Multivariate cointegration and temporal aggregation: some further simulation results

Jesús Otero
Facultad de Economía
Universidad del Rosario, Bogotá

Theodore Panagiotidis
Department of Economics
University of Macedonia, Greece

Georgios Papapanagiotou
Department of Economics
University of Macedonia, Greece

Abstract

We perform Monte Carlo simulations to study the effect of increasing the frequency of observations and data span on the JohansenJohansen (1995) maximum likelihood cointegration testing approach, as well as on the bootstrap and wild bootstrap implementations of the method developed by Cavaliere et al. (2012, 2014). Considering systems with three and four variables, we find that when both the data span and the frequency vary, the power of the tests depend more on the sample length. We illustrate our findings by investigating the existence of long-run equilibrium relationships among four indicators prices of coffee.

Keywords: Monte Carlo, span, power, cointegration, coffee prices.
JEL Codes: .

Corresponding author: T. Panagiotidis (tpanag@uom.edu.gr)

1 Introduction

Progress in information technology for collecting, organising, storing and retrieving data is leading to the availability of such a vast amount of economic and financial variables, that practitioners need to decide carefully the type of data (annual, quarterly, monthly, weekly, daily or even higher frequency) that they will use in their empirical work. This choice is important because tests of time series properties typically exhibit low power (as those for integration and cointegration), so that finding ways of improving power, either by increasing the observation frequency (the number of observations per unit of time) or the data span (the period between the first and last observation), becomes relevant.[1] Shiller and Perron (1985) found that the power of unit root tests depends solely on the total sample length; see also Lahiri and Mamingi (1995). Hooker (1993) examined the Engle and Granger (1987) cointegration test and showed that, contrary to unit root tests, it gains power from temporal disaggregation. Rossana and Seater (1995) showed that temporal aggregation leads to substantial information losses and increases long-term persistence. Marcellino (1999) offered theoretical results that while time-series properties such as integration and cointegration are invariant to time aggregation, the finite-sample power of testing procedures may decrease when using temporally aggregated data. Later, Otero and Smith (2000) examined the Johansen (1988, 1995) cointegration test applied to two time series, and found that the test ability to detect cointegration depended more on the sample length than the number of observations. Haug (2002) found that, cointegration tests gain power when, for a given data span, a higher frequency is used.

This paper aims to complement and extend existing Monte Carlo evidence on the effects of increasing the data span and the frequency of observation on the finite-sample properties of cointegration tests. We focus on the Johansen (1988, 1995) rank test in systems with three and four variables. Applied macroeconomic modelling often

[1]See Tiao (1972) and Brewer (1973) for some early studies on the effects of temporal aggregation on the time series properties of economic and financial data.

deals with more than two variables, and so the results contained in this paper should be appealing to practitioners. More importantly, in recent contributions to the cointegration literature Cavaliere et al. (2012, 2014) recommend bootstrap and wild bootstrap implementations of the Johansen rank test. These bootstrap implementations aim to improve the finite sample performance of the likelihood ratio test, which could be quite poor as indicated by Johansen (2002). Unlike previous bootstrap procedures, see e.g. Swensen (2006), in the implementations of Cavaliere et al. (2012, 2014), the resulting bootstrap data are I(1) and satisfy the null co-integration rank, regardless of the true rank. Hence, it is interesting to examine the performance of these new bootstrap procedures in the context of temporal aggregation. As an empirical application, we reexamine Vogelvang (1992) analysis of the existence of long-run relationships among the spot prices of four coffee varieties over an extended period of time.

The paper proceeds as follows. Section 2 describes the design of the Monte Carlo experiments. Section 3 summarises the main findings of the simulations. Section 4 presents an empirical illustration based on Vogelvang (1992). Section 5 concludes.

