

Can current macroeconomic models explain South African unemployment? Employment dynamics in a multiple-segment macro model

Paper presented at the ESSA Conference, Bloemfontein, September 2013

Philippe Burger (UFS) and Frederick Fourie (UFS)

1. Introduction

A recent survey of the South African unemployment research (Fourie 2011) reveals, inter alia, that there has been a limited amount of macroeconomic research on SA unemployment (compared to labour market studies, for instance). What there is, is on output elasticities of employment, sectoral shifts in employment, and possible labour-market related constraints on growth – and some work on increases in the long-run equilibrium rate of employment (the so-called NAIRU) after 1994.

A characteristic of all the macroeconomic work on unemployment in SA – and of macroeconomic policy analysis – is that it deals with the formal sector only. Meanwhile, evidence from unemployment research in the fields of labour economics and development indicate substantial segmentation in the South African economy: between the formal and the informal economies, within the informal sector, and between the survivalist segment and the informal and formal economies. Moreover, several labour market barriers exist that prevent people from improving their employment and earnings situation. Many such barriers impact especially on the poor and flow from the condition of poverty. This forms the basis of much marginalisation, inequality and continued poverty (Fourie 2011:10-44).

The labour market assumptions of standard macroeconomic models (and derived policy proposals) are at odds with such segmentation and barriers. The objective of this paper is to start bridging the divide between the macroeconomic discourse/models and the labour and development discourses on unemployment. Presenting a concise critical evaluation of the ability of standard macroeconomic theories to capture and explain the South African unemployment situation and provide a basis for appropriate policy, the paper finds that at best these theories provide a very partial explanation. It then proceeds to take up the

challenge of including the realities of segmentation, dualism and labour mobility barriers into a theoretical macroeconomic model.

Section 2 summarises and evaluates the labour component and unemployment analysis of mainstream theoretical macroeconomic models, in particular the currently dominant main alternatives: the New Classical model and the New Keynesian model (including multiple-equilibria and hysteresis models). Section 3 evaluates the applicability of these models in the South African, and possibly also developing country, context. The issue of segmentation appears to be critical, which leads to an evaluation of an often-quoted segmented macroeconomic model that has been proposed by Layard, Nickell and Jackman (LNJ). Section 4 takes stock with some thoughts on the road ahead. Section 5 then develops a three-segment model – an expanded LNJ-model of sorts – that illustrates how labour market segmentation and labour market entry barriers can be integrated. It provides novel findings on labour market equilibrium in a segmented context, on the presence of voluntary and involuntary unemployment in such an equilibrium, and on how demand or supply shocks differentially impact, through knock-on and flow effects, on employment and unemployment in the three segments (which include a survivalist segment). The impact of hysteresis in such a model is also considered. On the basis of the potential of such an expanded model, section 6 provides suggestions regarding priorities for macroeconomic research on unemployment and by extension poverty and inequality.

2. The labour market component of mainstream theoretical macroeconomic models

Like the traditional Classical/Keynesian divide of the 1930s to 1950s that turned into the Monetarist/Neo-Keynesian divide of the 1960s and 1970s, modern macroeconomics also is deeply divided on the existence of involuntary unemployment. The modern debate is characterised by the divide between the New Classical and the New Keynesian models. The New Classical model has succeeded in making itself the reference model, putting the onus on others to argue and substantiate deviations from it.¹ Hence, it is also called the New Classical

¹ This can be compared to the situation in an earlier era when the debate was between Monetarist and Neo-Keynesian models. Keynes showed the inadequacies of the Classical model, and his model became the dominant (or reference) model. The Monetarists had to argue deviations from, or limitations of, Keynesian/Neo-Keynesian models.

Benchmark Model (NCBM). This model posits, at its core, an economy in which there is no involuntary unemployment.

The New Keynesian model (NKM) finds itself on the other side of the divide, arguing for the existence of involuntary unemployment. It introduces imperfect product and labour markets as the key explanation for involuntary unemployment. More specifically, it allows for price- and wage-setting by firms and unions. Nevertheless, the NKM attempts to deviate as little as possible from New Classical model. Indeed, in several respects the New Keynesian model is but a variation of the NCBM. Pertinently, these models share a common analytical methodology for deriving macroeconomic behaviour from primary behaviour elements, i.e. the assumption of fully rational and optimising households and firms, linked to a labour market and a production function (see below).

The NCBM was also developed further into the Real Business Cycle Model (RBCM). The RBCM does not allow of involuntary unemployment. Thus, it finds itself on the same side of the divide as the NCBM.

The NCBM, the RBCM as well as the NKM are single-equilibrium models, i.e. models where, after a demand shock causes output to deviate from its pre-shock level, it returns to its initial level. In the case of the NCBM and RBCM equilibrium occurs at full employment, with no involuntary employment, while in the NKM the long-run employment level allows of, and incorporates, involuntary unemployment.

Some extensions of the NKM are multiple-equilibrium models, i.e. models where, after a demand shock causes employment and output to deviate from their pre-shock level, they do not return to any particular level. A prominent example is models characterised by *hysteresis*, which explains a long-run equilibrium with persistent involuntary unemployment.

It is clear that macroeconomists are deeply divided on the presence of involuntary unemployment. Together the NCBM and the RBCM constitute the view that unemployment is never involuntary. Economists who have adopted the NCBM and RBCM also view unemployment as much less of a problem than those who use the NKM. Those who accept the possibility of multiple equilibria see unemployment as even more of a problem.

However, as noted above, those models that incorporate involuntary unemployment equilibria define their models in terms of deviations from the NCBM. Therefore, the discussion of the range of models starts with a brief discussion of the NCBM, focusing only on its derivation and basic features.

2.1 New Classical Benchmark Model (NCBM) (The rational expectations model)

2.1.1 Deriving the NCBM

The NCBM is derived in a few simple steps. First a utility function and budget constraint is specified. Then the budget constraint is substituted into the utility function, which is then solved with respect to the employment variable. This produces the *labour supply* function. (Alternatively, setting wages in the labour supply function as left-hand variable for periods t and $t+1$ allows the derivation of the *Phillips curve*.) Using the production function to derive the marginal product of labour yields the *labour demand* function. The labour supply and demand functions together yield the equilibrium level of employment and wages.

2.1.2 Basic features of the NCBM

In the NCBM microfoundations are explicit. The model assumes rational expectations (i.e. all agents have full information on the ‘true’ model of the economy) and there is a total absence of any nominal rigidities of wages and prices (no contracts). Thus, the model implies an immediate adjustment to equilibrate supply and demand in markets.

The labour market component of the model is based on a labour supply (L^S) and labour demand (L^D) component where labour demand is determined by the marginal product of labour (Carlin and Soskice 2005: 568; 580). Long-run equilibrium in the model is supply-side determined, with a rapid return to long-run equilibrium (or NAIRU) after a disturbance. Thus, there is no distinction between short-run and long-run equilibria.

In equilibrium there is no involuntary unemployment – only frictional unemployment – and therefore no unemployment problem: “No unemployed persons in this model, only nonparticipants” (Cahuc and Zylberberg 2004:459).

Aggregate demand (and demand policy) can have no effect on equilibrium employment, not even in the short run. However, as pointed out by Cahuc and Zylberberg (2004:455): “This prediction is not verified by data”. A Phillips curve exists in the model, but due to rational expectations the relationship cannot be exploited systematically. Only unexpected changes in demand can lead to temporary Phillips curve effects (i.e. using demand policy to reduce unemployment in return for accepting higher inflation).

2.2 Real business cycle model (RBCM)

Deriving the RBC model is straightforward and based on the NCBM. It is the intertemporal re-specification of the classical model and therefore represents an extension of the NCBM. As such, it shares the same microfoundations with the NCBM.

A distinguishing feature of the RBCM is that business cycles mainly are modelled as fluctuations *of* the equilibrium (rather than deviation from equilibrium), meaning that business cycles are seen and modelled as equilibrium phenomena. It also means that another feature of RBCM is that of continual optimality. Equilibrium employment and output vary over time due to technological and other supply-side shocks. These variations in the equilibrium constitute the business cycle. In equilibrium (which virtually always exists) there is no involuntary unemployment – only frictional and search unemployment – and whatever unemployment level transpires, is optimal, since it reflects the optimising responses of agents to technology or preferences. Therefore, strictly speaking all unemployment is voluntary. Since all unemployment is voluntary, in this theoretical model there is no unemployment problem, no business cycle problem and hence no stabilisation role for macroeconomic policy.

2.3 New Keynesian model (NKM)

2.3.1 Deriving the NKM

As mentioned above, the NKM is defined in terms of its deviation from the NCBM. It is derived using the same steps as for deriving the NCBM. However, it introduces imperfect competition (monopolistic competition) through price-setting behaviour of firms, which is substituted into the constrained utility function as used in the NCBM. (It enters through the

equation “wage = price minus mark-up”, where the mark-up depends on the elasticity of demand – i.e. a typical monopoly model). Thus, the New Keynesian Model is derived by amending the NCBM – it is a “spanner in the works” model.

2.3.2 *Basic features explaining long-run involuntary unemployment*

The NKM also has explicit microfoundations. Like the NCBM and the RBCM, it assumes *rational expectations*, with rational intertemporal optimising behaviour displayed by all agents. However, the model differs from the NCBM and RBCM by accepting price setting and wage setting in *imperfectly* competitive markets. This implies rigid or slowly-adjusting prices and wages. Nevertheless, it can be said that the NKM wholly adopts major elements of the New Classical model and that there is not much ‘Keynesian’ about it (also with regard to aggregate demand effects; see below).

