

https://dspace.org/

Firm Wage Premia, Rent-Sharing and Monopsony When Underemployment is High

Bassier, Ihsaan

2021-02
Southern Centre for Inequality Studies (SCIS)
https://hdl.handle.net/10539/33953

Downloaded from a DSpace repository. https://dspace.org/

Firm Wage Premia, Rent-Sharing and Monopsony When Underemployment is High*

Ihsaan Bassier†

February 9, 2021

SCIS Working Paper

Southern Centre for Inequality Studies
University of the Witwatersrand

www.wits.ac.za/scis

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

SOUTHERN CENTRE FOR INEQUALITY STUDIES

Abstract

How important are firms in the labour markets of developing countries? Using matched employer-employee data from South Africa, I find firms explain a larger share of wages than in other, richer countries. I show this can be parsimoniously explained by the high degree of underemployment. Estimating separations elasticities by instrumenting wages of matched workers with firm wages, among other methods, I find a low separations elasticity which generates a high degree of monopsony. The correspondingly high estimated rent-sharing elasticity explains the important role of firm wage policies, even in an economy with a large labour surplus. This paper is a work in progress.

*Special thanks to Arindrajit Dube and Suresh Naidu for ongoing support and comments, as well as to Adam Aboobaker, Joshua Budlender, Aroop Chatterjee, Bridget Diana, Leila Gautham, Surbhi Kesar, Murray Leibbrandt, and participants in the SALDRU weekly seminar. Data access was only possible through support from the South African National Treasury and UNU-WIDER SA-TIED inequality work stream, along with the invaluable help of Amina Ebrahim and Junior Chiweza. Many thanks to the Southern Centre for Inequality Studies for financial support. An earlier version of this work is available online. All errors are my own.

†University of Massachusetts, Amherst. Room 114, Gordon Hall, Department of Economics, 412 North Pleasant Street, University of Massachusetts Amherst, MA 01002. Contact: ibassier@umass.edu

1 Introduction

How relevant are firm wage premia in economies with high underemployment? And are such wage premia best understood as arising from monopsony power in the labour market, as modelled for countries close to full employment? While a fast-growing literature has investigated the role of firms in explaining labour market inequality, much of the attention has focused on the United States (US) and a handful of European contexts, partly due to a paucity of matched employer-employee administrative data (Bonhomme et al. 2020; Card et al. 2018). This raises the question of whether we can extrapolate the findings of the literature into countries that are characterised by underemployment; and whether this underemployment may, in fact, exacerbate wage inequality through the channel of firm wage policies. To make progress on these questions, I use matched employer-employee data. I provide the first estimates of the role of firms in explaining labour market inequality in South Africa, as well as estimates of the mechanisms of rent-sharing and monopsony power.

South Africa contests the top spot in two undesirable world rankings: the highest rate of unemployment, with 25% of the labour force unemployed, and the most unequal country, with a Gini coefficient of 63 in household income (World Bank 2020). One may expect, as in classical models with competitive labour markets, that firm wage premia are negligible in economies with high underemployment, such as South Africa, since firms recruit workers at their reservation wage. A monopsonistic model with job search for the unemployed implies the same. However, incorporating on-the-job search introduces dispersion of firm wage premia (Burdett and Mortensen 1998). Monopsonistic models link high unemployment with high firm wage dispersion through low job-offer rates and a large gap between firm productivity and the reservation wage. This link remains untested, since there is little direct evidence of the role of firms in the labour markets of countries with high underemployment.

Using matched employer-employee tax data from 2011 to 2016 for the universe of South African formal sector workers, I find evidence of high variation in firm wage premia in two ways. Firstly, I follow the literature in decomposing earnings into AKM firm and worker effects (Abowd, Kramarz, and Margolis 1999). I find that firm wage premia explain 23% of the total wage variance. Sorting of high-wage workers to high-wage firms accounts for a further 11%. Secondly, I follow the event-study approach in Bassier, Dube, and Naidu (2020) and track matched workers in the same firm who switch to different firms in the post-period. I show a tight correlation between own wages and new average firm wages. The firm share in both cases is higher than most previous studies, highlighting the increased role of firms in wage setting in this high underemployment context.

One of the proximate causes of a greater role of firm wage premia is rent-sharing. Here, too, I find estimates from South Africa to lie toward the higher end of the literature. Regressions of AKM firm wage premia on proxies for rents, such as log value added per worker, following Card, Cardoso, and Kline 2016), yield rent-sharing elasticities of about 0.14. This variation in rents explains a quarter of the dispersion in firm wage premia. As an alternative approach, regressing within-firm changes in log wages to changes in firm-specific rents yields a rent-sharing elasticity of 0.23. Once again, these rent-sharing elasticities are large relative to the literature. This suggests a strong link between the dispersion in wage premia and firm production, rather than region, industry or amenities.

A range of models relate higher firm wage dispersion and rent-sharing to greater labour market power. As the primary measure of monopsony power, a low firm labour supply elasticity incentivises firms to mark down wages from marginal productivity, but this also means firm-specific productivity shifters affect wages more at the margin.[1] Moreover, monopsonistic competition also links high unemployment to high firm wage dispersion and rent-sharing.

Consistent with this explanation, I find that the firm labour supply elasticity for

[1] For example, Card et al. (2018), Lamadon, Mogstad, and Setzler (2019), and Manning (2003). See Online Appendix C1 for more details.

South Africa is low when applying a range of comparable methods to contexts with close to full employment. First, a regression of separations on firm wage premia yields a labour supply elasticity of about 0.85, which is substantially lower than firm labour supply elasticities of 1 to 4 estimated for several countries (Sokolova and Sorensen 2021). Second, regressing within-firm changes in average separations on changes in firm wages, instrumented by changes in firm-specific rents, gives a similar estimate. Third, following Bassier, Dube, and Naidu (2020), I use the matched event study of worker movers across firms, and find a low separations response to wide variations in wages at new firms. Overall, these estimates of a low labour supply elasticity provide a parsimonious link between underemployment and high firm wage dispersion.

While South Africa provides an extreme context that is ideal for testing the link between high underemployment and wage inequality through firms, a review of comparable estimates in Bonhomme et al. (2020) and Card et al. (2018) suggests, more generally, that higher unemployment is associated with higher variance in firm wage premia, sorting and rent-sharing. In turn, such higher variance, sorting and rent-sharing of firm wage premia are associated with higher total wage variance. An instructive comparison is Brazil: Alvarez et al. (2018) document similar patterns, where firm wage premia and sorting account for a high proportion of total wage variance. The rent-sharing elasticity is also about 0.14.

The large role firms play in South Africa and Brazil highlights a mechanism that may be relevant to emerging markets generally. Classical models of development predict high dispersion in firm productivity as some sectors lead the industrialisation process before equalising across the rest of the economy.[2] If under-employment (implying monopsony) and high dispersion in firm productivity (implying differential rent-sharing) are classical features of development, then firm wage

[2] Lewis (1954) narrates, "Capital and new ideas are [...] highly concentrated at a number of points, from which they spread outwards." More recently, Hsieh and Olken (2014) suggest from data on India, Indonesia, Mexico and USA that "development is associated with a decline in the skew of the firm size distribution."

premia may be an important channel through which wage inequality increases as economies tend towards full employment. This contribution of firms to inequality is not inevitable: collective bargaining or regulation via minimum wages are countervailing mechanisms. Alvarez et al. (2018) show that a reduction in the variance of firm wage premia accounts for 40% of the dramatic decline in Brazil's inequality between 1996 and 2012, a period over which the real minimum wage increased substantially. Future work on this paper will explore the links between monopsony power and development.

I investigate other explanations for my estimates in South Africa. Firstly, while my findings do not rely solely on the AKM method, I show that my estimates of firm wage premia are robust to several issues highlighted in the literature. My six-year panel establishes confidence against limited mobility bias in the AKM decomposition (Lachowska et al. 2020). I check my estimates with a bias-correction procedure (Andrews et al. 2008) and, alternatively, remain restricted to a set of firms with a large number of movers. I also provide evidence against substantial bias from endogenous matching by showing flat pre- and post-separation wage trends, symmetric wage gains/losses for workers that switch to different firm wage quartiles, and small AKM residuals (Card, Cardoso, and Kline 2016). In addition, I estimate an alternative set of firm wage premia from firm closures, following earlier work on industry differentials by Gibbons and Katz (1992). I provide evidence on compensating differentials by comparing firm premia estimated on wage income to firm premia estimated on the full wage package including medical aid, overtime payments, and annual bonus. The two sets of premia suggest that wage amenities do not offset wage income.

Secondly, South Africa has a substantial informal sector that is excluded from the main analysis that uses administrative tax data. I provide suggestive evidence from surveys that competitive dynamics are similar for informal labour, with similar labour supply and rent-sharing elasticities.

Thirdly, South Africa's system of collective bargaining in selected local labour markets may explain the dispersion of firm wage premia and rent-sharing. In contrast, I find that the contributions to inequality from the dispersion of firm wage

premia and sorting are *higher* in non-covered local labour markets. I argue that collective bargaining may still interact with monopsony power in nuanced but important ways. For example, I find suggestive evidence that collective bargaining increases rent-sharing across covered local labour markets, even though firm wage premia are more uniform within these markets.

Aside from shedding light on the link between underemployment and inequality through the role of firms, a final contribution of this paper is to use comparable methods in estimating firm wage premia, the rent-sharing elasticity, and the labour supply elasticity. The advantage of this is exemplified by my analysis of the gender wage gap in South Africa. Differences in firm wage premia account for 40% of the unconditional gender wage gap, which is higher than found elsewhere.[3] The labour supply elasticity for women is lower than for men and, using a simple markdown equation, the difference fully accounts for the gap in average firm wage premia by gender.

In the next section, I describe the data. Section 3 discusses the firm wage premia, Section 4 links these wage premia to rent-sharing and Section 5 provides estimates of monopsony power. Section 6 discusses the link between unemployment and firm wage premia internationally, and the informal sector, the role of unions and heterogeneity by gender. Section 7 concludes the paper.

2 Data

I use six years of South African administrative tax data between 2011 and 2016, made available through a confidential data-sharing agreement with South Africa's National Treasury and UNU-WIDER. A collection of papers describing and using this dataset appears in a special issue of the South African Journal of Economics in the first quarter of 2018. This is referenced in Appendix B along with a more detailed description of the sample construction.

