The Inability of Small Firms to Generate Employment in a High Wage Regime: the South African Wage Curve and Implications for Wage Subsidies

DRAFT: DO NOT QUOTE

Dieter von Fintel
Department of Economics, Stellenbosch University, South Africa

1 Introduction

Despite being Africa's largest economy, South Africa still has an uncharacteristically high unemployment rate, with the formal economy unable to absorb (predominantly unskilled) surplus labour. While early claims of jobless growth in the post-apartheid era have been refuted, the rise in unemployment over the early post-transition period remains one of the largest policy issues, particularly because the youngest generations have received more education, but have not been absorbed into the labour market and have added to the stock of the unemployed (Burger & von Fintel, 2009) (Burger, Van der Berg, & von Fintel, 2012)

The public debate on labour policy takes one of two angles: firstly, international competitiveness reports suggest that job creation is sluggish due to tight labour market regulations and the strong bargaining power of unions; secondly, proponents of strong labour market legislation (in particular from the Congress of South African Trade Unions, which is in alliance with the ruling African National Congress) are in favour of relaxing inflation targeting to lower interest rates (and employment) and the increased protection of workers' rights, with the thought that many unskilled workers are not receiving living wages. Most recently, the finance minister, Pravin Gordhan, has pointed to the employment debilitating effects of labour legislation. Nevertheless, labour laws are currently being revised to (inter alia) further prevent worker exploitation, regulate the role of so-called labour brokers who procure workers for short-term employment, and to streamline the process for workers to become union members. Each of these proposals can be seen as a further tightening of labour legislation. Other employment-promoting interventions, such as a wage subsidy, have been proposed by the Harvard Group of economists, but have been rejected by unions. One of the main reasons is that older (unionised) workers may be substituted with younger workers (who are targeted by the subsidy).

Magruder (2010) has shown that collective bargaining councils decrease employment by 8-13%, while increasing wages by 10-21%. In particular, small firms that are bound by these industry-wide agreements, are unable to pay these high wages, and are therefore limited in their employment creation potential.

Magruder notes that there is a "missing gap" in the number of small firms operating in the country. This apparently positive relationship between unemployment and wages stands in stark contrast to the wage curve - a negative relationship between regional unemployment and wages earned - (Blanchflower & Oswald, 1994), which has been put forward as an empirical law of economics. Lewis (2001) notes the long-run increases in unskilled wages (particularly in the public and manufacturing sectors) due to the deregulation of black unions since the 1970s. This has brought about a decrease in dispersion of labour market incomes, as unskilled wage growth has outpaced that of skilled wages (though overall inequality has remained persistent due to rising unemployment). This long-term decrease in the variance of wages can also be pinpointed to the fact that job opportunities in small firms (which generally pay lower wages) have also declined as a result. Indeed, Kingdon & Knight (2004) question why South Africa has such a small informal sector relative to the rest of Africa, and this in a context of one of the highest unemployment rates in the world. Usually the informal sector serves as a safety net for workers that are not absorbed by the formal sector. However, high barriers to entry are identified. Together, the evidence suggests that small firms struggle to create employment, remain in operation or face difficulties in paying sufficiently high wages.

However, Kingdon & Knight (2006), in another paper, establish empirically that a wage curve exists in South Africa, contradicting the abovementioned evidence that in some sectors wages rise with unemployment. While the occurrence of high unemployment regions with high wages might be reconciled with the Harris-Todaro (1970) model, this paper suggests that alternative mechanisms are driving this reality in South Africa. Their model suggests that individuals in a low-paying sector are attracted by the favourable wages in a high-paying sector; however, these workers cannot be absorbed due to a skills mismatch - a structural unemployment story that is also well-documented in South Africa (Bhorat & Hodge, 1999).

This paper first proceeds to reconcile the Kingdon & Knight (2006) evidence of a wage curve with other results showing a positive relationship between wages and unemployment (Magruder, 2010; Fedderke, 2012). It is evident that district-

level heterogeneity (which Kingdon & Knight's data is unable to fully account for) drives the result. Once appropriate district fixed effects are included in the specification, the relationship between unemployment and wages becomes positive. The first part of this paper explores under which circumstances this occurs and consults the international wage curve literature to establish the most credible way of dealing with spacial heterogeneity.

Given the district-level heterogeneity, it is therefore necessary to investigate which district-level differences in South Africa result in this peculiar positive relationship. As a first point of entry we consider the absence of small firms by controlling for concentration of various firm sizes; in addition, district level unionisation rates are introduced to understand the role of the potential upward wage rigidities introduced by these groups. We also control for the skills composition of districts to account for the type of vacancies that are available to job seekers (which allows the evaluation of a Harris-Todaro argument). One of the innovations of this paper is to account for district-level wage dispersion. It is evident that lower dispersion is found in districts where average wages are higher, suggesting that there is a compression of the lower tail of the earnings distribution. While unionisation is explicitly controlled for, this measure captures other factors that raise wages of the unskilled or destroy their jobs. One omitted factor that is proxied for by this variable is the role of district-level bargaining councils, which have been shown to limit the presence of small firms (Magruder 2010). Findings suggest that wage dispersion is associated with lower unemployment, which confirms that districts that have a thick lower tail (that is missing elsewhere) are able generate unskilled employment that is otherwise absent. This can be interpreted in light of "missing" small firms in South Africa. Furthermore, this measure is positively associated with both labour force participation and employment: low skilled individuals are likely to search for jobs in districts where they know that these types of “low tail” jobs have not been supplanted by large employers and their associated bargaining councils. Hence, the positive relationship between unemployment and wages is likely an artefact of the institutions that bargain up unskilled wages and limit the ability of small firms to create jobs in a high wage environment. Once controlling for each of these factors, the wage curve re-emerges, in line with the commonly

understood "empirical law" of Blanchflower and Oswald (1994). Together, these factors then explain why South Africa's labour market deviates from the empirical norm. The empirical strategy moves from a positive association between wages and unemployment; once conditioning on these variables, the association becomes negative, so that the channel of wage determination is established.

The implications of this study support the introduction of a wage subsidy as a policy measure. It is evident that the market is not willing to create jobs at high prevailing wage rates. While the policy proposal has been targeted towards young recipients who struggle to make the transition into the labour market (on the labour supply side), it is evident that an alternative effective route could be to target small firms with wage subsidies (on the labour demand side). These firms in particular experience difficulties in footing high wage bills under collective bargaining agreements and therefore reduce (unskilled) unemployment or close their doors altogether.

The rest of this paper is structured as follows. Section 2 reviews the wage curve literature, with a specific focus on high unemployment such as South Africa. It further goes on to illustrate why spatial heterogeneity should be dealt with in appropriate and specific manners, as it changes conclusions substantially. Section 3 then moves on to understanding the role of industrial structure in hindering the operation of the wage curve – in other words, the adjustment mechanism which allows wages to adjust downwards when unemployment is high. Section 4 concludes with some conjectures on policy.

2 Wage Curves in High Unemployment regions – the role of regional heterogeneity

In their recent entry in the Palgrave Dictionary of Economics, on the wage curve as an “empirical law of economics”, Blanchflower & Oswald (2008) express their surprise that the wage curve holds in a country with an unemployment rate as high as South Africa’s. In fact, it is uncanny that the elasticity is also as high as those measured in their original work on OECD countries (Blanchflower &

Oswald, 1994). The intuition for this apparent contradiction derives from the poor potential for declining wages to clear particularly large labour surpluses (specifically if wages have to drop to below reservation or subsistence levels). Should unemployment be persistent and high in the long run, it is possible that the flexibility implied by the wage curve does not conform to the "empirical law". Indeed, Fedderke's (2012) most recent contribution starts off with the premise that "South African labor market conditions are unusual by international standards. High and persistent unemployment do not prevent real labor costs from rising." This suggests that not only do wages fail to *fall* in some high unemployment scenarios (as suggested by a non-linear wage curve), but in some circumstances they are even able to *rise.*

However, to fully reconcile the conflicting predictions on the nature of the relationship between wages and unemployment, it is necessary to separate short-run adjustments from long-run ones. The rest of this paper therefore attempts to understand how empirical wage curves should ideally be estimated to characterise various modern labour markets, especially in developing countries with high levels of unemployment. In particular, we turn to the role of regional heterogeneity and fixed effects to account for long-run relationships. Optimal definitions of regions and labour markets are also briefly considered.