The New Keynesian model strives to deviate as little as possible from the New Classical model, but does insist on inserting the imperfect competition assumption. Using this assumption the model derives its important involuntary unemployment result.

As in NCBM the NKM long-run equilibrium employment level (NAIRU) is supply-side determined. However, assuming imperfectly competitive goods and labour market conditions (and thus inflexible prices and wages), the model results in a long-run equilibrium with involuntary unemployment (in addition to frictional and search unemployment). Therefore, persistent unemployment exists even in the face of rational expectations and maximising agents. Indeed, it results from such behaviour.

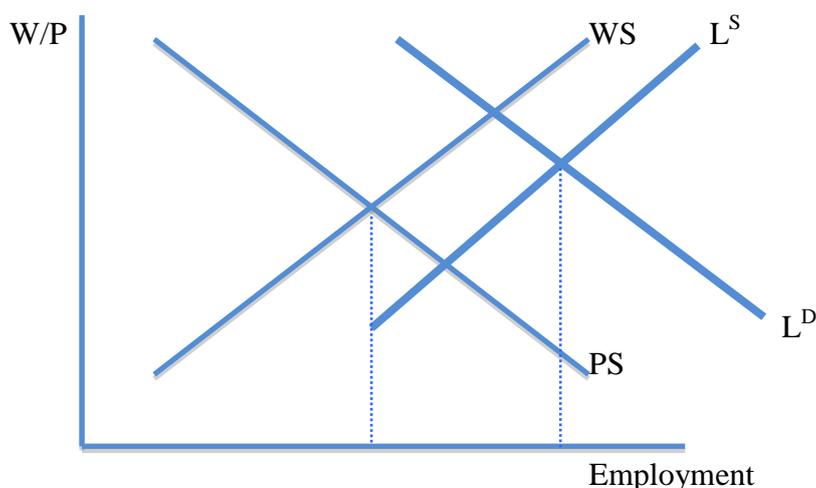
Two types of models can be differentiated on the basis of the imperfect market conditions included. In the first type firms are price setters while unions are wage setters. In the second type firms are both price *and* wage setters. Such price- and wage-setting behaviours change the model significantly. Instead of labour supply and demand, the NKM uses wage-setting (WS) and price-setting (PS) relationships to derive labour market equilibrium. Diagrammatically WS is upward-sloping, while the PS downward-sloping.² Figure 1

² That is if diminishing marginal returns to labour (MPL) exists; if MPL (and the constant mark-up) is constant, PS will be or horizontal (Carlin and Soskice 2005: 49; Cahuc and Zylberberg 2004:381).

demonstrates that the WS-curve lies above the Classical L^S -curve, and the PS-curve lies below the L^D -curve.

The long-run equilibrium level of employment is below the competitive/Classical equilibrium level. The difference between the two equilibrium employment levels represents long-run involuntary unemployment, caused by imperfect competition conditions.

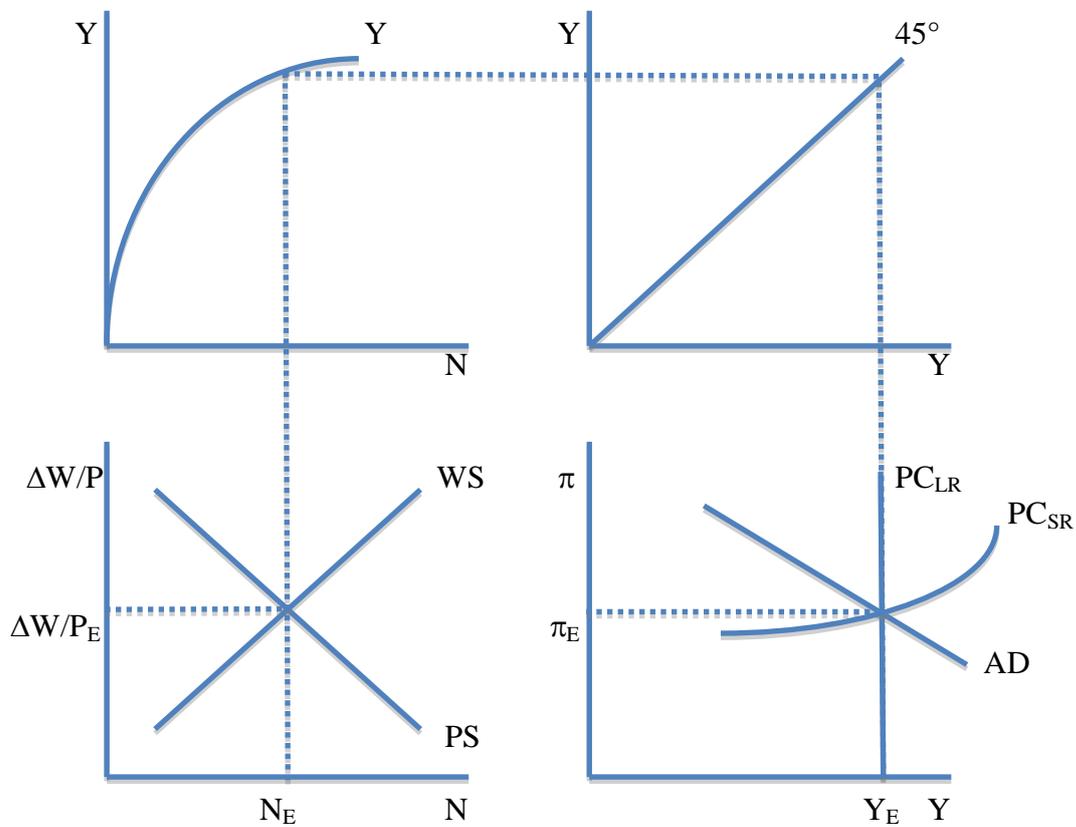
Figure 1: The NKM labour market and the PS-WS system



To complete the model, the labour market and resulting employment level are linked to output via a production function, as demonstrated in Figure 2.

- Note that, unlike Figure 1, where the real wage appears on the vertical axis of the WS-PS system, the WS-PS system in Figure 2 has the *rate of change* in the real wage rate on the vertical axis. This allows putting the inflation rate instead of the price level, in relation to output, in the bottom right-hand graph. With inflation on the vertical axis the AS curve becomes the Phillips Curve.
- Figure 2 also presents a short-run Phillips Curve that has a positive slope which gets steeper at higher output levels. The steepness results from the assumed decreasing marginal product of labour that also causes the production function in the top-left graph to increase at a decreasing rate, and the PS in the bottom-left graph to have a negative slope.

Figure 2: PS and WS, output and inflation



In the NKM, profit-maximising wage-setting and price-setting behaviour determines the equilibrium wage rate. Because of factors such as the menu costs of price setting, bargaining costs of wage setting, and the use of long-term, multi-period contracts, wages and prices adjust slowly (called wage and price stickiness). Therefore persistent involuntary unemployment is not a pure labour market problem caused by excessive wages or, for example, labour market regulations. It is the outcome of rational optimising behaviour of firms and consumers/workers in a specific competitive context (i.e. imperfect competition).

Contrary to earlier Keynesian and Neo-Keynesian models, aggregate demand deficiency cannot be the cause of a long-run equilibrium with persistent/long-run involuntary unemployment in the NKM. Should a demand deficiency occur, the short-run Phillips curve will adjust (as expectations catch up with actual inflation) and return the equilibrium to the

vertical, long-run Phillips-curve. In this respect the NPK fully accepts the NCBM result on aggregate demand impotency in the long run.

2.3.3 *The short run: business cycles and the Phillips curve*

The NKM differs significantly from the former two models in terms of its treatment of the short run. In the NKM the business cycle is explained as fluctuations *around* the equilibrium. Business cycles in the NKM are interpreted as *disequilibrium* phenomena (i.e. disequilibrium relative to the long-run equilibrium and WS-PS intersection) – in contrast to the RBCM where business cycles are seen as movements *of* the long-run equilibrium employment and output levels themselves (and therefore are seen as equilibrium phenomena).

In the NKM short-run or cyclical unemployment is due to demand or supply shocks; combined with expectational mismatches due to existing contracts. In addition, imperfect competition means that quantity (i.e. output and employment), rather than prices and wages, changes first (after a demand shock, say). Thus, adverse demand shocks reduce output and employment in the short run. The effects of these shocks might be either amplified through multiplier effects that operate through consumption, or dampened because of anticipatory consumption smoothing. There is a slow return to long-run equilibrium / NAIRU (due to, inter alia, inflation inertia (Carlin and Soskice 2005:595)). This slow return creates room for a short-run Phillips curve, with the NKM tradition containing various forms of the Phillips curve (including sticky prices with staggered price setting and sticky information resulting from delayed information on monetary policy).