The cleaned dataset used for my analysis consists of between eight and nine

[3] For example, Card, Cardoso, and Kline (2016) find that firm wage premia account for 20% of the gender wage gap in Portugal.

million workers each year, summarised in Table 1 with further summary statistics provided in Appendix Table B1. The median real annual wage is stagnant, growing about 0.2% per year, compared to much faster growth at the 90th percentile of about 1.5% per year. Nearly 40% of workers separate from firms each year, a high turnover rate in line with findings for the same data by Kerr (2018). Close to half of the workers who separate go to other firms (E-E separations) while the rest are not employed in the following year (E-N separations).

Overall, the administrative data contains no sampling errors, is probably more reliable for wages than surveys[4] and provides a unique opportunity to track a panel of workers over an extended period.[5] However, this dataset is not totally representative for all workers in South Africa. Informal, unreported work, such as domestic workers and informal traders, is excluded (together comprising a third of all employed, see Appendix Table B2). These workers likely have poorer payment and conditions. I focus on firms with more than 20 workers for reliable estimation of the wage premia. Some firms are missing data (e.g. 20% of firm profit/sales records) – these restrictions are unlikely to be random, although the direction of bias for this selection is unclear. Given all these caveats, the data still reflects the actual income of over half of all workers in South Africa for six years.

3 Firm wage premia

3.1 AKM estimation method

I estimate firm wage premia using the AKM wage equation, which imposes an additive structure of "firm effects" φ_j (used interchangeably with "firm wage premia" in this paper) and "worker effects" a_i as follows. The outcome is log annualised wages for individual i in firm j for year t. I control for $X_{i\,jt}$ as up to a cubic in age, fraction of the year employed and year fixed effects. All analysis is restricted to the largest connected set of firms.[6]

[4] Surveyed wages are typically dependent on respondents remembering exact figures across many months (including once-off payments like annual bonus) in the right definitions (e.g. net or before tax) and being willing to give up socially private information.

[5] There are no other matched employer-employee datasets available in South Africa. The National Income Dynamics Survey is the other possibility for panel data analysis for individuals in South Africa, but suffers from small sample size and is only collected every other year.

[6] Firms A, B and C are connected if the same worker is observed at A and B, and a (potentially

$$ln(wage_{i\,jt}) = \theta + \sum a_i + \sum \varphi_j + X_{i\,jt}\varrho + \upsilon_{i\,jt} \quad (1)$$

Identification of the firm effects relies on movers – workers who switch between firms. Imagine a simple two firm, two period case. Hull (2018) shows that assuming parallel trends (the counterfactual wage growth of a mover is that of a stayer) and impersistence (the mover's wage at the new firm is the same as if she had always been there), then the firm 2 effect is just a weighted average of the wage gain experienced by movers to firm 2 and the wage loss experienced by movers from firm 2. The additive structure of the AKM imposes that the wage gain and loss are equal.

Figure 1 presents a non-parametric justification for this structure, following Card, Cardoso, and Kline (2016). The figure depicts the average wages of workers before and after switching firms, without imposing any assumptions. Workers are classified by quartiles of the co-worker wage distribution, i.e. leaving out the wage of the worker to emphasize the firm wage.

In the three years before the switch, wages for movers across the distribution are stable, which is consistent with the parallel trends assumption. There is little evidence of a substantial "Ashenfelter Dip", i.e. the possibility that workers experience a negative event which registers as below-average wages in $T-1$ and causes them to switch firms, indicating falsely a firm wage gain on switching. The stability of wages after the move supports the impersistence assumption. The symmetric magnitude of wage changes for quartile *i* to *j* workers compared to quartile *j* to *i* workers supports the additive structure of the firm effects on wages (Appendix Figure A1 shows remarkable symmetry across all quartile transitions). Note the *large magnitude* of the change in wages associated with transition to different firms, for example workers who transition from quartile 1 to quartile 3 firms experience an average wage gain of 50%.

A second general check on the AKM structure is given by Appendix Figure A2,

different) worker is observed at both B and C. This enables comparison between wage premia of the firms, based on the differences in wages experienced by movers

which shows the median residuals by deciles of firm and worker effect for workers who move across firms. The imposed AKM additive structure seems to fit the empirical data worse at the bottom decile of the worker and firm effects distributions. But otherwise, the residuals are negligible in magnitude (0-2% of wages), especially in comparison to the magnitude of the changes in firm wage premia suggested by Figure 1.

I address three common concerns with the AKM identification of firm wage premia: limited mobility bias, endogenous mobility and compensating differentials. Firstly, regarding limited mobility bias, firm wage premia are more likely to be mis-measured with fewer movers. Mis-measurement of the firm effects increases its variance, falsely accounting for a larger proportion of the total wage variance. To investigate this, I follow Lamadon, Mogstad, and Setzler (2019) and show in Appendix Figure A4 that the variance of the firm wage premia increases when estimated on smaller shares of movers within each firm. Note, however, that the variance shows little movement once 60% or more of the movers are included. This is consistent with the analysis in Lachowska et al. (2020), who argue, using administrative data from Washington, that limited mobility bias is less of a concern when using longer time series. My six-year panel therefore lends credibility to the findings. For checks in robustness, I restrict the study to the set of firms with at least ten movers in each year. I also run the procedure provided by Gaure (2014) as a parametric correction.

Secondly, as discussed in Abowd, McKinney, and Schmutte (2019), the event analysis in Figure 1 may still be consistent with some endogenous mobility biasing the firm wage premia. Firm wage premia will be overestimated if wages depend on idiosyncratic firm-worker matches and workers who are at badly matched firms tend to move towards firms that offer better matches.[7] Reassuringly, Bonhomme,

[7] With mean-zero match effects, relative firm wage premia can be estimated as the average wage change for workers who move between two firms. However, if movers earn a below-average income at the original firm (negative match effect), or are attracted to an above-average income at the destination firm (positive match effect), then the wage change will over-estimate the firm effect.

Lamadon, and Manresa (2019) find using Swedish matched data that endogenous moves make little difference to their firm wage premia decompositions, despite strongly detecting the existence of such moves.

Nevertheless, one strategy to address this is by restricting the AKM sample of workers to separations from firms that have closed (to any firm) as an alternative set of firm wage premia. These separations are plausibly less endogenous, yet exhibit very similar estimates to the main set of firm wage premia. Moore and Scott-Clayton (2019) similarly consider firm pay premia and displaced workers, using matched administrative data from Ohio in the United States. Across the sample period, about 10 000 firms are not observed in the following year[8] and workers separate from these firms to a large network of firms. As a further measure of robustness, I estimate a set of firm wage premia on an AKM regression with relative year effects to address any systematic pre or post trends in wages, and on the sample of E-E separating workers in case a period of unemployment between working at different firms makes the respective wages less comparable.

Thirdly, I investigate the possibility that compensating differentials offset the wage premia with non-wage amenities (Lamadon, Mogstad, and Setzler 2019; Sorkin 2018). Since I observe different sources of income for wages in the tax data, I can compare the monthly wage to the total compensation package (including annual bonus, medical aid and overtime). As an immediate check, the AKM decomposition estimated on monthly wages only is very similar (Appendix Table A1).

If amenities offset wages in the firm premia, we would also expect the "total compensation" premia to be more compressed than the "monthly wage" premia. Appendix Figure A5(a) initially suggests that this may be the case, with a slope of 0.93 indicating that an increase in monthly wage premia is accompanied by less than a 1:1 increase in total compensation premia. However, the estimate may purely be a result of measurement error attenuation and, in fact, the reverse regression suggests this is the case. Without measurement error, the coefficient in the reverse

[8] A weakness of this approach is that some observed closures may in fact reflect other events, such as mergers.

regression should be the inverse, i.e. 1.08, yet the actual coefficient from the regression of wage on compensation is 0.9. The small role of amenities is affirmed by Appendix Figure A5(b), which shows the difference between the earnings and wage firm effects (a proxy for wage amenities) against the wage firm effects – there is no apparent pattern (slope of 0.07, standard error 0.14). While wage amenities do not reflect general amenities, this does suggest that the wage premia represent actual differences in firm value.

3.2 Decomposition of AKM firm wage premia

Table 2 summarises the variance decomposition of wages. The main set of firm wage premia (column 1) accounts for nearly a quarter of the total wage variance. This is positively correlated with the invariant worker component of wages. The correlation between worker and firm effects is about 0.2. The decomposition exercise suggests that, under a counterfactual of no firm wage premia, inequality would be a third lower whereas, under a counterfactual of no sorting, inequality would be 11% lower.

The importance of firm wage premia to inequality is robust to alternative methods. The proportions are similar when estimated from the set of closed firms[9] (32%, column 2), as well as when estimated from an AKM regression with relative year effects (27%, Appendix Table A1), and from the set of E-E separating workers (24%, Appendix Table A1). Table 2 column 3 also provides very similar bias-corrected estimates.[10]

By comparison, Song et al. (2018) find that firm effects explain 9% of the male wage variance in the US between 2007 and 2013. Card, Cardoso, and Kline (2016) explain 20% of the male variance in Portugal over the period 2002 to 2009, and Alvarez et al. (2018) explain 16% of the wage variance in Brazil between 2008 and

[9] Even though the regression is restricted to workers who leave closing firms, these workers subsequently join a wide network of other firms. The sample size is therefore not severely reduced compared to column 1, though a disadvantage is that it leverages fewer movers which may contribute to limited mobility bias.

[10] The bias-correction relies on errors that are independent and have common variance, as discussed by Kline, Saggio, and Sølvsten (2019) who show for Italian records that the correction reduces the share of variance explained by the firm fixed effects by 18% of the original share. It is reassuring in the South African context with a 6-year panel that the share explained remains the same after the correction

2012 (down from 24% for the period 1996 to 2000). Firm wage premia in South Africa appear to be especially prominent.

Appendix Figure A3 plots the distribution of firm and worker effects by decile, as a visualisation of the decompositions in Table 2. Workers with a high worker component in wages are disproportionately located in firms with high firm effects or wage premia. As firm effects decrease, the composition of worker effects correspondingly reverses so that those with low worker effects are at low-paying firms disproportionately. Still, *some* workers with low worker effects are still in high paying firms, which is a key requirement of the "job ladder" model where there exist "good" and "bad" jobs available to the same worker. Note that workers are disproportionately located at the highest decile of firm wage premia, which hints at some of the dynamics explored in a later section on the firm labour supply elasticity.

