

Forecasting self-employment in the UK

George Saridakis

Small Business Research Centre, Kingston University Business School,
Kingston Hill, Kingston Upon Thames, Surrey, KT2 7LB, UK.
E-mail: G.Saridakis@kingston.ac.uk

Grammatoula Papaioannou

School of Business and Economics, Loughborough University,
Ashby Road, Loughborough, Leicestershire, LE11 3TU, UK.
E-mail: G.Papaioannou@lboro.ac.uk

Abstract: In this paper we forecast UK self-employment, using annual data for five decades. We use the autoregressive moving average (ARMA) methodology to produce a forecast three periods into the future (2014-2016). We also express the ARMA model as a state space model and estimate one-step predictions and dynamic forecasts for the same period. We then compare the univariate forecasts with multivariate multi-step ahead forecasts using a vector autoregressive (VAR) methodology. Comparing the multivariate forecasts with the univariate forecasts, we observe that both point to an increase in UK enterprise activity in the future, with the increase being sharper in the former.

Keywords: self-employment, ARMA, state space, VAR.

JEL classification: E2, C53, L26, M21.

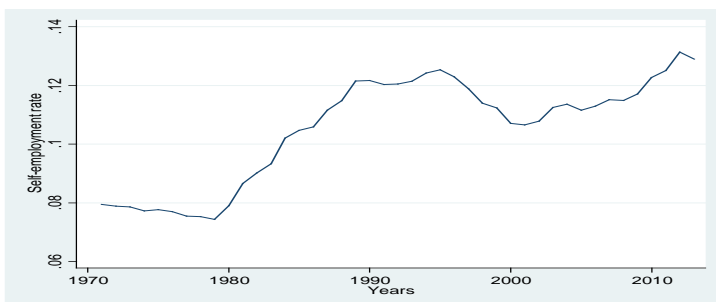
1. Introduction

Research on self-employment such as the relationship between tax evasion and self-employment, the role of unemployment on entrepreneurial decision, and the link between job satisfaction and occupational choice remains a rich source of academic debate and dispute (see Saridakis et al., 2014; Blanchflower and Oswald, 1998; Cowling and Mitchell, 1997). Whilst there are many published applied papers addressing the modelling of self-employment, they are not used for prediction. Recently research on self-employment has become even more important and interesting given the fact that many European countries are going through an unfavourable prolonged economic/financial crisis. On the one hand, financial constraints from the banking system and on the other, high public debts followed by Eurozone austere economic policies have affected the incomes of households and businesses, increased significantly unemployment rate (especially youth unemployment rate) and impacted on businesses growth and churning (see Saridakis, 2012). Hence, forecasts may help policymakers craft policies in this area and inform entrepreneurs and business analysts. This paper examines self-employment rates in a relatively liberal market economy – the UK. The UK is not a member of the euro-currency and has sought to promote entrepreneurship through self-employment initiatives.

Our aim is to predict the fluctuations of the UK self-employment rate, which it may be seen as an answer to increasing unemployment and generation of new ideas and innovative technologies that are essential components of returning to economic growth and prosperity. We forecasted the (log) total self-employment rate using annual data from 1971-2013 collected data from the National Statistics. Figure 1 shows that self-employment rose rapidly in the 1980s, decreased during the mid-1990s, and rose again in the 2000s (mean= 10.54% for the sample 1971-

2013). To forecast three periods into the future (2014-2016), we use the autoregressive moving average (ARMA) methodology and assume that our time-series process is a function of lagged random disturbances and its past values as well as a current disturbance term (Gujarati, 1995).¹ Finally, we consider fluctuations in unemployment rate that are likely to be related to fluctuations in self-employment rate and use an augmented VAR model to estimate multivariate multi-step ahead forecasts of self-employment growth (Patterson, 2000). Our models point to an increase in the self-employment rate in the future.

Figure 1. Self-employment rate, 1971-2013.



The rest of the paper organises as follows. Section 2 briefly presents the statistical methodology. Section 3 discusses the results. The last section concludes the paper.

¹ We also express the ARMA model as a state space model (see Hamilton, 1994) and estimate one-step predictions and dynamic forecasts for the same period.