Blanchflower & Oswald's (2008) prime example of a high unemployment region with a documented wage curve is based on the evidence of Kingdon & Knight (2006), who use the 1993 South African PSLSD data to estimate an elasticity of individual wages related with respect to magisterial district unemployment rates. The estimate is remarkably similar to the international "norm" of -0.10 that has emerged from a wide variety of studies. However, meta-analyses of the wage curve literature suggest that muted elasticities arise when unemployment is particularly high - though mostly non-linearities in unemployment are rejected statistically under normal circumstances (Nijkamp & Poot, 2005). Kingdon & Knight (2006) acknowledge this potential fact by adding a quadratic in unemployment for the South African specification. While in most international contexts, significant non-linearities result in a positive relationship only for

sections of the wage curve that are not found in the actual data, the high South African unemployment rate means that this potential phenomenon cannot be ignored.

Kingdon & Knight (2006) further explore this by segmenting their analysis into high and low unemployment regions. Their analysis focusses on differences between previous "homeland" and "non-homeland" regions. The former were small pseudo-independent "states" that were created by the apartheid government, where apparent black self-determination was instated. Black South Africans were effectively only citizens of these "states" (which were not recognised by the international community), even though the only viable labour market opportunities were available in the designated white areas of South Africa. A migrant labour system developed, whereby blacks were only allowed to work in white South Africa with permits, and were only recognised as temporary residents. As a consequence, a spatially divided labour market resulted, with few viable opportunities in the homelands and unemployment rates that were approximately double those of other regions. Kingdon & Knight (2006) find that an insignificantly positive wage curve relationship existed in homeland regions, while a strong negative wage curve holds in the rest of South Africa. Results such as these highlight the importance of regional heterogeneity where unemployment is not only high, but where it is widely dispersed along the spatial dimension.

Magruder (2012) investigates another feature of the South African labour market that has a strong spatial definition. Centralised bargaining in many industries sets up wage agreements that are often applicable at district level, regardless of whether firms have unionised workers or not. The consequence is that high wage districts tend to witness lower employment creation, especially amongst smaller firms, which tend to be in disproportionately low concentration in South Africa. Magruder uses these spatial discontinuities to identify what can indirectly be viewed as an antithesis to a wage curve in post-apartheid South Africa.

One simple way of accounting for these potential differences is to add regional fixed effects to wage curve specifications. However, their introduction raises multiple questions, which have been addressed in the literature in various manners, often dictated by data availability. Firstly, which is the appropriate geographic definition of a "local" labour market, and is this the correct level at which to measure unemployment for wage curve purposes? A related question is which accompanying fixed effect should be used in a wage curve specification? If labour markets (in the wage curve tradition) are fairly homogenous over large regions, then wages would respond to unemployment rates in much the same way in large parts of the country. However, in countries, where unemployment is highly dispersed, and where differences are large by region, the local labour market should be narrowly defined, and accounting for fixed effects is important. To the contrary, however, the more centralised bargaining is, the more homogenous labour markets are likely to be across large region, so that fewer regions of analysis could be used. Secondly, what is the empirical importance of accounting for fixed effects, and which type of data is required to support such an analysis.

Albaek, et al. (2000) draw attention to the importance of spatial heterogeneity in the wage specification, particularly because it matters substantially for the conclusions on the existence of the wage curve in their samples. They find that introducing regional fixed effects eliminate the otherwise large negative relationship between wages and unemployment in Nordic countries. While the groundbreaking work on the wage curve (Blanchflower & Oswald, 1994) advocated the introduction of regional fixed effects in the standard specification, it did not have the same impact as established by the Scandinavian researchers. In some cases the introduction of regional fixed effects amplifies the elasticity (for instance Papps (2001)), while in others they are muted (see for instance Baltagi, Blien, & Wolf (2000)). Blanchflower & Oswald (1994: 181) emphasised their inclusion to distinguish between long-run and short-run differences in the wage-unemployment relationship.

Albaek, et al. (2000) proceed to show theoretically why accounting for regional variation is important. They show that without fixed effects, the coefficient on

log(unemployment) represents a combination of long-run and short-run elasticities. Once they are accounted for, the coefficient represents the short-run responsiveness of wages to current changes in unemployment, which should theoretically be negative. Econometrically, the introduction of fixed effects is akin to a transformation of all variables to represent deviations from each region's long-run average in each period (if pure panel data exists), so that coefficients are interpreted as instantaneous rather than long-run. Theoretically, the fixed effects (the permanent, time-invariant regional component of *log(wages)*) should be positively correlated with unemployment (if migration models hold). By implication, coefficent estimates should be biased away from zero if fixed effects are not accounted for, thereby overstating the magnitude of the wage curve. The backing for this assertion is the well-known migration model of Harris & Todaro (1970), where in long-run equilibria expected wages and unemployment are positively associated. Migration between regions will only stop if *expected* wages are equal across sectors: that is, the wage rate scaled by the unemployment rate. By implication, high wage regions do not attract further individuals in equilibrium, because high unemployment exists concurrently. This positive *long-run* relationship is empirically verified in the United States (along with a negative transitory wage curve) (Blanchflower & Oswald, 1994), while a negative long-run relationship persists in the Nordic states (with an insignificant transitory wage curve) (Albaek, et al., 2000). A long-run negative relationship in the latter case perhaps explains why unemployment elasticities are overstated when not accounting for fixed effects (though small numbers of regions in their analysis could be confounding this analysis). Why, then do the Nordic results differ from the "empirical law"? Albaek, et al. (2000) emphasise the role of centralised bargaining in driving these results. Under highly centralised bargaining regimes, wages are likely to be unresponsive to higher *local* unemployment, as they are set at a more central regional level. The following section will follow up with a review of the literature, with the role of bargaining in mind in generating results different from the "empirical law".

2.1 A meta-analysis of wage curve studies and regional heterogeneity

Nijkamp & Poot's (2005) existing meta-analysis shows that including regional fixed effects in wage curve specifications reduces elasticities by an average of 3.5 percentage points, though this fall is not statistically significant across the specifications they survey. This suggests that accounting for long-run wage-setting dynamics does not detract from the wage curve as an empirical law. However, most of the studies they survey are taken from developed country evidence, where unemployment is not as high, and also not as dispersed across regions. Most utilise repeated cross sections, and present evidence of wage cuves with and without regional fixed effects; in these studies it is possible to include fixed effects at the same level at which unemployment rates are calculated due to the time dimension that is present in the series of data. In few cases a single cross section survey is used by necessity, and fixed effects are included for regions that are larger than those for which unemployment rates are calculated. This raises the question whether fixed effects for regions larger than those for which unemployment rates are estimated can effectively remove the long-run component of the wage-setting relationship.

Kingdon & Knight (2006) is a case in point. Due to the fact that they only resort to using one cross-sectional survey, they are not able to introduce regional fixed effects at the level of aggregation of the unemployment rates they measure. As a second best option, they include fixed effects for regions with larger geographic definitions (they include provincial fixed effects, while estimating unemployment rates at the level of the magisterial district, which are encapsulated by provinces). However, even within these regions, large heterogeneity exists in unemployment rates. Despite some provinces clearly having higher unemployment rates on average, it is evident that the heterogeneity within these geographic units is non-negligible. Notably, unemployment is highly correlated with the apartheid homeland demarcations. Figure 1 shows that this is true even in 2000, some years after the homeland system was abandoned by the post-transition regime. Not accounting for this level of heterogeneity is potentially important for wage curve analyses. It is firstly evident that a province cannot adequately define a "local labour market", purely by the differences in unemployment within these administrative regions. Secondly, the heterogeneity

within provinces will not be accounted for by introducing fixed effects at this higher level. Given that unemployment is of a long-term structural nature in South Africa, it is not sure that provincial fixed effects can purge the wage curve relationship from this feature adequately. Thirdly, Card (1995) highlights issues with degrees of freedom when regions are few: to illustrate, should 9 provincial unemployment rates be used to model individual wages, the wage curve elasticity is not based on the number of individuals, but by 9 multiplied by the number of surveys used. Kingdon & Knight (2006) take account of the first and third concerns in their study by estimating unemployment rates at the magisterial district level and using those in their wage curve analyses. However, their use of cross sectional data does not allow them to address the second concern, so that it is likely that sufficient long-term heterogeneity remains unabsorbed by fixed effects estimation. Other studies suffer the same limitations (Blanchflower & Oswald, 1990; Winter-Ebner, 1996). In each case, the introduction of fixed effects at a level of aggregation larger than the defined labour markets reduces wage curve elasticities, though does not render them insignificant (as is the case in other studies). However, the potential remains that the elasticity remains overstated. This possibility is discussed in more detail below.