2 Design of the Monte Carlo simulations

The design of our Monte Carlo experiments develops further the simulation setup considered by Engle and Granger (1987), Hooker (1993), Lahiri and Mamingi (1995) and Otero and Smith (2000) for the case of two time series. Accordingly, we consider two systems of equations consisting of three and four variables. First, we examine whether temporal aggregation causes size distortions (falsely rejecting the null hypothesis of no cointegration). Next, we examine the effect of temporal aggregation on the power of the tests. For the three-variables system we consider the cases of one and two cointegration vectors and for the four-variables system we consider the cases of one, two and three cointegration vectors. The specific data generating processes (DGPs) are presented in Table 1 along with their associated parameter values. In all simulations the error terms are assumed to follow independent standard normal distributions, and $u_{i,0} = 0$ for all values of $i = 1, 2, 3, 4$ where applicable. The following is an example of the DGP used to create a system of 4 variables with 3 cointegration relationships:

$$x_t = u_{3,t} + u_{4,t} - u_{1,t} - u_{2,t}, \quad u_1 = u_{1,t} + error$$

$$y_t = 2(u_{1,t} + u_{2,t}) - u_{3,t} - u_{4,t}, \quad u_2 = \rho_1 u_{2,t} + error$$

$$z_t = 2(u_{1,t} + u_{3,t} - u_{1,t}) - u_{2,t}, \quad u_1 = \rho_2 u_{3,t} + error$$

$$w_t = u_{2,t} + u_{4,t} - u_{1,t} - u_{3,t}, \quad u_1 = \rho_3 u_{4,t} + error$$

We consider two cases regarding the parameters ρ_i. In the frst case, we set $\rho_1 = 0.85$, $\rho_2 = 0.92$, $\rho_3 = 0.95$ and in the second case, we set $\rho_1 = 0.92$, $\rho_2 = 0.95$, $\rho_3 == 0.98$.

The generated series contain 1188 observations which is equivalent to 99 years of monthly observations. The Johansen test and the bootstrap algorithms are applied to the systems presented in Table 1 using sample sizes of 33, 66 and 99 years of monthly, quarterly and annual observations. We consider two methods of temporal aggregation. The first one keeps only the last observation of every quarter (year) and is often referred to as systematic sampling. The second one averages the three (twelve) non-overlapping observations corresponding to each quarter (year). In practice, the choice between the two methods ought to be guided by the type of variable under analysis, namely whether it is a stock (e.g. prices, wealth, foreign debt) or a flow (e.g. output, imports, exports). The size and power probabilities of the tests are computed at the 5% nominal level. We focus on the Trace test which tests the null hypothesis that the cointegration rank is less than or equal to r, denoted H(r), against the alternative of cointegration rank equal to p, denoted H(p), where p refers to the number of variables in the system. The null hypothesis is rejected when the trace statistic is greater than the corresponding critical value. We follow a sequential procedure which involves starting with $r = 0$ and testing $H(r)$ against $H(p)$ for $r = 0, 1, \ldots, p-1,$ until the null hypothesis is rejected for some value of r. The results are based on 5000 Monte Carlo replications, with 1000 bootstrap replications used to generate the bootstrap distributions of the tests. The R statistical language was employed.

3 Monte Carlo Simulation results

Table 2 reports size probabilities for the Trace test. The upper panel presents the probability of wrongly rejecting the null hypothesis of no cointegration in a system with three variables (DGP30), while the lower panel presents the results in a system with four variables (DGP40). Using the critical values from MacKinnon et al. (1999) MacKinnon et al. (1999) for inference, we observe that the trace test is approximately correctly sized for monthly data irrespective of the span of the data. However, there can be substantial size distortions that arise from temporal aggregation; the extent of the distortion diminishes as the data span increases though. For example, in DGP30 size is 17.5% when using 33 annual observations obtained through averaging; when the data span increases to 99 years the resulting size is 8.2%. The magnitude of the size distortions seems greater in DGP40. Size is controlled reasonably well for the two bootstrap methods (less than 5.8% for all samples).

Tables 3, 4 and 5 summarise the probability of identifying the exact number of existing cointegration vectors (where the numbers in parentheses indicate the number of observations in each sample). In the case of asymptotic inference, we use size-corrected critical values, while in the two bootstrap schemes we follow Cavaliere et al. (2012) (bootstrap) and Cavaliere et al. (2014) (wild bootstrap). Table 3 presents the results for the three-variables systems. The first two panels (DGP31) correspond to systems where a sole cointegration vector exists. The two lower panels correspond to systems with two cointegration vectors. For a given sample, reducing the data span yields substantial power losses than reducing the frequency of observations. For example, in all cases the test produces worse results when we use a sample of monthly data for 66 years (792 observations) instead of a sample of quarterly data for 99 years

(396 observations). In some cases, even a sample of annual data for 99 years (99 observations) yields better results than a sample of monthly data for 66 years (i.e. DGP31, $\rho = 0.98$ for all critical values).

