Journal of Multivariate Analysis 130 (2014) 376–386

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

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

Semiparametric efficiency for partially linear single-index regression models

Tao Chen, Thomas Parker *

Department of Economics, University of Waterloo, 200 University Avenue, Ontario, Canada N2L 3G1

ARTICLE INFO

Article history:
Received 16 April 2013
Available online 17 June 2014

AMS subject classifications:
62G20
62P20

Keywords:
Semiparametric efficiency
Partially linear single-index model
Single-index model
Conditional mean restriction
Conditional quantile restriction

ABSTRACT

We calculate semiparametric efficiency bounds for a partially linear single-index model using a simple method developed by Severini and Tripathi (2001). We show that this model can be used to evaluate the efficiency of several existing estimators.

1. Introduction

A straightforward and convenient method for the calculation of semiparametric efficiency bounds was proposed in [22] and illustrated using several example models. In this paper, we extend their method to investigate semiparametric efficiency for the finite-dimensional parameters of the model

$$y = g_0(X'\theta_0) + Z'\beta_0 + \varepsilon, \tag{1.1}$$

where θ_0 and β_0 are unknown, $g_0(\cdot)$ is an unknown function and ε has a distribution conditional on covariates X and Z. This model encompasses several interesting special cases — the linear-, partially linear-, single-index- and partially linear single-index models — and the method used here makes it simple to account for various identification conditions. Our method is that illustrated by [22] (explained in greater detail in the recent survey [23]), and we extend their results to semiparametric models of conditional quantiles, which were not considered by those authors.

A number of estimators of model (1.1) and related models have been proposed. Semiparametric efficient locally-linear quasi-likelihood estimation of this model was proposed in [2]; they also showed that their estimator reached the bound (and also verified that the estimator in [10] attained this bound). Sieve estimators of θ_0 in the single-index model have been proposed by [4,5]. Computationally attractive estimators were proposed in [27,12] for the single-index model under conditional quantile identification conditions. [28,26] proposed estimators of model (1.1) for a conditional mean identification condition. In addition, [9] have proposed an estimator for a closely-related group of models of conditional quantiles. It is of interest to know whether these estimators attain the relevant efficiency bounds.

* Corresponding author.
E-mail addresses: t66chen@uwaterloo.ca (T. Chen), tmparker@uwaterloo.ca (T. Parker).

http://dx.doi.org/10.1016/j.jmva.2014.06.006
0047-259X/

In this article we use a method due to [22] to derive the semiparametric efficiency bound for this model in a straightforward manner, independent of assumptions regarding identification or type of estimator. Model (1.1) is not addressed in [22], and we extend their method to this model without making any more restrictions on the conditional distribution of ε given covariates. We then use our bound to compare to results for conditional mean- and quantile location identification conditions. This method makes it easy to derive the general bound (i.e., for the "least-favorable parametric submodel") in a clear manner without going through the usual two-step style calculation, as represented for example by [20,7], and we view it as a complement to that model, which appears to be more well-suited to calculations for special cases (like efficiency bounds for the class of M-estimators, for example).

2. General assumptions

We assume the random variables $y \in \mathbb{R}$, $W = [X' \quad Z']' \in \mathbb{R}^p$, and $\varepsilon \in \mathbb{R}$ have densities $q_0^2(y|W)$, $b_0^2(W)$ and $\gamma_0^2(\varepsilon|W)$ respectively. The likelihood associated with an observation (y, W) is