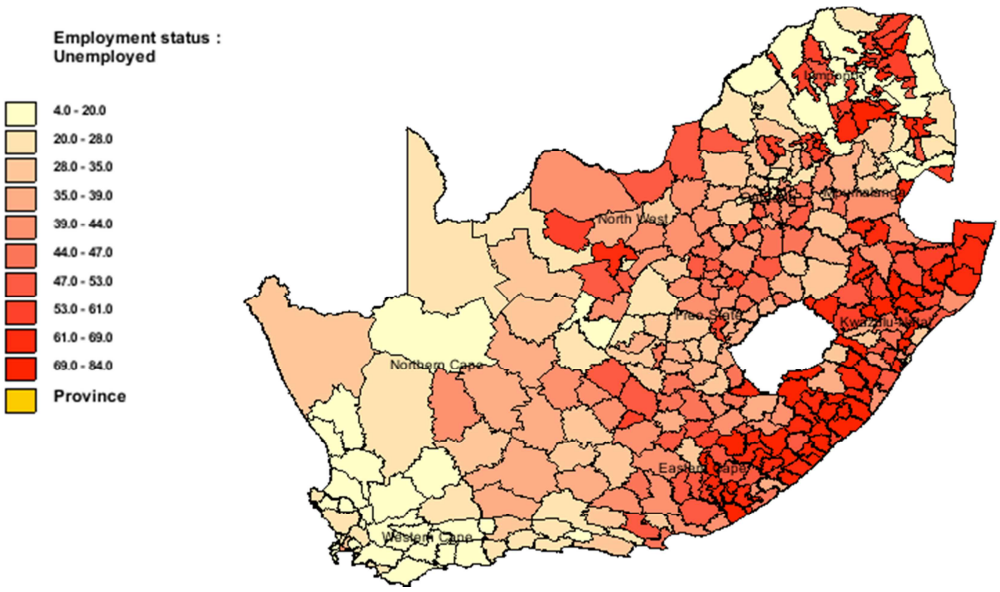

Figure 1 Broad Unemployment Rate - by Magisterial District - 2001 Census

While Nijkamp & Poot (2005) conclude that the introduction of fixed effects does not have an overall impact on the elasticities that they survey, some differences emerge. Table 1 draws together most of the estimates they review (bar for those from the book of Blanchflower & Oswald), and explicitly focusses on various study dimensons, namely the number of regional unemployment rates, the number of regional fixed effects, and also the potential for centralised bargaining to dampen short-run wage fluctuations. The latter is done by comparing countries across measures of the degree of centralisation of their collective bargaining systems (Iverson, 1998) and prevalence of central bargaining (Driffill, 2006). The purpose of this attempt is to highlight under which circumstances various empirical strategies generate wage curves that should be scrutinised more carefully, and to give potential explanations for some of the deviations from the norm, which suggests that regional fixed effects are not important for establishing a reflective wage curve (Nijkamp & Poot, 2005).

It is informative to first start with the studies of countries that have lower levels of centralised bargaining. US and UK estimates are robustly in line with the wage curve, even when accounting for regional fixed effects. The exception is the repeated cross section estimate of Blanchflower & Oswald (1990), which becomes statistically insignificant. However, this exposes a problem of limited degrees of freedom (Card, 1995), where only 11 regional unemployment rates are constructed. The specifications of Wagner (1994) suffer the same problem, so that the insignificance is not likely to result from the relatively high centralisation of bargaining that prevails in Germany. However, other estimates for (East and West) Germany are also rendered insignificant by *regional* fixed effects (Baltagi, Blien, & Wolf, 2000; Baltagi & Blien, 1998), despite defining larger numbers of regions, so that the bargaining argument may re-emerge. Other estimates for Germany (Pannenberg & Schwarze, 1998; Buettner, 1999), however, remain significant but small after conditioning on regional dummies. For Germany (with relatively highly centralised bargaining institutions), the evidence could nevertheless be consistent with the claims of Albaek, et al. (2000) that wage cruve elasticities are either small or insignificant in the presence of strongly centralised bargaining. However, antipodean studies (in regions with

only slightly lower degrees of centralised bargaining) do not follow this pattern. Australian repeated cross sectional data (Kennedy & Borland, 2000) yields smaller (yet significant) estimates of the elasticity, despite a small number of defined regions and moderate levels of centralised bargaining. In the case of New Zealand, the elasticity becomes more strongly negative (Papps, 2001). Belgian estimates (Janssens & Konings, 1998) follow a similar pattern to those from Australia, with similar limitations in terms of numbers of regions and within a similar bargaining environment. Turning to highly centralised bargaining regions, Albaek, et al. (2000) attribute their insignificant elasticities after introducing regional fixed effects to the centralised bargaining systems that operate in these countries. However, small degrees of freedom could just as likely be playing a greater role in this case than the reasons that they offer. Their choice of few large regions may, however, be justified on the grounds that "locl labour markets" could be potentially larger where central bargaining prevails. Furthermore, Austrian estimates (Winter-Ebner, 1996) – with a sufficient number of regions (though a small number of fixed effects) – remain statistically significant once controlling for region. This is true despite it being at the extreme of the centralised bargaining spectrum; however, the lack of additional cross sections entails that higher level regional fixed effects were not likely to mop up the full degree of heterogeneity that was present in the data. Overall, then, evidence for centralised bargaining as the mechanism for rendering estimates insignificant is limited. The evidence on the number of local labour markets and fixed effects is also not unambiguous from this meta-analysis. However, understanding levels of heterogeneity within these regions (perhaps also along the lines of bargaining councils, *inter alia*) is a potential route to understanding why region definitions are as important for wage curve analysis as their presence in the literature suggests.

Table 1 Summary of Wage Curve Studies and their Regional Fixed Effects

									Collective Bargaining Coverage (Driffill, 2006)		
Author	**Country**	**Period**	**Data type**	**Number of u% rates**	**FE Regions**	**Impact of introducing Regional FE**	**Centralisation Index (Iverson, 1998)**	**Centralisation Score (Driffill, 2006)**	**1980**	**1990**	**2000**
Blanchflower & Oswald (1990)	Britain	1981	CS	65	9	Elasticity remains significantly negative	0.177	3+	70	40	30
Blanchflower & Oswald (1990)	Britain	1983-1987	Repeated CS	11	11	Elasticity becomes insignificant	0.177	3+	70	40	30
Wagner (1994)	West Germany	1979;1985	Repeated CS	10	10	Elasticity becomes insignificant (except for log(U))	0.377	5--	80	80	90
Wagner (1994)	West Germany	1984-1990	Repeated CS	9	9	Elasticity becomes insignificant	0.377	5--	80	80	90
Bratsberg & Turunen (1996)	US	1979-1993	Panel	1376	1376	Elasticity becomes smaller, but remains significant	0.071	2	26	18	14
Winter-Ebmer (1996)	Austria	1983	CS	99 regions x 31 occupations	9	Elasticity becomes smaller, but remains significant	0.431	6	95	95	95
Partridge & Rickman (1997)	US	1972-1991	State-level panel	48	48	Elasticity is positive, but becomes smaller	0.071	2	26	18	14
Baltagi & Blien (1998)	Germany	1981-1990	Panel	142	142	Not significant for all workers, except when instrument added	0.377	5--	80	80	90
Janssens & Konings (1998)	Belgium	1985-1992	Individual Panel	11	11	Remains negatively signficiant	0.321	4	90	90	90
Pannenberg & Schwarze (1998)	East Germany	1992-1994	Individual Panel	35	35	Elasticity negatively significant with regional FE, but not once individual FE are incorporated					
Buettner (1999)	West Germany	1992	Regional Panel	325	325	Small but negatively significant	0.377	5--	80	80	90
Albaek et al (2000)	Nordic Countries	1989-1993	Repeated CS	8 to 19	8 to 19	Large regional unemployment rates become insignificant once controlling for these FE, though limited effect for district level unemployment rates	0.459 to 0.538	5- to 5	70 to 90	70 to 90	70 to 90
Baltagi et al (2000)	East Germany	1993-1998	Repeated CS	114	114	Fixed effects knock the significance, instruments bring it back					
Kennedy & Borland (2000)	Australia	1982-1995	Repeated CS	7	7	Elasticity becomes smaller, but remains significant		4	80	80	80
Papps (2001)	New Zealand	1986-1996	Repeated CS	60 (30 male, 30 female)	30	Elasticity becomes stronger		4	60	60	25
Blanchflower (2001)	Eastern Europe	1991-1997	Repeated CS	6 to 42	6 to 42	Elasiticty becomes smaller, but remains significant in most estimates; sometimes insignificant					