As an illustration of the importance of firm wage premia, Table 2 panel B shows average wage premia by group in the wage distribution. By construction, there is a large gap of 1.5 log points on average between workers in the bottom three deciles compared to workers in deciles 5-8. The proportion of this gap explained by the firm wage premia (about 40%) is about the same as the proportion explained by the worker fixed effect. This is a noteworthy finding for the South African literature which focuses on skills in explaining inequality (Leibbrandt, Ranchhod, and Green 2018), given that the worker fixed effect includes all invariant worker-specific characteristics such as education.

The firm wage premia provide further insight into group-level inequality. Table 2 panel B decomposes average wages into average worker and firm effects, by group. While roughly half of the gender wage gap for this sub-population (about 17%) is explained by differences in average worker effects[11], over 40% of the *unconditional* gender wage gap is explained through differences in sorting. Figure 2 illustrates the relative importance of worker and firm effects by gender and decile of the income distribution. The role of firm effects in generating inequality, whether

[11] The worker effects include both average differences in invariant characteristics that increase marginal productivity to the firm (such as skill and experience) *as well as* average economy-wide discrimination against women (such that women are underpaid no matter which firm they work at).

between genders or across the distribution, dramatically enhances the role of invariant worker characteristics in the total wage. Appendix Table A1 gives a number of alternative samples on the role of firm wage premia in explaining the gender wage gap, and the gap between the middle and bottom deciles of income distribution.

3.3 Matched event study of worker movers

As an alternative estimate of the relevance of firm wage premia, I follow Bassier, Dube, and Naidu (2020) in constructing a matched mover event-study (Appendix D provides details). In brief, this matches workers who leave the same firm at the same time with similar characteristics (wage, tenure, age, gender) and compares the change in their wage relative to the change in average firm wage at the new firms. This approach measures how much of the variation in own wage across firm switches is due to differences in firm wage policies, with the advantage over AKM of controlling for more confounders transparently.

Figure 3 shows a flat pre-trend, which is an anologous falsification check on the exogeneity assumption of the destination firm wage as in Figure 1. The wage trend after the move is also stable, indicating that concerns such as tenure profiles are not important. The coefficient of 0.65 from this regression of change in own wage to change in firm wage is high when compared to Bassier, Dube, and Naidu (2020) who find a comparable estimate of 0.32 for Oregon, USA.

Together, these results present robust evidence that substantial firm wage premia exist. Workers who switch firms experience large increases in wages while maintaining invariant characteristics such as race, sex, age and, plausibly, skill. What explains these effects? In the next section, I estimate the rent-sharing elasticity, which is a proximate cause for a greater role of firm wage premia, before considering the impact of monopsony power and collective bargaining.

4 A proximate explanation: rent-sharing

While the marginal revenue product of labour (MRPL) places a rough ceiling on the wage, MRPL may differ by firm – in turn allowing wages to differ for similar workers. Such differences in MRPL may be driven by anything from more productive technology or workflow to product market imperfections. I follow the literature in using value added per worker as a proxy for these "quasi-rents". In South Africa, the 75th percentile of the worker distribution is associated with value added per worker that is *ten times higher* than the 25th percentile. The link between this high dispersion in firm value added per worker and the AKM firm wage premia is estimated as the rent-sharing elasticity.

4.1 Empirical strategy

My main specification for the rent-sharing elasticity ε_{rent} follows Card, Cardoso, and Kline (2016) in regressing the firm wage premium φ_j on firm value added,
with controls X (such as fixed effects for time, industry and location).

$$\varphi_j = a + \varepsilon_{rent} ln(ValueAddedpp_j) + \Gamma X_{i\,jt} + v_{i\,j} \quad (2)$$

Assuming that individual wages can be decomposed into an *invariant* worker effect and a firm effect, as in the AKM model, any *firm specific* effect should reflect as differences in the firm component of the wage. The earlier literature on rent-sharing tended to use wages as the outcome, which yielded an upwardly biased rent-sharing coefficient since more profitable firms tend to employ more workers with higher invariant worker effects. Wages are then higher due to selection on worker effects *as well as* rent-sharing. The AKM firm wage premium effectively controls for individual characteristics, including unobservable characteristics. Specifications are run at the individual level and clustered by the firm.

Omitted variables correlate with profits and the firm wage premia may still bias the results. As an alternative specification that does not rely on the AKM modelling assumptions, I regress the change in log firm average wage to the change in log firm value added (at the firm level weighted by firm size and clustered by firm). Aside from providing a relatively independent empirical estimate, identification in a

differenced setting includes incumbent workers rather than just movers (as in AKM). One challenge is that annual movements in the value added are subject to measurement error, for example adjustments to the balance sheet related to asset purchases rather than actual profits. I therefore supplement the first differences specification by instrumenting the short difference ($ln(va_t) - ln(va_{t-1})$) by the long difference ($ln(va_{t+1}) - ln(va_{t-2})$) so as to isolate sustained changes to the value added. I also include fixed effects for industry by location to isolate variation from firm-specific shocks.

4.2 Results

Firm wage premia increase strongly with firm rents. The associated rent-sharing elasticity is high relative to other industrialised countries and it explains 25% of the total variance in firm wage premia. Figure 4 illustrates the association between firm value added per worker and wages (both in log terms). The association for wages is much steeper than for the firm component of wages, since workers with high wage worker effects tend to be in profitable firms. This upwardly biases the rent-sharing coefficient.

Table 3 presents estimates of the rent-sharing elasticity. These estimates demonstrate a strong correlation with p-values below 0.001. Column 1 shows an elasticity of 0.3 using the raw wage, which is upwardly biased compared to the coefficient of 0.14 in column 2 using the firm wage premium. Column 3 restricts the sample to firm closings using firm effects, finding a very similar coefficient to the full sample. Column 4 presents a first-differences specification, with a remarkably similar result. However, adjusting for measurement error by instrumenting across time periods, increases the coefficient to 0.24. This may still pick up industry-level adjustments, which should not be conflated with firm rent-sharing (Lamadon, Mogstad, and Setzler 2019). Specification 6 shows that the IV estimate is robust in accounting for market-level differences, by adding industry per location by year fixed effects.

Appendix Table A2 presents further alternative specifications, showing that the AKM estimates are strengthened by using gross profits as a measure of rent instead of the value added, or total factor productivity, which is proxied by the residual of a regression of sales on cubics in firm size, assets and material costs. When controlling for industry and geography fixed effects, the firm rent-sharing elasticity is 0.12 – very similar to the main AKM-based estimated (0.14). The change in wages of new hires is similarly increasing in the value added shock.

Appendix Table A1 shows that the rent-sharing elasticity is stable across different identifications of AKM wage premia: relative year effects, a minimum of 10 movers per firm, only E-E separations and using only wages rather than the full

earnings package (all estimates between 0.13 and 0.15).

Overall, these estimates fall towards the upper end of the range found in the AKM rent-sharing literature of between 0.05 and 0.15. These estimates are consistent across a variety of identification strategies.

5 Monopsony power: the firm labour supply elasticity

Why do firms pay different wages? A high dispersion of firm rents does not necessarily imply dispersion of firm wage premia – after all, the estimated rent-sharing elasticity is well below 1. In a competitive labour market, a more productive firm (perhaps due to non-reproducible technology) could employ workers at the market wage and retain the difference as profits. The evidence on rent-sharing, therefore, only provides a proximate cause for differences in firm wage premia. This section explores monopsony power as an explanation for the high dispersion in firm wage premia, and also as a link to the context of high underemployment.

In models of monopsonistic labour markets, rent-sharing is an optimal wage-setting outcome (Card et al. 2018; Lamadon, Mogstad, and Setzler 2019). Firms with higher marginal revenue product (or rents) gain more from employing more workers. In turn, this requires increases in the wage to attract more workers. Since wages are more sensitive to firm-specific rents in more monopsonistic markets (see Appendix C1), the relatively high rent-sharing elasticity estimated above suggests a low labour supply elasticity.

5.1 Empirical strategy

As a measure of monopsony power, I estimate the labour supply elasticity from worker separations, following the approach of Manning (2003). The intuition is that in a perfectly competitive market, workers leave low-paying firms immediately for high-paying firms. In general, a higher elasticity of separations from a firm with regard to wages indicates a more competitive market and less monopsony power. The direct estimate of the labour supply elasticity to the firm (ε_{LS}) combines estimated elasticities for the probability of separations to employment (ε_{EEsep}), to non-employment (ε_{ENsep}), and the recruitment elasticity from employment ($\varepsilon_{EErecruits}$),

each weighted by the proportion of hires from employment (θ).

$$ln(Separation_{it}) = a + \varepsilon_{sep} wage_{it} + \Gamma X_{i\,jt} + v_{it} \quad (3)$$

$$\varepsilon_{LS} = -(1 + \theta)\varepsilon_{EE sep} - (1 - \theta)\varepsilon_{EN sep} - \varepsilon_{EE recruits} \quad (4)$$

The primary regressor is the worker's wage. I estimate the firm labour supply elasticity using wages, but show that the firm wage premia give similar results as that drawn on related research in Bassier, Dube, and Naidu (2020). Controls indicated by $X_{i\,jt}$ include the worker fixed effect estimated in the AKM regression as a means of capturing invariant worker characteristics of the wage, as well as time effects. I cluster by firm. Note that E-N transitions include transitions from the formal sector to the informal sector, although such transitions are infrequent (see Appendix table B2).

As an alternative specification for the labour supply elasticity, I exploit the rent-sharing mechanism. I use changes in value added per worker as an instrument to isolate shocks to the firm average wage.[12] The pass-through to wages induces a change in the probability of separation from the firm. This is a relatively independent identification strategy that uses variation within each firm rather than the cross-section.

Finally, I alternatively estimate the separations elasticity using the matched event study of worker movers (see Appendix D, Bassier, Dube, and Naidu 2020). Identification arises from comparing matched workers from the same firm who separate to different firms, each being treated with a different average firm wage and correspondingly separating at different rates.

