THE BILATERAL RELATIONSHIP BETWEEN CONSUMPTION AND GDP IN MEXICO AND THE US: A COMMENT

GOMEZ-ZALDIVAR, Manuel*
VENTOSA-SANTAULARIA, Daniel

Abstract

This article presents a critical appraisal of three different econometric techniques commonly employed to analyze causal relationships among economic series. Our results indicate that the empirical application of the Granger causality test, the Engle-Granger cointegration test and the Hausman test for causality performed with small samples suffers severe size distortions, and therefore that the results should be taken with caution. Furthermore, we show that these tests produce better results if the series are differentiated. Our results are applied to the series for consumption and GDP in Mexico and the US and suggest that these series are cointegrated in the case of the US only (causality and cointegration are different). We comment upon these results in relation to the conclusions of Guisan (2004), and other related studies, in which several methods are used to analyze the bilateral causality between consumption and GDP in Mexico and the US and where it was found that cointegration and Granger causality tests may fail to detect the true causal relationships.

JEL classification numbers: C5, C52, E2, O51, O57

Keywords**:** Granger causality test, cointegration, Mexico, the US, consumption and Gross Domestic Product.

1. Introduction

Guisan (2004) analyzed the results of several tests to detect the causal relationship existing between real consumption and real GDP in Mexico and the United States: (i) the Granger causality test, (ii) the modified Granger causality test, (iii) the Engle-Granger cointegration test, and (iv) the Hausman test for causality. The main conclusions are:

1. Granger Causality:
   a. There is no evidence of Granger causality between consumption and GDP in Mexico. Hence, the Granger test failed to detect causality in this country.
   b. There is evidence of bilateral Granger causality between consumption and GDP in the US. Hence, the Granger test did not fail to detect causality in this country.
2. Modified Granger Causality:
   a. There is evidence of bilateral Granger causality between consumption and GDP in both countries. Therefore, the modified version of the Granger test leads to better results than the former test.
3. Cointegration:
   a. The results of the cointegration test are ambiguous and did not allow us to reject systematically the null hypothesis of no cointegration, although there is more evidence in favor of cointegration than there is against it.

* Manuel Gomez, Universidad de Guanajuato, Escuela de Economía. Email: manuel.gomez@ugto.org . Corresponding author: Daniel Ventosa.Universidad de Guanajuato, Escuela de Economía. Email: daniel@ventosa-santaularia.com. UCEA Campus Marfil, Col. El Establo, Guanajuato, Gto. CP. 36250

   b. There is evidence of a cointegrated relationship between consumption and GDP in the US.
4. Hausman Test for Causality:
   a. There is mixed evidence of causality "à la Hausman" in both countries.

In this paper, we discuss different features of Guisan's work. In particular, we analyze and extend her results in several directions: firstly, we apply a set of well-known unit-root tests to the variables under examination to determine which Data Generating Process (DGP), if any, best suits them. We do this because the performance and reliability of the tests employed depend upon the statistical properties of each variable. Secondly, we show by means of Monte Carlo experiments that the reliability of the tests employed decreases significantly when a short sample is used. Thirdly, we propose additional test procedures that strengthen the inference to be drawn from such tests. In particular, we perform Granger Causality (GC) tests with series in first differences [causality] and estimate an Error-Correction Model [Cointegration] for both countries.
Our results suggest that there is a cointegrated relationship between consumption and GDP in the US as well as an adjustment of its GDP when the variables deviate from their long-term equilibrium relationship.
The rest of the paper is organized as follows: In Section 1, we determine which DGP best fits the series for consumption and GDP in Mexico and the US. In Section 2, we present the results of the Monte Carlo experiments to show that the Granger causality test, Engle-Granger cointegration test (EG) and the Hausman test for causality suffer from severe size distortions when the series have a trending mechanism, whether the latter is stochastic or deterministic. Section 3 presents the results of the causality tests for the series for consumption and GDP in Mexico and the US in first differences. Section 4 shows the results of an Error-Correction Model (ECM) applied to these series. Conclusions are drawn in Section 5.

1. "DGP-ification" of the Consumption and GDP Series.

