regressions when assigning income to the randomly selected group of unemployed who become employed. The solid blue line reports results when using the 25th percentile income level and the dotted red line reports results when using the median income level.

In the NIDS, the unemployment rate is around 30 percent and the individual level Gini coefficient is 0.68 (top right point in the graph). As we simulate lower levels of unemployment, inequality declines (with the decline being more when using the median as opposed to the 25th percentile). Quantitatively, a decline in the unemployment rate of 10 percentage points leads to inequality falling from 0.68 to 0.645 when using the median.[32]

To allow for risk sharing within households, inequality is often measured at the household level using per-capita income level of the household as opposed to individual level income. We also consider this case. The sample now is not restricted to people in the labor force but includes all individuals in the survey. We compute per-capita household income (including

---

[30] An alternative to random selection would be to use the job-finding rate regression results (presented in Table 3) to decide the probability of an individual being employed. However, such an exercise would neglect the effect of different job-exit rates for various demographic groups on the equilibrium unemployment rate for these groups. Random selection on the other hand maintains the steady state distribution of unemployment across demographic groups in the counterfactual. It achieves an 'x' percent (not percentage points) reduction in overall unemployment by reducing the unemployment rate within each demographic group by the same 'x' percent. Therefore, with random selection, the simulations in effect produce new steady states where the relative unemployment rate between different demographic groups remains unchanged.

[31] Note that although we allow for income levels to differ for people with different education levels, race, age, and gender, we do not allow for any dispersion in income within a race, age, education, gender cell. This will result in lower levels of inequality in the counterfactual i.e. the fall in inequality may be overstated. However, we did robustness checks where we gave the newly employed the same distribution of income as the employed. While this thought experiment does not suffer the bias described above, it implies that the unemployed who we switch to employment get the same mean income as the currently employed, which might be too optimistic an assumption. This thought experiment results in larger drops in inequality than those reported in Figures 6.

[32] Note that these results are of a static nature. In a dynamic setting, the lower unemployment rate is likely to lead to further increases in income as the newly employed gain experience and skills. These dynamic effects may very well lead to larger declines in inequality that our counterfactual cannot capture.

imputed rent) by adding income across all household members and compute the Gini coefficient for this measure of household income.

In the simulation stage, we proceed exactly as before: we randomly select a subset of the unemployed and assign them income based on quantile regressions. We then recomputed household per-capita income for this simulated series and then re-compute the Gini.

The results are shown in the left panel of Figure 5. Inequality based on per-capita household income was 0.665 in the NIDS in 2012. A 10 percentage point drop in unemployment reduced the Gini to 0.645 in our simulations when using median incomes.[33]

It is hard to get a sense of economic magnitude from the results presented in Figure 5 i.e. is a fall in Gini from 0.665 to 0.645 large or small? To get a sense of economic magnitude, we ask the following question: By what percent will the government need to increase all transfer programs to achieve similar reductions in inequality i.e. how much additional redistribution is required to achieve the same fall in inequality associated with a fall in unemployment?

The rich information in the NIDS data regarding income from government grants allows us to conduct this thought experiment. We simply scale up the income from government grants for all individuals and then recomputed household per-capita income and the associated Gini. The results are presented in Figure 6. On the x-axis we have the percentage increase in transfer and on the y-axis the simulated Gini. When transfers are increased by zero percent, then of course we get the same Gini coefficient as in the data (0.665). As we increase the transfers by larger amounts, the computed Gini falls. To achieve a reduction in Gini from 0.665 to 0.645 would require an increase in transfers of about 40 percent, which was equivalent to 3.3 percent of GDP in 2012 or about 11.1 percent of government expenditure. Therefore, it would be prohibitively expensive to use redistributive policy alone to achieve the same fall in inequality that might arise from a fall in unemployment of 10 percentage points.

V. CONCLUSION

The analysis confirms some of our initial hypotheses about the determinants of high unemployment in South Africa. Large skill mismatches, poor educational outcomes, and the apartheid legacies have hurt job growth and perpetuated inequality. Unemployment, especially amongst youth, women, and blacks, has remained high.

While improving the quality of education remains key to address the long-term unemployment challenge, our analysis suggests that till that happens (for education to be a reliable signal of productivity) policies aimed specifically at providing experience to young

[33] Van der Berg (2010) does a similar exercise. They find that 2½ million additional jobs would reduce the Gini coefficient by only about 0.033, but would reduce the poverty headcount ratio by almost 9 percentage points. In contrast, an average wage increase of as much as 30% would only reduce the poverty headcount by about 4 percentage points, while leaving the Gini coefficient slightly higher (0.011 points).

first-time entrants to the labor force will be important to improve their employability. Therefore, implementation of programs, such as Employment Tax Incentive Act, which aims to provide necessary on-the-job training and work experience to youth should be beneficial. Also, we find that though creating informality may not be ideal, it may be a stepping stone into formal employment.

Moreover, as illustrated by our static exercise, reducing unemployment remains probably a more viable and sustainable way of reducing income inequality compared to redistribution alone (though redistribution has played a very important role in reducing poverty).

REFERENCES

Anand, Rahul, Roberto Perrelli, and Boyang Zhang. 2016. "South Africa's Exports Performance: Any Role for Structural Factors?" IMF Working Paper No. 16/24. Washington: International Monetary Fund.

Banerjee, Abhijit, Sebastian Galiani, James A. Levinsohn, Zoë McLaren, and Ingrid Woolard. 2008. "Why has unemployment risen in the New South Africa?" Economics of Transition 16 (4): 715–740.

Butcher, Kristin F., and Cecilia E. Rouse. 2001. "Wage Effects of Unions and Industrial Councils in South Africa." Policy Research Working Paper Series 2520. Washington: World Bank.