2.3.4 *Two subtypes of NKM models*

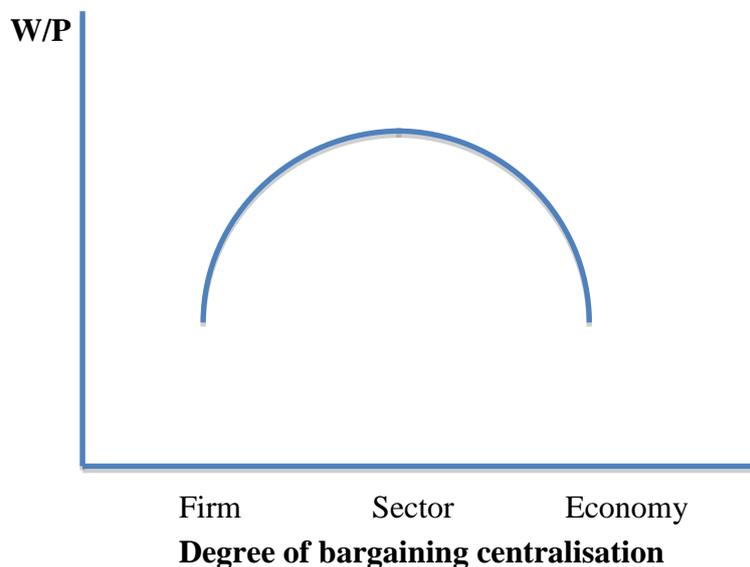
As noted above, in the NKM tradition two types of models can be differentiated on the basis of the imperfectly competitive market behaviour that they include.

- a. The first type is models with firms as price setters and unions as wage setters ('union models'). In the union model monopolistic firms use mark-up pricing while the union acts as a monopolistic wage setter.
- b. The second type is models with firms as both price and wage setters (efficiency wage models). In the efficiency wage models monopolistic firms use mark-up pricing *and* set efficiency wages.

A key feature of the efficiency wage models is the introduction of the assumption that worker productivity is a function of wage levels. At each level of employment firms set the wage above the wage at which workers are willing to supply their labour. In graphical terms, the wage offered by the firm at each level of employment is above the L^S curve and thus above any L^S-L^D equilibrium wage. Firms offer these higher wages because higher wages increase the potential loss that workers could incur if caught shirking. Hence, higher wages serve as an incentive to be more productive. Generally the focus is on the upper range of wages and skills, where the wage-productivity nexus is assumed to be important.

One well-known union model (and one that might be relevant to the South African situation, see discussion below) is the Calmfors-Driffill model. In this model the relationship between wages and the level at which wages are negotiated has an inverted U-shape (see Figure 3). Bargaining on economic-sector level leads to higher wages compared to the results obtained when bargaining occurs on firm or economy-wide levels.

Figure 3 – The Calmfors-Driffill model



2.4 Multiple equilibria models

2.4.1 *Basic features*

All the above models are single equilibrium models in the sense that those factors that might cause the WS and PS functions to change/shift (such as the mark-up by firms, the level of union power or the unemployment benefits paid by government), are assumed to be variable only in the long run. Thus, in the above models, long-run equilibrium (un)employment can vary/shift over time due to various types of structural or institutional changes which impact on WS and PS (creating a “time-variant NAIRU”).

Multiple equilibria models add a further dimension to structural or institutional changes that may impact on wage- and price-setting behaviour. However, the structural changes added by the multiple equilibrium models are assumed to have a higher frequency of occurrence than the structural or institutional factors noted above. It is this higher frequency that leads to these models being denoted as ‘multiple’ equilibrium models – the changes occur in what might be denoted the medium term. The best known multiple equilibria models are hysteresis models, which mainly were developed to help explain persistent unemployment (especially in the European context; see Blanchard (2005) and its list of references). Other models are based on procyclical labour productivity (“high-road/low-road” models) or on the role of fairness in the utility function of workers.³

2.4.2 *Hysteresis models*

The hysteresis model is an extension of the NKM model. The hysteresis model incorporates the monopolistic price setting behaviour of firms and typically also the wage setting behaviour of unions. As is the case with the standard New Keynesian model, the hysteresis

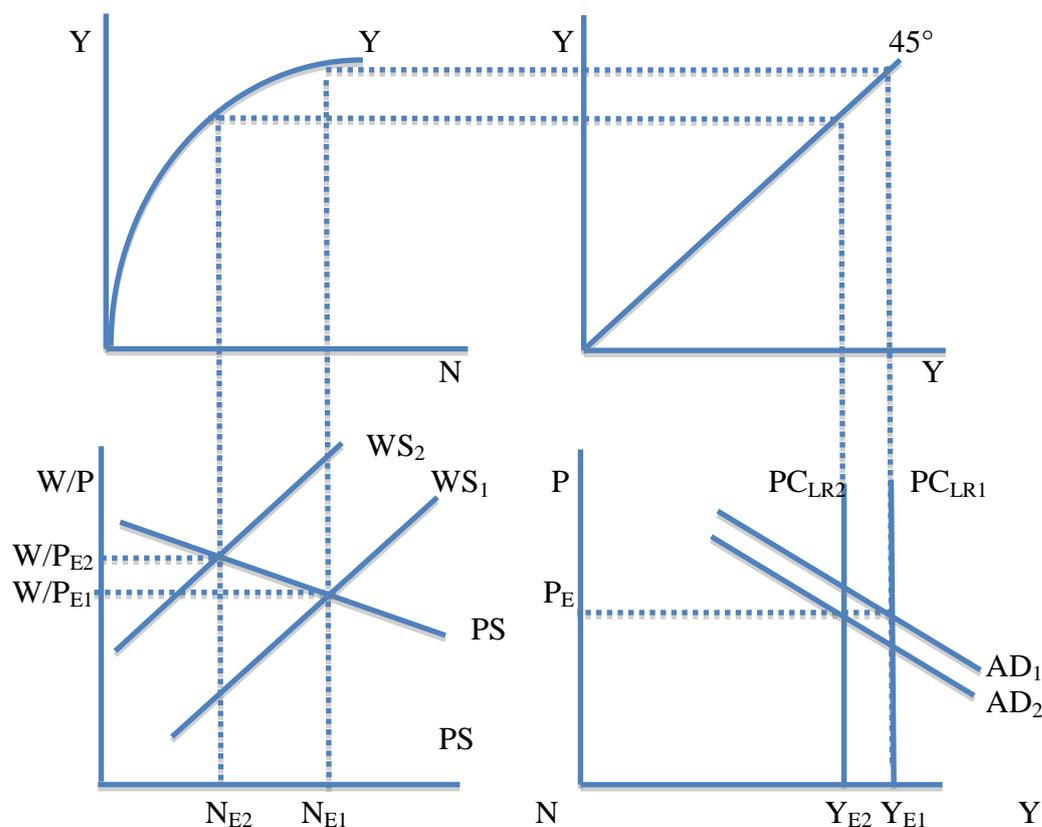
³ The “high-road/low-road” models analyse the existence of multiple equilibria on the basis of procyclical labour productivity. At high levels of income, productivity is high, and so are wages and employment, while at low levels of income, productivity is low, and so are wages and employment. (However, there is no mechanism that explains how an economy ends up in the high (low) income, productivity and employment range. However, one might postulate that exogenous shocks are responsible for where the economy ends up, while an efficiency wage argument could explain the changing productivity of workers.) In this model the economy can get stuck in a low-employment equilibrium (*or* a high-employment equilibrium). Another type of multiple-equilibria model postulates the existence of a range of equilibrium (un)employment rates that each depends on the role of fairness in the utility function of workers.

model allows of involuntary unemployment. In addition, it explicitly considers multiple equilibria by focusing on how the economy can move from one equilibrium to another. Once an equilibrium employment level is disturbed by a demand shock, employment can settle at a new equilibrium level and therefore display no tendency to return to its previous equilibrium level. Higher actual unemployment determines a new equilibrium unemployment level – therefore, the equilibrium is a function of the history of unemployment. This implies unemployment persistence. Moreover, not only does the hysteresis model explain the persistence of high unemployment and the presence of involuntary unemployment, it also implies that involuntary unemployment can occur at any level of unemployment.

There are various versions of the hysteresis model (Carlin and Soskice 2005: 117-20; 617-20). The first version is based on deskilling. Unemployment is initiated by a fall in aggregate and hence labour demand, but perpetuated by the loss of skills and productivity caused by the higher actual unemployment. The deskilled unemployed become less employable, which translates into less competition for jobs. Graphically, the WS shifts left and up (Figure 4, bottom left quadrant). Long-term equilibrium unemployment increases. There is no convergence or return to a stationary equilibrium (Cahuc and Zylberberg 2004: 480). When demand is restored, the unemployed do not necessarily find employment again. The equilibrium is at a lower level of employment.

Note that this process is asymmetric. While a negative demand shock might cause the supply of labour to decrease (because skills decrease), a positive demand shock will not cause the supply of labour to increase. Because of a lack of skills the supply of labour is constrained, meaning that, following the positive demand shock, there is a shortage of skilled workers to employ. This might thus explain why a restoration of demand after a slump does not necessarily translate into more employment.

Figure 4 – Hysteresis due to skills loss



The implication of this model is that changes in the wage depend not on the stock of unemployed, but only on changes in number of short-term unemployed (Cahuc and Zylberberg 2004: 480). Thus, the long-term unemployed do not affect wages, they are outside the labour market (Cahuc and Zylberberg 2004: 480) and they are involuntarily unemployed. Therefore, this is a form of segmentation, a case of the ‘heterogeneity of the unemployed’ (In extreme form it leads to search and matching models – Romer 2012: 458, 486-98). In so far as one can still call it a *long-run* Phillips curve, an increase in the long-term unemployed will cause a fall in the long-run output level, which will cause the long-run Phillips curve to shift left (Figure 4, bottom right) (it might now be more apt to call the vertical Phillips curve the medium to long-run Phillips curve).

A second version of the hysteresis model focuses on the scrapping of capital. The increase in the long-term unemployment is initiated by a fall in aggregate and hence labour demand, but perpetuated following the scrapping of capital, itself caused by the fall in demand. Graphically the PS curve shifts downwards (if it has a negative slope), or is truncated (if it is

flat). As demand picks up again, profits rise, but no new capacity is added in the immediate aftermath of the rise in demand. In the long-run, though, it will be profitable to add new capacity. Thus, this is seen as a quasi-hysteresis in that long-run unemployment increases in reaction to an increase in actual unemployment, but when in future companies start adding new capacity, long-run unemployment will decrease.

