Two-step generalised empirical likelihood inference for semiparametric models

Francesco Bravo*
University of York

November 2007
(Preliminary version)
Comments welcome

Abstract

This paper shows how generalised empirical likelihood can be used to obtain valid asymptotic inference for the finite dimensional component of semiparametric models defined by a set of moment conditions. The results of the paper are illustrated using two well-known semiparametric regression models: the partially linear single index model and the linear transformation model with random censoring.

Key words and Phrases. Empirical likelihood, Local linear smoother, Linear tranformation model, Partially linear single-index model, Random censoring.

*Address correspondence to: Francesco Bravo, Department of Economics and Related Studies, University of York, York, Y010 5DD, United Kingdom. email: fb6@york.ac.uk
1 Introduction

The generalised empirical likelihood (GEL henceforth) method introduced by Newey & Smith (2004) provides a general framework to obtain estimators and test statistics for the parameters of statistical models defined by moment conditions models. Well-known special cases of GEL that have been focus of recent attention in both the econometrics and statistical literature are empirical likelihood (EL) (Owen 1988), (Qin & Lawless 1994) and (Owen 2001), exponential tilting (Efron 1981) and (Imbens, Spady & Johnson 1998), and continuous updating (CU) (known also as Euclidean likelihood (EU)) (Owen 1991) and (Hansen, Heaton & Yaron 1996).

In this paper we consider GEL in the context of semiparametric models. To be specific we show how GEL can be used for semiparametric models that can be defined in terms of a set of moment conditions. This set-up is quite general and can be applied to a number of semiparametric regression models widely used in applied research both in economics and biostatistic, including partially linear, single index and transformation models (see for example Horowitz (1998)). Our main interest is to obtain inferences for the finite dimensional parameters. To do so we propose a two-step procedure in which in the first step we use the plug-in principle and replace any unknown nuisance parameters with a consistent estimate. In the second step we maximise the resulting profiled GEL criterion function, and use the resulting (centred) maximised criterion function as GEL test statistic. We also consider a second test statistic which is similar to a robust Lagrange multiplier statistic and is based on a direct by-product of the maximisation process.

In this paper we show that the GEL test statistic converges typically to a non-standard distribution that can be expressed as a weighted sum of chi-squared distributions, whereas the robust Lagrange multiplier-type statistic converges to a standard chi-squared distribution. This difference in the asymptotic behaviour of the proposed test statistics, which is reminiscent of the difference between a likelihood ratio and a Lagrange multiplier test statistic in misspecified parametric likelihood theory (White 1982), can be explained considering the internal studentisation property of GEL. GEL automatically estimates a covariance of the (profiled) moment indicators that is typically different from that characterising the asymptotic normality of the profiled moment indicator itself. On the other hand, exactly as in the case of mis-specified parametric likelihood models, the Lagrange multiplier type statistic can be robustified so as to take into account the difference between these two covariance matrices.

In this paper we make two main contributions: First we show that GEL can be used to construct tests and confidence regions for a possibly a subset of the finite dimensional parameters vector. Second we provide Monte Carlo evidence about the
finite sample properties of the proposed test statistics, and compare them with those based on commonly used Wald statistics. These results, which extend and complement those recently obtained by Hjort, McKeague & Keilegom (2004), Lu & Liang (2006) and Xue & Zhu (2006) among others, are particularly important from an applied point of view because most of the hypotheses of interest in applied work involve nuisance parameters.

It is important to note that instead of solving the saddlepoint problem typically associated with GEL inferences for composite hypothesis we use the plug-in principle and replace the finite dimensional nuisance parameters with consistent estimates obtained by an appropriate subset of the moment conditions themselves. This procedure is appealing from a computational point of view because solving saddlepoint problems with semiparametric models typically involves solving systems of nonlinear estimating equations containing possibly nonparametric estimates, which are in general difficult to handle numerically and time-consuming to program. On the other hand the GEL test statistics proposed in this paper are the result of two separate (simpler) optimisation problems. The first one can often be carried out using standard numerical methods. The second one involves maximising a globally concave function over a convex domain. Thus from a computational point of view the two step GEL method of this paper compares favourably to the standard GEL approach based on a saddlepoint estimator.

The rest of the paper is organised as follows: in Section 2 we review briefly GEL, describe the two-step procedure and develop the necessary asymptotic theory. In Section 3 we illustrate the main results with two examples: the single index partial linear model, and the transformation model with known distribution of the error and random censoring. In Section 4 we present the results of Monte Carlo simulations. In Section 5 we conclude and suggest some directions for future research.

The following notation is used throughout the paper: “a.s” stands for almost surely, \( \xrightarrow{a.s} \), \( \xrightarrow{p} \), \( \xrightarrow{d} \) denote convergence almost surely, in probability and in distribution, respectively, and \( \| \cdot \| \) denotes the Euclidean norm. Finally \( ^{\top} \) denotes transpose, while \( ^{m} \) denotes derivative.

2 Main results

Let \( \{ z_i \}_{i=1}^{n} \) denote an i.i.d. sample from an unknown distribution \( P \) whose support is \( Z \subset \mathbb{R}^d \). We denote \( \Theta \) and \( H \) for, respectively, a finite and an infinite dimensional parameter set, and \( \theta_0 \in \Theta, h_0 \in H \) as the true unknown finite and infinite dimensional parameters. Suppose that there exists a measurable vector valued function \( m : \mathbb{R}^d \times \)
$\mathbb{R}^k \times H \rightarrow \mathbb{R}^k$ such that

$$E \left[ m \left( z_i, \theta, h_0 \right) \right] = 0 \text{ if } \theta = \theta_0.$$ \hfill (1)

