

On the relationship between the business cycle and college enrollment in the U.S.: a time series approach

Pratikshya Sapkota* • Umesh Bastola

Washington State University, USA

Received: 7 August 2014

Revised: 19 December 2014

Accepted: 23 December 2014

Abstract

Causality between business cycle and college enrollment in the United States was examined using macro-level data from 1970 to 2011. Autoregressive distributed lag (ARDL) bounds test and Johansen test for cointegration showed a stable long-run cointegrating relationship among unemployment rate, per capita GDP, and college enrollment. Study results showed that the unemployment rate Granger caused college enrollment in both the short and the long run while GDP per capita Granger caused college enrollment only in the long run. Results indicated that, in the short run, college enrollment responds only to the change in opportunity costs over the business cycle. However, education is a normal good in the long run, and hence, college enrollment rises with increased incomes.

Keywords: college enrollment, unemployment, ARDL bounds test, Granger causality, business cycle

JEL Classification Codes: C22, E24, E32, I20

1. Introduction

The effect of business cycles on college enrollment has been examined widely in the literature. While college enrollment decisions are thought to be driven by variation in opportunity costs associated with business cycles, there are mixed findings in the literature as to whether college enrollment is affected significantly by business cycles. For example, Betts and MacFarland (1995), Dellas and Koubi (2003), and Dellas and Sakellaris (2003), among others, report that college enrollment decisions are counter-cyclical. Thus, under weaker economic conditions that

* Corresponding author. E-mail: p.sapkotabastola@email.wsu.edu.

Citation: Sapkota, P. and Bastola, U. (2015) On the relationship between the business cycle and college enrollment in the U.S.: a time series approach, *Economics and Business Letters*, 4(1), 7-16.

occur at the bottom of the business cycle that include higher unemployment rates and lower wages, college enrollment goes up. In other words, these studies found that economic downturns were associated with lower opportunity costs of investing in education which corroborates Becker's (1993) human capital theory. A New York Times article entitled "College enrollment falls as economy recovers" (Perez-Pena, 2013) reported a 2.3% decline in college enrollment in U.S. in Spring 2013 as compared to the previous spring (NSCRC, 2013). However, Johnson (2013) argued that college enrollment could potentially decline during times of economic recession because schools raise tuition rates, financial aid sources have shrunk, and students have difficulty obtaining loans. Bedard and Herman (2008) found this pattern to be counter-cyclical, procyclical, or acyclical depending on gender, GPA, and type of degree pursued. Moreover, Johnson (2013) found graduate enrollment to be countercyclical for females and acyclical for males. Other studies, such as Berger and Kostal (2002) and Card and Lemieux (2000) found an association between college enrollment and unemployment rates to be statistically non-significant.

Given the mixed findings in the literature on the relationship between college enrollment and business cycles, additional empirical investigation using more recent datasets and econometric methodologies would seem worth pursuing. The objectives of this study are to: i) investigate whether measures of business cycles such as unemployment rate and gross domestic product (GDP) are associated in a long-run relationship with college enrollment; and ii) examine the causal relationship between the variables. Specifically, this study contributes uniquely to existing literature by seeking to delineate short- and long-run causal effects. While previous studies relied on either cross-sectional or longitudinal data, we look at the relationships from a time-series econometric perspective and test the causality of business cycle on college enrollment in a Johansen's multivariate framework. A recently developed autoregressive distributed lag (ARDL) bounds testing approach and Johansen test for cointegration on annual time-series from 1970 to 2011 were used in the present study to show that college enrollment, unemployment rate, and GDP in the U.S. are associated in a long-run cointegrating relationship. Moreover, results showed the existence of a significant causal relationship between business cycle and college enrollment both in the short and the long run.

