Journal of Econometrics 165 (2011) 128–136

Contents lists available at SciVerse ScienceDirect

Journal of Econometrics

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

Linear programming-based estimators in simple linear regression

Daniel Preve[a,*], Marcelo C. Medeiros[b]

[a] Department of Statistics, Uppsala University and School of Economics, Singapore Management University, Box 513, 751 20 Uppsala, Sweden
[b] Pontifical Catholic University of Rio de Janeiro, Brazil

ARTICLE INFO

Article history:
Available online 12 May 2011

Keywords:
Linear regression
Endogeneity
Linear programming estimator
Quasi-maximum likelihood estimator
Exact distribution

ABSTRACT

In this paper we introduce a linear programming estimator (LPE) for the slope parameter in a constrained linear regression model with a single regressor. The LPE is interesting because it can be superconsistent in the presence of an endogenous regressor and, hence, preferable to the ordinary least squares estimator (LSE). Two different cases are considered as we investigate the statistical properties of the LPE. In the first case, the regressor is assumed to be fixed in repeated samples. In the second, the regressor is stochastic and potentially endogenous. For both cases the strong consistency and exact finite-sample distribution of the LPE is established. Conditions under which the LPE is consistent in the presence of serially correlated, heteroskedastic errors are also given. Finally, we describe how the LPE can be extended to the case with multiple regressors and conjecture that the extended estimator is consistent under conditions analogous to the ones given herein. Finite-sample properties of the LPE and extended LPE in comparison to the LSE and instrumental variable estimator (IVE) are investigated in a simulation study. One advantage of the LPE is that it does not require an instrument.

1. Introduction

The use of certain linear programming estimators in time series analysis is well documented. See, for instance, Davis and McCormick (1989), Feigin and Resnick (1994) and Feigin et al. (1996). LPEs can yield much more precise estimates than traditional methods such as conditional least squares (e.g. Datta et al., 1998; Nielsen and Shephard, 2003). The limited success of these estimators in applied work can be partially explained by the fact that their point process limit theory complicates the use of their asymptotics for inference (e.g. Datta and McCormick, 1995).

In regression analysis, it is well known that the ordinary least squares estimator is inconsistent for the regression parameters when the error term is correlated with the explanatory variables of the model. In this case an instrumental variables estimator or the generalized method of moments may be used instead. In economics, such endogenous explanatory variables could be caused by measurement error, simultaneity or omitted variables. To the authors' knowledge, however, there has so far been no attempt to investigate the statistical properties of LP-based estimators in a cross-sectional setting. In this paper we show that LPEs can, under certain circumstances, be a preferable alternative to LS and IV estimators for the slope parameter in a simple linear regression model. We look at two types of regressors which are likely to be of practical importance. First, we introduce LPEs to the simple case of a non-stochastic regressor. Second, we consider the general case of a stochastic, and potentially endogenous, regressor. For both cases we establish the strong consistency and exact finite-sample distribution of a LPE for the slope parameter.

The LPE can be used in situations where the regressor is strictly positive. For example, in empirical finance, we can consider regressions involving volatility and volume. In labor economics a possible application is the regression between income and schooling, for example.

The remainder of the paper is organized as follows. In Section 2, we establish the strong consistency and exact finite-sample distribution of the LPE when (1) the explanatory variable is non-stochastic, and (2) the explanatory variable is stochastic and potentially endogenous. In Section 3, we discuss how our results can be extended to other endogenous specifications and give conditions under which the LPE is consistent in the presence of serially correlated, heteroskedastic errors. We also describe how the LPE can be extended to the case with multiple regressors. Section 4 reports the simulation results of a Monte Carlo study comparing the LPE and extended LPE to the LSE and IVE. Section 5 concludes. Mathematical proofs are collected in the Appendix. An extended Appendix available on request from the authors contains some results mentioned in the text but omitted from the paper to save space.

* Corresponding author.
E-mail addresses: daniel.preve@statistics.uu.se (D. Preve), mcm@econ.puc-rio.br (M.C. Medeiros).