2.2 Regional heterogeneity and wage curve bias

Suppose that the true wage curve is represented by:

$$\log\left(earnings_{ir_j^*t}\right) = \beta_0 + \beta_1 \log(unemployment_{r_j^*t}) + \gamma' x_{ir_j^*t} + e_{ir_j^*t} \quad (1)$$

where $e_{ir_j^*t} = \mu_{r_j^*} + \lambda_t + u_{ir_j^*t}$, i and t index individuals and time respectively, and r_j^* defines an optimal geographic reference region or local labour market for industry *j*. The size of this region depends on how spatially segregated the labour market is. Should labour market circumstances be fairly homogenous across space, and workers able to migrate freely into another region or industry, the optimal region may be fairly large. In the South African case, these optimal regions are likely to be smaller than in other countries, owing to the spatial segregation introduced by the homeland system. While labour market movement between these regions has been liberalised after apartheid, the spatial patterns along very small regional lines still persist (as is evident in Figure 2). In addition, regional rigidities are contained by magisterial district boundaries, by which bargaining councils in various industries are defined, resulting in largely heterogeneous employment and wage effects in spatially contiguous regions (Magruder, 2012).

Should equation (1) be estimated without adding any regional fixed effects, the standard result holds that coefficients of interest may be biased and inconsistent if any of the covariates are correlated with the unobserved heterogeneity that is absorbed into the error term. In the current context; if some regions are home to institutions that foster characteristically high wages over the long-run (with high $\mu_{r_j^*}$) which limits the creation of employment in the short run, then the positive relationship between unemployment and the fixed effects means that the wage curve elasticity will be overstated. Regional bargaining councils that limit employment creation present a good example of such a scenario. The bias will be larger, the higher the correlation between unemployment and these specific institutions – which, by implication, means that great spatial diversity exists. In

countries where greater homogeneity in long-run district-level factors dominates, it may not be as important to account for fixed effects, which is the broad-stroke conclusion for OECD countries (Nijkamp & Poot, 2005).

2.2.1 Fixed effects for non-optimal regions

Suppose, however, that cross section data does not allow for fixed effects to be added at the same level of spatial aggregation as the unemployment rate. Where spatial differences are important, it may be preferable to account for spatial heterogeneity at a geographic level that is typically larger than the optimal labour market, a strategy followed by Kingdon & Knight (2006). Suppose that it is still possible to calculate unemployment rates for the optimal labour market size, but that fixed effects for larger regions and industries are included as follows:

$$\log\left(earnings_{ir_j^*t}\right) = \tilde{\beta}_0 + \tilde{\beta}_1 \log(unemployment_{r_j^*t}) + \tilde{\gamma}' x_{ir_j^*t} + \tilde{e}_{ir_j^*t} \quad (2)$$

where $\tilde{e}_{ir_j^*t} = \mu_{r_j} + \tilde{\lambda}_t + \tilde{u}_{ir_jt}$ and $\mu_{r_j} = \frac{1}{n_{r_j^* \subset r_j}} \sum_{r_j^* \subset r_j} \alpha_{r_j^*} \mu_{r_j^*}$ is now a weighted "average" fixed effect for all of the optimal sub-regions (r_j^*) of which the larger region (r_j) is comprised, with the optimal region specific deviations ($\xi_{r_j^*}$) from this average being absorbed into $\tilde{u}_{ir_jt} = u_{ir_j^*t} + \xi_{r_j^*}$.

$$\mu_{r_j^*} = \mu_{r_j} + \xi_{r_j^*} \Longrightarrow \xi_{r_j^*} = \mu_{r_j^*} - \mu_{r_j}$$

$$Cov\left[\log\left(unemployment_{r_j^*t}\right); \xi_{r_j^*}\right]$$

$$= Cov\left[\log\left(unemployment_{r_j^*t}\right); \mu_{r_j^*}\right] - \frac{1}{n_{r_j^* \subset r_j}} \sum_{r_j^* \subset r_j} \alpha_{r_j^*} Cov\left[\log\left(unemployment_{r_j^*t}\right); \mu_{r_j^*}\right] \quad (3)$$

Given that this fixed effect does, not "mop up" all of the unobserved heterogeneity in a region, a correlation between unemployment and the error term still arises, as in (3), which could result in biased and inconsistent estimates of the wage curve elasticity. The first term of (3) represents the degree of correlation between the error and explanatory variables that would arise if no

fixed effects were included in the specification. This also determines the magnitude of the inconsistency of the wage curve estimate. The second term represents the potential reduction in this bias that results from at least including the rougher fixed effects. However, suppose that large degrees of spatial heterogeneity exists within the larger district (r_j) than in each of its component districts (r_j^*), so that one district (say) with a large population (large $\alpha_{r_j^*}$) is a high unemployment-high wage district (in the Harris-Todaro (1970) sense), while other districts have (say) zero correlations with unemployment. The consequence of this would be a large between district correlation (large first term), and a small sum of within district correlations (small second term), so that the overall correlation remains. Essentially this suggests that if district fixed effects attempt to account for large levels of heterogeneity across sub-regions, it remains likely that estimates will remain biased and inconsistent if this heterogeneity within the larger regions is still large. Should μ_{r_j} account for much of the variation in $\mu_{r_j^*}$ as opposed to $\xi_{r_j^*}$, this strategy will result in minimal bias. However, if the sub-regional deviation displays substantial variation within the larger border, then accounting for these fixed effects will not be effective in capturing the true relationship in (1).

2.2.2 Optimal Labour Market Size for Wage Curve Analysis

Since most studies in the wage curve literature are silent on what should be the size of the "optimal" labour market, and no clear prescriptions exist at which level unemployment rates should be calculated, other potential (unnoticed) problems may arise as a result. Wagner (1994), for instance, points to the fact that unemployment rates should be measured at lower levels of aggregation to reflect "local" labour markets. Presumably, most applied work relies on the information of regions that is available in the data, which is often dictated by the sampling design, or political and administrative divisions. Where local labour markets naturally fill these boundaries (such as where minimum wages and bargaining are defined regionally), this is unproblematic.

However, extending the analysis above, it may be that unemployment rates that are calculated at high levels of geographic aggregation, could aggravate the estimation of the elasticity. Suppose now that unemployment is calculated for a larger region (r_j) than is optimal and used to estimate (4). Unemployment is now measured with error ($\omega_{(r_j^* \subset r_j)t}$) that is proportional to the unemployment rate of the unknown optimal labour market region (r_j^*):

$$\log\left(earnings_{ir_jt}\right) = \hat{\beta}_0 + \hat{\beta}_1 \log\left(unemployment_{r_jt}\right) + \hat{\gamma}' x_{ir_jt} + \hat{e}_{ir_jt}$$
$$= \hat{\beta}_0 + \hat{\beta}_1 \log\left(unemployment_{(r_j^* \subset r_j)t}\omega_{(r_j^* \subset r_j)t}\right) + \hat{\gamma}' x_{ir_jt} + \hat{e}_{ir_jt}$$
$$= \hat{\beta}_0 + \hat{\beta}_1 \log\left(unemployment_{(r_j^* \subset r_j)t}\right) + \hat{\gamma}' x_{ir_jt} + \hat{v}_{ir_jt} \quad (4)$$

where $\hat{v}_{ir_jt} = \mu_{r_j} + \lambda_t + \hat{\beta}_1 \log(\omega_{\left(r_j^* \subset r_j\right)t}) + u_{ir_jt}$

$= \frac{1}{n_{r_j^* \subset r_j}} \sum_{r_j^* \subset r_j} \alpha_{r_j^*} \mu_{r_j^*} + \xi_{r_j^*} + \lambda_t + \hat{\beta}_1 \log(\omega_{(r_j^* \subset r_j)t}) + u_{ir_jt}$ and

$\log(unemployment_{r_jt}) = \log(unemployment_{(r_j^* \subset r_j)t}) + \log(\omega_{(r_j^* \subset r_j)t})$ for all $r_j^* \subset r_j$ (5).