2. Data

Annual data from 1970 to 2011 were obtained for college enrollment, GDP, and unemployment rate from different sources. The data on college enrollment were obtained from the U.S. Census Bureau and constitute the total number of students enrolled in undergraduate, 2-year college, and graduate programs. Data on the US GDP (in per capita terms, in constant 2005 US dollars) were obtained from the World Development Indicators (The World Bank, 2012). Data on unemployment rates (for ages 16 years and above) were obtained from Bureau of Labor Statistics (BLS, 2013). Each time series was converted into an index by normalizing values for each variable by the corresponding initial year (1970) values, thus making each series begin with 100. Figure 1 shows that college enrollment and per capita GDP generally increased over time while unemployment rate has been more variable.

Figure 1. College enrollment, unemployment rate, and per capita GDP in the U.S. (1970 = 100)



3. Econometric methodology and results

The degree of integration of each variable was tested by employing augmented Dickey-Fuller (ADF) (Dickey and Fuller, 1979) and Phillips-Perron (PP) (Phillips and Perron, 1988) unit root tests. Following logarithmic transformation of each of the three series (unemployment rate, denoted as *unemp*; GDP, *gdp*; and college enrollment, *col*), each series was tested for non-stationarity (Table 1). Both ADF and PP tests suggested the presence of unit roots at levels for all three variables in cases: i) when only constant was used and ii) when both constant and trend were used. However, the null hypothesis of presence of unit roots was rejected in both cases when the tests were conducted for variables at their first difference. Thus, all variables were integrated of order one, i.e., $I(1)$.

Table 1. Augmented Dickey Fuller and Phillips-Perron unit root test

Variables	ADF Test		Phillips-Perron Test	
	Constant	Constant & Trend	Constant	Constant & Trend
Level				
col	-2.023	-3.076	-2.290	-3.116
unemp	-1.989	-1.949	-2.035	-1.992
gdp	-1.715	-2.459	-1.861	-1.590
First difference				
Δcol	-7.106***	-7.182***	-7.136***	-7.303***
$\Delta unemp$	-5.310***	-5.245***	-4.437***	-4.519***
Δgdp	-4.621***	-4.794***	-4.394***	-4.667***

*** indicates significant at the 1% level

3.1. Cointegration

The Autoregressive Distributed Lag (ARDL) bounds testing approach (Pesaran et al., 2001) was used to test for cointegration. The major advantages of the ARDL method are that: i) it can distinguish between dependent and independent variables; and ii) this approach has improved properties in that it allows flexibility in the structure of lags of the regressors as well as yielding robust results for small sample sizes (Pesaran et al., 2001). The ARDL technique involves investigating the presence of a long-run relationship using the following unrestricted error correction model (UECM) framework:

$$\Delta col_t = \beta_0 + \sum_{k=1}^n \beta_{1k} \Delta col_{t-k} + \sum_{k=0}^n \beta_{2k} \Delta unemp_{t-k} + \sum_{k=0}^n \beta_{3k} \Delta gdp_{t-k} + \delta_{11} col_{t-1} + \delta_{12} unemp_{t-1} + \delta_{13} gdp_{t-1} + \epsilon_{1t} \tag{1}$$

$$\Delta unemp_t = \theta_0 + \sum_{k=1}^n \theta_{1k} \Delta unemp_{t-k} + \sum_{k=0}^n \theta_{2k} \Delta col_{t-k} + \sum_{k=0}^n \theta_{3k} \Delta gdp_{t-k} + \delta_{21} \Delta unemp_{t-1} + \delta_{22} col_{t-1} + \delta_{23} gdp_{t-1} + \epsilon_{2t} \tag{2}$$

$$\Delta gdp_t = \gamma_0 + \sum_{k=1}^n \gamma_{1k} \Delta gdp_{t-k} + \sum_{k=0}^n \gamma_{2k} \Delta col_{t-k} + \sum_{k=0}^n \gamma_{3k} \Delta unemp_{t-k} + \delta_{31} gdp_{t-1} + \delta_{32} col_{t-1} + \delta_{33} unemp_{t-1} + \epsilon_{3t} \tag{3}$$