Tables 4 and 5 present the results for systems with four variables. Panels DGP41, DGP42 and DGP43 refer to the cases of one, two and three cointegration vectors, respectively. Similar to Table 3, the values denote the probability of identifying the exact number of existing cointegration vectors while the numbers in parentheses the number of observations for each sample. As in the case of three variables, decreasing the data span leads to greater power losses than decreasing the frequency of observations. These losses are more apparent when changing the number of years from 66 to 33. In this case, the probability of detecting the correct number of cointegration vectors decreases from 80% to 40%. It is also worth examining the power of the test when applied samples with the same number of observations. The samples of 33 years of monthly observations and 99 years of quarterly observations, both contain 396 observations. Independently of the DGP, the test produces better results when employed on the latter sample.

For all DGPs except DGP30 and DGP40 (used to calculate the size probabilities) we consider two cases regarding the parameters ρ_i. While size is not affected by the different values of these parameters, the power of the test is. The test performs worse when the value of ρ_i approaches one. These findings are in line with the findings in Otero and Smith (2000). One possible explanation for this result is that the when the value of ρ_i approaches one, the stationarity of the auto-regressive process $u_{i,t}$ is affected. Qualitatively similar results (not reported here, but available upon request) are obtained for the λ_{max} test.

4 Empirical example

To illustrate our findings, we examine the relationship between coffee prices using the four composite "indicator prices" constructed by the International Coffee Organisation (ICO). Vogelvang (1992) conducted a similar study using quarterly data over the period 1960–1982 and identified two cointegrating vectors.[2] The largest database corresponds to monthly observations over the 1964–2018 period. Quarterly and annual versions of the data are obtained by skip sampling and averaging techniques. We then test for cointegration in the following sampling periods: 1964–2018, 1983–2018 and 2000–2018. We use the Schwarz (1978) information criterion to select the order of the underlying VAR models. The augmented Dickey–Fuller (ADF) test suggested by Said and Dickey (1984), based on Dickey and Fuller (1979), revealed that all four variables are integrated of order one.

Table 6 summarises our findings. Using the information for the shortest data span, 2000–2018, we are not able to detect evidence of cointegration among coffee prices neither with monthly nor quarterly data (the tests are not applied on annual data due to insuficient number of observations). Extending the data span so that it runs from 1983 to 2018 all tests identify two vectors using monthly data and up to one vector using quarterly and annual data. Extending further to cover the longest sample period available the tests detect two cointegration vectors using monthly and quarterly data, and one using annual data (when the skip sampling method is used the bootstrap algorithms identify two vectors).

5 Conclusions

We perform a large set of Monte Carlo simulations to assess the effects of increasing the frequency of observations and the data span on the Johansen (1988, 1995) multi-

[2]It is likely that changes over time in the structure and regulation of the coffee market may have caused structural breaks. The effect of temporal aggregation on such breaks is left as a topic for further research though.

variate cointegration Trace test, and its bootstrap and wild bootstrap implementations advocated by Cavaliere et al. (2012, 2014). We find evidence supporting the view that both the frequency and span of the data affect the power of the test. Our findings indicate that the ability to detect the presence of long-term cointegration relationships depends more on the total sample length. The importance of data span over frequency of observations is highlighted by the fact that the test performs better when applied on a smaller set of observations collected over a long time period than a large number of observations gathered over a relatively small time period.

References

Brewer, K. (1973). Some consequences of temporal aggregation and systematic sampling for arma and armax models. *Journal of Econometrics 1,* 133–154.

Cavaliere, G., A. Rahbek, and R. Taylor (2012). Bootstrap determination of the co-integration rank in vector autoregressive models. *Econometrica 80,* 1721–1740.

Cavaliere, G., A. Rahbek, and R. Taylor (2014). Bootstrap determination of the co-integration rank in heteroskedastic var models. *Econometric Reviews 33,* 606–650.

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,* 427–431.

Engle, R. and C. Granger (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica 55,* 251–276.

Haug, A. (2002). Temporal aggregation and the power of cointegration tests: A monte carlo study. *Oxford Bulletin of Economics and Statistics 64,* 399–412.

Hooker, M. (1993). Testing for cointegration: Power versus frequency of observation. *Economics Letters 41,* 359–362.

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control 12,* 231–254.

Johansen, S. (1995). *Likelihood-based inference in cointegrated vector autoregressive models*. Oxford University Press.

