Dec. 2012, Volume 6, No. 12 (Serial No. 61), pp. 1641–1647
Journal of Civil Engineering and Architecture, ISSN 1934-7359, USA

Unit Root Analysis of Traffic Time Series in Toll Highways

Antonio Sánchez Soliño and Antonio L. Lara Galera
Construction Department, Madrid Polytechnic University, Madrid 28040, Spain

Abstract: Concession contracts in highways often include some kind of clauses (for example, a minimum traffic guarantee) that allow for better management of the business risks. The value of these clauses may be important and should be added to the total value of the concession. However, in these cases, traditional valuation techniques, like the NPV (net present value) of the project, are insufficient. An alternative methodology for the valuation of highway concession is one based on the real options approach. This methodology is generally built on the assumption of the evolution of traffic volume as a GBM (geometric Brownian motion), which is the hypothesis analyzed in this paper. First, a description of the methodology used for the analysis of the existence of unit roots (i.e., the hypothesis of non-stationarity) is provided. The Dickey-Fuller approach has been used, which is the most common test for this kind of analysis. Then this methodology is applied to perform a statistical analysis of traffic series in Spanish toll highways. For this purpose, data on the AADT (annual average daily traffic) on a set of highways have been used. The period of analysis is around thirty years in most cases. The main outcome of the research is that the hypothesis that traffic volume follows a GBM process in Spanish toll highways cannot be rejected. This result is robust, and therefore it can be used as a starting point for the application of the real options theory to assess toll highway concessions.

Key words: Real options, unit root analysis, investment, highway concession, traffic.

1. Introduction

Most of road traffic models are based on the relationship between traffic volume and a number of explicative variables for which available information and prediction capacity are greater than for traffic itself. However, the use of time-series models may be an alternative tool to predict the traffic volume and to build a confidence interval for the forecast, when there are available data for traffic in a given road during a enough long period.

In this case, it can be assumed, in principle, that variations of traffic volume follow a GBM (geometric Brownian motion), which can be described in the following way:

$$d\theta = a\theta\, dt + \sigma\theta\, dz \qquad (1)$$

where,

θ: traffic volume
$d\theta$: differential increment of traffic
a: growth rate of traffic
dt: differential time interval
σ: traffic volatility
dz: increment of a Wiener process

Starting from Eq. (1), and applying Itô's lemma [1], the process followed by the natural logarithm of θ can be described as:

$$d(\ln\theta) = a' dt + \sigma\, dz \qquad (2)$$

where, $\ln\theta$ is the natural logarithm of traffic and $a' = a - \sigma^2/2$.

On the right-hand side of Eq. (2), the parameter a' is a constant drift term or growth parameter. It means that the logarithm of traffic has a growth of a' per unit of time. Regarding the second term, dz is the increment of a standard Wiener process, so that $dz = \varepsilon_t (dt)^{1/2}$, where, ε_t is a variable which is normally distributed with zero mean and unit standard deviation [2]. This second term, σdz, adds a noise or variability to the path followed by

Corresponding author: Antonio Sanchez Soliño, professor, research fields: public economics and transportation. E-mail: asanoli@ciccp.es.

the logarithm of traffic. The amount of this noise is σ times a standard Wiener process, so the process represented by Eq. (2) has a standard deviation of σ. This means that the variance rate (the variance per unit of time) of this process is σ^2 [3]. It is assumed that the parameter σ, which is called the traffic volatility, is also a constant.

The discrete version of Eq. (2) would be the following:

$$\Delta(\ln\theta) = a'\Delta t + \sigma\,\Delta z \qquad (3)$$

where,

E(Δz) = 0 [expected value of Δz]

E[$\Delta(\ln\theta)$] = $a'\Delta t$ [expected value of $\Delta(\ln\theta)$]

V[$\Delta(\ln\theta)$] = $\sigma^2\Delta t$ [variance of $\Delta(\ln\theta)$]

This means that the change in the logarithm of traffic is normally distributed over any time interval Δt (with mean $a'\Delta t$ and standard deviation $\sigma\sqrt{\Delta t}$), following a random walk with a drift. This assumption is frequently made for economic and financial variables. For stock prices, for example, the hypothesis of GBM is generally accepted, and it has been used for the development of the theory of options' valuation, since the initial works carried out by Black and Scholes [4] and Merton [5]. In the field of road traffic, this assumption has been made by Zhao et al. [6] to analyze the decision-making process in highway development.

