Journal of Econometrics 178 (2014) 398–413

Contents lists available at ScienceDirect

Journal of Econometrics

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

Asymptotic refinements of a misspecification-robust bootstrap for generalized method of moments estimators

Seojeong Lee *

School of Economics, Australian School of Business, University of New South Wales, Sydney, NSW 2052, Australia

ARTICLE INFO

Article history:
Received 11 July 2011
Received in revised form
17 September 2012
Accepted 25 May 2013
Available online 2 September 2013

JEL classification:
C14
C15
C31
C33

Keywords:
Nonparametric iid bootstrap
Asymptotic refinement
Edgeworth expansion
Generalized method of moments
Model misspecification

ABSTRACT

I propose a nonparametric iid bootstrap that achieves asymptotic refinements for t tests and confidence intervals based on GMM estimators even when the model is misspecified. In addition, my bootstrap does not require recentering the moment function, which has been considered as critical for GMM. Regardless of model misspecification, the proposed bootstrap achieves the same sharp magnitude of refinements as the conventional bootstrap methods which establish asymptotic refinements by recentering in the absence of misspecification. The key idea is to link the misspecified bootstrap moment condition to the large sample theory of GMM under misspecification of Hall and Inoue (2003). Two examples are provided: combining data sets and invalid instrumental variables.

1. Introduction

This paper proposes a novel bootstrap procedure for the generalized method of moments (GMM) estimators of Hansen (1982). It extends the existing literature by establishing the same asymptotic refinements for t tests and confidence intervals (CI's) (i) without recentering the bootstrap moment function, and (ii) without assuming correct model specification. In contrast, the conventional bootstrap achieves the refinements only if recentering is done and the assumed moment condition is correctly specified. Thus, the contribution of this paper may look too good to be true at first glance, but it becomes apparent once we realize that those two eliminations are in fact closely related, because recentering makes the bootstrap non-robust to misspecification.

Bootstrapping has been considered as an alternative to the first-order GMM asymptotic theory, which has been known to provide poor approximations of finite sample distributions of test statistics especially when the model is highly non-linear or the number of moments is large, e.g., Blundell and Bond (1998), Bond and Windmeijer (2005), Hansen et al. (1996), Kocherlakota (1990), and Tauchen (1986).[1] Hahn (1996) proves the first-order validity of the bootstrap distribution of GMM estimators. Hall and Horowitz (1996) show asymptotic refinements of the bootstrap for t tests and the J test (henceforth the Hall–Horowitz bootstrap). Andrews (2002) proposes a computationally attractive k-step bootstrap procedure based on the Hall–Horowitz bootstrap. Inoue and Shintani (2006) extend the Hall–Horowitz bootstrap by allowing correlation of moment functions beyond finitely many lags. Brown and Newey (2002) suggest an alternative bootstrap procedure using the empirical likelihood (EL) probability (henceforth the Brown–Newey bootstrap).

In the existing bootstrap methods for GMM estimators, recentering is critical. Horowitz (2001) explains why recentering is important when applying the bootstrap to overidentified moment condition models, where the dimension of a moment function is greater than that of a parameter. In such models, the sample mean of the moment function evaluated at the estimator is not necessarily equal to zero, though it converges almost surely to zero if the model is correctly specified. In principle, the bootstrap considers the sample and the estimator as if they were the

* Tel.: +61 2 9385 3325; fax: +61 2 9313 6337.
E-mail address: jay.lee@unsw.edu.au.
URL: http://sites.google.com/site/misspecified/.

[1] The 1996 special issue of the *Journal of Business & Economic Statistics* deals with this problem in various contexts.

0304-4076/$ – see front matter
http://dx.doi.org/10.1016/j.jeconom.2013.05.008

population and the true parameter, respectively. This implies that the bootstrap version of the moment condition, that the sample mean of the moment function evaluated at the estimator should equal zero, does not hold when the model is overidentified. Recentering makes the bootstrap version of the moment condition hold. The Hall–Horowitz bootstrap analytically recenters the bootstrap moment function with respect to the sample moment condition. The Brown–Newey bootstrap recenters the bootstrap moment condition by employing the EL probability in resampling the bootstrap sample. Thus, both the Hall–Horowitz bootstrap and the Brown–Newey bootstrap can be referred to as *the recentered bootstrap*.

A naive bootstrap is to apply the standard bootstrap procedure as is done for just-identified models, without any additional correction, such as recentering. However, it turns out that this naive bootstrap fails to achieve asymptotic refinements for t tests and CI's, and jeopardizes the first-order validity of the J test. Hall and Horowitz (1996) and Brown and Newey (2002) explain that the bootstrap and sample versions of test statistics would have different asymptotic distributions without recentering, because of the violation of the moment condition in the sample.

Although they address that the failure of the naive bootstrap is due to the misspecification in the sample, they do not further investigate the conditional asymptotic distribution of the bootstrap GMM estimator under misspecification. Instead, they eliminate the misspecification problem by recentering. In contrast, I observe that the conditional asymptotic covariance matrix of the bootstrap GMM estimator under misspecification is different from the standard one. The conditional asymptotic covariance matrix is consistently estimable by using the result of Hall and Inoue (2003), and I construct the t statistic of which distribution is asymptotically standard normal even under misspecification.

