MODELING OF TIME SERIES CYCLICAL COMPONENT ON A DEFINED SET OF STATIONARY POINTS AND ITS APPLICATION ON THE U.S. BUSINESS CYCLE

Ilya Bolotov

Abstract

The purpose of the paper is to suggest an alternative approach to modeling of the cyclical component in time series for a defined set of stationary points, which can be used as means of validation of economic theories and empirical hypotheses on the business cycle. The method constitutes an ordinary least squares estimation (OLS) of a polynomial model, which is derived from defined stationary points on the basis of polynomial factorization and integral transform. The method is applied to de-trended seasonally adjusted time series. It suggests a simple model for the cyclical component, which allows to incorporate any definition of business cycle in econometric calculations, thus providing an alternative to the commonly used filters (such as the moving average) and simplifies decomposing of time series into individual components. Its drawback is however computational effort linked to higher degrees of polynomials in long time series. The method is tested on the definition of business cycle employed by the U.S. National Bureau of Economic Research (NBER) on the U.S. annual and quarterly gross domestic product data, 1947–2013.

Key words: time series, cyclical component, stationarity points, polynomials, U.S. gross domestic product

JEL Code: C02, C22, E32

Introduction

The purpose of this short paper is to suggest an approach to modeling of the cyclical component in time series, which can be adjusted to any economic theory or hypothesis concerning business cycle based on its proper definition of stationary points; and test it on the data of the U.S. gross domestic product (GDP). The method constitutes an ordinary least squares (OLS) estimation of a polynomial derived from a defined set of stationary points with the help of Vieta's formulas and integration. The method is a combination of a deterministic calculation, algebra and calculus, and a stochastic estimation, i.e. linear regression.

Why is this matter topical? The nature and causes of the business cycle continue to be disputed and examined in both theoretical and empirical studies, see e.g. (Egert, Sutherland, 2014) and (Kaihatsu, Kurozumi, 2014) or (Chari, Kehoe, McGrattan, 2007) and (Pomenkova, Marsalek, 2012)[1]. The vast majority of these studies derive cyclical fluctuations from time series, y_t, using a classical decomposition approach in an additive or multiplicative form[2]:

$$y_t = s_t + g_t + c_t + \varepsilon_t, \quad y_t = s_t g_t c_t \varepsilon_t \Leftrightarrow \log y_t = \log s_t + \log g_t + \log c_t + \log \varepsilon_t \qquad (1)$$

where s_t stands for seasonal component, g_t for trend, c_t for cyclical component and ε_t for residuals; a recent exemplary study is (Guillen, Rodriguez, 2014). Several studies replace classical decomposition with spectral analysis, see e.g. (Wang, 2013). The distinctive feature of these methods is that they derive cyclical fluctuations from properties of the examined time series rather than from predefined criteria. Economic theory however suggests definitions of the business cycle, e.g. two consecutive positive quarterly changes, $\Delta y_{t-2} < 0$, $\Delta y_{t-1} > 0$ and $\Delta y_t > 0$ for an expansion and negative ones, $\Delta y_{t-2} > 0$, $\Delta y_{t-1} < 0$ and $\Delta y_t < 0$, for a recession, which do not become reflected in econometric analysis. Therefore it becomes difficult to support theoretical assumptions using c_t derived from the use of trend and filters without further adjustment, as the number of theoretically defined and empirically found local extrema and their frequency may differ; see (Ravn, Uhlig, 2002), (Schenk-Hoppe, 2001), (Movshuk, 2003) and (Bildirici, Alp, 2012) on this matter.[3]

1 Method based on defined stationary points

One of the simplest ways to solve the described problem is to create a method for modeling cyclical components, $c_t = m_t + \gamma_t$, which can incorporate any necessary definition of the business cycle or of cyclical fluctuations in general. If stationary points (local extrema, minima and maxima) can be defined beforehand, a polynomial, i.e. $\alpha_0 t_0^n + \alpha_1 t_1^{n-1} + \ldots + \alpha^n$, where n is the number of stationary points plus one, or a dummy variable, $D_t \in \{0, 1\}$, can be used as m_t. The main advantage of polynomials over D_t is the ability to reflect weights of individual stationary points while being a single function, as cyclical fluctuations' amplitude may differ over time. Polynomials also often produce higher R^2 than dummies.