However, the GBM hypothesis is not always evident. Pindyck and Rubinfeld [7], for example, have analyzed whether commodity prices follow this process. They found that, for very long time series (more than 100 years), detrended prices of crude oil and copper do not follow a random walk, but a mean-reverting process. However, and to the contrary, the hypothesis of a random walk cannot be rejected for the detrended prices of lumber.

In this paper, a test is performed for the hypothesis of a GBM for the evolution of traffic volume on toll highways. Series available for Spanish toll highways have been used, which, in most cases, cover a thirty-year period. In the following section a description is given of the methodology used for the analysis of the existence of unit roots in time series in general. The Dickey-Fuller approach has been applied, which is the most widely used test for this kind of analysis. Then this methodology has been applied for traffic series in Spanish toll highways and the results obtained have been examined. The limitations of the analysis carried out are considered and the possible application of the results is discussed. Finally, the main conclusions of the paper are summarized.

2. Unit Roots Analysis of Time Series

Suppose that Y_t is a random variable which evolves over time following an autoregressive process that can be described as:

$$Y_t = \rho Y_{t-1} + u_t \qquad (4)$$

where, u_t is a random error term. Now, the parameter ρ can be analyzed. If ρ is equal to 1, then it is said that a unit root exists, which means that Y_t is a non-stationary variable. In the opposite case (if $\rho \neq 1$) the Y_t variable would be stationary.

A constant drift term α can be added to Eq. (4), without changing the reasoning. The equation would then be:

$$Y_t = \alpha + \rho Y_{t-1} + u_t \qquad (5)$$

Eq. (5) can be rewritten in the following way:

$$Y_t - Y_{t-1} = \alpha + (\rho - 1)Y_{t-1} + u_t \qquad (6)$$

The parameter ρ in Eq. (6) can be estimated by using OLS (ordinary least squares), and calculating the t-statistic to test whether ρ is significantly different from 1. If the hypothesis that $\rho = 1$ cannot be rejected, then it can be said that the process has a unit root, and therefore the Y_t variable is non-stationary after detrending. However, if the true value of ρ is 1, then the OLS estimator is biased toward zero [7]. Then the use of OLS could lead us to incorrectly rejecting the non-stationarity hypothesis.

To solve this problem, Dickey and Fuller [8, 9] used a Monte Carlo simulation to calculate the correct critical values for the distribution of the t-statistic when

$\rho = 1$. The DF (Dickey-Fuller) test is subsequently the most widely used test to analyze the existence of a unit root in a given process.

To apply the DF test, Eq. (6) can be written as follows:

$$\Delta Y_t = \alpha + \beta Y_{t-1} + u_t \quad (7)$$

where, $\beta = \rho - 1$.

Now, the OLS method is applied to estimate the value of the parameter β (where the null hypothesis is that $\beta = 0$) and to calculate its t-statistic. The t-statistic thus obtained is then compared with the critical values calculated by Dickey-Fuller. In fact, the critical values obtained by other authors based on the DF methodology are used. For example, McKinnon [10] obtained the following critical values.

If the t-statistic obtained in our estimation is greater than the critical value, the hypothesis that $\beta = 0$ cannot be rejected and then it is not possible to reject that the process is non-stationary after detrending. Observe that all critical values are negative. Therefore, if the t-statistic obtained in our estimation is positive, the null hypothesis cannot be rejected (i.e., it cannot be rejected that the process is non-stationary).

