SAJE South African Journal of Economics
ESSA The Economic Society Of South Africa (Est. 1925)

South African Journal of Economics Vol. 86:1 March 2018

UNEMPLOYMENT IN AFRICA: A FRACTIONAL INTEGRATION APPROACH

GUGLIELMO MARIA CAPORALE* AND LUIS ALBERIKO GIL-ALANA†

Abstract
This paper estimates long-memory models to analyse the stochastic behaviour of unemployment in eleven African countries (Botswana, Ethiopia, Ghana, Kenya, Malawi, Mauritius, Nigeria, Senegal, South Africa, Tanzania and Zambia) from the 1960s until 2010. The empirical results provide very strong evidence of lack of mean reversion in all series under examination. This suggests that hysteresis models are the most relevant for the African experience (not surprisingly, given the rigidities in their labour markets). Therefore in such countries shocks hitting the unemployment series will have permanent effects, and policy makers should take appropriate action to reverse the effects of negative shocks.
JEL Classification: C22, C32, E24
Keywords: Unemployment, Africa, fractional integration

1. INTRODUCTION

Measuring the level of persistence in unemployment rates is of vital importance for the African economies: they suffer some of the highest unemployment rates in the world and therefore it is essential for policy makers to be able to determine their degree of persistence in different countries, which will enable them to choose country-specific responses to external shocks. This paper applies fractional integration techniques to model unemployment in 11 African countries (see below for details). It is the first study to analyse the stochastic behaviour of unemployment in the African continent using state-of-the-art econometric methods. There is a scarcity of papers analysing African unemployment, the main reason being limited data availability and reliability. However, the University of Groniger has recently made a huge effort to construct the most complete dataset on African unemployment to date (see De Vries *et al.*, 2015), which we will use for our purposes.

There are two main ways of thinking about unemployment in the existing literature. The natural rate theory (see Phelps, 1967; Friedman, 1968) implies that it should fluctuate around a stationary equilibrium level, known as the natural rate or non-accelerating inflation rate of unemployment (NAIRU), which is determined by economic fundamentals. In "structuralist" models (see Phelps, 1994) this can shift over time as a result of infrequent shocks due to changes in economic fundamentals; once these shifts are taken

* Corresponding author: Professor, Department of Economics and Finance, Brunel University London, Uxbridge UB8 3PH, UK. E-mail: Guglielmo-Maria.Caporale@brunel.ac.uk
† Faculty of Economics and Navarra Center for International Development (NCID-ICS), University of Navarra.
Prof. Luis A. Gil-Alana gratefully acknowledges financial support from the Ministerio de Economía y Competitividad (ECO2014–55236). Comments from the Editor and two anonymous reviewers are gratefully acknowledged.

 doi: 10.1111/saje.12178

into account mean-reversion appears to characterise unemployment. This type of models has been found generally to be appropriate for the United States experience (see, *e.g.* Gordon, 1989; Song and Wu, 1998; Leon-Ledesma, 2002). By contrast, hysteresis models (see Blanchard and Summers, 1986, 1987; Barro, 1988) appear to fit better the European countries, where unemployment exhibits a high degree of persistence, and its dynamic behaviour can be captured by long-memory models with a (near) unit root. Blanchard and Summers (1986, 1987) define hysteresis as: "a case where the degree of dependence is very high, where the sum of coefficients is close but not necessarily equal to one." This is a feature, for instance, of "insider" models (see Lindbeck and Snower, 1988), or of models in which fixed and sunk costs make current unemployment a function of past labour demand (see Cross, 1994, 1995).