0304-4076/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jeconom.2011.05.011

2. Assumptions and results

2.1. Non-stochastic explanatory variable

The first regression model we consider is

$$\begin{cases} y_i = \beta x_i + u_i \\ u_i = \alpha + \varepsilon_i, \quad i = 1, \ldots, n \end{cases}$$

where the response variable y_i and the explanatory variable x_i are observed, and u_i is the unobserved non-zero mean random error. β is the unknown regression parameter of interest. We assume that $\{x_i\}$ is a nonrandom sequence of strictly positive reals, whereas $\{u_i\}$ is a sequence of independent identically distributed (i.i.d.) nonnegative random variables (RVs). For ease of exposition we assume that $E(u_i) = \alpha$. The potentially unknown distribution function F_u of u_i is allowed to roam freely subject only to the restriction that it is supported on the nonnegative reals. A well known continuous probability distribution with nonnegative support is the Weibull distribution, which can approximate the shape of a Gaussian distribution quite well.

A 'quick and dirty' estimator of the slope parameter, based on the nonnegativity of the random errors, is given by

$$\hat{\beta} = \min\left\{\frac{y_1}{x_1}, \ldots, \frac{y_n}{x_n}\right\}. \tag{1}$$

This estimator has been used to estimate β in certain constrained first-order autoregressive time series models, $y_i = \beta x_i + u_i$, with $x_i = y_{i-1}$ (e.g. Datta and McCormick, 1995; Nielsen and Shephard, 2003). As it happens, (1) may be viewed as the solution to the linear programming problem of maximizing the objective function $f(\beta) = \beta$ subject to the n constraints $y_i - \beta x_i \geq 0$. Because of this we will sometimes refer to $\hat{\beta}$ as a LPE. Regardless if the regressor is stochastic or non-stochastic, (1) is also the maximum likelihood estimator (MLE) of β when the errors are exponentially distributed. What is interesting, however, is that $\hat{\beta}$ consistently estimates β for a wide range of error distributions, thus the LPE is also a quasi-MLE.

Assumption 1 holds throughout the section.

Assumption 1. Let $y_i = \beta x_i + u_i$ $(i = 1, \ldots, n)$ where $u_i = \alpha + \varepsilon_i$ and

(i) $\{x_i\}$ is a nonrandom sequence of strictly positive reals,
(ii) 0 is not a limit point of $S \equiv \{x_1, x_2, \ldots\}$,
(iii) $\{u_i\}$ is an i.i.d. sequence of nonnegative RVs,
(iv) $\inf\{u : F_u(u) > 0\} = 0$,
(v) $E(\varepsilon_i) = 0$.

Note that β can be any real number and that conditions (iii) and (v) combined imply that the mean of u_i is $\alpha \geq 0$. Since $\hat{\beta}_n - \beta = R_n$, where $R_n = \min\{u_i/x_i\}$, it is clear that $P(\hat{\beta}_n - \beta \leq z) = 0$ for all $z < 0$ and, hence, the LPE is positively biased. Moreover, as (1) is nonincreasing in the sample size its accuracy either remains the same or improves as n increases. Proposition 1 gives the exact distribution of the LPE in the case of a non-stochastic regressor.

Proposition 1. *Under Assumption 1,*

$$P(\hat{\beta}_n - \beta \leq z) = 1 - \prod_{i=1}^{n}[1 - F_u(x_i z)].$$

The proof of the proposition follows from the observation that

$$P(\hat{\beta}_n - \beta \leq z) = P(R_n \leq z) \stackrel{\text{(i)}}{=} 1 - P(u_1 > x_1 z, \ldots, u_n > x_n z),$$

and condition (iii) of Assumption 1. By condition (iv), $F_u(u) > 0$ for every $u > 0$ implying that $\hat{\beta}$ consistently estimates β.[1]

Table 1
Ratio distributions with accompanying moments of $\hat{\beta}_n$. $F_z(z)$ is the cdf of the ratio $z = u_1/x_1$, with parameter $\theta = \theta_u/\theta_x$, on which the moments are based. Results hold for $\gamma = 0$ and $n > 2$. $\Gamma(\cdot)$ is the gamma function.