The third version of the hysteresis model focuses on insider-outsider union behaviour, where insiders take active steps to exclude outsiders from the work place. Long-term unemployment is, as in the above models, initiated by a fall in aggregate and hence labour demand, but perpetuated by the insider-outsider behaviour within labour unions. This is graphically represented by a kinked WS-curve that is positive up to the point of inflection, where after it becomes vertical.

Note that in all the hysteresis models an aggregate demand shock is the initiating factor, but the *persistence/perpetuation* of the higher unemployment (i.e. its hardening into a new, lower-employment equilibrium) is due to supply-side behaviour. Such supply-side behaviour comprises either firms that scrap capital, workers who lose skills or union that exclude the unemployed. The results imply that aggregate demand, by affecting the short-run equilibrium, can alter long-run equilibrium (un)employment. Transitory shocks may have permanent effects.

One can also consider examples where supply-side shocks are the initiating factors. For instance, industries that due to a liberalising economy face international competition and then close down because they are uncompetitive, may leave many unemployed. If these workers lose their skills, the effect might also be a form of hysteresis.

3. Applicability of the mainstream models to the South African situation

This section considers the applicability of the labour component of the above models to the South African and developing country situation.

3.1 NCB/RBC family

It is apparent that there are not many places in these models where characteristics peculiar to the South African labour market situation – noted in the Introduction – can be inserted or incorporated. In considering ‘applicability’, one must remember that in terms of the ‘accepted methodology’ these models are ‘not supposed to be realistic’ – they adopt a simplification intent on demonstrating that undisturbed markets yield optimal outcomes in terms of maximised utility. In these models there is only one (i.e. formal) sector; there is no secondary or informal or peripheral sector). Those who are unemployed are assumed to have *chosen* not to be employed at the going market wage – they are *voluntarily* unemployed. These unemployed simply are *not in the labour force*. In addition, those who cannot find employment can simply find self-employment.

The value of this model, which explains unemployment as a voluntary condition, seems rather limited in a country where according to the narrow definition 25% of the labour force is unemployed.

However, the RBC model might provide interesting insights as it highlights the impact of technological and supply-side shocks. It could find potentially interesting applications in a developing country that is subject to, for example, “technology import shocks” of which many are labour-saving and high-skills demanding.

3.2 NPK family

Models in this family have undergone much refinement to explain especially European unemployment and hysteresis (especially ‘Euroclerosis’ in the 1980s and 1990s). These refinements reflect imperfect competitive behaviour (price and wage setting behaviour) and more realistic behavioural assumptions (e.g. the wage-productivity link in efficiency wage models).

Given the oft-cited high levels of concentration in many South African goods markets, the price-setting model (with mark-up behaviour by companies at its core) appears relevant to the South African situation. The union wage-setting model also appears relevant to the South African situation, especially variations on the insider-outsider model or the Calmfors-Driffill model of bargaining centralization.

Efficiency wage models also appear to be relevant. With productivity as a function of wages, this may suggest a potentially important way to incorporate worker motivation and productivity *also at the lower end of wage spectrum*. Another question would be whether there could be something like “inefficiency wages”? Link to this is the question whether the fairness model is a potentially important way to incorporate issues of fairness and entitlement into South African labour market behaviour. However, given the existence of strong unions in many industries, it is a question whether firms are wage setters in South Africa (i.e. the union model may be more applicable)

3.3 Segmentation? Weaknesses of the NKMs

Although the New Keynesian models appear to be more applicable to a situation such as that of South Africa, they have several limitations that need to be noted. First, all these models are only applicable to the primary (e.g. formal) sector. Nevertheless, there are hints at segmentation:

- Efficiency wage setting by firms implies a primary and a secondary sector.
- Union wage setting implies insider-outsider situations.
- Skills loss by long-term unemployed (heterogeneity of the unemployed) implies that wages are not influenced by the stock of such long-term unemployed. This suggests that these *long-term unemployed are outside the labour market*.

However, even though such elements imply a kind of segmentation, no attention is paid to the implications of segmentation for employment and unemployment. These models are silent on the secondary market or the ‘outsider’ market or the survival of the long-term (involuntary) unemployed. The economic activities of those in the secondary sector are treated as a residual not requiring further analysis or explanation.

With little attention is paid to a secondary sector, there also is no place for behavioural innovations such as a wage-productivity nexus or a fairness-productivity link for those in the secondary sector. Are secondary sector agents not ‘rational and optimising’ like primary sector agents? If they are rational and optimising (and why would they not be?), why not analyse the secondary sector explicitly? If they are not, what determines their behaviour? Or is the lack of attention a case of “they do not matter because they are the ‘other’ ...”?

The NKM and hysteresis models both can explain why workers lose employment, whether in the short or long run. However, what these models do not consider or explain is why those who lose employment then become unemployed and not self-employed. To do so, these models, will have to generalise the basic skills argument of the hysteresis model and state that those employed have a different skills set than those who are self-employed. Thus, once workers lose their job, their skills set (aimed at employment) might not allow them to become self-employed. In addition, the continued depreciation of their employment skills set means that after a while they have no relevant skills set left, and hence remain unemployed. Note that this weakness is not only a weakness in explaining unemployment in South Africa, but also in a European/First World context.

A failure to explain the imperfect substitution between employment and self-employment is a weakness in most of these models. Indeed, the proper and refined treatment of self-employment is a substantive gap in most models. Some models ignore it, other summarily include it in the category of ‘workers’ (normally in a secondary sector) – ignoring the question, for example, whether self-employment is a labour market issue or an emerging firm issue, and what the dynamics between these two issues may be. This may be quite pertinent in a developing country context.

3.4 The segmented labour market macro model of Layard, Nickell and Jackman (LNJ)

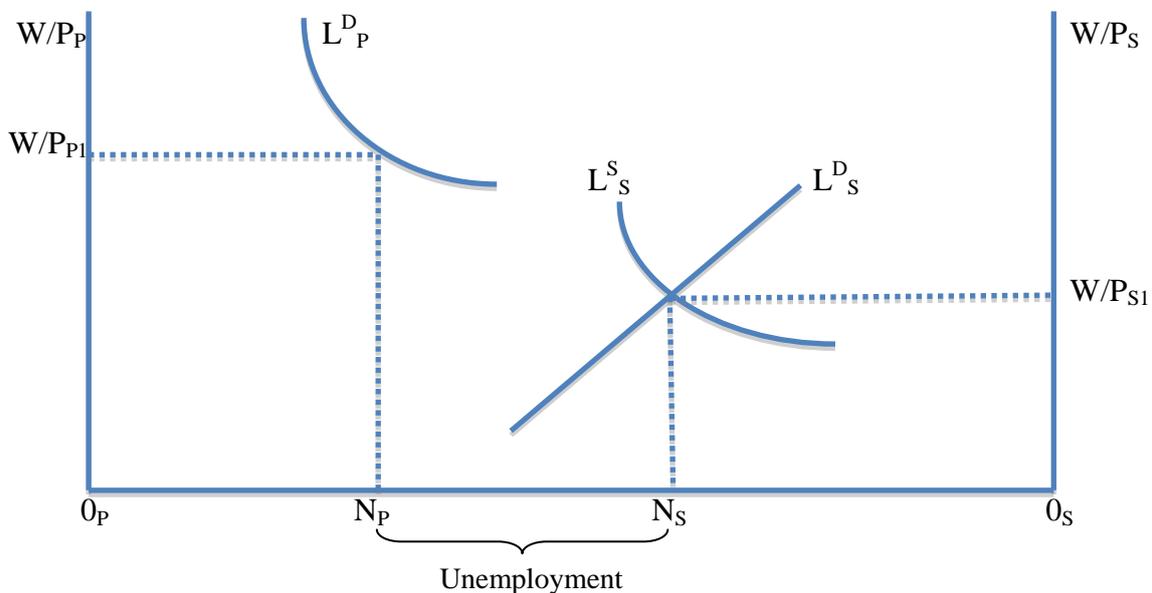
This section discusses an explicit segmented market macro model. It is the model of Layard, Nickell and Jackman (1991:41-44; also 2005), the most well-known attempt to incorporate labour market segmentation into a macroeconomic model.

3.4.1 Basic features

Segmented labour market models typically include a primary and a secondary sector. The LNJ model does the same. With respect to the primary sector its features are typically New Keynesian. They do not use the WS-PS framework, but their model can be translated into this framework without any harm (see section 5 below).

In the primary sector, wage setting is done by firms as efficiency wages or it occurs via union bargaining. In terms of Figure 5 wages in the primary sector are set at W/P_{P1} . Note that because the wage is set above the labour supply L^S_P , only points on L^D_P are relevant. Wage setting determines primary sector employment. Reducing primary sector wages would increase labour demand and employment. In equilibrium there are not many vacancies, and no shortage of labour for jobs. The excess primary sector labour supply – those excluded (or rationed) from the primary sector – go to the secondary sector. In the Figure 5 the secondary sector is measured on the horizontal axis from zero on the right-hand side. (The length of the horizontal axis indicates the entire labour force, i.e. the economically active.)