Johansen, S. (2002). A small sample correction for the test of cointegrating rank in the vector autoregressive model. *Econometrica 70,* 1929–1961.

Lahiri, K. and Mamingi (1995). Testing for cointegration: Power versus frequency of observation - another view. *Economics Letters 49,* 121–124.

MacKinnon, J., A. Haug, and L. Michelis (1999). Numerical distribution functions of likelihood ratio tests for cointegration. *Journal of Applied Econometrics 14,* 563–577.

Marcellino, M. (1999). Some consequences of temporal aggregation in empirical analysis. *Journal of Business and Economic Statistics 17,* 129–136.

Otero, J. and J. Smith (2000). Testing for cointegration: Power versus frequency of observation - Further Monte Carlo results. *Economics Letters 67,* 5–9.

Rossana, R. and J. Seater (1995). Temporal aggregation and economic time series. *Journal of Business and Economic Statistics 13,* 441–451.

Said, S. E. and D. A. Dickey (1984). Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika 71,* 599–607.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics 6,* 461–464.

Shiller, R. and P. Perron (1985). Testing the random walk hypothesis: Power versus frequency of observation. *Economics Letters 18,* 381–386.

Swensen, A. (2006). Bootstrap algorithms for testing and determining the cointegration rank in var models. *Econometrica 74,* 1699–1714.

Tiao, G. (1972). Asymptotic behaviour of temporal aggregates of time series. *Biometrika 59,* 525–531.

Vogelvang, E. (1992). Hypotheses testing concerning relationship between spot prices of various types of coffee. *Journal of Applied Econometrics 7,* 191–201.

Table 1: Design of the Monte Carlo simulations.

Name	Data generating process	Parameter values
DGP30	$x_t = x_{t-1} + error$	
	$y_t = y_{t-1} + error$	
	$z_t = z_{t-1} + error$	
DGP31	$x_t = u_{2,t} - u_{1,t}, \quad u_{1,t} = u_{1,t-1} + error$	
	$y_t = 2u_{1,t} - u_{2,t}, \quad u_{2,t} = \rho u_{2,t-1} + error$	
	$z_t = z_{t-1} + error$	$\rho = 0.95, 0.98$
DGP32	$x_t = u_{2,t} + u_{3,t} - u_{1,t}, \quad u_{1,t} = \rho_1 u_{1,t-1} + error$	$\rho_1 = 0.92, \rho_2 = 0.95$
	$y_t = u_{1,t} + u_{3,t} - u_{2,t}, \quad u_{2,t} = \rho_2 u_{2,t-1} + error$	$\rho_1 = 0.95, \rho_2 = 0.98$
	$z_t = 2(u_{1,t} + u_{2,t}) - u_{3,t}, \quad u_{3,t} = u_{3,t-1} + error$	
DGP40	$x_t = x_{t-1} + error$	
	$y_t = y_{t-1} + error$	
	$z_t = z_{t-1} + error$	
	$w_t = w_{t-1} + error$	
DGP41	$x_t = u_{2,t} + u_{3,t} - u_{1,t}, \quad u_{1,t} = u_{1,t-1} + error$	$\rho_1 = 0.95, \rho_2 = 1$
	$y_t = u_{1,t} + u_{3,t} - u_{2,t}, \quad u_{2,t} = \rho_1 u_{2,t-1} + error$	$\rho_1 = 0.95, \rho_2 = 0.98$
	$z_t = 2(u_{1,t} + u_{2,t}) - u_{3,t} u_{3,t} = \rho_2 u_{3,t-1} + error$	
	$w_t = w_{t-1} + error$	
DGP42	$x_t = 2u_{1,t} - u_{2,t}, \quad u_{1,t} = u_{1,t-1} + error$	$\rho_1 = 0.92, \rho_2 = 0.95$
	$y_t = u_{2,t} - u_{1,t}, \quad u_{2,t} = \rho_2 u_{2,t-1} + error$	$\rho_1 = 0.95, \rho_2 = 0.98$
	$w_t = \rho_1 w_{t-1} + error$	
	$z_t = z_{t-1} + error$	
DGP43	$x_t = u_{3,t} + u_{4,t} - u_{1,t} - u_{2,t}, \quad u_{1,t} = u_{1,t-1} + error$	$\rho_1 = 0.85, \rho_2 = 0.92, \rho_3 = 0.95$
	$y_t = 2(u_{1,t} + u_{2,t}) - u_{3,t} - u_{4,t}, \quad u_{2,t} = \rho_1 u_{2,t-1} + error$	$\rho_1 = 0.92, \rho_2 = 0.95, \rho_3 = 0.98$
	$z_t = 2(u_{1,t} + u_{3,t} - u_{2,t}) - u_{4,t}, \quad u_{3,t} = \rho_2 u_{3,t-1} + error$	
	$w_t = u_{2,t} + u_{4,t} - u_{1,t} - u_{3,t}, \quad u_{4,t} = \rho_3 u_{4,t-1} + error$	

