UCD CENTRE FOR ECONOMIC RESEARCH

WORKING PAPER SERIES

2010

Is Informality Bad? - Evidence from Brazil, Mexico and South Africa

Olivier Bargain and Prudence Kwenda, University College Dublin

WP10/03

January 2010

UCD SCHOOL OF ECONOMICS
UNIVERSITY COLLEGE DUBLIN
BELFIELD DUBLIN 4

Is Informality Bad? Evidence from Brazil, Mexico and South Africa*

Olivier Bargain and Prudence Kwenda

December 2009

Abstract

The informal sector plays an important role in the functioning of labor markets in emerging economies. To characterize better this highly heterogeneous sector, we conduct a distributional analysis of the earnings gap between informal and formal employment in Brazil, Mexico and South Africa, distinguishing between dependent and independent workers. For each country, we use rich panel data to estimate fixed effects quantile regressions to control for (time-invariant) unobserved heterogeneity. The dual nature of the informal sector emerges from our results. In the high-tier segment, self-employed workers receive a significant earnings premium that may compensate the benefits obtained in formal jobs. In the lower end of the earnings distribution, both informal wage earners and independent (own account) workers face significant earnings penalties vis-à-vis the formal sector. Yet the dual structure is not balanced in the same way in all three countries. Most of the self-employment carries a premium in Mexico. In contrast, the upper-tier segment is marginal in South Africa, and informal workers, both dependent and independent, form a largely penalized group. More consistent with the competitive view, earnings differentials are small at all levels in Brazil.

Key Words : self-employed, salary work, informal sector, earnings differential, quantile regression, fixed effects model.

JEL Classification : J21, J23, J24, J31, O17

*__Acknowledgements__ : The authors are affiliated to University College Dublin (UCD) and Bargain is affiliated to IZA. We are grateful to participants/discussants of the 2009 RIW conference and IZA workshop for useful comments. The usual disclaimers apply. Correspondence to: Olivier Bargain, UCD, Newman Building, Dublin 4, Ireland. Phone: +35317168357. Email: olivier.bargain@ucd.ie

1 Introduction

The existence of large informal sectors in developing countries has often been cited as a central factor underlying wage inequality, persistent poverty and labor market inefficiencies. According to the traditional view (Fields, 1975, Dickens and Lang, 1985), salary workers enter informality to escape unemployment or because they are rationed out of the formal sector as a result of an overly regulated labor market. They earn less than *identical* workers in the formal sector – wages in the latter are set above market-clearing prices because of minimum wages, higher unionization or efficiency-wage explanations. In a similar way, (informal) self-employment is seen as a means of overcoming economic hardships in developing countries (Leibenstein, 1968), and several authors report that on average the self-employed earn less than workers in paid employment (e.g., Aronson, 1991, Carrington et al., 1996, Sullivan and Smeeding, 1997). Some authors have recently questioned this paradigm, arguing that an important fraction of informal jobs may reflect the voluntary choice of workers given their preferences, skill endowments and competing earnings prospects. Evidence has been particularly compelling for Latin America, pointing to better earnings prospect in self-employment than in paid (formal) employment (e.g., Maloney, 1998a, 1998b, 1999, 2004, Yamada, 1996, Saavedra and Chong, 1999, among others).

In this context, accurate measures of earnings differentials across sectors represent an important aspect of the analysis of labor markets in developing countries. While earnings equalization should eventually occur in competitive labor markets, persisting earnings gaps may reflect compensating differentials across sectors, e.g., social security benefits in the formal sector, tax avoidance, independence and job flexibility in self-employment.[1] Hence earnings gaps across sectors do not allow testing directly the hypothesis of segmentation on the labor market, but may nonetheless provide useful information (see the discussion in Maloney, 1999). When undertaking this type of analysis, however, the most difficult problem may be the huge heterogeneity of the informal sector. Some studies find no significant earnings differences on average between formal paid workers and the self-employed (Arias and Khamis, 2007) or between formal and informal salary workers (Badaoui et al., 2008). Yet comparisons at the mean may conceal important differences

[1] Implicit gains for informal salary workers are less obvious. Yet one may see informal jobs as a labor market entry point and a training area for young workers, or a type of employment with more flexible hours for married women. Workers in informal employment also avoid taxes/social security contributions while possibly attaching a low value to formal sector benefits, either because these services are traditionally provided through family support or because workers may be aware of inefficiencies in formal social protection.

along the earnings distribution – or may not characterize the returns of the majority of informal sector workers if a handful of prominent entrepreneurs push up the average earnings (see the discussion in Hamilton, 2000). Also, recent labor market modeling suggests adopting a dual representation whereby a competitive/voluntary entry informal sector, often associated with self-employment (cf., Arias et al, 2005), coexists with a rationed/segmented group (Funkhouser, 1997, Blunch et al., 2001, Maloney, 2004, Fields, 2005).[2] While such a representation is convenient for modeling purposes, the informal earnings gap may change gradually along the earnings distribution or with workers' attributes, and the size and nature of these high- and low-tier segments remains an open empirical question.

The present paper suggests an attempt to capture the diversity of the informal sector by estimating earnings gaps along two main dimensions. Firstly, we carefully distinguish between informal self-employment, informal salary work and formal salary work. We estimate the earnings penalties/premia carried by the two first states, using formal sector wage earners as the reference point.[3] Secondly, to depart from estimations of the mean earnings gap, we use quantile regression techniques and unveil more complex patterns from which the lower-tier and upper-tier informal segments of the labor market can be characterized.

More specifically, we focus on three countries which have received a lot of attention in the literature on informality, namely Brazil, South Africa and Mexico. Defining informality in the most comparable way across countries, we exploit large (rotating) panels to account for workers' unobserved heterogeneity. That is, we estimate 'fixed effects' panel regressions at different points of the earnings distribution as suggest by Koenker (2004) and Canay (2008). The dual nature of the informal sector, with upper and lower-tier segments, emerges from our results. In the upper part of the distribution, self-employment carries a significant premium that may compensate the benefits obtained in formal jobs, while the wage gap between formal and informal salary workers tends to disappear. In the lower end, both independent (own account) workers and informal wage earners face earnings penalties vis-à-vis the formal sector. Interestingly, the dual structure is not balanced

[2]Several interesting studies show that the dual representation of the informal sector proves to be a better alternative than polar models (see in particular Cunningham and Maloney, 2001, and Günther and Launov, 2006).

[3]Some studies consider in turn the comparison between formal and informal salary workers and between the latter and independent workers (e.g., Arias and Khamis, 2008, Bosch and Maloney, 2007). Some other studies focus exclusively on formal versus informal salaried work (see our companion paper Bargain and Kwenda, 2009, and Badaoui et al., 2008, among others). At the other extreme, some studies approximate informality by self-employment (for instance Yamada, 1996, for Peru).

in the same way in all three countries. While most of the self-employed workers receive a premium in Mexico, possibly very large at the top of the distribution, the upper-tier segment is marginal in South Africa. Informal workers, both dependent and independent, form a largely penalized group in this country. More consistent with the competitive view, earnings differentials are small at all levels in Brazil.

The paper is organized as follows. Section 2 briefly presents the labor markets in the three countries under study and the related literature on informality. Section 3 describes the data, the selection and the identification of informality. The econometric approach is detailed in section 4. Section 5 discusses the empirical results and reports robustness checks. Section 6 concludes.

2 Informality in Brazil, South Africa and Mexico

The question of informality has received a lot of attention in the literature. A large amount of evidence is summarized in Leontaridi (1998), Perry et al. (2007), Jütting et al. (2007), Ruffer and Knight (2007), among others. Many specific references for Brazil, South Africa and Mexico are to be found in these surveys and throughout the present paper. In this section, we simply provide a brief background description for each country.

The Brazilian labor market is characterized by stringent labor legislation and has experienced a series of economic crises. These two factors are often blamed for the high rates of informal paid workers, which account for approximately 30% of urban employees, and a growing self-employment (Moro et al., 2003).[4] Carneiro and Henley (2001) and Menezes-Filho et al. (2004) show that for some workers, the informal sector may be a desirable form of employment in Brazil; they also find that the large informal wage gap can be explained by selection bias and consequently favor the competitive markets hypothesis. This view seems to be supported by studies on sectoral mobility. Barros et al. (1990) find high mobility rates between sectors in the Sao Paulo region while Ruffer and Knight (2007) argue that there cannot be segmentation if there is such free mobility between sectors. In contrast, other studies report evidence of significant earnings differentials –

[4]Noticeably, the constitutional changes of 1988 have led to an increase in labor costs, a reduction in hours worked and a more relaxed role of trade unions (Barros and Corseuil, 2001, Bosch et al., 2007). Concerning macroeconomic crises, alternating periods of recession and high inflation may have contributed to the expansion of the informal sector, which accounts for 87% of the jobs created between 1992 and 2002. Also, trade liberalization in the early 1990s must have put some pressure on the tradable good sector, resulting in large movements of labor out of the (formal) manufacturing sector and into the informal part of the service sector, with relatively contained unemployment (Hoek, 2007).

that may favor the segmentation hypothesis – in the lower part of the earnings distribution (Tannuri-Pianto and Pianto, 2002).

