Estimating cointegration parameters: an application of the double bootstrap

H.D. Vinod*, B.D. McCullough

Economics Department, Fordham University, Bronx, NY 10458-5198, USA

Abstract

The concept of cointegration, defined as long-run relationship between two or more random-walk variables, has spread widely in economics. Many methods have been offered for estimating the cointegrating parameter and its sampling distribution. Many of these estimations, while consistent, are biased, and have been criticized for reliance on asymptotic theory in samples which usually are of moderate size. We apply the double-bootstrap method to this problem. Specifically, we consider the cointegration between GNP and aggregate consumption, and study the marginal propensity to consume and the Keynesian multiplier.

1. Introduction

In economic terms, when C represents 'aggregate consumption' and Y represents 'aggregate disposable income,' the slope coefficient in the regression of C on Y has the interpretation dC/dY, which is often referred to as the 'marginal propensity to consume' (MPC). Recent advances in econometric theory (e.g. Engle and Granger, 1987) suggest that for certain time series data a fresh review of estimation and testing of regression coefficients may be needed. If C and Y are cointegrated, in the terminology of Engle and Granger, the traditional results do not apply.

Much research has been done on the topic of estimating and testing parameters in the context of cointegration (e.g. Engle and Granger, 1987; Phillips, 1986, 1987; Stock, 1987; Sims et al., 1990). These results rely on asymptotic theory, and have been criticized for reliance on asymptotic results in finite samples (e.g. Pagan and

* Corresponding author. E-mail: vinod@murray.fordham.edu.

This paper was presented in May 1991, at statistics 91 Canada meeting at Concordia University, Montreal, Canada. For brevity, we do not discuss some references.
Wickens, 1989, pp. 1008-1010). The bootstrap method offers a possible reply to such criticism.

Employing the terminology of Engle and Granger, in the following equation

\[ y = \beta_0 + \beta x \]  

if \( y \) and \( x \) are cointegrated, then \( \beta \) is referred to as the 'cointegrating parameter.' In this paper we consider the cointegration between \( C \) and \( Y \). Specifically, we wish to estimate the cointegrating parameter, \( \beta \), in the long-run relationship between consumption and income and to obtain good confidence intervals thereof. The various methods of estimating the cointegration parameter while biased, are consistent. However, obtaining appropriate sample sizes is problematic due to structural changes which occur in the economy. The cointegrating parameter might take on some value prior to a severe economic shock, for example the 1973 oil price increase, and another value afterwards. Suppose that before 1973 the cointegrating parameter assumed a value \( \beta_b \) and that after the recession it assumed a value \( \beta_a \). As the estimator of \( \beta \) is consistent, a large sample is desirable, but estimating across the structural breaks produces not an estimate of some \textit{true} \( \beta \) but, rather, a hybrid of the true parameters \( \beta_b \) and \( \beta_a \). There are no firm results on how large is large enough for a specific problem. A bootstrap can reveal the extent to which asymptotic normality is reasonable for inference, for example by using the method of Exploratory Data Analysis (Tukey, 1977; Velleman and Hoaglin, 1981).

The single bootstrap works best for an unbiased estimator. As it is well-known in the econometrics literature that many of the estimators of cointegrating parameters are biased but consistent; we therefore employ double bootstrap resampling methods to estimate the cointegrating parameter. Let \( \{x_n\} \) be a sample of size \( n \) with unknown distribution function \( F(\cdot) \). Let \( F_n \) be a consistent estimator of \( F(\cdot) \). We can obtain a bootstrap sample \( \{x_n^*\} \) whose components are conditionally independent, given \( \{x_n\} \). From each \( \{x_n^*\} \) we can further estimate \( F_n^* \), and thus obtain a second stage bootstrap sample \( \{x_n^{**}\} \), the components of which are conditionally independent, given \( \{x_n\} \) and \( \{x_n^*\} \).

Section 1 introduces the concept of cointegrated variables, Section 2 discusses the double bootstrap, Section 3 presents the results, and Section 4 offers the conclusions.