In this section we perform a well-known set of unit-root tests to determine which DGP, if any of those traditionally used in this literature, best suits the series for Mexico and the US. Equations 1–5 show the DGPs that are potential representations for consumption and GDP in the US and Mexico.

$$y_t = Y_0 + \sum_{i=1}^{t} u_{yi} \tag{1}$$

$$y_t = Y_0 + \mu_y t + \sum_{i=1}^{t} u_{yi} \tag{2}$$

$$y_t = Y_0 + \mu_y t + \theta_y DT_{yt} + \sum_{i=1}^{t} u_{yi} \tag{3}$$

$$y_t = \mu_y + \beta_y t + u_{yt} \tag{4}$$

$$y_t = \mu_y + \beta_y t + \gamma_y DT_{yt} + u_{yt} \tag{5}$$

where DT_{yt} is a dummy variable allowing changes in the slope, that is, $DT_{yt} = (t - T_b)1(t > T_b)$, where $1(\cdot)$ is the indicator function, and T_b is the unknown date of the break in y. We assume that the innovations, u_{yt}, are $i.i.d. N\left(0, \sigma_y^2\right)$.

DGP (1) represents a random-walk process, DGP (2) a process with stochastic and deterministic trends, DGP (3) a process with both stochastic and deterministic trends, and a break in the deterministic trend, DGP (4) a trend-stationary process and DGP (5) is a broken trend-stationary process.

1.1. Unit-Root Tests

Table 1 shows the results of applying the Augmented Dickey-Fuller tests (ADF), Dickey-Fuller GLS (DF-GLS), Phillips-Perron test (PP) and the Ng-Perron test. In all cases, the number of lags used to control for autocorrelation were automatically selected by the Schwarz Information Criterion (SIC). On the one hand, the results in columns 2, 3 and 4 show that it is not possible to reject the null of unit root for all the series. On the other, the Ng-Perron test finds mixed evidence regarding the stationarity of consumption in Mexico and GDP in US; furthermore, it shows that consumption in the US could be considered stationary. Finally, as in the previous three tests, it is not possible to reject the null of unit root for Mexican GDP. The inference drawn from these tests is not conclusive for the series except that of Mexican GDP, which seems to contain a unit root.

Table 1: Unit-Root Tests

Test / Variable	ADF[1]	DF-GLS[2]	PP[2]	Ng-Perron[2]			
				MZa	MZt	MSB	MPT
Consumption Mexico	1.48	-2.69	-2.07	-14.17*	-2.64*	0.18	6.53*
Consumption US	3.07	-2.37	-1.18	-18.57**	-2.90*	0.15**	5.76*
GDP Mexico	1.81	-1.90	-2.15	-5.24	-1.61	0.30	17.38
GDP US	4.93	-2.62	-1.94	-15.80*	-2.74	0.17*	6.12*

[1] Specification of the DF test: No drift; [2] With drift and trend. *,** and *** denote rejection of the null hypothesis at 10%, 5%, and 1%, respectively.

1.2. Unit-Root Tests allowing for Structural Breaks.

In this section, we employ unit-root tests that allow for structural breaks either under the null hypothesis (DGP 3) or under the alternative (DGP 5). Table 2 presents the results of applying the Zivot and Andrews test to our series; structural breaks are allowed in the intercept, the deterministic trend or both. This is a popular test that discriminates between the null of unit root and the alternative of stationarity with structural breaks. The last column of this table shows the results of the Gómez and Ventosa-Santaulària test (GVS).[2] Zivot and Andrews' test fails to reject the null hypothesis of unit root for all the variables; therefore, we can conclude that the series are not being generated by DGPs 4 and 5. The

[2] This formal statistical procedure distinguishes between the null hypothesis of unit root and that of unit root with drift (with a potential break). This procedure is asymptotically robust with regard to autocorrelation and takes into account a potential single structural break. See Gómez and Ventosa-Santaulària (2008).

GVS test identifies the presence of a drift for US consumption and US GDP: this suggests that DGP 2 could be generating both series. Finally, the series of consumption and GDP for Mexico do not have a deterministic trend; consequently, they seem to be better represented by DGP 1.

Table 2: Unit-Root Tests allowing for Structural Breaks

Test / Variable	Zivot and Andrews[1]			GVS[2] (R^2)
	Intercept[2]	Trend[2]	Both[2]	
Consumption Mexico	-3.76	-3.02	-3.53	0.65
Consumption US	-3.62	-4.06	-3.93	0.97***
GDP Mexico	-3.79	-2.63	-3.43	0.74
GDP US	-3.98	-4.35	-4.18	0.96***

[1] t-ratio associated to the autoregressive term; Critical Values provided by Zivot and Andrews (1992). [2] *** and *** denote rejection of the null hypothesis at 10%, 5%, and 1%, respectively.