where Δ is the first difference operator. The residuals $(\epsilon_{1t}, \epsilon_{2t}, \epsilon_{3t})$ in the above equations are assumed to be normally distributed and white noise. The F -test was used to examine whether the cointegrating relationships exist among variables. The null hypothesis of no cointegration among the variables in equation (1), $H_0 : \delta_{11} = \delta_{12} = \delta_{13} = 0$, was tested against the alternative, $H_a : \delta_{11} \neq \delta_{12} \neq \delta_{13} \neq 0$. A similar test was performed each for equations (2) and (3). If the computed F -statistic is higher than the upper bound critical value (UCB), the null hypothesis of no cointegration is rejected; if it is below the lower bound critical value (LCB), the null

hypothesis of no cointegration cannot be rejected; and if it lies between the LCB and UCB, the result is inconclusive. For sample sizes ranging from 30-80, Narayan (2005) provides critical values for the upper and the lower bounds.

Results from the ARDL bounds test for cointegration are presented in Table 2. The models estimated were ARDL (2,0,2), ARDL (1,0,1), and ARDL (1,0,2) when college enrollment, unemployment, and GDP were the dependent variables, respectively. The lag lengths selected for the models were based on AIC and SBC criteria both of which suggested the same lag lengths. The computed F -statistic was greater than the critical value at the 5% significance level (4.260) when college enrollment was used as the dependent variable. However, the null hypothesis of no cointegration could not be rejected when unemployment and GDP were used as the dependent variables. Hence, the ARDL bounds tests suggested presence of one cointegrating equation, i.e., when college enrollment was used as the dependent variable.

Table 2. ARDL bounds test for cointegration

Dependent variable	F-statistics ^a
Col	7.610
Unemp	2.861
GDP	3.366

^aF-critical values are obtained from Narayan (2005) for $k = 2$ and $n = 40$.
5% critical value bounds for $I(0) = 3.435$ and $I(1) = 4.260$

In order to check the robustness of the cointegration results obtained from the ARDL bounds test, Johansen cointegration test (Johansen, 1988; Johansen and Juselius, 1990) was also used. The Johansen cointegration test provides two likelihood ratio test statistics, maximum eigenvalue and trace, for the number of cointegrating vectors. The Johansen test also suggested the presence of one cointegrating vector. The null hypothesis of no cointegration was rejected by both the trace (p -value = 0.03) and maximum eigenvalue (p -value = 0.05) statistics while the null hypothesis of the presence of more than one cointegrating vector could not be rejected, supporting the ARDL test results¹.

3.2. Granger causality

Engle and Granger (1987) showed that if a pair of $I(1)$ series are cointegrated, any shocks in the long-run relationship tend to be corrected to bring the system back to the equilibrium. Hence, such a long-run relationship can be modeled with an error correction model (ECM). However, if a long-run cointegrating relationship does not exist, vector autoregression (VAR) model should be followed to examine the short-run causality among the variables. Therefore, the VAR model was estimated when unemployment and GDP were the dependent variables and multivariate Johansen cointegration methodology (which provides estimates of both the long- and the short-run relationships within a system of equations) was used when college enrollment was the dependent variable. The system of equations estimated in the Johansen method was a vector error correction model (VECM) derived from a standard unrestricted vector autoregressive model

¹ Detailed results from the Johansen cointegration tests are omitted for the purpose of brevity and are available upon request.

