Journal of Econometrics 174 (2013) 66–81

Contents lists available at SciVerse ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Low-frequency robust cointegration testing[☆]

Ulrich K. Müller[a], Mark W. Watson[a,b,*]

[a] Princeton University, Department of Economics, Princeton, NJ, 08544, United States
[b] Woodrow Wilson School, Princeton, NJ, 08544, United States

ARTICLE INFO

Article history:
Received 21 December 2009
Received in revised form
3 May 2012
Accepted 9 September 2012
Available online 9 February 2013

JEL classification:
C32
C12

Keywords:
Stochastic trends
Persistence
Size distortion
Interest rates
Term spread

ABSTRACT

Standard inference in cointegrating models is fragile because it relies on an assumption of an $I(1)$ model for the common stochastic trends, which may not accurately describe the data's persistence. This paper considers low-frequency tests about cointegrating vectors under a range of restrictions on the common stochastic trends. We quantify how much power can potentially be gained by exploiting correct restrictions, as well as the magnitude of size distortions if such restrictions are imposed erroneously. A simple test motivated by the analysis in Wright (2000) is developed and shown to be approximately optimal for inference about a single cointegrating vector in the unrestricted stochastic trend model.

1. Introduction

The fundamental insight of cointegration is that while economic time series may be individually highly persistent, some linear combinations are much less persistent. Accordingly, a suite of practical methods have been developed for conducting inference about cointegrating vectors, the coefficients that lead to this reduction in persistence. In their standard form, these methods assume that the persistence is the result of common $I(1)$ stochastic trends,[1] and their statistical properties crucially depend on particular characteristics of $I(1)$ processes. But in many applications there is uncertainty about the correct model for the persistence which cannot be resolved by examination of the data, rendering standard inference potentially fragile. This paper studies efficient inference methods for cointegrating vectors that is robust to this fragility.

We do this using a transformation of the data that focuses on low-frequency variability and covariability. This transformation has two distinct advantages. First, as we have argued elsewhere (Müller and Watson, 2008), persistence ("trending behavior") and lack of persistence ("non-trending, $I(0)$ behavior") are low-frequency characteristics, and attempts to utilize high-frequency variability to learn about low-frequency variability are fraught with their own fragilities.[2] Low-frequency transformations eliminate these fragilities by focusing attention on the features of the data that are of direct interest for questions relating to persistence. In particular, as in Müller and Watson (2008), we suggest focusing on below business cycle frequencies, so that the implied *definition* of cointegration is that error correction terms have a flat spectrum below business cycle frequencies. The second advantage is an important by-product of discarding high frequency variability. The major technical challenge when conducting robust inference about cointegrating vectors is to control size over the range of plausible processes characterizing the model's stochastic common trends. Restricting attention to low frequencies greatly reduces the dimensionality of this challenge.

The potential impact of non-$I(1)$ stochastic trends on standard cointegration inference has long been recognized. Elliott (1998) provides a dramatic demonstration of the fragility of standard cointegration methods by showing that they fail to control size

☆ We thank participants of the Cowles Econometrics Conference, the NBER Summer Institute and the Greater New York Metropolitan Area Econometrics Colloquium, and of seminars at Chicago, Cornell, Northwestern, NYU, Rutgers, and UCSD for helpful discussions. Support was provided by the National Science Foundation through grants SES-0518036 and SES-0617811.

* Corresponding author at: Princeton University, Department of Economics, Princeton, NJ, 08544, United States.
E-mail address: mwatson@princeton.edu (M.W. Watson).

[1] See, for instance, Johansen (1988), Phillips and Hansen (1990), Saikkonen (1991), Park (1992) and Stock and Watson (1993).

[2] Perhaps the most well-known example of this fragility involves estimation of HAC standard errors, see Newey and West (1987), Andrews (1991), den Haan and Levin (1997), Kiefer et al. (2000), Kiefer and Vogelsang (2005), Müller (2007) and Sun et al. (2008).

0304-4076/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
doi:10.1016/j.jeconom.2012.09.006

when the common stochastic trends are not $I(1)$, but rather are "local-to-unity" in the sense of Bobkoski (1983), Cavanagh (1985), Chan and Wei (1987) and Phillips (1987).[3] In a bivariate model, Cavanagh et al. (1995) propose several procedures to adjust critical values from standard tests to control size over a range of values of the local-to-unity parameter, and their general approach has been used by several other researchers; Campbell and Yogo (2006) provides a recent example. Stock and Watson (1996), Jansson and Moreira (2006) and Elliott et al. (2012) go further and develop inference procedures with specific optimality properties in the local-to-unity model.

An alternative generalization of the $I(0)$ and $I(1)$ dichotomy is based on the fractionally integrated model $I(d)$, where d is not restricted to take on integer values (see, for instance, Baillie, 1996 or Robinson, 2003 for introductions). Fractional cointegration is then defined by the existence of a linear combination that leads to a reduction of the fractional parameter. A well-developed literature has studied inference in this framework: see, for instance, Velasco (2003), Robinson and Marinucci (2001, 2003); Robinson and Hualde (2003) and Chen and Hurvich (2003a,b, 2006). As in the local-to-unity embedding, however, the low-frequency variability of the common stochastic trends is still governed by a single parameter, since (suitably scaled) fractionally integrated series converge to fractional Brownian motions, which are only indexed by d. In contrast to the local-to-unity framework, this decisive parameter can be consistently estimated, so that the uncertainty about the exact nature of the stochastic trend vanishes in this fractional framework, at least under the usual asymptotics.

