UNIVERSITÄT BERN

Faculty of Business, Economics and Social Sciences

Department of Economics

Cointegration Tests and the Classical Dichotomy

Luca Benati

17-04

April 2017

DISCUSSION PAPERS

Schanzeneckstrasse 1
Postfach
CH-3001 Bern, Switzerland
http://www.vwi.unibe.ch

Cointegration Tests and the Classical Dichotomy*

Luca Benati
University of Bern[†]

Abstract

Based on either Monte Carlo simulations, or several examples based on actual data, I show that the ability of Johansen's tests to detect a cointegration relationship significantly deteriorates under two empirically plausible circumstances: (*i*) when, in addition to a cointegration relationship, a system features one or more 'nuisance' series—i.e., series driven by permanent shocks *different* from those driving the cointegration relationship; and (*ii*) when a system features *multiple* cointegration relationships driven by *different* permanent shocks, as implied (e.g.) by the Classical Dichotomy (this being a special case of (*i*)).

These results suggest that performing Johansen's tests based on systems featuring both real and nominal series automatically biases the tests against rejecting the null. The substantive implication for applied research is that, when searching for cointegration based on Johansen's tests, a cointegration relationship should be tested based on the *smallest system* for which economic theory suggests cointegration should hold. I provide several illustrations of how failure to do so results in cointegration *not* being detected between (e.g.) either M_1 velocity and a short-term rate; real house prices and real rents; GDP and consumption; or short- and long-term interest rates.

Keywords: Unit roots; cointegration; Classical Dichotomy.

*I wish to thank Juan-Pablo Nicolini for kindly providing a quarterly version of Lucas and Nicolini's (2015) annual M_1 aggregate. Usual disclaimers apply. All of the data and the computer programs (written in MATLAB) used for this paper are available upon request.

[†]Department of Economics, University of Bern, Schanzeneckstrasse 1, CH-3001 Bern, Switzerland. Email: luca.benati@vwi.unibe.ch

1 Introduction

Economic theory suggests that, in several instances, alternative cointegration relationships should involve *different* permanent shocks. The Classical Dichotomy, for example, postulates that, in the very long run, the real and nominal sides of the economy should evolve independently,[1] with the former being driven by real forces such as permanent neutral and investment-specific technology shocks, and the latter depending instead on nominal forces such as permanent shocks to monetary aggregates (or their rates of growth).

In this paper I document and explore, both based on actual data, and *via* Monte Carlo simulations, a problem pertaining to Johansen's tests which arises systematically within this context. To the very best of my knowledge, the problem has not been identified and discussed before.

1.1 The problem

Specifically, I show that augmenting a cointegrated system with 'nuisance' I(1) series—i.e., series driven by *other* permanent shocks[2] (e.g., because they have been randomly generated in MATLAB)—makes it more and more difficult, in *small samples* of the size normally used in applied work, to detect cointegration based on Johansen's tests.[3]

A particular case, which is especially relevant for practical purposes, pertains to systems containing *multiple* cointegration relationships driven by *different* permanent shocks, as implied (e.g.) by the Classical Dichotomy. Within this context, either of the series pertaining to an individual cointegration relationship ends up playing the role of 'nuisance' series for the *other* cointegration relationships, thus decreasing the chances of detecting cointegration, sometimes dramatically so. E.g., I provide a simple example in which *two* cointegration relationships (between GDP and consumption, and a short- and a long-term rate, respectively) which are robustly identified based on the respective bivariate systems *disappear* when cointegration tests are performed based on the joint four-variable system, so that *no* cointegration relationship is identified. The problem is quite dramatic when working at the annual frequency—for which, even based on the entire post-WWI sample,[4] only (at most) about a hundred

[1] Benati (2015a) provides evidence compatible with this position. As he stresses, however, the extent of uncertainty is uniformly substantial, so that, strictly based on empirical evidence, it is not possible to make strong statements on this issue.

[2] So, to be clear, the 'nuisance' series are driven by *at least another* permanent shock unrelated to the original cointegration relationship.

[3] In a previous version of the paper I also reported Monte Carlo results showing that this problem does not materially distort results from Shin's (1994) tests of the null of cointegration. Ultimately, I decided to take these results out because they are not especially interesting. They are however available upon request.

[4] When working with nominal series, pre-WWI data cannot be 'lumped' together with data from subsequent years because of the radically different nature of the pre-WWI monetary regime, which makes the Lucas critique an especially relevant concern. E.g., Barsky (1987) and Benati

observations are available—but it is still often quite severe even based on quarterly post-WWII data, for which typical samples feature about 200-250 observations. In fact, I provide a simple example based on *monthly* data since 1963 in which cointegration between real house prices and real rents, which is robustly identified based on the bivariate system, once again disappears when the system is augmented with either real wages in the construction sector, or loans to real estate.

My results suggest that—assuming, as it seems reasonable, that the Classical Dichotomy holds—performing Johansen's tests based on systems featuring *both real and nominal series* automatically biases the tests against rejecting the null.

1.2 Implications for applied research

The substantive implication for applied research is therefore that, when searching for cointegration based on Johansen's tests, a specific cointegration relationship should be explored based on the *smallest system* for which economic theory suggests cointegration should hold. I provide several illustrations of how failure to follow this approach results in cointegration *not* being detected between either GDP and consumption; short- and long-term interest rates; M_1 velocity and a short-term rate; real house prices and real rents; labor productivity and real GDP *per capita*; or the multiplier of M_2-M_1 and a short-term nominal rate.

The paper is organized as follows. The next section discusses the methodology for bootstrapping critical and p-values for Johansen's tests I will use throughout the paper. Section 3 presents evidence from Monte Carlo simulations. Section 4 provides several illustrations of this problem based on actual data. Section 5 concludes.

2 Methodology

2.1 Bootstrapping Johansen's test statistics

Throughout the entire paper, I bootstrap critical and p-values for Johansen's trace and maximum eigenvalue tests[5] *via* the procedure proposed by Cavaliere *et al.* (2012, henceforth CRT). For either the trace test of the null hypothesis of no cointegration against the alternative of $h>0$ cointegration vectors, or the maximum eigenvalue test of zero *versus* one cointegration vectors, the process to be bootstrapped is a simple (i.e., non-cointegrated) VAR in differences. On the other hand, for the maximum eigenvalue tests of h versus $h+1$ cointegrating vectors, with $h>0$, the model to be

(2008) document dramatic differences between the stochastic properties of inflation between metallic standards and the post-WWI monetary regimes.

[5] The rationale for bootstrapping critical and p-values for Johansen's tests was provided by Johansen (2002) himself, who showed how, in small samples, trace and maximum eigenvalue tests based on asymptotic critical values typically tend to perform poorly.

bootstrapped is the VECM estimated under the null of h cointegrating vectors. All of the technical details can be found in CRT (2012), which the reader is referred to.

I select the VAR lag order as the maximum[6] between the lag orders chosen by the Schwartz and the Hannan-Quinn criteria[7] for the VAR in levels, with an allowed maximum number of lags equal to 12 when working at the monthly frequency; 4 when working at the quarterly frequency; and 2 when working at the annual frequency.

2.2 Monte Carlo evidence on the performance of the bootstrapping procedures

Either CRT (2012, Table I), Benati (2015, Table 1), or Benati, Lucas, Nicolini, and Weber (2016, Table 1; henceforth, BLNW) provide extensive Monte Carlo evidence on the excellent performance of CRT's bootstrapping procedure for Johansen's tests conditional on DGPs featuring *no cointegration*. In particular, either of the three papers shows that, essentially irrespective of the sample length, CRT's procedure produces correctly-sized tests, with empirical rejection frequencies (ERFs) very close to the test's nominal size.[8] If, however, the true DGP features *cointegration*, BLNW (2016) show that Johansen's tests, even bootstrapped as in CRT, perform well only if the persistence of the cointegration residual is sufficiently low, and/or the sample size is sufficiently large. On the other hand, if the cointegration residual is persistent, and/or the sample size is small, the tests fails to detect cointegration a non-negligible fraction of the times.[9] This is conceptually in line with some of the evidence reported by Engle and Granger (1987), and it has a straightforward explanation: As the cointegration residual becomes more and more persistent, it gets closer and closer to a random walk (in which case there would be no cointegration), and the procedure needs therefore larger and larger samples to detect the truth (i.e.: that the residual is highly persistent, but ultimately stationary).