Let $m_i (\theta, h) = m (z_i, \theta, h) , \theta = [\theta_1^\tau, \theta_2^\tau]^\tau$ where $\dim (\theta_j) = \mathbb{R}^{k_j}, \sum_j k_j = k \ (j = 1, 2)$; assume that $\Theta = \Theta_1 \times \Theta_2$, and suppose that we are interested to test the composite hypothesis $H_0 : \theta_1 = \theta_{10}$. If $h_0$ were known the standard GEL approach to test such hypothesis would be to compute the following test statistic

$$D_\rho = 2 \left( \tilde{P}_\rho \left( \theta_{10}, \hat{\theta}_2, h_0, \lambda \right) - \rho \left( 0 \right) \right) \hfill (2)$$

where $\hat{\theta}_2$ is a saddlepoint estimator defined as

$$\arg \min_{\hat{\theta}_2 \in \Theta_2} \tilde{P}_\rho \left( \theta_{10}, \hat{\theta}_2, h_0, \lambda \right)$$

where

$$\hat{\lambda} := \arg \max_{\lambda \in \hat{\Lambda}_n (\theta_{10}, h_0)} \tilde{P}_\rho \left( \theta_{10}, \hat{\theta}_2, h_0, \lambda \right),$$

$$\hat{\Lambda}_n (\theta_{10}, h_0) = \{ \lambda | \lambda^\tau m_i (\theta_{10}, \hat{\theta}_2, h_0) \in V \}, \ V \text{ is an open interval of the real line,}$$

$$\tilde{P}_\rho \left( \theta, h, \lambda \right) = \sum_{i=1}^n \rho \left( \lambda^\tau m_i (\theta, h) \right) / n, \hfill (3)$$

and $\rho : V \rightarrow \mathbb{R}$ satisfies certain regularity properties described in Assumption $\rho$ below. Examples of $\rho (v)$ are given in Section 5 below.

The test statistic $D_\rho$ is based on the difference in the GEL criterion function between the constrained estimator $\tilde{\theta} = [\theta_{10}^\tau, \hat{\theta}_2^\tau]^\tau$ and the unconstrained Z-estimator $\hat{\theta}$ that solves $\sum_{i=1}^n m_i \left( \hat{\theta}, h_0 \right) / n = 0$, which, because the model considered is exactly identified, results in $\tilde{P}_\rho \left( \hat{\theta}, h_0, \hat{\lambda} \right) = \rho \left( 0 \right)$. Note also that in the case of EL (and more generally for the Cressie-Read discrepancy (Baggerly 1998)) $D_\rho$ has an interesting interpretation as twice the logarithm of a nonparametric likelihood ratio (twice a nonparametric likelihood discrepancy) statistic, with the estimated auxiliary parameter $\hat{\lambda}$ as a Lagrange multiplier which ensures that the moment conditions (1) are satisfied in the sample. Under mild regularity conditions it can be shown that $D_\rho \xrightarrow{d} \chi^2_{k_1}$ (Guggenberger & Smith 2005).

Suppose now that $h$ is unknown, and we are still interested to obtain inferences about $\theta_1$. To construct an analogue of (2) we can estimate $h$ and $\theta_2$ either simultaneously or sequentially. In either cases the resulting test statistic requires the computation of the saddlepoint

$$\min_{\hat{\theta}_2 \in \Theta_2, h \in H} \max_{\lambda \in \hat{\Lambda}_n (\theta_{10})} \tilde{P}_\rho \left( \theta_{10}, \hat{\theta}_2, h, \lambda \right)$$

and $\lambda : V \rightarrow \mathbb{R}$ satisfies certain regularity properties described in Assumption $\lambda$ below. Examples of $\lambda (v)$ are given in Section 5 below.
which can be numerically difficult to solve and potentially unstable. To avoid these problems we propose a simple two-step procedure, in which the moment indicator \( (1) \) is partitioned as

\[
    m_i (\theta, h) = [m_{1i} (\theta_1, \theta_2, h)^\top , m_{2i} (\theta_1, \theta_2, h)]^\top
\]

where \( m_j : \mathbb{R}^d \times \mathbb{R}^{k_j} \times H \to \mathbb{R}^{k_j} \) \( (j = 1, 2) \). Let \( \hat{h} \) denote an estimator for \( h \) whose precise form depends on the structure of the problem under investigation (see Section 3 below). In the first step we use \( m_{2i} (\theta_1, \theta_2, \hat{h}) \) to estimate \( \theta_2 \). To be specific for a fixed \( \theta_{10} \) the estimator \( \hat{\theta}_2 \) solves

\[
    \sum_{i=1}^n m_{2i} (\theta_{10}, \hat{\theta}_2, \hat{h}) / n = 0.
\]

In the second step we use the profiled first moment condition \( m_{1i} (\theta_1, \hat{\theta}_2, \hat{h}) \) to compute the GEL statistic for \( H_0 : \theta_1 = \theta_{10} \).

Let \( \bar{\Omega}_1 (\theta_1, \hat{\theta}_2, \hat{h}) = \sum_{i=1}^n m_{1i} (\theta_1, \theta_2, h) m_{1i} (\theta_1, \theta_2, h) / n; \) assume that

\( (\rho) \rho \) is concave on \( V \), twice continuously differentiable in a neighbourhood of 0, and \( \rho_j = -1 \) \( (j = 1, 2) \) where \( \rho_j (v) = d^j \rho (v) / dv^j \) and \( \rho_j := \rho_j (0) \).

\( (C) \left\| \hat{\theta}_2 - \theta_{20} \right\| = o_p (1), \left\| \hat{h} - h_{10} \right\| = o_p (1) \)

\( (M) \max_i \left\| m_{1i} (\theta_{10}, \hat{\theta}_2, \hat{h}) \right\| = o_p \left( n^{1/2} \right) \)

\( (\Omega) \left\| \bar{\Omega}_1 (\theta_{10}, \hat{\theta}_2, \hat{h}) - \Omega_{10} \right\| = o_p (1) \) for some positive definite matrix \( \Omega_{10} \)

\( (N) \sum_{i=1}^n m_{1i} (\theta_{10}, \hat{\theta}_2, \hat{h}) / n^{1/2} \overset{d}{\longrightarrow} N (0, \Xi_{10}) \).