Evidence is relatively clear for Mexico. Maloney (1999) reports that moves into (out of) self-employment are associated with significant increases (decreases) in earnings whilst moves from informal salaried work are associated with increases in earnings. Studying mobility patterns across business cycles, Bosch and Maloney (2007, 2008) confirm that a substantial part of the informal sector in Mexico (and Brazil), particularly the self-employed, likely corresponds to voluntary entry while informal salaried work may correspond more closely to the standard queuing view, especially for younger workers. Gong et al. (2004) find that entry and exit rates for the formal sector are lower than for the informal sector; the probability of formal sector employment increases with the education level, possibly in response to higher returns to education attached to formal jobs. Gong and van Soest (2002) confirm this view, suggesting that the dual structure is supported for highly educated workers but not for low-educated ones.

South Africa is somewhat different from Latin American or other African countries. The presence of unemployment and the relatively small size of the informal sector are partly on account of the potential barriers to entry – or hidden costs – in informal employment for those who are unemployed (Chandra et al., 2002). These are due in particular to land/credit constraints, inhibition of entrepreneurial skills resulting from the apartheid era and high crime rate against business owners (cf., Devey et al., 2003, Fields, 2006). Another reason, which rather applies to salary workers, is that reservation wages may be higher in South Africa compared to lower income countries. The unemployed who receive some support from within or beyond the household may prefer to remain outside the low-tier informal sector where real income is very low (Kingdon and Knight, 2001). Evidence regarding the involuntary nature of the informal sector is mixed, however. Several authors point toward sharp segmentation between the formal and informal segments of the labor market (Hofmeyr, 2002, Kingdon and Knight, 2007), highlighting the role of trade unions, collective bargaining and labor standards in 'registered' employment. Informal sector wages, being more subject to market forces, are about 60% lower according to Kingdon and Knight (2007). Yet some dynamic segments of the informal labor market also exist according to some studies. For the region of KwaZulu-Natal, Valodia et al. (2006) and Cichello et al. (2005) find that, for many workers, the informal sector has generated more employment and shown faster wage progression in the 1990s.

3 Measuring the Raw Earnings Gap

3.1 Data, Selection and Informality Definition

For Brazil, we make use of the Monthly Employment Survey (Pesquisa Mensal de Emprego, PME) conducted by the *Instituto Brasileiro de Geografiae Estatistica* (IBGE). This is a monthly household survey on the six largest metropolitan areas of Brazil (i.e., Belo Horizonte, Porto Alegre, Recife, Rio de Janeiro, Salvador and Sao Paulo). Households are interviewed four months in a row and re-interviewed eight months later for another four months. We create a panel with observations that are a year apart, focusing on years 2002 to 2007. For South Africa, we use the Labor Force Survey (LFS), a bi-annual rotating panel conducted by Statistics South Africa (Stats SA) and covering all provincial areas. Twenty percent of the sampling units are rotated out of the survey and replaced with a new sample every six months; workers are therefore observed five times at most over a two-and-a-half year period. We use the waves of September 2001 to March 2007. For Mexico, we use the Mexican National Occupation and Employment Survey (ENOE) conducted by the *Instituto Nacional de Estadistica, Geographica e Informatica* (INEGI). This is a quarterly survey where workers are observed at most five times over a five-quarter period. We use data from the first quarter 2005 to the third quarter 2008.

These surveys provide information about job characteristics, incomes, work duration, demographics and education. Since households are identified over time but individuals are not, we construct panels of individual workers by linking persons within households over time on the basis of gender, race and age. For the baseline estimates, we select workers that are observed at least twice in the data. The attrition resulting from this procedure corresponds to 30% of the initial sample for Brazil, 19% for South Africa and 17% for Mexico. In the last section, we check for possibly non-random attrition that could bias our results. We restrict our sample to urban men aged 15 to 65 years, not engaged in any form of education, and in full time employment in the private sector. We focus on men because in all three countries a large proportion of women are not active or are engaged in unpaid work – accounting for selection into the labor market is not yet standard in quantile estimations (see Albrecht et al., 2004). We select only workers in the private sector, which excludes unpaid family workers (whose implicit earnings are difficult to evaluate) and public sector employees; for the latter, there are indeed important differences in institutional mechanisms regulating wages, both across countries and compared to the private sector.

We opt for the legalistic/social protectionist definition of informality which refers to the

lack/avoidance of formal registration, taxation and labor standards and the lack of social security protection. This definition is important for welfare considerations as informal sector workers may experience earnings penalties on top of "bad" work conditions (e.g., no social protection). This is also a broader concept of informality as it recognizes the possible presence of unregistered/unprotected workers in large firms (Perry et al., 2007).[5]

More specifically, the group of informal salaried workers is identified on the basis of lack of compliance with labor legislation, which is relatively straightforward to capture in the surveys at use. In Mexico employees have to contribute to the social security agency (IMSS). Similarly, employees in Brazil must hold a labor card (*carteira assinada*), the signing of which guarantees them access to formal labor protection. Therefore those wage employees not registered with the social security agency in Mexico or not holding a signed labor card in Brazil are consider as informal salaried (similar definitions are used for instance in Amuedo-Dorantes, 2004, Tannuri-Pianto and Pianto, 2002, Bosch and Maloney, 2007, 2008). For South Africa, the LFS contains several questions regarding fringe benefits and other aspects of the job that can be used to identify the sector, in particular questions regarding whether the firm provides medical aid and deducts unemployment insurance contributions (see also Badaoui et al., 2008).

The group of independent or self-employed workers also belongs to a large extent to the informal sector as defined above or as characterized by the ILO. For Brazil, self-employed do have the legal obligation to pay social security contributions and Henley et al. (2006) report that around 95% do not do so. A relatively small group of self-employed in Mexico (less than 6% of all self-employed in our survey) satisfy the IMSS registration although under no legal obligation to do so. As in Bosch and Maloney (2008), we treat them as formal sector workers and drop them from the sample. In South Africa, although self-employed workers can make contributions to social security, we find that only 3% do so. The data allows us to identify those owners of a registered firm and who pay taxes (19.5% of all self-employed). Those who are registered/pay taxes or make social security contributions are excluded from our sample. As a further check, we find that very few self-employed workers own firms of more than five employees (15% in Brazil, less than 5% in Mexico and 10.5% in South Africa). Note that we also check the validity of the self-reported employment state with relevant information for each country.[6]

[5]ILO traditionally recommends classifying informal as workers in small establishments of fewer than 5-10 employees, who tend to be informal along different dimensions. Henley et al. (2006), Perry et al. (2007) and Bosch and Maloney (2008) show there is substantial overlap in these definitions. We have checked this for the datasets at use and find that it is broadly the case, except for Brazil where we found many informal (unregistered) salaried workers in large firms – see Bargain and Kwenda (2009) for more details.

[6]For instance, holding a working permit in Brazil should only apply to those in paid employment. We

This selection leaves an unbalanced panel of 22,186 men with 44,372 observations in Brazil; 9,237 men with 22,757 observations in South Africa and 107,465 men with 363,911 observations in Mexico. Summary statistics are reported in Table 1 and discussed below. We categorize workers in one of the three states, namely self-employed, formal and informal paid work. Self-employment accounts for 34% of total employment in Brazil, 10% in South Africa and 26% in Mexico. Informal salary work accounts for 12% of total employment in Brazil, 11% in South Africa and 33% in Mexico.

3.2 Earnings and Sample Description

We construct a measure of hourly earnings for all workers, using monthly gross earnings and reported work hours in the primary job. For the self-employed, information on monthly earnings does not allow distinguishing between returns to labor and to capital – we discuss this point in the concluding section. Earnings are adjusted over time using the national consumer price indices provided by the IBGE, Stats SA and the Central Bank of Mexico.

Table 1 presents the descriptive statistics for Brazil, South Africa and Mexico and table 2 completes the description of the samples by reporting the estimates of a multinomial logit of the workers' states (the reference is formal salary worker). In line with previous studies, we show in table 1 that self-employed workers earn more on average than wage earners in Latin America but earn less than formal salaried workers in South Africa. Other things being equal, self-employed workers are substantially older in Latin America and informal paid workers tend to be younger than all other workers in all three countries, which is consistent with the view that informal salaried work acts an entry point into the labor market. Another standard result is that highly educated workers are more concentrated in formal employment and, to a lesser extent, in self-employment. Figure 1 plots the proportions of workers' types in each decile of the pooled sample. All employment states are represented in all deciles but only formal sector employees are spread relatively equally over the earnings distribution. Informal paid workers are mainly concentrated in the lower end while self-employed workers are more often in the upper tails in Mexico and Brazil and in the lower deciles in South Africa.

drop the few self-employed worker who declare having such signed labor card (3.4% of them), interpreting it as an indication of misclassification. For South Africa, we use a question on whether a worker runs his/her own business. The marginal fraction declaring not to have their own enterprise is excluded from the sample. Data also allows distinguishing between owners of firms (i.e., employers) and individual independent workers (i.e., own account workers). Hence, we verify that own account workers are consistently located in firms of size equal to one (errors are of an order less than 1%).