This specification now highlights a composite error term with two additional components that can be potentially correlated with unemployment. The first is $\xi_{r_j^*}$, which has been analysed above, and is attributable to diversity within the regions for which finer fixed effects are supposed to account. The second is $\hat{\beta}_1 \log(\omega_{(r_j^* \subset r_j)t})$, a component that results from measurement error, by not calculating employment for an optimally sized local labour market. Equation (5) holds as an identity within each "large" region, and can be analysed in the classical measurement error context: sub-region specific errors are correlated with the included covariate by definition within the chosen region. Sub-regions with low unemployment will have large positive measurement errors at the higher regional level and visa versa, inducing a negative correlation between measured unemployment and the composite error term. By definition, therefore, the higher the level of aggregation, the closer the estimate of the (negative) wage

curve elasticity will tend to zero, or become positive. The latter is amplified by the second term in (6), which suggests that the larger the variation in measurement error resulting from the choice of aggregation, the potential exists to induce a positive correlation with unemployment measured at the higher level. This is particularly problematic if sub-regional heterogeneity in unemployment is large, and is also introduced by not defining heterogenous regions finely enough.

$$Cov\left[\log(unemployment_{r_j t}); \log(\omega_{(r_j^* \subset r_j)t})\right]$$
$$= Cov\left[\log\left(unemployment_{(r_j^* \subset r_j)t}\right); \log(\omega_{(r_j^* \subset r_j)t})\right] + Var\left[\log(\omega_{(r_j^* \subset r_j)t})\right] \quad (6)$$

2.3 An Illustration

To illustrate these potential problems with wage curve estimates, we turn to the context of Kingdon & Knight (2006), who find a wage curve for South Africa using cross sectional data. While they are able to estimate unemployment rates at the fairly low level of the magisterial district (as in Figure 1), their use of a cross section survey limits them in controlling for fixed effects at the same level. They therefore estimate approximately 330 district level unemployment rates and include only 9 provincial fixed effects. This study employs a series of cross sections from South Africa's Labour Force Surveys, enumerated twice yearly in March and September. The period spans September 2000 to March 2004, and is therefore able to capture short-run changes in earnings and unemployment. These surveys have been chosen, because they allow the identification of individuals' place of residence by various classifications (the smallest being magisterial districts[1], then larger district councils and finally provinces), and this is a period for which consistent measurement of labour market status and

[1] These are administrative boundaries that were adhered to prior to 1994. However, many sampling frames still use these small units and data in many years can be linked at this level. District councils are the new administrative boundaries that were introduced in the post-apartheid era for the purposes of local government. These regions are larger than the magisterial districts, and are mostly comprised of a number of former MDs. South Africa has a second tier of governance at the provincial level. Since 1994 the country has ben divided into 9 of these regions.

earnings occurred (Burger & Yu, 2006). While this period was marked by high economic growth, unemployment continued to rise concurrently. It is therefore of interest to know whether wages responded to growing unemployment, and whether existing high long-run unemployment influences the wage setting relationship. However, the specific task of this analysis is to understand the role of geographic definitions in establishing this wage-setting relationship.

We focus our attention on black males only, as this group typically has a high unemployment rate, is the most highly unionized, and has been affected most by the creation of the apartheid homelands and the migrant labour system. Unemployment rates are based on the broad definition, following Kingdon & Knight (2006). Wage curves are estimated with a log-log specification, so that the coefficient of interest represents an elasticity. Following the original Blanchflower & Oswald (1990) specifications, we also include a cube of the *log(unemployment rate).*

Table 2 presents wage curve estimates, modeled with unemployment rates of magisterial districts, the most disaggregate level possible[2]. Without fixed effects or controls the elasticity is small, but negatively significant (column 1). Including controls in column 2 raises the elasticity, while the addition of provincial fixed effects to this specification in column 3 hardly changes the result. Together this suggests that wage curve elasticity estimates for black males are small; furthermore, following the procedure of Kingdon & Knight (2006) - by introducing fixed effects for fairly large regions - does not absorb much of the heterogeneity that is present across districts, nor does it remove long-run wage-setting behaviour from wage curve estimates. In column 4 district council fixed effects (of which there are 55) are introduced, and the estimate again drops slightly. However, once the 331 magisterial district fixed effects are introduced in column 6, the wage curve becomes positively insignificant. This suggests that the negative result obtained by Kingdon & Knight (2006) does not represent

[2] This section will only discuss elasticities from the log-log specifications, the inclusion of non-linearities generally supports that the higher unemployment rates are, the smaller is the wage curve effect. However, the analysis of the impact of fixed effects is similar for the non-linear specifications.

transitory responses of wages to unemployment, but that the relationship is long-run in nature. Because the sign of the coefficient on unemployment changed, it is evidence for the fact that the fixed effects (the permanent component of wages) are negatively correlated with unemployment. Therefore, in the long-run, regions with high wages are persistently associated with low unemployment regions. This contrasts with the Harris-Todaro (1970) prediction that migration should facilitate the equalization of expected wages between high and low unemployment regions. What this may suggest is that spatial labour market segmentation has prevented sufficient numbers of (unemployed) black migrants to move to high wage regions, but they remained unemployed in low wage regions. Low wage regions could not absorb these individuals into the workplace by downward adjustment in wages. This narrative matches the empirical evidence of Kingdon & Knight (2006) with homeland (high unemployment, low wage) and non-homeland (lower unemployment, higher wages) segmentation. However, this suggests that the national (transitory) wage curve is not manifest in South Africa, but that long-run regional segmentation is a stronger determinant of wage-setting rather than fluctuations in unemployment. Column 5 controls for district-level covariates, to proxy for overall labour demand conditions – this includes proportions of individuals in various sectors and occupations, unionization rates and average education levels within the region. It is evident that these demand responses do not play the same role as the district level fixed effects, so that other unobserved factors are playing their part in this non-standard result. Long-run persistent factors that divide the labour market into "good" and "bad" regions are the more likely candidates.

Table 3 repeats the analysis, but this time with unemployment rates measured at the district council level. All observations in the survey are contained within only 55 district councils, as opposed to the 331 more finely divided magisterial districts. The patterns are very similar to those above, with the wage curve being eliminated once the district level fixed effects are included at the same level at which unemployment was estimated. However, it is notable that the coefficients are substantially larger (and remain statistically highly significant) than when magisterial district unemployment rates were used. Two potential reasons could

be offered for this. Firstly, individual labour market behaviour might be more responsive to conditions in a larger district. Secondly, equation (6) above highlights that the larger the variation of measurement error resulting from the aggregation, the larger the correlation with unemployment, and by implication the greater the bias of coefficients. These estimates are also closer to the norm of -0.1 recorded widely in the literature. This raises the question whether most studies suffer from aggregation issues, or whether all estimates are measuring "appropriate" labour market responses. It is notable that most studies do not record as many regions as is used in Table 2 (see Table 1), but are closer in number to estimates in Table 3. It is, therefore, also possible that the unemployment rates are measured with error when small regions are used, and statistics cannot be calculated with precision, or be representative of small regions. It is not clear from this analysis which of these scenarios is the truth. However, it is certain that using fewer districts to construct unemployment rates raises the elasticity.

This process is taken further by aggregating unemployment to the level of the province in Table 4. The elasticities again grow remarkably as the level of aggregation increases. These elasticities are, however, implausibly high. This then suggests that it is measurement error that dominates the increase in the elasticity, rather than actual responsiveness. Indeed, once provincial fixed effects are entered, the negative correlation between the fixed effect and unemployment appears again, in that the elasticity becomes significantly positive. This time the coefficient is highly statistically significant and large in magnitude. However, as pointed out by many authors in this literature, standard errors are underestimated when using aggregate variables, and may lead to inappropriately significant estimates (Moulton, 1990).

These sets of estimates highlight that the higher the level of aggregation, the further elasticities move from zero. Given that, with a moderate number of districts, elasticities resemble those in the literature (Table 3), some consistency with other estimates is found. However, fixed effects at the appropriate level remove the wage curve completely. If regional labour markets in other studies

were chosen to represent a greater number of smaller regions, the “normal” elasticity found in the literature (of -0.1) would likely also become smaller. It is therefore an imperative to understand whether this elasticity is an artifact of the labour market delineations that have been chosen, or whether it is indeed as high as documented in the literature. Besides that, it is notable that in a country where great variation in unemployment exists along spatial dimensions, it may be important to disaggregate to smaller regions.