**Theorem 1** Assume that \( \rho, C, M, \Omega \) and \( N \) hold. Then under \( H_0 : \theta_1 = \theta_{10} \)

\[
    D_\rho = 2 \left( \tilde{P}_\rho (\hat{\theta}_{10}, \hat{\theta}_2, \hat{h}, \hat{\lambda}) - \rho (0) \right) \overset{d}{\longrightarrow} \sum_{j=1}^{k_1} \omega_j \chi^2_{1,j}, \quad (4)
\]

\[
    LM_\rho = n \hat{\lambda}^\top \Omega_{10}^{-1} \Xi_{10} \hat{\lambda} \overset{d}{\longrightarrow} \chi^2_{k_1}
\]

where \( \chi^2_{1,j} \) are independent chi-squared random variables with one degree of freedom and the weights \( \omega_j \) are the eigenvalues of \( \Omega_{10}^{-1} \Xi_{10} \).
Remark 1. Theorem 1 implicitly assumes the existence of the maximiser $\hat{\lambda}$. This follows (with probability approaching 1) as long as $\lambda^T m_{1i} (\theta_{10}, \theta_2, h)$ is in the domain $\mathcal{V}$ of the function $\rho$ for all $\lambda \in \Lambda_n$, $\theta_2 \in \Theta_2$, $h \in H$ and $1 \leq i \leq n$. Given assumption $M$ it suffices for the theory here that $\Lambda_n$ places a bound on $\lambda$ that shrink with $n$ slower than $n^{-1/2}$ (see the proof of Theorem 1 for an example of $\Lambda_n$).

Remark 2. To compute both statistics we need to obtain consistent estimates of $\Omega_{10}$ and $\Xi_{10}$. Moreover to avoid resorting to simulations to compute $D_\rho$ we can use two adjusted statistics. The first one is based on results of Rao & Scott (1981) and is given by
\begin{equation}
D_{\rho_1} = \left(p/\text{trace}\left(\Omega_{10}^{-1}\Xi_{10}\right)\right) D_\rho \overset{d}{\to} \chi^2_{k_1}.
\end{equation}

The second one is as in Xue & Zhu (2006) and is given by
\begin{equation}
D_{\rho_2} = \left[\left(\text{trace}\left(\Xi_{10}^{-1} M_{10}\right)\right)/\text{trace}\left(\Omega_{10}^{-1} M_{10}\right)\right] D_\rho \overset{d}{\to} \chi^2_{k_1}
\end{equation}
where $M_{10} = \sum_{i=1}^n m_{1i} (\hat{\theta}_{10}, \hat{\theta}_2, h) \sum_{i=1}^n m_{1i} (\hat{\theta}_{10}, \hat{\theta}_2, h)^\top$.

Remark 3. Note that in case of simple hypotheses the conclusion of the theorem are still valid with $k$ replacing $k_1$, and $\Omega_{10}$, $\Xi_{10}$ replaced by their full parameters analogues $\Omega_0 = E \left[ m_i \left( \theta_{10}, \theta_{20}, \hat{h} \right) m_i \left( \theta_{10}, \theta_{20}, \hat{h} \right)^\top \right]$, $\Xi_0 = \text{COV} \left( m_i \left( \theta_{10}, \theta_{20}, \hat{h} \right) \right)$.

Remark 4. In case $\Omega_{10}^{-1} \Xi_{10} = I$ under the same assumptions of Theorem 1 it is easy to see that the distance statistic $D_\rho$ converges in distribution to a standard $\chi^2_{k_1}$.

3 Examples

We begin this section with a brief discussion about possible estimators $\hat{h}$ of $h$. Given that $h$ is an unknown function the typical estimator $\hat{h}$ will be based on nonparametric estimation methods such as kernels, local polynomials, splines etc. However other choices are available. For example when $h$ is the unknown distribution of a random censoring variable $\hat{h}$ can be based on the Kaplan-Meier estimator Kaplan & Meier (1958). In certain circumstances martingale methods can be used so that $\hat{h}$ is the solution of an appropriate set of estimating equations. The two examples considered in this section illustrate this point since estimation of $h$ is based on local polynomials and on martingale methods, respectively.

3.1 Partially linear single index model

Let $\{z_i\}_{i=1}^n = \{y_i, x_i\}_{i=1}^n$ denote an i.i.d. sample from an unknown distribution $F$ with support is $Z = Y \times X \subset \mathbb{R} \times \mathbb{R}^{d_x}$. The partially linear single index model is
\begin{equation}
y_i = g_0 \left( x_{1i}^T \theta_{10} \right) + x_{2i}^T \theta_{20} + \varepsilon_i
\end{equation}
where \( g_0 : \mathbb{R} \rightarrow \mathbb{R} \) is an unknown function, \( \{x_i\}_{i=1}^n \) are unobservable i.i.d. random errors with \( E(x_i | x_i) = 0 \) a.s. and \( \|\theta_{10}\| = 1 \) for identifiability. Model (7) covers two important cases: the single index model with \( \theta_{20} = 0 \), and the partially linear model with \( \theta_{10} = 1 \).

Note that because of the identifiability restriction on \( \theta_{10} \) \( g \) does not have a derivative at \( \theta_{10} \). Thus as in Xue & Zhu (2006) we can use the so-called delete-one-component and write

\[
\theta_{10} = \left[ \theta_{110}, \theta_{120}, \ldots, \theta_{1(p-1)0}, 1 - \left( \|\theta_{00}\| \right)^{1/2}, \theta_{1(p+1)0}, \ldots, \theta_{1k_10} \right]',
\]

and since \( \|\theta_{1(j)}\| < 1 \) by the implicit function theorem \( \theta \) is differentiable in a neighbourhood of \( \theta_{1(0)} \) with Jacobian matrix \( \partial\theta_1/\partial\theta_1' = J_{\theta(j)} = [\gamma_1, \ldots, \gamma_{k_1}] \) and

\[
\gamma_l = - \left( 1 - \left( \|\theta_{1(j)}\| \right)^{1/2} \right) \left[ \theta_{110}, \theta_{120}, \ldots, \theta_{1(l-1)0}, - \left( 1 - \left( \|\theta_{00}\| \right)^{1/2} \right), \theta_{1(l+1)0}, \ldots, \theta_{1k_10} \right].
\]