Figure 5 – The LNJ two-sector model



While the primary sector is New Keynesian in nature, the secondary sector is surprisingly very New Classical. The secondary sector labour market is assumed to be market clearing “in the sense that wages are not high enough to attract a queue of job-seekers, nor do vacancies

last long since skill requirements are low” (Layard et al 1991: 42). In Figure 5 the wage in the secondary sector adjusts to clear at W/P_{S1} , where labour supply in the secondary sector, L^S_s , equals labour demand L^D_s . Layard *et al* include self-employment in their definition of ‘employment’, so the voluntarily unemployed are not self-employed. If secondary sector wages were lower, employment would fall (due to a reduced supply of labour). The rising labour supply curve is generated by a range of reservation wages of different people in the secondary sector.

Equilibrium produces an employment level in each sector, presumably feeding into corresponding production functions and a combined aggregate output. Employment also produces a group of economically active people that are “between the two sectors” – the distance between N_P and N_S – that are *both voluntarily and involuntarily unemployed*. They are “involuntarily unemployed with respect to primary sector” at the going wage there, but simultaneously they are “voluntarily unemployed with respect to the secondary sector” (i.e. not willing to work at the going wage in the secondary sector).

If shocks occur that affect primary sector employment, that will be reflected in changes in secondary sector labour supply and employment as well as in unemployment. Thus there are spill-over effects from the primary sector. (Such dynamics will be analysed further in section 5 below.)

Compared to the standard New Keynesian models the LNJ model introduces the possibility that labour market conditions are not homogenous across the economy. More specifically, their model allows of the existence of a persistent wage differential in the economy, with wages in the primary sector being higher than wages in the secondary sector. The New Keynesian features of the primary sector cause a lower employment rate in the primary sector, with those who would have been employed in the primary sector now supplying their labour in the secondary sector. This additional labour supply in the secondary sector causes wages in that sector to be lower than those in the primary sector. (In the absence of the New Keynesian features, wages theoretically would be expected to be uniform across the two sectors, rendering the distinction between the two sectors redundant).

3.4.2 *Applicability of the LNJ type of segmented model to SA and developing country situations*

The LNJ model appears to be a promising improvement on the standard New Keynesian model. The model explicitly recognises segmentation and shows the existence of a secondary labour market, unlike all the other models. But there is limited analysis of the labour market behaviour of those in the secondary sector. And the LNJ model suffers from the same New Classical critique: Why do many of those who become unemployed in the primary sector, stay unemployed and not all become (self-)employed in the secondary sector?

Even though the LNJ theoretical model contains a sector with New Keynesian features, these features do not, as in the one-sector New Keynesian models, produce or explain the existence of involuntary unemployment on the aggregate level (though workers can be involuntary unemployed with respect to the primary sector). The assumption of a perfectly competitive, market-clearing secondary labour market delivers this result. In this sense the Layard *et al* model ends up being quite close to the NCBM in which there is also no aggregate involuntary unemployment. However, as will be demonstrated below, the unemployment rate in the model exceeds the unemployment rate that would exist in the absence of New Keynesian features.

In addition, while Layard *et al* introduce their much-quoted segmented-model diagram in the first, overview part of book, they immediately *abandon* it and proceed to in-depth analysis only of the primary sector. They are short-lived segmentationists. There is no further analysis of labour market behaviour of those in the secondary sector. Also, there is no analysis of the flows of labour between the two sectors – which, potentially, is crucial to understanding employment and unemployment. What are the determinants and dynamics (and possible hindrances) of such transitions?

Moreover, there is no analytical mention, appreciation or consideration of the economic activities of those excluded even from the secondary market: the *tertiary* segment (the ‘default’ labour market, e.g. a survivalist segment) and, obviously, no analyses of their transitions to and from the secondary or primary sectors. They simply are seen as voluntarily unemployed and thus unproblematic.

4. Overall evaluation and road ahead: Two key issues

At present macroeconomic theory from whichever school of thought analytically more or less ignores the secondary sector – and it has absolutely no theory of the tertiary segment (i.e. those who drop out of the first two segments altogether). Consequently there is no theory of inter-sector labour flows and their determinants and dynamics. In addition, there also is no theory to explain the imperfect substitutability between employment and self-employment that can help answer the New Classical critique, which is: why do workers who lose employment become unemployed and not self-employed? Is it simply a choice due to a too low available wage/earnings – or are there barriers and constraints that explain the situation?

Two key issues arise. The first relates to the need to explicitly account for segmentation. It is necessary to open the analytical box with regard to the *secondary* and *tertiary* segments, i.e. to analyse and model labour market behaviour within all segments. At the very least the secondary sector should be analysed in depth.

The proper treatment of the self-employed, notably outside the formal sector/professions, remains a major analytical gap and inconsistency in the various models. It reflects the failure to deal with those that are unemployed (but obviously still stay alive in some way...)

There is also a need to analyse flows of labour between the two sectors – potentially crucial to understanding employment and unemployment. What are the determinants and dynamics of such flows? Can they be modelled?

The second issue relates to models that can account for a non-clearing secondary labour market, and thus long-run involuntary unemployment. The latter implies the consideration of barriers to entry and mobility barriers between the sectors. None of the models discussed above allow of the possibility that entry and mobility barriers might exist for those who might wish to supply their labour services to either of the sectors.

This is particularly relevant for application to the South African context given that the existence of a range of entry and mobility barriers is central to the findings from much South African labour market and development literature on unemployment and employment (cf. Fourie 2011: 41-44). These include skills mismatches, geographical-spatial factors such as transport

costs, lack of work experience, household culture with respect to work experience, work ethic and search; lack of information about jobs and jobs environment, lack of labour market networks, lack of resources to support search. Entry and mobility barriers mean that labour supply is not merely a function of wages. Any job search activity implies a balancing of risks/expectations and costs. Search barriers and costs raise the reservation wage for job searchers, restrain and, beyond some level, truncate labour supply.

Thus it is necessary to incorporate entry and mobility barriers that affect transitions between the three segments – notably from tertiary to secondary, and then from secondary to primary segments. In this way the analysis of barriers links up with the analysis of inter-sectoral flows.

The next section develops a model, roughly along the lines of the LNJ model, that attempts to include some of these dimensions that apply to the South African labour market.

5. A three-segment ‘barriered’ macroeconomic model: Augmenting the model of Layard *et al*

5.1 Outlines of the model structure

The basic framework developed here is a *three-segment model* comprising

- (a) a *primary sector*, the employment sector of choice for all workers;
- (b) a *secondary sector* that serves as the second-choice employment (and self-employment?) sector for workers (often seen as a ‘bad jobs’ sector), and
- (c) a *tertiary segment* to approximate a last-resort, survivalist or subsistence segment (frequently found in developing countries).⁴

The stylised characteristics of these segments can be described as follows:

- (a) *Primary sector (as in New Keynesian one-sector & LNJ two-sector models)*: Firms as monopolistic competitors and therefore price setters, and wage setting by either unions or firms (efficiency wages). This is seen as the ‘*good jobs*’ sector, with work stability and security. However, this labour market does not clear, leaving some workers involuntarily

⁴ Otherwise conceptualised segmentations (or sub-segmentations) could be entertained and probably be dealt with in similar ways.

unemployed. They fall back on the secondary sector for employment (whilst perhaps continuing to search for primary sector jobs). Low wages or fairness/entitlement aspects may impact on productivity and motivation of unskilled workers (inefficiency wages?).

- (b) *Secondary sector/segment*: Not assumed to be perfectly competitive – no *a priori* association between e.g. informality and perfectly competitive conditions. On both the labour demand and labour supply sides, the secondary sector economy faces obstacles and lacks continuity and security. Thus, the derived demand for labour is insecure and fluctuating – it is the *bad jobs sector*. This impacts on magnitude, continuity and quality of labour supply. There also are distance, network, information and other barriers to entry. Low wages or fairness/entitlement aspects may impact on productivity and motivation (inefficiency wages?). Human skills depreciate due to infrequent employment. As a result, labour markets in the secondary sector do not function smoothly to clear. Many workers remain involuntarily unemployed (or underemployed).
- (c) *Tertiary/survivalist segment*: Those remaining under- or unemployed have to fall back on the tertiary or survivalist segment. It is the default or last-resort segment, the '*no-jobs*' segment. There is no labour market or market-oriented self-employment to speak of. Livelihoods rely on subsistence farming or similar urban activities, survivalist activities, or scattered trading or work opportunities (in addition to social or intra-family transfers, gifts, etc.) whilst continuing to search for jobs, or alternatively becoming non-searching/discouraged (whilst remaining part of broader labour force).

This model is developed by augmenting the Layard *et al* model by including barriers to entry with regard to the secondary sector – and explicitly recognising the tertiary segment. Such labour market barriers will limit the effective supply of labour to the secondary sector. A limited labour supply causes the unemployment rate to be higher than what it would be in the absence of the barriers. Thus, unemployment in the augmented model will be higher than in the LNJ model, which in turn is higher than in the NCBM. Because (in the augmented model) these additional unemployed would have supplied their labour bar these barriers, they can be characterised as involuntarily unemployed.

Thus, a key feature of the barrier-augmented model is to demonstrate the possibility of the existence of involuntary unemployment in long-run equilibrium. It will also be shown that the reduced secondary sector labour supply implies a smaller wage differential between the two sectors.

These involuntarily unemployed constitute a third, or tertiary, segment characterised by the absence of employment earnings and, consequently, poverty for many (although poor workers can and do occur in all three segments). This segment can be conceptualised under either a narrow or a broad definition of unemployment. It may also be defined to include persons who technically are classified as ‘not economically active’, but do engage in survivalist or subsistence economic output activities mostly for own consumption.