2. Monte Carlo Simulations

The previous section showed that Mexican variables can be seen as driftless unit roots whilst the US series behave more like unit roots with drift. These results are used in the present section to design Monte Carlo experiments to analyze the accuracy of the tests employed by Guisan (2004) when the empirical application is performed in small samples. The tables presented below assume different DGPs for the series; these DGPs were chosen according to the findings of the previous section as well as to previous results in this literature.[3] There is evidence that the GC test, EG cointegration test and Hausman test for causality may suffer from severe size distortions when applied in small samples.

2.1. Monte Carlo Evidence with Trended Series.

Table 3 shows the results of applying the GC test using the DGPs found in section one. According to the previous section, the Mexican variables behave as unit roots, whilst the US series appear to be unit root with drift. Performing this test with such DGPs generates severe size distortions, especially in small samples (*T=30*). We should expect the rejection rates to be around 5%, but the Monte Carlo simulation exhibits rejection rates of between 12% and 17%.

[3] The parameter values of the DGPs employed in all the Monte Carlo experiments as well as the number of replications can be found in the appendix of this article.

Table 3: Granger Causality Test*

DGP x \ DGP y	Driftless Unit Root	Unit Root with Drift	Trend Stationary	Trend Stationary with Break
Sample Size, T=30				
Driftless Unit Root	0.1705	0.0927	0.1516	0.2072
Unit Root with Drift	0.2986	0.1196	0.2594	0.4481
Trend Stationary	0.0867	0.0816	0.0908	0.1009
Trend Stationary with Break	0.0827	0.0827	0.0857	0.0934
Sample Size, T=50				
Driftless Unit Root	0.1529	0.0740	0.2853	0.3856
Unit Root with Drift	0.2827	0.1048	0.6885	0.9498
Trend Stationary	0.0911	0.0693	0.1777	0.2166
Trend Stationary with Break	0.0825	0.0694	0.1258	0.1412
Sample Size, T=100				
Driftless Unit Root	0.1481	0.0594	0.5571	0.4926
Unit Root with Drift	0.2751	0.0858	1.0000	1.0000
Trend Stationary	0.1140	0.0575	0.8344	0.7225
Trend Stationary with Break	0.1001	0.0582	0.5502	0.4064
Sample Size, T=150				
Driftless Unit Root	0.1451	0.0603	0.5740	0.4574
Unit Root with Drift	0.2709	0.0841	1.0000	1.0000
Trend Stationary	0.1408	0.0563	0.9967	0.9502
Trend Stationary with Break	0.1172	0.0546	0.9004	0.6442

* Number of replications: 10,000; rejection rate of the null hypothesis of no Granger causality; Level: 0.05

When the adequate DGP is a deterministic trend with a break, this test suffers from similar size distortions. In fact, when the sample size is larger, using one (or both) of the variables generated by DGPs (4) or (5) would aggravate such size distortions. These simulations are in line with the findings of Ventosa-Santaulària and Vera-Valdés (2008). The authors studied the asymptotic properties of the GC test for similar specifications when the variables are mean-stationary with level breaks and trend-stationary with trend breaks processes; they found that the GC test may lead to erroneous inference and reject (asymptotically) the null hypothesis of no Granger causality between otherwise independent variables.

A potential solution for this problem—as will be proposed in the next section—is to perform the GC tests with variables in first differences. Table 4 shows the Monte Carlo results of performing the Modified Granger Causality (MGC) test using similar DGPs. The size distortions that occur with the MGC test are even more significant than they are with the GC test. In this case, rejection rates are 16% and 9%, and do not decrease, even for samples as large as T=150. Size distortions are even more important when the variables are trend-stationary or/and broken trend-stationary processes. In this case, rejection rates reach 100%.