4 Econometric Approach

We first estimate standard Mincer earnings equations at the mean and at various quantiles using pooled years data for each country. Explanatory variables comprise standard human capital information (age, age squared, education) and other individual/household characteristics as reported in table 1 (race, number of children, marital status, region) as well as broad industry dummies to control for the possible structural differences between sectors.

Next, we rely on (unbalanced) panel data to identify time-invariant unobserved heterogeneity. We first estimate a fixed effects model (FE) for each country and compare the result to standard OLS. We also extend this approach to the whole distribution by estimating fixed effects quantile regression (FE-QR), to be compared to the results of the standard quantile regression (QR). Denote I_{it} (resp. S_{it}) a dummy taking value one if person i observed at time t is informal salary worker (resp. self-employed). Denote x_{it} a set of controls, α_i the time-invariant heterogeneity (the individual fixed effect) and ε_{it} an i.i.d. normally distributed stochastic term accounting for possible measurement error. The FE model is simply written:

$$y_{it} = \alpha_i + \gamma_t + x_{it}\beta + \delta I_{it} + \rho S_{it} + \varepsilon_{it}$$

where $E\left[\varepsilon_{it}\left|\alpha_i, x_{it}, I_{it}, S_{it}\right.\right] = 0$ for all individuals i and periods t. The FE estimator is consistent even if unobserved characteristics are correlated with both selection and earnings, as long as those characteristics are constant over time. The estimated coefficients $\widehat{\delta}$ and $\widehat{\rho}$ are interpreted as a measure of the conditional earnings premium/penalty experienced by informal salary workers and self-employed workers respectively, compared to formal wage earners. These conditional earnings gaps are derived from the comparison between those who move between employment states and the 'stayers'. Denote $C = 1, 2, 3$ the three different states, respectively self-employed, informal salary workers and formal salary workers. Let us illustrate the identification by a simple two-period example and three of the possible cases:

$$\begin{aligned}
E\left[y_{i2} - y_{i1} | C_{i1} = k, C_{i2} = k\right] &= \Delta \text{ for } k = 1, 2, 3 \\
E\left[y_{i2} - y_{i1} | C_{i1} = 1, C_{i2} = 3\right] &= \Delta - \rho \\
E\left[y_{i2} - y_{i1} | C_{i1} = 2, S_{i2} = 3\right] &= \Delta - \delta \\
\text{with } \Delta &= \gamma_2 - \gamma_1 + (x_{i2} - x_{i1})\beta
\end{aligned}$$

The change in earnings for those moving into formal employment and coming from self-employment (second line) or informal salary work (third line) contribute to identify the

premium/penalty of the two latter states compared to formal sector remunerations, together with the stayer of all types (first line). Identification is completed by the movers obtained by all the other possible permutations between states. Note that at this stage, we do not account for possible differences in the earnings gaps whether it is identified on workers moving from formal to informal sectors or on those moving in the other direction, but allow for asymmetrical effects in the last section.

The extension of the standard QR model to longitudinal data goes as follows. For any worker i, we can write the τ^{th} quantile of the y distribution conditionally on observables as:

$$F^{-1}_{y_{it}}(\tau \mid x_{it}) = \alpha_i + \gamma_t(\tau) + x_{it}\beta(\tau) + \delta(\tau)I_{it} + \rho(\tau)S_{it}, \forall\tau \in [0,1].$$

Fixed effects α's have a pure *location* shift effect on the conditional quantiles of the response (i.e., they affect all quantiles in the same way). We can use Koenker (2004)'s approach to estimate this model or the alternative and simpler approach suggested by Canay (2008). The latter exploits the assumption that α terms are pure location shifters, so that they can be estimated in a first step by traditional mean estimations (for instance by OLS estimator in first differences). Then it is possible to use the estimated $\widehat{\alpha}_i$ in order to regress corrected earnings $\widehat{y}_i = y_i - \widehat{\alpha}_i$ on the other covariates by traditional QR.

5 Empirical Results

5.1 Main Results

Our main results are represented in figures 2, 3 and 4 and commented below. For each country, we report the estimated coefficient $\widehat{\delta}$ and $\widehat{\rho}$, i.e., the earnings penalties/premia from informal salary work and informal self-employment compared to the formal sector. The left panel of each graph shows the estimates from OLS and QR on pooled years while the right panel depicts the estimates of the FE and FE-QR on panel data. Dashed lines and empty diamonds represent the bootstrapped 95% confidence intervals. For each country, we can see that QR (resp. FE-QR) estimates are not all contained in the interval surrounding the OLS (resp. FE) coefficient and reveal important differences along the earnings distribution. In table 3, we also report the earnings penalty at the mean, the median and two extreme quantiles as well as the bootstrapped standard errors.[7]

For *Brazil*, OLS estimates indicate that on average self-employed workers receive an earnings premium of 11% compared to formal sector wage earners, while informal salary

[7] The full estimation tables, not reported because of space limitation, are available from the authors.

workers face a mean penalty of 9%. The left panel of Figure 2 shows a more complex pattern, revealing that the self-employed workers face a moderate penalty in the bottom of the distribution but a premium in the upper 75%, which increases with earning levels. The informal sector penalty for salary workers is decreasing with the earning levels, i.e., it is significant at the bottom (around 15%) but disappears at the top. Results in the right panel show that when accounting for fixed effects, the penalty for informal salary workers decreases by around a third. The self-employment premium decreases by around 10 points in the upper half while the penalty at the bottom increases slightly – significant penalties for self-employed workers now concern the first half of the distribution. Thus, it appears that formal salary workers have "better" unobserved skills than informal sector counterparts in the first half of the distribution but "worse" unobserved characteristics than self-employed workers in the second half.

For *South Africa*, we find an average informal sector penalty of around 62% for salary workers and 30% for self-employed. The left panel of figure 3 shows little variation along the earnings distribution for the former group but a decreasing penalty with earning levels for the latter, turning into a premium at the top. When accounting for unobserved heterogeneity, the earning penalty faced by both dependent and independent workers of the informal sector decreases at all levels and especially in the lower part of the distribution, with a drop of $20 - 30$ percentage points. Independent workers of the top quarter of the distribution receive a moderate premium while the penalty faced by informal salary workers decreases with earnings levels and tends to vanish at the top.

Results of pooled QR for *Mexico* are relatively similar to what is found in Brazil, with a premium of around 13% on average for self-employed workers and a penalty of around 15% for informal salary workers. The left panel of figure 4 shows that for both types of worker, the earnings differential to formal employment increases with earnings levels. When accounting for fixed effects, the self-employed premium increases slightly (up to 16% on average) and the penalty faced by informal salary workers becomes very moderate. The self-employment premium, very large at the top, is observed at all points of the distribution except the lower 15%.[8]

[8] Very similarly, Cunningham and Maloney (2001) report that 13% of the self-employed workers are in the lower-tier informal sector, using cluster analysis techniques and defining lower- and upper-tiers on the basis of the capital intensity per worker, earnings, hours worked, firm life, and education.

5.2 Discussion and Additional Results

Most interestingly, results above show a similar pattern in all three countries. Firstly, time-invariant unobservables are an important factor behind the observed earnings gaps, even after controlling for a rich set of characteristics. Secondly, formal salary workers have "better" unobserved skills than informal dependent and independent workers in all three countries, with the exception of independent workers in Brazil. Thirdly, when controlling for unobservables, it appears that the upper-tier segment of the informal sector is comprised of self-employed workers and is the most rewarding type of employment of all. In contrast, the lower end of the earning scale is mainly composed of informal salary workers and, to some extent, of self-employed workers that we examine below.

Fundamental differences across countries appear in the relative size of each segment, mostly reflecting the specific nature of independent employment in the three countries under study. Mexico and South Africa can be seen as two polar cases while Brazil lies somewhere in-between. Self-employment appears to be a desirable segment of the Mexican labor market, dominating salary work, both formal and informal, at almost all points of the earnings distribution. This is line with previous studies and conveys that there may be a 'Mexican exception' (Maloney, 1999, Gonzalez and Maloney, 1999, Marcouiller et al., 1997). In contrast, the upper-tier segment is marginal in South Africa and earnings penalties for both dependent and independent workers in the informal sector can be substantial, certainly too large to be justified by compensating differentials. This is consistent with the traditional view that formal employment is the desirable outcome in this country and that informal sector workers may suffer from segmentation (Kingdon and Knight, 2001, 2007, Hofmeyr, 2002, Ruffer and Knight, 2007). This is also in line with the fact that unemployment is a frequent substitute to informal employment for those queuing for formal jobs. More consistent with the competitive view, earnings differentials are small at all levels in Brazil.