2. Cointegration case

Let \( x_t \) be a random variable and let \( \varepsilon_t \) be a stationary i.i.d. sequence with mean zero and variance \( \sigma^2 \), and suppose that \( \Delta x_t = x_t - x_{t-1} = \varepsilon_t \). Then \( x_t \) is non-stationary and is said to be \textit{integrated of order one}, denoted \( x_t \sim I(1) \) because its first-differences are stationary. An \( I(1) \) variable is often called a 'random-walk' variable.
Let $x_t$ and $y_t$ each be integrated of order one ($I(1)$). Then the equation

$$y_t = c + \beta x_t + \varepsilon_t$$

(2)

will lead to a *spurious regression* of the kind empirically documented by Granger and Newbold (1974, p. 115). Their ordinary least squares OLS estimator of $\beta$ yields significant $t$-statistic roughly three-quarters of the time in their Monte Carlo study when, in fact, the nominal size of the tests is 0.05 and the true parameter $\beta \equiv 0$ by construction. Intuitively, this type of regression ‘fails’ because the applicable central limit theorems require stationary variables. This phenomenon was explained by Phillips (1986), who showed that in this case, the limiting distribution of the *studentized* statistic, i.e. the ratio of the OLS estimator to its standard error, does not converge. Therefore the ratio of OLS estimator to its standard error is not a Student’s $t$ variable, even asymptotically. Since it diverges with non-zero probability, it leads to spuriously significant test results.

However, there are occasions when two seemingly random-walk variables are related, but not in the short run and a regression such as (2) does not lead to spurious results. Consider the spot and future prices for a commodity, each of which appears to be a random-walk variable. However, they cannot ‘walk’ too far apart, or arbitrage opportunities will arise. So, in the short run the series are unrelated, but over the long run they move together. Other pairs of cointegrated series are: stock prices and dividends; short and long-term interest rates; price levels and nominal wages; housing starts and population: to mention just a few.

Suppose that $x_t$ and $y_t$ are $I(1)$ variables, and define a linear combination $z_t$

$$z_t = x_t + \gamma y_t$$

(3)

Normally one would expect that a linear combination of $I(1)$ variables would also be $I(1)$. However, if there exists some $\gamma$ such that $z_t$ is $I(0)$, then the variables $x_t$ and $y_t$ are said to be *cointegrated of order one*, denoted $CI(1)$ (Engle and Granger, 1987). In this case $\beta$ of Eq. (2) is not a spurious parameter.

More formally, consider the following system of equations:

$$
\begin{pmatrix}
x_t \\
y_t
\end{pmatrix} = 
\begin{pmatrix}
x_{t-1} \\
y_{t-1}
\end{pmatrix} + v_t
$$

(4)

where $x_0 = y_0 = 0$, $\{v_t\}$ is a two-dimensional stationary process defined by

$$v_t = A(L)\eta_t = \sum_{j=0}^{x} A_j \eta_{t-j}, \quad \sum_{j=0}^{x} \|A_j\| < \infty, \quad A \equiv \begin{pmatrix} a_1^1 \\ a_2^1 \\ a_2^2 \end{pmatrix} = \sum_{j=0}^{x} A_j \neq 0$$

where $\{\eta_t\}$ is an i.i.d. $(0, I_2)$ sequence.

Since both variables are $I(1)$, the two eigenvalues of the system typically will be unity and OLS applied to either equation will be subject to the spurious regression phenomenon. On the other hand, if the variables are $CI(1)$, then one eigenvalue is unity and the other is zero, and OLS will be not only $\sqrt{T}$ consistent, but *super consistent* in the sense that the estimator has $T$-convergence (Stock, 1987).
When $y_t$ is the economic variable 'consumption' and $x_t$ is (disposable) 'income' then $\beta$ is interpreted as the MPC. The standard Keynesian model leads to a 'multiplier effect' on real (not just disposable) income of government-sponsored increases in consumption in the following fashion:

$$\text{MULR} = \frac{1}{(1 - \text{MPC})(1 - \lambda)} = 1 + \beta(0.7) + [\beta(0.7)]^2 + \cdots$$  \hfill (5)

where $\lambda$ is the average tax rate, e.g. 30% in the example at hand, assumed to be constant.