Burger, Philippe. 2015. "Wages, Productivity and Labour's declining income share in post-apartheid South Africa." South African Journal of Economics 83 (2): 159-173.

Cortes, Guido Matias, Nir Jaimovich, Christopher J. Nekarda, and Henry E. Siu. 2014. "The Micro and Macro of Disappearing Routine Jobs: A Flows Approach." NBER Working Paper No. 20307.

Elsby, Michael, Bart Hobijn, and Aysegul Sahin. 2013. "Unemployment Dynamics in the OECD." The Review of Economics and Statistics, MIT Press 95 (2): 530-548.

Essers, Dennis. 2014. "South African Labour Market Transitions During the Global Financial and Economic Crisis." UNU-WIDER Working Paper 115/2014. United Nations University: World Institute for Development Economics Research.

Hall, Robert E.. 2005. "Employment Fluctuations With Equilibrium Wage Stickiness." American Economic Review 95(1): 50-65.

Kerr, Andrew, Martin Wittenberg, and Jairo Arrow. 2013. "Job Creation and Destruction in South Africa." SALDRU Working Paper Number 92. Cape Town: Southern Africa Labour and Development Research Unit, University of Cape Town.

Kingdon, Geeta, and John Knight. 2004. "Unemployment in South Africa: The Nature of the Beast." World Development 32(3): 391–408.

Kingdon, Geeta, and John Knight. 2009. "Unemployment: South Africa's Achilles' Heel." In South African Economic Policy Under Democracy edited by Janine Aron, Brian Kahn, and Geeta Kingdon. Oxford: Oxford University Press.

Leung, Ronald, Marco Stampini, and Désiré Vencatachellum. 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.

Leibbrandt, Murray, Ingrid Woolard, Arden Finn, and Jonathan Argent. 2009. “Trends in South Africa Income Distribution and Poverty Since the Fall of Apartheid.” OECD Social, Employment and Migration Working Papers No. 101. OECD: Paris.

Levinsohn, James. 2008. “Two Policies to Alleviate Unemployment in South Africa.” CID Working Paper No. 166. Harvard University.

Magruder, Jeremy 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.

Michaud, Pierre-Carl, and Désiré Vencatachellum. 2003. “Human Capital Externalities in South Africa.” Economic Development and Cultural Change 51(3): 603-628.

Ngarachu, Maria, Volker Schöer, and Axel Schimmelpfennig. 2016. “The Costly Road to Work? Wages and Transport Costs in South Africa.” IMF Working Paper (forthcoming). Washington: International Monetary Fund.

National Treasury. 2011. “Confronting Youth Unemployment: Policy Options for South Africa”.

Ranchhod, Vimal, and Taryn Dinkelman. 2007. “Labour Market Transitions in South Africa: What Can We Learn from Matched Labour Force Survey Data?” SALDRU Working Paper No. 14. Cape Town: SALDRU, University of Cape Town.

Shimer, Robert. 2005. “The Cyclical Behavior of Equilibrium Unemployment and Vacancies.” American Economic Review 95 (1): 25-49.

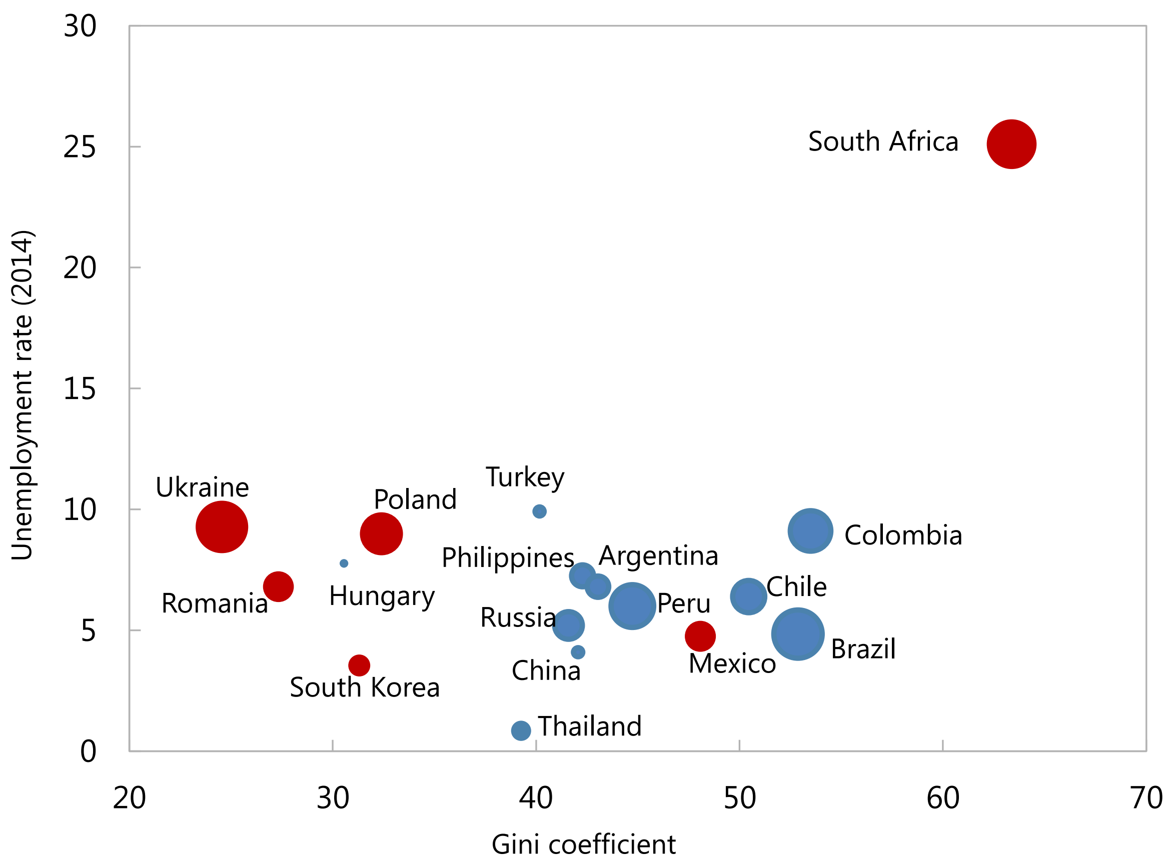
Shimer, Robert. 2012. "Reassessing the Ins and Outs of Unemployment." Review of Economic Dynamics 15(2): 127-148.