The mathematical derivation of a polynomial c_t model was performed e.g. in (Bolotov, 2012). In this paper it will be re-elaborated and performed in five successive steps:

[1] This overview reflects the works on business cycle in the SSCI World of Science database, which were the closest to the topic of this paper at the date of its completion.
[2] In this paper additive form will be preferred.
[3] It is also possible to mention Howrey's criticism of Kuznets swing, the ca. 18-year-long cycles.

1) Deriving the cyclical component c_t from y_t by seasonal adjustment and de-trending:

$$c_t = y_t - s_t - g_t \qquad (2)$$

2) Defining the set of stationary points by applying the definition-specific criteria to time series y_t:

$$SP = \{t_1, t_2, t_3, t_4 \dots, t_5\} \qquad (3)$$

3) Formulating a polynomial from the set SP, i.e. the first derivative of m_t, on the basis of fundamental theorem of algebra, the Fermat's theorem and polynomial factorization rules[4]:

$$dm_t/dt = \alpha_0(t - t_1)(t - t_2) \dots (t - t_n) = \alpha_0 \prod_{i=1}^{i=n}(t - t_i) \qquad (4)$$

4) Integrating dm_t/dt to obtain the m_t equation with α_0 and α_n as unknown parameters:

$$m_t = \int \alpha_0 \prod_{i=1}^{i=n}(t - t_i)dt + \alpha_n \qquad (5)$$

5) Formulating the model of c_t and performing an ordinary least squares (OLS) estimation of α_0 and α_n, a_0 and a_n, where $\int \prod_{i=1}^{i=n}(t - t_i)dt$ is the independent variable (regressor), $t_i \in SP$.

$$c_t = \alpha_n + \alpha_0 \int \prod_{i=1}^{i=n}(t - t_i)dt + \gamma_t \qquad (6)$$

Corrections for GDP growth and large numbers

The method also requires corrections for the following two problems to be succesfully applied: GDP time series usually show strong trend, which implies growth in absolute values of c_t in time, and bigger *SP* sets lead to large numbers (N > 10^9, often N > 10^{50} or 10^{100}), which make calculations either complicated or infeasible for standard software. In this paper the author proposes a mathematical substitution for m_t (an alteration of the model) with the following compensations: a smoothing factor of $1/10^{(n-t)}$ penalizing older observations and an *n*th or (*n*+1)th root[5] of dm_t/dt to decrease the large numbers:

$$c_t = \mu_t + \xi_t = \beta_n + \beta_0 \int \sqrt[N=n\ or\ n+1]{\prod_{i=1}^{i=n}(t - t_i) \cdot 1/10^{(n-t)}}\, dt + \xi_t \qquad (7)$$

Both compensations are based on logic of the method without ground research.

[4] According to this theorems every polynomial equation with complex coefficients and degree equal to or greater than 1, e.g. $dm_t/dt = 0$, has at least one complex root. Fermat's theorem states that for every stationary point m_t'equals zero. The equation of dm_t/dt can therefore be written in the described form as a product of α_0 and $(t - t_i)$, where *i* ranges from 1 to *n*.

[5] The choice of root depends on which number, *n* or (*n*+1), is odd so that negative values of m_t can be preserved.

Simplification of calculations

The simplest way to increase precision of calculations of the independent variable $\int \sqrt[N]{\prod_{i=1}^{i=n}(t - t_i) \cdot 1/10^{(n-t)}}\,dt$ is to simplify the part $\prod_{i=1}^{i=n}(t - t_i)$ to the form $a_0 n t_0^{n-1} + \alpha_1(n-1)t_1^{n-2} + \dots + a_{n-1}$ and afterwards derive the antiderivative. The coefficients α_i, where i ranges from 1 to n – 1, can be calculated using the Vieta's formulas based on elementary symmetric polynomials s_i:

$$\alpha_i = (-1)^i s_i \alpha_0, 1 \le i \le n-1, s_1 = \sum_{j=1}^{j=n} t_j\,,\ s_2 = \sum_{j=1}^{j=n}\sum_{k>j}^{k=n} t_j t_k\,, \dots, s_{n-1} = \prod_{j=1}^{j=n} t_j \quad (8)$$

2 Modelling cyclical fluctuations of the U.S. GDP

To empirically verify the described model this paper will use the data for time series of the U.S. real gross domestic product (GDP) in two versions: annual (1820–2013), prices of 1990, and quarterly (1947Q3–2013Q4), prices of 2009. The data sources are Groningen University, Home Maddison database and Maddison Project database (historical data on GDP per capita and population) and U.S. Bureau of Economic Analysis (BEA) (population). U.S. National Bureau of Economic Research (NBER) complex (multicriterial) definition of the business cycle is used to identify stationary points, SP, Peak and Trough (data is published by NBER).