(VAR) of lag length k . Thus, specifying either a multivariate VECM or VAR of lag length k depicted as follows tests the Granger causality:

$$\Delta col_t = \beta_0 + \sum_{k=1}^n \beta_{1k} \Delta col_{t-k} + \sum_{k=1}^n \beta_{2k} \Delta unemp_{t-k} + \sum_{k=1}^n \beta_{3k} \Delta gdp_{t-k} + \lambda ECT_{t-1} + \mu_{1t} \tag{4}$$

$$\Delta unemp_t = \theta_0 + \sum_{k=1}^n \theta_{1k} \Delta unemp_{t-k} + \sum_{k=1}^n \theta_{2k} \Delta col_{t-k} + \sum_{k=1}^n \theta_{3k} \Delta gdp_{t-k} + \mu_{2t} \tag{5}$$

$$\Delta gdp_t = \gamma_0 + \sum_{k=1}^n \gamma_{1k} \Delta gdp_{t-k} + \sum_{k=1}^n \gamma_{2k} \Delta col_{t-k} + \sum_{k=1}^n \gamma_{3k} \Delta unemp_{t-k} + \mu_{3t} \tag{6}$$

In the above equations, μ_{it} 's are the serially uncorrelated random errors with mean zero and finite covariance matrix. ECT_{t-1} is the error correcting term which represents the cointegrating vectors and λ is the adjustment coefficient which indicates speed of adjustment towards the equilibrium path of any deviations from the long-run cointegrating relationship. While the long-run causal relationship is established through the significance of the coefficient of lagged error correction term as tested by a t -test, the short-run causal relationship is established through the joint significance of the lagged terms of each of the explanatory variables which is tested with a F - or Chi-square-test.

Given that time-series models are often sensitive to the lag lengths, more than one criterion was used to select lag lengths to use. The Likelihood-ratio Statistic (LR), Final Prediction Error (FPE), Akaike Information Criterion (AIC), Schwarz Criterion (SC) and Hannan-Quinn Information Criterion (HQIC) identified two lag lengths to be the optimal lag length for the Granger causality tests (Table 3).

Table 3. Lag length selection

Lag	LR	FPE	AIC	SC	HQIC
0	NA	3.21e-06	-4.135	-4.006	-4.089
1	253.916	2.95e-09	-11.130	-10.613	-10.946
2	33.235*	1.64e-09*	-11.728*	-10.823*	-11.406*
3	7.894	2.04e-09	-11.536	-10.244	-11.076

* indicates lag length selected by the criterion

The results of the Granger causality tests suggest that there is a significant causal link (p-value = 0.053) from the unemployment rate to college enrollment in the short run (Table 4). However, no evidence of any other short-run causal relationships was found. As expected, the coefficient on the lagged error correction term was significant at the 1% level and had the correct (negative) sign. This implies that, in the long run, both the unemployment rate and the GDP Granger cause college enrollment. Moreover, the error correction coefficient of -0.481 indicated that about 48% of deviation in the long-run relationship is corrected in the current period to bring the system back to the equilibrium path.

Table 4. Granger causality test

Dependent Variable	Short-run results			Long-run results
	χ^2 statistics [prob]			ECT(t-stats)
	Δcol	$\Delta unemp$	Δgdp	
Δcol	-	5.875 [0.053]	3.093 [0.212]	-0.481*** (-4.279)
$\Delta unemp$	2.909 [0.233]	-	2.205 [0.332]	-
Δgdp	0.962 [0.618]	3.282 [0.193]	-	-

*** indicates significant at the 1% level.

Values in brackets and parentheses represent probability and *t*-statistics, respectively

3.3. Long-run estimates and stability

The long-run coefficient estimates based on ARDL (2,0,2) for the cointegrating equation are presented in Table 5. The coefficient for GDP was positive and highly significant which implies that GDP and college enrollment move in the same direction and suggests a procyclical relationship in the long run. On the other hand, the coefficient for the unemployment rate was positive but statistically non-significant.

Table 5. Estimated long-run coefficients based on ARDL (2,0,2).

