South African Reserve Bank Working Paper Series WP/23/04

Investigating unemployment hysteresis in South Africa

Nicola Viegi and Vincent Dadam

Authorised for publication by Konstantin Makrelov

7 June 2023

South African Reserve Bank

South African Reserve Bank Working Papers are written by staff members of the South African Reserve Bank and, on occasion, by consultants under the auspices of the South African Reserve Bank. The papers deal with topical issues and describe preliminary research findings, and develop new analytical or empirical approaches in their analyses. They are solely intended to elicit comments and stimulate debate.

South African Reserve Bank Working Papers are externally refereed.

Information on South African Reserve Bank Working Papers can be found at
https://www.resbank.co.za/en/home/publications/Papers/working-papers

Enquiries relating to the Working Paper Series can be addressed to:
Head: Economic Research Department
South African Reserve Bank
P O Box 427
Pretoria 0001

Tel. +27 12 313 3911

Investigating unemployment hysteresis in South Africa

Vincent Dadam* and Nicola Viegi†

Abstract

This paper investigates hysteresis in South Africa's unemployment. First we test the presence of hysteresis in unemployment using traditional stationarity tests and non-linear transformation methods to identify two further characteristics of hysteresis, namely remanence and selective memory. In the second part of the paper we estimate a simple insider-outsider model using a Bayesian vector autoregression methodology to identify the shocks driving unemployment dynamics. The main finding is that mark-up shocks and negative productivity shocks are the main drivers of unemployment, with demand shocks playing a secondary role. Nominal wages are not responsive to real shocks and are an important component of inflation. These results point to the difficulty of absorbing the current level of unemployment without a significant increase in the flexibility of goods and labour markets. At the same time, the evidence suggests that, if reforms are being implemented, demand policies can play a significant role in improving employment and growth, reversing the structural unemployment evident in the data.

JEL classification

E24, E32, J64

Key words

Unemployment, hysteresis, nominal wages, mark-up, insider-outsider dynamics.

---

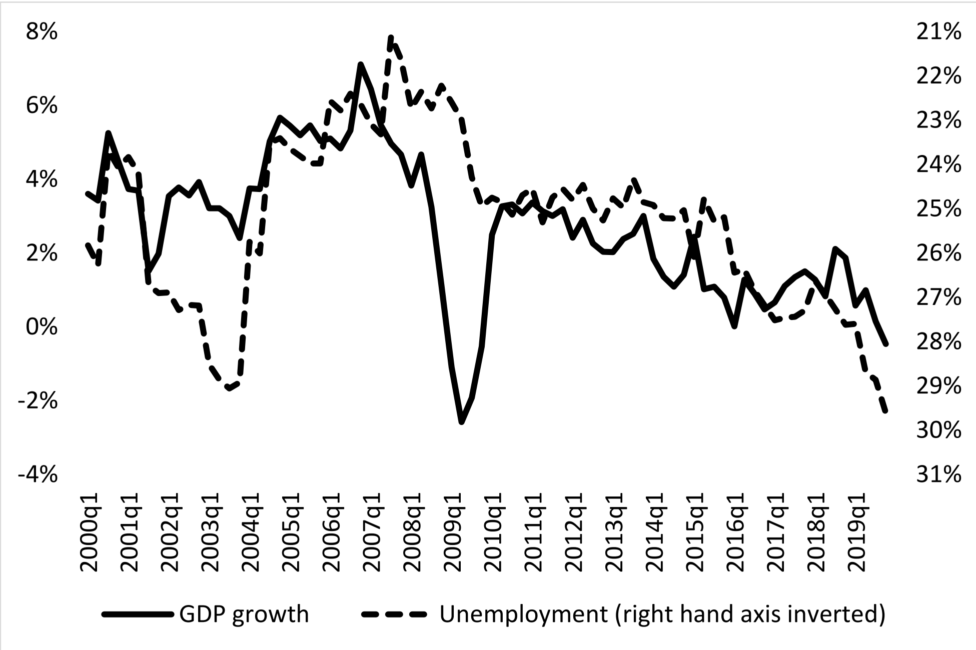
* Corresponding author, University of Pretoria and South African Modelling Network, vincent.dadam@up.ac.za.

† University of Pretoria and Economic Research Southern Africa, nicola.viegi@up.ac.za.

1. Introduction

The unemployment rate in South Africa has been on an upward trend for the last 10 years, going from 21.1% just before the global financial crisis in 2007 to 29.6% just before the start of the COVID-19 pandemic at the end of 2019.[1] In the 2000–2019 period, the pattern of unemployment follows quite closely that of gross domestic product (GDP) growth – a consequence of the economy's inability to absorb newcomers in the labour market, combined with slow growth performances. Figure 1 shows this correlation, with the scale for unemployment inverted in order to make the common negative trend after the 2008 global financial crisis more evident.

Figure 1: South Africa's unemployment and GDP growth 2000–2019



This dynamic hardly conforms to standard models of a macroeconomic cycle as stochastic fluctuations around a stable long-term equilibrium growth path (Woodford 2003), which is the theoretical framework behind modern monetary policy. More likely, we are observing an economy where stochastic short-run shocks have long-term effects on the equilibrium values of unemployment and potential output. The economy shows hysteresis, that is, the equilibrium rates are history dependent (Blanchard and

1 The COVID-19 pandemic shock has dramatically accelerated the trend of growing unemployment, with the official unemployment rate reaching 35.3% in the fourth quarter of 2021. While we do not include this shock in our analysis, our conclusions can help in understanding some of the causes of the slow recovery from the shock itself.

Summers 1987). Therefore a rise in unemployment following a contractionary shock leads to an increase in the underlying equilibrium unemployment (O'Shaughnessy 2011). This has fundamental implications for monetary policy because the Phillips curve specification used in most monetary policy models will poorly capture the underlying economic structure. In fact, the Phillips curve will appear 'flat' in the sense that the link between the output gap and inflation will most likely be insignificant. Subsequent increases in the demand for labour will generate inflationary wage and price increases before unemployment returns to pre-shock levels.

Given the implications of this hypothesis, this paper aims to formally identify hysteresis in South Africa's unemployment data. Once we have determined that unemployment is characterised by strong hysteresis, we use a Bayesian vector autoregression (VAR) methodology to identify the causes of this hysteresis. We focus on three main determinants: shocks to nominal wages, following Blanchard and Summers (1987), shocks to mark-up by firms (Gambetti and Pistoresi 2004) and shocks to the real effective exchange rate as a proxy for changes in costs of intermediate and investment goods affecting firms' decisions (Darby et al. 1999). We also account for monetary policy shocks in our robustness checks through the interest rate.

The term hysteresis has a long history in economics and became widely applied in the context of explaining the persistence of European unemployment in the 1980s (Blanchard and Summers 1987; Ball 2009). The concept has found new life after the 2008 global financial crisis to explain the decade-long persistence of economic stagnation in industrialised countries (Gali 2015, 2020; Garga and Singh 2021). The focus of the literature is identifying the microeconomic mechanism behind the aggregate hysteresis effects. Gali (2015), for example, focuses on the role of labour market bargaining institutions and wage rigidities in a New Keynesian macroeconomic model, but this is just one of the many possible mechanisms that can permanently 'scar' an economy.

Cerra, Fatas and Saxena (2020) review the current debate and evidence and point to two main mechanisms. One is a labour market mechanism, where temporary shocks can have permanent effects because of the interaction of unemployment with labour market institutions (Layard, Nickell and Jackman 1991; Dolado and Jimeno 1997) or because of the effect of shocks on skill accumulation and depreciation (Acharya et al. 2022). In Dadam and Viegi (2015), we discuss the possible relevance of this channel for the South African economy, by showing that wages do not respond significantly to real economic conditions. This partly explains the big losses in employment following the 2008 global financial crisis and the slowness of the recovery. An alternative mechanism is when hysteresis appears because of the effect of shocks on investment and innovations (Darby et al. 1999), and the entry/exit of innovative firms that affect the long-run growth of the economy (Dosi et al. 2018; Decker et al. 2016). Kerr, Wittenberg and Arrow (2014) show that in the South African context, firm deaths constitute about 25% of job destruction while firm births account for a relatively small 11% of job creation. These results confirm earlier findings in Davis, Haltiwanger and Schuh (1996) which indicate that high rates of job destruction combined with steadily low rates of job creation may constitute a long-term source of hysteresis in South African unemployment.

The rest of the paper is organised as follows. Section 2 defines the concept of hysteresis and tests for its presence in the South African unemployment series. Section 3 investigates the causes and consequences of hysteresis in unemployment by setting up a simple model of a labour market with insider-outsider dynamics and hysteresis. We estimate the model using Bayesian VAR and then discuss the results. Section 4 presents our closing observations.

2. Hysteresis in unemployment: definition and evidence

The term hysteresis is widely used in economics to cover different concepts of persistence in economic dynamics. For example, Gali (2020) interprets hysteresis as a long-lasting deviation of unemployment from its natural rate that is compatible with 'flexible wages', while Garga and Singh (2021) interpret it as a permanent change in potential output, that is, a unit root in the underlying equilibrium values. Both approaches try to mimic the unit root in unemployment or economic growth observed for many countries.

The first test for hysteresis is then a test of persistence in the series, that is, a simple unit root test. In Table 1 we report the results of both the Augmented Dickey-Fuller

(ADF) and Phillips-Perron tests for a unit root on the South African unemployment series. The tests cannot reject the hypothesis of a unit root at any confidence level, a result confirmed by the other available unit root tests.