Table 4: Modified Granger Causality Test*

DGP y / DGP x	Driftless Unit Root	Unit Root with Drift	Trend Stationary	Trend Stationary with Break
Sample Size, T=30				
Driftless Unit Root	0.1608	0.0525	0.0398	0.0754
Unit Root with Drift	0.3482	0.0875	0.0882	0.2258
Trend Stationary	0.0132	0.0119	0.0114	0.0146
Trend Stationary with Break	0.0120	0.0120	0.0105	0.0117
Sample Size, T=50				
Driftless Unit Root	0.1632	0.0532	0.1477	0.3027
Unit Root with Drift	0.3501	0.0910	0.4884	0.8969
Trend Stationary	0.0195	0.0088	0.0370	0.0573
Trend Stationary with Break	0.0145	0.0085	0.0229	0.0297
Sample Size, T=100				
Driftless Unit Root	0.1682	0.0485	0.5652	0.5505
Unit Root with Drift	0.3490	0.0842	1.0000	1.0000
Trend Stationary	0.0519	0.0071	0.6358	0.5913
Trend Stationary with Break	0.0337	0.0086	0.3040	0.2506
Sample Size, T=150				
Driftless Unit Root	0.1715	0.0517	0.6309	0.5521
Unit Root with Drift	0.3518	0.0949	1.0000	1.0000
Trend Stationary	0.0969	0.0061	0.9938	0.9491
Trend Stationary with Break	0.0635	0.0083	0.8499	0.6008

* Number of replications: 10,000; rejection rate of the null hypothesis of no Granger causality; Level: 0.05

Tables 5, 6 and 7 show the Monte Carlo results of performing the EG cointegration test with and without one lag and the Hausman test for causality, respectively. The Monte

Carlo experiments confirm that these tests draw correct inference when the variables are unit root and/or unit root with drift, even when the sample size is small, for example, 30 observations. Nevertheless, there are severe distortions whenever one or both variables include a deterministic trend. Such size distortions worsen the larger the sample size.

Table 5: Engle-Granger Cointegration Test, No Lags*

DGP y / DGP x	Driftless Unit Root	Unit Root with Drift	Trend Stationary	Trend Stationary with Break
Sample Size, T=30				
Driftless Unit Root	0.0590	0.0320	0.9879	0.9737
Unit Root with Drift	0.0752	0.0752	0.9949	0.9946
Trend Stationary	0.0259	0.0017	0.9772	0.9479
Trend Stationary with Break	0.0235	0.0002	0.9753	0.9419
Sample Size, T=50				
Driftless Unit Root	0.0477	0.0277	0.9999	0.9979
Unit Root with Drift	0.0636	0.0665	1.0000	1.0000
Trend Stationary	0.0434	0.0359	0.9997	0.9978
Trend Stationary with Break	0.0269	0.0052	0.9997	0.9956
Sample Size, T=100				
Driftless Unit Root	0.0508	0.0257	1.0000	0.9993
Unit Root with Drift	0.0641	0.0614	1.0000	1.0000
Trend Stationary	0.4841	0.9990	1.0000	1.0000
Trend Stationary with Break	0.3181	0.6906	1.0000	1.0000
Sample Size, T=150				
Driftless Unit Root	0.0516	0.0282	1.0000	0.9951
Unit Root with Drift	0.0642	0.0617	1.0000	1.0000
Trend Stationary	0.6255	1.0000	1.0000	1.0000
Trend Stationary with Break	0.5760	0.9999	1.0000	1.0000

* Number of replications: 10,000; rejection rate of the null hypothesis of no cointegration; Level: 0.05.

As stated in Noriega and Ventosa-Santaulària (2007), we should bear in mind that the EG cointegration test between variables that include deterministic trends and/or breaks may provide spurious results, that is, there may be considerable size and power distortions. In this case, we should further bear in mind that the concept of cointegration refers to a long-run equilibrium relationship between the variables. There is no evident link between causality and cointegration.

We would therefore suggest the estimation of an Error-Correction Model (ECM). With the ECM, we should be able to draw inference concerning which variables adjust whenever there is a short-run disequilibrium. Although this could not be formally regarded as causality, we would at least know which variable "moves first" after a shock occurs.