In the rest of the paper, we focus essentially on self-employed workers – a detailed analysis of the informal salary workers is conducted in Bargain and Kwenda (2009). To complement the results above, we pinpoint the independent workers at the two ends of the distribution. For each country, we estimate a multinomial logit on the sub-sample of self-employed, with three categories defined as the bottom segment ($0.1 - 0.2$ quantiles of the pooled sample), the middle segment ($0.3 - 0.7$ quantiles) and the top segment ($0.8 - 0.9$ quantiles). Our estimated coefficients refer to the probability of being located at the top or the bottom of the earnings distribution relative to the middle segment, used as the reference group. Table 4 shows that independent workers located at the bottom of

the distribution in Mexico and Brazil are more likely to be own account workers, young, with primary or no education and in service work or elementary occupations. For these two countries, those at the top are more often employers (firm owners, as opposed to own account worker), older and with higher education. These patterns are much less clear in South Africa.

The FE-QR model simply uses dummy variables for informal salary work and self-employment, and may be seen as misspecified. While it is well known that, in case of misspecification, least square regression provides a minimum mean squared error linear approximation to the true functions, Angrist et al. (2006) provide a similar result for quantile regression. Therefore our findings have meaningful interpretation even if the true informal wage penalty depends on the covariates. Nonetheless, we can examine the heterogeneity of the informal earnings gaps by simply interacting the informal sector dummies with workers' age and education levels. Results are reported in the panel A of figure 6 for the self-employment penalty/premium. It shows that the between-group variation is not important compared to within-group heterogeneity (i.e., the variation of the earnings gap along the distribution) in Mexico. Variation in age and education can affect the earnings premium/penalty more significantly in the other countries. In particular, having low education increases the penalty by $3-7$ points in the first half of the distribution in Brazil, and being a younger worker (age 25) increases the penalty by $16-20$ points in the first half in South Africa.

5.3 Robustness Checks

The identification of FE on movers is standard but one must verify that the number of moves across sectors is large enough for a valid use of this estimator. We calculate the proportion of all panel transitions (i.e., all pairs of observations for the same workers) that correspond to a move. We find that 5%, 8% and 10% of all transitions are moves between formal and informal salary work (either ways) for Brazil, South Africa and Mexico respectively, which correspond to $1,117$, $1,088$ and $25,028$ movers respectively. Moves between formal salary work and self-employment represent 4%, 4% and 3% of all transitions, corresponding to 807, 500 and $6,650$ movers respectively. These are reassuring numbers regarding the possibility to identify the parameter of interest.[9]

We also check that movers are not too specific. Firstly, one may expect that cross-sector moves are limited to specific groups of workers, for instance those with the least

[9] Note that the identification of the earnings gaps is completed by the moves between informal salaried workers, which correspond to 5%, 3% and 9% of all transitions respectively.

earnings who are in search of better job prospects. In this case, our estimates could be biased. Figure 5 depicts the proportion of movers per quintiles of the "initial" earnings distribution, i.e., using the periods where workers are first observed in the pooled sample. This picture shows that moves are relatively spread over the whole distribution in Brazil and South Africa, although a larger number of moves between formal and informal salary work occurs at the bottom. The high frequency of moves in Mexico (around 20% of all panel transitions) occurs at all earnings level – even if moves between formal sector and self-employment are more frequent in the upper part. This overall picture is reassuring since moves are not overly concentrated in some parts of the earnings scale.[10]

Another aspect of the identification strategy that merits discussion is the assumption that the earnings penalty is the same for those that move from the informal sector to the formal sector as it is for those that move in the opposite direction. If all the unobservable heterogeneity is time-invariant, as assumed in the FE estimator, then this is not an issue. However, with the traditional view of self-employment as a safety net for those losing preferred formal sector jobs, one would expect that moves into informal self-employment are more often the result of time-specific negative shocks (e.g., productivity shocks), and that moves in this direction capture larger penalties (or smaller premia) than moves in the opposite direction. We replicate our results when including only one type of move at a time. Graphs in panel B of figure 6 show that results are not fundamentally asymmetrical.

Finally, we check that panel attrition does not lead to some bias by selecting out a specific type of workers. Indeed our baseline estimation excluded all workers observed only once in the data. However, it might be expected that workers in the informal sector are more likely to exit from the panel because of higher migration or higher misreporting. To check for possible non-random attrition, we simply estimate QR on pooled years for those observed only once in the data and compare the estimated earnings gaps to baseline results. The panel C of figure 6 shows that results are very similar in both cases, conveying that sample attrition does not relate to labor market states.

6 Concluding Discussion

Using large panel data for Brazil, South Africa and Mexico, we estimate the conditional earnings gaps between formal and informal sectors along the earnings distribution, distin-

[10] We refrain from drawing any conclusions based on these 'raw' transitions. A more in-depth interpretation of inter-sector flows would require some adjustments for turnover and job creation as performed in Bosch and Maloney (1997) and Maloney (1999), which is naturally beyond the scope of the present paper.

guishing informal wage earners from self-employed workers. Time-invariant unobserved heterogeneity seems to play an important role in explaining the earnings differentials, even after controlling for a rich set of characteristics. A consistent result across countries is that informal salaried workers are to a large extent the least paid group in the urban workforce. The lower-tier segment of the informal sector is also composed of young and unskilled own-account owners. The earnings penalty faced by these groups is moderate in Latin America but very large in South Africa, in line with the traditional view. The upper segment of the informal labor market is composed of firm owners in all three countries, who fare better than formal sector workers at the top of the earnings distribution. Yet this segment is marginal in South Africa while most of the entrepreneurs in Mexico receive earnings premia. More consistent with the competitive view, earnings differentials are small at all levels in Brazil.

We conclude with a series of comments on the present approach. Firstly, we have used panel information to purge our estimations from time-invariant unobservable heterogeneity. Extending the approach to time-varying unobservables would require to model selection explicitly. Yet it seems extremely challenging to find proper instruments, i.e., instruments that can convincingly explain selection into a given sector (but not earnings) and that also vary over time. Secondly, as in many studies, we have compared self-employment income to formal sector wages on the basis of hourly earnings. To distinguish between wages and profits for the independent workers, Headen (1990) suggests to predict the returns to labor for the self-employed using wage estimations on employees. Under the assumption of equal returns to labor for both dependent and independent workers in the informal sector, the earnings gap between self-employed and informal employees would give a measure of the returns to capital. It is obtained simply by comparing the two estimated earnings gaps obtained in the paper (that is, relative to the formal sector). This difference is not significantly different from zero at the bottom of the distribution, which is consistent with little capital-intensive businesses in the lower-tier, and increases gradually with earnings levels. At the top of the distribution, this difference amounts to around 10% of the informal sector wages in Brazil, 20% in South Africa and 40% in Mexico. Yet the assumption of equal returns to labor remains to be discussed. One may also argue that similar corrections should be made for salary workers, i.e., adjusting wages of salary workers for private investment in human capital. Finally, gross earnings gaps could be corrected for income taxes and social contributions paid in the formal sector (as in Badaoui et al., 2008, and Bargain and Kwenda, 2009) – yet informal self-employment should be better identified for that purpose. Accounting for non-pecuniary advantages attached to a particular sector, and above all for medical benefits and pensions paid in

the formal sector, represents a considerable challenge, given data limitation. Yet this is a fundamental and necessary improvement for more comprehensive welfare analyses.[11]

References

[1] Albrecht, J.W., A van Vuuren and S. Vroman (2004): "Decomposing the gender wage gap in the Netherlands with sample selection adjustments", forthcoming in *labor Economics*

[2] Amuedo-Dorantes, C. (2004): "Determinants and Poverty Implications of Informal Sector Work in Chile", in *Economic Development and Cultural Change*, 52(2), 349-368.

[3] Angrist, J., V. Chernozhukov and I. Fernández-Val (2006): "Quantile Regression under Misspecification, with an Application to the U.S. Wage Structure", *Econometrica*, 74(2), 539-563

[4] Arias, O. and M. Khamis (2008): "Comparative Advantage, Segmentation and Informal Earnings: A Marginal Treatment Effects Approach", IZA DP No. 3916

[5] Arias, O., A. Blom, M. Bosch, W. Cunningham, A. Fiszbein, G. Lopez Acevedo, W. Maloney, J. Saavedra, C. Sanchez-Paramo, M. Santamaria, Siga (2005): "Pending issues in protection, productivity growth, and poverty reduction," Policy Research Working Paper Series 3799, The World Bank.

[6] Aronson, R. L. (1991), *Self-Employment: A Labor Market Perspective*. Ithaca, N.Y. ILR Press

[7] Badaoui, E. Strobl, E. and Walsh, F (2008): "Is There an Informal Employment Wage Penalty? Evidence from South Africa", *Economic Development and Cultural Change*, 56, 683–710

[8] Bargain, O. and P. Kwenda (2009): "The informal wage gap - new evidence using panel data", IZA working paper.

[9] Barros, R.P. and Corseuil, C.H. (2001): "The impact of regulations on Brazilian labor market performance", Research Network Working Paper No. R-427, Inter-American Development Bank.

[10] Blau, D.M., (1985): "Self-employment and Self-Selection in Developing Country Labor Markets", *Southern Economic Journal*, 51, 2, 351-63

[11] For instance, see the discussion in Bourguignon et al. (2007) on the role of pensions on income inequality in Brazil.