2. Time-series analysis

We use a traditional Box-Jenkins model to produce short-run forecasts. Briefly, an ARMA (p, q) model, with p autoregressive lags and q moving-average lags can be written as:

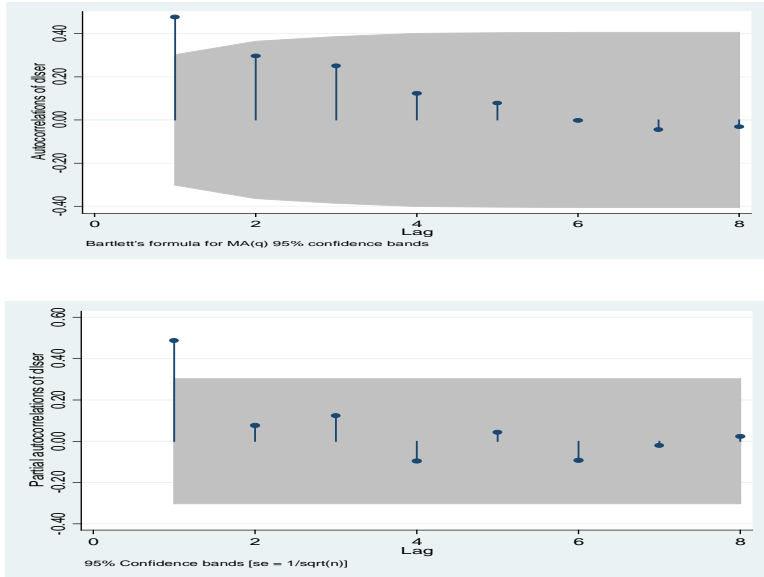
$$y_t = \theta_0 + \theta_1 y_{t-1} + \theta_2 y_{t-2} + \dots + \theta_p y_{t-p} + a_1 e_{t-1} + \dots + a_q e_{t-q} + e_t \quad (1)$$

Since the (log) total self-employment rate series is non-stationary², we created a new time-series by taking the first differences of the logged data, i.e. $y_t^* = (y_t) - (y_{t-1})$, which are stationary and become the input for our analysis and for which a mixed ARMA model can be constructed.

To determine an appropriate ARMA (p, q) model we examine the autocorrelation function (ACF) and the partial autocorrelation function (PACF) plots of the adjusted model. The difference between the ACF and the PACF is that the former gives the sequence of correlations between $(y_t^*$ and $y_{t-1}^*)$, $(y_t^*$ and $y_{t-2}^*)$ and so on, without holding the effects of intermediate lags constant (see Gujarati, 1995). In Figure 2, we observe a weak decay in the ACF and a single spike in the PACF for the first twelve lags of the differenced self-employment rate data.

² The Augmented-Dickey Fuller (ADF) test was applied to test the order of integration of the variables (i.e. to test the number of times a variable needs to be differenced in order to make it a stationary series: see Dickey and Fuller (1979)). Also, the Ng-Perron test confirms results from ADF test (Ng and Perron, 2001).

Figure 2. ACF and PACF of differenced self-employment.



To address this ARMA(1,1) or an AR(1) model may be considered. We estimate the models and compare the values of the AIC (see Akaike, 1974). Briefly, the AIC is used to determine if a particular model with p and q parameters is a good statistical fit. The model specification with the lower value is AR(1) with AIC equal to -178.768.

3. Empirical findings

The estimated AR model for the (log) total self-employment rate, with $p=1$, is given below:

$$\hat{y}_t^* = 0.006 + 0.487 y_{t-1}^* \quad (2)$$

(0.005) (0.141)

The standard errors of the coefficient are given below in parentheses. The estimated lagged coefficient is significant at a 1% significance level. The correlation between actual and predicted values is found to be 0.484 and statistically significant at 1% ($p\text{-value}=0.001$). Examining the residual correlogram for autocorrelations up to lag 8 shows that the autocorrelations are not statistically different from zero (plus and minus signs are balanced and also there are no patterns left). Finally, the Ljung-Box Q -statistic³ is insignificant with large p -value (results are available upon request). We also carry out an over-specification test, however, adding further autoregressive terms do not statistically improve the model (the coefficients of higher lagged values are found to be insignificant, and produced larger mean squares).

The AR(1) model is now used to forecast the (log) total self-employment rate for 3 years into the future. The forecasts obtained for three years into the future along with the forecast errors and the estimated 95% confidence intervals are presented in Table 1⁴. To obtain the forecast of self-employment rate level (rather than its changes), we integrate the first-differenced series⁵. We plot our forecasting values in Figure 3. Our results suggest that the total self-employment rate will increase within the next three years. Considering the annual growth of workforce population (estimated around 0.525%) and adjusted its values

³ The Q -statistic at lag k is a test statistic for the null hypothesis that there is no autocorrelation up to order k .