Table 2 Black Male Wage Curve Estimates - with Magisterial District Unemployment Rates

	log(Individual Black Male Monthly Earnings)											
	1	2	3	4	5	6	7	8	9	10	11	12
log(Broad Black Male MD Unemployment Rate)	-0.039**	-0.068***	-0.059***	-0.050***	-0.066***	0.011	-0.148***	-0.171***	-0.104***	-0.070***	-0.145***	0.021
log(Broad Black Male MD Unemployment Rate)^3							0.010***	0.009***	0.004***	0.002	0.006***	-0.001
Controls	N	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y
MD averages of controls	N	N	N	N	Y	N	N	N	N	N	Y	N
Time FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Prov FE (x9)	N	N	Y	N	N	N	N	N	Y	N	N	N
DC FE (x55)	N	N	N	Y	N	N	N	N	N	Y	N	N
MD FE (x331)	N	N	N	N	N	Y	N	N	N	N	N	Y
Constant	7.060***	4.015***	4.305***	4.547***	3.644***	4.712***	6.964***	3.922***	4.264***	4.527***	3.565***	4.720***
R-squared	0.002	0.479	0.504	0.515	0.506	0.532	0.004	0.48	0.505	0.515	0.507	0.532
N	69156	66717	66717	66717	66717	66717	69156	66717	66717	66717	66717	66717
F statistic	3.853	1092.958	987.498	4123.083	815.568	.	6.799	1074.57	966.517	4088.144	804.306	.

NOTES: * p<0.1, ** p<0.05, *** p<0.01. Own Calculations from LFS September 2000 to LFS March 2004. Controls include union membership, firm size dummies, age, age squared, education dummies, a marriage dummy, occupational skill dummies, a public sector dummy and indicators of primary, secondary and tertiary sector work. Where district level covariates are included, the same covariates are averaged, except for age and marriage. The 331 Magisterial Districts are each contained within one of 55 larger District Councils, which are in turn part of one of 9 provinces respectively.

Table 3 Black Male Wage Curve Estimates - with District Council Unemployment Rates

	log(Individual Black Male Monthly Earnings)									
	1	2	3	4	5	6	7	8	9	10
log(Broad Black Male DC Unemployment Rate)	-0.138***	-0.156***	-0.078***	-0.002	-0.151***	-0.252***	-0.218***	-0.132***	0.002	-0.183***
log(Broad Black Male DC Unemployment Rate)^3						0.016***	0.010***	0.008***	0	0.004***
Controls	N	Y	Y	Y	Y	N	Y	Y	Y	Y
DC averages of controls	N	N	N	N	Y	N	N	N	N	Y
Time FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Prov FE (x9)	N	N	Y	N	N	N	N	Y	N	N
DC FE (x55)	N	N	N	Y	N	N	N	N	Y	N
Constant	6.961***	3.948***	4.312***	4.605***	4.323***	6.869***	3.902***	4.275***	4.608***	4.313***
R-squared	0.003	0.481	0.506	0.517	0.506	0.005	0.482	0.506	0.517	0.506
N	70182	67729	67729	67729	67729	70182	67729	67729	67729	67729
F statistic	5.721	1146.848	1021.956	503.444	850.16	11.193	1110.532	995.481	498.422	832.447

NOTES: * p<0.1, ** p<0.05, *** p<0.01. Own Calculations from LFS September 2000 to LFS March 2004. Controls include union membership, firm size dummies, age, age squared, education dummies, a marriage dummy, occupational skill dummies, a public sector dummy and indicators of primary, secondary and tertiary sector work. Where district level covariates are included, the same covariates are averaged, except for age and marriage. The 331 Magisterial Districts are each contained within one of 55 larger District Councils, which are in turn part of one of 9 provinces respectively.

Table 4 Black Male Wage Curve Estimates - with Magisterial District Unemployment Rates

	log(Individual Black Male Monthly Earnings)							
	1	2	3	4	5	6	7	8
log(Broad Black Male Prov Unemployment Rate)	-0.319***	-0.299***	-0.033	0.204***	-2.568***	-1.968***	-0.129	0.000
log(Broad Black Male Prov Unemployment Rate)^3					0.847***	0.630***	0.034	0.061
Controls	N	Y	Y	Y	N	Y	Y	Y
Prov averages of controls	N	N	N	Y	N	N	N	Y
Time FE	Y	Y	Y	Y	Y	Y	Y	Y
Prov FE (x9)	N	N	Y	N	N	N	Y	N
Constant	6.798***	3.844***	4.376***	5.360***	5.455***	2.876***	4.315***	5.240***
R-squared	0.004	0.481	0.506	0.504	0.01	0.485	0.506	0.504
N	70268	67812	67812	67812	70268	67812	67812	67812
F statistic	9.142	1140.232	1014.824	851.457	24.855	1125.34	987.839	832.167

NOTES: * p<0.1, ** p<0.05, *** p<0.01. Own Calculations from LFS September 2000 to LFS March 2004. Controls include union membership, firm size dummies, age, age squared, education dummies, a marriage dummy, occupational skill dummies, a public sector dummy and indicators of primary, secondary and tertiary sector work. Where district level covariates are included, the same covariates are averaged, except for age and marriage. The 331 Magisterial Districts are each contained within one of 55 larger District Councils, which are in turn part of one of 9 provinces respectively.

3 Industrial Structure, the Wage Curve and Unemployment

We now proceed to show how industrial structure masks the transitory wage curve in South Africa. Table 5 presents selected statistics to highlight the structure of the labour market in September 2000. It is evident that both larger firms and unionization generally carry premia on wages. In fact, these are correlated features of the labour market, with more than half of unionized workers employed by firms employing 50 or more individuals. This trend is not as stark in the non-unionised sector. However, this provides an indication that unionization is associated with a dearth of small firms, with less than 15% of unionized workers being found in firms that employed less than 10 workers. As noted in the introduction, wage dispersion serves as a proxy for the presence of small firms in a district. This notion is evident in the data: particularly non-unionised workers working in smaller firms live in districts with higher wage dispersion; those that work in larger firms tend to live in districts with lower wage dispersion. Hence, we continue with with this measure as a rough proxy for collective bargaining.

Table 5 Industrial Structure and Labour Market Outcomes

		Firm Size					
LFS2000b		1 worker	2-4 workers	5-9 workers	10-19 workers	20-49 workers	50 or more
Non-union	Mean Earnings	1174.075	1519.266	2414.686	7148.336	10975.810	5080.486
	Broad Unempl	0.356	0.319	0.305	0.297	0.292	0.290
	St Dev[log(Earnings)]	1.013	1.018	0.997	0.979	0.973	0.984
	Workers	965851	894651	824516	917989	831733	1389341
	Proportion	0.166	0.154	0.142	0.158	0.143	0.239
Union	Mean Earnings	2741.365	2302.828	23671.040	9390.330	5500.280	6406.283
	Broad Unempl	0.327	0.302	0.321	0.329	0.319	0.302
	St Dev[log(Earnings)]	1.039	0.993	0.993	1.006	1.013	0.974
	Workers	30265	118218	216076	365083	516080	1320063
	Proportion	0.012	0.046	0.084	0.142	0.201	0.514

Table 6 presents estimates of a wage curve based on a panel of magisterial district averages over time. In this set of estimates, explicit controls for industrial structure are introduced to tease out the mechanism that masks the wage curve in South Africa. Column 1 is the point of departure, where no significant wage curve exists when adding district fixed effects at the same geographic level at which unemployment was measured (as above). The second model controls for unionization, to see whether the lack of the effect is simply because of wage

rigidity caused by high levels of union activity in South Africa. This by itself does not generate a wage curve. Introducing the measure for district-level wage dispersion in column 3 also has no influence in this regard. Controlling for industrial structure by itself (Column 5) still does not yield a negatively significant coefficient. However, other scenarios do. Interacting the unemployment rate with unionization rates of the district uncovers an important feature: a wage curve arises in districts with (hypothetical) unionization rates of 0%, but is soon counteracted as this rate increases. Column 6 controls for *both* firm structure and district-level wage dispersion. Since both represent the dispersion of firm size, and by implication, the absence of small firms, this model more fully controls (though indirectly) for the effects of collective bargaining agreements and other factors that limit the prevalence of small business. In this case, a negatively significant wage curve emerges, highlighting that the "standard" result of a wage curve is masked by rigidities that push workers into high wage large firms and low dispersion districts. By conjecture, the higher wages are correlated with the absence of small firms that would otherwise have created employment. Interestingly, this result is also not far from the 10% elasticity that is almost universally reported. Column 7 again differentiates this measured effect along highly unionized and non-unionized districts. It again becomes clear that this is strongest in districts with low levels of unionization.