Statistics South Africa. 2013. “Labour Market Dynamics in South Africa.”

Van der Berg, Servaas. 2010. “Current Poverty and Income Distribution in the Context of South African History.” Department of Economics Working Papers 22/2010. South Africa: Stellenbosch University.

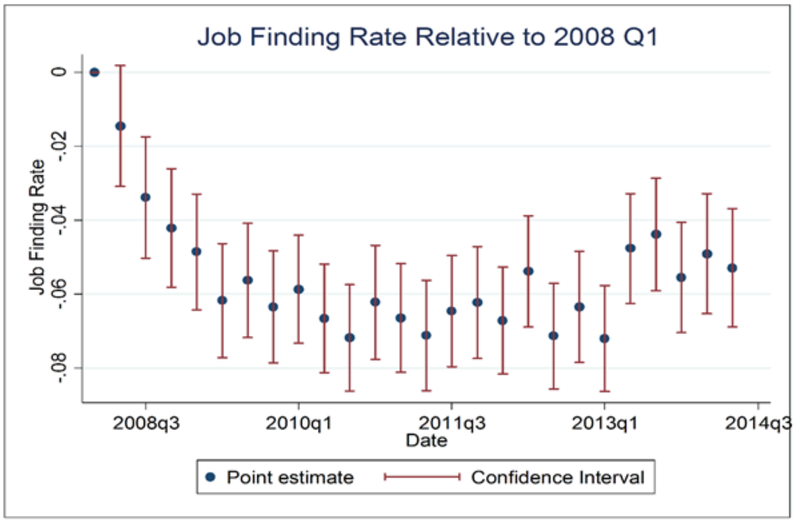
Verick, Sher. 2012. “Giving Up Job Search During a Recession: The Impact of the Global Financial Crisis on the South African Labour Market.” Journal of African Economies 21(3): 373-408.

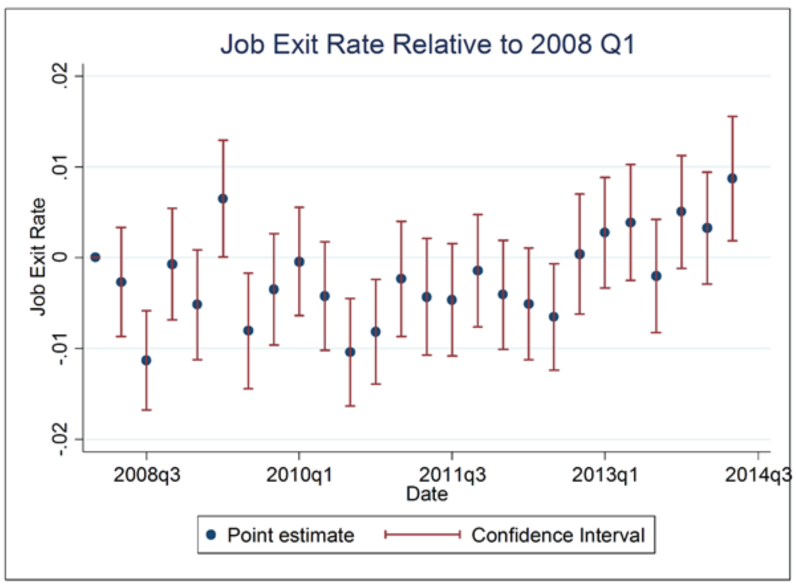
Figure 1. Unemployment versus Inequality Across Countries



Sources: WEO; and World Bank World Development Indicators
1/ Blue represents a decrease and red an increase in unemployment between 2008 and 2014. The size of the bubble illustrates the magnitude of the change in unemployment.

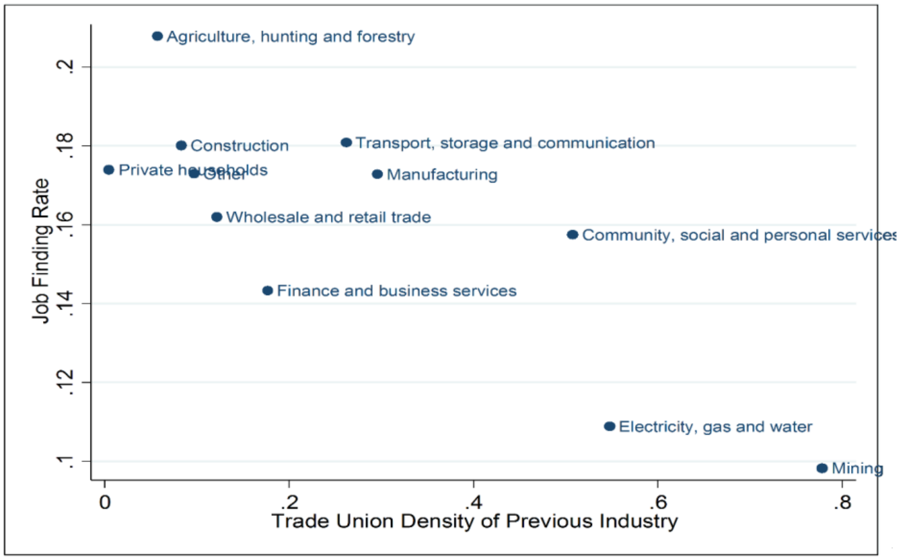
Figure 2. Finding and Exit Rate Over Time