⁴ The forecast values are given from the following equations: $\hat{y}_{T+1}^* = \hat{\delta} + \hat{\theta}_1 y_T^*$;
 $\hat{y}_{T+2}^* = \hat{\delta} + \hat{\theta}_1 y_{T+1}^*$; $\hat{y}_{T+3}^* = \hat{\delta} + \hat{\theta}_1 y_{T+2}^*$.

⁵ $\Delta y_t = \delta + \theta_1 \Delta y_{t-1} \Leftrightarrow y_t - y_{t-1} = \delta + \theta_1 (y_{t-1} - y_{t-2}) \Leftrightarrow y_t = y_{t-1} + \delta + \theta_1 y_{t-1} - \theta_1 y_{t-2}$.

accordingly, we predict that there will be an increase in the stock of self-employment by 2.6% (i.e. self-employed up about 107,341 in three years).⁶

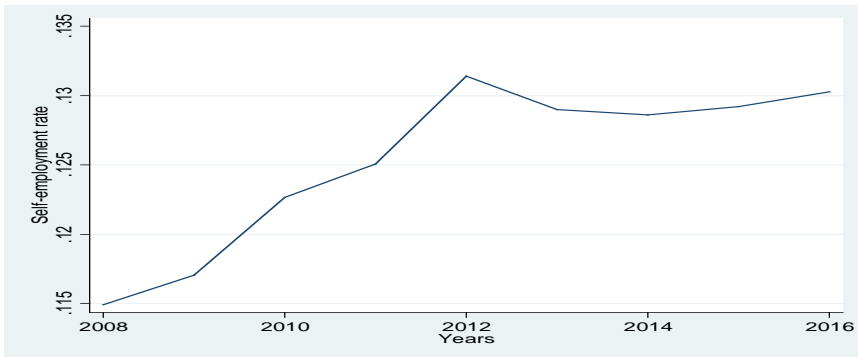
Table 1. Forecasts and Forecast Intervals for Changes in (log) Total Self-employment.

Year	Forecasts	Std. Error ¹	Forecast interval ²
2014	-0.003	0.028	(-0.058, 0.052)
2015	0.005	0.031	(-0.057, 0.066)
2016	0.008	0.032	(-0.055, 0.071)

¹The forecast errors for 1 year, 2 years and 3 years ahead can be written as: $v_1=e_{t+1}$, $v_2=\theta_1e_{t+1}+e_{t+2}$ and $v_3=(\theta_1^2+\theta_2)e_{t+1}+\theta_1e_{t+2}+e_{t+3}$ (see Hill et al., 2008).

²The 95% confidence interval is given by: $(\hat{y}_{T+j}^*-t_c\sigma_j, \hat{y}_{T+j}^*+t_c\sigma_j)$ where $j=1,2,3$ and $\sigma_1^2=\sigma_v^2$, $\sigma_2^2=\sigma_v^2(1+\theta_1^2)$, $\sigma_3^2=\sigma_v^2(1+\theta_1^4+\theta_1^2)$ (see Hill et al., 2008).

Figure 3. Self-employment rate in future (2014-2016) using AR(1) model.



⁶ Finally, following Hamilton (1994) and letting the state be $m_t = y_t^*$ and we write the AR(1) model as a state space model with the observation equation: $y_t^* = am_t + e_t$ and the state equation: $m_t = bm_{t-1}$. However, the estimation of a local-level model (Drukker and Gates, 2011) suggests similar predictive pattern.

Finally we use an augmented VAR(3) model to estimate a multivariate dynamic forecast of the self-employment rate:

$$z_t = \theta_0 + \sum_{i=1}^3 \Phi_i z_{t-i} + u_t \quad (2)$$

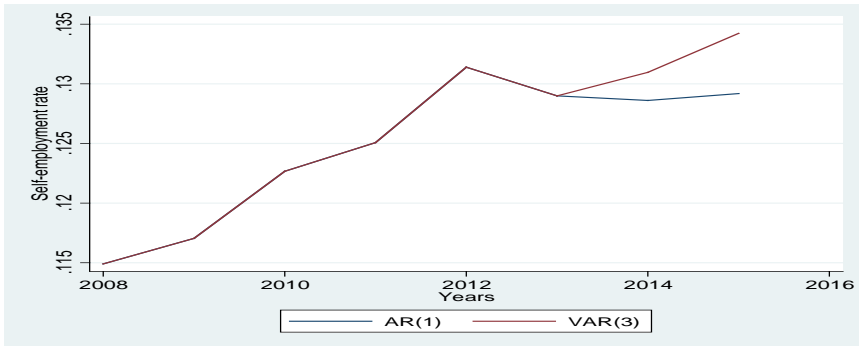
where z_t is a 2x1 vector of jointly determined dependent variables. In this model, we included the unemployment variables in the unrestricted VAR, since unemployment has been found to be closely connected to entrepreneurial activity (see, for example, Saridakis et al., 2014; Baumgartner and Caliendo, 2008; Storey, 1991; Hamilton, 1989). Firstly, we find that a residual based cointegration test based on Johansen's ML procedure suggests that the variables are cointegrated.⁷ Secondly, we make an assumption about the behaviour of the explanatory variable in the self-employment equation. For this reason, information about the behaviour of the economic variable was extracted based on reported forecasts for the years 2014-2015 by the National Institute of Economic and Social Research (NISESR).⁸ The forecast results are presented in Figure 4. The results also suggest an increase in self-employment rate for 2014-2015. Specifically it is estimated an increase in the stock of self-employment by 5.2% within the next two years (i.e. self-employed up about 214,000). Comparing the multivariate forecast with the univariate forecast presented earlier, we observe that both point to an increase in the self-employment rate with the increase being sharper in the multivariate forecast.⁹

⁷ The null hypothesis ($r = 0$) is rejected at the 5% level assuming an unrestricted intercept and restricted trend in the VAR.

⁸ NISESR unemployment projection for the year 2016 is not available.

⁹ An out-of-sample forecast for the period 2010-2013 suggests high correlation between actual and predicted values (0.840), especially with forecasts very close to the first two years suggesting that the proposed model may be useful for forecasting the target variable.

Figure 4. Multivariate dynamic forecast (2014-2015).



4. Conclusions

This paper has sought to predict the rate of self-employment drawing upon unemployment rates and past values of self-employment. We used an AR and VAR models to predict self-employment activity in the UK. Our forecast suggests a significant rise in self-employment rate over the next years. To more accurately predict self-employment rate additional factors into the regression model might be considered, since the decision to enter into self-employment is a complex phenomenon. Future research should deal with these issues by building on the multivariate model of self-employment rate including a variety of macroeconomic and social type variables.

References

Baumgartner, H.J. and M. Caliendo, 2008, Turning Unemployment into Self-Employment: Effectiveness of Two Start-Up Programmes, *Oxford Bulletin of Economics and Statistics*, 70(3), pp. 347-373.

Blanchflower, D.G. and A. Oswald, 1998, What Makes an Entrepreneur?, *Journal of Labor Economics*, 16, pp. 26-60.

Cowling, M. and P. Mitchell, 1997, The Evolution of UK Self Employment: A study of government Policy and the Role of the Macro-Economy, *Manchester School*, 64 (4), pp. 427-442.

Dickey, D. A. and W. A. Fuller, 1979, Distribution of the estimators for autoregressive time series with a unit root, *Journal of the American Statistical Association*, 74, pp. 427-31.

Drukker, D. M. and Gates, R. B., 2011, Stata Space Methods in Stata, *Journal of Statistical Software*, 41(10), pp. 1-25.

Gujarati, D., 1995, *Basic Econometrics*, Third Edition, McGraw-Hill International Editions.

Hamilton J.D., 1994, *Time Series Analysis*, Princeton University Press, Princeton.

Hamilton, R. T. 1989, Unemployment and business formation rates: reconciling time series and cross section evidence, *Environment and Planning*, 21, pp. 249–255.

Hill, C. R., W. E. Griffiths and G. C. Lim, 2008, *Principles of Econometrics*, Third Edition, Wiley.

Ng, S. and P. Perron, 2001, Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power, *Econometrica*, 69, pp. 1519-1554.

Office for National Statistics 2013, Self-employed up 367,000 in Four Years, Mostly Since 2011,
http://www.ons.gov.uk/ons/dcp171776_298533.pdf.

Patterson, K. 2000, *An Introduction to Applied Econometrics*, Macmillan press LTD.

Saridakis, G, S. Marlow and D.J. Storey, 2014, Do different factors explain male and female self-employment rates? *Journal of Business Venturing*, 29(3), pp. 345-362.

Saridakis, G. 2012, Introduction to the special issue on enterprise activity, performance and policy during times of crisis, *International Small Business Journal*, 30(7), pp. 733-735.

Storey, D. J. 1991, The birth of new firms – does unemployment matter? A review of the Evidence, *Small Business Economics*, 3, pp. 167-178.