5.2 Results

The labour supply elasticity is low across the three estimation strategies, which is consistent with high rent-sharing as well as high unemployment. Figure 5 shows a

[12] This allows me to trace out the firm labour supply curve through isolating shifts in the firm-specific marginal revenue curve. This avoids changes in firm wages associated with, for example, composition (an increase in average skill of workers) which may be correlated with separations (high-skill workers tend to have lower separations for reasons exogenous to the wage), but is not illustrative of monopsony power

negative correlation between separations and wages when measured using actual wages or firm wage premia.

Table 4, column 1 gives the unconditional ε_{LS} = 0.5. This may be biased for many reasons, for example, we expect a high skill worker who is paid poorly to have a similar separation response to a low skill worker who is paid poorly (relative to expected wage), yet the unconditional regression treats their separation-relevant wages as different. To adjust for this, I control for invariant worker characteristics as proxied by the AKM worker effects. Column 2 shows ε_{LS} = 0.85, where reassuringly E-E separations in particular are more responsive to the wage. Using firm wage premia similarly focuses on the firm component of the wage rather than worker characteristics (Bassier, Dube, and Naidu 2020), with a similar estimate of ε_{LS} = 0.89 (column 3). Controlling for industry by geography fixed effects, which may also proxy for potentially different amenities by industry and location (Sorkin 2018), decreases the estimate to ε_{LS} = 0.77 (column 4).

The second alternative estimation strategy using first differences at the firm level gives similar results: regressing the change in the firm average separation rate to the instrumented change in firm wage gives an estimate of ε_{LS} = 0.74 (column 5). In the third estimation strategy using the event study (reported in Appendix D), I find a higher labour supply elasticity than the AKM-based approach of 1.2 to 1.6. This is still about half as large as the comparable elasticity of 3 found for Oregon, USA (Bassier, Dube and Naidu 2020).

In support of robustness, Appendix Table A1 shows ε_{LS} is similar using relative year effects for the AKM wage premia (0.8), higher when ensuring more movers or estimated off E-E separators (1.1), and similar when using monthly wages rather than total compensation (0.85). Appendix Table A3 shows the estimate is lower using wage premia identified from closed firms (0.5), but similar when controlling for the AKM wage residual and worker effects (0.9). A concern is that the separations elasticity may be inaccurately estimated if firms are bound to minimum wages: the last column restricts to wages above ZAR50 000, which is above most sectoral minimum wages in South Africa. The estimate is lower with ε_{LS} = 0.63.

Considered together, the range of estimates using different methods and samples indicates confidently that the labour supply elasticity based on separations

implies substantially more wage-setting monopsony power in South Africa than for other settings. For example the meta-analysis of Sokolova and Sorensen (2018) provides a median ε_{LS} of 2.7 for estimates in the literature.

In addition to providing an explanation for the high dispersion in firm wage premia and rent-sharing, monopsony power also provides a tight link to unemployment. Search models such as those of Burdett and Mortensen (1998) provide a monotonic relationship between lower job-offer rates for the unemployed and higher, steady state unemployment. This is not to say that high underemployment in South Africa or elsewhere is primarily explained by search models (structural explanations are clearly more salient), yet monopsony power may exacerbate unemployment. More generally, underemployment is associated with a lower reservation wage *b*. In monopsonistic labour markets where firm wages vary between *b* and productivity *p*, a decrease in *b* mechanically increases firm wage dispersion. Monopsony models that explicitly include reservation wages (as noted by Card et al. (2018)) draw an explicit link between lower reservation wages and higher rent-sharing. Empirically, Appendix Figure A6 suggests a negative relationship between local labour supply elasticities and unemployment in South Africa.

Finally, does monopsony power actually translate into lower wages? As Alan Manning writes in his recent review of the literature, even if firms have a high degree of labour market power, there remains the question as to "whether employers exercise this monopsony power as a simple profit-maximising model would suggest or whether other factors act as a constraint" (Manning 2020, p.9).

In the next section, I investigate one such constraint – collective bargaining. I document here that firm wage premia are indeed responsive to the estimated markdown pertaining to monopsony power. I estimate the markdown parameter $\varepsilon_{LSE}/(1 + \varepsilon_{LSE})$ at the industry by location level. Controlling for location and industry fixed effects, the coefficient of 0.54 implies that wages decrease with more monopsony power (Appendix Table A4).[13] The predicted magnitude of the

[13] These regressions use firm wage premia, φ_j, as suggested by the AKM framing above. Incidentally, this result also suggests another source of dispersion in firm wage premia directly from heterogeneity

relationship between the markdown and wages depends on the model, but this does suggest there is substantial use of potential monopsony power.[14]

6 Discussion

6.1 Firm wage premia and unemployment internationally

In terms of a model of the labour market, this paper provides evidence on the link between South Africa's high unemployment and its high inequality through firm monopsony power. In this subsection, I review its consistency with existing international evidence, and then consider its applicability to developing countries more generally.

Comparable estimates of firm wage premia and rent-sharing are largely confined to higher income settings without labour surpluses. I compile estimates of AKM decompositions reviewed by Bonhomme et al. (2020), and the 16 firm- level estimates of rent-sharing elasticities reviewed by Card et al. (2018). For the 12 countries with estimates, Figure 6 shows that the share of wage variance explained by firm wage premia, the share explained by sorting, and the rent-sharing elasticities each appear to increase with unemployment. In turn, the total wage variance is positively associated with each of these three statistics. One country-specific study that explicitly considers the relationship between unemployment and monopsony power is that of Hirsch, Jahn, and Schnabel (2018). These authors estimate 0.15 decrease in the firm labour supply elasticity for every percentage point increase in unemployment for West Germany (over their period of estimation, cyclical unemployment ranges from 6 to 12%). While the number of countries in this collected review is small, meaning the relationships are by no means dispositive, the corresponding estimates for South Africa appear to be close to the extrapolated

in monopsony power across different local labour markets. However, the R^2 on these regressions is low.

[14] In the simplest model, the labour supply elasticity relates to the wage through a markdown, $w = \varepsilon LSE \,/\, 1+\varepsilon_{LSE}\, p$, where p is the worker's marginal revenue product. If productivity and the markdown are independent, then the predicted log-log coefficient is 1 (the estimated coefficient of 0.54 coverted to log-log terms is 0.99). See Appendix C1 for other models.

linear trends, despite far higher unemployment and wage inequality.

To consider the applicability of this literature to developing countries generally, we should address structural unemployment (or more generally, underemployment) directly as a defining characteristic of such contexts. From equation 4, consider setting $1 - \theta$ (the proportion of hires from unemployment) to the unemployment rate u; the elasticity of separation to non-employment $\varepsilon_{ENsep} = 0$; and the elasticity of recruits from employment compared to non-employment $\varepsilon_{EErecruits} = 0$. Setting these elasticities to 0 follows if we think the value of the lowest firm wages are sufficiently above the value of unemployment, that changing that wage does not induce more exit from unemployment. Then, $\varepsilon_{LS} = -(2 - u)\varepsilon_{EEsep}$, the labour supply elasticity decreases with the unemployment rate, whatever the source of that unemployment is (including surplus labour *and* frictional unemployment).

The possibility of firm-level monopsonistic dynamics contrasts with much of the analysis in classical development models which typically take place at the sectoral level. It is at this level that an industrial sector draws on surplus labour at a constant subsistence wage (Basu 2003). This case is nested in standard monopsony models. More generally, the introduction of firm-level heterogeneity – and in particular on-the-job search by workers between these firms – sharply changes the optimal posted wage. This incentivises higher-productivity firms to post higher wages in order to attract more workers, thus resulting in a large dispersion in firm wages (Burdett and Mortensen 1998).

There are indications in the literature that such firm heterogeneity may turn out to be a relevant aspect of the development process. In terms of comparable estimates of separations elasticities which indicate monopsony power, Sokolova and Sorensen (2021) find only two that are in developing contexts (Brazil and Mexico) out of all estimates between 1977 and 2019. Both Brazil and Mexico have a high dispersion in firm wage premia and a low labour supply elasticity, similar to my analysis for South Africa (Alvarez et al. 2018; Dal Bo, Finan, and Rossi 2013). Moreover, evidence from other firm-level studies highlight the potential importance of monopsonistic competition in developing contexts. For example, Soderbom, Teal, and Wambugu (2005) find a strong correlation between firm wage and size in Ghana and Kenya which, interpreted literally, suggests a low labour supply elasticity

with high monopsony power. Fafchamps and So¨derbom (2006) explain a large share of the gender wage gap of 11 African countries through selection of low-wage firms while Hsieh and Klenow (2009) document a much larger dispersion in productivity for China and India relative to the US, implying large effects on wage inequality of any pass-through onto wages.

6.2 Imperfect competition in the informal sector

The main analysis has focused on formally employed labour in large firms, even though a characteristic feature of developing countries is a substantial proportion of informally employed labour. This subsection explores the competitive dynamics in the informal sector and its interaction with the formal sector in South Africa. Using quarterly panel data from household labour-force surveys, I begin by describing the transitions between the formal and informal sectors (Online Appendix E, tables 1 and 2). Nearly one in five workers transition from the informal to formal sector every quarter, compared to only 1 in 20 in the reverse direction. Moreover, workers who transition from the informal to the formal sector report substantially larger wages, while workers transitioning in the reverse direction report wage losses of similar magnitude. The patterns are similar to those referred to in Figure 1 above comparing high and low wage firms. They are consistent with low productivity informal sector enterprises offering low wages towards the bottom of a job ladder.

What are the competitive dynamics within the informal sector? I estimate separation elasticities using panel survey data, with the advantage of tracking informal sector transitions. However, the disadvantage of measurement error resulting in less credible estimates is also considered. Nonetheless, the relative magnitudes in Appendix E Table 3 show that the labour supply elasticities for formal and informal workers are extremely similar. Next, I use enterprise level data on informal sector firms, and find an elasticity of wages to sales of about 0.26. While this estimate includes limited controls, it is remarkably close to the rent-sharing elasticity using raw wages in the tax data for formal firms reported in Table 3 of 0.3. Together, the labour supply and rent-sharing elasticities, though only indicative

given the quality of data, suggest that the imperfect competitive dynamics in the informal sector may not be too different from the formal sector, as investigated in my main analysis. Two recent papers analyse the interaction between the formal and informal sectors with search frictions in the context of Brazil (Meghir, Narita, and Robin 2015; Ulyssea 2018). These papers also find that the informal and formal sectors share a common labour market, compete for similar workers and face similar competitive dynamics. The differences in the sectors are that informal sector firms are less productive and face fewer administrative costs (but are scale- constrained).[15] Given the similarities in competitive dynamics between the sectors, the attention given to the distinction between them may be unwarranted. In a review of the informal sector literature, Rothenberg et al. (2016) motivate the division through a concern with inefficient allocation. Yet, there is large dispersion in productivity even within formal sector firms, and informal firms may simply be part of that distribution. As Hsieh and Olken (2014) comment for India, Indonesia and Mexico, the firm size distributions (including informal firms) are smooth which is inconsistent with a bimodal distribution motivating a particular emphasis on the formal-informal divide.