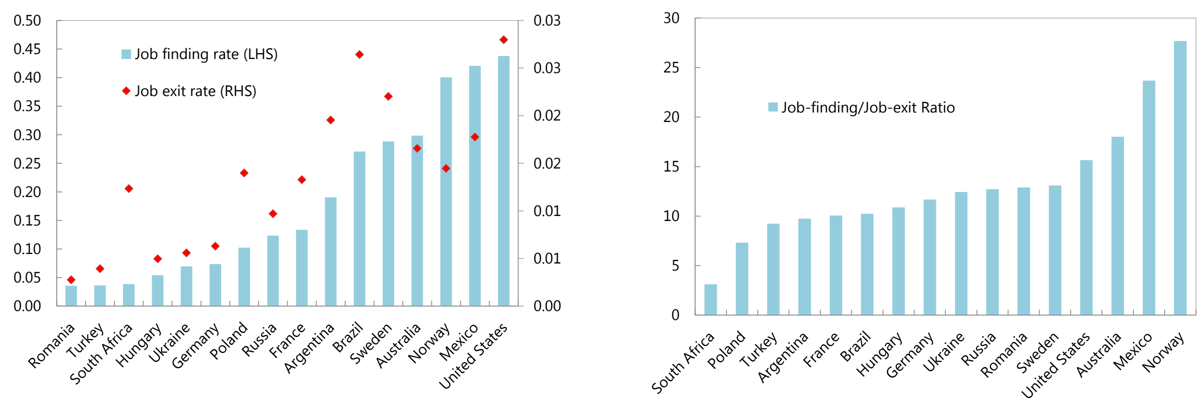
Notes: Uses data from QLFS 2008Q1 to 2014Q3. Plots the job finding rate and job exit rate relative to 2008Q1. The rates are recovered from probit regressions reported in column 1 of Tables 3 and 4 respectively.

Figure 3. Job Finding Rate vs. Trade Union Density of Previous Industry



Notes: Uses data from QLFS 2010Q3 to 2014Q3. Plots the finding rate of unemployed individuals by their previous industry employment against the trade union density of the previous industry.

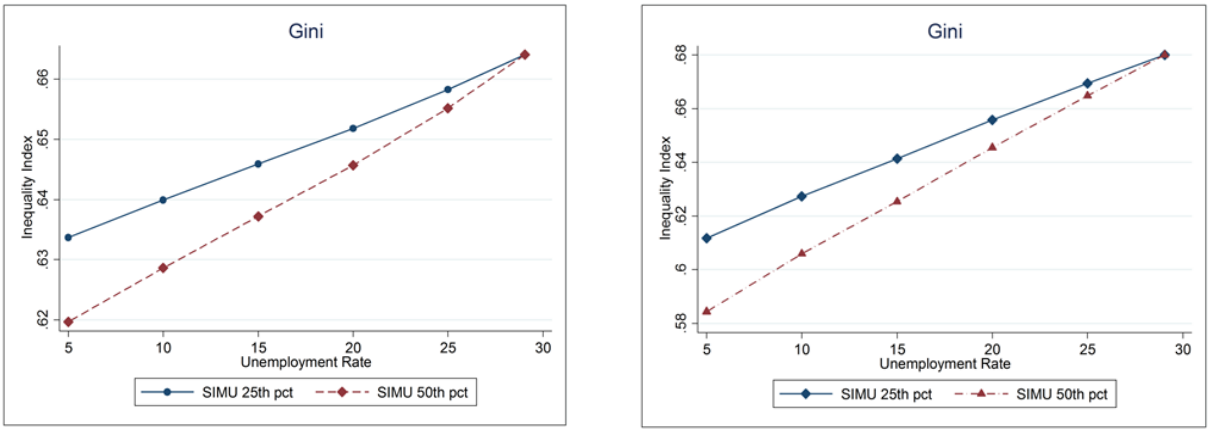
Figure 4: Cross-country Comparison of Job-Finding and job-Exit Rate



Note: Figures are based on ILO-KILM data and represent average of 2001-2012 wherever data is available.

