Available online at www.sciencedirect.com

ScienceDirect

Journal of Econometrics 139 (2007) 217–236

JOURNAL OF Econometrics

www.elsevier.com/locate/jeconom

Approximating the distribution of the two-stage least squares estimator when the concentration parameter is small

D.S. Poskitt[a,*], C.L. Skeels[b]

[a]Department of Econometrics and Business Statistics, Monash University, Vic 3800, Australia
[b]Department of Economics, The University of Melbourne, Vic 3010, Australia

Available online 18 July 2006

Abstract

This paper presents a new approximation to the exact sampling distribution of the instrumental variables estimator in simultaneous equations models. It differs from many of the approximations currently available, Edgeworth expansions for example, in that it is specifically designed to work well when the concentration parameter is small. The approximation is remarkable in that simultaneously: (i) it has an extremely simple final form; (ii) in situations for which it is designed it is typically much more accurate than is the large sample normal approximation; and (iii) it is able to capture most of those stylized facts that characterize lack of identification and weak instrument scenarios. The development leading to the approximation is also novel in that it introduces techniques of some independent interest not seen in this literature hitherto.

JEL classification: C13; C16; C39

Keywords: Concentration parameter; IV estimator; Two-stage least squares; Simultaneous equations model; *t* approximation; Weak instruments

1. Introduction

In this paper, we present a new approximation to the exact sampling distribution of the instrumental variables (IV) estimator of the coefficients on the endogenous regressors in a

*Corresponding author. Tel.: +61 3 9905 9378; fax: +61 3 9905 5474.
E-mail address: don.poskitt@buseco.monash.edu.au (D.S. Poskitt).

0304-4076/$ - see front matter
doi:10.1016/j.jeconom.2006.06.011

single equation from a linear system of simultaneous equations. More specifically, we examine the properties of the two-stage least squares (2SLS) estimator and, as will be seen, obtain a particularly simple approximation which (i) is much more accurate than the large sample normal approximation, at least in those situations for which it is designed, and (ii) has the ability to capture many of the stylized facts that constitute the current state of knowledge. Manipulation of our results provides simple demonstrations of many of the qualitative characteristics that have been obtained under the different paradigms used to analyse weak identification, the related issue of weak instruments, and simultaneous equations models more generally.

Recent years have seen much exploration of the consequences of weak identification and weak instruments for estimation and inference in simultaneous equations models. The literature exploring this model reveals a variety of perspectives from which the problem has been considered, ranging from exact finite sample theory for totally or partially unidentified models (Phillips, 1989; Nelson and Startz, 1990a, b; Choi and Phillips, 1992), to local-to-zero asymptotics for identified (but asymptotically unidentified) models (Staiger and Stock, 1997; Wang and Zivot, 1998; Zivot et al., 1998, 2003), through to the many-instrument asymptotics of Bekker (1994).[1] More recently, a body of literature has developed that seeks to combine the many-instrument asymptotics of Bekker (1994) with the local-to-zero asymptotics of Staiger and Stock (1997), resulting in many-weak-instruments asymptotics; see, for example, Chao and Swanson (2005a, b) and Stock and Yogo (2005). These asymptotic approaches differ essentially in the structure of the sequence in which they nest the model of interest and, although the exact details may differ with the approach, certain stylized facts emerge from these studies as characterizing the sampling behaviour of IV estimators; including (i) sampling distributions that are complicated mixtures of Normal distributions, typically asymmetric about the parameter of interest, and (ii) non-standard asymptotic results with non-degenerate limiting distributions.

At the risk of getting ahead of ourselves, we find that the distributions of certain functions of the 2SLS estimator can be approximated by various members of the family of Student-t distributions.[2] Our approximation provides a framework that goes some way towards unifying the qualitatively similar but technically distinct results of Staiger and Stock (1997), Wang and Zivot (1998), Zivot et al. (1998, 2003), on the one hand, and Phillips (1989), Nelson and Startz (1990b) and Choi and Phillips (1992) on the other. For example, t distributions can be thought of as mixed-Normal distributions, a feature of many existing results. Similarly, the asymptotic normality implied by the many-instrument asymptotics of Bekker (1994) can also be obtained as a special case. Quite apart from its simplicity and its explanatory power, the approximation is of independent interest in view of the novelty of its development which, as far as we are aware, has not appeared in the econometrics literature heretofore.

The remainder of the paper has the following structure. In the next section, we will introduce the model and establish notation whilst considering a canonical transformation.

[1]Given the close relationship between weak instruments and a lack of identification, this literature can be traced back through to the work of inter alia Sargan (1983), Sims (1980) and Basmann (1963). For a more comprehensive treatment of the literature in this area see the survey by Stock et al. (2002).

[2]It has been known for some time that the distribution of the IV estimator is approximately multivariate t; see, for example, Phillips (1980, p. 870). However, the approximations presented here involve different parameterizations and, as we show below, they only reduce to existing results in certain special cases.