In this kind of test, it is assumed that there is no serial correlation in the error term u_t. However, the process described by Eq. (7) may be non-stationary, even if there is serial correlation in u_t. As an extension of the methodology, serial correlation can be allowed now, by using the so-called ADF (augmented Dickey-Fuller) test. For that purpose, the model is expanded by adding the lagged dependent variable to the right-hand side of the equation, as follows:

$$\Delta Y_t = \alpha + \beta Y_{t-1} + \sum_{j=1}^{m} \lambda_j \Delta Y_{t-j} + u_t \quad (8)$$

where, λ_j represent the m parameters obtained in the regression analysis between the dependent variable ΔY_t and the same dependent variable with a lag of j periods (i.e., ΔY_{t-j}). For example, for annual data, if two lags are considered, the following expression would apply:

$$\Delta Y_t = \alpha + \beta Y_{t-1} + \lambda_1 \Delta Y_{t-1} + \lambda_2 \Delta Y_{t-2} + u_t \quad (9)$$

where, two terms have been added, on the right-hand side of the equation, that include the dependent variable with a lag of one year and two years (ΔY_{t-1} and ΔY_{t-2}, respectively). The number of lags considered in the analysis depends on the decision of the analyst and the kind of problem being analyzed.

The regression analysis to determine the parameters in Eq. (8) is made using OLS. The t-statistic obtained for the parameter β is then compared with the same critical values contained in the former (Table 1). Again, if the t-statistic obtained in our estimation is greater than the critical value, it cannot be rejected that $\beta = 0$ and that the process is non-stationary after detrending.

3. Results Obtained for Spanish Toll Highways

In this section, the methodology described above is applied, in both versions (the Dickey-Fuller and the Augmented Dickey-Fuller tests), for traffic series in Spanish toll highways. As a starting point, data collected by the public authority [11] which is in charge of the supervision of national toll highways are used. These highways have an average length of 134 km, and all of them are managed by private companies under concession contracts. These private companies are obliged to provide the relevant data to the said public authority, and this is published, and is available for researchers or for any person with an interest in the matter.

The AADT (annual average daily traffic) has been used in the research. By using annual data, the problem of seasonality in traffic volumes is avoided. The collected data are included in Appendix 1 in this paper.

Table 1 Critical values for t-statistic in DF unit roots tests.

Sample size	Significance level = 5%	Significance level = 10%
25	-3.00	-2.63
50	-2.93	-2.60
100	-2.89	-2.58
∞	-2.86	-2.57

In order to perform the DF test, the following variable is used: $Y_t = ln\ (\theta_t)$, where θ_t is the volume of traffic, in terms of AADT. Therefore, Eq. (7) is applied, where $\Delta Y_t = ln\ (\theta_t / \theta_{t-1})$. A regression analysis has been performed, using OLS to obtain the estimation of the parameter β and the t-statistic for that estimation for each highway. The results for the relevant t-statistics are included in the third column of Table 2.

These results can be compared with the critical values in Table 1. As the period of analysis is around thirty years in most cases, the critical values can be taken for a sample size equal to 25 in Table 1. It can then be seen that for significance levels of 5% and 10%, the null hypothesis (i.e., $\beta = 0$) cannot be rejected for any of the highways analyzed. This means that, according to the DF test, the hypothesis that traffic in Spanish toll highways follows a GBM process cannot be rejected.

The ADF test has also been performed, by using Eq. (8), where once again $\Delta Y_t = ln\ (\theta_t / \theta_{t-1})$. One lag and two lags have been taken for the analysis, which is considered to be sufficient in view of the results obtained.

With one lag, the regression analysis is applied using the following expression:

$$\Delta(\ln \theta_t) = \alpha + \beta(\ln \theta_{t-1}) + \lambda_1 \Delta(\ln \theta_{t-1}) + u_t \quad (10)$$

Here the parameter β is estimated and its t-statistic is then calculated.

With two lags, the relevant expression is analogous, and again, the estimation of the parameter β is carried out.