The cyclical component c_t is derived according to formula (2).

Seasonal adjustment for quarterly data

The seasonal adjustment for quarterly data had been already performed by BEA with the help of X12-ARIMA method, therefore this paper skips this step.

De-trending for annual and quarterly data

To increase the quality of results, the author uses 4 de-trending techniques: an exponential trend (OLS) $g_t = \beta_0 t^{\beta_1} \Leftrightarrow \log g_t = \log \beta_0 + \beta_1 \log t$ (best fit according to analysis of differences, Δy_t / y_t, R^2 and DW statistics), moving average (MA) (of 3 and 4 observations for annual and quarterly data), Hodrick-Prescott (HP) filter with parameters λ =100 (annual) and λ = 1600 (quarterly) (Hodrick, Prescott, 1997) and Baxter-King (BK) filter with frequency limits 2 to 8, k = 3, for annual data and 6 to 32, k = 12, for quarterly data (Baxter, King, 1995).[6]

Polynomial generation tools correction for big numbers

The author uses the GNU Regression, Econometrics and Time-series Library (gretl), MS Excel and proper JavaScript code to derive μ_t, which is based on Vieta's formulas.

[6] The parameter/ limit values for BK filter were suggested as "standard" by statistical / econometric software.

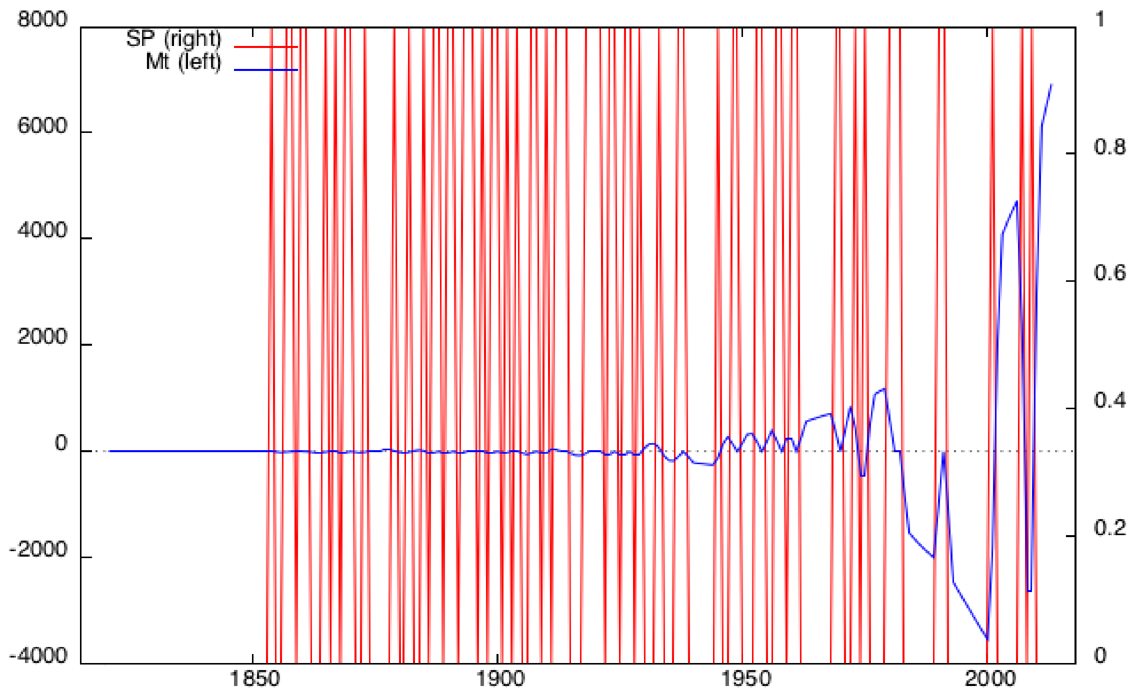
Modeling cyclical component for annual data, 1820–2013