The empirical literature testing unemployment theories initially relied on standard unit root tests (such as Dickey and Fuller, 1979; Augmented Dickey Fuller or Phillips and Perron, 1988), and subsequently used panel approaches to deal with the well-known problem of the low power of standard unit root tests (see, *e.g.* Leon-Ledesma, 2002), or autoregressive fractionally integrated moving average models (ARFIMA) to test for long memory in unemployment (see, for instance, Gil-Alana, 2001, 2002). Caporale and Gil-Alana (2007, 2008) also allowed for breaks in a fractional integration framework, and Caporale *et al.* (2017) took into account the possible correlation between the unemployment series. The advantage of a fractional integration framework compared to the classical I(0)/I(1) integration dichotomy is that the fractional parameter can take any real value, and therefore no arbitrary restrictions are imposed on the stochastic behaviour of the series and the model allows for a much richer dynamic structure that might suit the unemployment series particularly well. Estimating the differencing parameter d in a fractional integration framework is crucial to obtain information on the relevance of alternative unemployment theories. Specifically, an order of integration equal to 0 supports the NAIRU hypothesis, while a positive value of d (measuring the degree of persistence in unemployment) represents evidence in favour of hysteresis.

Our analysis improves considerably on previous studies using less sophisticated empirical methods. Most of them concern South Africa. For instance, Weir-Smith and Ahmed (2013) look at it from the point of view of economic geography taking a more descriptive approach focussing on the spatiality of economic activities. Phiri (2014) estimates a momentum threshold autoregressive model is to analyse nonlinear equilibrium reversion between unemployment and economic growth for South African data between the periods 2000 and 2013, but unlike us does not allow for possible long-memory properties. Vermeulen (2015) investigates the effects of inflation on (un) employment in South Africa through those of inflation on output, but again his standard cointegration framework imposes strong restrictions on the dynamic behaviour of the series, in contrast to our approach. Sam and Pokharival (2016) is an example of a study about unemployment in another African country; they analyse the determinants of youth unemployment in Kenya, but again their autoregressive distributed lag model does not allow for long memory, which is a serious limitation compared to our method. Therefore the contributions of this paper is twofold: first, it uses a new, much more extensive data set on African unemployment not previously analysed in empirical works, thereby offering new evidence; second, it applies for the first time long-memory techniques in this context. The layout is the following: Section 2 outlines the methodology, Section 3 discusses the data and the empirical results, Section 4 offers some concluding remarks.

2. METHODOLOGY

As mentioned before this study applied fractional integration methods; more specifically it uses both parametric and semiparametric techniques. For the former we use a version of the Lagrange multiplier (LM) test of Robinson (1994) that is based on the Whittle function in the frequency domain. The set-up is the following:

$$y_t = \theta^T z_t + x_t; \qquad (1-L)^d x_t = u_t, \qquad t = 1, 2, \ldots \tag{1}$$

where z_t is a (kx1) vector of deterministic terms, θ is a (kx1) vector of unknown coefficients and x_t is assumed to be I(d) where d is also estimated from the data. Given the parametric nature of this method, one needs to specify a structure for the error term, *i.e.* u_t in equation (1); we assume uncorrelated (*i.e.* white noise) and autocorrelated errors in turn, in the latter case using the exponential spectral model of Bloomfield (1973). This model approximates highly parameterized ARMA models with a small number of parameters and performs very well in the context of fractional integration. Robinson (1994) proposed a test statistic for testing the null hypothesis:

$$H_o: \ d = d_o \tag{2}$$

in equation (1) where d_o can be any real value. This allows to test for a number of hypotheses of interest, such as short memory ($d_o = 0$) against long memory behaviour ($d_o > 0$), but also alternatives such as stationarity ($d_o < 0.5$)/nonstationarity ($d_o > 0.5$) or mean reversion ($d_o < 1$) vs. lack of mean reversion behaviour ($d_o \geq 1$). As for the semiparametric approach, we also use a Whittle method but with a band of frequencies degenerating to zero (Robinson, 1995).

3. DATA AND EMPIRICAL RESULTS

The series analysed is the number of unemployed (in thousands) in 11 African countries, namely Botswana, Ethiopia, Ghana, Kenya, Malawi, Mauritius, Nigeria, Senegal, South Africa, Tanzania and Zambia. The frequency is annual and the starting date is 1960 for Ghana, Nigeria, South Africa and Tanzania; 1961 for Ethiopia; 1964 for Botswana; 1965 for Zambia, 1966 for Kenya and Malawi and 1970 for Mauritius and Senegal; the sample ends in 2010 for all countries. The data source is the African Sector Database (Output and Labour Data) from the University of Groninger (for details, see De Vries *et al.*, 2015). Data are only available until 2010. Only countries with at least 40 observations have been included.