This result then shows that the standard levels of wage flexibility would exist in South Africa if the industrial structure did not favour large firms that can afford to meet the wage demands of union workers and workers covered by collective bargaining agreements. By implication, small firms are driven out by unaffordable wage bills, generating high levels of unemployment with rigidly high wages. This argument is in no way causal, as the unemployment rate is not driving high wages, but this descriptive tool shows why the association is not negatively significant in South Africa. For this purpose, we now turn to absorption and participation functions to more directly understand this link[3]. It

[3] We omit unemployment functions here, as they would simply represent the inverse of the wage curve presented above, since we do not yet account for the issue of simultaneity in wage and unemployment determination. Future work

is well-documented that unemployment rose in the post-apartheid period due to a faster increase in participation relative to expansion in employment (Burger & von Fintel, 2009), so that these constituent functions will inform our view of unemployment.

Tables 7 and 8 reveal that wages are positively associated with both participation (always) and employment (in most cases). Naturally, potential workers enter the labour market, attracted by higher earnings potential. The positive coefficient on in the employment equation suggests that jobs are created in high wage districts, which suggests that these districts are also those with a concentration of larger, well-paying firms. In column 3 of the employment function (Table 8), the introduction of wage dispersion yields the coefficient negatively insignificant, which points to this explanation.

From table 7 it is evident that unionisation tends to attract individuals into the labour market (columns 2 and 3), though column 4 uncovers that this is true in high wage districts. In low wage districts, the effect is the opposite. Hence, it is the unionization wage premium that stimulates participation, rather than the resulting lower levels of employment (Table 8, column 2) that would otherwise have deterred it.

Firm size does not influence participation, though the prevalence of large employers in a district ironically reduces employment. This concurs with the evidence of Magruder (2010). Once controlling for firm size (Table 8, column 6 and 7), wage dispersion is associated with higher employment rates; this also indicates that the presence of lower-paying, smaller firms stimulates overall employment rates. Wage dispersion robustly also stimulates labour force participation (Table 7). Smaller firms with lower-paying vacancies are likely to attract a job queue that would not be present if only "big business" were in operation in a particular district. This evidence also becomes clearer when

---

considers an instrumentation strategy exploiting lags as instruments, using the Arrelano-Bond GMM System estimator. Alternatively, geographic data, based on distances from homeland borders will also be used as instruments for labour market outcomes.

looking at the occupational skills and sectoral composition of districts. Black males are less likely to enter the labour force where districts are composed of high proportions of skilled workers, or where higher concentrations of individuals are employed in the tertiary and primary sectors (Table 7). This is precisely because this group is predominantly employed in unskilled occupations and more likely to work in the primary sector (Table 8). If small firms are to create this unskilled work, their absence in certain districts contributes to the unemployment problem.

Together, this body of evidence suggests that a dearth of small firms is generating higher inflows into the labour market (with the attraction of high large firm wage premia), with lower absorption (as these individuals that would otherwise be absorbed by small firms are not necessarily employed). Hence, wage rigidity and industrial structure tends to be contributing to high unemployment, high wage regions, which are dominated by large firms. This concurs with a Harris-Todaro (1970) view of the world, where potential participants are attracted by high wages (in this case, offered by large firms within regions that are subject to strong collective bargaining), but not necessarily absorbed. As pointed out by Magruder (2010), this can be largely attributed to the absence of small firms in South Africa.

Table 6 Wage Curve Estimates Controlling for Industrial Structure

Mean[log(African Male Earnings)]	1	2	3	4	5	6	7
log(Broad Unempl - African Male)	-0.035	-0.061	-0.091	-0.127*	-0.094	-0.115*	-0.190**
Union		1.134***	1.145***	1.012***			0.856***
log(Broad Unempl - African Male) x Union				0.323*			0.324*
St Dev[log(Black Male Earnings)]			-0.103***			-0.073**	-0.100***
Firm size: 2-4 workers					0.104	0.096	0.129*
Firm size: 5-9 workers					0.406***	0.404***	0.313***
Firm size: 10-19 workers					0.659***	0.655***	0.444***
Firm size: 20-49 workers					0.725***	0.723***	0.508***
Firm size: 50+ workers					0.857***	0.852***	0.553***
Educ: Primary	0.279***	0.300***	0.308***	0.290***	0.288***	0.294***	0.296***
Educ: Incomplete Secondary	0.184	0.247**	0.261**	0.243**	0.184*	0.194*	0.247**
Educ: Matric	0.405***	0.450***	0.481***	0.444***	0.399***	0.422***	0.463***
Educ: Tertiary	1.256***	0.861***	0.917***	0.870***	1.072***	1.115***	0.855***
Occup: Highly Skilled	1.376***	1.419***	1.430***	1.407***	1.378***	1.385***	1.409***
Occup: Skilled	1.391***	1.066***	1.091***	1.040***	1.239***	1.258***	1.019***
Occup: Semi-skilled	0.288***	0.227***	0.218***	0.222***	0.285***	0.279***	0.218***
Sector: Secondary	0.017	-0.05	-0.04	-0.045	-0.07	-0.062	-0.075
Sector: Tertiary	0.134**	-0.008	0.008	-0.006	0.179***	0.191***	0.065
Constant	6.266***	6.119***	6.179***	6.154***	5.820***	5.868***	5.919***
District FE	Y	Y	Y	Y	Y	Y	Y
Time FE	Y	Y	Y	Y	Y	Y	Y
R-squared	0.156	0.295	0.299	0.296	0.232	0.234	0.327
N	2393	2393	2393	2393	2393	2393	2393

Table 7 Labour Force Participation Function - African Males

Broad LFP - African Male	1	2	3	4	5	6	7
log(African Male Earnings)	0.010*	0.014**	0.016***	0.019***	0.014**	0.015***	0.020***
Union		-0.031*	-0.036**	0.168			0.055
log(African Male Earnings) x Union				-0.028			-0.012
St Dev[log(Black Male Earnings)]			0.032***			0.031***	0.031***
Firm size: 2-4 workers					0.013	0.016	0.015
Firm size: 5-9 workers					-0.02	-0.019	-0.018
Firm size: 10-19 workers					-0.031	-0.03	-0.026
Firm size: 20-49 workers					-0.023	-0.022	-0.018
Firm size: 50+ workers					-0.023	-0.021	-0.015
Educ: Primary	0.026	0.024	0.022	0.024	0.025	0.023	0.022
Educ: Incomplete Secondary	-0.022	-0.025	-0.028	-0.025	-0.024	-0.027	-0.029
Educ: Matric	0.103***	0.100***	0.091***	0.100***	0.101***	0.093***	0.090**
Educ: Tertiary	0.166***	0.171***	0.149**	0.183***	0.168***	0.147**	0.156***
Occup: Highly Skilled	-0.083**	-0.091***	-0.096***	-0.088**	-0.089***	-0.093***	-0.098***
Occup: Skilled	-0.090***	-0.087***	-0.095***	-0.086***	-0.087***	-0.096***	-0.093***
Occup: Semi-skilled	-0.019	-0.019	-0.017	-0.018	-0.022*	-0.02	-0.019
Sector: Secondary	-0.032*	-0.030*	-0.033**	-0.033**	-0.028*	-0.031*	-0.031*
Sector: Tertiary	-0.034**	-0.031**	-0.035**	-0.032**	-0.034**	-0.039***	-0.036**
Constant	0.662***	0.638***	0.612***	0.608***	0.650***	0.626***	0.595***
District FE	Y	Y	Y	Y	Y	Y	Y
Time FE	Y	Y	Y	Y	Y	Y	Y
R-squared	0.026	0.028	0.035	0.029	0.03	0.037	0.038
N	2393	2393	2393	2393	2393	2393	2393