Observe that macroeconomists are interested in estimating the long-run multiplier, $\text{MULR} = [1 - \beta(1 - \lambda)]^{-1}$. In this example the statistic $[1 - \hat{\beta}(0.7)]^{-1}$ is of considerable interest. It has a random variable $\hat{\beta}$ in the denominator and the sampling distribution of this statistic can be shown to have the Cauchy distribution.

If two variables are cointegrated, then (2) does not lead to a spurious regression, as shown by Engle and Granger (1987). On the other hand, neither does the studentized statistic coverage to a unit normal. The asymptotic theory shows that cointegration leads to asymptotically non-normal distributions. The traditional asymptotic standard errors and confidence intervals are unreliable here, and the asymptotic results available for estimating cointegration parameters might be severely biased due to the small sample size. Bootstrapping is potentially valuable, especially for the problem of estimating sampling distributions which depend on unknown parameters. In any case, some calibration is useful.

In this paper we suppose that $A$ is of full rank and consider the sampling distribution of the OLS estimator of the MPC $= \beta$ from Eq. (2)

$$\hat{\beta} = \sum_{t=0}^{T} \hat{\xi}_t \hat{\eta}_t / \sum_{t=1}^{T} \hat{\xi}_t^2, \quad \hat{\xi}_t = x_t - \bar{x}, \quad \hat{\eta}_t = y_t - \bar{y}.$$

(6)

a root based our $\hat{\beta}$ and explained in the following sections. We are also interested in the corresponding long run multiplier defined in (5).

Before we turn to the following section, some clarification is in order. Consider a system of two equations comprising (2) and the first part of (4):

$$y_t = c + \beta x_t + \epsilon_t$$  \hfill (7)

and

$$x_t = x_{t-1} + v_{1t}.$$  \hfill (8)

In this paper we are considering OLS estimation of $\beta$ under the assumption that $\epsilon_t$ and $v_{1t}$ are uncorrelated. The possibility of nonzero correlation is an important reason why nonstandard (non-normal) distributions for hypothesis tests regarding cointegrating vectors are proposed in the literature. The possibility of using parametric bootstrap or Kunsch-type bootstrap for dependent data mentioned in Vinod’s (1991) survey is not considered here, for lack of software. Since the double bootstrap of the
following section is new, we wish to explore its potential for assessing the nonnormality. Although it cannot directly help for assessing the dependent data structure implied by the above correlation, it should reveal the nonnormality, which in turn may encourage development of software for dependent data.

3. The double bootstrap

Beran (1987) suggested pre-pivoting as a possible alternative to studentizing followed by a known monotone transformation. The idea was expanded in Beran (1990) who renamed pre-pivoting as the double bootstrap. The term pivoting in Sir Fisher's terminology is appropriate when the sampling distribution of a transformed statistic does not depend on unknowns. For example a studentized statistic. Beran's term root allows for a possible parameter dependence of the sampling distribution.

For clarity, let us consider the regression example, where the first stage bootstrap computes pseudo-\( y \) values denoted by a superscript \( * \). They are defined by

\[
y_j^* = Xb + e_j^*,
\]

where \( b = (X'X)^{-1}X'y \) and where the \( e_j^* \) are computed from the original \( e = y - Xb \) by sampling with replacement. It is customary to scale up the original residuals by \( \sqrt{T/(T-k)} \). For convenience we may treat \( b \) itself as a root and parameter dependence may arise due to the presence of lagged dependent variables in \( X \) or biased estimation of \( \beta \), etc. The estimates \( b_j^* \) from \( j = 1, \ldots, J \) (where \( J = 1000 \), say) bootstrap iterations are computed first. We store their (possibly rescaled) residual vectors as \( e_j^* \).