Notes: In DGPXY, X refers to the number of variables and Y to the number of cointegration vectors. We con- sider two sets of parameters for all DGPs except DGP30 and DGP40 which are used for size.

Table 2: Size of the $Trace$ test.

Asymptotic inference			Bootstrap			Wild bootstrap			
33 years	66 years	99 years	33 years	66 years	99 years	33 years	66 years	99 years	
DGP30									
M	0.056(396)	0.051(792)	0.047(1188)	0.013(396)	0.017(792)	0.025(1188)	0.052(396)	0.052(792)	0.057(1188)
Q skip	0.059(132)	0.054(264)	0.055(396)	0.008(132)	0.018(264)	0.024(396)	0.046(132)	0.049(264)	0.056(396)
A skip	0.133(33)	0.078(66)	0.066(99)	0.023(33)	0.020(66)	0.019(99)	0.038(33)	0.048(66)	0.045(99)
Q avg	0.067(132)	0.059(264)	0.054(396)	0.012(132)	0.019(264)	0.023(396)	0.043(132)	0.054(264)	0.048(396)
A avg	0.175(33)	0.097(66)	0.082(99)	0.023(33)	0.021(66)	0.022(99)	0.045(33)	0.047(66)	0.046(99)
DGP40									
M	0.058(396)	0.057(792)	0.047(1188)	0.015(396)	0.009(792)	0.012(1188)	0.047(396)	0.046(792)	0.037(1188)
Q skip	0.065(132)	0.061(264)	0.052(396)	0.009(132)	0.012(264)	0.016(396)	0.060(132)	0.046(264)	0.045(396)
A skip	0.119(33)	0.082(66)	0.065(99)	0.016(33)	0.015(66)	0.012(99)	0.037(33)	0.046(66)	0.034(99)
Q avg	0.106(132)	0.111(264)	0.094(396)	0.008(132)	0.011(264)	0.014(396)	0.058(132)	0.049(264)	0.049(396)
A avg	0.214(33)	0.167(66)	0.131(99)	0.022(33)	0.018(66)	0.019(99)	0.041(33)	0.043(66)	0.041(99)

Notes: The table reports the probability of falsely identifying a cointegration relationship. The top panel refers to the three-variables system and lower panels refer to four-variables systems. Critical values are based on MacKinnon et al. (1999) (Asymptotic inference), Cavaliere et al. (2012) (Bootstrap) and Cavaliere et al. (2014) (Wild bootstrap). Numbers in parentheses denote the number of observations for each sample. M, Q, A stand for monthly, quarterly and annual data. Skip refers to systematic sampling method for aggregating data and avg to the averaging with non-overlapping observations method. The numbers in parentheses refer to the number of observations for each sample.

Table 3: SEmpirical power of the $Trace$ test for systems with three variables.