I now turn to the Monte Carlo evidence for the problem discussed in the present work.

---

[6] I consider the maximum between the lag orders chosen by the SIC and HQ criteria because the risk associated with selecting a lag order smaller than the true one (model mis-specification) is more serious than the one resulting from choosing a lag order greater than the true one (over-fitting).

[7] On the other hand, I do not consider the Akaike Information Criterion since, as discussed (e.g.) by Luetkepohl (1991), for systems featuring I(1) series the AIC is an inconsistent lag selection criterion, in the sense of not choosing the correct lag order asymptotically.

[8] E.g., in BLNW's Table 1, with sample lengths ranging from T=50 to T=1000, the ERFs at the 10 per cent level range between 0.098 and 0.119.

[9] For example, with $T = 100$, cointegration will be detected, at the 10 per cent level, 43.3 per cent of the times if $\rho = 0.75$, and just 12.0 per cent of the times if $\rho = 0.95$, with ρ being the persistence of the cointegration residual.

3 Exploring the Problem *via* Monte Carlo Simulations

Tables 1 and 2 report results from the following Monte Carlo experiments.

3.1 The Monte Carlo experiment

I stochastically simulate the process

$$y_t = y_{t-1} + v_t, \quad \text{with } v_t \sim i.i.d.\ N(0,1) \tag{1}$$

$$x_t = y_t + u_t \tag{2}$$

$$u_t = \rho u_{t-1} + \epsilon_t, \text{ with } 0 \leq \rho < 1,\ \epsilon_t \sim i.i.d.\ N(0,1) \tag{3}$$

As for ρ, I consider three possible values, corresponding to alternative extents of persistence of the cointegration residual between y_t and x_t, that is, $\rho = 0.25$, 0.75, 0.95. I simulate the system (1)-(3) 1,000 times for three alternative sample lengths, $T = 50$, 100, 200. Based on each simulation, I then perform Johansen's trace and maximum eigenvalue tests of the null of no cointegration, bootstrapping the critical and p-values based on CRT's (2012) procedure. Bootstrapping is performed based on 1,000 replications.

I then perform the same simulations, and, based on the simulated data, the same cointegration tests, but this time augmenting the system (1)-(3) with either

(1) k additional random walks orthogonal to (1)-(3)—which I generate as in (1)—with $k = 1$, 2, 3, 4; or

(2) k additional series cointegrated with y_t—which I generate as in (2)-(3)—with $k = 1$, 2, 3, 4.

These simulations are designed to explore how the performance of Johansen's tests depends on

(I) the sample length;

(II) the persistence of the cointegration residual;

(III) the fact that the original cointegrated system $[y_t,\ x_t]'$ is augmented with either (i) k random walks orthogonal to the two original series, or (ii) k additional processes cointegrated to y_t and x_t; and

(IV) the number of additional random walks or cointegrated processes, k.

3.2 Evidence

The results are reported in Tables 1-2. The following main findings emerge from the tables:

first, in line with the Monte Carlo evidence reported in BLNW's (2016) Table 1, based on the simple bivariate cointegrated system $[y_t,\ x_t]'$, Johansen's tests perform well—in the sense of detecting cointegration—if and only if the sample period is

Table 1 Monte Carlo evidence on the performance of Johansen's tests: Fractions of replications for which no cointegration is rejected at the 10 per cent level[a] with k additional random walks

I: Based on the trace test

Persistence of the cointegration residual:	0 additional random walks			1 additional random walk			2 additional random walks		
	$T=50$	$T=100$	$T=200$	$T=50$	$T=100$	$T=200$	$T=50$	$T=100$	$T=200$
$\rho=0.25$	0.852	1.000	1.000	0.589	0.992	1.000	0.388	0.928	1.000
$\rho=0.75$	0.289	0.711	0.998	0.172	0.393	0.940	0.125	0.276	0.781
$\rho=0.95$	0.141	0.164	0.273	0.106	0.101	0.142	0.102	0.107	0.128
	3 additional random walks			4 additional random walks					
	$T=50$	$T=100$	$T=200$	$T=50$	$T=100$	$T=200$			
$\rho=0.25$	0.246	0.798	1.000	0.186	0.617	0.997			
$\rho=0.75$	0.116	0.200	0.589	0.103	0.176	0.448			
$\rho=0.95$	0.111	0.102	0.112	0.116	0.099	0.108			

II: Based on the maximum eigenvalue test

Persistence of the cointegration residual:	0 additional random walks			1 additional random walk			2 additional random walks		
	$T=50$	$T=100$	$T=200$	$T=50$	$T=100$	$T=200$	$T=50$	$T=100$	$T=200$
$\rho=0.25$	0.870	1.000	1.000	0.630	0.999	1.000	0.434	0.977	1.000
$\rho=0.75$	0.267	0.711	0.998	0.148	0.412	0.976	0.112	0.270	0.891
$\rho=0.95$	0.127	0.144	0.218	0.088	0.089	0.133	0.097	0.104	0.130
	3 additional random walks			4 additional random walks					
	$T=50$	$T=100$	$T=200$	$T=50$	$T=100$	$T=200$			
$\rho=0.25$	0.251	0.888	1.000	0.182	0.760	1.000			
$\rho=0.75$	0.121	0.185	0.688	0.102	0.160	0.512			
$\rho=0.95$	0.120	0.103	0.103	0.104	0.099	0.101			

[a] Based on 1,000 Monte Carlo replications, and, for each of them, on 1,000 bootstrap replications.

Table 2 Monte Carlo evidence on the performance of Johansen's tests: Fractions of replications for which no cointegration is rejected at the 10 per cent level[a] with k with additional cointegrated series

I: Based on the trace test

Persistence of the cointegration residual:	0 additional cointegrated series			1 additional cointegrated series			2 additional cointegrated series		
	$T=50$	$T=100$	$T=200$	$T=50$	$T=100$	$T=200$	$T=50$	$T=100$	$T=200$
$\rho=0.25$	0.852	1.000	1.000	0.967	1.000	1.000	0.977	1.000	1.000
$\rho=0.75$	0.289	0.711	0.998	0.327	0.858	1.000	0.283	0.884	1.000
$\rho=0.95$	0.141	0.164	0.273	0.124	0.155	0.306	0.108	0.139	0.298
	3 additional cointegrated series			4 additional cointegrated series					
	$T=50$	$T=100$	$T=200$	$T=50$	$T=100$	$T=200$			
$\rho=0.25$	0.953	1.000	1.000	0.930	1.000	1.000			
$\rho=0.75$	0.244	0.856	1.000	0.157	0.804	1.000			
$\rho=0.95$	0.122	0.111	0.227	0.116	0.094	0.184			

II: Based on the maximum eigenvalue test

Persistence of the cointegration residual:	0 additional cointegrated series			1 additional cointegrated series			2 additional cointegrated series		
	$T=50$	$T=100$	$T=200$	$T=50$	$T=100$	$T=200$	$T=50$	$T=100$	$T=200$
$\rho=0.25$	0.870	1.000	1.000	0.845	1.000	1.000	0.690	1.000	1.000
$\rho=0.75$	0.267	0.711	0.998	0.204	0.613	0.999	0.170	0.511	0.993
$\rho=0.95$	0.127	0.144	0.218	0.101	0.131	0.179	0.097	0.112	0.149
	3 additional cointegrated series			4 additional cointegrated series					
	$T=50$	$T=100$	$T=200$	$T=50$	$T=100$	$T=200$			
$\rho=0.25$	0.557	1.000	1.000	0.445	0.999	1.000			
$\rho=0.75$	0.150	0.374	0.990	0.116	0.322	0.958			
$\rho=0.95$	0.109	0.093	0.131	0.105	0.108	0.127			

[a] Based on 1,000 Monte Carlo replications, and, for each of them, on 1,000 bootstrap replications.

sufficiently long, and the persistence of the cointegration residual is sufficiently low. If, on the other hand, either T is small, or ρ is large, Johansen's tests fail to correctly reject the null of cointegration a non-negligble, and sometimes large fraction of the times.