Table 6: Engle-Granger Cointegration Test, One Lag*

DGP y / DGP x	Driftless Unit Root	Unit Root with Drift	Trend Stationary	Trend Stationary with Break
Sample Size, T=30				
Driftless Unit Root	0.0469	0.0215	0.7052	0.6028
Unit Root with Drift	0.0721	0.0692	0.7962	0.7955
Trend Stationary	0.0160	0.0000	0.5920	0.4292
Trend Stationary with Break	0.0153	0.0000	0.5800	0.4097
Sample Size, T=50				
Driftless Unit Root	0.0409	0.0175	0.9194	0.7258
Unit Root with Drift	0.0551	0.0587	0.9928	0.9925
Trend Stationary	0.0158	0.0000	0.8836	0.6123
Trend Stationary with Break	0.0143	0.0000	0.8455	0.4957
Sample Size, T=100				
Driftless Unit Root	0.0483	0.0246	0.9660	0.6176
Unit Root with Drift	0.0601	0.0620	1.0000	1.0000
Trend Stationary	0.0744	0.3638	0.9998	0.9941
Trend Stationary with Break	0.0309	0.0054	0.9953	0.7785
Sample Size, T=150				
Driftless Unit Root	0.0498	0.0247	0.9261	0.4887
Unit Root with Drift	0.0619	0.0598	1.0000	1.0000
Trend Stationary	0.3231	1.0000	1.0000	1.0000
Trend Stationary with Break	0.2067	0.5862	1.0000	0.9921

* Number of replications: 10,000; rejection rate of the null hypothesis of no cointegration; Level: 0.05.

Table 7: Hausman Test for Causality*

DGP x \ DGP y	Driftless Unit Root	Unit Root with Drift	Trend Stationary	Trend Stationary with Break
Sample Size, T=30				
Driftless Unit Root	0.0561	1.000	0.7845	0.7846
Unit Root with Drift	0.0593	0.6909	0.7632	0.7644
Trend Stationary	0.0584	1.0000	0.7788	0.7795
Trend Stationary with Break	0.0585	1.0000	0.7788	0.7790
Sample Size, T=50				
Driftless Unit Root	0.0493	1.0000	0.9643	0.9649
Unit Root with Drift	0.0536	0.9124	0.9598	0.9602
Trend Stationary	0.0525	1.0000	0.9642	0.9641
Trend Stationary with Break	0.0522	1.0000	0.9642	0.9643
Sample Size, T=100				
Driftless Unit Root	0.0499	1.0000	1.0000	1.0000
Unit Root with Drift	0.0508	0.9972	0.9999	1.0000
Trend Stationary	0.0507	1.0000	1.0000	1.0000
Trend Stationary with Break	0.0507	1.0000	1.0000	1.0000
Sample Size, T=150				
Driftless Unit Root	0.0508	1.0000	1.0000	1.0000
Unit Root with Drift	0.0520	0.9998	1.0000	1.0000
Trend Stationary	0.0510	1.0000	1.0000	1.0000
Trend Stationary with Break	0.0509	1.0000	1.0000	1.0000

* Number of replications: 10,000; rejection rate of the null hypothesis of no Hausman causality; Level: 0.05.

2.2. Monte Carlo Evidence with Variables in First Differences

It is not surprising that causality tests—when applied to integrated variables—yield poor results; however, in practice, in the case of variables integrated of order one, such tests applied to stationary first differences, may also fail to accept true causal relationships and reject untrue ones, as is the case in the following results presented by Guisan (2001):

Percentages of cointegration acceptation for models in levels and first differences between real consumption and real GDP in 25 OECD countries for the period 1961-97

Table 7bis.*,**

Summary of results**	Levels	First Differences
% of Own Cointegration McKinnon EG	0%	88%
% of Own Cointegration McKinnon ADF	84%	100%
% of Cross Cointegration McKinnon EG	19%	23%
% of Cross Cointegration McKinnon ADF	66%	96%

* Source: Guisan (2001). ** The author notes that both EG and ADF tests perform better with variables in first differences—compared to levels—in detecting true causality between the consumption and GDP of the own country (i.e. the percentages of acceptance of the true hypothesis are higher), but performs worse in first differences than in levels to reject the untrue hypothesis of a causal relationship between the crossed variables of different countries (i.e. the percentages of acceptance of the untrue hypothesis are higher).

Tables 8, 9 and 10 show the Monte Carlo results for each of the tests; the variables have been first-differenced. These results show that size distortions are considerably reduced for the GC and MGC tests for all DGP combinations. There is no relevant improvement in working with differenced variables when the Hausman test for causality test is used.