The annual time series of the U.S. GDP, which consists of 194 observations and is most likely the longest break-free time series of GDP in today macroeconomic statistics, has, according to NBER, 64 stationary points: 33 peaks and 34 troughs (peaks and troughs occurred in the same year are aggregated), the latest being 2007 and 2009. Based on the data, the estimation of dm_t/dt for the U.S. GDP, $\widehat{dm_t/dt}$, has the following non-simplified form before correction[7]:

$$\widehat{dm_t/dt} = a_0 \int_1^{194} (t-35)(t-38)(t-39)(t-41)(t-42)(t-46)(t-48)(t-50)(t-51)(t-54)(t-60)(t-63)(t-66)(t-68)(t-69)(t-71)(t-72)(t-74)(t-75)(t-76)(t-78)(t-80)(t-81)(t-83)(t-85)(t-88)(t-89)(t-91)(t-93)(t-94)(t-95)(t-99)(t-100)(t-101)(t-102)(t-104)(t-105)(t-107)(t-108)(t-110)(t-114)(t-118)(t-119)(t-126)(t-129)(t-130)(t-134)(t-135)(t-138)(t-139)(t-141)(t-142)(t-150)(t-151)(t-154)(t-156)(t-161)(t-162)(t-163)(t-171)(t-172)(t-182)(t-188)(t-190) \qquad (9)$$

The corresponding estimation of independent variable (the integral), designated as M_t, is presented in Fig. 1 together with a dummy variable D_t for stationary points, *SP*.[8]

Fig. 1: Estimation of independent variable (the integral) for annual data



Source: author

[7] The simplified form proved to be too long for this paper. The reader can recur to online simplification tools.
[8] The variable takes two values, 1 for a stationary point and 0 for the rest.

Correlation between M_t, annual U.S. GDP (prices of 1990) and cyclical components OLS_C, MA_C, HP_C and BK_C is presented in Tab 1.

Tab. 1: Pearson correlation results for Mt, annual data

GDP_1990	OLS_C	MA_C	HP_C	BK
0.1798	-0.1742	0.0616	0.0993	0.0646

Source: author

The results show that *Mt* is mostly positively correlated with the cyclical components derived by other methods with a weak correlation coefficient (this is predictable due to a different definition of fluctuations) and shows stronger positive relation to the real U.S. GDP than HP_C (0.0406) and and BK_C (0.0252), which is a positive result for a new method. The long-term negative fluctuation observed since 1970s is nevertheless puzzling and seems to be generated by the polynomial itself. The results can be improved by fully estimating c_t using an OLS method, see equation 7.

Modeling cyclical component for quarterly data, 1947Q1–2013Q4

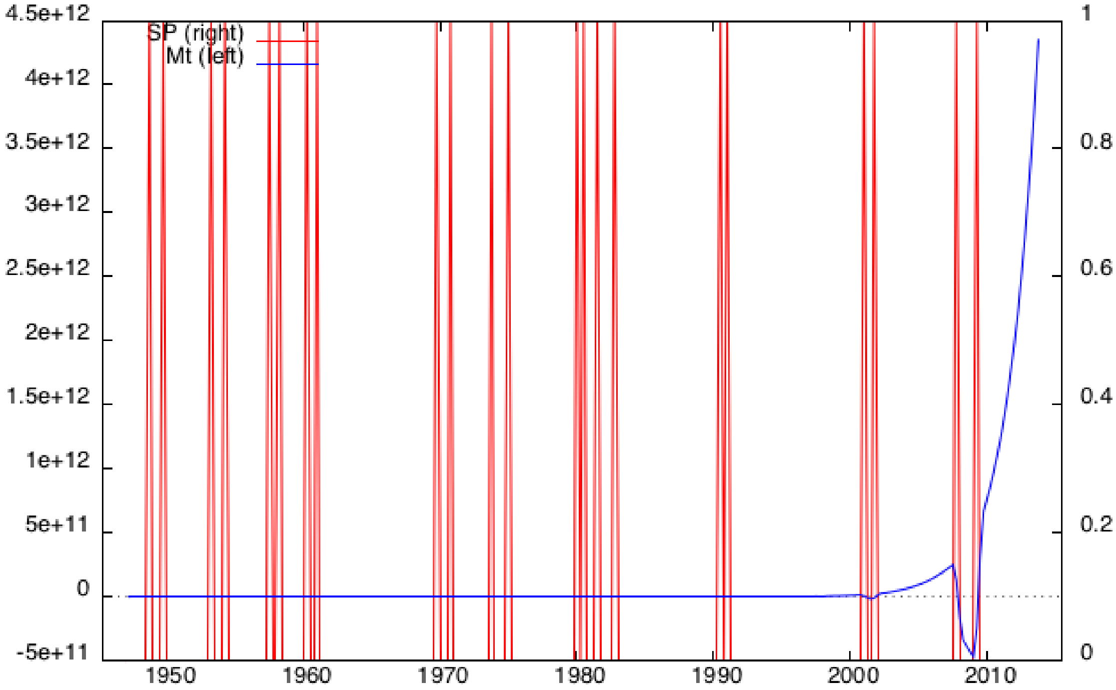
The quarterly time series of the U.S. GDP, which consists of 268 observations is also one of the longest break-free time series on GDP with, according to NBER, 22 stationary points: 11 peaks and 11 troughs, the latest being 2007Q4 and 2009Q2. Based on the data, the estimation of dm_t/dt for the U.S. GDP, $\widehat{dm_t/dt}$, has the following non-simplified form before correction[9]:

$$\widehat{dm_t/dt} = a_0 \int (t-8)(t-12)(t-26)(t-30)(t-43)(t-46)(t-54)(t-57)(t-92)(t-96)(t-108)(t-113)(t-133)(t-135)(t-139)(t-144)(t-175)(t-177)(t-217)(t-220)(t-244)(t-250) \qquad (10)$$

The corresponding estimation of independent variable (the integral), also designated as M_t, is presented in Fig. 2 together with a similar dummy variable D_t for stationary points, *SP*.

[9] The simplified form proved to be too long for this paper as well.

Fig. 2: Estimation of independent variable (the integral) for quarterly data



Source: author

Correlation between *Mt*, seasonally adjusted quarterly U.S. GDP (prices of 2009) and cyclical components OLS_C, MA_C, HP_C and BK_C is presented in Tab 2.

Tab. 2: Pearson correlation results for Mt, quarterly data

GDP_2009QSA	OLS_C	MA_C	HP_C	BK
0.4248	-0.4198	0.2067	0.0807	-0.0817

Source: author

The results are similar to the ones for the annual data with even stronger correlation between *Mt* and the U.S. GDP, second only to OLS_C (-0.9942) and even stronger when compared with MA_C (0.3457), which can be explained, among other, by a greater number of observations. Negative fluctuation is also observed only for the recent recession (2008–2009). There is however a problem with the smoothing factor, which proved to be insufficient leading to large values for 2010s. The results can be improved by fully estimating c_t as well.

Some thoughts on quality assessment

The author suggests that one of the methods to assess the quality of M_t with regard to a specific business cycle definition (the described method serves only this purpose), specifically when large numbers do not allow exact OLS fitting and precise estimates of R^2, DW and other criteria, is the correlation analysis, mostly correlation between the differences Δc_t (as approximation of first derivatives) derived from different methods and the dummy variables for stationary points (*SP*), peaks (*Peak*) and troughs (*Trough*), consult Tab. 3.

Tab. 3: Pearson correlation results for first differences

c_t	Annual			Quarterly		
	SP	**Peak**	**Trough**	**SP**	**Peak**	**Trough**
d_OLS_C	0,0794	0,0273	0,0285	0,0756	0,0592	0,0454
d_MA_C	-0,2617	-0,1017	-0,3338	-0,0282	-0,1050	0,0660
d_HP_C	-0,2917	-0,0484	-0,3877	-0,1865	-0,0765	-0,1815
d_BK_C	-0,2561	0,0223	-0,4065	-0,2356	-0,1535	-0,1724
d_Mt	**-0.0585**	**-0.0551**	**0.0038**	**-0.0198**	**0.0299**	**-0.0578**

Source: author

From the table it is perceivable that correlation between dummy variables *SP*, *Peak* and *Trough* for both annual and quarterly data is mostly inverse and weaker than 0.25 for all d_c_t (the estimations of Δc) with d_Mt being similar to d_OLS_C and d_MA_C rather than to HP and BK filters. M_t is therefore not substantially worse off than the other methods and can be used for cyclical fluctuations modeling as well after OLS estimations, which are left out in this short paper.