Second, in line with this paper's main theme about the impact of 'nuisance' I(1) series on the performance of Johansen's tests, for each single combination of values of ρ and T, the more orthogonal random walks are added to the original cointegrated system $[y_t,\ x_t]'$, the worse the performance of the trace and maximum eigenvalues tests becomes. E.g., focusing, for purely illustrative purposes, on the combination $T = 100$ and $\rho = 0.75$, the fraction of Monte Carlo simulations for which the trace test rejects the null of no cointegration is equal to 0.711 based on the bivariate system $[y_t,\ x_t]'$, but it falls to 0.393 with the addition of a *single random walk*. With two random walks the fraction of rejections further decreases to 0.276, whereas with three and four random walks it falls to 0.200 and 0.176, respectively. These results provide a straightforward explanation for what we will see in Section 4: A key point to stress here is that, based on the results reported in Table 1, the 'greatest damage' (so to speak) is inflicted by the *first additional random walk*. Further augmenting the system with additional random walks produces smaller and smaller deteriorations of the performance of Johansen's tests. This explains results such as those of Table 4, where the addition of log GDP to the bivariate system featuring M_1 velocity and the short rate causes the previously identified cointegration relationship between these two series to 'disappear'.

Third, augmenting the original bivariate system $[y_t,\ x_t]'$ with additional cointegrated processes produces mixed results. Specifically, starting with the *trace* test, its ability to detect cointegration first improves, and then somehow worsens, if the persistence of the cointegration residual is low. For example, with $T = 50$ and $\rho = 0.25$, the fraction of Monte Carlo simulations for which the trace test rejects the null of no cointegration is equal to 0.852 based on the bivariate system $[y_t,\ x_t]'$, it increases to 0.977 with two additional cointegrated processes, and it falls to 0.930 with four additional processes. By the same token, with $T = 100$ and $\rho = 0.75$, the fraction of simulations for which no cointegration is rejected is equal to 0.711 based on the system $[y_t,\ x_t]'$, it increases to 0.884 with two additional processes, and it falls to 0.804 with four additional processes. With $\rho = 0.95$, however, the fractions of rejections are almost uniformly decreasing in the number of additional processes, although the fractions are much smaller than those obtained with smaller values of ρ. Turning to the *maximum eigenvalue* test of 0 *versus* 1 cointegration vectors, its performance gets uniformly worse the more additional cointegrated processes are added to the bivariate system system $[y_t,\ x_t]'$. This result, however, can be easily rationalized simply by noting that with k additional cointegrated processes, the test of 0 *versus* 1 cointegration vectors is obviously not the most appropriate, and therefore the most powerful.

4 Illustrations Based on Actual Data

I now turn to illustrations based on actual data. The recurring theme among these examples is that a cointegration relationship which has been detected based on the smallest system suggested by economic theory 'disappears' when the system is augmented with series whose I(1) component, according to theory, should *not* be driven by the permanent shock driving the original cointegration relationship (e.g., several examples will involve the Classical Dichotomy).

All of the series I work with in this section are I(1) according to the unit root tests of Elliot, Rothenberg, and Stock (1996). (I do not report this evidence for reasons of space, but all of these results are available upon request.) All of the data are described in detail in Appendix A.

4.1 Example I: Long-run money demand

In their investigation of long-run money demand since the mid-XIX century, BLNW (2016) document the existence of a cointegration relationship between M_1 velocity and a short-term nominal interest rate (henceforth, $V_{1,t}$ and R_t, respectively) for the United States since 1915, Canada since 1926, and many other countries over the post-World War I period. Since this relationship involves permanent interest rates shocks[10]—which, by the Classical Dichotomy, should not drive the unit root component, of (e.g.) real consumption and GDP—based on the discussion in Section 3 we should expect that, once augmenting the system $[V_{1,t}\ R_t]'$ with series such as GDP, it should become more difficult to detect the original cointegration relationship. As I will now show, this is indeed the case, and starkly so. Specifically, I will show that, for both the United States and Canada, augmenting the system $[V_{1,t}\ R_t]'$ with log real GDP causes the original cointegration relationship to 'disappear'.[11] On the other hand, augmenting $[V_{1,t}\ R_t]'$ with the long-term rate—i.e., with a series which, according to both economic theory, and results from cointegration tests,[12] is driven by the very same permanent shocks driving the short rate—does not cause the original cointegration relationship to 'disappear': Rather, a second cointegration relationship is identified.

Table 3 reports results from Johansen's trace and maximum eigenvalue tests for the three alternative systems for the United States, for the periods 1915-2014 (based on annual data) and 1959Q1-2012Q4 (based on quarterly data), and for Canada,

[10] I.e., by the Fisher effect, the permanent inflation shocks associated with the post-World War I monetary regimes.

[11] The same holds if the system $[V_{1,t}\ R_t]'$ is augmented with log consumption. These results are not reported here for reasons of space, but they are available upon request.

[12] For the United States, results from the trace and maximum eigenvalue tests are (bootstrapped p-values in parentheses) 23.2 (0.01) and 20.9 (0.01), respectively based on annual data for the period 1915-2014, and 26.143 (0.006) and 24.091 (0.004), respectively based on quarterly data for the period 1959Q1-2012Q4. For Canada they are 36.3 (0.00) and 34.2 (0.00), respectively.

Table 3 Results from Johansen's cointegration tests for long-run money demand[a]

Tested system:	*Trace tests of the null of no cointegration against the alternative of h or more cointegration vectors:*					
	United States (1915-2014)		United States (1959Q1-2012Q4)		Canada (1926-2006)	
	$h=1$	$h=2$	$h=1$	$h=2$	$h=1$	$h=2$
$Y_t = [V_{1,t},\, R_t]'$	20.769 (0.039)	–	18.012 (0.080)	–	23.244 (0.015)	–
$Y_t = [V_{1,t},\, R_t,\, \ln(GDP_t)]'$	22.468 (0.444)	–	28.504 (0.101)	–	22.725 (0.526)	–
$Y_t = [V_{1,t},\, R_t,\, r_t]'$	52.100 (0.001)	25.759 (0.000)	43.568 (0.011)	15.483 (0.022)	58.424 (0.000)	21.988 (0.002)
	Maximum eigenvalue tests of h versus h+1 cointegration vectors:					
	United States (1915-2014)		United States (1959Q1-2012Q4)		Canada (1926-2006)	
	0 *versus* 1	1 *versus* 2	0 *versus* 1	1 *versus* 2	0 *versus* 1	1 *versus* 2
$Y_t = [V_{1,t},\, R_t]'$	16.557 (0.050)	–	17.283 (0.049)	–	21.714 (0.007)	–
$Y_t = [V_{1,t},\, R_t,\, \ln(GDP_t)]'$	17.327 (0.269)	–	17.929 (0.166)	–	18.126 (0.284)	–
$Y_t = [V_{1,t},\, R_t,\, r_t]'$	26.341 (0.023)	21.640 (0.008)	28.086 (0.018)	14.320 (0.088)	36.436 (0.001)	20.212 (0.013)

[a] Bootstrapped p-values (in parentheses) are based on 10,000 bootstrap replications. $V_{1,t} = M_1$ velocity; R_t = short-term nominal rate.; GDP_t = real GDP; r_t = long-term nominal rate.

for the period 1926-2006 (based on annual data). Throughout all of Section 4, results will be based on 10,000 bootstrap replications of CRT's procedure. For either country, three results stand out:

(I) in line with the results reported in BLNW (2016), evidence of cointegration between M_1 velocity and the short rate is uniformly very strong for either country.