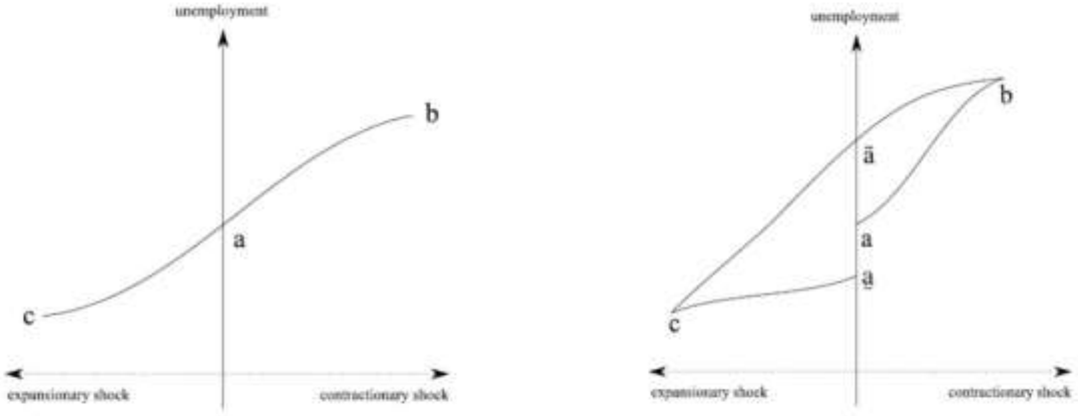
Table 1: Unit root tests on South Africa's unemployment 2000Q1–2019Q4

Null hypothesis: unemployment has a unit root	Adj. t-stat	Prob.*
Augmented Dickey-Fuller test statistic	-1.384906	0.5856
Phillips-Peron test statistic	-1.457863	0.5497

This is the same result found in Pikoko and Phiri (2019) for South Africa and in most of the literature reviewed by Cerra, Fatas and Saxena (2020) looking at hysteresis in most Organisation for Economic Co-operation and Development (OECD) countries. While this simple concept of hysteresis is widely applied in the literature, the unit root approach has been criticised because it does not capture three properties of hysteresis, namely *non-linearity*, *selective memory* (only non-dominated past shocks are remembered by the system), and *remanence* (temporary shocks permanently change the equilibrium of the system) (Hallett and Piscitelli 2002; Amable et al. 2004; Bassi and Land 2016).

To illustrate the difference, Figure 2 compares what would be a typical dynamic response to a shock in a natural rate model versus a model with hysteresis. In panel (a), a contractionary shock moves unemployment from point *a* to point *b*: if the model is stationary, the shock is absorbed and unemployment goes back to point *a* after a certain time, depending on persistence. In a unit root process, unemployment would remain at point *b* until an equal shock in the opposite direction brings the system back to *a*. Expansionary and contractionary shocks will have the same dynamic effect except for a change in sign, from point *a* to point *c*.

Figure 2: Shocks and unemployment in natural rate versus hysteresis models



(a) Shocks in a Natural Rate Model

(b) Shocks with Hysteresis

In contrast, panel (b) shows the dynamics of unemployment in a hysteresis model. After the contractionary shock moving unemployment from *a* to *b*, the shock has changed the properties of the system, so that the new equilibrium is going to be $\bar{a}$, with a permanent increase in the 'equilibrium' unemployment. Notice that any shock can generate these dynamics: in particular, strong demand shocks will have a permanent effect as much as supply shocks. Also, contractionary and expansionary shocks can have very different effects, depending on the underlying economic and institutional structure. An expansionary effect can move the system from $\bar{a}$ to $\underline{a}$, with a much lower equilibrium unemployment rate, depending on the way the system reacts to the specific shock.

It is therefore important to account for the characteristics of hysteresis that take its definition beyond the barriers of a simple unit root property. As such, the *remanence* feature of hysteresis is particularly relevant. This feature is emphasised in the Preisach model of hysteresis in electromagnetism (Mayergoyz 1986) to more generally define a non-linear input-output system with memory. The emphasis is on the non-linear characteristic of the process, relative to the linear unit root process widely used in economics. For example, a large shock will have proportionally more influence than a small shock in the following system dynamics.

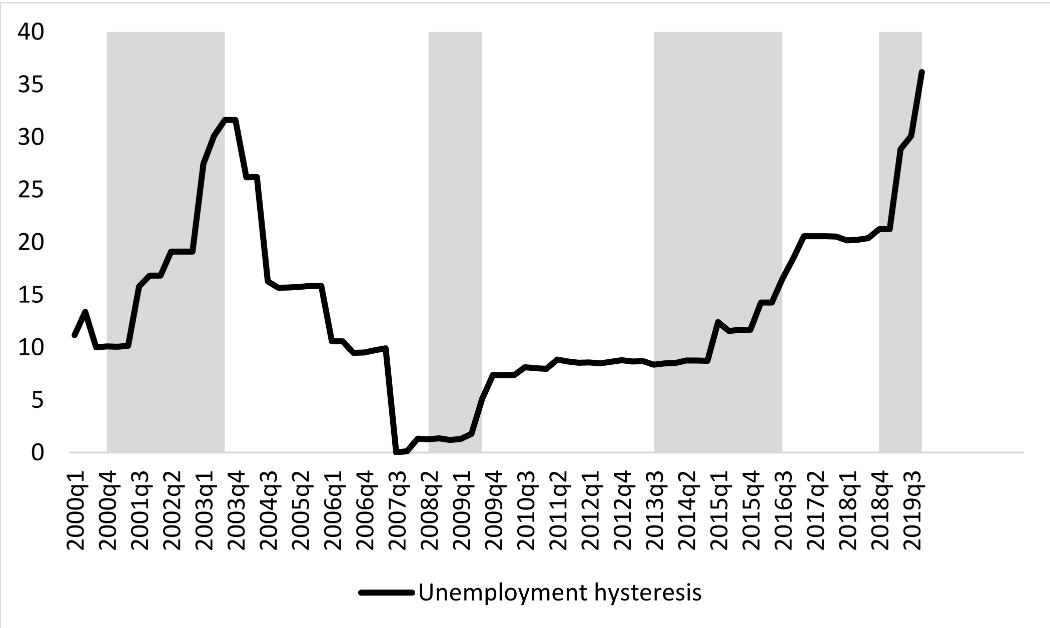
Selective memory refers to the other property of hysteresis that only non-dominated past shocks affect the current equilibrium of the system. Contrary to unit root processes, in hysteresis models only important shocks are remembered. Therefore, the current equilibrium of the system is the result of an accumulation of past salient shocks, and identifying these shocks can provide a narrative of the important events that have produced the present state of the economy.

In economics, Piscitelli et al. (2000) develop a test for strong hysteresis by identifying the dominant shocks in a series and calculating a non-linear transformation of the series where each shock is weighted for its degree of remanence (that is, how much the shock is remembered in the series history). The test is based on the Preisach model of hysteresis (Cross 1994; Göcke 2002), where the output of a system (in our case unemployment) responds in a non-linear way to input shocks because the micro units or relays (in our case, the firms) respond differently to a common shock. These two requirements encompass the essence of a hysteretic system that has a selective memory – a system with erasable memory where only the non-dominated extremum values of shocks in the past can impact the current output of the system.

To illustrate the framework following Cross (1994), assume that aggregate employment *N* is generated by the aggregation of individual firms' decisions that vary according to a price level (the shock variable which can be any exogenous process affecting the profitability of the firm). This price level determines the entry and exit of firms in the market. Therefore, the price level acts as a shock to the employment series and also as a barrier of entry/exit to individual firms in a given market. Firms will enter the market if the price level is above a certain threshold $\beta_{,j}$ and will exit the market when the price is below $\alpha_{,j}$, with $\alpha_{,j} < \beta_{,j}$, that is, each firm having different thresholds (for example because of different productivity levels, as in Melitz (2003)). When the price level is between $\alpha_{,j}$ and $\beta_{,j}$, the firm will be inactive, either remaining in the market or remaining outside depending on what its position was in the previous period, like in the model of investment under uncertainty of Dixit and Pindyck (1994). Each firm is therefore defined by the pair of switching values to the evolution of the aggregate price level and the aggregate price level determines which firms exit, enter or remain in the market. Thus, the switching values are defined as exit and entry triggers such that $\alpha_j = P_{exit,j}$ and $\beta_j = P_{entry,j}$. The number of active firms at a certain price level determines the dynamics of aggregate employment. After a large shock to the price level, the total number of firms in the market will change while small shocks will not change the number as more firms will be in the 'neutral' zone where decisions about entry and exit are postponed. This heterogeneity at the micro level is enough to generate a non-linear aggregate system where employment has selective memory of past shocks, that is, the system has remanence.[2]

To determine if South Africa's unemployment series has these characteristics, we apply the test in Piscitelli et al. (2000). Figure 3 displays the results, with the shaded area denoting the OECD South Africa recession indicator which is in fact well picked up by the index.