Conclusion

The method described in this paper derives a polynomial model of the cyclical component in time series, c_t, from a set of pre-defined stationary points, which can fit the majority of business cycle definitions; and consists of deterministic calculations and OLS-fitting that under the conditions of homoscedastic and serially uncorrelated errors γ_t is an optimal estimator of c_t. The test on the U.S. GDP, annual and quarterly data, for a part of the method showed that the method does not produce substantially worse results compared with OLS estimations, moving averages, Hodrick-Prescott and Baxter-King filters while being more similar to the first two. Certain calibration is nevertheless needed as part of results were puzzling and imprecise, as well as corrections for GDP growth and big numbers also require robustness checks and may need modifications. The main equation of the method takes the following form:

$$c_t = \beta_n + \beta_0 \int \sqrt[n \, or \, n+1]{\prod_{i=1}^{i=n} (t - t_i) \cdot 1/10^{(n-t)}} \, dt + \xi_t \quad (11)$$

In total, the method represents a certain innovation in polynomial modelling of time series and may also be used in theoretical explanations. For example, the residuals γ_t can be examined to assess a business cycle theory's consistency or, if held a priori valid, to analyze impacts of random influences on c_t. The author would like to encourage readers to perform certain estimations on their own if they find the concept interesting.

Acknowledgment

This paper was elaborated in the framework of the University of Economics, Prague, Faculty of International Relations' institutional support, No. IP0200040.

References

Baxter, M., & King, R. G. (1995). Measuring Business Cycles Approximate Band-Pass Filters for Economic Time Series (NBER Working Paper No. 5022). National Bureau of Economic Research, Inc. Retrieved from http://ideas.repec.org/p/nbr/nberwo/5022.html

Bildirici, M. E., & Alp, S. (2012). Measurement of Filters' Efficiencies and Application of Nnsfdi Method. Ekonomska Istrazivanja-Economic Research, 25(4), 937–957.

Bolotov, I. (2012). The Problem of Relationships Between Conditional Statements and Arithmetic Functions. Mundus symbolicus, 20, 5–12.

Chari, V. V., Kehoe, P. J., & McGrattan, E. R. (2007). Business Cycle Accounting. Econometrica, 75(3), 781–836. doi:10.1111/j.1468-0262.2007.00768.x

Egert, B., & Sutherland, D. (2014). The Nature of Financial and Real Business Cycles: The Great Moderation and Banking Sector Pro-Cyclicality. Scottish Journal of Political Economy, 61(1), 98–117. doi:10.1111/sjpe.12036

Guillen, A., & Rodriguez, G. (2014). Trend-cycle decomposition for Peruvian GDP: application of an alternative method. Latin American Economic Review, 23(1). doi:10.1007/s40503-014-0005-3

Hodrick, R. J., & Prescott, E. C. (1997). Postwar U.S. Business Cycles: An Empirical Investigation. Journal of Money, Credit and Banking, 29(1), 1–16.

Kaihatsu, S., & Kurozumi, T. (2014). Sources of business fluctuations: Financial or technology shocks? Review of Economic Dynamics, 17(2), 224–242. doi:10.1016/j.red.2013.08.001

Movshuk, O. (2003). Does the choice of detrending method matter in demand analysis? Japan and the World Economy, 15(3), 341–359. doi:10.1016/S0922-1425(02)00030-0

Pomenkova, J., & Marsalek, R. (2012). Time and frequency domain in the business cycle structure. Agricultural Economics-Zemedelska Ekonomika, 58(7), 332–346.

Ravn, M. O., & Uhlig, H. (2002). On adjusting the Hodrick-Prescott filter for the frequency of observations. Review of Economics and Statistics, 84(2), 371–376. doi:10.1162/003465302317411604

Schenk-Hoppe, K. R. (2001). Economic growth and business cycles: A critical comment on detrending time series. Studies in Nonlinear Dynamics and Econometrics, 5(1), 75–86. doi:10.1162/108118201753421141

Wang, P. (2013). Business Cycle Phases and Coherence-a Spectral Analysis of Uk Sectoral Output. Manchester School, 81(6), 1012–1026. doi:10.1111/j.1467-9957.2012.02304.x

Contact

Ilya Bolotov

University of Economics, Prague

Nám W. Churchilla 4

130 67 Praha 3

Czech Republic

Ilya.bolotov@vse.cz