Land Use Policy 61 (2017) 466–474

Contents lists available at ScienceDirect

Land Use Policy

journal homepage: www.elsevier.com/locate/landusepol

Forecasting agricultural production using co-integration analysis

Stilianos Alexiadis[1]

Ministry of Rural Development & Food, Department of Strategic Planning of Rural Development, Evaluation & Documentation, Division of Documentation & Agricultural Statistics, Room 519, 2 Acharnon St., 104 32, Athens, Greece

ARTICLE INFO

Article history:
Received 17 March 2016
Received in revised form 25 November 2016
Accepted 29 November 2016
Available online 9 December 2016

Keywords:
Forecasting
Co-integration
Agriculture

ABSTRACT

A forecasting method, easily understood and applied by administrators/policy-makers, is developed in this paper. Based on a simple, but quite realistic assumption, that production in a given region 'pulls' or 'shapes' the volume of production nationally a co-integration forecasting method is proposed in the paper. Using data from administrative regions, the empirical analysis suggests that the property of co-integration can be used to forecast agricultural production in the short-run.

1. Introduction

Many agricultural and regional economic development policies view the performance of *regions* as being linked to national economic performance. A 'region' is, indeed, a vital element in the performance of the national economy. In practice, however, several development polices assume a dichotomy between regional and national level. While all regions are part of the performance of the economy as a whole, the degree at which each spatial unit is linked to the performance of the national economy varies substantially. Some regions may exhibit strong links with total (national) activity while for other regions this link is weak or even absent. It becomes of crucial importance, therefore, to determine quantitatively that link. This estimate is of particular assistance to a wide range of applications. Predominately, among them is *forecasting*, a valuable apparatus in policy-making especially in agriculture. Forecasting models have a long tradition in agricultural economics (e.g. Taylor, 1924; Hopkins, 1927; Martin and Garcia, 1981; Brandt, 1985; Privette et al., 2015; Yazdanpanah et al., 2015). Although numerous methods have been proposed, ranging from parametric tests (Schmitz and Watts, 1970) to Normalized Quadratic Inverse Demand models (Klaiber and Holt, 2010), co-integration analysis is considered as an appropriate way to obtain short-run forecasts in agriculture.

It is undeniable that 'regional' factors have a substantial impact on shaping activities in the economy as a whole. It would be reasonable to assume that the level of a certain activity (e.g. agricultural production) in a region has a long-run relation with total (national) production. An intriguing problem is how to identify a region that 'shapes' national production. A useful approach is provided by co-integration analysis.

The purpose of this paper is to put forward the idea that (short-run) forecasting is feasible using the co-integration property between regional and national production in agriculture. Imposing co-integrating restrictions improves forecasting power, especially in models which exhibit strong evidence of co-integration between variables. This argument will be examined empirically in the agricultural sector of Greece. Using data for selected products, co-integration analysis will be applied across the administrative divisions of Greece (NUTS-2 regions). In order to achieve that, the paper is divided into four further sections. The second section provides an outline of the co-integration method and highlights the paucity of work in the area of forecasting using co-integrated techniques; it is this gap the paper seeks to fill. The forecasting technique is presented in the third section. The obtained results are discussed in the fourth section; while a fifth section concludes the paper by suggesting avenues for future research.

E-mail address: salexiadis7@aim.com

[1] I would like to acknowledge the helpful and supportive comments of two anonymous referees who positively, yet critically, helped refine the paper. The findings, interpretations and conclusions are those entirely of the author and do not necessarily represent the official position, policies or views of the Ministry of Rural Development & Food and/or the Greek Government.

http://dx.doi.org/10.1016/j.landusepol.2016.11.038
0264-8377/

2. Forecasting using the co-integration property

Usually, forecasting employs time-series analysis. Empirically this is indistinguishable from the issue of stationarity. A stochastic process is said to be stationary if its mean and variance are constant over time and the covariance between the two time periods depends only on the lag between the two time periods and not the actual time at which the covariance is colf a time series is stationary, its mean, variance and auto-covariance remain the same no matter at what point they are measured; that is, they are time invariant (Gujarati, 1995). Such a time series will tend to return to mean (mean reversion) and fluctuations around this mean (measured by its variance) will have broadly constant amplitude (Cuthbertson et al., 1995). In several cases, no relationship between two variables is expected, yet a regression of one on the other variable often shows a significant relationship. This situation exemplifies the problem of spurious regression. Assume that two (discrete) time-series, Y_t and X_t are individually non-stationary, $I(1)$, and the residuals, obtained from regressing $Y_t = a_1 + a_2X_t + u_t$, namely $u_t = Y_t - \hat{a}_1 - \hat{a}_2X_t$ are subjected to a unit-root test.[1] Regress $\Delta\breve{u}_t = \hat{a}u_{t-1}$ and if the computed |t| exceeds the critical Engle-Granger value,[2] then the conclusion is that the residuals are stationary, $I(0)$. Although Y_t and X_t are individually $I(1)$, their linear combination might be $I(0)$. Consequently, the linear combination cancels out the stochastic trends in Y_t and X_t. In this case, a regression of Y_t on X_t would be meaningful (not spurious). In econometric jargon, Y_t and X_t are co-integrated and $\hat{a}_2$ is known as the 'co-integrating parameter'.[3] Using a non-spurious (co-integrating) regression, even if individually the two variables are non-stationary, forecasts can be obtained. In this context 'one might wish to take more properly into account the nature of the relationships between variables (for example through co-integration analysis), as it would help the forecaster to present an economic story' (Carnot et al., 2005, p.74). The majority of research involving co-integration, however, has focused on hypothesis testing, not forecasting, because the presence of long-run relationships among variables is often a prediction of a theoretical model (Duy and Thoma, 1998, p. 292).