[11] Blunch, N-H., S. Canagarajah., and D. Raju (2001): "The Informal Sector Revisited: A Synthesis Across Space and Time", Social Protection Discussion Paper Series No. 0119. Social Protection Unit Human Development Network The World Bank.

[12] Bosch and W. F. Maloney (2007): "Comparative Analysis of Labor Market Dynamics Using Markov Processes: An Application to Informality," IZA Discussion Papers 3038, Institute for the Study of Labor (IZA).

[13] Bosch, M. and W. Maloney (2008): "Cyclical Movements in Unemployment and Informality in Developing Countries", World Bank

[14] Bosch, M., E. Goni and W. Maloney (2007): "The Determinants of Rising Informality in Brazil: Evidence from Gross Worker Flows", IZA Discussion Paper 2970

[15] Bourguignon, F., F. Ferreira and P. Leite (2007): "Beyond Oaxaca–Blinder: Accounting for differences in household income distributions", *Journal of Economic Inequality*.

[16] Canay, I. A. (2008): "A Note on Quantile Regression for Panel Data Models", Department of Economics, Northwestern University

[17] Carneiro, F.G. and A. Henley (2001): "modeling formal vs informal employment and earnings: microeconometric evidence for Brazil", *Annals of the XXIX National Meeting of Economics of ANPEC*.

[18] Carrington, W. J., M. Kristin, and P. Brooks (1996): "The Role of Employer-Employee Interactions in Labor Market Cycles: Evidence from the Self-Employed", *Journal of Labor Economics*, 14 (4), 571-602

[19] Chandra, V., J. Nganou and C. Noel (2002): "Constraints to growth in Johannesburg's black informal sector: Evidence from the 1999 informal sector survey", World Bank Report No. 24449-ZA.

[20] Cichello, P.L., G.S. Fields, M. Leibbrandt (2005): "Earnings and Employment Dynamics for Africans in Post-apartheid South Africa: A Panel Study of KwaZulu-Natal", *Journal of African Economies*, 1-48.

[21] Cunningham, W. and W.F. Maloney (2001): "Heterogeneity among Mexico's microenterprises: an application of factor and cluster analysis", *Economic Development and Cultural Change*, 50(1) 131-56.

[22] Devey, R., C. Skinner, and I. Valodia (2003): "Informal Economy Employment Data in South Africa: A critical Analysis", Report prepared for the Employment Data Research Group, Human Sciences Research Council.

[23] Dickens, W. and K. Lang (1985): "A Test of the Dual labor Market Theory", *American Economic Review*, Vol. 75, No. 4, pp.792-805

[24] Fields,G. S. (1975): "Rural-Urban Migration, Urban Unemployment and Underemployment, and job search activity in LDCs" *Journal of Development Economics,* Vol. 2, pp. 165-187.

[25] Fields,G. S. (2005): "A Guide to Multisector labor Markets Models", Paper prepared for the World Bank labor Market Conference.

[26] Fields, G. S. (2006): "modeling labor market policy in developing countries: A selective review of the literature and needs for the future", mimeo., December, Cornell University.

[27] Funkhouser, E (1997): "Mobility and labor Market Segmentation: The Urban labor Market in El Salvador", *Economic Development and Cultural Change*, Vol. 46, No.1, pp.123-153

[28] Gong, X., A. van Soest, and E. Villagomez (2004): "Mobility in the urban labor market: a panel data analysis for Mexico", *Economic Development and Cultural Change*

[29] Gong, X. and A. van Soest (2002): "Wage differentials and mobility in the urban labor market: a panel data analysis for Mexico", *Labor Economics*, Vol. 9, Issue 4, pp. 513-529

[30] Gonzalez, P.A. and Maloney, W. F (1999): "Logit analysis in a rotating panel context and an application to self-employment decisions," Policy Research Working Paper Series 2069, The World Bank

[31] Günther, I. and A. Launov (2007): "Competitive and Segmented Informal labor Markets", IZA Discussion Papers 2349, Institute for the Study of labor (IZA).

[32] Hamilton, B.H. (2000) : "Does entrepreneurship pay? An empirical analysis of the returns to self-employment", *Journal of Political Economy*, Vol.108 No.3, pp. 604-631.

[33] Henley, A., G.R. Arabsheibani, and F.G. Carneiro (2007): "On defining and measuring the informal sector", Policy research working paper 3866, World Bank.

[34] Hofmeyr, J.F. (2002): "The importance of segmentation in the South African labor market", University of Natal.

[35] Jütting, J., T. Xenogiani and J. Parlevliet (2007): "Work and Well-Being: Informal employment revisited", OECD Development Centre

[36] Kingdon, G., and Knight, J. (2001) "Why high open unemployment and small informal sector in South Africa?" Centre for the Study of African Economies, Department of Economics, University of Oxford, October 2001.

[37] Kingdon, G. and J. Knight (2007): 'Unemployment in South Africa, 1995-2003: Causes, problems and policies', *Journal of African Economies*, forthcoming.

[38] Koenker, R. (2004): "Quantile Regression for Longitudinal Data", *Journal of Multivariate Analysis*, 91, 74-89.

[39] Leibenstein, H. (1968), "Entrepreneurship and Development", *American Economic Review*, 58 72-83

[40] Leontaridi, M. (1998): "Segmented labor Markets: Theory and Evidence." *Journal of Economic Surveys* 12(1): pp. 103–109.

[41] Maloney, W. F. (1998a): "The structure of labor markets in developing countries: Time series evidence on competing views. Working Paper 1940, World Bank, Washington, DC

[42] Maloney, W. F. (1998b): "Are labor markets in developing countries dualistic?," Policy Research Working Paper Series 1941, The World Bank.

[43] Maloney,W.F. (1999): "Does Informality Imply Segmentation in Urban labor Markets? Evidence from Sectoral Transitions in Mexico", *World Bank Economic Review,* Vol.13, No.3, pp.275-302

[44] Maloney,W.F. (2004): "Informality revisited", *World Development*, Vol 32 , No7, pp.1159-78

[45] Marcouiller, D. and V. Ruiz de Castilla and C. Woodruff (1997): "Formal Measures of the Informal-Sector Wage Gap in Mexico, El Salvador, and Peru", *Economic Development and Cultural Change*, University of Chicago Press, Vol. 45, No.2, pp 367-92.

[46] Menezes-Filho, N., N. Mendes, and E. Almeida (2004): "O diferencial de salarios formal-informal no brasil: Segmentação ou vies de seleção?", *Revista Brasileira de Economia* 58 (2).

[47] Moro, S., F. Chein, and A. F. Machado (2003): "Self-employment in Brazil and its determinants: a spatial analysis," Textos para DiscussÃ£o Cedeplar-UFMG td204, Cedeplar, Universidade Federal de Minas Gerais.

[48] Packard, T.G. G. (2007): "Do Workers in Chile Choose Informal Employment? A Dynamic Analysis of Sector Choice", World Bank Policy Research Working Paper No. 4232.

[49] Perry, G., W. Maloney, O. Arias, P. Fajnzylber, A. Mason, J. Saavedra-Chanduvi (2006, eds.), *Informality: Exit and Exclusion*, World Bank latin America and Caribbean Studies.

[50] Ruffer, T. and J. Knight (2007): "Informal sector labor markets in developing countries", University of Oxford

[51] Saavedra, J. and A. Chong (1999): "Structural reforms, institutions and earnings: evidence from the formal and informal sectors in urban Peru", *Journal of Development Studies*, 35(4), 95-116

[52] Sullivan, D.H. and T.M. Smeeding (1997): "All the World's entrepreneurs: The Role of Self-employment in Nineteen Nations", Luxembourg Income Study Working Paper Series No.163

[53] Tannuri-Pianto, M.E. and D.M. Pianto (2002): "Informal Employment in Brazil – A Choice at the Top and Segmentation at the Bottom: A Quantile Regression Approach", Department of Economics Working Paper 236, University of Brasilia.

[54] Yamada, G. (1996) : "Urban Informal Employment and Self-employment in Developing Countries: Theory and Evidence." *Economic Development and Cultural Change*, 44, 289-314.