Asymptotic inference			Bootstrap			Wild bootstrap			
33 years	66 years	99 years	33 years	66 years	99 years	33 years	66 years	99 years	
DGP31, $\rho = 0.95$									
M	0.271(396)	0.813(792)	0.950(1188)	0.164(396)	0.704(792)	0.960(1188)	0.252(396)	0.806(792)	0.947(1188)
Q skip	0.254(132)	0.759(264)	0.934(396)	0.130(132)	0.618(264)	0.942(396)	0.208(132)	0.731(264)	0.939(396)
A skip	0.223(33)	0.503(66)	0.836(99)	0.028(33)	0.305(66)	0.709(99)	0.036(33)	0.402(66)	0.786(99)
Q avg	0.291(132)	0.682(264)	0.902(396)	0.151(132)	0.629(264)	0.922(396)	0.221(132)	0.729(264)	0.931(396)
A avg	0.261(33)	0.515(66)	0.825(99)	0.025(33)	0.304(66)	0.695(99)	0.033(33)	0.392(66)	0.767(99)
DGP31, $\rho = 0.98$									
M	0.084(396)	0.175(792)	0.382(1188)	0.037(396)	0.101(792)	0.270(1188)	0.074(396)	0.178(792)	0.382(1188)
Q skip	0.082(132)	0.181(264)	0.359(396)	0.033(132)	0.089(264)	0.252(396)	0.066(132)	0.162(264)	0.357(396)
A skip	0.084(33)	0.166(66)	0.302(99)	0.016(33)	0.066(66)	0.177(99)	0.021(33)	0.117(66)	0.263(99)
Q avg	0.083(132)	0.118(264)	0.328(396)	0.042(132)	0.104(264)	0.240(396)	0.076(132)	0.164(264)	0.340(396)
A avg	0.109(33)	0.103(66)	0.170(99)	0.017(33)	0.080(66)	0.205(99)	0.023(33)	0.125(66)	0.286(99)
DGP32, $\rho_1 = 0.92, \rho_2 = 0.95$									
M	0.486(396)	0.943(792)	0.947(1188)	0.079(396)	0.330(792)	0.664(1188)	0.119(396)	0.416(792)	0.756(1188)
Q skip	0.393(132)	0.932(264)	0.949(396)	0.065(132)	0.300(264)	0.637(396)	0.105(132)	0.383(264)	0.724(396)
A skip	0.146(33)	0.732(66)	0.945(99)	0.006(33)	0.193(66)	0.486(99)	0.013(33)	0.259(66)	0.571(99)
Q avg	0.425(132)	0.909(264)	0.921(396)	0.077(132)	0.336(264)	0.636(396)	0.116(132)	0.413(264)	0.718(396)
A avg	0.137(33)	0.701(66)	0.826(99)	0.005(33)	0.195(66)	0.515(99)	0.010(33)	0.256(66)	0.594(99)
DGP32, $\rho_1 = 0.95, \rho_2 = 0.98$									
M	0.064(396)	0.366(792)	0.729(1188)	0.052(396)	0.325(792)	0.675(1188)	0.083(396)	0.405(792)	0.757(1188)
Q skip	0.058(132)	0.337(264)	0.678(396)	0.039(132)	0.289(264)	0.636(396)	0.066(132)	0.360(264)	0.718(396)
A skip	0.053(33)	0.228(66)	0.549(99)	0.004(33)	0.151(66)	0.465(99)	0.009(33)	0.210(66)	0.553(99)
Q avg	0.059(132)	0.478(264)	0.720(396)	0.045(132)	0.312(264)	0.623(396)	0.073(132)	0.386(264)	0.702(396)
A avg	0.057(33)	0.248(66)	0.573(99)	0.003(33)	0.154(66)	0.475(99)	0.007(33)	0.210(66)	0.561(99)

Notes: The Table reports the probability of detecting the correct number of cointegration vectors existing in each system. In the case of asymptotic inference we use size-adjusted critical values, while in the two bootstrap schemes we follow Cavaliere et al. (2012) (Bootstrap) and Cavaliere et al. (2014) (Wild bootstrap). Numbers in parentheses denote the number of observations for each sample. M, Q, A stand for monthly, quarterly and annual data. Skip refers to systematic sampling method for aggregating data and avg to the averaging with non-overlapping observations method. The numbers in parentheses refer to the number of observations for each sample

Table 4: Empirical power of the $Trace$ test for systems with four variables (case 1).