Table 8 Employment Function - African Males

Employment - African Male	1	2	3	4	5	6	7
log(African Male Earnings)	0.008	0.014**	-0.011	0.024***	0.015**	0.017***	0.026***
Union		-0.042**	0.037*	0.428***			0.196
log(African Male Earnings) x Union				-0.066***			-0.032
St Dev[log(Black Male Earnings)]			-0.067***			0.060***	0.058***
Firm size: 2-4 workers					-0.003	0.004	0.003
Firm size: 5-9 workers					-0.056**	-0.054**	-0.053**
Firm size: 10-19 workers					-0.052**	-0.050**	-0.045*
Firm size: 20-49 workers					-0.054**	-0.052**	-0.047*
Firm size: 50+ workers					-0.069***	-0.065***	-0.056**
Educ: Primary	-0.014	-0.017	0.052	-0.017	-0.012	-0.017	-0.018
Educ: Incomplete Secondary	-0.117***	-0.121***	0.130***	-0.120***	-0.116***	-0.123***	-0.126***
Educ: Matric	-0.073*	-0.077*	0.171***	-0.077*	-0.071*	-0.088**	-0.090**
Educ: Tertiary	0.205***	0.212***	-0.182**	0.240***	0.213***	0.171**	0.192***
Occup: Highly Skilled	-0.096**	-0.106***	0.093**	-0.101**	-0.105***	-0.112***	-0.116***
Occup: Skilled	-0.106***	-0.102***	0.137***	-0.100***	-0.105***	-0.121***	-0.116***
Occup: Semi-skilled	0.012	0.013	-0.030*	0.014	0.01	0.014	0.015
Sector: Secondary	-0.040**	-0.038**	0.044**	-0.044**	-0.034*	-0.040**	-0.041**
Sector: Tertiary	-0.057***	-0.053***	0.088***	-0.056***	-0.061***	-0.070***	-0.067***
Constant	0.497***	0.465***	0.367***	0.394***	0.487***	0.441***	0.382***
District FE	Y	Y	Y	Y	Y	Y	Y
Time FE	Y	Y	Y	Y	Y	Y	Y
R-squared	0.054	0.056	0.093	0.06	0.061	0.08	0.083
N	2393	2393	2393	2393	2393	2393	2393

4 Conclusion and Policy Implications

The Harvard Group of economists has proposed a wage subsidy to target the severe youth unemployment problem in South Africa. Labour market entry has occurred in a large wave amongst the youngest entrants, without being absorbed (Burger, Van der Berg, & von Fintel, 2012), and requires concerted policy effort to solve. That evidence suggests that targeting labour supply with the subsidy may be appropriate. However, the simulated evidence presented by Go *et. al.* (2010) suggests that the policy will largely be ineffective if the labour market remains rigid, and if various skills categories cannot substitute for each other. Each of these still remain dominant features of South Africa's labour market. Essentially this suggests that the policy cannot solve the issue of structural unemployment, nor of labour market inflexibility. The models presented in this paper show that black males tend not to enter the labour market, nor are absorbed, if district employment is highly concentrated in skilled occupations. This is a sector which traditionally has not employed this particular group. Hence, stimulating the creation of this type of work would potentially absorb the long queue of young workers who have recently joined the ranks of the unemployed. At the same time, unionisation fuels rigidity, with high concentrations of employment in large firms, which pay high wages and drive out small firms. This is evidenced by lower wage dispersity in districts with higher concentrations of large firms, suggesting that there is an absent "lower tail", which concurs with the notion that small firms have exited the market. Furthermore, the regression analysis above uncovers the mechanism that leads to both high unemployment and high wages in the same regions: only after controlling for firm size and wage dispersion does a standard wage curve emerge, which suggests that industrial structure is limiting the ability of small firms to lower wages below those set by bargaining councils and to absorb additional workers. Hence, the labour market is limited from clearing by way of wage adjustment, and unemployment ensues. An alternative to focussing on labour supply, therefore, is to consider directing the wage subsidy at small firms, who cannot afford to stay in business due to high wage bills driven up by centralised bargaining. This proposal requires finer attention, as some efforts at supporting small businesses with tax incentives have not necessarily enjoyed success. However, this measure would influence job creation directly, and support small businesses in one of their main costs, as opposed to a tax break that targets costs more generally. This also circumvents the criticism of unions that older workers will be substituted by younger workers under the current proposal; rather, no substitution will occur with the alternative target, as it does not influence the choice of labour supply, nor will it take away workers

from large employers. Substitution effects are likely to play a smaller role, as this approach will likely create new jobs (and employuers) that did not exist or were at great risk of being destroyed, rather than influencing the behaviour of existing employers.

5 Works Cited

Blanchflower, D. G., & Oswald, A. J. (2008). Wage Curve. In S. N. Durlauf, & L. E. Blume (Eds.), *The New Palgrave Dictionary of Economics* (2nd Edition ed.). Palgrave Macmillan.

Kingdon, G., & Knight, J. (2006). How Flexible Are Wages in Response to Local Unemployment in South Africa? *Industrial and Labour Relations Review , 59* (3), 471-495.

Nijkamp, P., & Poot, J. (2005). The Last Word on the Wage Curve? *Journal of Economic Surveys , 19* (3), 421-450.

Magruder, J. (2012). High Unemployment Yet Few Small Firms: The Role of Centralized Bargaining in South Africa . *American Economic Journal: Applied Economics* (forthcoming).

Albaek, K., Asplund, R., Blomskog, S., Barth, E., Gumundsson, B. R., Karlsson, V., et al. (2000). Dimensions of the wage-unemployment relationship in the nordic countries: wage flexibility without wage curve. In S. W. Polachek (Ed.), *Worker Well-being - Research in Labour Economics* (Vol. 19, pp. 345-381).

Blanchflower, D. G., & Oswald, A. J. (1994). *The Wage Curve.* Cambridge: MIT Press.

Harris, J., & Todaro, M. P. (1970). Migration, Unemployment, and Development: A Two-Sector Analysis. *The American Economic Review , 60* (March), 126-142.

Wagner, J. (1994). German Wage Curves, 1979-1990. *Economics Letters , 44*, 307-311.

Papps, K. L. (2001). Investigating a Wage Curve for New Zealand. *New Zealand Economic Papers , 35* (2), 218-239.

Baltagi, B. H., Blien, U., & Wolf, K. (2000). The East German wage curve 1993–1998. *Economics Letters , 69*, 25-31.

Card, D. (1995). The Wage Curve: A Review. *Journal of Economic Literature , XXXIII* (June), 785-799.

Blanchflower, D. G., & Oswald, A. J. (1990). The Wage Curve. *NBER Working Paper , 3181*.

Winter-Ebner, R. (1996). Wage curve, unemployment duration and compensating differentials . *Labour Economics , 3*, 425-434.

Iverson, T. (1998). Wage Bargaining, Central Bank Independence, and the Real Effects of Money. *International Organization , 52* (3), 469-504.

Driffill, J. (2006). The Centralization of Wage Bargaining Revisited: What have we Learnt? *Journal of Common Market Studies , 44* (4), 731-756.

Baltagi, B. H., & Blien, U. (1998). The German wage curve: evidence from the IAB employment sample. . *Economics Letters , 61*, 135-142.

Pannenberg, M., & Schwarze, J. (1998). Labor market slack and the wage curve. *Economics Letters , 58*, 351-354.

Buettner, T. (1999). The effect of unemployment, aggregate wages and spatial contiguity on local wages: an investigation with German district level data . *Papers in Regional Science , 78*, 47-67.

Kennedy, S., & Borland, J. (2000). A wage curve for Australia? *Oxford Economic Papers , 52*, 774-803.

Janssens, S., & Konings, J. (1998). One more wage curve: the case of Belgium. *Economics Letters , 60*, 223-227.

Fedderke, J. (2012). The Cost of Rigidity: the case of the South African labor market. *Economic Research Southern Africa Working Paper , 290*.
Burger, R., & Yu, D. (2006). Wage trends in post-apartheid South Africa: Constructing an earnings series for South Africa from household survey data . *Labour Market Frontiers , 8*, 1-8.
Moulton, B. R. (1990). An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units. *The Review of Economics and Statistics , 72* (2), 334-338.
Burger, R. P., & von Fintel, D. P. (2009). *Determining the Causes of the Rising South African Unemployment Rate: An Age, Period and Generational Analysis.* Stellenbosch Economic Working Paper 24/2009.
Burger, R. P., Van der Berg, S., & von Fintel, D. P. (2012). *The unintended consequences of education policies on South African participation and unemployment.* Stellenbosch Economic Working Papers 11/2012.
Lewis, J. D. (2001). *"The unintended consequences of education policies on South African participation and unemployment.* World Bank Informal Discussion Paper on Aspects of the Economy of South Africa 16.
Kingdon, G., & Knight, J. (2004). Unemployment in South Africa: The Nature of the Beast. *Wordl Development , 32* (3), 391-408.
Bhorat, H., & Hodge, J. (1999). Decomposing Shifts in Labour Demand in South Africa. *The South African Journal of Economics , 67* (3), 349-380.
Go, D., Kearney, M., Korman, V., Robinson, S., & Thierfelder, K. (2010). Wage subsidy and labour market flexibility in South Africa. *Journal of Development Studies , 46* (9), 1481-1502.