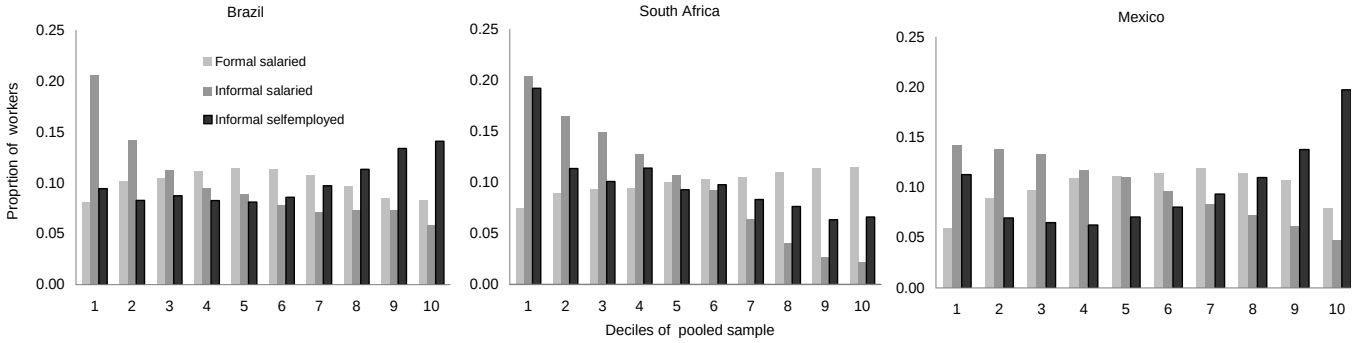
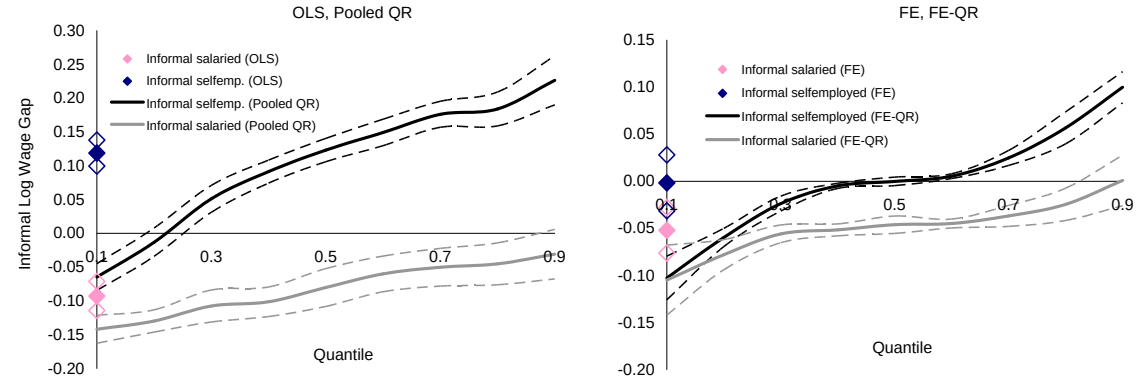
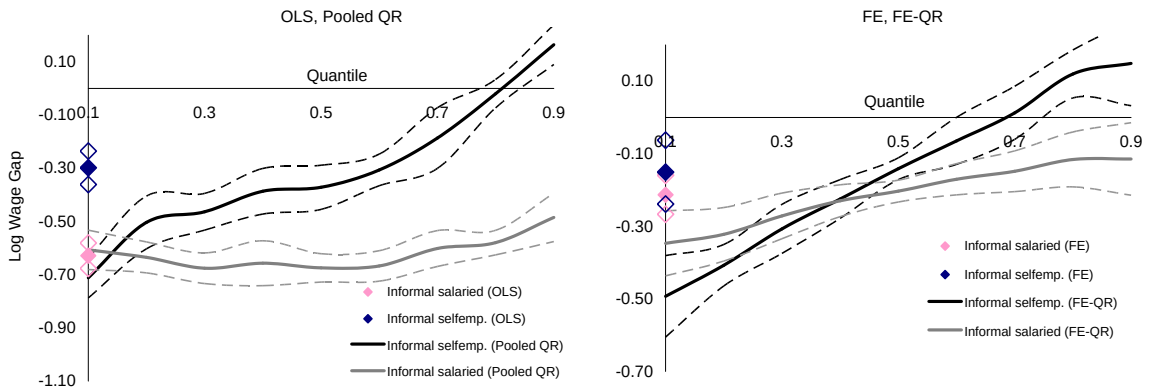


Figure 1: Distribution of Workers across Sectors



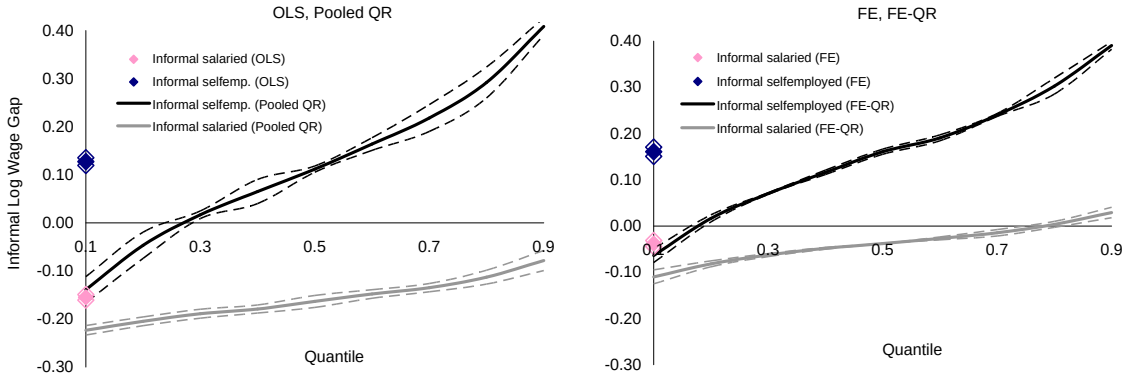
Note: bootstrapped 95% confidence intervals represented by dashed lines and empty diamonds

Figure 2: Estimated Wage Gaps (Brazil)



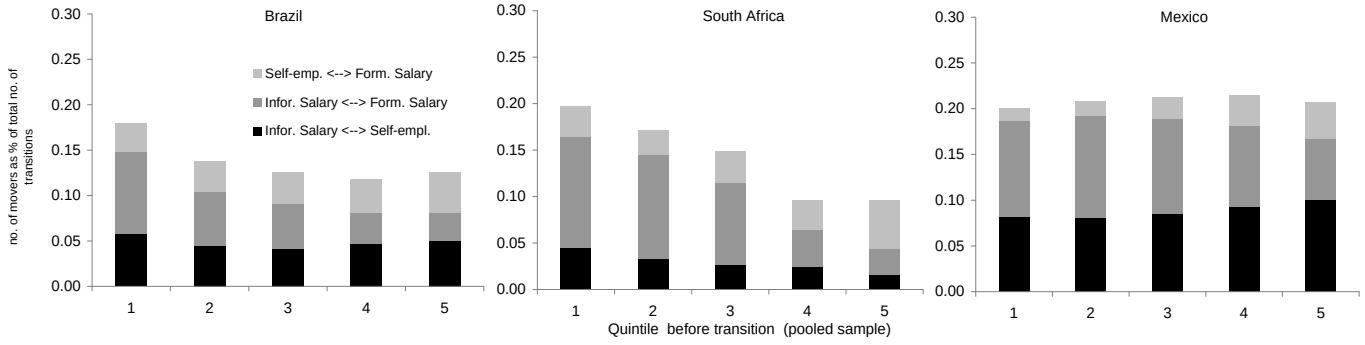
Note: bootstrapped 95% confidence intervals represented by dashed lines and empty diamonds

Figure 3: Estimated Wage Gaps (South Africa)



Note: bootstrapped 95% confidence intervals represented by dashed lines and empty diamonds

Figure 4: Estimated Wage Gaps (Mexico)



The graphs represent the number of movers in % of the total number of transitions in the panels (including both stayers and movers-type of transitions). Quintiles are calculated on the basis of the period where workers are first observed in the pooled sample . Total number of transitions = 22,186 in Brazil; 14,026 in South Africa; 263,082 in Mexico