Asymptotic inference			Bootstrap			Wild bootstrap			
33 years	66 years	99 years	33 years	66 years	99 years	33 years	66 years	99 years	
DGP41, $\rho_1 = 0.95, \rho_2 = 0.98$									
M	0.202(396)	0.640(792)	0.850(1188)	0.111(396)	0.583(792)	0.698(1188)	0.166(396)	0.627(792)	0.581(1188)
Q skip	0.193(132)	0.604(264)	0.786(396)	0.088(132)	0.506(264)	0.705(396)	0.131(132)	0.568(264)	0.607(396)
A skip	0.186(33)	0.432(66)	0.635(99)	0.015(33)	0.237(66)	0.595(99)	0.019(33)	0.298(66)	0.593(99)
Q avg	0.109(132)	0.644(264)	0.798(396)	0.107(132)	0.553(264)	0.687(396)	0.143(132)	0.591(264)	0.578(396)
A avg	0.132(33)	0.442(66)	0.591(99)	0.016(33)	0.244(66)	0.580(99)	0.020(33)	0.285(66)	0.569(99)
DGP42, $\rho_1 = 0.92, \rho_2 = 0.95$									
M	0.173(396)	0.795(792)	0.949(1188)	0.092(396)	0.685(792)	0.961(1188)	0.165(396)	0.782(792)	0.951(1188)
Q skip	0.131(132)	0.713(264)	0.934(396)	0.057(132)	0.565(264)	0.938(396)	0.105(132)	0.686(264)	0.942(396)
A skip	0.085(33)	0.358(66)	0.912(99)	0.003(33)	0.167(66)	0.620(99)	0.006(33)	0.253(66)	0.726(99)
Q avg	0.111(132)	0.560(264)	0.901(396)	0.067(132)	0.607(264)	0.939(396)	0.114(132)	0.708(264)	0.941(396)
A avg	0.099(33)	0.356(66)	0.724(99)	0.003(33)	0.148(66)	0.577(99)	0.005(33)	0.218(66)	0.681(99)
DGP42, $\rho_1 = 0.82, \rho_2 = 0.92, \rho_3 = 0.95$									
M	0.475(396)	0.939(792)	0.960(1188)	0.427(396)	0.950(792)	0.961(1188)	0.522(396)	0.944(792)	0.950(1188)
Q skip	0.364(132)	0.933(264)	0.958(396)	0.291(132)	0.929(264)	0.959(396)	0.377(132)	0.934(264)	0.949(396)
A skip	0.078(33)	0.681(66)	0.947(99)	0.010(33)	0.583(66)	0.929(99)	0.018(33)	0.670(66)	0.936(99)
Q avg	0.390(132)	0.922(264)	0.945(396)	0.309(132)	0.929(264)	0.954(396)	0.402(132)	0.934(264)	0.946(396)
A avg	0.067(33)	0.623(66)	0.923(99)	0.005(33)	0.508(66)	0.912(99)	0.010(33)	0.595(66)	0.922(99)

Notes: The Table reports the probability of detecting the correct number of cointegration vectors existing in each system. In the case of asymptotic inference we use size-adjusted critical values, while in the two bootstrap schemes we follow Cavaliere et al. (2012) (Bootstrap) and Cavaliere et al. (2014) (Wild bootstrap). Numbers in parentheses denote the number of observations for each sample. M, Q, A stand for monthly, quarterly and annual data. Skip refers to systematic sampling method for aggregating data and avg to the averaging with non-overlapping observations method. The numbers in parentheses refer to the number of observations for each sample

Table 5: Empirical power of the $Trace$ test for systems with four variables (case 2).

	Asymptotic inference			Bootstrap			Wild bootstrap		
	33 years	66 years	99 years	33 years	66 years	99 years	33 years	66 years	99 years
DGP41, $\rho_1 = 0.95, \rho_2 = 1$									
M	0.179(396)	0.547(792)	0.877(1188)	0.069(396)	0.386(792)	0.814(1188)	0.146(396)	0.535(792)	0.877(1188)
Q skip	0.179(132)	0.503(264)	0.831(396)	0.053(132)	0.302(264)	0.733(396)	0.114(132)	0.449(264)	0.821(396)
A skip	0.277(33)	0.371(66)	0.616(99)	0.015(33)	0.134(66)	0.390(99)	0.019(33)	0.209(66)	0.523(99)
Q avg	0.199(132)	0.627(264)	0.867(396)	0.062(132)	0.358(264)	0.752(396)	0.113(132)	0.490(264)	0.823(396)
A avg	0.326(33)	0.381(66)	0.637(99)	0.014(33)	0.137(66)	0.396(99)	0.018(33)	0.196(66)	0.507(99)
DGP42, $\rho_1 = 0.95, \rho_2 = 0.98$									
M	0.031(396)	0.160(792)	0.332(1188)	0.011(396)	0.090(792)	0.258(1188)	0.025(396)	0.149(792)	0.362(1188)
Q skip	0.029(132)	0.156(264)	0.311(396)	0.009(132)	0.072(264)	0.230(396)	0.022(132)	0.128(264)	0.336(396)
A skip	0.031(33)	0.086(66)	0.220(99)	0.001(33)	0.030(66)	0.131(99)	0.002(33)	0.057(66)	0.197(99)
Q avg	0.017(396)	0.082(264)	0.157(396)	0.009(132)	0.089(264)	0.259(396)	0.023(132)	0.143(264)	0.358(396)
A avg	0.022(33)	0.046(66)	0.109(99)	0.001(33)	0.030(66)	0.142(99)	0.002(33)	0.054(66)	0.213(99)
DGP42, $\rho_1 = 0.92, \rho_2 = 0.95, \rho_3 = 0.98$									
M	0.052(396)	0.383(792)	0.729(1188)	0.046(396)	0.339(792)	0.677(1188)	0.070(396)	0.413(792)	0.747(1188)
Q skip	0.047(132)	0.334(264)	0.695(396)	0.032(132)	0.285(264)	0.636(396)	0.053(132)	0.364(264)	0.712(396)
A skip	0.024(396)	0.178(66)	0.522(99)	0.001(396)	0.125(66)	0.441(99)	0.003(396)	0.174(66)	0.524(99)
Q avg	0.050(132)	0.271(264)	0.717(396)	0.037(132)	0.330(264)	0.671(396)	0.059(132)	0.407(264)	0.738(396)
A avg	0.031(33)	0.177(66)	0.530(99)	0.001(33)	0.122(66)	0.463(99)	0.003(33)	0.167(66)	0.533(99)