Table 8: Granger Causality Test, Variables in First Differences*

DGP y / DGP x	Driftless Unit Root	Unit Root with Drift	Trend Stationary	Trend Stationary with Break
Sample Size, T=30				
Driftless Unit Root	0.0861	0.0833	0.0832	0.0827
Unit Root with Drift	0.0836	0.0834	0.0863	0.0859
Trend Stationary	0.0825	0.0807	0.0973	0.0970
Trend Stationary with Break	0.0828	0.0809	0.0967	0.0969
Sample Size, T=50				
Driftless Unit Root	0.0669	0.0689	0.0669	0.0665
Unit Root with Drift	0.0689	0.0678	0.0729	0.0724
Trend Stationary	0.0679	0.0666	0.0870	0.0869
Trend Stationary with Break	0.0674	0.0665	0.0869	0.0869

Number of replications: 10,000; rejection rate of the null hypothesis of no Granger causality; Level: 0.05.

Table 9: Modified Granger Causality Test, Variables in First Differences*

DGP y / DGP x	Driftless Unit Root	Unit Root with Drift	Trend Stationary	Trend Stationary with Break
Sample Size, T=30				
Driftless Unit Root	0.0119	0.0115	0.0118	0.0118
Unit Root with Drift	0.0111	0.0110	0.0099	0.0100
Trend Stationary	0.0108	0.0129	0.0344	0.0346
Trend Stationary with Break	0.0110	0.0130	0.0344	0.0343
Sample Size, T=50				
Driftless Unit Root	0.0087	0.0077	0.0082	0.0083
Unit Root with Drift	0.0092	0.0111	0.0086	0.0085
Trend Stationary	0.0088	0.0071	0.0302	0.0302
Trend Stationary with Break	0.0090	0.0070	0.0301	0.0301

* Number of replications: 10,000; rejection rate of the null hypothesis of no Granger causality; Level: 0.05.

Table 10: Hausman Causality Test, Variables in First Differences*

DGP y / DGP x	Driftless Unit Root	Unit Root with Drift	Trend Stationary	Trend Stationary with Break
Sample Size, T=30				
Driftless Unit Root	0.0551	0.7618	0.4118	0.4047
Unit Root with Drift	0.0552	0.7612	0.4122	0.4066
Trend Stationary	0.0523	0.7522	0.4170	0.4110
Trend Stationary with Break	0.0524	0.7522	0.4170	0.4113
Sample Size, T=50				
Driftless Unit Root	0.0561	0.9627	0.6600	0.6514
Unit Root with Drift	0.0538	0.9619	0.6564	0.6479
Trend Stationary	0.0857	0.9604	0.6614	0.6542
Trend Stationary with Break	0.0567	0.9604	0.6612	0.6540

* Number of replications: 10,000; rejection rate of the null hypothesis of no Hausman causality; Level: 0.05.

3. Causality Tests with Variables in First Differences

We now use Guisan's (2004) data set to draw inference concerning causality. The strategy was advanced earlier in this work: differencing the series is appropriate when dealing with non-stationarity.

Table 11: Causality Tests, Variables in First Differences [Guisan data set]

Country	Independent - Dependent	Hausman Causality		Granger Causality		Modified Granger Causality	
		t-stat.	p-value	F-stat.	p-value	F-stat.	p-value
Mexico	Consumption-GDP	-0.79	0.21	0.37	0.69	0.09	0.76
	GDP-Consumption	-1.21	0.11	0.08	0.91	0.59	0.44
US	Consumption-GDP	-2.41	0.00	4.08	0.03	4.18	0.05
	GDP-Consumption	-1.41	0.07	3.05	0.06	0.74	0.39

The results in Table 11 show that the tests, using variables in first differences, fail to detect causality: there is no evidence of Granger causality or Hausman causality between the Mexican variables at the 5% level. In the case of the US variables, all the tests reveal evidence in favor of causality from consumption to GDP, but not vice-versa.