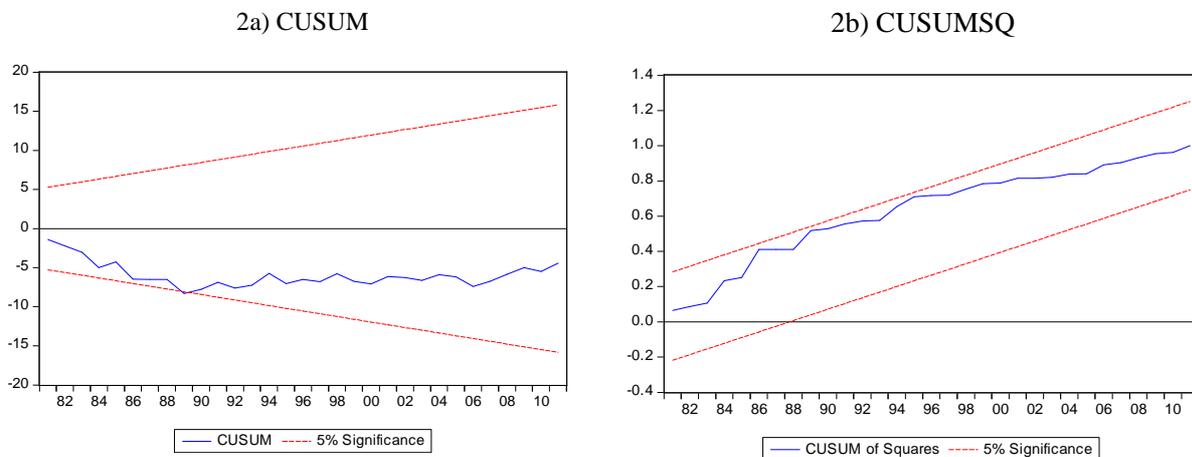
Regressor	Coefficient	Standard error	T-ratio [p-value]
<i>Dependent variable: Col</i>			
Unemp	0.168	0.110	1.525 [0.137]
GDP	0.893	0.126	7.088 [0.000]
Intercept	0.340	1.508	0.226 [0.823]

Next, we examine the stability of the estimated long-run parameters because unstable parameters can result in model misspecification and biased results (Narayan and Smyth, 2005). Following Pesaran and Pesaran (1997), the test for stability involved estimating the following error correction model:

$$\Delta col_t = \beta_0 + \sum_{k=1}^n \beta_{1k} \Delta col_{t-k} + \sum_{k=1}^n \beta_{2k} \Delta unemp_{t-k} + \sum_{k=1}^n \beta_{3k} \Delta gdp_{t-k} + \lambda ECT_{t-1} + \mu_{1t} \quad (7)$$

where all variables are as previously defined. The above model was estimated by ordinary least squares and the residuals were subjected to applying the cumulative sum of recursive residuals (CUSUM) and the CUSUM of square (CUSUMSQ) tests proposed by Brown et al. (1975) to determine the long-run parameter stability. The results from the CUSUM and the CUSUMSQ tests are plotted in Figures 2 and 3, respectively. The horizontal axes, the vertical axes, and the straight dotted lines represent year, corresponding statistics, and the 5% critical bounds for the statistics, respectively. The plots clearly indicate that the estimated parameters are stable since both the CUSUM and the CUSUMSQ statistics are well within the 5% critical bounds delineated by the dotted lines.

Figure 2. Plot of CUSUM and CUSUMSQ statistics. The dotted lines represent critical bounds at the 5% significance level.



4. Conclusions

Concern over increased college enrollment due to increased unemployment during the economic recession has increased in recent years. Relationships among college enrollment, unemployment rate, and GDP were examined for the United States with an ARDL bounds approach and Johansen test for cointegration and VECM and VAR models for Granger causality. This study contributes uniquely to the literature by explicitly delineating short run versus long run causal effects. The tests for cointegration suggested the presence of a stable long-run equilibrium relationship among the variables when college enrollment was the dependent variable. Results from the Granger causality tests indicated existence of causal links from the unemployment rate to college enrollment in both the short and the long run and from GDP to college enrollment in the long run. Moreover, the estimated long-run parameters indicated that college enrollment was procyclic when GDP was used as a measure of the business cycle. Intuitively, these findings suggest that in the short run, college enrollment responds only to the change in opportunity costs (as reflected by the unemployment rate) over the business cycle. However, in the long run, education is a normal good, and college enrollment rises with increased income levels.