Notes: The Table reports the probability of detecting the correct number of cointegration vectors existing in each system. In the case of asymptotic inference we use size-adjusted critical values, while in the two bootstrap schemes we follow Cavaliere et al. (2012) (Bootstrap) and Cavaliere et al. (2014) (Wild bootstrap). Numbers in parentheses denote the number of observations for each sample. M, Q, A stand for monthly, quarterly and annual data. Skip refers to systematic sampling method for aggregating data and avg to the averaging with non-overlapping observations method. The numbers in parentheses refer to the number of observations for each sample

Table 6: Cointegration analysis of cofee prices ($p-$values only).

Sample period	H_0	2000-2018			1983-2018			1964-2018		
		$Trace$	B	BW	$Trace$	B	BW	$Trace$	B	BW
M	$r = 0$	0.187	0.134	0.280	0.000***	0.005***	0.018**	0.000***	0.001***	0.001***
	$r \leq 1$	0.484	0.341	0.335	0.009***	0.037**	0.096**	0.001***	0.002***	0.002***
	$r \leq 2$	0.578	0.505	0.492	0.188	0.197	0.158	0.274	0.278	0.238
	$r \leq 3$	0.512	0.650	0.616	0.233	0.257	0.267	0.226	0.194	0.153
A skip	$r = 0$				0.010***	0.111	0.134	0.000***	0.004***	0.005***
	$r \leq 1$				0.302	0.489	0.624	0.027	0.062*	0.038**
	$r \leq 2$				0.487	0.391	0.468	0.413	0.499	0.405
	$r \leq 3$				0.392	0.462	0.495	0.338	0.348	0.476
Q avg	$r = 0$	0.514	0.600	0.778	0.008***	0.010***	0.026**	0.000***	0.001***	0.002***
	$r \leq 1$	0.460	0.569	0.586	0.154	0.134	0.164	0.007***	0.008***	0.012**
	$r \leq 2$	0.335	0.398	0.391	0.103	0.085	0.111	0.221	0.240	0.178
	$r \leq 3$	0.580	0.576	0.504	0.207	0.210	0.258	0.263	0.248	0.151
A avg	$r = 0$				0.013**	0.101	0.086*	0.000***	0.004***	0.006***
	$r \leq 1$				0.344	0.582	0.540	0.167	0.285	0.239
	$r \leq 2$				0.368	0.352	0.331	0.276	0.327	0.254
	$r \leq 3$				0.323	0.377	0.439	0.249	0.225	0.225

Notes: B and WB denote the bootstrap and wild bootstrap versions of the test, respectively. All models are estimated for the restricted constant case. The table reports the p values for the Trace test. The asymptotic critical values are tabulated by MacKinnon et al. (1999) and the number of replications used in both bootstrap algorithms is set to 1000. ***, ** and * denote signifcance at the 1, 5 and 10% signifcance levels. We do not apply the tests on annual data from 2000 to 2018 due to insufcient number of observations.