The relevant t-statistics for each highway are included in the fourth and fifth columns in Table 2. As it can be observed, making a comparison with the critical values in Table 1, the null hypothesis cannot be rejected for any of the highways and, subsequently, it cannot be rejected that traffic follows a GBM process. On the other hand, there is not a clear pattern in the values of the t-statistic with one lag and with two lags. For some highways, the t-statistic is nearer the critical value with two lags than with one lag, and in other cases it is the other way round.

4. Limitations of the Analysis and Application of the Results

According to the results obtained in the research described in this paper, the GBM hypothesis for traffic volume cannot be rejected. However, one should be aware of the limitations of the analysis. These results provide only weak evidence in favor of the hypothesis that traffic actually follows a GBM. In fact, the results could be different for longer periods of analysis, as the results obtained by Pindyck and Rubinfeld [7] show for the case of commodity prices. Unfortunately, longer traffic series are not normally available.

Table 2 Results of unit roots tests for traffic series.

Name of highway	Period of analysis	DF test t-statistic	ADF test (one lag) t-statistic	ADF test (two lags) t-statistic
Villalba-Adanero	1974-2007	0.6692	0.7843	0.8895
Zaragoza-Mediterráneo	1976-2007	-1.5419	-0.8040	-1.2278
Sevilla-Cádiz	1974-2007	1.1856	0.3834	0.0468
Montmeló-La Junquera	1974-2007	0.6792	-0.5228	-0.2624
Barcelona-Tarragona	1974-2007	-1.4503	-1.7321	-1.2479
Montmeló-Papiol	1978-2007	-0.7704	-1.2694	-1.4408
Bilbao-Zaragoza	1978-2007	0.8923	-0.4104	-0.1090
Burgos-Armiñón	1978-2007	-2.0183	-0.6922	-0.3636
León-Campomanes	1983-2007	0.2918	-1.3950	-1.1522
Tarragona-Valencia	1974-2007	-0.3193	-0.9579	-0.8513
Valencia-Alicante	1976-2007	-1.3442	-0.8213	-0.1557

Nevertheless, the results are robust, in the sense that the relevant tests have been applied to all the national toll highways in Spain, and the hypothesis could not be rejected in any of these. Furthermore, it would be possible to generalize the results, since there are various types of highways in the sample used: some of these are coastal highways (with a clearly tourist nature), others are interurban highways and, finally, other highways have some of the features of metropolitan transportation networks.

Another limitation of the analysis is the assumption of a constant volatility of traffic. For the estimation of this volatility, historic data in Spanish toll highways have been used. A simple procedure to calculate traffic volatility is as follows:

Suppose a traffic series for a certain highway: θ_1, θ_2, θ_3, … θ_n, where θ_i is the traffic volume in year *i*. Then the following variable is defined: $x_i = \Delta \ln(\theta_i) = \ln(\theta_i / \theta_{i-1})$, and $\bar{x}$ is obtained as the mean of $x_1, x_2, \ldots x_n$. The volatility of traffic, defined as the standard deviation of the sample $x_1, x_2 \ldots x_n$, would then be as follows:

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{i=1}^{i=n}(x_i - \bar{x})^2} \qquad (11)$$

Using this definition, the volatility for traffic in each toll highway in Spain has been obtained, starting from data contained in Appendix 1. It has been assumed that the volatility in each highway remains constant, but it may in fact change over time. However, it has been observed that traffic volatility in toll highways is greater over the first years of the concession, before becoming smaller and more stabilized. This means that, if with sufficiently long time series (say twenty years) the hypothesis of a constant volatility in the future can be assumed. In the present case, it has been obtained that the volatility of traffic in Spanish toll highways (for annual data) tends towards an average value close to 0.075.

The hypothesis of the Geometric Brownian Motion given by Eq. (1) can be applied for the valuation of toll highways concessions. In this kind of concession, both the forecast of future traffic and the measure of the risks involved are essential for the appraisal of the business. The calculation of the value of the volatility of traffic (probably the most important source of uncertainty in a toll highway) allows for using the model to build a confidence interval for the traffic forecast.