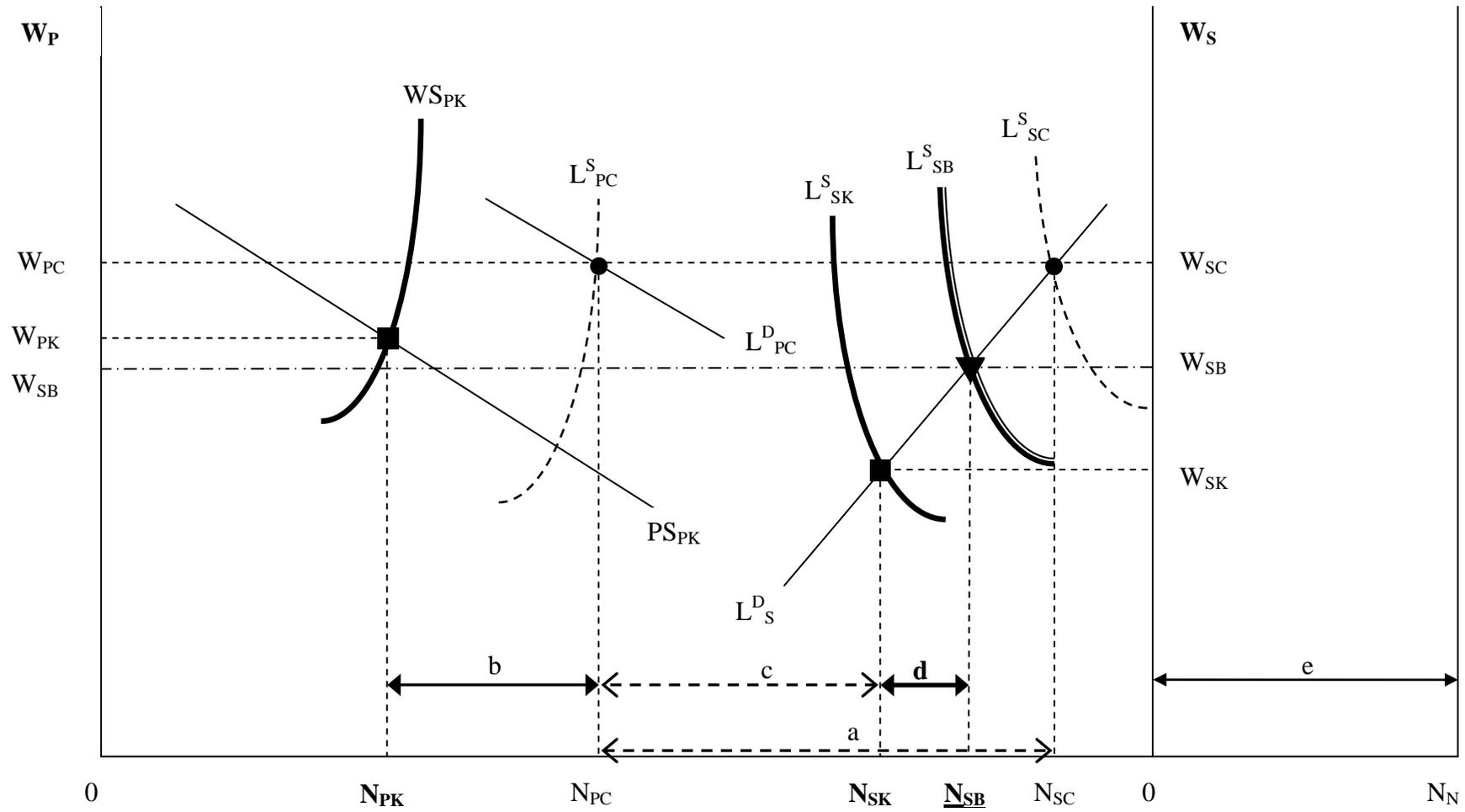
5.2 The mechanics of the three-segment model in comparison to the NCBM and the LNJ model

This subsection develops the three-segment model sequentially with reference to two other ‘benchmark’ models, comparing the unemployment and wage levels that they imply. It does so in an adaptation, but also a deeper exploration, of the Layard *et al* (LNJ) two-sector model presented above. The primary sector mainly represents firms (and unions) that behave largely as described by the New Keynesian models discussed above. The secondary sector does not display such behaviour (simply because its firms may not possess market or face union power). Following Layard *et al*, the secondary sector is defined to include workers who are, or become, self-employed. (Our reason for including them in the secondary sector, and not in the primary sector, is that self-employed workers are unlikely to display efficiency wage behaviour or face unions when setting their own required wage. Other options are possible, of course.)

5.2.1 The impact of New Keynesian conditions in the primary sector on aggregate unemployment: a deeper look at the LNJ model

The first step is to reconstruct the LNJ model to analyse its characteristics more rigorously. As a point of reference, consider a perfectly competitive (NCBM) two-segment model of an economy. The wage level will be the same in both sectors, brought about by perfect labour mobility between sectors. In Figure 6 this reference condition is shown by the intersection of the supply and demand curves for labour in respectively the primary and secondary sectors (L_{PC}^S and L_{PC}^D for the primary sector and L_{SC}^S and L_{SC}^D for the secondary sector). The equilibrium wage rate in this completely competitive model is $W_{PC} = W_{SC}$. At equilibrium, employment in the primary sector will be N_{PC} and in the secondary sector N_{SC} .

Figure 6 – Unemployment in the three-segment ‘barriered’ model



Unemployment will equal the distance represented by the double-headed arrow entitled a = the distance between N_{PC} and N_{SC} . All unemployment is voluntary. (Note that the secondary sector is measured on the horizontal axis from zero on the right-hand side. The length of the horizontal axis indicates the entire labour force, i.e. the economically active. The horizontal axis has been extended to also show those who are not economically active, numbering N_N . They are outside the labour market.)

Next we introduce New Keynesian features to the primary sector, i.e. imperfectly competitive conditions. Accordingly, firms have price-setting power, while either firms or unions have wage-setting power. Therefore, the analysis uses the price-setting (PS) and wage-setting (WS) relationships for the primary sector, as before. In the primary sector labour market equilibrium will occur at the intersection of WS_{PK} and PS_{PK} , at wage W_{PK} and employment level N_{PK} . Compared to the NCBM equilibrium, a number of workers equal to the distance b will not be employed in the primary sector as a result of the presence of imperfectly competitive conditions and New Keynesian behaviour such as efficiency wages.

However, these workers might not be unemployed. They have two other options. They can be economically inactive, meaning that N_N , will be larger than in the NCBM reference case (the double-headed arrow e will lengthen). Or, they can work in the secondary sector, in which case labour supply in the secondary sector will be higher than in perfectly competitive reference case. Hence, labour supply will lie to the left of L^S_{SC} at L^S_{SK} . If all those not employed in the primary sector (due to the introduction of the New Keynesian features) work in the secondary sector, the horizontal displacement from L^S_{SC} to L^S_{SK} will equal distance b . However, because of the slope of L^S_{SK} , employment in the secondary sector will be higher by less than b . Equilibrium in the secondary sector occurs at employment level N_{SK} .

Unemployment in this case (which is the LNJ model) will be higher than in the NCBM case: b plus c in this model will exceed a in the NCBM. The ‘loss’ of employment in the primary sector will exceed the ‘gain’ in employment in the secondary sector.

Note that, although the primary sector is now characterised by imperfectly competitive conditions, the secondary sector still is assumed to be perfectly competitive (in keeping with the LNJ model). There is no price-setting and wage-setting behaviour in this sector.

Since distance b in Figure 6 represents those who are not employed in the primary sector because of its imperfectly competitive conditions, and because these workers would be willing to supply their labour in the absence of these features, they might be termed “involuntarily unemployed in the primary sector”. This is the standard New Keynesian single-sector result.

If these workers do not become economically inactive, they supply their labour services in the secondary sector. Given that the secondary sector is assumed to be perfectly competitive, the residual, unabsorbed group who are unemployed (i.e. b plus c) are “voluntarily unemployed in the secondary sector”. This also leaves them voluntarily unemployed in the final instance. In the aggregate there are no involuntarily unemployed left. This is the basic LNJ two-sector result – a perhaps surprisingly New Classical outcome. It also means that the tertiary segment in this case is empty – the model cannot explain the existence of a tertiary segment, i.e. of persistent involuntary unemployment.

It is to be noted that, even though aggregate unemployment in the LNJ two-sector model is voluntary, the unemployment rate in the model exceeds the unemployment rate that would exist in the absence of the imperfectly competitive conditions (New Keynesian features) of the model.

Also note that some individuals might also decide to leave the labour market altogether and join the economically inactive, thus increasing N_N , thus lengthening e and shifting the secondary sector vertical axis left and reducing c .

5.2.2 The impact of labour market barriers in the secondary sector

The discussion above highlighted several reasons why the secondary sector might be an imperfectly competitive sector. Several barriers to entry might exist for those wishing to work in the secondary sector. Since the secondary sector is defined to include the self-employed, these barriers also imply barriers to self-employment. Entry barriers mean that not all of those who cannot find employment in the primary sector will be able to supply their labour in the secondary sector. This also applies to individuals who have not been employed in either of the sectors, e.g. persons in the tertiary segment.

Thus, due to barriers such as a lack of skills, large distances to the labour market, or lack of resources to finance job search, the *effective* supply of labour in the secondary sector is less than the no-barriers labour supply or ‘*desired*’ supply. In Figure 6 this is represented by a labour supply curve, L_{SB}^S , that lies to the right of L_{SK}^S . Secondary sector employment in the presence of barriers, N_{SB} , will be lower than employment in the absence of the barriers, N_{SK} .