First, we estimate the parametric model given by equation (1) with $z_t = (1, t)^T$, *i.e.* including a constant and a linear (time) trend. Table 1 displays the estimates of the fractional differencing parameter d in the following model

$$y_t = \theta_1 + \theta_2 t + x_t; \qquad (1-L)^d x_t = u_t, \qquad t = 1, 2, \ldots \tag{3}$$

where y_t represents the total number of unemployed in each country, θ_1 and θ_2 are unknown coefficients on the intercept and a linear time trend, respectively, and x_t is

Table 1. Estimates of d

(i) No autocorrelation			
	No det. terms	**A constant**	**A linear time trend**
Botswana	0.89 (0.72, 1.23)	**1.48 (1.18, 1.88)**	1.50 (1.23, 1.89)
Ethiopia	1.03 (0.87, 1.26)	1.43 (1.32, 1.62)	**1.52 (1.41, 1.66)**
Ghana	1.08 (0.89, 1.33)	1.28 (1.09, 1.55)	**1.31 (1.14, 1.54)**
Kenya	0.86 (0.54, 1.19)	**1.57 (1.40, 1.81)**	1.55 (1.38, 1.70)
Malawi	0.98 (0.76, 1.26)	1.34 (1.11, 1.62)	**1.34 (1.13, 1.61)**
Mauritius	0.96 (0.66, 1.30)	1.16 (0.91, 1.47)	**1.14 (0.95, 1.42)**
Nigeria	1.07 (0.92, 1.29)	1.40 (1.23, 1.67)	**1.40 (1.23, 1.67)**
Senegal	0.97 (0.76, 1.28)	1.76 (1.50, 2.10)	**1.69 (1.42, 2.04)**
South Africa	0.83 (0.65, 1.08)	1.08 (0.97, 1.27)	**1.09 (0.93, 1.30)**
Tanzania	1.00 (0.84, 1.26)	1.47 (1.32, 1.75)	**1.60 (1.48, 1.81)**
Zambia	0.97 (0.79, 1.23)	1.38 (1.19, 1.68)	**1.44 (1.26, 1.69)**
(ii) Autocorrelated (Bloomfield)			
	No det. terms	**A constant**	**A linear time trend**
Botswana	0.67 (0.56, 1.06)	0.99 (0.83, 1.48)	**0.93 (0.44, 1.49)**
Ethiopia	0.96 (0.73, 1.38)	1.45 (1.18, 1.83)	**1.52 (1.27, 1.79)**
Ghana	0.86 (0.56, 1.39)	1.20 (0.74, 2.01)	**1.20 (0.53, 1.96)**
Kenya	0.47 (0.39, 1.23)	**1.52 (1.19, 1.94)**	1.47 (1.18, 1.91)
Malawi	0.75 (0.45, 1.47)	0.90 (0.57, 1.86)	**0.93 (0.10, 1.83)**
Mauritius	0.45 (0.36, 1.33)	1.18 (0.79, 1.87)	**1.11 (0.69, 1.86)**
Nigeria	1.11 (0.78, 1.52)	1.10 (0.33, 1.53)	**1.09 (0.57, 1.51)**
Senegal	0.75 (0.53, 1.39)	0.93 (0.47, 1.98)	**0.93 (0.02, 1.81)**
South Africa	0.69 (0.46, 1.13)	**1.23 (0.98, 2.31)**	1.34 (0.95, 2.36)
Tanzania	0.89 (0.67, 1.31)	1.38 (1.19, 1.86)	**1.51 (1.25, 1.82)**
Zambia	0.86 (0.57, 1.35)	1.02 (0.65, 1.51)	**1.03 (0.46, 1.52)**

Note: The values in bold refer to the selected models. In parentheses the 95% confidence bands of the non-rejection values of *d*.
Source: Parametric methods (Dahlhaus, 1989; Robinson, 1994).

assumed to be I(*d*) where *d* can be any real value. We report the estimates of *d* for the three cases of *(i)* no deterministic terms (when θ_1 and θ_2 are assumed to be equal to 0 in equation (1)), *(ii)* with an intercept (θ_1 is unknown and θ_2 is assumed to be equal to 0) and *(iii)* with an intercept and a linear trend (with both θ_1 and θ_2 unknown), and assume in turn that the error term, u_t in equation (3) is a white noise process (in the upper part of Table 1) or autocorrelated as in the model of Bloomfield (1973) (in the lower part of the table). Also reported are the 95% confidence bands of the non-rejection values are based on the LM tests of Robinson (1994).