Figure 3: Hysteresis transformation of unemployment and OECD recession indicator (shaded area)



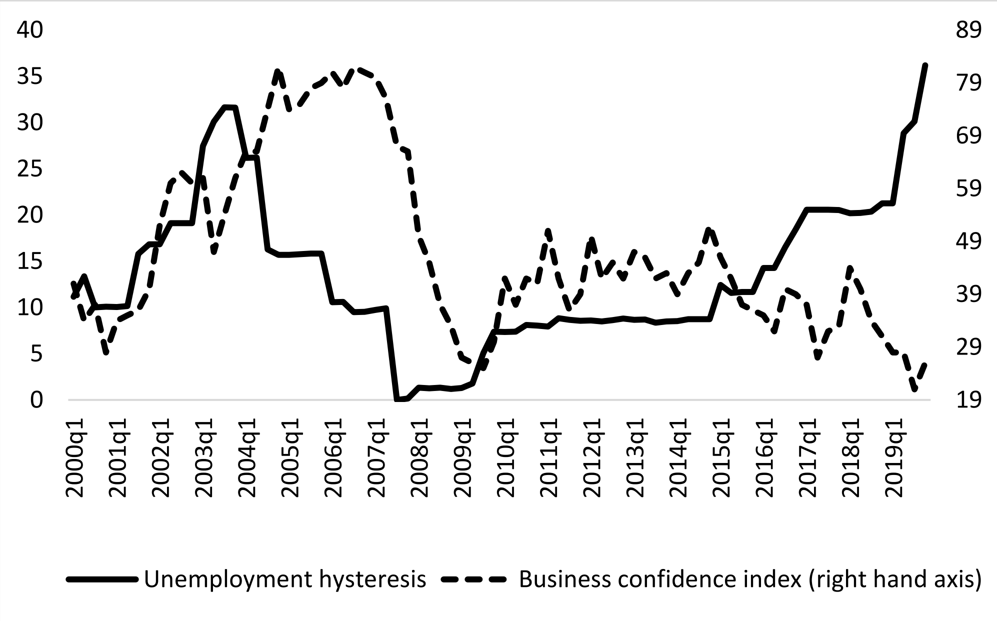
The hysteresis transformation emphasises shocks that are not locally dominated, so that the series selectively remembers shocks that were relevant in changing the 'equilibrium' unemployment rate. For example, unemployment was certainly affected by the 2008 global financial crisis but it seems to pick up considerably only after 2014, when strong fiscal support after the 2008 crisis reached its limits and the economy

2 See Appendix A for further details on the dynamics in a Preisach model of hysteresis.

entered the first post-crisis recession that has resulted in economic stagnation since then.

We now compare the hysteresis transformation with a measure of business confidence in South Africa compiled by the Bureau for Economic Research. The goal here is to investigate the response of unemployment to domestic shocks. Figure 4 presents these two series for the whole sample of interest.

Figure 4: Hysteresis transformation of unemployment and business confidence index

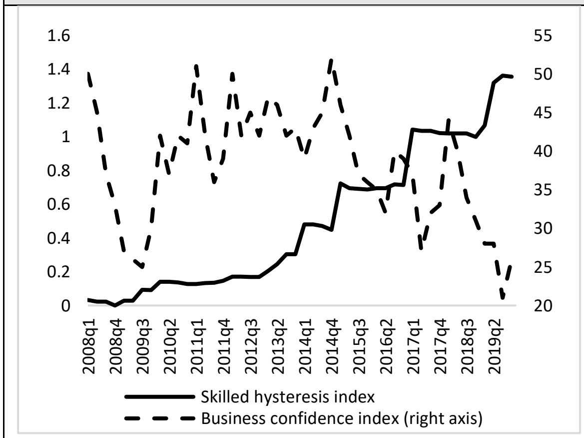


In the first part of the sample there is a positive correlation between the business confidence index and the hysteresis transformation of unemployment. This can mainly be explained by the labour force expanding at the end of the apartheid era; however, this labour force was predominantly unskilled and therefore couldn’t be absorbed right away. In contrast, the end of apartheid had a positive impact on business sentiment. The hysteresis transformation then declines as business confidence peaks. This is followed by a co-movement between the two indicators during the 2008 global financial crisis that is consistent with our expectations. Moreover, the plateau noticed in the hysteresis transformation between 2009Q4 and 2014Q4 aligns almost perfectly with the behaviour of the business confidence index. The year 2015 in South Africa was marked by significant social and political turmoil that began with the surprise dismissal of the Minister of Finance, which acutely and negatively impacted economic growth prospects and investor and business confidence. The persistent and negative correlation between the two variables that begins in 2016 remarkably captures these events, as evident in Figure 4.

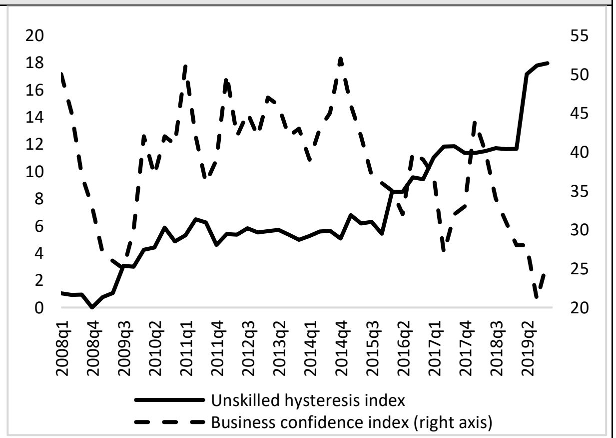
To assess the dynamics of unemployment further, we consider the skill gap among the unemployed. In particular, we calculate a disaggregated hysteresis transformation to define indices for skilled and unskilled unemployment rates. As a proxy for the unskilled, we consider among the unemployed those individuals who have completed high school, those who have dropped out of high school and those who have no schooling. The skilled unemployed are individuals with tertiary level education. Because we are using the Quarterly Labour Force Survey as the data source, the data is constrained to begin in 2008Q1. The results are reported in Figure 5.

Figure 5: Hysteresis transformation of unemployment – skilled versus unskilled

(a) Skilled



(b) Unskilled



Consistent with our expectations, we find that the hysteresis phenomenon is more pronounced among the unskilled. In the beginning of the sample period, the 2008 global financial crisis shock seemed to have had a bigger impact on unskilled unemployment, evident in the jump in the index as the series reaches a higher plateau. The skilled hysteresis index however shows that the shock did not significantly affect the skilled unemployed. In contrast, the deterioration of business conditions in 2015 contributed to an increase in unemployment for both skilled and unskilled, with hysteretical behaviour. This interesting finding shows that adverse domestic shocks that have recently dampened economic prospects have left both the skilled and unskilled struggling to find jobs, therefore highlighting the slowness of the recovery following a shock in South Africa. It is important to highlight that the hysteretical effect remains more pronounced with the unskilled unemployed group.

We conclude our test for strong hysteresis with a simple forecasting exercise in Table 2 to compare the performances of an autoregressive specification with the generated hysteresis index in forecasting unemployment. We find that an estimation of unemployment using the lagged hysteresis transformation, as opposed to lagged unemployment as dictated by a unit root process, provides a better out-of-sample forecast by more than halving the root mean squared error (RMSE) of the forecast.

Table 2: Hysteresis ordinary least squares test

Variable	Coefficient	Probability	Coefficient	Probability
Unemployment rate (-1)	0.927	0.00		
GDP growth	-0.116	0.02	-0.384	0.00
Inflation	0.079	0.03	0.22	0.00
Hysteresis index (-1)	-	-	0.22	0.00
Constant	1.812	0.18	22.16	0.00
RMSE (out of sample forecast sample 2018Q1–2019Q4)	1.135		0.54	

The findings are compelling and suggest a better understanding of the dynamics that generate heterogeneous responses of unemployment to various shocks. The second part of this paper focuses on answering the following question: what variables are most likely to generate shocks that may induce a hysteretical response in unemployment? We are interested in evaluating the importance of shocks often discussed in the literature – that is, demand and nominal wages shocks (Blanchard and Summers 1987), shocks to mark-up (Gambetti and Pistoresi 2004) and real effective exchange rate shocks (Darby et al*.* 1999). We also include monetary policy shocks as a robustness check.

3. Hysteresis in unemployment: causes and consequences

3.1 The model

To shed some light on the drivers of unemployment dynamics in South Africa, we use a similar framework to that used in Maidorn (2003) and Gambetti and Pistoresi (2004) where wage bargaining and productivity developments produce hysteresis in the labour and product markets, with big shocks to nominal wages and productivity having long-term effects on unemployment and GDP growth. This allows us to define specific dynamics between the variables considered in this framework in order to then identify various shocks.

The model assumes imperfect competition in both product and labour markets à la Nickell (1988). Each firm *i* uses labour and technology at time *t* in a production function with constant returns to scale of the following form:

$$Y_{it} = A_{it}N_{it} \tag{1}$$

where N_{it} employment in firm *i* at time *t*, Y_{it} is the firm's output and A_{it} is a random labour-augmented technology that is known at time *t*. Real aggregate demand is exogenously determined at the beginning of each period. We define P_{it} as the output price, which allows us to write the demand function for good *i* that takes the form of a typical Dixit-Stiglitz manner

$$Y_{it}^{d} = \omega_i \left(\frac{P_{it}}{P_t}\right)^{-\phi} D_t \tag{2}$$

Here, $\phi > 0$ is the elasticity of demand, P_t denotes the price index of aggregate output, and ω_i represents the share of total nominal aggregate demand for firm *i* when $P_{it} = P_t$. $\sum \omega_i = 1$. $P_{it}$is the monopoly price and is set as a mark-up over marginal cost, that is:

$$P_{it} = \frac{\phi}{\phi - 1}\frac{W_{it}}{A_{it}} \tag{3}$$

where W_{it} represents the known wage rate at time$t$ ￼.

The aggregate values are derived assuming homogenous goods, technology and uniform wages. This means $A_{it} = A_t$, $N_{it} = N_t$ and $W_{it} = W_t$. Therefore the aggregate relations are:

$$Y_t = A_t N_t \tag{4}$$

$$Y_t = D_t^{\phi} \tag{5}$$

$$\frac{P_t}{W_t} = \frac{1}{A_t}\frac{\phi}{\phi - 1} \tag{6}$$

In log linear form, (4), (5) and (6) are as follows:

$$y_t = n_t + a_t \tag{7}$$

$$y_t = \phi d_t \tag{8}$$

$$p_t = w_t - a_t + \mu_t \tag{9}$$

where μ_t is the representation of price shocks. The system of equations (7), (8) and (9) represent the goods market in the economy. Technology, demand and mark-up are all exogenously determined random walk processes. Therefore, we can write:

$$\Delta a_t = \epsilon_{st}$$

$$\Delta d_t = \epsilon_{dt}$$

$$\Delta \mu_t = \epsilon_{pt}$$

in which ϵ_{st}, ϵ_{dt} and ϵ_{pt} are respectively independent and identically distributed (*i.i.d.)* uncorrelated aggregate productivity, demand and price shocks.