Yet, Müller and Watson (2008) demonstrate that relying on below business cycle variation, it is a hopeless endeavor to try to consistently discriminate between, say, local-to-unity and fractionally integrated stochastic processes from data spanning 50 years. Similarly, Clive Granger discusses a wide range of possible data generating processes beyond the $I(1)$model in his Frank Paish Lecture (Granger, 1993) and argues, sensibly in our opinion, that it is fruitless to attempt to identify the exact nature of the persistence using the limited information in typical macro time series. While local-to-unity and fractional processes generalize the assumption of $I(1)$ trends, they do so in a very specific way, leading to worries about the potential fragility of these methods to alternative specifications of the stochastic trend.

As demonstrated by Wright (2000), it is nevertheless possible to conduct inference about a cointegrating vector without knowledge about the precise nature of the common stochastic trends. Wright's idea is to use the $I(0)$ property of the error correction term as the identifying property of the true cointegrating vector, so that a stationarity test of the model's putative error correction term is used to conduct inference about the value of the cointegrating vectors. Because the common stochastic trends drop out under the null hypothesis, Wright's procedure is robust in the sense that it controls size under any model for the common stochastic trend. But the procedure ignores the data beyond the putative error correction term, and is thus potentially quite inefficient.

Section 2 of this paper provides a formulation of the cointegrated model in which the common stochastic trends follow a flexible limiting Gaussian process that includes the $I(1)$, local-to-unity, and fractional/long-memory models as special cases. Section 3 discusses the low-frequency transformation of the cointegrated model. Throughout the paper, inference procedures are studied in the context of this general formulation of the cointegrated model. The price to pay for this generality is that it introduces a potentially large number of nuisance parameters that characterize the properties of the stochastic trends and the relationship between the stochastic trends and the model's $I(0)$ components. In our framework, none of these nuisance parameters can be estimated consistently. The main challenge of this paper is thus to study efficient tests in the presence of nuisance parameters under the null hypothesis, and Sections 4–6 address this issue.

Using this framework, the paper then makes six contributions. The first is to derive lower bounds on size distortions associated with trend specifications that are more general than those maintained under a test's null hypothesis. For example, for tests constructed under a maintained hypothesis that the stochastic trends follow an $I(1)$ process, we construct lower bounds on the test's size when the stochastic trends follow a local-to-unity or more general stochastic process. Importantly, these bounds are computed not for a specific test, but rather for *any* test with a pre-specified power. The paper's second contribution is an upper bound on the power for any test that satisfies a pre-specified rejection frequency under a null that may be characterized by a vector of nuisance parameters (here the parameters that characterize the stochastic trend process). The third contribution is implementation of a computational algorithm that allows us to compute an approximation to the lowest upper power bound and, when the number of nuisance parameters is small, a feasible test that approximately achieves the power bound.[4] Taken together these results allow us to quantify both the power gains associated with exploiting restrictions associated with specific stochastic trend processes (for example, the power gains associated with the specializing the local-to-unity process to the $I(1)$ process), and the size distortions associated with these power gains when the stochastic trend restrictions do not hold. Said differently, these results allow us to quantify the benefits (in terms of power) and costs (in terms of potential size distortions) associated with restrictions on the stochastic process characterizing the stochastic trend. Section 4 derives these size and power bounds in a general framework, and Section 5 computes them for our cointegration testing problem.

The fourth contribution of the paper takes up Wright's insight and develops efficient tests based only on the putative error-correction terms. We show that these tests have a particularly simple form when the alternative hypothesis restricts the model's stochastic trends to be $I(1)$. The fifth contribution of the paper is to quantify the power loss associated with restricting tests to those that use only the error-correction terms rather than all of the data. This analysis shows that, in the case of single cointegration vector, a simple-to-compute test based only on the error-correction terms essentially achieves the full-data power bound for a general stochastic trend process, and is thus the efficient test. These results are developed in Section 6.

The paper's sixth contribution is empirical. We study the post-WWII behavior of long-term and short-term interest rates in the United States. While the levels of the interest rates are highly persistent, a suitably chosen linear combination of them is not, and we ask whether this linear combination corresponds to the term spread, the simple difference between long and short rates. More specifically we test whether the cointegrating coefficient linking long rates and short rates is equal to unity. This value cannot be rejected using a standard efficient $I(1)$ test (Wald or LR versions of Johansen's (1991) test), and we show that this result continues to hold under a general trend process. Of course, other values of the cointegrating coefficient are possible both in theory and in the data, and we construct a confidence set for the value of the cointegrating coefficient allowing for a general trend process and compare it to the confidence set constructed using standard $I(1)$ methods. These results are presented in Section 7.

2. Model

Let $p_t,\ t = 1, \ldots, T$ denote the $n \times 1$ vector of variables under study. This section outlines a time domain representation of

[3] Also see Elliott and Stock (1994) and Jeganathan (1997).

[4] The second and third contributions are applications of general results from a companion paper, Elliott et al. (2012), applied to the cointegration testing problem of this paper.