6.3 Collective bargaining: a complementary explanation?

This paper has linked imperfect competition in the labour market, as indicated by rent-sharing and dispersion in firm wage premia, to monopsony. However, there are many other institutional constraints on firm wage-setting. Extensive literature on collective bargaining provides an example (Farber et al. 2018; Lee and Mas 2012). South Africa's collective bargaining coverage is similar to the average for the OECD at about a third of workers (OECDstat 2016).

We can immediately rule out the possibility that unions completely account for the firm wage premia dispersion by considering firms not covered by collective bargaining. Figure 7 shows that the dispersion in firm wage premia and sorting is

[15] On the other hand, they suggest there is substantial poaching by informal firms from formal firms, which conflicts with the flow of workers across sectors in South Africa, and the associated asymmetrical within-worker wage changes, suggesting that informal sector jobs can largely be considered inferior to formal sector jobs.

higher for uncovered firms. This can be rationalised, for example, if collective bargaining increases the non-rent component of the firm wage pre- mium (e.g. a classic union premium), or if workers select more homogeneous, higher value added firms (as in monopsonistic models of minimum wages (Manning 2003)). More directly, the Labour Force Survey asks workers how their wages are set (QLFS, 2016): in 2016 quarter 1, over half had wages set unilaterally by employers, compared to about 30% whose wages were collectively bargained.

The co-existence of bargaining and unilateral wage-setting in the same labour market (even in industry by location cells) hints towards an environment of latent monopsony power. Collective bargaining and monopsony power may interact in nuanced ways. On the one hand, they may decrease inequality *within* collectively bargained industry regions as wages are standardised. On the other hand, collective bargaining may force firms to share more rents, resulting in higher inequality *across* these local labour markets. In Appendix C2, I outline a model of bargaining between workers and firms in the context of a monopsonistic labour market. Nash bargaining results in a wage that is between the monopsonistic wage and the competitive wage, where the outcome on this locus is determined by union bargaining power. Intuitively, unions' interest in wages augments the firm's existing best response of raising wages to attract more workers in response to a productivity shock. In Table A5, I provide evidence of a positive correlation between the rent-sharing elasticity and union density at the industry-location level.

Such interactions are also hinted at in the current literature. Card, Devicienti, and Maida (2014) consider the relationship between firm wage premia, bargaining and investment in Germany. In one specification, they report higher rent-sharing for firms with a high probability of union contract (0.16) compared to a low probability (0.13). Lagos (2019) estimates the causal effects of collective bargaining on firm compensation in Brazil. Although his paper does not focus on how rent-sharing varies with union strength, he does find a union premium on amenities, as well as a corresponding increase in hires to the firm. His findings suggest a complementarity between monopsony and collective bargaining. The broader point is that institutional constraints may influence the rent-sharing elasticity, in addition

to monopsony power.[16] My evidence on monopsony is therefore a partial explanation of the rent-sharing and dispersion in firm wage premia but, more generally, explains the existence of other constraints in developing country settings that do not preclude the relevance of monopsony power.

6.4 Heterogeneity by gender

Finally, I provide a decomposition in wage premia by gender, showing that the difference in average firm wage premia for men compared to women is about 7.2%. This difference accounts for about 40% of the gender wage gap. Can this be linked to rent-sharing and monopsony power, as was indicated in the main results?

Literature has developed regarding the role of firm wage premia in explaining discrimination. Discrimination in this context is usually defined as systematically lower wages for equally productive workers of one group compared to another, where these groups are typically defined by race or gender. One channel through which this may operate is if one group has greater access to high-premium firms than another, allowing a wage gap (conditional on worker productivity) to persist through networks or social norms constraining search (Caldwell and Oehlsen 2018). For example, women who are responsible for more unpaid care work may be more constrained to jobs that are more flexible or closer to home, narrowing the pool of alternatives. More directly, Lang and Lehmann (2012) show that a taste-discrimination model with just a small proportion of prejudiced firms, combined with search frictions, confers substantial monopsony power by group to non-prejudiced firms. Manning (2003) for the United Kingdom (ch. 7), Hirsch, Schank, and Schnabel (2010) for Germany and Gerard et al. (2018) for Brazil find a substantial role of firm wage-premia in explaining such group wage gaps. I take this further by empirically linking this gap in firm wage premia to monopsony power.

Appendix Table 7 presents the rent-sharing and labour supply elasticities by gender.[17] The rent-sharing elasticity is *higher* for women compared to men. This

[16] Another example of a constraint is provided in Saez, Schoefer, and Seim (2019) who find a high rent-sharing elasticity which they rationalise in a model of monopsonistic wage-setting constrained by fairness concerns.
[17] The results for the elasticities are similar using gender-specific firm wage premia or homogeneous firm wage premia, and the correlation between male and female firm wage premia is 0.93. One caveat is that the average difference in firm wage premia uses the homogeneous firm wage premia. It is

contrasts with the result in Card, Cardoso, and Kline (2016), who find higher rent-sharing for men in Portugal: they suggest a regression of the female firm wage premium on the male firm wage premium instrumented by the log value added, as a test of the ratio of rent-sharing elasticities, and find a coefficient of 0.88. I find a coefficient of 1.04 (standard error of 0.018). Note that, on average, women are at firms with substantially lower value add than men and, as a result, the implied wage premia from rent-sharing is 1% greater for *men* compared to women.

Higher rent-sharing is consistent with more monopsony power, as I suggest in my main analysis. Indeed, the labor supply elasticity is lower for women compared to men. Using the simple markdown relation $\varphi j = \frac{\varepsilon}{1+\varepsilon} p$ the differences in labor supply elasticity imply a 7.4% difference in firm wage premia – fully accounting for the difference in firm wage premia by gender. This exercise also illustrates the usefulness of comparable estimates of inequality, rent-sharing and monopsony power through the framing of firm wage premia, which is a contribution of this paper.

7 Conclusion

The main empirical result of this paper is that firms in South Africa play a large role in determining wages, relative to the literature. A large portion of variation in wages is due to firm wage premia and, as a proximate cause, high-rent firms paying higher wages explains a quarter of this variation in firm wage premia. Estimates of the rent-sharing elasticity fall within a high range of 0.12 to 0.23, across a range of checks in robustness and alternative estimation strategies.

What drives this rent-sharing and ultimately the variation in firm wage pre- mia? One channel is that firms operate in a strongly monopsonistic labour market. I estimate a low firm labour supply elasticity, implying a large markdown on firm

unclear how to compare gender-specific firm wage premia given that the magnitudes of firm wage premia are arbitrary up to a constant. Card, Cardoso, and Kline (2016) suggest normalising a set of "no surplus" firms, and this procedure yields higher firm-specific firm wage premia for women: using their schema, differences in log wages by gender can be attributed to 3% for bargaining and sorting for 8% (overall men have 5% higher firm wage premia). Note that men and women may still be paid differently in low value-added firms if the value added is an imperfect proxy for rents. Indeed, adding a firm by gender interaction term in the AKM earnings equations yields average interaction terms very close to zero.

wage premia – a correlation I document at the industry by location level. This low labour supply elasticity implies that firm-specific shifters (such as rents) affect wages more than in more competitive markets. Moreover, monopsonistic models link higher labour market power with high unemployment.

The role of firm wage-setting power is conspicuously underemphasised in the academic literature and policy in South Africa, which typically focus on the supply side to address the country's high levels of inequality and unemployment. Leibbrandt, Finn, and Woolard (2012) estimate that income from the labour market accounts for 85% of the Gini coefficient in household income, 62% of which is due to earnings inequality (rather than inequality between the employed and unemployed). The AKM wage decomposition therefore implies that firms account for a fifth of overall income inequality in South Africa.[18] Firm wage premia, sorting and rent-sharing all indicate that a worker's wage depends on firm-level dynamics, which in turn become important contributors to overall inequality. Anti-monopsony policy could substantially decrease this firm-level inequality, while also increasing employment and wages (Naidu, Posner, and Weyl 2018). Such policy includes addressing labour market concentration, reducing barriers to employment mobility, as in the search-based dynamic monopsony model, and institutionalising counterbalances to monopsonistic wage setting such as collective bargaining.

This study of South Africa links high unemployment to high inequality through firms, a channel that may generalise to developing contexts of high underemployment. Under conditions of labour surplus, an upwards-sloping labour supply curve may be counter-intuitive, since one may expect firms to simply recruit unemployed workers at the "subsistence" wage. Yet, on the job search combined with costs of migration, poor transport infrastructure and a limited supply of skills may lead to a larger role for frictions and imperfect competition generally. Indeed, in a remarkable large scale experimental evaluation of the spill-over effects from India's public employment programme NREGS, Muralidharan, Niehaus, and

[18] Assuming the percentage explained from the Gini is close to the percentage explained of total wage variance, firms roughly explain 34% (23% variance plus 11% sorting) of 62% of 85%, or 18% of total household income inequality.

Sukhtankar (2017, p. 31) note that "finding positive effects on employment forced us to question the default assumption of competitive labour markets, and look for credible ways to test this assumption." Monopsony power based on search frictions are consistent with recent experimental evidence in South Africa (Abel, Burger, and Piraino 2020) and the high cost of searching for work (Mlatsheni and Ranchhod 2017). This role of monopsony power may turn out to be a pervasive and important feature in the development process, but needs to be further modelled and tested in developing countries.