(II) When the system $[V_{1,t}\ R_t]'$ is augmented with log real GDP, the cointegration relationship originally identified in the bivariate system disappears.

(III) On the other hand, when the system $[V_{1,t}\ R_t]'$ is augmented with the long-term nominal rate, not only the previously identified cointegration relationship does not disappear, but rather—as we would expect based on economic theory—a second cointegration relationship is identified.

In the light of the Monte Carlo evidence we saw in Section 3, the explanation for such starkly contrasting results is straightforward: Whereas in example (II) the original system is augmented with a nuisance series—i.e., a series which is *not* driven by the same permanent shock driving $[V_{1,t}\ R_t]'$—in example (III) the additional series is instead driven by the very same shock, and it does not therefore play the role of a nuisance series.

When a researcher first encounters this kind of results, (s)he might be excused for just shrugging them off, and interpreting them as possibly due to a combination of small sample issues, and the 'luck of the draw':[13] After all, precisely because we are dealing with statistical tests, even under ideal conditions a certain number of fluke results is to be expected. When one keeps systematically obtaining these results within widely disparate contexts, however, a pattern starts to emerge. As I will show in the rest of the paper, the pattern is indeed a very consistent one.

4.2 Example II: Real house prices and real rents

Table 4 reports results from Johansen's cointegration tests for four alternative systems featuring housing market variables, for the United States for the period January 1963-December 2014. In line with previous evidence—see, in particular, Gallin (2008) and Benati (2015b)—the null of no cointegration between log real house prices and log real rents (henceforth, hp_t and re_t, respectively) is strongly rejected. In line with the previous discussion, however, augmenting the bivariate system $[hp_t,\ re_t]'$ with I(1) series which—according to either economic theory, or previous empirical evidence—should *not* be expected to be driven by the same shock driving the permanent component of house prices and rents, causes the previously identified cointegration relationship to disappear. Specifically,

(i) augmenting the system $[hp_t,\ re_t]'$ with log real wages in construction produces p-values for the trace and maximum eigenvalue tests equal to 0.358 and 0.149, respectively.

[13] This was indeed my initial reaction when, a few years ago, I started noticing this kind of results.

Table 4 United States, January 1963-December 2014: Results from Johansen's cointegration tests for the housing market[a]

	Trace tests of the null of no cointegration against the alternative of h or more cointegrating vectors:	
	$h = 1$	$h = 2$
Log real house price and log real rent	16.527 (0.041)	–
Log real house price, log real rent, and log industrial production	27.899 (0.098)	
Log real house price, log real rent, and log real wages in construction	21.858 (0.358)	
Log real house price, log real rent, and log real loans to real estate *per capita*	26.633 (0.148)	5.430 (0.666)
	Maximum eigenvalue tests of h versus h+1 cointegrating vectors:	
	0 *versus* 1	1 *versus* 2
Log real house price and log real rent	14.499 (0.052)	
Log real house price, log real rent, and log industrial production	18.209 (0.141)	
Log real house price, log real rent, and log real wages in construction	18.112 (0.149)	
Log real house price, log real rent, and log real loans to real estate *per capita*	21.203 (0.064)	3.651 (0.938)

[a] Bootstrapped p-values (in parentheses) are based on 10,000 bootstrap replications

Table 5 United States, 1953Q2-2016Q2: Results from Johansen's cointegration tests for labor productivity and GDP *per capita*[a]

	Trace tests of the null of no cointegration against the alternative of h or more cointegrating vectors:	
	$h = 1$	$h = 2$
Log labor productivity and log real GDP *per capita*	17.749 (0.030)	–
Log labor productivity, log real GDP *per capita*, and 3-month Treasury bill rate	28.136 (0.132)	–
Log labor productivity, log real GDP *per capita*, and 10-year government bond yield	26.303 (0.174)	–
Logarithms of labor productivity, real GDP *per capita*, and real consumption *per capita*	48.898 (1.0×10^{-4})	22.274 (1.0×10^{-4})
	Maximum eigenvalue tests of h versus h+1 cointegrating vectors:	
	0 *versus* 1	1 *versus* 2
Log labor productivity and log real GDP *per capita*	14.543 (0.054)	–
Log labor productivity, log real GDP *per capita*, and 3-month Treasury bill rate	18.437 (0.170)	–
Log labor productivity, log real GDP *per capita*, and 10-year government bond yield	18.290 (0.166)	–
Logarithms of labor productivity, real GDP *per capita*, and real consumption *per capita*	26.624 (0.013)	17.906 (0.017)

[a] Bootstrapped p-values (in parentheses) are based on 10,000 bootstrap replications

(*ii*) Augmenting it with log industrial production produces marginally better results, with the p-values now being equal to 0.098 and 0.141, respectively.

(*iii*) Finally, augmenting the system $[hp_t, re_t]'$ with log real loans to real estate *per capita* identifies one cointegration vector based on the maximum eigenvalue statistic, whereas based on the trace tests it does not identify any cointegration relationship.

4.3 Example III: Labor productivity and GDP *per capita*

Table 5 reports results from Johansen's tests for the logarithms of labor productivity and real GDP *per capita* for the United States for the period 1953Q2-2016Q2.[14] As we should expect based on basic economic logic, tests performed based on the bivariate system detect strong evidence of cointegration. When the bivariate system is augmented with either the 3-month Treasury bill rate, or the 10-year government bond yield—i.e., series whose I(1) component should be driven by permanent inflation shocks, which, according to the Classical Dichotomy, should be orthogonal to the permanent shocks driving labor productivity and GDP *per capita*—evidence of cointegration however disappears, with p-values ranging between 0.132 and 0.174. On the other hand, when the bivariate system is augmented with the logarithm of real consumption *per capita*, we strongly identify—as it should be—*two* cointegration vectors, with p-values ranging between 1.0×10^{-4} and 0.017.

4.4 Example IV: GDP and consumption

Cointegration between real GDP and real consumption is one of the most robust and best-known stylized facts in economics.[15] And indeed—see Table 6—for both the United States and Canada, for the periods 1915-2014 and 1926-2006, respectively, Johansen's tests reject the null of no cointegration between the two series at least at the 10 per cent level. Conceptually in line with what we saw in the previous sub-sections, however, augmenting the system $[\ln(GDP_t),\ \ln(C_t)]'$ with the short rate causes the original cointegration relationship to disappear for the United States. As for Canada, evidence of cointegration becomes markedly weaker, with the p-value for the trace test being now equal to 0.136, and that for the maximum eigenvalue test being instead equal to 0.051. Finally—and again, in line with what we saw in the previous sub-sections—when the original system is instead augmented with another series driven by the very same permanent shock (within the present context, investment), for the United States two cointegration relationships are robustly identified based on either the trace or the maximum eigenvalue tests. As for Canada, the

[14] Both series are available (at least) since 1948Q1, but we work with the period 1953Q2-2016Q2 because one of the I(1) 'nuisance' series we will use is the 10-year government bond yield, which is only available since April 1953.

[15] See, first and foremost, Cochrane (1994).