The moment indicator is

\[
m_i (\theta_1, \theta_2, h) = [g' (x_{i1}^T \theta_1) x_{i1}^T J_{\theta(j)}, x_{i2}^T] (y_i - g (x_{i1}^T \theta_1) - x_{i2}^T \theta_2)
\]

where \( g' (x_{i1}^T \theta_1) = \partial g (x_{i1}^T \theta_1) / \partial \theta(j) \) and \( h = [g, g'] \). To obtain an estimator \( h \) of \( h \) we use the local linear smoother Fan & Gijbels (1996), which has the advantage over the Nadaraya-Watson kernel estimator of estimating \( g \) and \( g' \) simultaneously. Let \( \hat{\alpha}, \hat{\beta} \) solve the local (weighted) least squares problem

\[
\min_{\alpha, \beta} \sum_{i=1}^{n} (y_i - x_{i2}^T \theta_2 - \alpha - \beta (x_{i1}^T \theta_1 - t))^2 K_h (x_{i1}^T \theta_1 - t)
\]

where \( K_h (\cdot) = K_h (\cdot/h) / h, K (\cdot) \) is a kernel function with bandwidth \( h = h (n) \). Simple calculations show that

\[
\hat{\alpha} = \sum_{i=1}^{n} u_i (\theta_1, t) (y_i - x_{i2}^T \theta_2) / \sum_{j=1}^{n} u_j (\theta_1, t)
\]

\[
\hat{\beta} = \sum_{i=1}^{n} u_i (\theta_1, t) (y_i - x_{i2}^T \theta_2) / \sum_{j=1}^{n} u_j (\theta_1, t)
\]

where

\[
u_i (\theta_1, t) = K_h (x_{i1}^T \theta_1 - t) [s_{2n} (\theta_1, t) - (x_{i1}^T \theta_1 - t) s_{1n} (\theta_1, t)]
\]

\[
u_i (\theta_1, t) = K_h (x_{i1}^T \theta_1 - t) [(x_{i1}^T \theta_1 - t) s_{0n} (\theta_1, t) - s_{1n} (\theta_1, t)]
\]

\[
s_{jn} (\theta_1, t) = \sum_{i=1}^{n} (x_{i1}^T \theta_1 - t)^j K_h (x_{i1}^T \theta_1 - t) / n \quad j = 0, 1, 2.
\]
Then the estimators $\hat{g}(\theta, t)$ and $\hat{g}'(\theta, t)$ are defined as
\[
\hat{g}(\theta, t) = \sum_{i=1}^{n} w_i(\theta_1, t) (y_i - x_{2i}^T \theta_2),
\]
\[
\hat{g}'(\theta, t) = \sum_{i=1}^{n} \tilde{w}_i(\theta_1, t) (y_i - x_{2i}^T \theta_2)
\]
where
\[
w_i(\theta_1, t) = u_i(\theta_1, t) / \sum_{j=1}^{n} u_j(\theta_1, t),
\]
\[
\tilde{w}_i(\theta_1, t) = \tilde{u}_i(\theta_1, t) / \sum_{j=1}^{n} \tilde{u}_j(\theta_1, t).
\]

Suppose that we are interested to test $H_0 : \theta_1 = \theta_{10}$. Partition
\[
m_i \left( \theta_{10}, \theta_2, \hat{h} \right) = \begin{bmatrix} m_{1i} \left( \theta_{10}, \theta_2, \hat{h} \right) \end{bmatrix}, \quad m_{2i} \left( \theta_{10}, \theta_2, \hat{h} \right)
\]
; an estimator $\hat{\theta}_2$ for the nuisance parameter $\theta_2$ can be defined as
\[
\sum_{i=1}^{n} m_{2i} \left( \theta_{10}, \hat{\theta}_2, \hat{h} \right) / n = 0
\]
which admits a simple closed form solution. Then the profiled moment indicator to be used in the GEL criterion function is
\[
m_{1i} \left( \theta_{10}, \hat{\theta}_2, \hat{h} \right) = \left[ \hat{g}'(x_{1i}^T \theta_{10}) x_{1i}^T \Omega_{ij} \right] (y_i - \hat{g}(x_{1i}^T \theta_{10}) - x_{2i}^T \hat{\theta}_2)
\]
(8)

Let
\[
\Omega_{12} (\theta) = E \left[ (x_{1i}^T - E (x_{1i}^T | x_{1i}^T \theta_{10})) (x_{2i} - E (x_{2i}^T | x_{1i}^T \theta_{10}))^T \right],
\]
\[
\Omega_{22} (\theta) = E \left[ x_{2i} - E (x_{2i}^T | x_{1i}^T \theta_{10}) x_{2i} - E (x_{2i}^T | x_{1i}^T \theta_{10})^T \right],
\]
\[
q_{ii} (\theta, t) =\hat{g}'(x_{1i}^T \theta_{10}) x_{1i} - \Omega_{12} (\theta) \Omega_{22} (\theta)^{-1} x_{2i}
\]
\[
\Omega_{1i} (\theta, t) = g' \hat{g}(x_{1i}^T \theta_{10}) - \Omega_{12} (\theta) \Omega_{22} (\theta)^{-1} (x_{2i} - E (x_{2i} | x_{1i}^T \theta_{10}))
\]

Assume as in Xue & Zhu (2006)

A1 the density function $f(t)$ of $x_{1i}^T \theta_1$ is bounded away from 0 and satisfies a Lipschitz condition of order 1 on $T_x$ where $T_x = \{ x | t = x_{1i}^T \theta_1, x \in X_1 \}$ and $X_1$ is the bounded support of $x_1$,