We now move on to define the labour market component of the model. This follows the formalism first introduced by Dolado and Jimeno (1997). The labour force evolves in log terms according to the following:

$$l_t = u_t + n_t \tag{10}$$

where l_t is the labour force and u_t denotes unemployment. We can also define the labour force as follows:

$$l_t = \alpha(w_t - p_t) - bu_t + \tau_t \tag{11}$$

in which α and b are constant parameters and τ_t denotes a labour supply shift factor that captures changes in the participation rate and population growth. τ_t follows a random walk in a manner similar to a_t, d_t and μ_t. As such, we may write:

$$\Delta\tau_t = \epsilon_{lt}$$

with ϵ_{lt} and *i.i.d.* uncorrelated labour shock. We assume an insider-outsider framework with hysteresis in which a targeted nominal wage w_t^* determines the actual nominal wage. In particular:

$$w_t = w_t^* + \epsilon_{wt} + \gamma_1\epsilon_{dt} + \gamma_2\epsilon_{pt} \qquad (12)$$

$$w_t^* = arg\{n_t^e = (1-\lambda)n_{t-1} + \lambda l_{t-1}\} \qquad (13)$$

where n_t^e is the expected employment which evolves according to the level of hysteresis prevailing in the economy, $\lambda \in [0,1]$ denotes the hysteresis parameter, ϵ_{wt} is an *i.i.d.* shock to wages which also reflects the bargaining power of unions, and γ_1 and γ_2 are constant parameters.

The current level of nominal wages is determined by the previous period, which therefore suggests that the expected employment level is dependent on the previous period weighted average of the labour force l_{t-1} and employment n_{t-1}. Two scenarios are considered in determining the nominal wage:

- If $0 < \lambda < 1$, the unions bargain a wage such that the expected employment level n_t^e is larger than the employment in the previous period n_{t-1}, therefore increasing the size of the workforce.
- If $\lambda = 0$, full hysteresis prevails in the economy. In this scenario, the segmentation of the labour market between insiders and outsiders emphasises the dominant position of the former over the latter in the determination of the nominal wage. Simply put, the insiders decide the nominal wage that ensures their employability, with virtually no weight given to the unemployed in the wage-bargaining process.

Assuming full hysteresis, we can express the model as a moving average representation in first differences:[3]

$$\Delta y_t = \phi \epsilon_{dt} \tag{14}$$

$$\Delta n_t = \phi \epsilon_{dt} - \epsilon_{st} \tag{15}$$

$$\Delta w_t = \gamma_1 \epsilon_{dt} + \epsilon_{wt} + \gamma_2 \epsilon_{pt} \tag{16}$$

$$\Delta p_t = \gamma_1 \epsilon_{dt} - \epsilon_{st} + \epsilon_{wt} + (1 + \gamma_2)\epsilon_{pt} \tag{17}$$

$$\Delta u_t = \frac{1}{1-b}\left[\phi \epsilon_{dt} + (1+\alpha)\epsilon_{st} - \alpha \epsilon_{pt} + \epsilon_{lt}\right] \tag{18}$$

This model is the base for our empirical estimation in the next section. Equation (14) suggests that demand shocks drive output; from equation (15) we derive that productivity shocks drive employment, with increasing productivity reducing the demand for labour; from (16), wages are driven by an exogenous process of wage formation and respond to increases in aggregate demand and increases in prices. Equation (17) is a classical Phillips curve, where inflation is driven by shocks to demand, to wages and to supply, plus an independent mark-up shock ϵ_{pt}. Finally, from (18) we derive labour supply shocks ϵ_{lt}.

3.2 Data and estimation

In this section we take the model to the data using a Bayesian structural VAR. We use quarterly data from 2000Q1 to 2019Q4. Employment, real GDP, nominal wages, prices and unemployment are all transformed in logs. We end the sample at 2019Q4 because the data thereafter include the extremely large COVID-19 shock that would have dominated the estimation unless we dampened its effect in a somehow arbitrary way. Besides the consumer price index and real GDP obtained from Statistics South Africa, the source of the remaining variables is the South African Reserve Bank (SARB). Specifically, we use the data from the Quarterly Projection Model (QPM), the SARB’s main forecasting tool. This provides us with a coherent unemployment series for the whole of the estimation period.

---

3 See Appendix B.

We estimate a Bayesian VAR model with two lags with the data in annual log differences, so that everything is scaled in term of annual growth rates. We use Normal-Wishart priors with hyper-parameters as in Giannone, Lenza and Primiceri (2015), searched on a grid to maximise the posterior probability distribution conditional on the data. To identify the shocks, we use the factorisation suggested in equations (14) to (18). We also experiment with classical diagonal factorisation and with different ordering of the variables without significant changes in the results.

We use the system of equations (14)–(18) as our baseline specification before adding endogenous and exogenous variables to the model such as the real effective exchange rate and the monetary policy instrument.

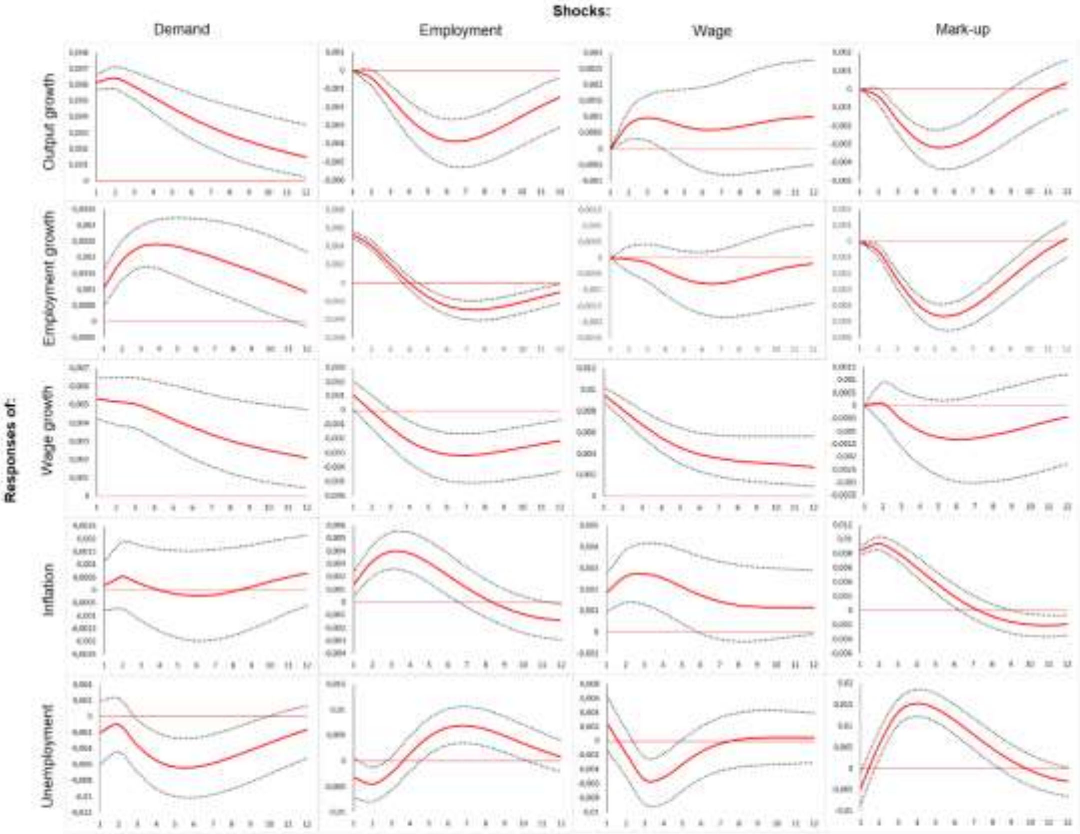
3.3 Estimation results

3.3.1 Baseline

In Figure 6 we show the impulse response functions of the Bayesian VAR containing the five variables in the model (14)–(18) above. The shocks we consider are: (i) demand shocks, such as fiscal and monetary policy and international demand for domestic goods; (ii) employment shocks, which, from equation (15), are a proxy for productivity shocks including fluctuations in commodity prices, electricity constraints common after 2008 and any other change in the production process that changes the demand for labour; (iii) nominal wage shocks, which reflect exogenous changes in nominal wages not connected to other shocks, such as the change in wages caused by wage negotiations rather than changes in the inflation rate or productivity; and (iv) mark-up shocks, which reflect changes in prices reflecting the monopoly power of firms and their ability to leverage their market power to increase their profit rates.[4]

[4] Thakoor (2020) finds that South Africa's mark-up has gone up by 25% since 2000, with levels higher than emerging and developing countries in Asia. Since 2008, these levels have been higher than sub-Saharan African peers. Interestingly, sectors exposed to external competition such as the manufacturing sector were not exempt from these high levels between 2010 and 2012.

Figure 6: Impulse responses – baseline



The impulse responses follow expected patterns. Demand shocks have an expansionary effect with relatively little inflationary effect. The most inflationary shocks are employment (or negative supply shocks) and wage shocks, with both putting pressure on wages and prices and negatively affecting employment and growth. The mark-up shock has the strongest negative effect on the real side of the economy and strong inflationary effects on prices. Unemployment is clearly more sensitive to mark-up shocks. From the Bayesian VAR we can also estimate the implied steady state value of the growth variables in the system, which gives the results shown in Table 3.