Figure 5: Transitions across Employment States

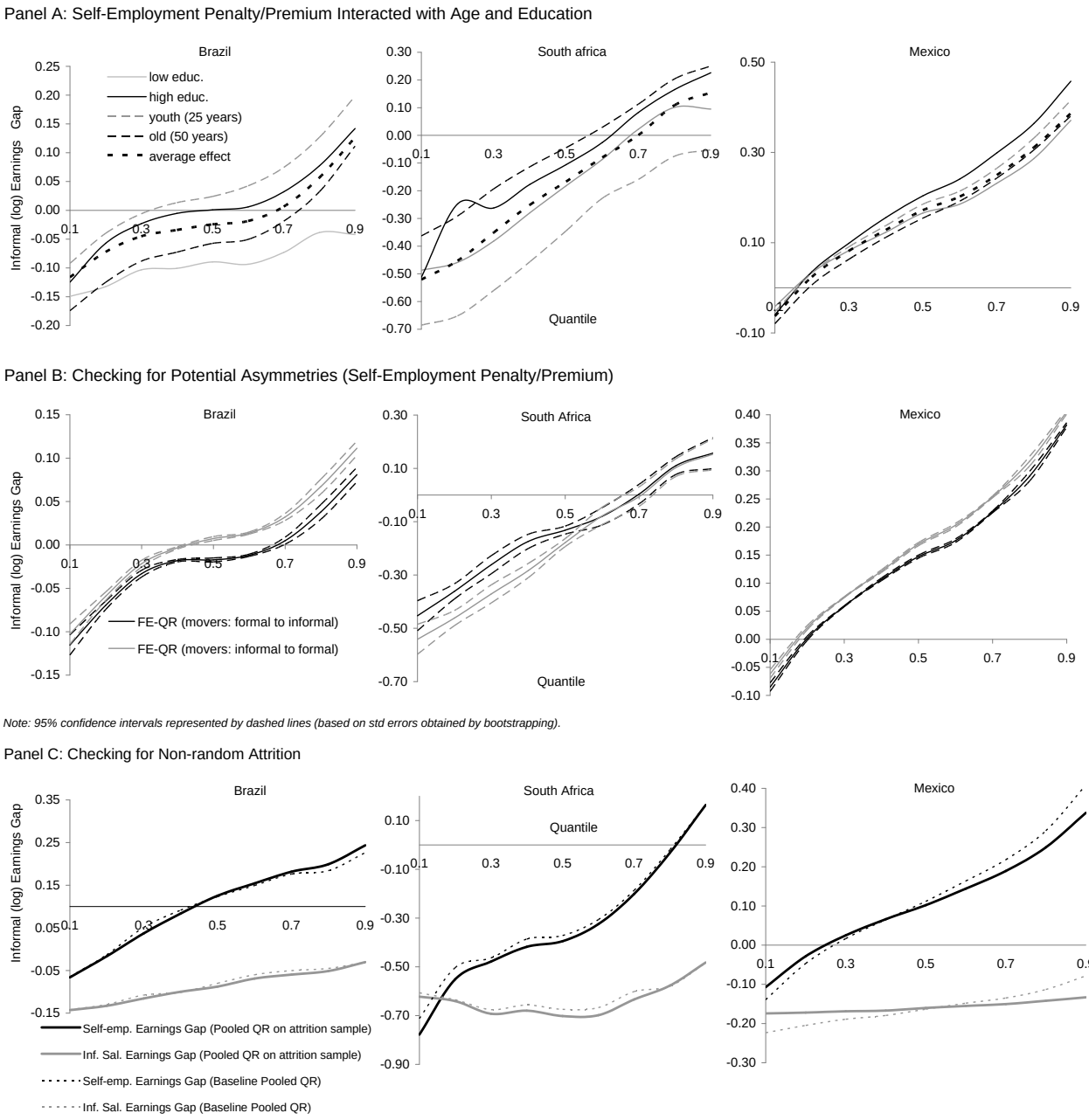


Figure 6: Extensions and Robustness Checks

Table 1: Selected Samples: Descriptive Statistics

	Brazil			South Africa				Mexico		
	Formal salaried	Informal salaried	Self-employed		Formal salaried	Informal salaried	Self-employed		Formal salaried	Informal salaried	Self-employed
Hourly gross earnings	4.79	3.79	6.39		2.54	1.06	1.79		2.78	2.02	3.82
Std dev.	(6.97)	(5.39)	(10.76)		(3.69)	(1.71)	(3.43)		(0.50)	(0.54)	(0.75)
Demographics											
Age	36.6	36.6	43.3		38.5	39.2	41.6		34.6	32.5	41.8
household size	3.8	3.9	3.7		5.9	6.4	7.7		4.6	4.9	4.5
# children (0-10 years)	3.2	3.9	3.2		1.7	2.2	2.7		1.8	1.6	2.6
Married	0.64	0.56	0.73		0.63	0.49	0.66		0.62	0.45	0.66
Black	0.07	0.07	0.06	Black	0.74	0.86	0.94				
Brown	0.32	0.31	0.27	Coloured	0.26	0.14	0.06				
White	0.61	0.61	0.67								
Education											
No Schooling	0.01	0.02	0.02	No schooling	0.09	0.15	0.09	No education	0.02	0.05	0.05
1-3 Years	0.04	0.06	0.06	Primary	0.31	0.40	0.33	1-3 Years	0.04	0.08	0.10
4-7 Years	0.24	0.31	0.29	Secondary	0.53	0.42	0.52	4-7 Years	0.24	0.34	0.34
8-10 Years	0.18	0.19	0.16	Vocational	0.07	0.03	0.05	8-10 Years	0.45	0.40	0.32
11+ Years	0.53	0.43	0.48	University	0.01	0.00	0.01	11+ Years	0.25	0.13	0.19
Region											
Recife	0.06	0.05	0.05	Western Cape	0.21	0.10	0.06	>100,000 Inhab.	0.72	0.56	0.64
Salvador	0.07	0.06	0.07	Eastern Cape	0.09	0.16	0.12	15,000-99,999 Inhab.	0.11	0.16	0.15
Belo Horizonte	0.16	0.12	0.16	Northern Cape	0.08	0.05	0.02	2,500-14,999 Inhab.	0.08	0.14	0.12
Rio de Janeiro	0.27	0.32	0.31	Free State	0.11	0.07	0.09	< 2,500 Inhab.	0.08	0.13	0.10
Sao Paulo	0.25	0.29	0.24	Kwazulu-Natal	0.11	0.15	0.16				
Porto Alegre	0.19	0.16	0.17	North West	0.11	0.13	0.10				
				Gauteng	0.12	0.12	0.17				
				Mpumalanga	0.10	0.10	0.15				
				Limpopo	0.05	0.12	0.14				
Economic sector											
Manufacturing	0.31	0.18	0.12		0.35	0.08	0.10		0.36	0.17	0.14
Construction	0.07	0.15	0.21		0.12	0.26	0.20		0.10	0.34	0.25
Trade & Retail	0.23	0.29	0.35		0.28	0.18	0.52		0.24	0.14	0.23
Services	0.24	0.23	0.19		0.17	0.33	0.10		0.19	0.20	0.27
Transport and Comm	0.13	0.11	0.11		0.07	0.14	0.09		0.09	0.15	0.11
# panel observations	24,105	5,200	15,067		18,102	2,444	2,211		151,370	118,357	94,184
share of workers	54%	12%	34%		79%	11%	10%		42%	33%	26%

Statistics concern the selected sample of male aged 15-65, neither in education nor in the public sector. Data covers the period 2002-2007 for Brazil, 2001-2007 for South Africa and 2005-2008 for Mexico. Hourly earnings in 2002 PPP international $.

Table 2: Selected Samples: Multinomial Logit of Workers' Status

	Brazil					South Africa					Mexico			
	Informal salaried		Self-employed			Informal salaried		Self-employed			Informal salaried		Self-employed	
Demographics	*Ref: white, single*					*Ref: black, sinlge*					*Ref: sinlge*			
Age	-0.141	(0.012)	0.121	(0.010)		-0.072	(0.016)	-0.021	(0.021)		-0.106	(0.004)	0.138	(0.005)
Age squared	0.002	(0.000)	-0.001	(0.000)		0.001	(0.000)	0.001	(0.000)		0.001	(0.000)	-0.001	(0.000)
Married	-0.340	(0.045)	-0.044	(0.036)		-0.510	(0.071)	-0.114	(0.083)		-0.524	(0.016)	-0.260	(0.019)
Household size	0.068	(0.024)	0.008	(0.020)		-0.049	(0.014)	-0.062	(0.014)		0.035	(0.003)	-0.014	(0.004)
# of children (0-10 years)	-0.063	(0.027)	-0.080	(0.022)		0.156	(0.027)	0.164	(0.029)		-0.001	(0.004)	0.027	(0.005)
Black	-0.170	(0.082)	-0.445	(0.066)	Coloured	-0.291	(0.102)	-0.857	(0.142)					
Brown	-0.105	(0.049)	-0.334	(0.040)										
Education	*Ref: no schooling*					*Ref: no schooling*					*Ref: no schooling*			
1-3 Years	0.142	(0.172)	0.006	(0.138)	Primary	-0.366	(0.093)	-0.168	(0.139)	1-3 Years	-0.298	(0.050)	-0.175	(0.056)
4-7 Years	0.164	(0.154)	0.091	(0.122)	Secondary	-1.162	(0.101)	-0.397	(0.146)	4-7 Years	-0.574	(0.044)	-0.378	(0.049)
8-10 Years	0.033	(0.158)	0.051	(0.125)	Vocational	-1.769	(0.175)	-0.683	(0.201)	8-10 Years	-0.995	(0.044)	-0.662	(0.050)
11+ Years	-0.280	(0.156)	0.153	(0.122)	University	-2.325	(1.003)	-0.009	(0.693)	11+ Years	-1.330	(0.046)	-0.592	(0.051)
Economic Sector	*Ref: construction*					*Ref: construction*					*Ref: construction*			
Manufacturing	-1.212	(0.068)	-2.121	(0.057)		-2.070	(0.111)	-1.782	(0.122)		-1.793	(0.020)	-1.698	(0.025)
Trade & Retail	-0.626	(0.068)	-1.417	(0.055)		0.184	(0.107)	-0.260	(0.138)		-0.204	(0.025)	-0.655	(0.030)
Services	-0.771	(0.076)	-1.313	(0.061)		-0.027	(0.089)	-1.086	(0.134)		-1.438	(0.024)	-1.022	(0.029)
Transport and Comm	-0.437	(0.064)	-0.674	(0.051)		-1.136	(0.093)	0.119	(0.095)		-0.495	(0.025)	0.083	(0.028)
Region	*Ref: Recife*					*Ref: Western Cape*					*Ref: >100,000 Inhab.*			
Salvador	0.007	(0.117)	0.051	(0.090)	Eastern Cape	1.040	(0.121)	0.971	(0.171)	15,000-99,999 Inhab.	0.608	(0.021)	0.429	(0.024)
Belo Horizonte	-0.086	(0.100)	0.050	(0.078)	Northern Cape	0.147	(0.147)	-0.145	(0.257)	2,500-14,999 Inhab.	0.789	(0.023)	0.624	(0.027)
Rio de Janeiro	0.286	(0.095)	0.039	(0.074)	Free State	0.206	(0.146)	0.722	(0.189)	< 2,500 Inhab.	0.550	(0.024)	0.346	(0.029)
Sao Paulo	0.428	(0.096)	0.030	(0.076)	North West	0.692	(0.136)	0.609	(0.193)					
Porto Alegre	0.033	(0.103)	-0.067	(0.081)	Gauteng	0.433	(0.135)	0.825	(0.176)					
					Mpumalanga	0.276	(0.142)	0.957	(0.181)					
					Limpopo	1.270	(0.151)	1.595	(0.196)					
					Kwazulu Natal	0.532	(0.134)	0.906	(0.175)					
Constant	1.655	(0.294)	-2.609	(0.256)		1.023	(0.363)	-1.308	(0.462)		3.704	(0.079)	-2.549	(0.100)
Pseudo-R2	0.103					0.186					0.142			