4. Error-Correction Model with Variables in First Differences

To formally implement the ECM, consider equation (6), which represents the long-run equilibrium relationship between consumption and GDP, $c_{z,t}$ and $y_{z,t}$ respectively,

$$\begin{aligned} c_{z,t} &= \alpha_z + \beta_z y_{z,t} + \varepsilon_{z,t} \\ ECM_{z,t} &= c_{z,t} - \alpha_z - \beta_z y_{z,t} \end{aligned} \tag{6}$$

where $z = mx, usa$,

If $c_{z,t}$ y $y_{z,t}$ are $CI(1,1)$, the variables have an error-correction form

$$\begin{aligned} \Delta c_{z,t} &= \gamma_{z,1} + \theta_{z,1} ECM_{z,t-1} + \sum_{s=1}^{m_1^i} \delta_{z1s} \Delta c_{z,t-s} + \sum_{s=1}^{m_2^i} \phi_{z1s} \Delta y_{z,t-s} + u_{z,1t} \\ \Delta y_{z,t} &= \gamma_{z,2} + \theta_{z,2} ECM_{z,t-1} + \sum_{s=1}^{m_3^i} \delta_{z2s} \Delta y_{z,t-s} + \sum_{s=1}^{m_4^i} \phi_{z2s} \Delta c_{z,t-s} + u_{z,2t} \end{aligned} \tag{7}$$

where $\theta_{z,1}$ and $\theta_{z,2}$ are interpreted as "speeds of adjustment to a short-run disequilibrium"; $u_{z,1t}$ and $u_{z,2t}$ are white noise disturbances. The ECM allows us to verify that changes in consumption and GDP at period t depend upon the deviation from their long-run equilibrium relationship in period $t-1$. For instance, if the level of consumption at t is above the level determined by (6), then we would expect that at *t+1* its level would decrease or GDP would increase to return to the long-run level. The last two terms that appear in both equations in (7) are included to take into account the potential problem of autocorrelation. In order to be cointegrated, at least one parameter, either $\theta_{z,1}$ or $\theta_{z,2}$ should be statistically significant. If both were zero, the long-run equilibrium relationship would not exist and consumption and GDP would not be cointegrated. Table 12 shows the results of estimating an ECM for Mexico and the US.

The number of lags included was selected in each case by optimizing the Akaike Information Criterion (AIC).

Table 12: Error-Correction Model

Country	Independent - Dependent	$\hat{\theta}_{z,i}$	$t_{\hat{\theta}_{z,i}}$	Constant and Lags	$Q_{16,d.f.}$	$LM_{2,lags}$
Mexico	Consumption-GDP	-0.357	-1.119	NO/0	7.156	1.947
	GDP-Consumption	-0.271	-0.638	NO/0	4.176	1.216
U.S.	Consumption-GDP	0.390	1.954	YES/m1=1	2.167	2.024
	GDP-Consumption	1.024	2.923	NO/m4=1	4.825	0.047

The ECM suggests that consumption and GDP in Mexico are not cointegrated. In these cases, both speed of adjustment parameters, $\theta_{z,i}$, are statistically equal to zero. This implies that either consumption or GDP is unresponsive to the previous period's deviations from the long-run equilibrium between these two variables. Furthermore, the results imply that consumption and GDP in the US are cointegrated. Whenever consumption at time *t*, $c_{us,t}$, exceeds the long-run equilibrium value, $\alpha_{us} + \beta_{us} y_{us,t}$ ($\varepsilon_{us,t} > 0$), the income, $y_{us,t+1}$ is "corrected" (augmented) in the following period at a speed of 1.024.

5. Conclusions

We show by means of Monte Carlo experiments that severe size distortions arise when working with small sample-size series in the case of the Granger causality test, the modified version of the Granger causality test, the Engle-Granger cointegration test, and the Hausman test for causality. Furthermore, the results obtained from these tests are unreliable if the series are not stationary, for which reason we chose to work with the series in first differences. Our empirical results reveal that the methodological improvements did not lead to the detection of a causal relationship between consumption and GDP: there is no evidence of either causality or cointegration between the Mexican series for consumption and GDP [this may be due to the small sample used; further research, with larger samples, should be carried out]. These results are similar to those in Guisan (2004). In the case of the US series, we find evidence of causality from consumption to GDP. We also find evidence of cointegration between these variables. The estimated ECM model states that the GDP variable is that which adjusts to short-run disequilibria. Notwithstanding these findings, it should be clear that cointegration does not imply causality (it is rather a long-run—equilibrium—relationship between the variables) and that causality in small samples is difficult to detect with the available tests.