Table 3: Steady state growth of variables

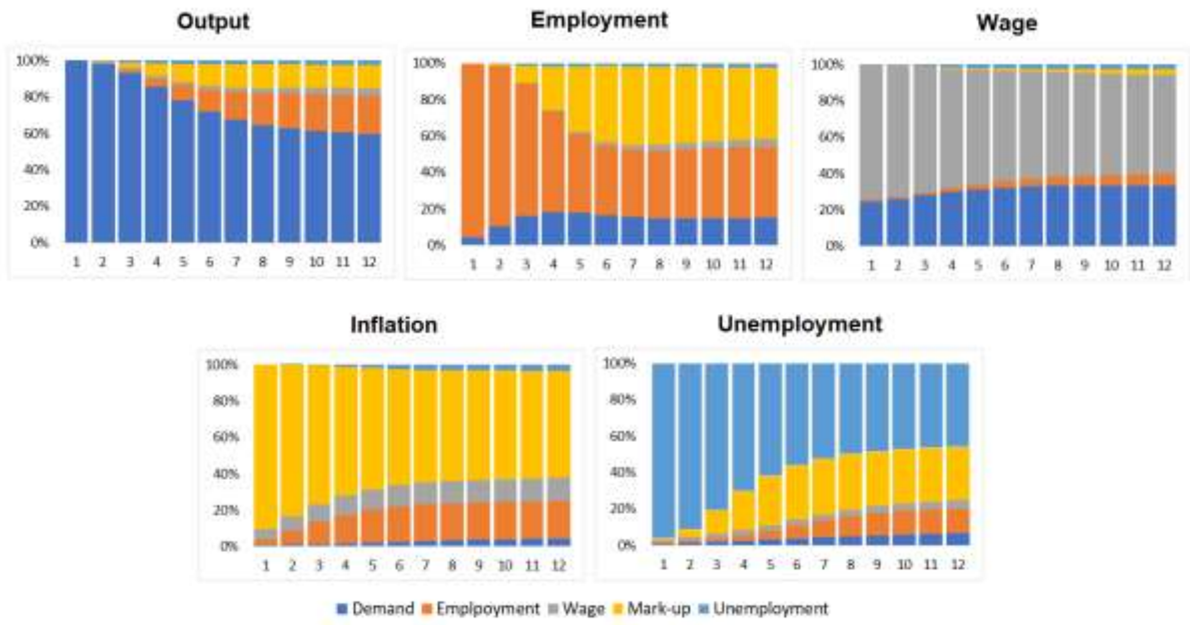
Variables	Steady state
Output growth	2.3%
Employment growth	0.9%
Nominal wage growth	0.9%
Inflation	4.9%
Unemployment growth	1.16%

Clearly, the data imply a growth in unemployment that is above that of employment, and therefore a continuous expansion of the unemployment rate. Further, nominal wages increase faster than prices, suggesting a sustained increase in real wages, above productivity growth (the difference between growth in output and growth in employment).

These findings are in line with our expectations and have long been debated in the South African literature. The structural nature of unemployment is founded upon three main pillars: nominal wages, prices and employment. An increase in wages due to the high bargaining power of labour unions reduces the employable pool while keeping the insiders safe. As a response, firms pass on this increase to the workers through a rise in prices given firms' monopolistic structure in South Africa, which inevitably translates into high and persistent unemployment. This hysteresis effect is further emphasised by the considerable skill gap prevailing in the labour market.

The forecast error variance decomposition (FEVD) consolidates these results. This exercise investigates the contributions of all considered shocks to a one-step forecast error variance of each variable. We report the results in Figure 7.

Figure 7: Forecast error variance decomposition



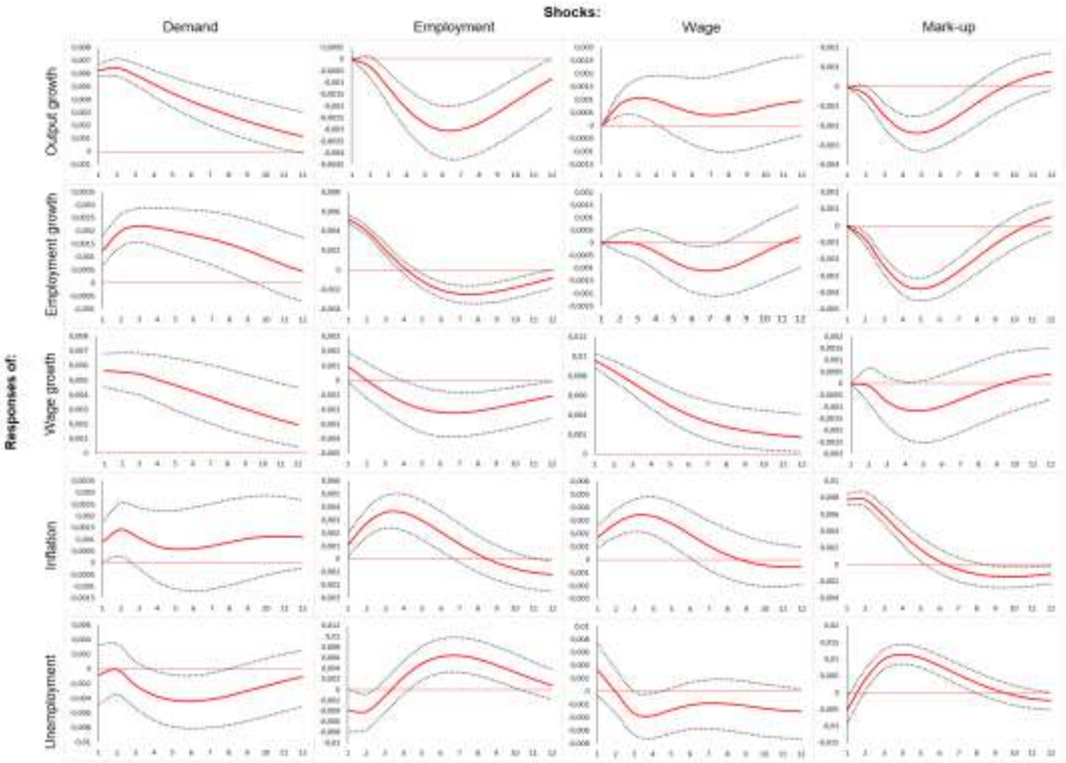
The main contributors to variations in unemployment are mark-up and negative productivity shocks. Nominal wage FEVD is mostly explained by exogenous wage shocks and increases in labour demand due to demand shocks. The exogenous nature of the wage process confirms the finding in Dadam and Viegi (2019) that nominal wages in South Africa do not react to labour market conditions: in fact, neither employment nor unemployment affects wage dynamics in the above FEVD, as in the pure insider-outsider model of Blanchard and Summers (1987). Employment, on the other hand, is very little influenced by aggregate demand and is mostly driven by productivity shocks. Prices respond to mark-up shocks, wages and productivity shocks.

The combination of the impulse responses, the steady state and the FEVD analysis describes an economic environment where mark-up, productivity and wage shocks have a strong role in determining employment and unemployment in the country. While demand shocks have a strong influence on output in the short run, they tend to be absorbed by wage increases, with little or no effect on employment and especially unemployment. The structural nature of unemployment is thus confirmed in the data.

3.3.2 Robustness analysis

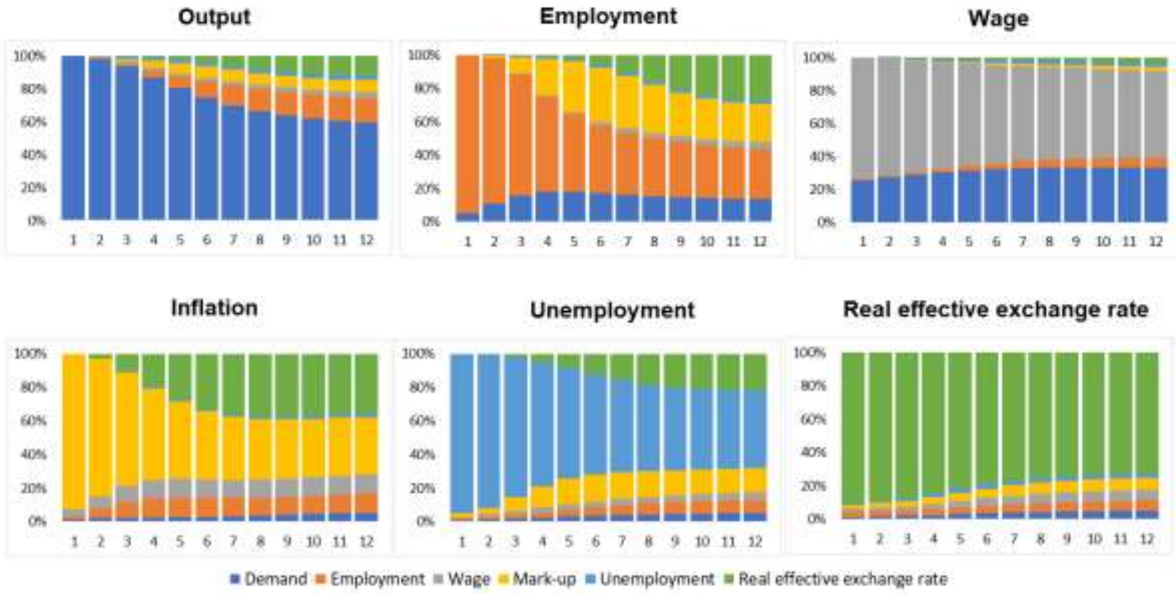
In this section we augment the baseline framework with variables aimed at capturing the influence of external shocks on the dynamics of a small open economy. We are also interested in the transmission of monetary policy. We begin by including the real exchange rate in the set of endogenous variables in Figure 8.