Table 6 Results from Johansen's cointegration tests for GDP and consumption[a]

Tested system:	*Trace tests of the null of no cointegration against the alternative of h or more cointegration vectors:*			
	United States (1915-2014)		Canada (1926-2006)	
	$h = 1$	$h = 2$	$h = 1$	$h = 2$
$Y_t = [\ln(GDP_t),\, \ln(C_t)]'$	15.682 (0.081)	–	30.510 (0.009)	–
$Y_t = [\ln(GDP_t),\, \ln(C_t),\, R_t]'$	21.781 (0.472)	–	30.510 (0.136)	–
$Y_t = [\ln(GDP_t),\, \ln(C_t),\, \ln(I_t)]'$	54.294 (0.001)	19.977 (0.001)	96.004 (0.000)	27.314 (0.007)
	Maximum eigenvalue tests of h versus h+1 cointegration vectors:			
	United States (1915-2014)		Canada (1926-2006)	
	0 *versus* 1	1 *versus* 2	0 *versus* 1	1 *versus* 2
$Y_t = [\ln(GDP_t),\, \ln(C_t)]'$	14.984 (0.064)	–	20.167 (0.055)	–
$Y_t = [\ln(GDP_t),\, \ln(C_t),\, R_t]'$	17.259 (0.263)	–	24.918 (0.051)	–
$Y_t = [\ln(GDP_t),\, \ln(C_t),\, \ln(I_t)]'$	34.316 (0.004)	19.269 (0.017)	68.690 (0.000)	17.883 (0.169)

[a] Bootstrapped p-values (in parentheses) are based on 10,000 bootstrap replications. GDP_t = real GDP; C_t = real consumption; I_t = real investment.

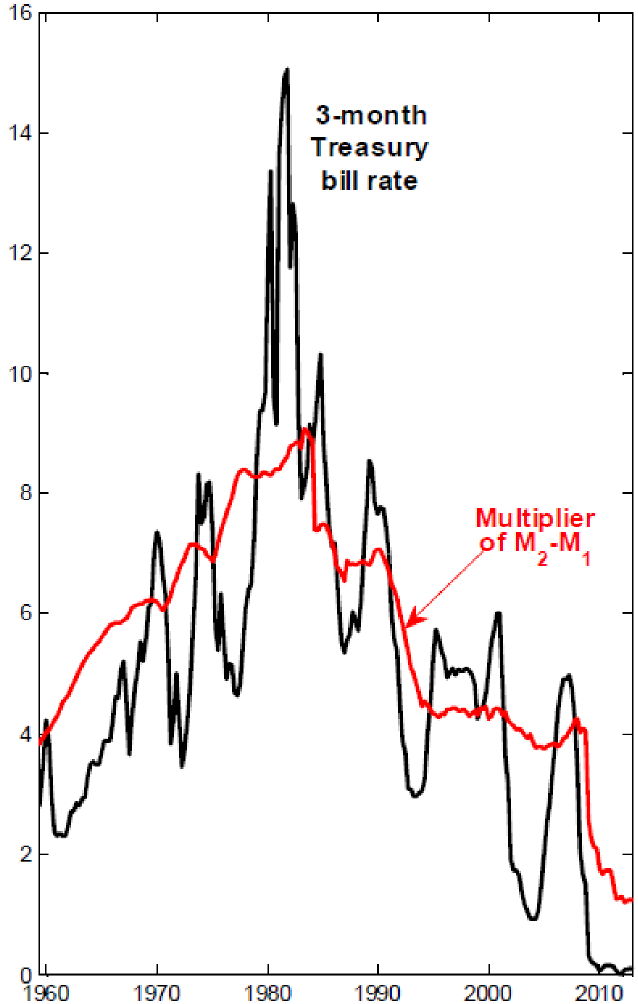

Figure 1 United States, 1959Q1-2012Q4: 3-month Treasury bill rate, and multiplier of M_2-M_1

Table 7 United States, 1959Q1-2012Q4: Results from Johansen's cointegration tests for the multiplier of M_2-M_1, and the 3-month Treasury bill rate[a]

	Trace tests of the null of no cointegration against the alternative of h or more cointegrating vectors:	
	$h = 1$	$h = 2$
$Y_t = [\mu_{2-1,t},\ R_t]'$	18.970 (0.058)	–
$Y_t = [\mu_{2-1,t},\ R_t,\ \ln(GDP_t)]'$	27.210 (0.145)	–
$Y_t = [\mu_{2-1,t},\ R_t,\ F_t]'$	45.982 (0.004)	13.886 (0.046)
	Maximum eigenvalue tests of h versus h+1 cointegrating vectors:	
	0 *versus* 1	1 *versus* 2
$Y_t = [\mu_{2-1,t},\ R_t]'$	17.916 (0.033)	–
$Y_t = [\mu_{2-1,t},\ R_t,\ \ln(GDP_t)]'$	14.755 (0.365)	–
$Y_t = [\mu_{2-1,t},\ R_t,\ F_t]'$	32.096 (0.006)	13.785 (0.092)

[a] Bootstrapped p-values (in parentheses) are based on 10,000 bootstrap replications. $\mu_{2-1,t}$ = multiplier of M_2-M_1; R_t = 3-month Treasury bill rate; GDP_t = real GDP *per capita*; F_t = Federal Funds rate.

trace tests robustly identify two cointegration relationships, whereas the maximum eigenvalue tests identify one.

With the exception of this last, and comparatively minor difference, the results in Table 6 are therefore uniformly in line with those discussed up until now, and suggest that augmenting a cointegration relationship with nuisance I(1) series (as previously defined) makes it more and more difficult to detect the originally identified relationship based on Johansen's tests.

4.5 Example V: The multiplier of M_2-M_1 and a short-term rate

As documented by Benati and Ireland (2017), in the United States, over the entire period since World War I, the multiplier of the monetary aggregate M_2-M_1 has exhibited a very strong correlation with the 3-month Treasury bill rate, and, in fact, has been cointegrated with it.[16] Figure 1 plots the two series for the period 1959Q1-2012Q4: The strong correlation between them is readily apparent. As stressed by Benati and Ireland (2017), the obvious explanation for such a pattern has to do with the permanent portfolio shifts out of non interest-bearing M_1, and into interest-bearing M_2-M_1, caused by the permanent interest rate (i.e., inflation) shocks associated with the post-WWI U.S. monetary regimes.

The first row of Table 7 reports results from Johansen's cointegration tests between the 3-month Treasury bill rate and the multiplier of M_2-M_1: With the p-values equal to 0.058 and 0.033, respectively, both the trace and the maximum eigenvalue tests confirm the visual impression from Figure 1 of a long-run equilibrium relationship between the two series. The cointegration relationship 'disappears', however, once the bivariate system is augmented with log real GDP *per capita*, i.e. a series which, according to the Classical Dichotomy, should be thought of as orthogonal to inflation in the infinite long run. Finally, once again in line with what we have seen up until now, when instead the bivariate system is augmented with the Federal Funds rate—which both theory, and results from cointegration tests,[17] suggest should be cointegrated with the 3-month Treasury bill rate—not only the original cointegration relationship does not disappear, but a second cointegration vector is identified.

4.6 Example VI: Short- and long-term nominal interest rates

Economic theory suggests that the permanent shocks impacting upon nominal interest rates (i.e., permanent inflation shocks, and permanent shocks to the natural rate of interest) should have an identical impact on nominal rates at all maturities, so

[16] As detailed in Appendix A, the multiplier of M_2-M_1 has been computed as the ratio between M_2-M_1 and the monetary base.

[17] Results from trace and maximum eigenvalue tests (bootstrapped p-values in parentheses) are 35.209 (8.0×10^{-4}) and 31.526 (4.0×10^{-4}) respectively.

Table 8 Results from Johansen's cointegration tests for short- and long-term rates[a]

Tested system:	*Trace tests of the null of no cointegration against the alternative of h or more cointegration vectors:*			
	United States (1915-2014)		United Kingdom (1922-2014)	
	$h = 1$	$h = 2$	$h = 1$	$h = 2$
$Y_t = [R_t,\, r_t]'$	23.220 (0.016)	–	29.523 (0.002)	–
$Y_t = [R_t,\, r_t,\, \ln(GDP_t)]'$	26.368 (0.235)	–	20.889 (0.535)	
$Y_t = [R_t,\, r_t,\, V_{1,t}]'$	52.100 (0.001)	25.759 (0.000)	42.564 (0.010)	15.400 (0.035)
	Maximum eigenvalue tests of h versus h+1 cointegration vectors:			
	United States (1915-2014)		United Kingdom (1922-2014)	
	0 *versus* 1	1 *versus* 2	0 *versus* 1	1 *versus* 2
$Y_t = [R_t,\, r_t]'$	20.901 (0.010)	–	28.253 (0.001)	–
$Y_t = [R_t,\, r_t,\, \ln(GDP_t)]'$	22.933 (0.068)	–	16.527 (0.324)	–
$Y_t = [R_t,\, r_t,\, V_{1,t}]'$	26.341 (0.022)	21.640 (0.007)	27.164 (0.024)	14.332 (0.102)

[a] Bootstrapped p-values (in parentheses) are based on 10,000 bootstrap replications.
R_t = short-term nominal rate.; GDP_t = real GDP; r_t = long-term nominal rate.

that short- and long-term nominal rates should be cointegrated.[18] Table 8 reports evidenced from Johansen's cointegration tests for the United States (1915-2014) and the United Kingdom (1922-2014) for three alternative systems featuring short-term (R_t) and long-term (r_t) nominal interest rates. Evidence is in line with what we have seen up until now. Specifically,

(*i*) For both countries, evidence from either test strongly points towards cointegration between short- and long-term nominal interest rates.