Besides, the terms of reference in toll highway concessions (and the concession contracts) often contain certain clauses that allow for a degree of operational flexibility in the management of the business. The valuation of this kind of clauses in contracts can be carried out using a real options approach, a methodology based on the development of the theory of financial options. Under this approach, traffic volume on the highway (for which a GBM process is assumed) is used as the underlying asset in an option contract. Options that are embedded in the concession agreement are thus calculated as a derivative of the traffic volume. This means that traffic is treated as the source of uncertainty that determines the value of the options.

The possible exercise of this series of rights represents an added value for the project which is not captured by the traditional procedures of valuation. The habitual practice of calculating the NPV (net present value) of the project by means of the discount of cash flows, leads to erroneous results when the project incorporates a certain degree of flexibility.

Therefore, the theory of real options is an alternative tool for the correct valuation of toll highway concessions, under the hypothesis that the variations of traffic volume follow a GBM like the one described in former Eq. (1).

5. Conclusions

The main result of the research is that the hypothesis that traffic follows a generalized Wiener process (or so-called Geometric Brownian Motion) in Spanish toll highways cannot be rejected. In other words, the evidence found leads to the conclusion that the non-stationarity hypothesis for traffic cannot be rejected, but have to bear in mind that this is only a weak evidence in favor of the hypothesis that traffic actually follows a non-stationary process.

The GBM hypothesis can be applied to the valuation of toll highway concessions. The terms of contracts in toll highway concessions often contain certain clauses that allow for a degree of operational flexibility in the management of the business. The valuation of these kinds of clauses in contracts can be carried out using a real options approach. The full description of this methodology is beyond the scope of this paper [12], but some of the options that usually appear in concession contracts have been quoted: minimum traffic guarantees (traffic floors), maximum traffic limitations (traffic caps), extension of the concession, anticipated reversion, granting of public subsidies, public participation loans, etc.. These mechanisms reduce the variability of the project cash-flows, and allow for more flexibility and a better management of the concession based on the contingency of future events. The theory of real options is an alternative tool for the correct valuation of highway concessions when these kinds of rights are present in concession contracts, and the results of our research allow for the application of this methodology under the assumption that the evolution of traffic volume follows a Geometric Brownian Motion.

References

[1] K. Itô, On stochastic differential equations, Memoirs of the American Mathematical Society 4 (1951) 1–51.

[2] A. K. Dixit and R. S. Pindyck, Investment under Uncertainty, Princeton University Press, Princeton, New Jersey, 1994.

[3] J. C. Hull, Options, Futures and Other Derivatives, Pearson Prentice Hall, 6th ed., Upper Saddle River, New Jersey, 2006.

[4] F. Black and M. Scholes, The pricing of options and corporate liabilities, Journal of Political Economy 81 (1973) 637–654.

[5] R. C. Merton, Theory of rational option pricing, Bell Journal of Economics and Management Science 4 (1973) 141–183.

[6] T. Zhao, S. K. Sundararajan and C. Tsen, Highway development decision-making under uncertainty: a real options approach, Journal of Infrastructure Systems 10 (1) (2004) 23–32.

[7] R. S. Pindyck and D. L. Rubinfeld, Econometric Models and Economic Forecasts (4th ed.), McGraw-Hill, Boston, 1998.

[8] D. A. Dickey and W. A. Fuller, Distribution of the estimators for autoregressive time-series with a unit root, Journal of the American Statistical Association 74 (1979) 427–431.

[9] D. A. Dickey and W. A. Fuller, Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root, Econometrica 49 (1981) 1057–1072.

[10] J. McKinnon, Critical Values for Cointegration Tests, Discussion Paper 90.4, University of California, San Diego, 1990.

[11] Government Office for Toll Highways, 2007 Report, Transportation and Public Works Department, Madrid, 2008. (in Spanish)

[12] A. L. L. Galera and A. S. Soliño, A real options approach for the valuation of highway concessions, Transportation Science 44 (3) (2010) 416–427.