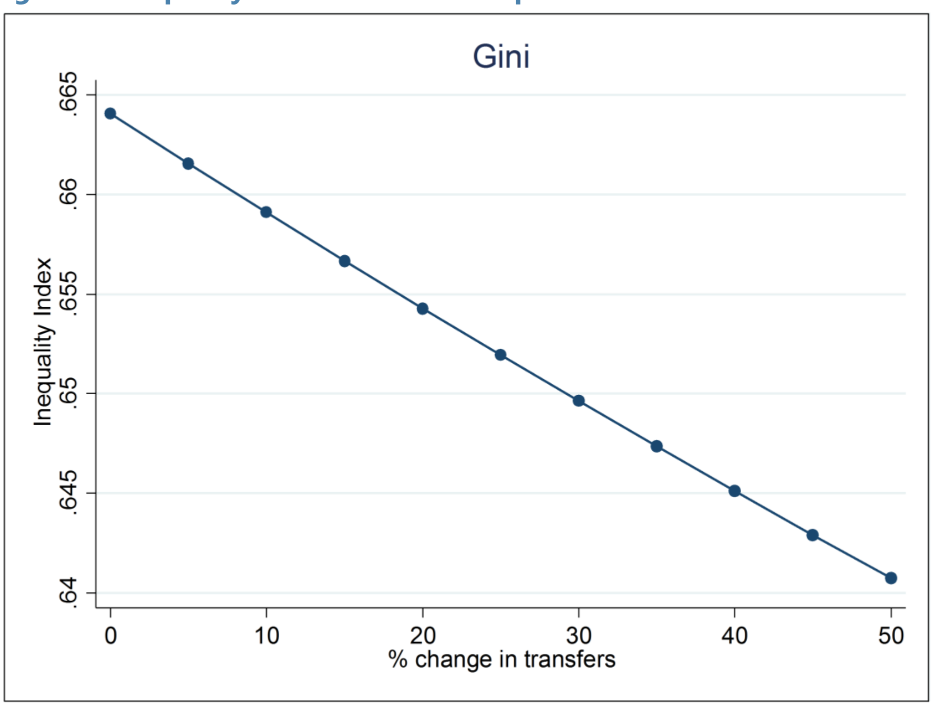
Figure 5. Inequality for Different Unemployment Rates

Household per-capita income *Individual Income*



Notes: Uses data from NIDS wave 3. Plots the Gini coefficient for different simulated levels of the unemployment rate. The left panel uses household per-capita income to measure inequality while the right panel uses individual income to compute inequality.

Figure 6. Inequality for Different Assumptions on Government Transfers



Notes: Uses data from NIDS wave 3. Plots the Gini coefficient for different levels of government transfers.

Table 1. Comparison of Matching Algorithm to Official Panel—2013Q3 to 2013Q4

		Matched by Algorithm	
		No	Yes
Matched by StatsSA	No	33.98	0.36
	Yes	2.95	62.7

Notes: Uses data for 2013Q3 and 2013Q4 of the QLFS. Reports the percentage of the sample that is matched or not by the algorithm described in section II A and the official panel released by StatsSA.

Table 2. Comparison of Full Sample to Matched Sample

		Full Sample	Matched to next quarter only
Gender	Male	48.9	48.31
	Female	51.1	51.69
Race	African/Black	78.34	78.6
	Coloured	9.36	9.1
	Indian/Asian	2.75	2.76
	White	9.54	9.54
Education	Less than primary	14.25	14.66
	Primary but not secondary	49.42	49.3
	Secondary	24.98	24.51
	University	10.41	10.6
	Other	0.95	0.92
Age	15 to 24	29.88	28.48
	25 to 34	26.52	25.8
	35 to 44	20.35	20.88
	45 to 54	14.14	14.99
	55 to 64	9.11	9.85
Employment Status	Employed	43.01	43.39
	Unemployed	13.87	13.42
	Discouraged job seeker	5.94	5.93
	Other not economically active	37.19	37.26

Notes: Uses data from QLFS 2008Q1 to 2014Q3. Reports the share of the population by different individual characteristics. The column marked “Full Sample” reports the share for the entire sample. The column labeled “Matched to next quarter only” reports the share for when only considering the observations which our algorithm matched to an observation in the next quarter.

Table 3. Transition Regression: Job-Finding Rate (dy/dx)

	(1)	(2)	(3)	(4)
Unemployment Defintion	Broad	Broad	Narrow	Narrow
Education (<primary)				
Primary but not secondary	-0.005*	-0.005*	-0.007*	-0.007*
Secondary but not university	-0.005	-0.004	-0.009*	-0.008*
University or more	0.003	0.004	0.003	0.003
Other	0.017	0.016	-0.006	-0.007
Race (African/Black)				
Colored	0.000	0.000	-0.004	-0.004
Indian/Asian	-0.008	-0.008	-0.008	-0.008
White	-0.013*	-0.013*	-0.003	-0.003
Gender (Male)				
Female	-0.031***	-0.032***	-0.036***	-0.036***
Age (15-24)				
Age 25-34	0.025***		0.027***	
Age 35-44	0.030***		0.035***	
Age 45-54	0.021***		0.029***	
Age 55-64	-0.004		0.003	
Experience (Emp before)				
Never employed	-0.061***		-0.052***	
Age (15-24, employed)				
Age 15-24, not employed		-0.070***		-0.059***
Age 25-34, employed		0.015***		0.020***
Age 25-34, not employed		-0.044***		-0.031***
Age 35-44, employed		0.021***		0.029***
Age 35-44, not employed		-0.037***		-0.019**
Age 45-54, employed		0.010**		0.020***
Age 45-54, not employed		-0.026***		0.019
Age 55-64, employed		-0.020***		-0.009
Age 55-64, not employed		-0.038**		0.033
Unemp duration (>=1 yr)				
Short term unemployed(<1 year)	0.072***	0.072***	0.078***	0.078***
Observations	173,626	173,626	105,887	105,887

Notes: Uses data from the QLFS 2008Q1 to 2014Q3. The dependent variable is a dummy which takes value 1 if an individual transtions from unemployment to employment ($f_{i,t}$). The figures reported are the marginal effects from probit regressions relative to the excluded group (which is noted in brackets in the first column).Columns 1 and 2 use the broad definition of unemployment to define the LHS variable while 3 and 4 use the narrow definition. Columns 1 and 3 include age and experience (ever worked before) as separate variables, while columns 2 and 4 include the interaction between the two. Standard errors were clustered at the household level. We do not report standard errors but indicate whether a coefficient was statistically significantly different from zero. *** p<0.01, ** p<0.05, * p<0.1

Table 4. Transition regression: Job-Exit Rate (dy/dx)