If we were to stop at the first stage bootstrap, and the estimator were unbiased, we would have a 'good' estimate of the cointegrating parameter in a finite sample, since \( E(b^*-b) \rightarrow 0 \) for an unbiased estimator. More generally, let us assume that the estimator is consistent, but possibly biased in small samples, and therefore we suggest using the double-bootstrap. The second stage bootstrap is designed to refine the first-stage estimates to take account of the finite-sample bias, nonnormality and parameter dependence of the sampling distribution of the consistent estimator. For each \( j \) we compute a \( T \times 1 \) vector of pseudo-\( y \) values defined by the equation:

\[
y_k^* = Xb_j^* + e_{jk}^*
\]

where \( e_{jk}^* \) are found from the \( (T) \)-vector of \( e_j^* \) values, by sampling with replacement. This is repeated for \( k = 1, 2, \ldots, K \) (\( K = 1000 \), say) second stage bootstrap iterations. Regressing these pseudo-\( y \) values on \( X \) yields \( b_{jk}^{**} \): \( k = 1, \ldots, K \) for each \( j \). and \( E(b_{jk}^{**} - b_{jk}^*) \rightarrow 0 \). Each second-stage estimate is expected to provide additional information beyond the first-stage estimate.

Let \( Z_j \) denote the fraction of \( b_{jk}^{**} \) values which are less than \( b_j^* \) in the second stage bootstrap over the \( K \) iterations. Note that \( Z_j \in [0,1] \) and uses a bootstrap estimate of the sampling distribution evaluated at \( b_j^* \). Each point is based on \( K \) iterations, and is expected to be more reliable, the larger is the value of \( K \).
The empirical CDF of $Z_j$ is uniform in $[0,1]$ if the statistic is a normal pivot and its sampling distribution does not depend on unknowns.

For sufficiently large $K$ and $J$, the empirical CDF of $Z_j$ over $j = 1, 2, ..., J$ approximates $H_{n,1}$, the second stage bootstrap sampling distribution of $Z_j$. Beran (1990) shows that the inverse denoted by $H_{n,1}^{-1}$ and evaluated at $(1 - \alpha)$, where $\alpha$ is the significance level, yields $c_{n,1}$, the preliminary critical value. We revise the preliminary value by using the inverse of $H_n$ the sampling distribution of the first stage bootstrap estimates $b^*_j$.

4. Results

We use the CITIBASE annual data for aggregate US consumption $C$ in 1982 dollars and the personal disposable income $Y$ in 1982 dollars, from 1948 to 1988. Since it is well known that consumption and income are cointegrated, Stock (1987), we omit details regarding OLS estimation. Let us turn directly to the possibly new information provided by the double bootstrap. Using annual data the OLS regression of $C$ on $Y$ yields an estimate of the marginal propensity to consume (MPC), which is now called the cointegrating parameter to be: 0.801503. The 1000 first-stage estimates, have a mean $b^* = 0.803042$. Now,

$$\text{bias}(b) = E[b - \beta] = E[(X'X)^{-1}X'e] = EW_e$$

is unobservable, where we have introduced the notation $W = (X'X)^{-1}X'$ for brevity. When the estimator is biased in small samples, one can estimate the finite sample bias (b) by considering $E[b^* - b] = E[We^*]$ and using its observable right-hand side (conditional on the sample). Efron and Tibshirani (1986) have shown that with the single bootstrap one can obtain a point estimate of the sampling bias which, on the assumption of the model, has zero expectation.