$$\mathscr{L}(\theta; y, W) = q_0^2(y|W) b_0^2(W)$$

which could alternatively be expressed using the conditional density of the additive error term ε, because the model (1.1) implies q_0 and γ_0 satisfy the equation $q_0(y|W) = \gamma_0(y - g_0(X'\theta_0) + Z'\beta_0|W)$. Assume $g_0 \in \mathrm{L}^2(\mathbb{R}, \lambda)$ and has a derivative $g_0' \in \mathrm{L}^2(\mathbb{R}, \lambda)$, where the notation $\mathrm{L}^2(\mathcal{A}, \mu)$ denotes the space of square-integrable functions on some domain $\mathcal{A}$ with respect to some measure μ, and λ is Lebesgue measure. Because g_0 is unknown, we make the definitions $X_0 = X'\theta_0$ (where X' denotes "X transpose" below) and $W_0 = [X_0 \quad Z']' \in \mathbb{R}^{p_0}$, and impose the "index restriction" $q_0(y|W) = q_0(y|W_0)$ (equivalently, $\gamma_0(\varepsilon|W) = \gamma_0(\varepsilon|W_0)$). This differs slightly from previous partial-index models in the literature (e.g., [20]) that assumed the variance function was a fully nonparametric function of W. As is pointed out by [25], when considering estimation, such models may suffer from the curse of dimensionality. Our results are relevant for models with variance functions that generally depend only on W_0. This restriction has some precedence in the literature; for example, [10] restrict their attention to similar cases. Finally, we note that we implicitly assume the model is identified. In practice, this would mean for example, assuming the first element of θ_0 is normalized to 1 and the first element of X is continuously distributed, as well as using trimming in an estimator to ensure the positivity of the density of $X'\theta_0$; however, we abstract away from these details to focus on the technique used to derive the efficiency bound.

We make minimal assumptions regarding b_0^2, the marginal density of W: we assume that b_0 is a member of the space $\mathcal{B}$, where

$$\mathcal{B} = \left\{ b \in \mathrm{L}^2(\mathbb{R}^p, \lambda) : b^2(w) > 0, \int_{\mathbb{R}^p} b^2(w) \mathrm{d}w = 1 \right\},$$

and the additional identification assumption that $\mathrm{E}\left[(W - \mathrm{E}[W|X_0])(W - \mathrm{E}[W|X_0])'\right]$ exists and is nonsingular (one could assume only a generalized inverse, as in [26] for a model similar to (1.1)). Assume that $\gamma_0 \in \Gamma$, where

$$\Gamma = \left\{ \gamma : \mathbb{R} \times \mathbb{R}^p \to \mathbb{R} : \gamma(u|W) = \gamma(u|X'\theta_0, Z), \gamma^2(u|W) > 0, \int_{\mathbb{R}} \gamma^2(u|W) \mathrm{d}u = 1, \right.$$
$$\left. \gamma(u|W) \text{ is bounded and continuous, and } \int_{\mathbb{R}} \left(\gamma'(u|W)\right)^2 \mathrm{d}u < \infty, \text{ all } w.p.1 \right\}$$

where γ' refers to a partial derivative with respect to u — that is, $\gamma'(u|w) = \partial\gamma(u|w)/\partial u$. Further conditions that γ_0 must satisfy will be specified below, depending on the identification condition imposed on ε. Because of the aforementioned equation of q_0 with γ_0, the space of functions $\mathcal{Q} \ni q$ is essentially the same as Γ described here, and we simply rely on Γ as the relevant space.

To derive a semiparametric efficiency bound for this model, we follow the strategy of [22] — in order to consider the likelihood functions of one-dimensional submodels local to the true model, we organize local deviations from the model using real-valued $t \in [0, t_0]$ for some $t_0 > 0$. Let $\xi := (\theta', \beta')'$, and consider a curve $t \mapsto (\xi_t, g_t, \gamma_t, b_t)$ from $[0, t_0]$ into $\mathbb{R}^p \times \mathrm{L}^2(\mathbb{R}, \lambda) \times \Gamma \times \mathcal{B}$ that passes through $(\xi_0, g_0, \gamma_0, b_0)$ at $t = 0$. The score function for the model with respect to t (treating t as if it were the parameter to be estimated) is

$$S_0 = \frac{2\dot{q}(y|W)}{q_0(y|W)} + \frac{2\dot{b}(W)}{b_0(W)} \tag{2.1}$$

where $\dot{q} = \frac{\mathrm{d}}{\mathrm{d}t} q_t(y|W)|_{t=0}$ is tangent to q_t at $t = 0$ and all other "dotted" quantities are defined analogously. The Fisher information for the parameter ξ is

$$i_F = \mathrm{E}\left[S_0^2\right] = \mathrm{E}\left[\frac{4\dot{q}^2(y|W)}{q_0^2(y|W)} + \frac{4\dot{b}^2(W)}{b_0^2(W)}\right]$$