	(1)	(2)	(3)	(4)
Unemployment Defintion	Broad	Broad	Narrow	Narrow
Sample	Full	Trade Union	Full	Trade Union
Education (<primary)				
Primary but not secondary	-0.001	-0.001	0.004***	0.004*
Secondary but not university	-0.013***	-0.012***	-0.002	-0.003
University or more	-0.030***	-0.028***	-0.015***	-0.014***
Other	-0.006	-0.007	-0.001	-0.002
Race (African/Black)				
Colored	-0.002	0.003	-0.003*	0.000
Indian/Asian	-0.028***	-0.026***	-0.016***	-0.015***
White	-0.032***	-0.028***	-0.022***	-0.019***
Age (15-24)				
Age 25-34	-0.025***	-0.020***	-0.014***	-0.012***
Age 35-44	-0.039***	-0.034***	-0.025***	-0.023***
Age 45-54	-0.050***	-0.044***	-0.033***	-0.030***
Age 55-64	-0.060***	-0.056***	-0.041***	-0.039***
Gender (Male)				
Female	-0.001	-0.002*	-0.006***	-0.006***
Sector (Formal sector)				
Informal sector	0.024***	0.017***	0.014***	0.010***
Private households	0.003**	0.000	0.003**	0.000
Type of Job (Limited Duration)				
Permanent	-0.077***	-0.066***	-0.055***	-0.046***
Unspecified Duration	-0.026***	-0.023***	-0.020***	-0.017***
Not Applicable	-0.049***		-0.039***	
Trade Union (member)				
Not a member		0.018***		0.013***
Don't know		0.016***		0.013***
Observations	328,063	174,683	328,065	174,683

Notes: Uses data from the QLFS 2008Q1 to 2014Q3. The dependent variable is a dummy which takes value 1 if an individual transtions from employment to unemployment ($d_{i,t}$). The figures reported are the marginal effects from probit regressions relative to the excluded group (which is noted in brackets in the first column).Columns 1 and 2 use the broad definition of unemployment to define the LHS variable while 3 and 4 use the narrow definition. Columns 2 and 4 include whether the individual is a trade union member or not. As the question of trade union membership was only asked after 2010Q3 in the QLFS, these columns use a smaller sample. Standard errors were clustered at the household level. We do not report standard errors but indicate whether a coefficient was statistically significantly different from zero. *** p<0.01, ** p<0.05, * p<0.1

Table 5. Transition regression: Unemployment/Informal to formal employment (dy/dx)

	(1)	(2)	(3)	(4)
Unemployment Defintion	Broad	Broad	Narrow	Narrow
Unemployed				
Work in Informal sector	0.052***		0.049***	
Unemployed with experience				
Unemployed never worked		-0.039***		-0.040***
Work in Informal sector		0.040***		0.036***
Education (<primary)				
Primary but not secondary	0.013***	0.013***	0.018***	0.017***
Secondary but not university	0.036***	0.038***	0.044***	0.046***
University or more	0.071***	0.071***	0.083***	0.085***
Other	0.025***	0.025***	0.030***	0.029***
Race (African/Black)				
Colored	0.028***	0.023***	0.026***	0.022***
Indian/Asian	0.043***	0.039***	0.048***	0.044***
White	0.080***	0.074***	0.099***	0.093***
Gender (Male)				
Female	-0.042***	-0.040***	-0.050***	-0.048***
Age (15-24)				
Age 25-34	0.018***	0.005**	0.012***	-0.001
Age 35-44	0.011***	-0.007***	0.006**	-0.012***
Age 45-54	-0.001	-0.020***	-0.007***	-0.025***
Age 55-64	-0.013***	-0.030***	-0.019***	-0.034***
Observations	270,203	270,203	196,021	196,021

Notes: Uses data from the QLFS 2008Q1 to 2014Q3. The dependent variable is a dummy which takes value 1 if an individual transtions from unemployment or informal employment to formal employment ($i_{i,t}$). The figures reported are the marginal effects from probit regressions relative to the excluded group (which is noted in brackets in the first column).Columns 1 and 2 use the broad definition of unemployment to define the LHS variable while 3 and 4 use the narrow definition. Columns 1 and 3 include a RHS variable for whether the individual was unemployed or in the informal sectoe. Columns 2 and 4 include a RHS variable of whether an individual was unemployed with no experience, unemployed with experience, or in the informal sector. Standard errors were clustered at the household level. We do not report standard errors but indicate whether a coefficient was statistically significantly different from zero. *** p<0.01, ** p<0.05, * p<0.1*** p<0.01, ** p<0.05, * p<0.1

Table 6. Transition Probabilities

		t+1		
		Employed	Unemployed	OLF
t	**Employed**	0.931	0.048	0.020
	Unemployed	0.106	0.755	0.138
	OLF	0.025	0.095	0.880

Notes: Use data from the QLFS 2008Q1 to 2014Q3. Reports the transition probability from states along the rows to states along the column.

Table 7. Results of Counterfactuals

Counterfactuals	
Finding Rate	
	Unemployment
Original	0.32
Experience	
If everyone has some experience before	0.28
If no one has experience before	0.39
Education	
If everyone moves to the next level of education	0.32
If all are graduates	0.31
Unemployment duration	
If everyone was temporarily unemployed	0.25
If everyone was long term unemployed	0.35
Exit Rate	
Education	
If everyone moves to the next level of education	0.30
If all are graduates	0.24

Notes: Uses data from the QLFS 2008Q1 to 2014Q3. The table reports results for the counterfactual exercises described in Sec III B. The top panel reports results for when we change the job-finding rate, while the bottom panel reports results for changes in the job-exit rate.