Bibliography

Brown, R., J. Durbin, and J. Evans (1975). "Techniques for Testing the Constancy of Regression Relationships Over Time", *Journal of the Royal Statistical Society*, 37, 149–192.

Dickey, D., and W. Fuller (1979). “Distribution of the Estimators for Autoregressive Time Series with a Unit Root,” *Journal of the American Statistical Association*, 74(366), 427–431.
Enders, W. (2004). *Applied Econometric Time Series*. Wiley, Second Edition.
Engle, R., and C. Granger (1987). “Cointegration and Error Correction: Representation, Estimation, and Testing,” *Econometrica*, 55, 251–276.
Elliot, Grahwa, Rothenberg, Thomas J., and Stock, James H., (1996). “Efficient tests for an autoregressive unit root”, *Econometrica* 64 (4), pp. 813-836.
Gómez, M. and Ventosa-Santaulària D (2008). “Testing for a deterministic trend when there is evidence of Unit-Root”. Guanajuato School of Economics Working Paper Series No EC200802.
Guisan, M.C. (2001). “Causality and cointegration between consumption and GDP in 25 OECD countries: limitations of the cointegration approach”, *Applied Econometrics and International Development*, vol. 1 pp 36-61.[1]
Guisan, M.C. (2003). “Causality tests, interdependence and model selection: application to OECD countries, 1960-1997”. Working Paper Series of *Economic Development*, No. 63.[1]
Guisan, M.C. (2004). “A comparison of causality tests applied to the bilateral relationship between consumption and GDP in the USA and Mexico”, *International Journal of Applied Econometrics and Quantitative Studies*, Vol. 1, pp. 115-130.[1]
Guisan, M.C., Malacon, C., and Exposito, P. (2003). “Effects of the integration of Mexico into NAFTA on Trade, Industry, Employment and Economic Growth”. Working Paper Series of *Economic Development*, No. 63.[1]
Kwiatkowski, Denis, Phillips, Peter and Schmidt, Peter (1992). “Testing the null hypothesis of stationarity against the alternative of a unit root”, *Journal of Econometrics* 54 (1), pp. 159-178.
Ng, Serena and Perron, Pierre, (2001). “Lag Length Selection and the Construction of Unit Root tests with Good Size and Power”, *Econometrica* 69(6), pp. 1519-1554.
Noriega A. and Ventosa-Santaulària D. (1997). “Spurious Regression and Trending Variables”, *Oxford Bulletin of Economics and Statistics*, Vol. 69 (3), pp. 439-444.
Phillips, P.C.B. and Perron, P. (1988). “Testing for a unit root in time series regression”, *Biometrica*, vol. 75 (2), pp. 335-346.
Phillips, P. C. B. (1986) “Understanding spurious regressions in econometrics”, *Journal of Econometrics*, vol. 33, pp. 311–40.
Ventosa-Santaulària D. and Vera-Valdés, J.E. (2008). “Granger-Causality in the presence of structural breaks”, Economics Bulletin, Vol. 3, No 61 pp. 1-14
Zivot, E., and D. Andrews (1992). “Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis,” *Journal of Business and Economic Statistics,* vol. 10, 251–270.

Appendix: Data Generating Processes of the Simulations

The parameter values used for all the simulations included in this article are as follows:

DGP	Parameters [var. *y*]	Parameters [var. *x*]
1 [Driftless Unit Root]	$\sigma_y^2 = 1$	$\sigma_x^2 = 1$
2 [Unit Root with Drift]	$\sigma_y^2 = 1$; $\mu_y = 7$	$\sigma_x^2 = 1$; $\mu_x = 2$
4 [Trend Stationary]	$\sigma_y^2 = 1$; $\mu_y = 7$; $\beta_y = 0.03$	$\sigma_x^2 = 1$; $\mu_x = 2$ $\mu_y = 7$; $\beta_x = -0.03$
5 [Trend Stationary with Break]	$\sigma_y^2 = 1$; $\mu_y = 7$; $\beta_y = 0.03$; $\gamma_y = 0.02$	$\sigma_x^2 = 1$; $\mu_x = 2$ $\mu_y = 7$; $\beta_x = -0.03$; $\gamma_x = 0.04$

Notes: Innovations an *iid* normally distributed with zero mean and constant variance. The number of replications is 10,000.

[1] Articles on line at the EAAEDS Web site: http://www.usc.es/economet/eaa.htm