A2 $g_0(t)$ has two continuous derivatives on $T_x$ and $E (x_1^T x_{1i}^T \theta_1 = t)$ satisfies a Lipschitz condition of order 1,
A3  the kernel $K(u)$ is a bounded probability density function satisfying $\int_{-\infty}^{\infty} uK(u) \, du = 0$, $\int_{-\infty}^{\infty} u^2K(u) \, du \neq 0$, $\int_{-\infty}^{\infty} u^8K(u) \, du < \infty$,

A4  $E(\varepsilon_i | x_i) = 0$ a.s., $\sup_x E(\varepsilon_i^4 | x_i = x) < \infty$, $\sup_{t \in T_x} E(\|x_{2i}\|^2 | x_{1i}, \theta_1 = t) < \infty$,

A5  $nh^2 \to \infty$, $nh^4 \to 0$, $nh h_1^3 \to \infty$, $\lim \sup_{n \to \infty} nh h_1^5 < \infty$,

A6  $\Omega_{10} = E[\varepsilon_i^2 q_{1i}(\theta_0, t) q_{1i}(\theta_0, t)^\top]$, and $\Xi_{10} = E[\varepsilon_i^2 \tilde{q}_{1i}(\theta_0, t) \tilde{q}_{1i}(\theta_0, t)^\top]$ are positive definite.

Condition A5 introduces another bandwidth $h_1 = h_1(n)$ to control for the variability of $g'$. The convergence rate of the estimator of $g'$ is slower than that of $g$ and thus using the same bandwidth would slower the $n^{1/2}$ rate for $\tilde{\theta}_1$ unless a third order kernel, undersmoothing and the more stringent condition $nh^6 \to 0$ are used.

**Proposition 2** Assume that $\rho$ and A1 – A6 hold. Then under $H_0 : \theta_1 = \theta_{10}$ the conclusions of Theorem 1 are valid for the profile moment indicator (8).

**Remark 5.** It should be noted that because of the restriction $\|\theta_1\| = 1$ the actual dimension of $\theta_1$ is $k_1 - 1$. Therefore as long as the $j$th component of $\theta_1$ is positive, Proposition 2 can be reformulated in terms of $\theta_{10}^{(j)}$ to produce an asymptotic $\chi^2$ approximation with $k_1 - 1$ degrees of freedom. Such approximation, which is used in Xue & Zhu (2006), can improve the finite sample accuracy of the GEL statistic.

Consistent estimators for the matrices $\Omega_{10}$ and $\Xi_{10}$ are

$$
\hat{\Omega}_1 = \sum_{i=1}^{n} \hat{q}_{1i}(\hat{\theta}) \hat{q}_{1i}(\hat{\theta})^\top \varepsilon_i^2 / n
$$

$$
\hat{\Xi}_1 = \sum_{i=1}^{n} \hat{q}_{1i}(\hat{\theta}) \hat{q}_{1i}(\hat{\theta})^\top \varepsilon_i^2 / n
$$

where

$$
\hat{q}_{1i}(\hat{\theta}) = g' J_{\theta_{10}}(x_{1i} - \hat{E}(x_{1i}|x_{1i}, \hat{\theta}_1)) - \hat{\Omega}_{12} \hat{\theta}_1^{-1} (x_{2i} - \hat{E}(x_{2i}|x_{1i}, \hat{\theta}_1))
$$

$$
\hat{E}(x_{1i}|x_{1i}, \hat{\theta}_1) = \sum_{i=1}^{n} w_i(\hat{\theta}_1, x_{1i}, \hat{\theta}_1) x_{1i}, \quad \hat{E}(x_{2i}|x_{1i}, \hat{\theta}_1) = \sum_{i=1}^{n} w_i(\hat{\theta}_1, x_{1i}, \hat{\theta}_1) x_{2i},
$$

$$
\hat{\Omega}_{j2}(\theta) = \sum_{i=1}^{n} \left[ (x_{ji} - \hat{E}(x_{ji}|x_{1i}, \hat{\theta}_1)) (x_{2i} - \hat{E}(x_{2i}|x_{1i}, \hat{\theta}_1))^\top \right] / n \ (j, k = 1, 2),
$$

$$
\hat{\varepsilon}_i^2 = (y_i - \hat{g}(x_{1i}, \hat{\theta}_1) - x_{2i}, \hat{\theta}_2)^2
$$
and $\hat{\theta}$ is a consistent estimator of $\theta_0$. For example we can use the method recently proposed by Xia & Hardle (2006) to compute

$$\hat{\theta} = \arg \min_{\theta_1, \theta_2} \sum_{i=1}^{n} (y_i - \hat{g}(x_{1i}^\tau \theta_1) - x_{2i}^\tau \theta_2)^2 \text{ s.t. } \|\theta_1\| = 1.$$

### 3.2 Linear transformation models with random censoring

Let $\{z_i\}_{i=1}^{n} = \{y_i^*, x_i\}_{i=1}^{n}$ denote an i.i.d. sample from an unknown distribution $F$ with support $Z = Y \times X \subset \mathbb{R}^+ \times \mathbb{R}^{d_x}$. With right random censoring the available sample is $\{y_i, x_i, \delta_i\}_{i=1}^{n}$ where $y_i = \min(y_i^*, c_i)$, $\delta_i = I\{y_i^* < c_i\}$ and let $\{c_i\}_{i=1}^{n}$ is an i.i.d. sample from an unknown distribution $G$ with support $C \subset \mathbb{R}^+$, assumed independent from $F$. The linear transformation models is

$$g_0(y_i) = -x_i^\tau \theta_0 + \varepsilon_i \quad (9)$$

where $g_0(\cdot)$ is an unknown monotone increasing function and $\{\varepsilon_i\}_{i=1}^{n}$ are i.i.d. random errors from a known distribution $\phi_\varepsilon$. For example if $\phi_\varepsilon$ is the extreme value distribution (9) becomes the well-known proportional hazard model of Cox (1972). Note that since $g_0$ is unknown the parametric assumption on $\varepsilon$ should not be viewed as restrictive.