If the sampling distribution of the estimator is approximately normal and does not depend on unknown parameters, the single-bootstrap may be used to obtain an initial estimate of the bias (and sampling error), conditional on the sample. In our example, the initial estimate of the bias is:

$$E(b^* - b) = 0.803042 - 0.801503 = -0.001539,$$

which suggests a possibly biased estimator. However, the first-stage bootstrap estimate of the bias is itself possibly biased, conditional on the sample. The double-bootstrap can offer an improved estimate of the bias of a consistent estimator. It may be computed from the grand mean of all estimates of the cointegrating parameter. Note that it is not necessary to retain all $JK$ (one million, in our case) coefficients in the computer memory. One can simply retain the one thousand averages of the sampling distributions of each of the second stage iterations in memory and then compute a grand average of the $J(=1000)$ first stage averages. Now, from each first
Fig. 1.

stage bootstrap estimate, \( \hat{b}_j^* \) we obtain 1000 estimates \( \hat{b}_{jk}^{**} \) of \( \hat{b}_j^* \). These are averaged to produce 1000 values of \( \hat{b}_j^{**} \). The average of these, in turn, produces an estimate of \( \beta \) which is 0.8030463. This may be expected to improve upon both \( \hat{b}_j \), the OLS estimate, and the average \( = 0.8030422 \) of the first stage \( \hat{b}_j^* \) values. For this example, the two averages are virtually identical. Thus double bootstrap is useful when one is interested in estimating the bias of a consistent estimator for the finite-sample case.

Now we show that the double bootstrap permits an assessment of the extent of possible nonnormality of the sampling distribution of the estimated MPC, and the complication arising from parameter dependence of the sampling distribution. Fig. 1 plots the sampling distribution of \( \hat{b}_j^* \), which appears to be not too different from the normal curve plotted there. The usual numerical measures of skewness and kurtosis are respectively \(-0.077\) and \(3.073\) compared to 0 and 3 for the normal. By contrast, when we consider the sampling distribution of \( Z_j \), defined following (9) and plotted in Fig. 2 it does not appear to be remotely near the uniform distribution defined over the range \([0, 1]\). From the normal density shown in the same figure, the density appears to be a bit closer to the normal than the uniform. The mean and standard deviation of the uniform is 0.5 and 0.288675 respectively. These are compared to the mean 0.50711 and the standard deviation 0.05393 for the distribution of the \( Z_j \)'s. Similarly, the skewness of the uniform should be zero and kurtosis should be (below 3) 1.8, instead of 0.670 and 6.944 for the distribution of \( Z_j \)'s. The obvious lack of closeness to the uniform distribution arises from the nonnormality and/or parameter dependence of the sampling distribution. This confirms the results of the asymptotic theory for cointegration. The double bootstrap offers a convenient method of solving the problem.
Assume that we are interested in a 95% confidence interval. With $J = 1000$ realizations of the single bootstrap distribution, the lower limit of the 95% confidence interval is obviously the simple average of the 25th and 26th ordered values. The 2.5% quantile for the estimated distribution of $Z_j$'s is 0.4285. This suggests that the lower limit of the 95% interval should be at 428th ordered value. Similarly, the upper limit is at the 844th ordered value. This yields a narrow confidence interval of 0.802848 to 0.803458. The OLS value of 0.801503 is slightly below the lower limit of the 95% confidence interval from the double bootstrap, suggesting underestimation by the OLS, or a negative bias.

Improved understanding of the sampling distribution can obviously improve parameter estimation. For example, one can use a linear combination of ranked values including the trimean, trimmed means, R-estimation, L-estimation, etc. For our example the trimean $= 0.25(Q_1 + Q_3) + 0.5(\text{median})$ where $Q_1$ and $Q_3$ denote the first and the third quartile respectively, is 0.803060.

5. Conclusions

We demonstrate the potential uses of the double bootstrap for econometrics. This paper reports that the sampling distribution of the OLS estimate of the cointegrating relation between consumption and income is indeed nonnormal and parameter dependent. This confirms the prediction of the asymptotic theory. An appeal of the double bootstrap is that we can adjust for such nonnormality and/or parameter...
dependence in a particular problem quite simply. First, we locate the confidence interval of $Z_j$'s computed in the second stage bootstrap and use it to locate improved estimates of the quantiles of the first stage bootstrap. We have shown that the double bootstrap offers a new method of estimating the parameters of the basic (long run) cointegrating relations in economics. Clearly, more work is needed to evaluate and develop the double bootstrap in conjunction with non-iid errors.

References