The secondary sector wage in the presence of barriers will be at W_{SB} , higher than W_{SK} in the LNJ model, thereby, ironically, implying a smaller wage differential compared to the LNJ model. The secondary sector equilibrium wage will, however, always be below the primary sector wage.⁵

What about unemployment? First, the unemployment rate in the augmented model is higher than in the no-barriers/LNJ model. In the presence of barriers in the secondary sector unemployment will equal b plus c plus d , compared to only b plus c in the LNJ model. Secondly, in long-run labour market equilibrium a proportion of the unemployed are involuntarily unemployed. While b equals the number of workers who are ‘involuntarily unemployed in the primary sector’, and b plus c represents the number of ‘voluntarily unemployed in the final instance’, d represents the number of those who are, in the aggregate, ‘involuntarily unemployed in the final instance’. In the absence of the barriers these workers would have been able to supply their labour to the secondary sector.

These individuals constitute the tertiary segment, as defined. Thus the barriered model explains the existence of a tertiary segment containing persons who are, in the final instance, involuntarily unemployed persons in long-run labour market equilibrium.

⁵ W_{SB} in Figure 6 has been drawn to be lower than the wage in the primary sector. Should it be higher, one would expect workers employed in the primary sector to shift their supply to the secondary sector until wages are equalised. However, when wages are higher in the primary sector, because of the New Keynesian features introduced in the primary sector, workers cannot shift their supply from the secondary to the primary sector. Hence, wages in the secondary sector will always be lower or equal to wages in the primary sector.

Notes:

- Some individuals might also decide to leave the labour market and become economically inactive, thus increasing N_N , lengthening e , shifting the secondary sector vertical axis left and reducing c , the size of the tertiary segment.
- One can also introduce labour market entry barriers in the primary sector. It would analogously reduce effective labour supply in the primary sector and produce a higher wage and a lower level of employment in equilibrium. More workers would be excluded from this sector; resultant impacts on the secondary and tertiary segments follow logically. Importantly, this also means that there will be more persons involuntarily unemployed in the primary sector. (At least some of them will also be barred (or barred) from the secondary sector, thus pushing them into the tertiary segment.)

5.3 Dynamics in the three-segment model: shocks and hysteresis

The previous subsection compared the NCBM, LNJ and three-segment, “augmented LNJ” models in long-run equilibrium. An important result in the three-segment model with labour market barriers was the presence of higher total unemployment and of involuntary unemployment in long-run equilibrium.

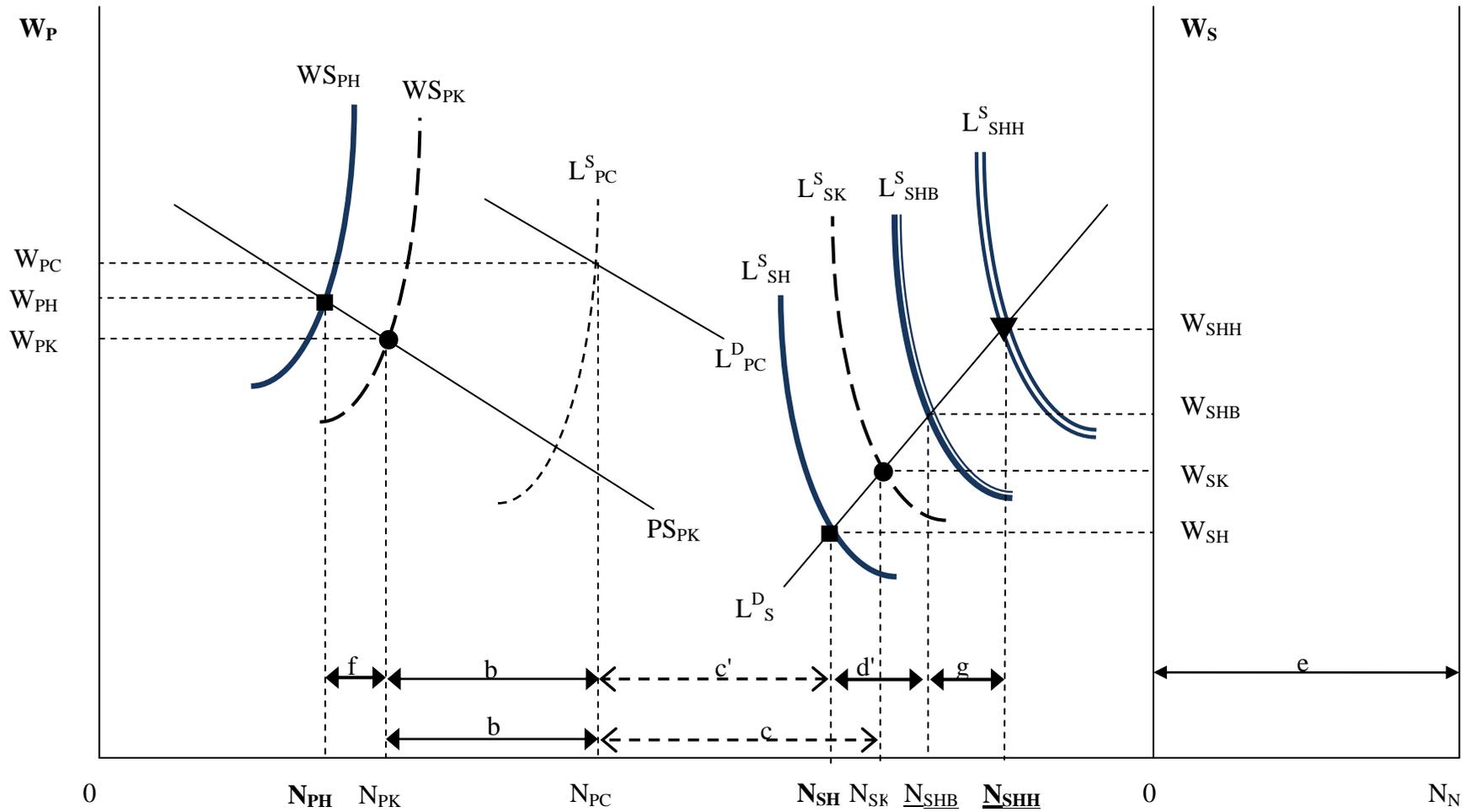
This section discusses the effect of a demand shock and the sequence of events that follow if hysteresis occurs. We first consider the model without labour market barriers.

5.3.1 *Hysteresis but no labour market barriers*

While hysteresis can be introduced to both sectors, we will first introduce it to the primary sector and then to the secondary sector. As before we use the LNJ model as the point of departure. In Figure 7 the model starts with a primary sector equilibrium at the intersection of WS_{PK} and PS_{PK} and a secondary sector equilibrium at the intersection of L^S_{SK} and L^D_S .

Should a demand shock affect the primary sector – contracting output, labour demand and employment in the short to medium run – and should those who become unemployed as a result of the demand shock become unemployable in the primary sector due to e.g. a loss of skills, we have hysteresis: WS moves from WS_{PK} to WS_{PH} . The long-run employment equilibrium (in the primary sector) shifts to a lower level and remains stuck at N_{PH} .

Figure 7 – The three-segment ‘barriered’ model with hysteresis



If these unemployed move their supply of labour to the secondary sector, L^S moves from L^S_{SK} to L^S_{SH} . Whereas unemployment equals b plus c in the LNJ New Keynesian model, it will equal b plus c' plus f in the LNJ New Keynesian model with hysteresis effects. If all those who become unemployed as a result of hysteresis shift their labour supply to the secondary sector (they could also leave the labour market), then the horizontal (leftward) shift from L^S_{SK} to L^S_{SH} will equal f .

The increase in employment in the secondary sector will be less than f (assuming that the supply of labour in the secondary sector has a positive slope). Thus, the reduction in employment in the primary sector due to hysteresis effects is not offset by the increase in employment in the secondary sector.

Therefore, the total number of unemployed in the LNJ-type segmented model with primary-sector imperfectly competitive conditions *and* hysteresis effects (i.e. c' plus b plus f in Figure 7) will exceed the number of total unemployed in the LNJ New Keynesian model (i.e. b plus c in Figures 6 and 7), which is already larger than the number of unemployed in the New Classical reference case (i.e. a in Figure 6). Similarly, involuntary unemployment in the primary sector will equal b plus f , which is higher than in the other cases. In the aggregate, involuntary unemployment ‘in the final instance’ still is zero (because at this stage the secondary sector still is assumed to be market-clearing).

5.3.2 *Hysteresis plus labour market barriers*

If the secondary sector is characterised by barriers to entry, then the effective supply of labour in the secondary sector⁶ will be constrained (say at L^S_{SHB}) and therefore less than the desired supply of labour (L^S_H). Employment in the secondary sector will be at N_{SHB} , meaning that a number of workers equal to d' in Figure 7 will be added to the unemployed already identified (thus totalling d' plus c' plus b plus f in Figure 7). They (d') will be involuntarily unemployed in the secondary sector, and d' then also is the total involuntarily unemployed in

⁶ The barriers will constrain both the labour supply in the secondary sector as such and the absorption of unemployed individuals who flow there from the primary sector due to hysteresis. Effective labour supply L^S_{SHB} is drawn to show the total effect of the barriers, i.e. after some overflow from the primary sector.

the final instance.⁷ This completes the analysis of hysteresis in the primary sector. Note that it affects both sectors.