Let $\lambda_\varepsilon(\cdot)$ and $\Lambda_\varepsilon(\cdot)$ denote the hazard and cumulative hazard function, respectively, and let $N_i(t) = \delta_i I\{y_i \leq t\}$ and $Y_i(t) = I\{y_i \geq t\}$ denote the counting process and the at-risk process, respectively. The moment indicator is the (counting process) martingale integral

$$m_i(\theta, h) = \int_0^\infty x_i [dN_i(t) - Y_i(t) d\Lambda_\varepsilon(g(t) + x_i^\tau \theta)],$$

where and $h = g$ (Lu & Liang 2006). Partition $\theta = [\theta_1^\tau, \theta_2^\tau]^\tau$ and suppose that we are interested to test $H_0 : \theta_1 = \theta_{10}$. Estimators for $g$ and $\theta_2$ can be defined as in Chen, Jin & Jing (2002) as the solution to

$$\sum_{i=1}^{n} (dN_i(t) - Y_i(t) d\Lambda_\varepsilon(g(t) + x_{1i}^\tau \theta_{10} + x_{2i}^\tau \theta_2)) = 0$$

$$\sum_{i=1}^{n} x_{2i} [dN_i(t) - Y_i(t) d\Lambda_\varepsilon(g(t) + x_{1i}^\tau \theta_{10} + x_{2i}^\tau \hat{\theta}_2)] = 0,$$

where $g$ is an nonincreasing function satisfying $g(0) = -\infty$. This requirement ensures that $\Lambda_\varepsilon(C + g(0)) = 0$ for any finite $C$. 

10
The profiled moment indicator to be used in the GEL criterion function is

\[ m_{1i} (\theta_{10}, \hat{\theta}_2, \hat{h}) = \int_0^\infty x_{1i} \left[ dN_i(t) - Y_i(t) d\Lambda_\varepsilon \left( \hat{g}(t) + x_{1i}^T \theta_{10} + x_{2i}^T \hat{\theta}_2 \right) \right]. \quad (10) \]

Let \( \tau = \inf (t : \Pr (y_i > t) = 0) \), and for any \( s, t \in (0, \tau) \) let

\[ B(t, s) = \exp \left\{ \int_s^t \frac{E [\lambda'_\varepsilon (g_0(u) + x_i^T \theta) Y_i(u)]}{E [\lambda'_\varepsilon (g_0(u) + x_i^T \theta) Y_i(u)]} dg_0(u) \right\} \]

\[ \mu_j(t) = \frac{E [x_{ji} \lambda_\varepsilon (g_0(y_i) + x_i^T \theta) Y_i(t) B(t, y_i)]}{E [\lambda_\varepsilon (g_0(y_i) + x_i^T \theta) Y_i(t) B(t, y_i)]} \quad j = 1, 2. \]

Moreover let

\[ \Omega_{12}(\theta) = \int_0^\tau \frac{E [(x_{1i} - \mu_1(t)) x_{2i}^T \lambda'_\varepsilon (g_0(t) + x_i^T \theta) Y_i(t)]}{d\lambda_\varepsilon(t)} \]

\[ \Omega_{22}(\theta) = \int_0^\tau \frac{E [(x_{2i} - \mu_2(t)) x_{2i}^T \lambda'_\varepsilon (g_0(t) + x_i^T \theta) Y_i(t)]}{d\lambda_\varepsilon(t)} \]

\[ q_i(\theta, t) = x_{1i} - \mu_1(t) - \Omega_{12}(\theta) \Omega_{22}(\theta)^{-1} (x_{2i} - \mu_2(t)) \]

Let \( \psi_\varepsilon(t) = \partial \log \lambda_\varepsilon(t) / \partial t = \lambda'(t) / \lambda(t) \); assume that

B1 \( \lambda_\varepsilon(\cdot) > 0 \), \( \psi_\varepsilon(\cdot) \) is a continuous function, \( \lim_{s \to -\infty} \lambda_\varepsilon(s) = 0 = \lim_{s \to -\infty} \psi_\varepsilon(s) \)

B2 \( \tau \) is finite, \( \Pr (y_i > \tau) > 0 \), \( \Pr (c_i = \tau) > 0 \)

B3 \( x_i \) is compactly supported, that is \( \Pr (\|x_i\| > X_0) = 0 \) for some \( X_0 > 0 \)

B4 \( g_0 \) has continuous and positive derivatives

B5 \( \Omega_{10} = \int_0^\tau \frac{E [x_{1i} x_{1i}^T \lambda_\varepsilon (g_0(t_i) + x_i^T \theta_0) Y(t)]}{d\lambda_\varepsilon(t)} \), \( \Xi_{10} = \int_0^\tau \frac{E [q_i(\theta_0, t) q_i(\theta_0, t)^T \lambda_\varepsilon (g_0(t) + x_i^T \theta_0) Y(t)]}{d\lambda_\varepsilon(t)} \) are positive definite

Proposition 3 Assume that \( \rho \) and B1 – B5 hold. Then under \( H_0 : \theta = \theta_{10} \) the conclusions of Theorem 1 are valid for the profile moment indicator (10)

Consistent estimators for the matrices \( \Omega_{10} \) and \( \Xi_{10} \) are

\[ \hat{\Omega}_{10} = \int_0^\tau \sum_{i=1}^n \left[ x_{1i} x_{1i}^T \lambda_\varepsilon \left( \hat{g}(t) + x_{1i}^T \hat{\theta} \right) Y_i(t) \right] d\hat{g}(t) / n \]

\[ \hat{\Xi}_{10} = \int_0^\tau \sum_{i=1}^n \left[ \hat{q}_i(\hat{\theta}, t) \hat{q}_i(\hat{\theta}, t)^T \lambda_\varepsilon \left( \hat{g}(t) + x_{1i}^T \hat{\theta} \right) Y_i(t) \right] d\hat{g}(t) / n, \]
where

\[ \hat{q}_i(\hat{\theta}, t) = x_{1i} - \hat{\mu}_1(t) - \hat{\Omega}_{12}(\hat{\theta}) \hat{\Omega}_{22}(\hat{\theta})^{-1}(x_{2i} - \hat{\mu}_2(t)), \]