Figure 8: Impulse responses – real effective exchange rate



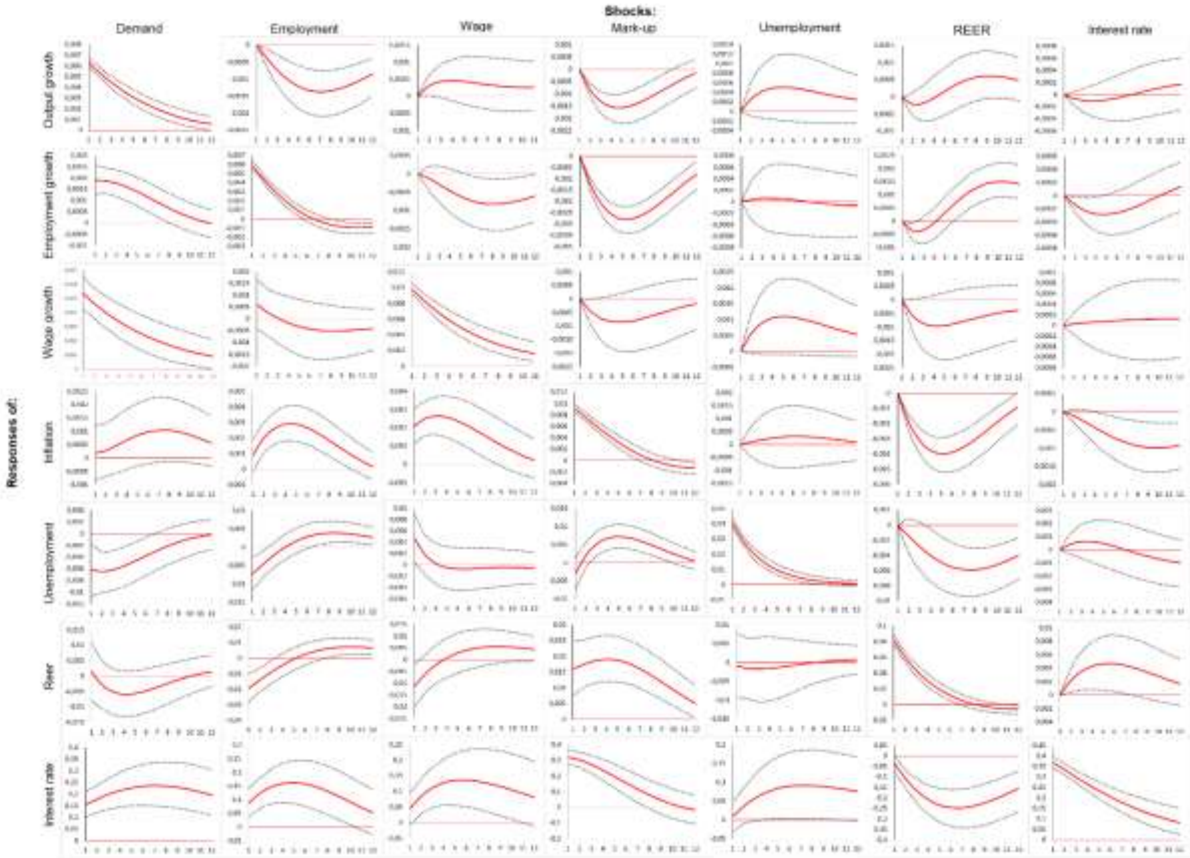
Accounting for the real effective exchange rate in the model does not change the main results. Interestingly, it does sharpen the findings for the impulse responses, as shown in Figure 8, and for the FEVD, as shown in Figure 9. Noticeably, however, real exchange rate variations explain part of what was previously named 'mark-up' shocks, which can now also be interpreted as the level of pass-through of exchange rate variation to prices. These exchange rate variations are important for all variables, except for wages, which remain totally unaffected, confirming the largely exogenous process that they represent.

Figure 9: FEVD – real effective exchange rate



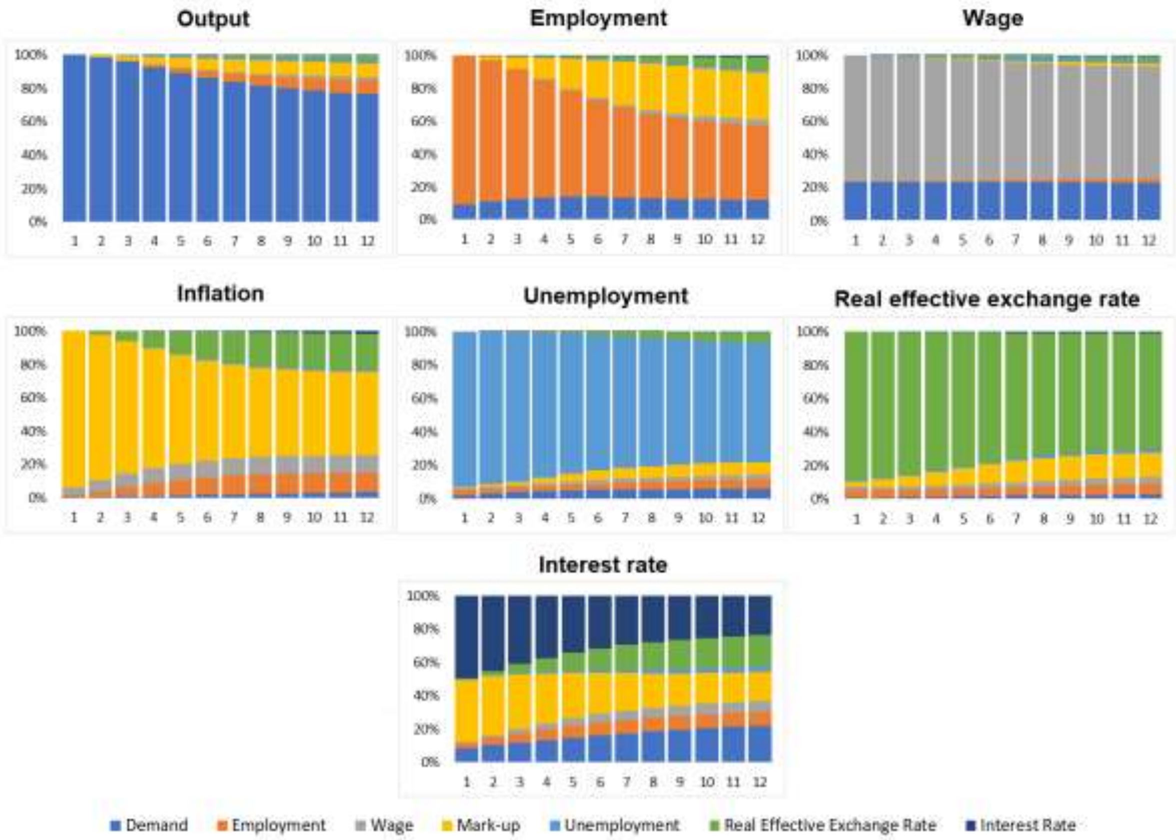
We move on to the response of the variables when we account for monetary policy. The impulse responses are shown in Figure 10.

Figure 10: Impulse responses – monetary policy



Similar to the inclusion of real effective exchange rate to the baseline scenario, we find in this case that the results remain consistent, which consolidates our findings. Furthermore, a contractionary monetary policy successfully reduces both employment and inflation. However, wages and unemployment do not significantly respond to a monetary policy shock, which seems to be a recurrent theme especially regarding the response of wages.

Figure 11: FEVD – monetary policy



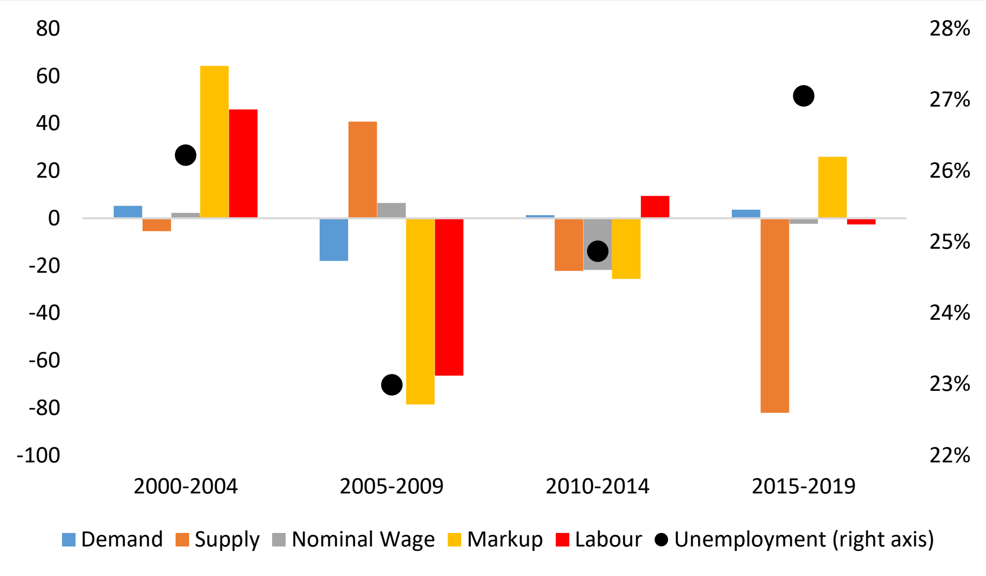
It is evident from this analysis that in the South African context there is a strong connection between nominal wages, prices and unemployment. This is exacerbated by the high bargaining power of workers, combined with the significant skill gap in the labour market that tends to favour skilled insiders when the economy emerges from a recession. Therefore, unemployment is kept consistently high, emphasised by the slowness of the recovery following a contractionary shock. Considering the monopolistic structure of firms in the economy does not improve the environment because of the constant ability to transfer the cost of increasing wages to workers through an increase in prices.