Multinomial Logit with base outcome = being a formal salaried worker

Table 3: Conditional Earnings Gaps: Summary

		Mean		Q=0.2		Q=0.5		Q=0.8	
		coef.	std.err.	coef.	std.err	coef.	std.err	coef.	std.err
Brazil									
pooled OLS and QR									
	Informal Salaried	-0.093	(0.011)	-0.130	(0.008)	-0.080	(0.014)	-0.045	(0.016)
	Informal Selfemployed	0.119	(0.010)	-0.013	(0.011)	0.123	(0.009)	0.184	(0.013)
FE and FE-QR									
	Informal Salaried	-0.052	(0.012)	-0.078	(0.004)	-0.046	(0.002)	-0.024	(0.004)
	Informal Selfemployed	-0.002	(0.015)	-0.060	(0.002)	0.000	(0.001)	0.057	(0.004)
South Africa									
pooled OLS and QR									
	Informal sector	-0.629	(0.024)	-0.635	(0.030)	-0.674	(0.027)	-0.579	(0.024)
	Informal Selfemployed	-0.298	(0.032)	-0.503	(0.051)	-0.372	(0.042)	-0.020	(0.027)
FE and FE-QR									
	Informal sector	-0.213	(0.027)	-0.322	(0.019)	-0.203	(0.008)	-0.116	(0.019)
	Informal Selfemployed	-0.151	(0.045)	-0.407	(0.015)	-0.141	(0.008)	0.120	(0.017)
Mexico									
pooled OLS and QR									
	Informal Salaried	-0.155	(0.003)	-0.205	(0.003)	-0.163	(0.002)	-0.113	(0.003)
	Informal Selfemployed	0.127	(0.004)	-0.047	(0.003)	0.112	(0.002)	0.292	(0.003)
FE and FE-QR									
	Informal Salaried	-0.038	(0.003)	-0.082	(0.002)	-0.038	(0.000)	0.004	(0.002)
	Informal Selfemployed	0.161	(0.005)	0.017	(0.002)	0.161	(0.000)	0.303	(0.002)

Earnings gap = estimated coefficient of the informal salary / self employment dummy. All estimations based on the variables reported in the descriptive statistics, except time-invariant characteristics (race, education and region) in the fixed effects estimations. Std. errors in brackets.

Table 4: A Multinomial Logit Estimation of the Top and Bottom Self-employment

	Brazil				South Africa					Mexico			
	Bottom		Top			Bottom		Top		Bottom		Top	
Type of selfemployment	*Ref: selfemployed employers*					*Ref: selfemployed employers*				*Ref: selfemployed employers*			
Own account worker	0.372	(0.068)	-0.129	(0.095)		0.832	(0.224)	0.343	(0.521)	0.243	(0.025)	-0.459	(0.019)
Demographics	*Ref: white, single*					*Ref: black, sinlge*				*Ref: sinlge*			
Age	-0.088	(0.018)	0.356	(0.067)		-0.053	(0.062)	-0.001	(0.001)	-0.056	(0.007)	0.026	(0.006)
Age squared	0.001	(0.000)	0.000	(0.559)		0.001	(0.001)	-0.031	(0.319)	0.001	(0.000)	0.000	(0.000)
Married	-0.503	(0.059)	2.661	(0.503)		-0.598	(0.236)	0.076	(0.529)	-0.156	(0.028)	0.097	(0.022)
Household size	0.129	(0.032)	-0.497	(0.155)		0.065	(0.064)	0.288	(0.172)	-0.010	(0.007)	-0.030	(0.005)
# of children (0-10 years)	-0.116	(0.036)	-0.718	(0.175)		0.015	(0.120)	1.176	(0.992)	-0.006	(0.006)	-0.001	(0.005)
Black	0.386	(0.099)	0.855	(0.504)	Coloured	-0.038	(0.731)	-0.476	(0.468)				
Brown	0.344	(0.065)	1.393	(0.507)									
Education	*Ref: no schooling*					*Ref: no schooling*				*Ref: no schooling*			
1-3 Years	-0.123	(0.159)	0.856	(0.118)	Primary	-0.396	(0.331)	-0.217	(0.450)	1-3 Years	-0.282	(0.060)	0.227	(0.058)
4-7 Years	-0.431	(0.141)	0.858	(0.101)	Secondary	-0.864	(0.344)	0.897	(0.708)	4-7 Years	-0.492	(0.053)	0.330	(0.052)
8-10 Years	-0.818	(0.149)	1.100	(0.134)	Vocational	-0.685	(0.693)	-31.377	(1.644)	8-10 Years	-0.637	(0.055)	0.495	(0.053)
11+ Years	-1.484	(0.150)	1.285	(0.108)	University	-32.073	(1.192)	-0.022	(0.515)	11+ Years	-0.777	(0.061)	0.721	(0.056)
Economic Sector	*Ref: construction*					*Ref: construction*				*Ref: construction*			
Manufacturing	-0.132	(0.101)	0.571	(0.241)		0.065	(0.443)	-0.242	(0.455)	1.747	(0.049)	0.021	(0.032)
Trade & Retail	0.066	(0.081)	0.634	(0.166)		0.464	(0.382)	0.244	(0.681)	1.598	(0.076)	0.096	(0.059)
Services	-0.549	(0.116)	0.333	(0.140)		-0.099	(0.558)	-0.050	(0.790)	1.517	(0.067)	0.216	(0.049)
Transport and Comm	-0.315	(0.107)	-0.117	(0.136)		1.002	(0.566)	-1.622	(1.108)	1.716	(0.051)	0.156	(0.033)
Occupation	*Ref: Professionals*					*Ref: Professionals*				*Ref: Professionals*			
Director	-0.659	(0.254)	1.346	(0.159)		0.171	(0.771)	1.176	(0.992)	0.011	(0.137)	0.186	(0.082)
Service work	0.471	(0.174)	-0.027	(0.152)		0.812	(0.598)	0.226	(0.681)	0.365	(0.080)	-0.847	(0.057)
Elementary work	0.248	(0.186)	-0.718	(0.175)		1.007	(0.586)	0.100	(0.689)	1.262	(0.184)	-1.063	(0.246)
Other	-0.220	(0.179)	0.096	(0.151)		0.407	(0.644)	0.418	(0.757)	0.690	(0.083)	-0.899	(0.061)
Region	*Ref: Recife*					*Ref: Western Cape*				*Ref: >100,000 Inhab.*			
Salvador	0.239	(0.133)	0.103	(0.140)	Eastern Cape	1.250	(0.981)	-0.429	(0.892)	5,000-99,999 Inhab.	0.334	(0.036)	-0.161	(0.029)
Belo Horizonte	-0.574	(0.115)	-0.075	(0.040)	Northern Cape	-0.752	(1.390)	-2.825	(1.253)	',500-14,999 Inhab.	0.538	(0.037)	-0.381	(0.033)
Rio de Janeiro	-0.284	(0.108)	0.014	(0.043)	Free State	0.981	(1.013)	-0.496	(0.774)	< 2,500 Inhab.	0.584	(0.041)	-0.256	(0.037)
Sao Paulo	-0.655	(0.117)	1.346	(0.159)	North West	0.560	(1.024)	-0.150	(0.673)					
Porto Alegre	-0.655	(0.122)	-0.027	(0.152)	Gauteng	0.801	(0.988)	-1.121	(0.774)					
					Mpumalanga	0.658	(1.014)	-1.796	(0.849)					
					Limpopo	0.707	(1.047)	-0.801	(0.738)					
					Kwazulu Natal	0.710	(0.995)	-0.088	(0.103)					
Constant	1.640	(0.439)	-5.497	(0.672)		-1.716	(1.763)	-4.218	(2.578)	-1.407	(0.177)	-0.425	(0.147)
Pseudo-R2	0.239					0.122				0.070			

Multinomial Logit with base outcome = middle quintiles