\[ \hat{\mu}_j(t) = \sum_{i=1}^{n} \frac{[x_{ji}\lambda_{\varepsilon}(\hat{g}(y_i) + x_i^\top\hat{\theta})Y_i(t)\hat{B}(t, y_i)]}{\sum_{i=1}^{n} \lambda_{\varepsilon}(\hat{g}(y_i) + x_i^\top\hat{\theta})Y_i(t)\hat{B}(t, y_i)} \quad j = 1, 2 \]

\[ \hat{B}(t, s) = \exp \left\{ \int_{t}^{s} \frac{\sum_{i=1}^{n} \lambda_{\varepsilon}'(\hat{g}(u) + x_i^\top\hat{\theta})Y_i(u)}{\sum_{i=1}^{n} \lambda_{\varepsilon}(\hat{g}(u) + x_i^\top\hat{\theta})Y_i(u)} d\hat{g}(u) \right\}, \]

\[ \hat{\Omega}_{j2}(\hat{\theta}) = \int_{0}^{t} \sum_{i=1}^{n} \left[ (x_{ji} - \hat{\mu}_j(t)) x_{2i}^\top \lambda_{\varepsilon}'(\hat{g}(t) + x_i^\top\hat{\theta})Y_i(t) \right] d\hat{g}(t) / n \]

and \( \hat{\theta} \) is the consistent estimator for \( \theta_0 \) (Chen et al. 2002) defined as the solution to

\[ \sum_{i=1}^{n} \left( dN_i(t) - Y_i(t) dA_{\varepsilon}(\hat{g}(t) + x_i^\top\hat{\theta}) \right) = 0. \]

### 4 Monte Carlo results

In this section we use simulations to assess the finite sample properties of GEL based statistics (4) for the two examples discussed in the previous section. In the simulations we consider the three GEL statistics that are most used in practice, namely the Empirical likelihood (EL), Euclidean distance (EU), the exponential tilting (ET). These are given respectively by

\[ D_{EL} = 2 \sum_{i=1}^{n} \log \left( 1 - \hat{\lambda}^\top m_{1i}(\theta_{10}, \hat{\theta}_2, \hat{h}) \right) \]

\[ D_{EU} = \sum_{i=1}^{n} \left( 1 + \hat{\lambda}^\top m_{1i}(\theta_{10}, \hat{\theta}_2, \hat{h}) \right)^2 / 2 \]

\[ D_{ET} = 2 \sum_{i=1}^{n} \left( 1 - \exp \left( \hat{\lambda}^\top m_{1i}(\theta_{10}, \hat{\theta}_2, \hat{h}) \right) \right) \]

In general to compute \( \hat{\lambda} \) one can apply the multivariate Newton’s algorithm to \( \sum_{i=1}^{n} \rho \left( \lambda^\top m_{1i}(\theta_{10}, \hat{\theta}_2, \hat{h}) \right) \); this amounts to Newton’s method for solving the nonlinear system of \( q \) first-order conditions \( \sum_{i=1}^{n} \rho_1 \left( \lambda^\top m_{1i}(\theta_{10}, \hat{\theta}_2, \hat{h}) \right) m_{1i}(\theta_{10}, \hat{\theta}_2, \hat{h}) = 0 \) with starting point in the iterative process set to \( \lambda_0 = 0^\top \). For such choice of starting point, the convergence of the algorithm is typically quadratic. Note also that
the case of EU there is no need to use any numerical optimisation method to find the maximiser $\hat{\lambda}$ since the latter can be obtained in closed form and is given by $\hat{\lambda} = \hat{\Omega}_1 \left( \theta_1, \hat{\theta}_2, \hat{h} \right)^{-1} \sum_{i=1}^{n} m_{ii} \left( \theta_1, \hat{\theta}_2, \hat{h} \right)$. In the simulations we consider a number of statistics: the distance $D_\rho$ and Lagrange multiplier $LM_\rho$ given in (4), the adjusted distance statistic $D_\rho^{a2}$ (6), and a Wald statistic. The results are based on 5000 replications, while the critical values of the statistic $D_\rho$ are based on 50000 replications.

4.1 Single index model

We consider as in Xue & Zhu (2006) the single index model

$$y_i = \left( x_i^T \theta_0 - 1/3^{1/2} \right)^2 + 1 + \varepsilon_i$$

where $x_i$ is a trivariate vector with independent uniform $[0,1]$ components, $\theta_0 = [2^{1/2}/2, 3^{1/2}/3, 6^{1/2}/6]^T$, and $\varepsilon_i \sim N (0, 0.04)$.

The hypothesis of interest is $H_0 : \theta = \theta_0$ so that there are no nuisance parameters. As in Xue & Zhu (2006) we use as the kernel function $K(t) = 15 (1 - t^2) I \{|t| \leq 1\}/16$, and select the bandwidths $h, h_1$ as

$$h = \hat{h}_{opt} n^{-2/15}, \quad h_1 = \hat{h}_{opt}$$

where $\hat{h}_{opt}$ is chosen by least squares cross-validation. The simulations results are presented in Table 1.

Table 1 approx. here

4.2 Partially linear model

We consider the partial linear model

$$y_i = \exp \left( x_{1i} \right) + x_{2i}^T \theta_0 + \varepsilon_i$$

where $x_{1i}$ is uniform $[0,1]$, $x_i$ is bivariate vector with standard normal components, $\theta_0 = [0.5, -2]$ and $\varepsilon_i \sim N (0, 2)$. We use the same kernel and the same selection method for the bandwidths $h, h_1$ as those used in the previous example. The simulations results for the hypothesis of interest is $H_0 : \theta = \theta_0$ are presented in Table 2.

Table 2 approx. here
4.3 Linear transformation

We consider as Chen et al. (2002) and Lu & Liang (2006) the transformation model

$$\log (y_i) = x_i^\tau \theta_0 + \varepsilon_i$$

where $x_i^\tau = [x_{1i}, x_{2i}]$, $x_{1i}$ is a Bernoulli random variable with probability 0.5, $x_{2i}$ is uniform $[0, 1]$, and $\theta_0 = [0, 0]$.