The evidence of hysteresis in unemployment shows the importance of reassessing the context in which policy in general is conducted in South Africa, and particularly how forecasting is dealt with. This is especially relevant since, as the data show, unless there is an improved economic outlook, the unemployment rate rises at a consistent pace and tends to reach a plateau at a higher equilibrium than previous periods. The FEVD shows consistent evidence of the importance of how the variables interact with each other. Ignoring the strong linkages between nominal wages, inflation and unemployment is bound to yield biased forecast results which will in turn influence policy decisions in the context of an inflation-targeting regime. Ultimately, the persistent rise in wages translating into overall higher prices could be problematic for a central bank committed to inflation targeting. This therefore puts the bargaining power of unions at the centre of this network because wage demands are not necessarily compatible with economic conditions and are not always productivity driven. The optimal policy recommendation would consequently be for the trade unions to internalise the inflation objective of the central bank.

3.4 Historical decomposition

We conclude this analysis by reporting the historical decomposition to assess the main contributions of shocks to the fluctuations in unemployment. Rather than reporting the full contributions of shocks at each point in time, we aggregate them over a period of five years throughout the sample by taking a five-year average of each shock contribution. We focus on the baseline scenario, with the results shown in Figure 12.

Figure 12: Historical decomposition



The high level of unemployment at the start of the sample is dominated by mark-up shocks and labour shocks, while negative supply shocks contribute noticeably less during the 2000–2004 period. Between 2005 and 2009, South Africa experienced a period of growth that significantly contributed to the decrease in unemployment. Importantly, favourable mark-up shocks created a competitive environment where firms were able to absorb the labour force, therefore contributing to the fall in unemployment. This was supported by positive supply shocks as the country experienced a period of robust growth. However, this was abruptly halted by the 2008 global financial crisis. Fast-forward to the last five years of the sample and we see how dramatic protests throughout the country as a result of political and social unrest led to the South African economy plunging into a recession. Unemployment increased drastically mainly as a result of negative supply shocks and unfavourable mark-up shocks.

4. Conclusion

In this paper we investigated hysteresis in the South African labour market as an explanation of the structural unemployment prevailing in the economy. We conducted various tests for hysteresis, including the construction of an index of strong hysteresis using the unemployment series. The second part of the analysis shed some light on the causes and consequences of a persistently high unemployment rate. To achieve that objective, we use a simple of model of hysteresis with insider-outsider dynamics, which we estimate using a Bayesian VAR methodology on South African data.. The model can only be solved if there is evidence of a unit root in the unemployment series. We found this to be the case for South Africa.

We found that demand shocks play a secondary role in explaining unemployment, with the main drivers being mark-up and negative productivity shocks. The results also suggest that nominal wages do not respond to real shocks. These findings are sharpened by our various robustness checks (by controlling for the real effective exchange rate and the monetary policy instrument) and are consolidated by the FEVD in which nominal wage shocks contribute significantly to explaining inflation forecast errors. Moreover, mark-up shocks become prominent contributors to unemployment forecast errors as the horizon increases.

These results are in line with the literature that explains the persistence in unemployment after an adverse shock through the linkages between nominal wages and prices. Specifically, workers benefiting from an increase in nominal wages via the bargaining power of labour unions bear the inflationary cost transferred to them by firms operating in a monopolistic environment, while struggling firms revert to significant labour shedding. Reforms to improve flexibility in the goods and labour markets could ease the difficulty of absorbing the high level of unemployment in South Africa’s economy.

Appendix A: Illustrating the aggregate dynamics in the Preisach model of hysteresis

We follow Cross (1994) for an application of the Preisach model in economics. Assume that aggregate employment *N* is generated by the aggregation of individual firms' decisions that varies according to a price level (the shock variable which can be any exogenous process affecting the profitability of the firm). This price level determines the flow of entry and exit of firms in the market. Therefore, the price level acts as a shock to the employment series and also as a barrier of entry/exit to individual firms in a given market. Firms will enter the market if the price level is above a certain threshold $\beta_{,j}$ and will exit the market when the price is below $\alpha_{,j}$, with $\alpha_{,j} < \beta_{,j}$, that is, each firm has a different threshold (for example, because of different productivity levels, as noted in Melitz (2003)). When the price level is between $\alpha_{,j}$ and $\beta_{,j}$, the firm will be inactive, either remaining in the market or remaining outside depending on what its position was in the previous period, like in the model of investment under uncertainty of Dixit et al. (1994).

Each firm is therefore defined by the pair of switching values to the evolution of the aggregate price level and the aggregate price level determines which firms exit, enter or remain in the market. Thus, the switching values are defined as exit and entry triggers such that $\alpha_j = P_{exit,j}$ and $\beta_j = P_{entry,j}$. The number of active firms at a certain price level determines the dynamics of aggregate employment. After a large shock to the price level, the total number of firms in the market will change while small shocks will not change the number as more firms will be in the 'neutral' zone where the decision to enter or exit has been postponed. This heterogeneity at the micro level is enough to generate a non-linear aggregate system where employment has selective memory of past shocks – that is, the system has remanence.

Given these characteristics, the Preisach model can be summarised by assuming a continuum in P of heterogeneous firms that behave hysteretically as follows:

$$N(t) = \iint_P u(\alpha,\beta) R_{\alpha,\beta} P(t) d\alpha d\beta \qquad (A.1)$$

in which $N(t)$ denotes aggregate employment (the output variable), $u(\alpha, \beta)$ is the density function of the individual firm (also known as the Preisach Function), $R_{\alpha,\beta}$ represents the individual relays defining the relationship between employment and aggregate demand at the firm level, and $P(t)$ is the input variable that is a proxy for product aggregate demand.

To illustrate how the selective memory works in a Preisach model, assume that $\alpha_0 = min\{P(t)|t = 1,2, \dots, n\}$ and $\beta_0 = max\{P(t)|t = 1,2, \dots, n\}$. α_0 and β_0 denote a heterogeneous set of firm's hysteretic relays that forms the limiting triangle T in Figure A.2 such that $T = \{(\alpha, \beta)| \geq \alpha \wedge \alpha \geq \alpha_0 \wedge \beta \leq \beta_0\}$. The limiting triangle T also represents the area where the density $u(\alpha, \beta)$ exists.

Figure A.1: Input function

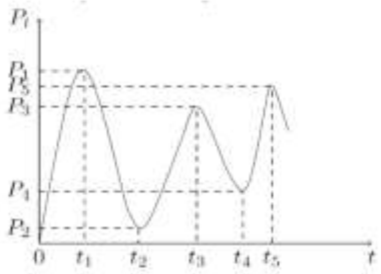


Figure A.2: Memory map in the Preisach model of hysteresis

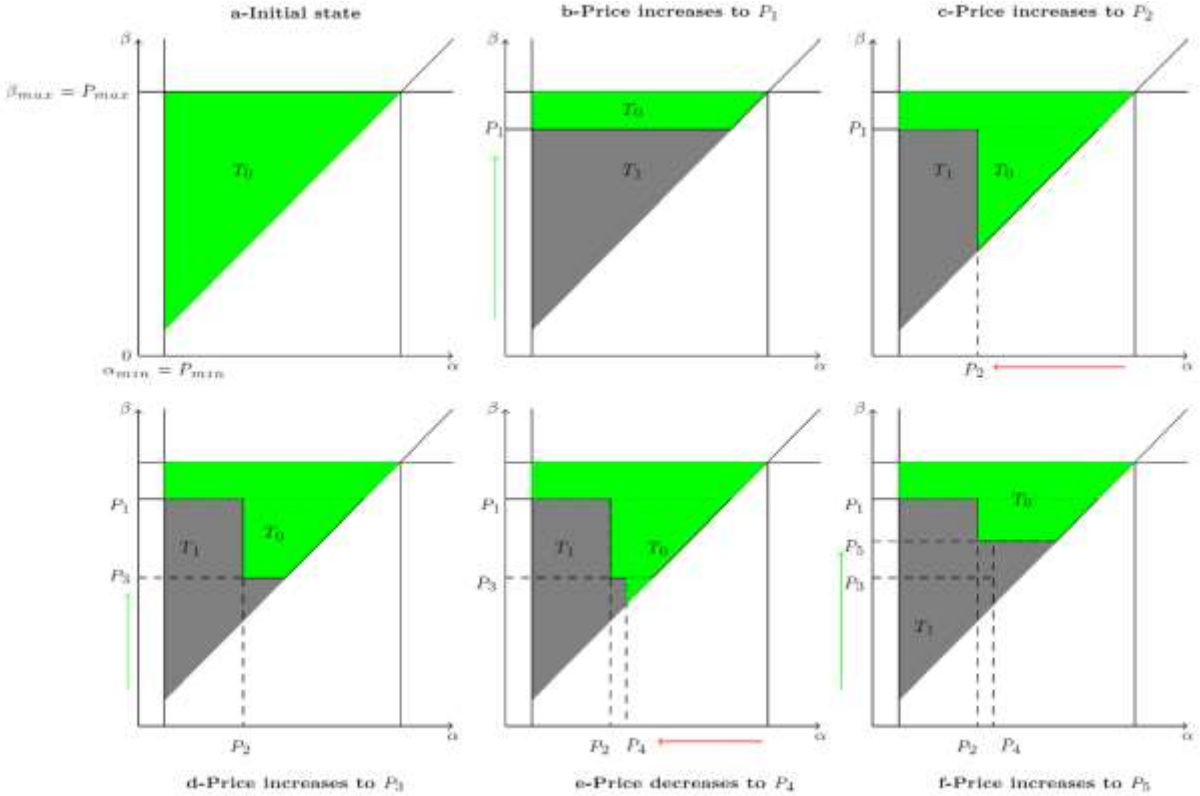


Figure A.1 represents a series for the evolution of aggregate price over time. The responses of employment to the cyclical changes in aggregate price are recorded in Figure A.2. An increase in price in Figure A.1 will lead to a shift in Figure A1 from $\beta = \alpha_{min}$ to the corresponding position where $\beta = P_t$, switching relays from the original limiting triangle T_0 (now a trapezoid) to the new triangle T_1. Conversely, a decrease in price will cause a shift from $\alpha = \beta_{max}$ to the corresponding $\alpha = P_t$, therefore switching from T_1 to T_0.