Acknowledgements. We would like to thank Dr. Carole Engle and two anonymous referees for their insightful comments and suggestions.

References

- Becker, G.S. (1993) *Human capital: A theoretical and empirical analysis, with special reference to education*, University of Chicago Press, Chicago.
- Bedard, K. and Herman, D.A. (2008) Who goes to graduate/professional school? The importance of economic fluctuations, undergraduate field, and ability, *Economics of Education Review*, 27(2), 197-210.

- Berger, M.C. and Kostal, T. (2002) Financial resources, regulation, and enrollment in US public higher education, *Economics of Education Review*, 21(2), 101-110.
- Betts, J.R. and McFarland, L.L. (1995) Safe port in a storm: the impact of labor market conditions on community college enrollments, *Journal of Human Resources*, 30(4), 741-765.
- BLS (2013) *Bureau of Labor Statistics*, United States Department of Labor, available at www.data.bls.gov (accessed 5 October 2013).
- Card, D. and Lemieux, T. (2001) *Dropout and enrollment trends in the postwar period: what went wrong in the 1970s?*, in J. Gruber (ed.) *Risky Behavior Among Youths: An Economic Analysis*, University of Chicago Press, Chicago, 439-482.
- Dellas, H. and Koubi, V. (2003) Business cycles and schooling, *European Journal of Political Economy*, 19(4), 843-859.
- Dellas, H. and Sakellaris, P. (2003) On the cyclicity of schooling: theory and evidence, *Oxford Economic Papers*, 55(1), 148-172.
- Dickey, D.A. and Fuller, W.A. (1979) Distribution of the estimators for autoregressive time series with a unit root, *Journal of the American Statistical Association*, 74(336a), 427-31.
- Engle, R.F. and Granger, C.W.J. (1987) Cointegration and error correction: representation, estimation, and testing, *Econometrica*, 55(2), 251-276.
- Johansen, S. (1988) Statistical analysis of cointegration vectors, *Journal of Economic Dynamics and Control*, 12(2), 231-254.
- Johansen, S. and Juselius, K. (1990) Maximum likelihood estimation and inference on cointegration—with applications to the demand for money, *Oxford Bulletin of Economics and Statistics*, 52(2), 169-210.
- Johnson, M.T. (2013) The impact of business cycle fluctuations on graduate school enrollment, *Economics of Education Review*, 34, 122-134.
- Masih, A.M. and Masih, R. (1997) On the temporal causal relationship between energy consumption, real income, and prices: some new evidence from Asian-energy dependent NICs based on a multivariate cointegration/vector error-correction approach, *Journal of Policy Modelling*, 19(4), 417-440.
- Narayan, P.K. and Smyth, R. (2005) Electricity consumption, employment and real income in Australia evidence from multivariate Granger causality tests, *Energy Policy*, 33(9), 1109-1116.
- Narayan, P.K. and Smyth, R. (2006) Female labour force participation, fertility and infant mortality in Australia: some empirical evidence from Granger causality tests, *Applied Economics*, 38(5), 563-572.
- NSCRC (2013) *Term enrollment estimates Spring 2013*, National Student Clearinghouse Research Center, available at <http://research.studentclearinghouse.org> (accessed 24 September 2013).
- Perez-Pena, R. (2013) *College enrollment falls as economy recovers*, The New York Times, July 25, 2013, www.nytimes.com
- Pesaran, M.H., Shin, Y. and Smith, R.J. (2001) Bounds testing approaches to the analysis of level relationships, *Journal of Applied Econometrics*, 16(3), 289-326.
- Phillips, P.C.B. and Perron, P. (1988) Testing for a unit root in time series regression, *Biometrika*, 75(2), 335-346.

World Bank (2012) *World Development Indicators*, available at <http://data.worldbank.org> (accessed 7 October 2013).