(*ii*) Augmenting the system $[R_t,\ r_t]'$ with log real GDP causes the previously identified cointegration relationship to disappear for the United Kingdom. As for the United States evidence is ambiguous, with the maximum eigenvalue test detecting one cointegration vector, and the trace test instead detecting none.

(*iii*) Augmenting the system $[R_t,\ r_t]'$ with M_1 velocity—i.e., with a series which, as shown by BLNW (2016), is, for either country, cointegrated with the short-term rate—leads to the identification of *two* cointegration vectors for the United States (as it should be). As for Canada, the trace tests identify two cointegration vectors, whereas the maximum eigenvalue tests identify one, with the p-value for the test of 1 versus 2 cointegration vectors, at 0.102, being just marginally insignificant.

I now turn to systems featuring multiple cointegration relationships driven by different permanent shocks.

4.7 Example VII: Multiple cointegration relationship driven by different permanent shocks

A stark illustration of the problem under investigation is provided by the results reported in Table 9. As previously discussed, for the United States for the period 1915-2014 results from either the trace or the maximum eigenvalue tests strongly point towards cointegration between either M_1 velocity and the short-term rate (see Table 3), or GDP and consumption (see Table 6). In contrast with these results, *neither* the trace *nor* the maximum eigenvalue tests for the four-variables system $[V_{1,t},\ R_t,\ \ln(GDP_t),\ \ln(C_t)]'$ identify *any* cointegration relationship, with the p-values ranging between 0.226 and 0.555.

In the light of the previously discussed results, the most logical explanation for this evidence is that, within the present context, the series pertaining to either of the two cointegration relationships embedded in the system $[V_{1,t},\ R_t,\ \ln(GDP_t),\ \ln(C_t)]'$—that is, $V_{1,t}$ and R_t, and $\ln(GDP_t)$ and $\ln(C_t)$, respectively—end up playing the role of 'nuisance' series for the *other* cointegration relationships (i.e., for the relationship driven by the other permanent shock).

The results reported in Table 9 are especially stark not only because *two cointegration relationships disappear*, but also because this example involves the Classical Dichotomy, which most economists would likely regard as a bedrock of economics.

[18] Empirical evidence in support of this notion is provided by Benati (2016). In particular, Benati (2016) shows that, in the post-WWII United States, all interest rates' spreads are strongly stationary.

Table 9 United States, 1915-2014: Results from Johansen's cointegration tests for GDP, consumption, M_1 velocity, and a short-term rate[a]

	Trace tests of the null of no cointegration against the alternative of h or more cointegrating vectors:	
	$h = 1$	$h = 2$
$Y_t = [\ln(GDP_t), \ln(C_t), V_{1,t}, R_t]'$	42.527 (0.385)	22.359 (0.226)
	Maximum eigenvalue tests of h versus h+1 cointegrating vectors:	
	0 *versus* 1	1 *versus* 2
$Y_t = [\ln(GDP_t), \ln(C_t), V_{1,t}, R_t]'$	20.168 (0.555)	16.791 (0.295)

[a] Bootstrapped p-values (in parentheses) are based on 10,000 bootstrap replications.
GDP_t = real GDP; C_t = real consumption; $V_{1,t}$ = M_1 velocity; R_t = short-term nominal rate.

Table 10 United States, 1959Q1-2012Q4: Results from Johansen's cointegration tests for real GDP and consumption *per capita*, labor productivity, the 3-month Treasury bill rate, the 10-year government bond yield, and M_1 velocity[a]

Tested system:	*Trace tests of the null of no cointegration against the alternative of h or more cointegrating vectors:*			
	$h=1$	$h=2$	$h=3$	$h=4$
$Y_t = [\ln(GDP_t), \ln(C_t), \ln(PR_t)]'$	40.647 (0.005)	15.899 (0.007)	–	–
$Y_t = [\ln(GDP_t), \ln(C_t), \ln(PR_t), V_{1,t}, R_t, r_t]'$	117.700 (0.013)	82.458 (0.002)	49.030 (0.003)	23.166 (0.032)
	Maximum eigenvalue tests of h versus h+1 cointegrating vectors:			
	0 *versus* 1	1 *versus* 2	2 *versus* 3	3 *versus* 4
$Y_t = [\ln(GDP_t), \ln(C_t), \ln(PR_t)]'$	24.748 (0.026)	13.821 (0.068)	–	–
$Y_t = [\ln(GDP_t), \ln(C_t), \ln(PR_t), V_{1,t}, R_t, r_t]'$	35.241 (0.308)	33.428 (0.119)	25.864 (0.149)	14.695 (0.485)

[a] Bootstrapped p-values (in parentheses) are based on 10,000 bootstrap replications.

GDP_t, C_t = real GDP and consumption *per capita*; PR_t = labor productivity; R_t = 3-month Treasury bill rate; r_t = 10-year government bond yield; $M_{1,t}$, = M_1 velocity.

Table 11 United States, 1953Q2-2016Q2: Results from Johansen's cointegration tests for real GDP and hours *per capita*, inflation, the 3-month Treasury bill rate, and the 3-, 5-, and 10-year government bond yields[a]

Tested system:	*Trace tests of the null of no cointegration against the alternative of h or more cointegrating vectors:*			
	$h = 1$	$h = 2$	$h = 3$	$h = 4$
$Y_t = [\pi_t, R_t, r_t^1, r_t^3, r_t^{10}]'$	121.619 (0.000)	71.177 (0.000)	43.393 (0.000)	18.950 (0.000)
$Y_t = [\ln(GDP_t), \ln(H_t), \pi_t, R_t, r_t^3, r_t^5, r_t^{10}]'$	196.049 (0.000)	139.896 (0.000)	84.899 (0.000)	55.603 (0.000)
	Maximum eigenvalue tests of h versus h+1 cointegrating vectors:			
	0 *versus* 1	1 *versus* 2	2 *versus* 3	3 *versus* 4
$Y_t = [\pi_t, R_t, r_t^1, r_t^3, r_t^{10}]'$	50.442 (8.0×10^{-4})	27.784 (0.061)	24.443 (0.021)	14.586 (0.065)
$Y_t = [\ln(GDP_t), \ln(H_t), \pi_t, R_t, r_t^3, r_t^5, r_t^{10}]'$	56.152 (0.013)	54.997 (0.003)	29.296 (0.272)	23.760 (0.203)

[a] Bootstrapped p-values (in parentheses) are based on 10,000 bootstrap replications.
GDP_t = real GDP *per capita*; H_t = hours *per capita*; π_t = inflation; R_t = 3-month Treasury bill rate; r_t^3, r_t^5, r_t^{10} = 3-, 5-, and 10-year government bond yield.

If the real and nominal sides of the economy are indeed orthogonal in the infinite long-run, the results discussed in this section suggest that testing for the null of no cointegration based on multivariate systems featuring *both real and nominal series* will automatically bias the tests' results against rejecting the null even if cointegration truly is there.