Appendix I: Derivation of Counterfactual Equations

To get intuition on how the stock flow equations work, consider the case where there are no transitions from out of the labor force into employment and unemployment. In this case, we have two states e and u which are the share of people in the labor force who are employed and unemployed respectively.[34] These evolve according to:

$$\begin{pmatrix} e_{t+1} \\ u_{t+1} \end{pmatrix} = \begin{pmatrix} \rho_{ee,t} & \rho_{ue,t} \\ \rho_{eu,t} & \rho_{uu,t} \end{pmatrix} \begin{pmatrix} e_t \\ u_t \end{pmatrix}$$

where $e_t = 1 - u_t$, $\rho_{ee,t} = 1 - \rho_{eu,t}$ and $\rho_{uu,t} = 1 - \rho_{ue,t}$.

Substituting these into the equations and assuming a steady state we get

$$\begin{pmatrix} 1-u \\ u \end{pmatrix} = \begin{pmatrix} 1-\rho_{eu} & \rho_{ue} \\ \rho_{eu} & 1-\rho_{ue} \end{pmatrix} \begin{pmatrix} 1-u \\ u \end{pmatrix}$$

The second equation simplifies to

$$u = \rho_{eu}(1-u) + (1-\rho_{ue})u$$
$$\Rightarrow u = \frac{\rho_{eu}}{\rho_{eu} + \rho_{ue}} = \frac{d}{d+f}$$

Therefore, the steady state unemployment rate is an increasing function of the destruction rate and a decreasing function of the finding rate.

For this simplified case, the counterfactual unemployment rate when considering a different job-finding rate is simply given by:

$$u^c = \frac{\rho_{eu}}{\rho_{eu} + \rho_{ue}^c} = \frac{d}{d+f^c}$$

When we have three states as in the text, the unemployment rate cannot be written in the form shown above, but rather takes the form in equation (2). A short derivation is provided below.

The steady state version of equation (1) is

$$\begin{pmatrix} e \\ u \\ 1-u-e \end{pmatrix} = \begin{pmatrix} \rho_{ee} & \rho_{ue} & \rho_{ne} \\ \rho_{eu} & \rho_{uu} & \rho_{nu} \\ \rho_{en} & \rho_{un} & \rho_{nn} \end{pmatrix} \begin{pmatrix} e \\ u \\ 1-u-e \end{pmatrix}$$

The first two equations simplify and can be written as

$$(1-\rho_{ee}-\rho_{ne})e + (\rho_{ne}-\rho_{ue})u = \rho_{ne}$$
$$(\rho_{nu}-\rho_{eu})e + (1-\rho_{uu}-\rho_{nu})u = \rho_{nu}$$

[34] Note that in the text, because we had three states, the population included all individuals between the age of 15 and 64. Now that we are assuming only 2 states, we restrict the sample to the labor force and therefore redefine *e* and *u* accordingly.

Using the fact that $\rho_{ee} = 1 - \rho_{eu} - \rho_{en}$ and $\rho_{uu} = 1 - \rho_{ue} - \rho_{un}$ and putting into matrix notation again we get

$$\begin{pmatrix} \rho_{eu} + \rho_{en} + \rho_{ne} & \rho_{ne} - \rho_{ue} \\ \rho_{nu} - \rho_{eu} & \rho_{ue} + \rho_{un} + \rho_{nu} \end{pmatrix} \begin{pmatrix} e \\ u \end{pmatrix} = \begin{pmatrix} \rho_{ne} \\ \rho_{nu} \end{pmatrix}$$

Appendix II: Regression Results in Levels

Appendix Table 1. Transition Regression: Job-Finding Rate (Levels)

	(1)	(2)	(3)	(4)
Unemployment Defintion	Broad	Broad	Narrow	Narrow
Education				
Less than primary	0.110***	0.110***	0.126***	0.126***
Primary but not secondary	0.106***	0.105***	0.119***	0.118***
Secondary but not university	0.105***	0.106***	0.117***	0.117***
University or more	0.114***	0.114***	0.129***	0.129***
Other	0.127***	0.126***	0.119***	0.119***
Race				
African/Black	0.107***	0.107***	0.120***	0.120***
Colored	0.107***	0.107***	0.116***	0.116***
Indian/Asian	0.100***	0.099***	0.113***	0.112***
White	0.094***	0.094***	0.117***	0.117***
Gender				
Male	0.123***	0.123***	0.137***	0.137***
Female	0.092***	0.092***	0.101***	0.101***
Age				
Age 15-24	0.088***		0.098***	
Age 25-34	0.114***		0.125***	
Age 35-44	0.118***		0.133***	
Age 45-54	0.109***		0.127***	
Age 55-64	0.084***		0.101***	
Experience				
Employed Before	0.131***		0.138***	
Never employed	0.071***		0.086***	
Age				
Age 15-24, employed		0.123***		0.124***
Age 15-24, not employed		0.053***		0.065***
Age 25-34, employed		0.138***		0.144***
Age 25-34, not employed		0.079***		0.093***
Age 35-44, employed		0.144***		0.153***
Age 35-44, not employed		0.087***		0.104***
Age 45-54, employed		0.133***		0.144***
Age 45-54, not employed		0.097***		0.143***
Age 55-64, employed		0.103***		0.115***
Age 55-64, not employed		0.085***		0.157***
Unemp duration				
Long term unemployed (>=1 yr)	0.088***	0.089***	0.091***	0.091***
Short term unemployed(<1 yr)	0.160***	0.160***	0.169***	0.169***
Observations	173,626	173,626	105,887	105,887

dependent variable is a dummy which takes value 1 if an individual transtions from unemployment to employment ($f_{i,t}$). The figures reported are the levels of the transition probability from the corresponding probit regression (the levels counterpart of Table 3 in the main text). Columns 1 and 2 use the broad definition of unemployment to define the LHS variable while 3 and 4 use the narrow definition. Columns 1 and 3 include age and experience (ever worked before) as separate variables, while columns 2 and 4 include the interaction between the two. Standard errors were clustered at the household level. We do not report standard errors but indicate whether a coefficient was statistically significantly different from zero. *** p<0.01, ** p<0.05, * p<0.1