The most appropriate model specification for each series is shown in bold in the tables. In the white noise case, the time trend is required for all countries except for Botswana and Kenya; under the assumptions of autocorrelated disturbances the exceptions are instead Kenya and South Africa. As for the differencing parameters, in the white noise case the estimates of *d* are above 1 for all countries, and the unit root null (*i.e.* $d = 1$) cannot be rejected only for South Africa (1.09) and Mauritius (1.14); in the remaining cases, the estimates are significantly higher than 1 ranging from 1.31 (Ghana) to 1.69 (Senegal). With autocorrelated disturbances the values become substantially smaller, and only for three countries are the estimates significantly higher than 1 (Tanzania, Kenya and Ethiopia), while for the remaining eight countries the unit root null cannot be rejected.

Table 2 displays the estimates of *d* for different bandwidth parameters using a semi-parametric "local" Whittle method proposed in Robinson (1995). The order of integration is found to be equal to or higher than 1 in all cases. Evidence of unit roots is found

Table 2. Semiparametric estimates of d

m	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
BOT	**0.824**	**0.923**	**1.040**	**1.179**	1.351	1.395	1.392	1.420
ETH	1.500	1.500	1.500	1.500	1.500	1.500	1.500	1.500
GHA	**0.681**	**1.052**	1.470	1.500	1.500	1.500	1.500	1.500
KEN	1.416	1.444	1.500	1.500	1.500	1.500	1.500	1.500
MLW	**0.751**	**1.031**	**1.198**	1.336	1.335	1.490	1.489	1.500
MAT	**1.077**	**1.239**	1.322	1.417	1.295	1.321	1.375	1.401
NIG	1.414	1.494	1.500	1.500	1.332	1.332	1.368	1.362
SNG	**1.114**	1.337	1.439	1.500	1.500	1.500	1.500	1.500
SAF	1.375	**1.134**	**1.168**	**1.219**	1.288	1.375	1.459	1.500
TAN	1.500	1.500	1.500	1.500	1.500	1.500	1.500	1.500
ZAM	**1.051**	**1.235**	1.429	1.500	1.493	1.500	1.500	1.500
Low 5%	0.632	0.664	0.689	0.709	0.725	0.739	0.752	0.762
Up 5%	1.367	1.335	1.310	1.290	1.274	1.260	1.247	1.237

Note: m is the bandwidth parameter. In bold evidence in favour of the unit root null, *i.e.* $d = 1$.

in the cases of Botswana, Ghana, Malawi, Mauritius, Senegal and South Africa; for the remaining countries the orders of integration are significantly higher than 1.

4. CONCLUSIONS

This paper estimates long-memory models to analyse the stochastic behaviour of unemployment in eleven African countries (Botswana, Ethiopia, Ghana, Kenya, Malawi, Mauritius, Nigeria, Senegal, South Africa, Tanzania and Zambia) from the 1960s until 2010; it is the first academic study to do so. The empirical results provide very strong evidence of lack of mean reversion in all series under examination. This suggests that hysteresis models are the most relevant for the African experience, which is not a very surprising result if one considers the low degree of economic (financial development) of most of the countries in the sample as well as the existence of various types of rigidities in their labour markets. Therefore in such countries shocks hitting the unemployment series will have permanent effects, and policy makers should take appropriate action to reverse the effects of negative shocks.

REFERENCES

BARRO, R. (1988). The natural rate theory reconsidered: The persistence of unemployment. *American Economic Review, Papers and Proceedings*, 78: 32-37.

BLANCHARD, O. J. and SUMMERS, L. H. (1986). *Hysteresis and the European Unemployment Problem*. NBER Working Paper Series No. 1950, MIT Press.