Table 10 reports a less dramatic, but more relevant set of results, being based not on long-run annual data—which are comparatively less investigated in macroeconomics—but rather on U.S. quarterly series for the period 1959Q1-2012Q4. As reported in Table 3, for the trivariate system comprising M_1 velocity, the 3-month Treasury bill rate, and the 10-year government bond yield, Johansen's tests identified two cointegration relationships. As Table 10 shows, for the trivariate system comprising the logarithms of labor productivity, and of real GDP and real consumption *per capita*, the tests also identify (as we would expect) two cointegration vectors. When the tests are performed based on the joint six-variables system, however, things change quite significantly: Whereas the trace tests do indeed identify the expected four cointegration vectors, the maximum eigenvalue tests identify *none*. In particular, none of the p-values for either of the maximum eigenvalue tests, from 0 *versus* 1 up to 3 *versus* 4, is smaller than 10 per cent.

Table 11 reports similar evidence for a system featuring the logarithms of real GDP and hours worked *per capita*, GDP deflator inflation, and four interest rates series, estimated based on U.S. data for the period 1953Q2-2016Q2. As basic economic logic would have it, both the trace and the maximum eigenvalue tests identify four cointegration vectors in the five-variables sub-system featuring inflation and the four nominal interest rates series, reflecting the impact on nominal rates of permanent inflation shocks. In line with the results discussed in the previous paragraph, however, based on the largest, seven variables system the trace tests still identify, as expected, four cointegration vectors, whereas the maximum eigenvalue tests identify only two.[19]

4.8 Implications

I could go on providing several additional examples, but by now my point should be clear: When you start with an I(1) system for which both economic theory, and results from cointegration tests, suggest that there is a cointegration relationship driven by a specific permanent shock, and you augment it with one or more I(1) series which, according to theory, should be driven by a *different* permanent shock, there is a very good chance that the original cointegration relationship will 'disappear'.

[19] The results reported in Tables 10 and 11 would appear to suggest that, when facing the problem discussed in the present work, the trace test is more robust than the maximum eigenvalue test: In fact, in both instances, trace tests identified the correct number of cointegration vectors, whereas the maximum eigenvalue tests fell short. Although these two examples would seem to point towards this conclusion, some caution is in order, as the Monte Carlo results in Table 1 do not point towards a clear superiority of trace tests compared to maximum eigenvalue tests.

This has the following important implication: It is common practice, in macroeconomics, to work with multivariate systems featuring *both* real *and* nominal I(1) series, i.e., series which, according to the Classical Dichotomy, should be driven by *different* permanent shocks. The previously discussed results suggest that, under these circumstances, performing Johansen's tests based on the full system will automatically bias the test results against rejecting the null. The starkest illustration of this problem is the just-mentioned example in which, in the United States for the period 1915-2014, two cointegration relationships just 'vanish'. The example is an extreme one, but it brings home, in the starkest possible way, the point that under these circumstances things can easily go very badly wrong, and inference based on cointegrated (structural) VARs can turn out to be seriously distorted.

The logical implications for applied research is therefore that, when testing the null of no cointegration, a specific cointegration relationship should be explored based on the *smallest system* for which economic theory suggests cointegration should hold.

5 Conclusions

In this paper I have documented a specific problem with tests of the null of no cointegration which, to the very best of my knowledge, had not been previously noticed and analyzed. Specifically, I have shown—based on both several examples based on actual data, and Monte Carlo simulations—that augmenting a cointegration relationship with 'nuisance' I(1) series (i.e., series driven by other permanent shocks) makes it more and more difficult, in small samples of the size typically used in applied work, to detect such relationship based on Johansen's tests. The substantive implication for applied research is that, when testing the null of no cointegration, a cointegration relationship should be tested based on the smallest system for which economic theory suggests cointegration should hold. I provide several illustrations of how failure to follow this approach results in cointegration not being detected between either GDP and consumption; short- and long-term interest rates; M_1 velocity and the short-term rate; or real house prices and real rents. A particular case, which is especially relevant for practical purposes, pertains to systems containing multiple cointegration relationships driven by different permanent shocks, as implied (e.g.) by the Classical Dichotomy. Within this context, either of the series pertaining to an individual cointegration relationship ends up playing the role of 'nuisance' series for the other cointegration relationships, thus decreasing the chances of detecting cointegration, sometimes dramatically so.

References

Balke, N., and R. Gordon (1986): "Appendix B: Historical Data," *in Gordon, R.J., ed. (1986), The American Business Cycle: Continuity and Changes, The University of Chicago Press.*

Barsky, R. (1987): "The Fisher Hypothesis and the Forecastability and Persistence of Inflation," *Journal of Monetary Economics*, 19(1), 3–24.

Benati, L. (2008): "Investigating Inflation Persistence Across Monetary Regimes," *Quarterly Journal of Economics*, 123(3), 1005–1060.

——— (2015a): "The Long-Run Phillips Curve: A Structural VAR Investigation," *Journal of Monetary Economics*, 76(November), 15–28.

——— (2015b): "Permanent and Transitory Components of House Prices Fluctuations," University of Bern, mimeo.

——— (2016): "A New Approach to Estimating the Natural Rate of Interest," University of Bern, mimeo.

Benati, L., and P. Ireland (2017): "What Drives the Dynamics of the Money Multiplier?," University of Bern and Boston College, in progress.

Benati, L., R. E. Lucas Jr., J.-P. Nicolini, and W. Weber (2016): "International Evidence on Long-Run Money Demand," *NBER Working Paper 22475, July 2016.*

Cavaliere, G., A. Rahbek, and R. Taylor (2012): "Bootstrap Determination of the Co-Integration Rank in Vector Autoregressive Models," *Econometrica*, 80(4), 1721–1740.

Cochrane, J. H. (1994): "Permanent and Transitory Components of GNP and Stock Prices," *Quarterly Journal of Economics*, 109(1), 241–265.

Elliot, G., T. Rothenberg, and J. H. Stock (1996): "Efficient Tests for an Autoregressive Unit Root," *Econometrica*, 64(4), 813–836.

Engle, R., and C. Granger (1987): "Cointegration and Error Correction: Representation, Estimation, and Testing," *Econometrica*, 55, 251–276.

Gallin, J. (2008): "The Long-Run Relationship between House Prices and Rents," *Real Estate Economics*, 36(4), 635Ű–58.

Hills, S., and R. T. N. Dimsdale (2010): "The UK Recession in Context: What Do Three Centuries of Data Tell Us?," *Bank of England Quarterly Bulletin*, (2010 Q4), 277–291.

JOHANSEN, S. (2002): "A Small Sample Correction of the Test for Cointegrating Rank in the Vector Autoregressive Model," *Econometrica*, 70, 1929–1961.

KENDRICK, J. W. (1961): *Productivity Trends in the United States.* Princeton University Press.

LUCAS JR., R. E., AND J.-P. NICOLINI (2015): "On the Stability of Money Demand," *Journal of Monetary Economics*, forthcoming.

LUETKEPOHL, H. (1991): *Introduction to Multiple Time Series Analysis, II edition.* Springer-Verlag.

METCALF, C., A. REDISH, AND R. SHEARER (1996): "New Estimates of the Canadian Money Stock: 1871-1967," *University of British Columbia Discussion Paper No. 96-17*, July 1996.

SHEARER, R. A., AND C. CLARK (1984): "Canada and the Interwar Gold Standard, 1920-35: Monetary Policy without a Central Bank," *in Michael D. Bordo and Anna J. Schwartz, editors, A Retrospective on the Classical Gold Standard, 1821-1931, University of Chicago Press*, pp. –.

SHIN, Y. (1994): "A Residual-Based Test of the Null of Cointegration against the Alternative of No Cointegration," *Econometric Theory*, 10(1 (Mar., 1994)), 91–115.

A The Dataset

Here follows a detailed description of the dataset. All of the *annual* series used in Section 2 are from Benati, Lucas, Nicolini, and Weber (2016), and they have been either updated ot extended when possible.