Main Tables and Figures

Table 1: Summary statistics of tax panel data

	Workers	Real earnings (ZAR)			Separations
	(freq.)	(p50)	(p90)	(mean)	(mean)
2011	8,353,791	87,426	357,735	169,446	36.5%
2012	8,681,995	87,805	365,590	174,496	36.8%
2013	8,900,366	86,377	366,006	173,343	36.3%
2014	8,981,113	86,158	370,980	173,164	36.1%
2015	9,150,558	87,527	381,560	174,745	37.4%
2016	8,999,547	88,632	385,823	178,232	

Note: Wages are annualised and adjusted for inflation (base year 2016). Earnings include wage benefits such as overtime and annual bonus. A separation occurs when a worker is no longer recorded at the same firm in the following year. Observations are restricted to workers at firms with more than 20 workers. Source: Own calculations, South African tax records, 2011-2016.

Table 2: Decomposition of firm and worker effects

	(1)	(2)	(3)
Observations	45,944,987	41,591,152	45,944,763
Var(Log Wage)	1.32	1.42	1.40
% Var(Firm FE)	23%	32%	23%
% Var(Worker FE)	43%	41%	40%
% 2 × Cov(Firm FE, Worker FE)	10.8%	9.2%	12%
Method	AKM	Closed firms	Bias Corr.

(a) Overall variance decomposition

	Gender		Income decile		
	Men	Women	Dec. 10	Dec. 5-8	Dec. 1-4
Observations	30,200,000	22,900,000	5,307,887	21,200,000	21,200,000
lnWage	11.448	11.276	13.262	11.712	10.256
Worker FE	0.040	-0.053	1.177	0.029	-0.551
Firm FE	0.031	-0.041	0.375	0.181	-0.423
lnWage gap	0.172			1.550	1.455
% Worker FE	54%			74%	40%
% Firm FE	42%			13%	41%

(b) Mean components of wage, by group

Note: Panel A shows the variance decomposition using an AKM regression on all workers (col 1); an AKM regression on the connected sample of closed firms (col 2); and the method suggested by Andrews (2008) to correct for limited mobility bias (col 3). Panel B shows averages of the full sample AKM fixed effects. Deciles of the income distribution are calculated by year. The Dec. 5-8 gap compares to Dec. 10, and the Dec. 1-4 gap compares to Dec. 5-8. Workers in both panels are limited to the largest connected set of firms with more than 20 employees. Source: Own calculations, South African tax records, 2011-2016.

Table 3: Rent-sharing elasticities

	(1)	(2)	(3)	(4)	(5)	(6)
ln(VApe)	0.301	0.141	0.140	0.141	0.237	0.225
	(0.011)	(0.008)	(0.010)	(0.036)	(0.061)	(0.041)
Obs	28 mill.	26 mill.	26 mill.	127,289	62,004	60,443
R^2	.21	.25	.21	.13	.08	.08
ln(wage)	Y					
Firm FE		Y	Y			
All firms		Y				
Closings			Y			
Firm wage first diff.				Y	Y	Y
OLS				Y		
IV					Y	Y
Market × Year FE						Y

Notes. VApe is value added per employee as a proxy for rent, calculated as sales minus non- labor expenses. The final 3 specifications are run at the firm-level (weighted by number of workers), comparing differenced outcomes within firms over time. The IV is the long difference ($lnva_t$+1 $lnva_t$ 2). Market↵FE refer to industry by location fixed effects. Workers are limited to connected firms with more than 20 employees. Standard errors are given in parentheses. Source: Own calculations, South African tax records, 2011-2016.

Table 4: Firm labour supply elasticities

	(1)	(2)	(3)	(4)	(5)
Separations	-0.264	-0.308	-0.382	-0.261	-0.319
	(0.007)	(0.014)	(0.024)	(0.005)	(0.127)
E-E separations	-0.298	-0.509	-0.548	-0.424	-0.34
	(0.012)	(0.018)	(0.034)	(0.010)	(0.171)
E-N separations	-0.392	-0.294	-0.376	-0.292	-0.221
	(0.010)	(0.022)	(0.031)	(0.008)	(0.086)
E-E recruits	0.163	0.041	0.117	0.002	-0.129
	(0.008)	(0.013)	(0.026)	(0.003)	(0.051)
E-E recruits (%)	0.451	0.451	0.451	0.451	0.442
LSE	0.484	0.858	0.885	0.773	0.742
	(0.020)	(0.032)	(0.058)	(0.015)	(0.256)
					14.961
Observations	38.2 mill.	36.2 mill.	36.2 mill.	36.2 mill.	108,047
Fstat					14.961
First Difference					Y
Regressor					
lnWage	Y	Y		Y	Y
Firm FE			Y		
Controls					
Worker type		Y		Y	
Indus X Geo FE				Y	

Notes. The top 4 rows represent separate regressions per column. E-E indicates employment- to-employment and E-N indicates employment to non-employment. The worker type control adds the worker fixed effect as a continuous variable control regressor. Colum 4 includes 221 geographic by 20 industry fixed effects. Column 5 is run at the firm-level (weighted by number of workers), and compares the change in separations within a firm to the change in average wages, instrumented by change in log value added per worker. Workers are limited to connected firms with more than 20 employees. Standard errors are given in parentheses. Source: Own calculations, South African tax records, 2011-2016.

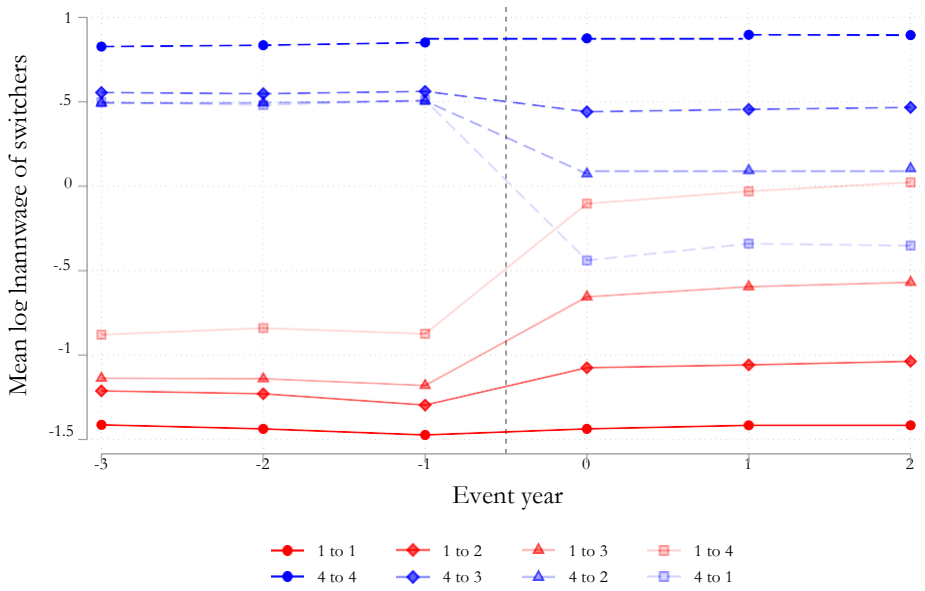


Figure 1: Wage profiles of workers who move firms

Notes: The legend shows worker moves from origin quartile to destination quartile. Only origin quartiles 1 and 4 are plotted. Quartiles are calculated as the mean co-worker quartile in the firm, i.e. leaving own-wage out of mean firm wage. Only employment-to-employment movers are included, such that each worker stayed at the same firm from 2011 to 2013, then moved to a new firm and stayed there from 2014 to 2016. Event year 0 (or tax year 2014) represents wages at the new firm. Wages of the full sample (including stayers) are residualised on year effects before plotting.



Figure 2: Decomposition of worker and firm fixed effects, by income group and gender

Notes. Worker and firm effects are estimated from the full sample AKM regression. Decile of income is calculated by year. For each decile-gender group, the average worker and firm effects are plotted. M is Male and F is Female. Workers are limited to those at connected firms with more than 20 employees.

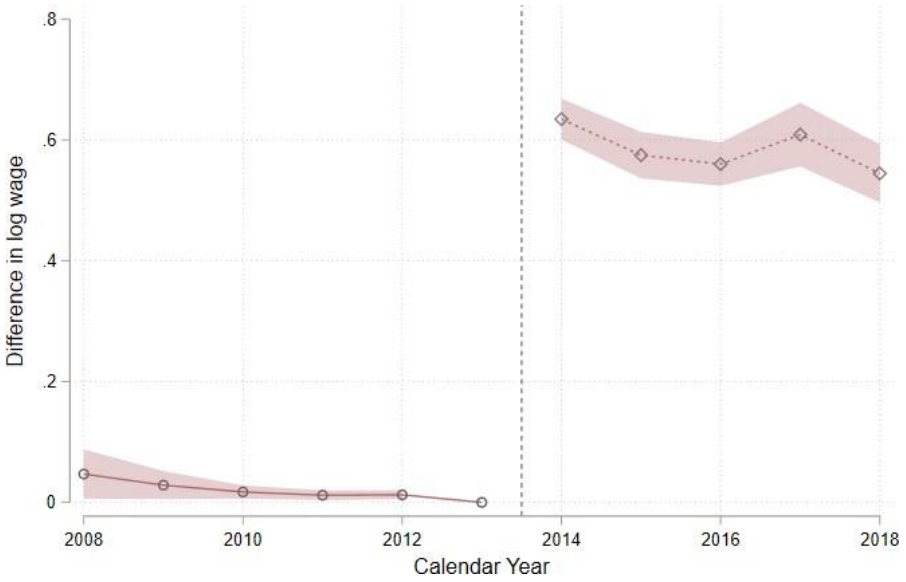


Figure 3: First stage: Difference in log wage on difference in firm average wages

Notes. See Appendix D for sample construction and specification details. The regression includes fixed effects for Origin firm, salary bins (12), fraction of year employed in bins (8), tenure bins (3), gender bins (2), and age bins (8). Origin firms are restricted to a firm size of at least 20 workers. The difference in individual and firm wages are censored at the 1% tails. Observations are restricted to Origin firm (before the event) and Destination firm (after the event).