If (in addition to barriers to entry) hysteresis also occurs in the secondary sector, a demand shock will cause persistent/long-run unemployment in the secondary sector. Unlike the primary sector workers who moved to the secondary sector upon losing their jobs, job-losers in the secondary sector have no other employment sector to move to. Their options are to remain (involuntarily) unemployed (i.e. to enter the tertiary segment) or move out of the labour force. If their continued unemployment causes these workers to lose relevant skills, the labour supply in the secondary sector will decrease and L^S will move from L^S_{SHB} to L^S_{SHH} ; equilibrium employment will be at N_{SHH} . Total as well as involuntary unemployment will increase with a further number of workers equals to g , bringing the total involuntarily unemployed in the final instance to d' plus g in Figure 7. The tertiary segment will be significantly larger than in any previous case. (The number of purely voluntarily unemployed is c' . Recall that b plus f represents those ‘involuntarily unemployed in the primary sector’, but ‘voluntarily employed with regard to the secondary sector’. Thus those in b and f are not part of aggregate involuntary unemployment)

Therefore, if hysteresis is present in both sectors, a demand shock will detrimentally impact on equilibrium employment in both the sectors. The secondary sector is affected twice, first through the spill-over of individuals from the primary sector and secondly through the direct impact of the shock on the secondary sector.

With hysteresis in both sectors, the post-shock long-run equilibrium will exhibit higher total unemployment and higher involuntary unemployment ‘in the final instance’ than without hysteresis.

⁷ If the proportion of workers who desire to supply labour to the secondary sector who also face a barrier to entry is the same in the LNJ and the LNJ-with-hysteresis models, then d' in Figure 7 will be larger than d in Figure 6.

5.4 Main results of the three-segment barriered model

- In a basic segmented model with Post-Keynesian features in the primary sector, in the aggregate there is no persistent involuntary unemployment. (This is the basic LNJ two-sector result.) The model does not explain the existence of a tertiary segment, defined as comprising the involuntarily unemployed.
- If there are labour market barriers in the secondary (and/or primary) sector, the model explains the presence of higher total unemployment and specifically of involuntary unemployment in long-run equilibrium. Thus the barriered model explains the existence of a tertiary segment comprising the involuntarily unemployed.
- Similarly, the presence of hysteresis in either sector (or in both) results in higher unemployment and, specifically, involuntary unemployment ‘in the final instance’ in long-run labour market equilibrium. A tertiary segment is explained.
- The presence of both labour market barriers and hysteresis results in even higher unemployment and, specifically, higher involuntary unemployment ‘in the final instance’ in long-run labour market equilibrium – and a larger tertiary segment than with only one of these phenomena.
- Inter-segment flows and knock-on labour market effects – ‘downward cascading’ of unemployed individuals – determine the size of the tertiary segment.

6. Conclusion: Towards a macroeconomic research agenda

The existence of labour market segmentation is a well-known finding in labour market research, and poverty research frequently highlights labour market marginalisation. Given the interest in the unemployment-poverty nexus, interest has focused on segmentation between the formal sector and informal sector/economy. Therefore, in referring to Layard-type models in the SA context, the primary and secondary sectors routinely are taken to be the formal and informal sectors respectively (e.g. Kingdon and Knight 2004).

This paper has argued for a macroeconomic framework that incorporates the diversity of economic activities ranging from the formal sector to the informal sector *and survivalist activities* and thereby provides a suitable basis for macroeconomic policy in the peculiar South African context. It illustrated such a framework by deriving a novel three-segment model that explicitly incorporates a primary sector (formal sector?) the secondary sector

(informal economy?) and a tertiary (survivalist?) segment as well as labour market entry and mobility barriers.

With respect to long-run labour market equilibrium, the model explains the existence of involuntary unemployment and a tertiary segment in such an equilibrium. Labour market entry and mobility barriers are central to this conclusion, in addition to assumptions of imperfectly competitive product and labour markets (in New Keynesian fashion).

With respect to short-term shocks and dynamics, the three-segment model enables analyses of the impact of cyclical aggregate demand disturbances or aggregate supply shocks on employment conditions in *both* sectors – *and* on the tertiary, last-resort segment. This provides a systematic framework for situating inter-segment transitions and labour-market flows in a segmented macroeconomic model. Phenomena like hysteresis can also be introduced.

More specifically, as this model is developed further, it should be able to assist with analyses of following:

1. Disturbances in either labour market segment impact on employment and total output along differentiated channels – but a net impact on total output (GDP) can be derived.
2. A differentiated analysis of shocks and dynamics in different segments.
3. Aggregate demand or supply shocks can impact differentially on *sectoral* AD or AS, and backwards to different labour segments.
4. Demand shock and hysteresis in the primary sector spill workers over to secondary or tertiary segments, or outside labour force.
5. Depending on the mix (determined by barriers, elasticities, etc.), there will be different impacts on total employment, GDP and unemployment. Note that the shift in secondary sector L^S (increase in effective labour supply) – and thus absorption of labour in secondary sector – depend on the severity of entry and mobility barriers faced by spill-over workers.
6. In turn this impacts on the spill-over to the tertiary segment, or movements out of the labour force.
7. Such a model would allow an analysis of spill-over dynamics with regard to the flows, or transitions, of labour between different segments or livelihood/work states.

8. These inter-segment labour flows – and their determinants and dynamics – can then be modelled.

Simultaneously these aspects begin to constitute an unemployment-oriented macroeconomic research agenda. The possibility of having such a model suggests a research agenda on issues such as the following:

1. The differential impact of demand and supply shocks (macroeconomic cycles) on employment and unemployment in different segments and sectors, and
2. How labour market structural conditions (segmentation; entry barriers) can soften or aggravate this impact.
3. The impact of macroeconomic growth on employment (and unemployment) in different labour market segments and sectors, and
4. The impact of labour market structural conditions (segmentation; entry barriers, wage differentials) on the employment effect of macroeconomic growth.
5. The extent to which labour market structural conditions are a constraint on macroeconomic growth.

In this way one can begin to respond to the following challenges (listed in Fourie 2011:85-6):

1. Can a macroeconomic analysis of employment and unemployment proceed legitimately without engaging with, and incorporating, the information, search and access problems caused by poverty and various segmentations? Or the broader implications of being in a developing country, of dealing with human and societal development?⁸
2. Can an analysis of growth, and constraints on growth, proceed without taking account of the implications of segmentation, poverty conditions and marginalisation for the assumed free flow of labour into a formal sector with flexible labour markets?
3. Can a growth-oriented employment analysis (or a growth strategy) proceed legitimately without dealing with the constrained employment-creation capacity of formal sector growth *and* the intrinsically-linked worlds of informal production and employment, various types of subsistence and survivalist activities and marginalised poor people?
4. Can marginalisation analysts and anti-poverty policy designers – *and* inequality analysts *and* macroeconomists – continue to take little notice of the impact of cyclical or other

⁸ A South African macroeconomics textbook that is explicit in its attempt to integrate some realities of poverty, development, involuntary and structural unemployment into the supply side of its model, is Fourie and Burger (2009).

‘macroeconomic’ shocks and policy steps on the vulnerable and the poor – particularly if there is hysteresis, which could significantly prolong the impact of shocks on employment?

5. Can a poverty-oriented or inequality-oriented analysis of unemployment and wages proceed legitimately without engaging with the relative importance of formal sector growth (and the implications of a negative wage elasticity of the demand for labour)?
6. To what extent can macroeconomic policy measures – e.g. interest rate and exchange rate policy – shoulder the burden of explaining and resolving unemployment)?

Other issues that need to be tackled are:

1. The proper treatment of the self-employed, notably outside the formal sector/professions, remains a major analytical gap and inconsistency in the various models. It reflects the failure to deal with those that are unemployed (but obviously still stay alive in some way...)
2. New theory is needed on price setting/formation and wage setting/formation in the secondary sector under various conditions.
3. Efficiency wages present a potentially important way to incorporate worker motivation and productivity *also at the lower end of wage spectrum*. Another question would be whether there could be something like “inefficiency wages”, i.e. wages that are set too low and induce inefficiency and low labour productivity?
4. Linked to this is the question whether the fairness model may be a potentially important way to incorporate perceived fairness and entitlement as determinants of South African labour market behaviour.
5. Is there a substantive and important segmentation within the formal sector, e.g. between a corporatist and a market sector – and what could be its macroeconomic and employment implications?

* * *

Bibliography

- Blanchard, O. 2005. *European unemployment: The evolution of facts and ideas*. Working Paper 11750, National Bureau of Economic Research, Cambridge, Mass.
- Cahuc, P and Zylberberg, A. 2004. *Labor Economics*. MIT Press.
- Carlin, W and Soskice, D.2005. *Advanced Macroeconomics*. Oxford University Press.
- Fourie, F.C.v.N. 2011. *The South African unemployment debate: three worlds, three discourses?* Working Paper 63 (June), SALDRU, University of Cape Town.
- Fourie, F.C.v.N. and Burger, P. 2009. *How to think and reason in macroeconomics*. 3rd edition. Cape Town: Juta.
- Kingdon, G.G. and Knight, J.B. 2004. Unemployment in South Africa: the nature of the beast. *World Development*, 32(3): 391-408.
- Layard, R., Nickell, S and Jackman, R. 1991 and 2005. *Unemployment: macroeconomic performance and the labour market*. Oxford: Oxford University Press.
- Romer, D. 2012. *Advanced Macroeconomics*. 4th ed. McGraw-Hill Irwin. New York.