Hall and Inoue (2003) show that the asymptotic distributions of GMM estimators under misspecification are different from those of the standard GMM theory.[2] In particular, the asymptotic covariance matrix has additional non-zero terms in the presence of misspecification. Hall and Inoue's formulas for the asymptotic covariance matrix encompass the case of correct specification as a special case. The variance estimator using their formula is denoted by the Hall–Inoue variance estimator, hereinafter. Imbens (1997) also describes the asymptotic covariance matrices of GMM estimators robust to misspecification by using a just-identified formulation of overidentified GMM. However, his description is general, rather than being specific to the misspecification problem defined in this paper.

I propose a bootstrap procedure that uses the Hall–Inoue variance estimators in constructing the sample and the bootstrap t statistics. It ensures that the bootstrap t statistic satisfies the asymptotic pivotal condition without recentering. Moreover, the sample t statistic is also asymptotically pivotal regardless of misspecification in the population. In other words, my bootstrap applies to the robust t statistic which is studentized with the Hall–Inoue variance estimator. Therefore, it works without assuming correct model specification in the population, and is referred to as the misspecification-robust (MR) bootstrap. In contrast, the conventional first-order asymptotics as well as the recentered bootstrap would not work under misspecification, because the conventional t statistic is not asymptotically pivotal anymore.

The MR bootstrap achieves asymptotic refinements, a reduction in the error of test rejection probability and CI coverage probability by a factor of n^{-1} for symmetric two-sided t tests and symmetric percentile-t CI's, over the asymptotic counterparts. The magnitude of the error is $O(n^{-2})$, which is sharp. This is the same magnitude of error shown in Andrews (2002), that uses the Hall–Horowitz bootstrap for independent and identically distributed (iid) data with slightly stronger assumptions than those of Hall and Horowitz (1996).

I note that the MR bootstrap is not for the J test. To get the bootstrap distribution of the J statistic, the bootstrap should be implemented under the null hypothesis that the model is correctly specified. The recentered bootstrap imposes the null hypothesis of the J test because it eliminates the misspecification in the bootstrap world by recentering. In contrast, the MR bootstrap does not eliminate the misspecification and thus, it does not mimic the distribution of the J statistic under the null. Since the conventional asymptotic and bootstrap t tests and CI's are valid only in the absence of misspecification, it is important to conduct the J test and report the result that the model is not rejected. However, even a significant J statistic would not invalidate the estimation results if possible misspecification of the model is assumed and the validity of t tests and CI's is established under such an assumption, as is done in this paper.

Three papers in the literature are in a similar vein in terms of bootstrap methods under misspecification. Corradi and Swanson (2006) show the first-order validity of the block bootstrap for conditional distribution tests under dynamic misspecification. Kline and Santos (2012) examine the higher-order properties of the wild bootstrap in a linear regression model when the mean independent assumption of the error term is misspecified. In particular, a referee suggested to clarify the marginal contribution of this paper with respect to the work of Gonçalves and White (2004) which proves the first-order validity of the bootstrap for t tests based on the quasi-maximum likelihood (QML) estimators studentized with the misspecification-robust variance estimator of White (1982).

First, the QML estimator is a special case of the GMM estimator when one uses the first-order condition of the QML as the moment condition. This also puts an additional restriction that the model is just-identified. Therefore, this paper covers a broader class of models than Gonçalves and White (2004). For example, the proposed bootstrap applies to the two-stage least squares (2SLS) estimator. In addition, the definition of misspecified moment condition model should be distinguished from that of misspecified likelihood function. The former arises only when the model is overidentified, which implies that the first-order condition of the QML forms a correctly specified moment condition even if the likelihood function is misspecified. Thus, the misspecification-robust QML variance estimator corresponds to the conventional GMM variance estimator under correct specification, rather than the Hall–Inoue variance estimator.[3]

Second, Gonçalves and White (2004) neither provide a guidance whether to recenter or not, nor explain the relationship between recentering and misspecification. One of the contributions of Hall and Horowitz (1996) is that bootstrapping for GMM is nonstandard so that one should recenter the moment function to achieve asymptotic refinements. I argue that recentering can be detrimental and is not even needed if we use the Hall–Inoue variance estimator. The key idea is to link the misspecified moment condition in the bootstrap world to the large sample theory of GMM under misspecification of Hall and Inoue (2003).

The remainder of the paper is organized as follows. Section 2 discusses theoretical and empirical implications of misspecified models and explains the advantage of using the MR bootstrap t tests and CI's. Section 3 outlines the main result. Section 4 defines

[2] Hall and Inoue (2003) do not deal with bootstrapping, however.

[3] Hall and Inoue (2003) explain their marginal contribution over Gallant and White (1988), White (1996), and Maasoumi and Phillips (1982) in this regard.