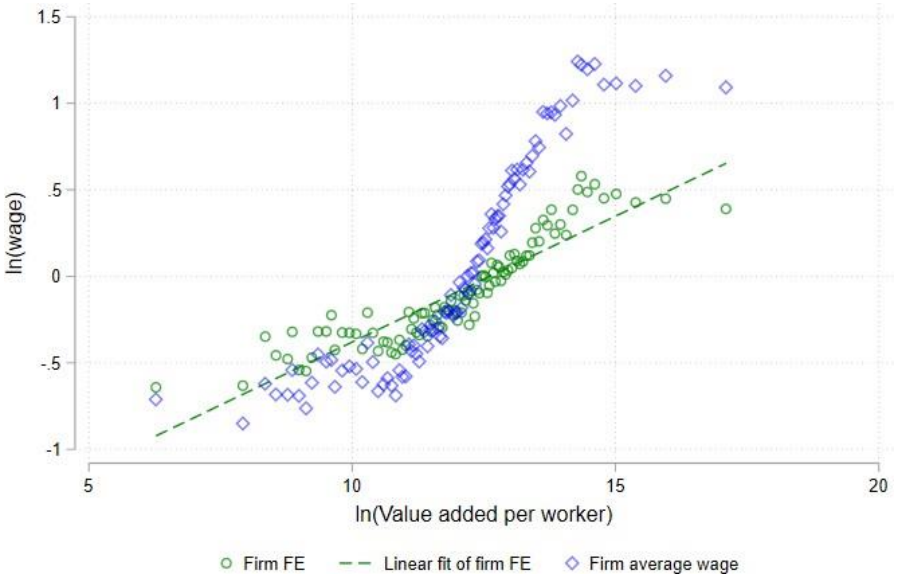


Figure 4: Non-parametric scatter of rent-sharing elasticity

Notes. Firm average wage refers to directly recorded wages for each worker, and are centred on 0 for plotting. Firm fixed effects are estimated using the AKM regression. The plot is generated from firm-level data weighted by number of workers and limited to connected firms with more than 20 employees.

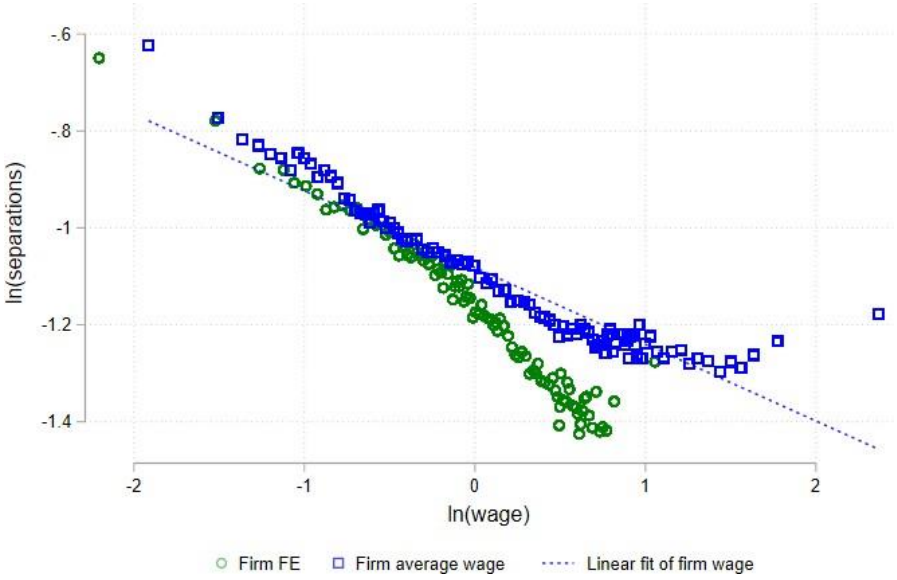


Figure 5: Non-parametric scatter of labour supply elasticity

Notes. Firm average wage refers to directly recorded wages for each worker, and are centred on 0 for plotting. Firm fixed effects are estimated using the AKM regression. Firm fixed effects and firm average wage are each divided into 100 quartiles, with the average ln separations plotted for each. The plot is generated from firm-level data weighted by number of workers and limited to connected firms with more than 20 employees.

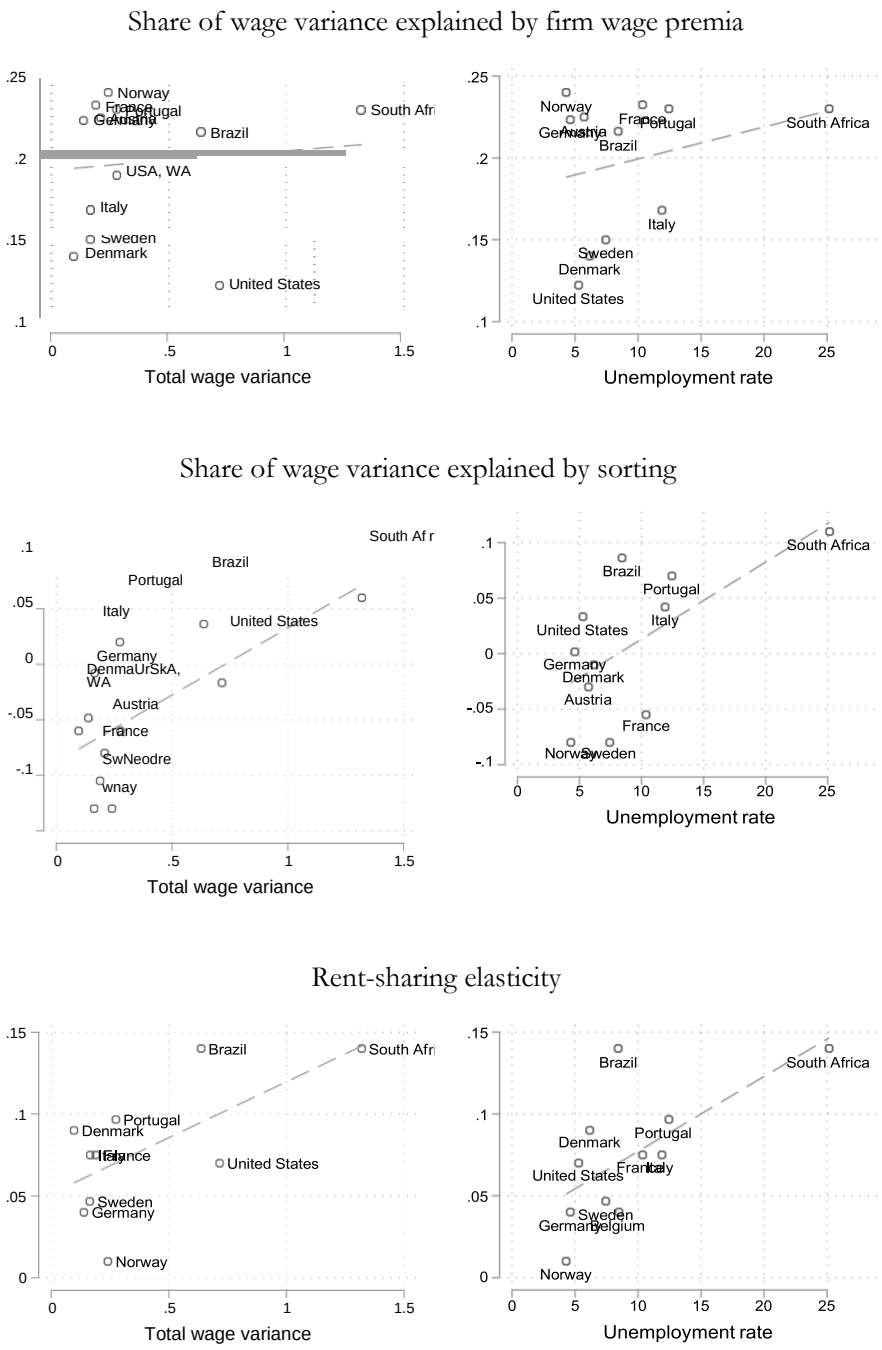
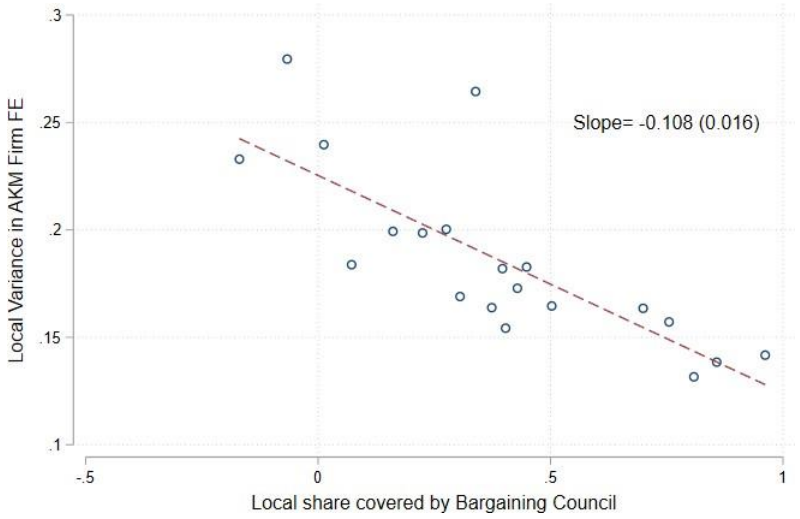
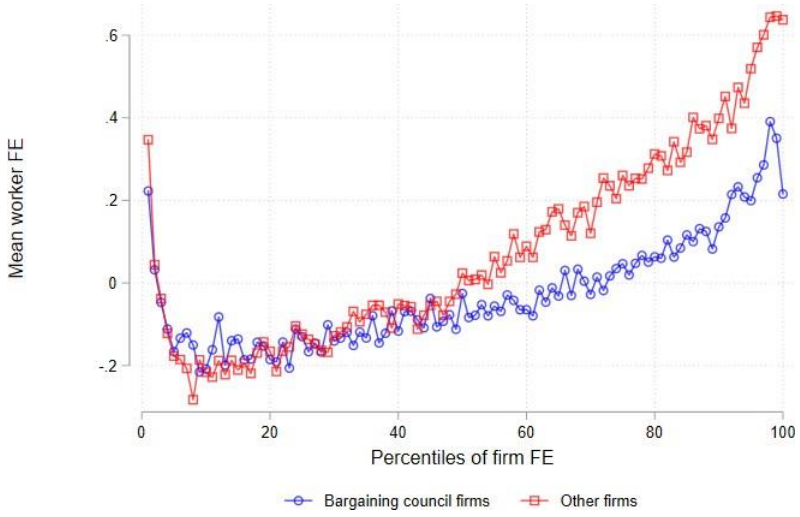


Figure 6: International comparison of firm effects, sorting and rent-sharing

Notes. Firm effects, sorting and total variance estimates are compiled from Bonhomme et al. (2020); 3 estimates for France are excluded as outliers with share of variance explained by firm effects greater than 50%. Rent-sharing estimates are compiled from Card et al. (2018) using their preferred firm level estimates. The estimates from this paper are added for South Africa, and each point represents the average estimate from studies of that country. Unemployment is compiled from ILO data for each country for 2015.