Appendix 1 Traffic Data

AADT (annual average daily traffic) in Spanish Toll Highways

Year /Highway	Villalba-Adanero	Zaragoza-Mediterráneo	Sevilla-Cádiz	Montmeló-La Junquera	Barcelona-Tarragona	Montmeló-Papiol	Bilbao-Zaragoza	Burgos-Armiñón	León-Campomanes	Tarragona-Valencia	Valencia-Alicante
1974	7,258		3,171	14,728	15,377					5,603	
1975	7,817		3,382	13,354	15,367					5,776	
1976	8,168	5,276	3,017	13,002	16,630					6,002	3,563
1977	6,690	6,179	3,039	13,925	19,760					6,870	4,148
1978	7,796	6,439	3,470	15,823	22,811	9,389	4,689	2,479		7,524	5,183
1979	8,455	7,001	3,681	15,859	23,659	6,875	4,169	3,604		7,828	5,874
1980	8,326	7,053	3,774	15,026	24,565	7,480	4,606	4,060		7,773	6,059
1981	8,380	6,920	3,999	15,557	23,575	6,470	4,681	5,622		7,590	6,258
1982	8,355	6,761	3,929	15,948	23,613	6,723	4,754	4,966		7,455	6,147
1983	8,283	6,607	3,629	15,934	23,166	6,861	4,374	4,611	2,494	7,233	6,071
1984	8,452	6,489	3,417	16,478	23,597	6,944	4,281	4,970	2,049	7,178	6,124
1985	8,810	6,659	3,632	17,099	24,857	7,352	4,275	5,142	2,141	7,596	6,933
1986	9,478	7,181	3,959	18,892	27,154	27,404	4,433	5,487	2,275	8,514	7,240
1987	10,360	8,119	4,525	21,282	30,793	31,558	4,874	5,994	2,445	9,707	8,316
1988	11,420	9,387	5,282	23,671	34,963	42,998	5,617	6,832	2,768	10,873	9,376
1989	12,929	11,423	6,350	26,296	39,624	51,004	6,494	7,777	3,233	12,336	10,563
1990	14,005	12,127	6,835	26,660	40,618	52,226	6,870	8,294	3,661	12,501	12,027
1991	15,610	12,327	7,791	27,802	42,080	54,489	7,118	8,954	4,254	13,043	12,663
1992	16,415	12,174	9,214	28,488	41,379	49,997	7,052	9,403	4,256	12,894	12,595
1993	16,504	11,425	8,005	28,124	40,152	45,884	6,956	9,680	4,199	12,336	12,085
1994	16,628	10,958	7,978	28,554	41,123	46,960	6,930	10,172	4,583	12,469	12,301
1995	17,358	11,309	7,648	28,509	43,270	48,724	7,013	11,026	4,680	12,907	12,313
1996	17,866	11,027	7,434	27,076	43,530	52,453	7,038	11,430	4,718	13,070	12,423
1997	18,687	11,423	7,828	29,021	45,677	58,635	7,343	12,198	4,995	14,186	13,207
1998	20,715	12,377	10,101	30,717	47,799	63,220	8,082	13,696	5,659	16,692	16,271
1999	22,918	13,350	11,825	33,815	47,089	70,219	9,002	15,161	6,320	19,092	18,987
2000	24,325	14,870	13,300	35,955	51,278	83,935	10,623	16,605	6,642	20,453	21,225
2001	25,482	15,206	15,218	37,901	53,721	90,218	11,742	18,062	7,433	22,004	23,409
2002	27,238	15,594	16,534	40,464	55,994	92,636	12,196	19,348	7,679	22,796	24,968
2003	28,662	15,464	17,897	41,756	57,782	95,712	12,844	20,101	8,048	23,396	26,640
2004	30,301	15,350	19,642	43,324	59,053	99,460	13,503	21,072	8,736	23,932	27,302
2005	30,770	14,744	21,859	44,918	60,342	111,353	13,542	21,206	9,006	23,482	28,180
2006	32,998	15,273	24,244	47,122	63,683	115,607	14,177	22,209	9,683	25,215	29,207
2007	34,414	15,541	24,951	49,180	66,217	118,519	14,712	23,937	10,288	25,110	29,411