Assume the initial state with $P_t < \alpha_0$ where all the relays are switched off and all the firms that cannot afford the price prevailing in the product market face an entry barrier and therefore employ no workers. When the price increases to P_1 ($\beta \leq P_1$), only firms with prices satisfying the condition $P_{entry} \leq P_1$ start to hire. However, firms to which the price P_1 constitutes a barrier, start shedding labour and exit the market to form (with the firms already outside) the trapezoid T_0 in panel (b).

A decrease to price P_2 from P_1 generates different dynamics. Now, firms with prices set such that $P_{exit} \geq P_2$ will start firing workers, therefore decreasing the number of firms inside the market set by T_1. Fast-forward to panel (e), the staircase separating outsiders (T_0) and insiders (T_1) illustrates the property of hysteresis that only extremum values of price changes count. In other words, aggregate employment will be denominated by the sequence of non-dominated maximums and minimums of the aggregate price level.

Panel (d) shows another feature of hysteresis that is central to this study – the erasable, selective memory characteristic of the series. The idea is that some shocks have a longer lifespan as their lingering effects are felt long after the point of impact. To illustrate this, assume an increase in prices from P_4 to P_5. This makes P_5 the new local maximum because $P_5 > P_3$, but not the global maximum which remains P_1 that is greater than P_5. As a result, the effect of the previous price increase to P_3 is wiped out from the memory bank of the employment series. This is illustrated by a shift upward of the staircase separating insiders and outsiders, which serves to illustrate the selective memory property of hysteresis in the employment series. Aggregating these outputs in a sequential manner allows us to derive an index that captures maximums and minimums in a series to assess its hysteretical feature.

Appendix B: An insider-outsider dynamic model with hysteresis

Assume the long-linearised model:

$$y_t = n_t + a_t \quad (B.1)$$

$$p_t = w_t - a_t + \mu_t \quad (B.2)$$

$$y_t = \phi d_t \quad (B.3)$$

$$l_t = \alpha(w_t - p_t) - bu_t + \tau_t \quad (B.4)$$

$$w_t = w_t^* + \epsilon_{wt} + \gamma_1\epsilon_{dt} + \gamma_2\epsilon_{pt} \quad (B.5)$$

$$w_t^* = arg\{n_t^e = (1-\lambda)n_{t-1} + \lambda l_{t-1}\} \quad (B.6)$$

From (B.3):

$$\Delta y_t = \phi\Delta d_t \quad (B.7)$$

$$\Delta y_t = \phi\epsilon_{dt} \quad (B.8)$$

From (B.1):

$$y_t - n_t = a_t \quad (B.9)$$

$$\Delta y_t - \Delta n_t = \Delta a_t \quad (B.10)$$

Substituting in (B.8):

$$\phi\epsilon_{dt} - \Delta n_t = \epsilon_{st} \quad (B.11)$$

$$\Delta n_t = \phi\epsilon_{dt} - \epsilon_{st} \quad (B.12)$$

From (B.2):

$$\Delta p_t = \Delta w_t - \Delta a_t + \Delta\mu_t \quad (B.13)$$

$$\Delta p_t = \epsilon_{wt} - \epsilon_{st} + \epsilon_{pt} \quad (B.14)$$

From (B.5):

$$\Delta w_t = \epsilon_{wt} + \gamma_1\epsilon_{dt} + \gamma_2\epsilon_{pt} \quad (B.15)$$

From (B.4):

$$u_t + n_t = \alpha(w_t - p_t) - bu_t + \tau_t$$

$$(1+b)u_t = \alpha(w_t - p_t) - n_t + \tau_t$$

$$(1+b)\Delta u_t = \alpha\left(\epsilon_{wt} - \epsilon_{wt} + \epsilon_{st} + \epsilon_{pt}\right) - \phi\epsilon_{dt} + \epsilon_{st} + \epsilon_{lt}$$

$$(1+b)\Delta u_t = \alpha\epsilon_{st} - \phi\epsilon_{dt} + \epsilon_{st} + \epsilon_{lt}$$

$$\Delta u_t = \frac{1}{1+b}\left[(1+\alpha)\epsilon_{st} - \phi\epsilon_{dt} + \epsilon_{lt}\right] \qquad (B.16)$$

References

Acharya, S, Bengui, J, Dogra, K and Wee, S L. 2022. 'Slow recoveries and unemployment traps: monetary policy in a time of hysteresis'. *The Economic Journal* 132(646): 2000–2047.

Amable, B, Henry, J, Lordone, F and Topol, R. 2004. 'Complex remanence vs. simple persistence: are hysteresis and unit-root processes observationally equivalent?' *Economic Complexity* 14: 67–89.

Ball, L M. 2009. 'Hysteresis in unemployment: old and new evidence'. National Bureau of Economic Research (NBER) working paper no. 14818.

Bassi, F and Land, D. 2016. 'Investment hysteresis and potential output: a post-Keynesian-Kaleckian agent-based approach'. *Economic Modelling* 52: 35–49.

Blanchard, O J and Summers, L H. 1987. 'Hysteresis in unemployment'. *European Economic Review* 31(1–2): 288–295.

Cerra, V, Fatas, A and Saxena, S C. 2020. 'Hysteresis and business cycles'. International Monetary Fund (IMF) working paper, WP/20/73.

Cross, R B. 1994. 'The macroeconomic consequences of discontinuous adjustment: selective memory of non-dominated extrema'*. Scottish Journal of Political Economy* 41(2): 212–221.

Dadam, V and Viegi, N. 2015. 'Labour market and monetary policy in South Africa'. In *Fourteen years of inflation targeting in South Africa: the challenge of a changing mandate.* South African Reserve Bank Conference Series 2014.113–138

Dadam, V and Viegi, N. 2019. 'Estimating a new Keynesian wage Phillips curve for South Africa'. *Journal of Development Perspectives* 3(1–2): 137–163.

Darby, J, Hallett, A H, Ireland, J and Piscitelli, L. 1999. 'The impact of exchange rate uncertainty on the level of investment'*. The Economic Journal* 109(454): 55–67.

Davis, S, Haltiwanger, J and Schuh, S. 1996. *Job creation and destruction*. Cambridge, MA: The MIT Press.

Decker, R A, Haltiwanger, J, Jarmin, R S and Miranda, J. 2016. 'Declining business dynamism: implications for productivity'. Hutchins Center working paper no. 23. The Brookings Institution.

Dixit, A K and Pindyck, R S. 1994. *Investment under uncertainty*. Princeton University Press.

Dolado, J J and Jimeno, J F. 1997. 'The causes of Spanish unemployment: a structural VAR approach'. *European Economic Review* 41(7): 1281–1307.

Dosi, G, Pereira, M, Roventini, A and Virtillito, M. 2018. 'Causes and consequences of hysteresis: aggregate demand, productivity and employment'. *Industrial and Corporate Change* 27(6): 1015–1044.

Gali, J. 2015. 'Hysteresis and the European unemployment problem revisited'. NBER working paper no. 21430.

Gali, J. 2020. 'Insider-outsider labour markets, hysteresis and monetary policy'. NBER working paper no. 27385.

Gambetti, L and Pistoresi, B. 2004. 'Policy matters. the long-run effects of aggregate demand and mark-up shocks on the Italian unemployment rate'. *Empirical Economics* 29(2): 209–226.

Garga, V and Singh, S R. 2021. 'Output hysteresis and optimal monetary policy'. *Journal of Monetary Economics* 117: 871–886.

Giannone, D, Lenza, M and Primiceri, G E. 2015. ‘Prior selection for vector autoregressions’. *Review of Economics and Statistics* 97(2): 436–451.

Göcke, M. 2002. ‘Various concepts of hysteresis applied in economics’. *Journal of Economic Surveys* 16(2): 167–188.

Hallett, A H and Piscitelli, L. 2002. ‘Testing for hysteresis against nonlinear alternatives’. *Journal of Economic Dynamics and Control* 27(2): 303–327.

Kerr, A, Wittenberg, M and Arrow, J. 2014. ‘Job creation and destruction in South Africa’. *South African Journal of Economics* 82(1): 1–18.

Layard, R, Nickell, S and Jackman, R. 1991. *Unemployment: macroeconomic performance and the labour market*. Oxford: Oxford University Press.

Maidorn, S. 2003. ‘The effects of shocks on the Austrian unemployment rate: a structural VAR approach’. *Empirical Economics* 28(2): 387–402.

Mayergoyz, I. 1986. ‘Mathematical models of hysteresis’. *IEEE Transactions on Magnetics* 22(5): 603–608.

Melitz, M J. 2003. ‘The impact of trade on intra-industry reallocations and aggregate industry productivity’. *Econometrica* 71(6): 1695–1725.

Nickell, S. 1988. ‘Imperfect competition and the labour market’. In *Modelling the labour market*, edited by M Beenstock. Dordrecht: Springer. 145–167.

O’Shaughnessy, T. 2011. ‘Hysteresis in unemployment’. *Oxford Review of Economic Policy* 27(2): 312–337.

Pikoko, V and Phiri, A. 2019. ‘Is there hysteresis in South African unemployment? Evidence from the post-recessionary period’. *Acta Universitatis Danubius. Economica* 15(3): 365–387.

Piscitelli, L, Cross, R, Grinfeld, M and Lamba, H. 2000. 'A test for strong hysteresis'. *Computational Economics* 15(1): 59–78.

Thakoor, V. 2020. 'Market power, growth and inclusion: the South African experience'. IMF working paper, WP/2020/206.

Woodford, M. 2003. *Interest and prices: foundations of a theory of monetary policy.* Princeton University Press.