A.1 Canada

An annual series for nominal GDP, available since 1870, has been constructed by linking the Urquhart series (available from *Statistics Canada*, which is Canada's national statistical agency), for the period 1870-1924; series 0380-0515, v96392559 (1.1) from *Statistics Canada*, for the period 1925-1980; and series 0384-0038, v62787311 (1.2.38) from *Statistics Canada*, for the period 1981-2013. A series for the official discount rate, available since 1926, has been constructed as follows. Since 1934, when the Bank of Canada was created, it is simply the official bank rate ('Taux Officiel d'Escompte') from the Bank of Canada's website. Before that, I use the Advance Rate, which had been set by the Treasury Department for the discounting of bills, from Table 6.1 of Shearer and Clark (1984). To be precise, Shearer and Clark (1984) do not provide the actual time series for the Advance Rate, but rather the dates at which the rate had been changed (starting from August 22^{nd}, 1914), together with the new value of the rate prevailing starting from that date. Based on this information, I constructed a daily series for the rate starting on January 1^{st}, 1915 *via* a straightforward MATLAB program, and I then converted the series to the annual frequency by taking annual averages. A long-term interest rate available since 1919 ('Government of Canada Marketable Bonds - Average Yields - Over 10 Years') is from the Bank of Canada. A monthly series for M_1 starting in January 1872 is from Metcalf, Redish, and Shearer (1996), and it has been extended as follows. We use the series from Metcalf *et al.* (1996) until December 1952; after that, I link it *via* splicing to the series labelled as 'Currency and demand deposits, M_1 (x 1,000,000), v37213' until November 1981 from *Statistics Canada*; finally, from December 1981 until December 2006, I use the series from *Statistics Canada* labelled as 'M_1 (net) (currency outside banks, chartered bank demand deposits, adjustments to M_1 (continuity adjustments and inter-bank demand deposits) (x 1,000,000), v37200'. An important point to stress is that over the periods of overlapping, the three series are near-identical (up to a scale factor), which justifies their linking. On the other hand, for the period after December 2006 I was not able to find an M_1 series which could be reliably linked to the one I use for the period December 1981-December 2006 (over the last several decades, Canada's monetary aggregates have undergone a number of redefinitions, which complicates the task of constructing consistent long-run series for either of them). As a result, I decided to end the sample period in 2006. We convert the monthly series to the annual frequency by taking simple annual averages. Series for real GDP, real consumption, and real investment, available since 1926, are from

CANSIM (Canada's national statistical agency), and they have been constructed as follows. As for GDP, I linked the series v96730304 (available for the period 1926-1986) to the series v62471340 (available since 1981). Over the period of overlapping, the two series' rates of growth are very close, which justifies their linking. By the same token, for consumption I linked the series v96730305 and v62471303, which are available for the same periods as the corresponding GDP series.

A.2 United States

A.2.1 Annual data

The series for the 3-month Treasury Bill rate, nominal GDP, and the M_1 aggregate are from Benati *et al.* (2016). They were originally, from Lucas and Nicolini (2015), and they have been updated based on either series' updated original data sources. The original source for the 3-month Treasury Bill rate is the *Economic Report of the President* (henceforth, *ERP*), whereas the ones for nominal GDP are Kuznets and Kendrick's Table Ca184-191 before 1929, and Table 1.1.5 of the *National Income and Product Accounts* (henceforth, *NIPA*) after that. The series for the long-term rate is the yield on corporate bonds' series from Balke and Gordon (1986) until 1918. Since 1919 it is *Moody*'s Seasoned Baa Corporate Bond Yield series, which I downloaded from FRED II at the St. Louis FED's website (Balke and Gordon also used this series for the period since 1919). The series for real GDP is from Table 1.1.6 of the *NIPA* since 1929. Before that, it is from Officer and Williamson's website, http://www.measuringworth.com/. The series for real personal consumption expenditures and real gross private domestic investment are from Table 1.1.6 of the *NIPA* since 1929. Before that, they are from Table A-II*a* ('Gross National Product, Commerce Concept, Derivation from Kuznets Estimates, 1869-1957') of Kendrick (1961).

A.2.2 Quarterly data

Seasonally adjusted series for real and nominal GDP (GDPC96 and GDP, respectively) sre from the U.S. Department of Commerce, Bureau of Economic Analysis. The seasonally adjusted series for real chain-weighted investment and real chain-weighted consumption of non-durables and services have been computed based on the data found in Tables 1.1.6, 1.1.6B, 1.1.6C, and 1.1.6D of the National Income and Product Accounts. Whereas real consumption pertains to non-durables and services, real investment has been computed by chain-weighting the relevant series pertaining to durable goods; private investment in structures, equipment, and residential investment; Federal national defense and non-defense gross investment; and State and local gross investment. A seasonally adjusted series for real output per hour of all persons in the non-farm business sector (OPHNFB) is from the US. Bureau of Labor Statistics. In line with Lucas and Nicolini (2015), a quarterly, seasonally adjusted M_1

series is equal to M1SL from the St. Louis Fed's website (converted to the quarterly frequency by taking averages within the quarter) until 1983Q4, and it is equal to M1SL plus Money Market Deposit Accounts since then.

The remaining variables are originally available at the monthly frequency, and they have been converted to the quarterly frequency by taking averages within the quarter. The Federal funds rate (FEDFUNDS) and the 3-month Treasury bill rate (TB3MS) are taken from the St. Louis Fed's website. Civilian non-institutional population (CNP16OV) is from the U.S. Department of Labor, Bureau of Labor Statistics. The St. Louis Adjusted Monetary Base (AMBSL) and M_2 (M2SL) are from the St. Louis Fed's website.

A.2.3 Monthly data

A monthly seasonally adjusted series for the core personal consumption expenditure deflator ('PCEPILFE: Personal Consumption Expenditures Excluding Food and Energy, Chain-Type Price Index, Seasonally Adjusted, Monthly, Index 2009=100') is from the U.S. Department of Commerce: Bureau of Economic Analysis. A monthly seasonally adjusted series for the core CPI ('CPILFESL: Consumer Price Index for All Urban Consumers: All Items Less Food & Energy, Seasonally Adjusted, Monthly, Index 1982-84=100') is from the U.S. Department of Labor: Bureau of Labor Statistics. As for the rent series, the ideal one to use would be the owner's equivalent rent component of the CPI ('CUSR0000SEHC: Consumer Price Index for All Urban Consumers: Owners' equivalent rent of residences, Seasonally Adjusted, Monthly, Index December 1982=100'). The problem with this series is that it is comparatively short, as it starts in January 1983, thus limiting the analysis to just 32 years. In the paper I have therefore used, as rent series, the shelter component of the CPI ('CUSR0000SAH1: Consumer Price Index for All Urban Consumers: Shelter, Consumer Price Index, Seasonally Adjusted, Monthly, Index 1982-84=100'), which starts in January 1953, and over the common sample period has been remarkably close to the owner's equivalent rent component (both series are from the U.S. Department of Labor: Bureau of Labor Statistics). A monthly seasonally unadjusted series for the median sales price of new homes sold in the United States is from the website of the Manufacturing and Construction Division, Residential Construction Branch, of the U.S. Census Bureau (http://www.census.gov/construction/nrs/historical_data/historic_releases.html). The original seasonally unadjusted series has been seasonally adjusted *via* ARIMA X-12.

A.3 United Kingdom

All of the data for the United Kingdom are from version 2.3 of the *Bank of England*'s dataset of long-run historical statistics, which is available from the *Bank of England*'s website (the Excel spreadsheet is called threecenturies_v2.3.xlsx; henceforth, TC). The first version of the dataset was discussed in detail in Hills and Dimsdale (2010). Specifically, the following series are from the sheet 'A1. Headline series': nominal

GDP at market prices (column L), real GDP at market prices, chained volume measure, 2013 prices (column B), the *Bank of England*'s official discount rate (column AE), M_1 (column AW), and a '10 year/medium-term government bond yield' (column AG).