Appendix Table 2. Transition Regression: Job-Exit Rate (Levels)

	(1)	(2)	(3)	(4)
Unemployment Defintion	Broad	Broad	Narrow	Narrow
Sample	Full	Trade Union	Full	Trade Union
Education				
Less than primary	0.055***	0.053***	0.033***	0.032***
Primary but not secondary	0.055***	0.053***	0.037***	0.036***
Secondary but not university	0.043***	0.041***	0.031***	0.030***
University or more	0.025***	0.025***	0.019***	0.019***
Other	0.049***	0.046***	0.032***	0.030***
Race				
African/Black	0.051***	0.048***	0.035***	0.033***
Colored	0.049***	0.051***	0.032***	0.033***
Indian/Asian	0.024***	0.022***	0.019***	0.018***
White	0.020***	0.020***	0.013***	0.014***
Gender				
Male	0.048***	0.047***	0.035***	0.034***
Female	0.047***	0.045***	0.030***	0.028***
Age				
Age 15-24	0.082***	0.074***	0.054***	0.050***
Age 25-34	0.057***	0.054***	0.040***	0.038***
Age 35-44	0.043***	0.040***	0.029***	0.027***
Age 45-54	0.032***	0.030***	0.021***	0.020***
Age 55-64	0.022***	0.018***	0.013***	0.011***
Sector				
Formal sector	0.042***	0.044***	0.029***	0.030***
Informal sector	0.066***	0.061***	0.044***	0.040***
Private households	0.045***	0.044***	0.032***	0.030***
Type of Job				
Limited Duration	0.098***	0.087***	0.070***	0.061***
Permanent	0.021***	0.022***	0.015***	0.016***
Unspecified Duration	0.072***	0.065***	0.050***	0.044***
Not Applicable	0.049***		0.031***	
Trade Union				
Member		0.031***		0.021***
Not a member		0.049***		0.033***
Don't know		0.047***		0.033***
Observations	328,063	174,683	328,065	174,683

Notes: Uses data from the QLFS 2008Q1 to 2014Q3. The table reports results from a regression where the dependent variable is a dummy which takes value 1 if an individual transtions from employment to unemployment ($d_{i,t}$). The figures reported are the levels of the transition probability from the corresponding probit regression (the levels counterpart of Table 4 in the main text). Columns 1 and 2 use the broad definition of unemployment to define the LHS variable while 3 and 4 use the narrow definition. Columns 2 and 4 include whether the individual is a trade union member or not. As the question of trade union membership was only asked after 2010Q3 in the QLFS, these columns use a smaller sample. Standard errors were clustered at the household level. We do not report standard errors but indicate whether a coefficient was statistically significantly different from zero. *** p<0.01, ** p<0.05, * p<0.1

Appendix Table 3. Unemployment/Informal to Formal Employment (Levels)

	(1)	(2)	(3)	(4)
Unemployment Defintion	Broad	Broad	Narrow	Narrow
Unemployed/Informal				
Unemployed	0.056***		0.064***	
Work in Informal sector	0.108***		0.113***	
Unemployed/Informal/Experience				
Unemployed with experience		0.072***		0.080***
Unemployed never worked		0.033***		0.040***
Work in Informal sector		0.112***		0.116***
Education				
Less than primary	0.053***	0.053***	0.060***	0.060***
Primary but not secondary	0.066***	0.066***	0.078***	0.077***
Secondary but not university	0.089***	0.091***	0.104***	0.106***
University or more	0.124***	0.124***	0.143***	0.144***
Other	0.078***	0.078***	0.090***	0.089***
Race				
African/Black	0.067***	0.068***	0.079***	0.079***
Colored	0.095***	0.091***	0.104***	0.101***
Indian/Asian	0.111***	0.107***	0.127***	0.123***
White	0.147***	0.142***	0.178***	0.172***
Gender				
Male	0.095***	0.094***	0.111***	0.110***
Female	0.053***	0.054***	0.061***	0.062***
Age				
Age 15-24	0.065***	0.078***	0.081***	0.095***
Age 25-34	0.082***	0.083***	0.093***	0.094***
Age 35-44	0.076***	0.071***	0.087***	0.083***
Age 45-54	0.064***	0.058***	0.074***	0.070***
Age 55-64	0.051***	0.048***	0.063***	0.060***
Observations	270,203	270,203	196,021	196,021

Notes: Uses data from the QLFS 2008Q1 to 2014Q3. The table reports results from a regression where the dependent variable is a dummy which takes value 1 if an individual transtions from unemployment or the informal sector to formal employment ($i_{i,t}$). The figures reported are the levels of the transition probability from the corresponding probit regression (the levels counterpart of Table 5 in the main text). Columns 1 and 2 use the broad definition of unemployment to define the LHS variable while 3 and 4 use the narrow definition. Columns 1 and 3 include a RHS variable for whether the individual was unemployed or in the informal sectoe. Columns 2 and 4 include a RHS variable of whether an individual was unemployed with no experience, unemployed with experience, or in the informal sector. Standard errors were clustered at the household level. We do not report standard errors but indicate whether a coefficient was statistically significantly different from zero. *** p<0.01, ** p<0.05, * p<0.1