——— and SUMMERS, L. H. (1987). Hysteresis in unemployment. *European Economic Review*, 31: 288-295.

BLOOMFIELD, P. (1973). An exponential model in the spectrum of a scalar time series. *Biometrika*, 60: 217-226.

CAPORALE, G. M. and GIL-ALANA, L. A. (2007). Non-linearities and fractional integration in the US unemployment rate. *Oxford Bulletin of Economics and Statistics*, 69: 521-544.

——— and GIL-ALANA, L. A. (2008). Modelling the US, UK and Japanese unemployment rates: Fractional integration and structural breaks. *Computational Statistics and Data Analysis*, 52: 4998-5013.

———, GIL-ALANA, L. A. and LOVCHA, Y. (2017). The PPP hypothesis revisited: Evidence using a multivariate long memory model. forthcoming, Empirical Economics Letters.

CROSS, R. B. (1994). The macroeconomic consequences of discontinuous adjustment: Selected memory of non-dominated extrema. *Scottish Journal of Political Economy*, 41: 212-221.

——— (1995). *The Natural Rate of Unemployment: Reflections on 25 Years of the Hypothesis*. Cambridge, UK: Cambridge University Press.

DAHLHAUS, R. (1989). Efficient parameter estimation for self-similar process. *Annals of Statistics*, 17: 1749-1766.

DE VRIES, G. J., TIMMER, M. P. and DE VRIES, K. (2015). Structural transformation in Africa: Static gains, dynamic losses. *The Journal of Development Studies*, 51: 674-688.
DICKEY, D. and FULLER, W. (1979). Distribution of the estimators for autoregressive time series with unit root. *Journal of the American Statistical Association*, 74: 427-431.
FRIEDMAN, M. (1968). The role of monetary policy. *American Economic Review*, 58: 1-17.
GIL-ALANA, L. A. (2001). The persistence of unemployment in the USA and Europe in terms of fractionally ARIMA models. *Applied Economics*, 33: 1263-1269.
——— (2002). Modelling the persistence of unemployment in Canada. *International Review of Applied Economics*, 16: 465-478.
GORDON, R. J. (1989). Hysteresis in history. Was there ever a Phillips curve? American Economic Review. *Papers and Proceedings*, 79: 220-225.
LEON-LEDESMA, M. (2002). Unemployment hysteresis in the US states and the EU: A panel approach. *Bulletin of Economic Research*, 54: 95-103.
LINDBECK, A. and SNOWER, D. J. (1988). *The Insider-Outsider Theory of Employment and Unemployment*. Cambridge, MA: MIT Press.
PHELPS, E. S. (1967). Phillips curve, expectations of inflation and optimal unemployment. *Economica*, 34: 254-281.
——— (1994). *Structural Slumps: The Modern Equilibrium Theory of Unemployment, Interest, and Assets*. Cambridge, MA, Harvard University Press.
PHILLIPS, P. C. B. and PERRON, P. (1988). Testing for a unit root in time series regression. *Biometrica*, 75: 335-346.
PHIRI, A. (2014). Nonlinear co-integration between unemployment and economic growth in South Africa. *Managing Global Transitions*, 12: 303-324.
ROBINSON, P. M. (1994). Efficient tests of nonstationary hypotheses. *Journal of the American Statistical Association*, 89: 1420-1437.
——— (1995). Gaussian semiparametric estimation of long range dependence. *Annals of Statistics*, 23: 1630-1661.
SAM, S. O. and POKHARIVAL, G. P. (2016). "Modelling Economic Determinants of Youth Unemployment in Kenya". *Journal of Emerging Trends in Economics and Management Sciences (JETEMS)*, 7: 31-38.
SONG, F. M. and WU, Y. (1998). Hysteresis in unemployment? Evidence from OECD countries. *Quarterly Review of Economics and Finance*, 38: 181-192.
VERMEULEN, C. (2015). *Inflation, Growth and Employment in South Africa: Trends and Trade*. ERSA Working Paper No. 547, Economic Research Southern Africa, South Africa.
WEIR-SMITH, G. and AHMED, F. (2013). "Unemployment in South Africa: Building a spatio-temporal understanding". *South African Journal of Geomatics*, 2: 218-230.