Assume that for a particular time period, let $[t_0 \ldots T]$, a dataset of agricultural products, $p = 1,\ldots,m$, for each spatial unit (region) of a country, $r = 1,\ldots,n$, is available. The problem at hand is to forecast the volume of national production ($Y_{N,p}$) at $T+1$. The forecasting process can be portrayed as follows. Since $Y_{N,p} = \sum^{N}{}_{r=1} Y_{r,p}$ $\forall t \in [t_0 \ldots T]$ as a first step estimate equation $Y_{N,p,t} = a_{p,r} + c_{p,r}Y_{r,p,t}$ and then apply the co-integration test. Having identify $\forall p$, the region in which production is co-integrated with national production and minimizes the deviations (r^*), it is possible to forecast national production. Suppose that experts (e.g. agronomists) in r^* estimate that production is expected to change, in relation to the previous year, by ±g%; hence $\hat{Y}_{r*,T+1} = \pm\gamma\% Y_{r*,T}$. Since $\hat{c}_{r*}$ gives the 'sensitivity' of $Y_{N,p}$ to changes in $Y_{r,p,t}$, national production at $T+I$ can be estimated using a simple formula $Y_{N,p,T+1} = \hat{a}_{r*} + \hat{c}_{r*}\hat{Y}_{r*,T+1}$.

3. Empirical application

The approach outlined in Section 2 will be applied for four categories of agricultural products ($p = 1,\ldots,12$), using data from the NUTS-2 regions of Greece ($r = 1,\ldots,13$).[4] The first category is 'Main Cereals', including Common Wheat, Durum Wheat, Rye, Burley, Oats, Grain Maize and Rice. The second category includes 'Industrial Plants', viz. Cotton-seed and Tobacco. The third category, 'Root Crops', includes Potatoes and Sugar Beets. A final category is 'Aromatic Plants'.[5] The data cover the period 2000–2013. At this point, it is worth mentioning that due to geomorphologic and climate factors, the contribution of each region to the national production vary considerably (Table 1).

A unit-root test was conducted for each individual time-series. As perhaps anticipated, few time-series turned to be $I(0)$. According to the unit-root test, the time series for each product at the national level are $I(1)$. Similarly, for Sugar-Beets, Cotton-seed, Tobacco and 'Aromatic Plants', this test indicates that the time-series are $I(1)$. The hypothesis of $I(0)$ at 1% level of significance is confirmed in only 8 cases. Regressing the level of national production against the level of production in each region suggests a long-run relation between national and regional production for several products. The unit-root test on the residuals confirms the co-integration hypothesis for most products at 1% level of significance, with the exception of Oats and Rice (at 5% and 10% level of significance, respectively).[6] Regressing national production against regional production, produces the results on Table 2.

Regressions for Oats using data from regions R1, R2, R5 and R9 gave spurious results.[7] A similar situation is evident for regions R2 and R5 with respect Cotton production. Three regions are of particular importance (R2, R4 and R7) since production in these regions, for almost half of the products included in the analysis, exhibit the co-integration property. Burley, Common Wheat, Durum Wheat and Potatoes are co-integrated with national production in several regions.

The co-integrated regions, at the usual levels of significance, are shown in Fig. 1. Spatial dependence between the co-integrated regions is evident, especially, for Durum Wheat, Rye, Burley and Tobacco.

The co-integrated region that minimises the deviations from the actual values may be chosen to perform the forecasting experiments. The choice is made using the U-coefficient, defined as $U = \sqrt{(1/m)\sum^{m}{}_{i=1}(A_t - F_t)^2} / \sqrt{(1/m)\sum^{m}{}_{i=1}(A_t)^2}$, where A and F are the actual and the fitted (predicted) values, respectively. This coefficient provides a measure of how well a time series of estimated values compares to a corresponding time series of observed values (Intriligator, 1978). This ratio can be extremely useful for comparing different forecast methods. The closer the value of U is to zero, the better the forecast method. Table 3 sets out the U-coefficient using the estimates for each region.

For the set of co-integrated regions, the parameters obtained yield minimum values of the U-coefficient in eight cases. For Durum Wheat most regions gave a U-coefficient in the range 0.002–0.008. Forecasts for this particular product, however, were conducted using the parameters for region R2 since the unit-root test for the residuals implies acceptance of $I(0)$ at 1% level of significance. Applying a similar reasoning region R3 is chosen for Rye. For Cotton, region R2 yields the minimum value of the U-coefficient. Using data for this particular region, however, produces spurious results and the hypothesis of co-integration is not accepted at the usual levels of significance. Consequently, the co-integration parameters obtained from region R1, for which the hypothesis of

[1] For any variable, this test appears as follows: $X_t = \rho X_{t-1} + u_t$.

[2] Since the residuals are based on the estimated co-integrated parameter, the Dickey-Fuller and the Augmented Dickey-Fuller critical significance values are not quite appropriate; instead the Engle-Granger critical values are used.

[3] A test for co-integration, according to Granger (1986) 'can be thought of as a pre-test to avoid 'spurious regression' situations' (p. 226).

[4] Data were obtained from the Ministry of Rural Development & Food of Greece. A list with the NUTS-2 regions of Greece is provided in the Appendix.

[5] Systematic cultivation of these products is rather rare (mainly scattered).

[6] See Table A1 in the Appendix.

[7] As a rule of thumb, if R^2 exceeds the Durbin-Watson Statistic, then the regression is spurious (Table A1).