Ratio	$F_z(z), z > 0$	$E(\hat{\beta}_n - \beta)$	$\mathrm{Var}(\hat{\beta}_n - \beta)$
$\frac{\mathrm{Exp}(\theta_u)}{\mathrm{Exp}(\theta_x)}$	$1 - \frac{1}{1+\theta^{-1}z}$	$\frac{\theta}{n-1}$	$\frac{\theta^2 n}{(n-2)(n-1)^2}$
$\frac{U(0,\theta_u)}{U(0,\theta_x)}$	$\frac{1}{2\theta}z, z \leq \theta$; $1 - \frac{\theta}{2z}, z > \theta$	$\frac{2\theta}{n+1}\left[1 + \frac{1}{(n-1)2^n}\right]$	$O\left(\frac{1}{n^2}\right)$
$\frac{\mathrm{Ra}(\theta_u)}{\mathrm{Ra}(\theta_x)}$	$1 - \frac{1}{1+\theta^{-2}z^2}$	$\frac{\theta\sqrt{\pi}}{2}\frac{\Gamma(n-1/2)}{\Gamma(n)}$	$\theta^2\left[\frac{1}{n-1} - \frac{\pi}{4}\frac{\Gamma^2(n-1/2)}{\Gamma^2(n)}\right]$

Intuitively, this is because the left-tail condition on u_i implies that the probability of obtaining an error arbitrarily close to 0 is non-zero and, hence, that (1) is likely to be precise in large samples.

Corollary 1. *Under Assumption 1, $\hat{\beta}_n \stackrel{a.s.}{\to} \beta$ as $n \to \infty$.*

From Corollary 1 it follows that α (the unknown mean of the error term) can be consistently estimated by

$$\hat{\alpha} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{\beta}x_i), \tag{2}$$

the sample mean of the residuals, under fairly weak conditions.[2]

It is worth noting that the MLE of β satisfies the stochastic inequality $\hat{\beta}_{\mathrm{ML}} \leq \hat{\beta}$. Regardless if x_i is stochastic or non-stochastic, in some cases the LPE will be equal to $\hat{\beta}_{\mathrm{ML}}$. For instance, it is readily verified that if the random errors are (1) exponentially distributed with non-zero density function $(1/a)\exp\{-u/a\}$ for $u \geq 0$

$$\hat{\beta}_{\mathrm{ML}} = \hat{\beta}, \qquad \hat{a}_{\mathrm{ML}} = \hat{\alpha}, \tag{3}$$

and (2) uniformly distributed on the interval $[0, b]$

$$\hat{\beta}_{\mathrm{ML}} = \hat{\beta}, \qquad \hat{b}_{\mathrm{ML}} = \max\{y_i - \hat{\beta}x_i\}. \tag{4}$$

As an illustration of Proposition 1 in action, Corollary 2 shows that the exact distribution of $\hat{\beta} - \beta$ when the errors are Weibull distributed, and the regressor is non-stochastic, is also Weibull. The Weibull distribution, with distribution function $1 - \exp\{-(u/a)^b\}$ for $u \geq 0$, nests the well known exponential $(b = 1)$ and Rayleigh $(b = 2)$ distributions.

Corollary 2. *Let the regression errors be Weibull distributed. Then, under Assumption 1,*

$$P(\hat{\beta}_n - \beta \leq z) = 1 - \exp\left\{-\left[\frac{z}{a\left(\sum_{i=1}^{n} x_i^b\right)^{-1/b}}\right]^b\right\},$$

if $z \geq 0$ and 0 otherwise. Hence, $\hat{\beta}_n - \beta$ is Weibull with scale parameter $a\left(\sum_{i=1}^{n} x_i^b\right)^{-1/b}$ and shape parameter b.

For example, in view of Corollary 2 with $b = 1$ it is clear that

$$\sum_{i=1}^{n} x_i(\hat{\beta}_n - \beta),$$

is exponentially distributed with scale parameter a. Moreover, by (3) and basic results of large sample theory, the statistic

$$\frac{1}{\hat{\alpha}_n}\sum_{i=1}^{n} x_i(\hat{\beta}_n - \beta),$$

is asymptotically standard exponential.

[1] If x_i instead is assumed to be strictly negative then the estimator $\max\{y_i/x_i\}$ is strongly consistent for β.

[2] If, under Assumption 1, $\alpha < \infty$ and if $n^{-1}\sum_{i=1}^{n} x_i$ is $O(1)$ as $n \to \infty$.