(a) Variance of AKM firm wage premia



(b) Sorting of worker and firm wage premia

Figure 7: Firm wage premia by collective bargaining council coverage

Notes. Panel A: For each industry (2-digit) by location, I estimate the share of firms covered by bargaining councils and the variance of AKM firm wage premia. The regression is weighted by firm size and includes fixed effects for industry (1-digit) and location. Panel B: The full distribution of AKM firm wage premia are ranked by 100 quantiles, and the average AKM worker fixed effects are calculated by quantile and bargaining council coverage.

References

Abel, Martin, Rulof Burger, and Patrizio Piraino (2020). "The value of reference letters: Experimental Evidence from South Africa". In: *American Economic Journal: Applied Economics* 12.3, pp. 40–71.

Abowd, John M, Francis Kramarz, and David N Margolis (1999). "High wage workers and high wage firms". In: *Econometrica* 67.2, pp. 251–333.

Abowd, John M, Kevin L McKinney, and Ian M Schmutte (2019). "Modeling endogenous mobility in earnings determination". In: *Journal of Business & Economic Statistics* 37.3, pp. 405–418.

Alvarez, Jorge, Felipe Benguria, Niklas Engbom, and Christian Moser (2018). "Firms and the decline in earnings inequality in Brazil". In: *American Economic Journal: Macroeconomics* 10.1, pp. 149–89.

Andrews, M. J, L. Gill, T. Schank and R. Upward (2008). "High wage workers and low wage firms: negative assortative matching or limited mobility bias?" In: *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 171.3, pp. 673–697.

Bassier, Ihsaan, Arindrajit Dube, and Suresh Naidu (2020). *Monopsony in Movers: The Elasticity of Labor Supply to Firm Wage Policies*. Tech. rep. National Bureau of Economic Research.

Basu, Kaushik (2003). *Analytical development economics: The less developed economy revisited*. MIT press.

Bonhomme, Stephane, Thibaut Lamadon, and Elena Manresa (2019). "A distributional framework for matched employer employee data". In: *Econometrica* 87.3, pp. 699–739.

Bonhomme, S., K. Holzheu, T. Lamadon, E. Manresa, M. Mogstad, and B. Setzler (2020). *How Much Should We Trust Estimates of Firm Effects and Worker Sorting?* Tech. rep. National Bureau of Economic Research.

Burdett, Kenneth and Dale T Mortensen (1998). "Wage differentials, employer size, and unemployment". In: *International Economic Review*, pp. 257–273.

Caldwell, Sydnee and Emily Oehlsen (2018). "Monopsony and the Gender Wage Gap: Experimental Evidence from the Gig Economy". In: *Massachusetts Institute of Technology Working Paper*.

Card, David, Ana R. Cardoso, and Patrick Kline (2016). "Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women". In: *The Quarterly Journal of Economics* 131.2, pp. 633–686.

Card, David, Francesco Devicienti, and Agata Maida (2014). "Rent-sharing, holdup, and wages: Evidence from matched panel data". In: *Review of Economic Studies* 81.1, pp. 84–111.

Card, David et al. (2018). "Firms and labor market inequality: Evidence and some theory". In: *Journal of Labor Economics* 36.S1, S13–S70.

Dal Bo, Ernesto, Frederico Finan, and Martın A Rossi (2013). "Strengthening state capabilities: The role of financial incentives in the call to public service". In: *The Quarterly Journal of Economics* 128.3, pp. 1169–1218.

Fafchamps, Marcel and Mans Soderbom (2006). "Wages and labor management in African manufacturing". In: *Journal of Human Resources* 41.2, pp. 356– 379.

Farber, Henry S, D. Herbst, I. Kuziemko, and S. Naidu (2018). *Unions and inequality over the twentieth century: New evidence from survey data*. Tech. rep. National Bureau of Economic Research.

Gaure, Simen (2014). "Correlation bias correction in two-way fixed-effects linear regression". In: *Stat* 3.1, pp. 379–390.

Gerard, Francois et al. (2018). *Assortative Matching or Exclusionary Hiring? The Impact of Firm Policies on Racial Wage Differences in Brazil*. Tech. rep. National Bureau of Economic Research.

Gibbons, Robert and Lawrence Katz (1992). "Does unmeasured ability explain inter-industry wage differentials?" In: *The Review of Economic Studies* 59.3, pp. 515–535.

Hirsch, Boris, Elke J Jahn, and Claus Schnabel (2018). "Do Employers Have More Monopsony Power in Slack Labor Markets?" In: *ILR Review* 71.3, pp. 676–704.

Hirsch, Boris, Thorsten Schank, and Claus Schnabel (2010). "Differences in labor supply to monopsonistic firms and the gender pay gap: An empirical analysis

using linked employer-employee data from Germany". In: *Journal of Labor Economics* 28.2, pp. 291–330.

Hsieh, Chang-Tai and Peter J Klenow (2009). "Misallocation and manufacturing TFP in China and India". In: *The Quarterly journal of economics* 124.4, pp. 1403–1448.

Hsieh, Chang-Tai and Benjamin A Olken (2014). "The missing "missing middle"". In: *Journal of Economic Perspectives* 28.3, pp. 89–108.

Hull, Peter (2018). *Estimating treatment effects in mover designs*. Tech. rep.

Kerr, Andrew (2018). "Job flows, worker flows and churning in South Africa". In: *South African Journal of Economics* 86, pp. 141–166.

Kline, Patrick, Raffaele Saggio, and Mikkel Sølvsten (2019). *Leave-out estimation of variance components*. Tech. rep. National Bureau of Economic Research.

Lachowska, M., A. Mas, R. Saggio, and S.A. Woodbury (2020). *Do Firm Effects Drift? Evidence from Washington Administrative Data*. Tech. rep. National Bureau of Economic Research.

Lagos, Lorenzo (2019). *Labor Market Institutions and the Composition of Firm Compensation: Evidence from Brazilian Collective Bargaining*. Tech. rep.

Lamadon, Thibaut, Magne Mogstad, and Bradley Setzler (2019). *Imperfect Competition, Compensating Differentials and Rent Sharing in the US Labor Mar- ket*. Tech. rep. National Bureau of Economic Research.

Lang, Kevin and Jee-Yeon K Lehmann (2012). "Racial discrimination in the labor market: Theory and empirics". In: *Journal of Economic Literature* 50.4, pp. 959–1006.

Lee, David S and Alexandre Mas (2012). "Long-run impacts of unions on firms: New evidence from financial markets, 1961–1999". In: *The Quarterly Journal of Economics* 127.1, pp. 333–378.

Leibbrandt, Murray, Arden Finn, and Ingrid Woolard (2012). "Describing and decomposing post-apartheid income inequality in South Africa". In: *Development Southern Africa* 29.1, pp. 19–34.

Leibbrandt, Murray, Vimal Ranchhod, and Pippa Green (2018). *Taking stock of*

South African income inequality. Tech. rep. UNU-WIDER.

Lewis, Arthur W (1954). "Economic development with unlimited supplies of labour". In: *The Manchester School* 22.2, pp. 139–191.

Manning, Alan (2003). *Monopsony in motion: Imperfect competition in labor markets*. Princeton: Princeton University Press.

— (2020). "Monopsony in labor markets: a review". In: *ILR Review*.

Meghir, Costas, Renata Narita, and Jean-Marc Robin (2015). "Wages and informality in developing countries". In: *American Economic Review* 105.4, pp. 1509–46.

Mlatsheni, Cecil and Vimal Ranchhod (2017). "Youth labour market dynamics in South Africa: evidence from NIDS 1-2-3". In: *SALDRU, University of Cape Town*.

Moore, Brendan and Judith Scott-Clayton (2019). *The Firm's Role in Displaced Workers' Earnings Losses*. Tech. rep. National Bureau of Economic Research.

Muralidharan, Karthik, Paul Niehaus, and Sandip Sukhtankar (2017). *General equilibrium effects of (improving) public employment programs: Experimental evidence from India*. Tech. rep. National Bureau of Economic Research.

Naidu, Suresh, Eric A Posner, and Glen E Weyl (2018). "Antitrust Remedies for Labor Market Power". In: *Harvard Law Review*.

OECDstat (2016). *OECD-Stat Collective bargaining coverage*. URL: https://stats.oecd.org/Index.aspx?DataSetCode=CBC (visited on 11/14/2020).

Rothenberg, A. D, A. Gaduh, N. Burger, C. Chazali, I. Tjandraningsih, R. Radikun, C. Sutera, and S. Weilant (2016). "Rethinking Indonesia's informal sec- tor". In: *World Development* 80, pp. 96–113.

Saez, Emmanuel, Benjamin Schoefer, and David Seim (2019). "Payroll taxes, firm behavior, and rent sharing: Evidence from a young workers' tax cut in Sweden". In: *American Economic Review* 109.5, pp. 1717–63.

Soderbom, Mans, Francis Teal, and Anthony Wambugu (2005). "Unobserved heterogeneity and the relation between earnings and firm size: evidence from two developing countries". In: *Economics letters* 87.2, pp. 153–159.

Sokolova, Anna and Todd Sorensen (2021). "Monopsony in labor markets: A

meta-analysis". In: *ILR Review*

74.1, pp. 27–55.

Song, Jae et al. (2018). "Firming up inequality". In: *The Quarterly Journal of Economics* 134.1, pp. 1–50.

Sorkin, Isaac (2018). "Ranking firms using revealed preference". In: *The quarterly journal of economics* 133.3, pp. 1331–1393.

Ulyssea, Gabriel (2018). "Firms, informality, and development: Theory and evidence from Brazil". In: *American Economic Review* 108.8, pp. 2015–47.

World Bank (2020). *World Development Indicators*. URL: https : / / data . worldbank.org/indicator (visited on 09/22/2020).