The hazard function for the error term $\varepsilon$ is specified as

$$\lambda_{\varepsilon} (t) = \exp (t) / (1 + \gamma \exp (t))$$

for $\gamma = 0, 1, 2$, which includes the proportional hazard model for $\gamma = 0$ and the proportional odds model for $\gamma = 1$. The censoring times, assumed to be independent of $x_i$, are uniform $[0, c]$ where $c$ takes different values according to the expected proportion of censoring.

Note that to compute $\hat{g}$ and $\hat{\theta}_2$ we use the same iterative algorithm suggested by Chen et al. (2002). To be specific for observed (failure) times $t_1, ..., t_k$ and fixed initial value $\theta_2^{(0)}$ we first obtain $\hat{g}^{(0)} (t_1)$ as the solution of

$$\sum_{i=1}^{n} Y_i (t_1) \Lambda_\varepsilon \left( \hat{g}^{(0)} (t_1) + x_{1i}^\tau \theta_{10} + x_{2i}^\tau \theta_{2}^{(0)} \right) = 1. \quad (11)$$

Then obtain $\hat{g}^{(0)} (t_j)$, $j = 2, ..., k$ recursively using

$$\hat{g}^{(0)} (t_j) = \hat{g}^{(0)} (t_{j-}) + 1 / \sum_{i=1}^{n} Y_i (t_j) \Lambda_\varepsilon \left( \hat{g}^{(0)} (t_{j-}) + x_{1i}^\tau \theta_{10} + x_{2i}^\tau \theta_{2}^{(0)} \right). \quad (12)$$

Next obtain $\hat{\theta}_2^{(1)}$ as the solution of

$$\sum_{i=1}^{n} \int_{0}^{\tau} x_{2i} \left[ dN_i (t) - Y_i (t) d\Lambda_\varepsilon \left( \hat{g}^{(0)} (t) + x_{1i}^\tau \theta_{10} + x_{2i}^\tau \hat{\theta}_2^{(1)} \right) \right] = 0, \quad (13)$$

and repeat $(11) - (13)$ until convergence.

The simulations results for the hypothesis of interest is $H_0 : \theta = \theta_0$ are presented in Table 3.

Table 3 approx. here

The results of Tables 1-3 can be summarised as follows: First all three test statistics based on the GEL approach have good size properties, typically better than those based on a standard Wald statistic, with the Empirical likelihood and exponential tilting having an edge over the Euclidean likelihood statistic. Second the adjusted distance statistics $D_{\rho}^{\alpha_2}$ seem to be more accurate than their original versions $D_{\rho}$. 

14
5 Conclusion

In this paper we show how GEL can be used to obtain inferences for the parametric component of semiparametric models. In particular we propose a computationally simple two-step method in which nuisance parameters (the infinite dimensional and possibly part of the finite dimensional ones) are profiled out using standard (semi-parametric techniques) methods (first step) and then the resulting profiled estimating equation are used for inference using standard GEL approach. We investigated the finite sample properties of the proposed test statistic using simulations. The latter seem to suggest that the GEL approach compares favourably with existing methods and demonstrates its potential in the analysis of semiparametric models.

References


### 6 Appendix

Throughout the Appendix C denotes a generic positive constant that maybe different in different uses, “M”, “CS”, “T” denote Markov, Cauchy-Schwarz and Triangle inequalities, “CMT”, “LLN” and “CLT” denote Continuous Mapping Theorem, Law of Large Numbers and Central Limit Theorem, respectively.
Proof of Theorem 1. We use the same arguments of Guggenberger & Smith (2005). Let \( c_n = n^{-1/2} \max_i \left\| m_{1i} (\theta_{10}, \hat{\theta}_2, \hat{h}) \right\| \) and \( \Lambda_n = \left\{ \lambda : \| \lambda \| \leq n^{-1/2} c_n^{-1/2} \right\} \). Then \( \sup_{\lambda \in \Lambda_n} \left| \lambda^T m_{1i} (\theta_{10}, \hat{\theta}_2, \hat{h}) \right| = o_p (1) \) and by (\( \rho \)), a second order Taylor expansion about \( \lambda = 0 \) and (\( \Omega \))

\[
0 \leq \hat{P}_\rho \left( \theta_{10}, \hat{\theta}_2, \hat{h} \right) - \rho (0) \leq -2 \hat{\lambda}^T \hat{m}_1 \left( \theta_{10}, \hat{\theta}_2, \hat{h} \right) - \hat{\lambda}^T \hat{\Omega} \left( \theta_{10}, \hat{\theta}_2, \hat{h} \right) \hat{\lambda} \\
\leq 2 \| \hat{\lambda} \| \left\| \hat{m}_1 \left( \theta_{10}, \hat{\theta}_2, \hat{h} \right) \right\| - C \| \hat{\lambda} \|^2.
\]

By (\( N \)) we have that \( \| \hat{\lambda} \| \leq O_p (n^{-1/2}) \) and \( \hat{\lambda} \in \Lambda_n \) a.s. By construction

\[
\sum_{i=1}^n \rho_1 \left( \hat{\lambda}^T m_{1i} \left( \theta_{10}, \hat{\theta}_2, \hat{h} \right) \right) m_{1i} \left( \theta_{10}, \hat{\theta}_2, \hat{h} \right) = 0
\]

so that by mean value expansion about 0

\[
0 = -\hat{m}_1 \left( \theta_{10}, \hat{\theta}_2, \hat{h} \right) - \hat{\Omega} \left( \theta_{10}, \hat{\theta}_2, \hat{h} \right) \hat{\lambda}.
\]

By (\( \Omega \)) it follows that \( n^{1/2} \hat{\lambda} = -\Omega_{10}^{-1} n^{1/2} \hat{m}_1 \left( \theta_{10}, \hat{\theta}_2, \hat{h} \right) + o_p (1) \) and the conclusion follows by (\( N \)) and CMT. The second conclusion of the theorem follows immediately by CMT, noting that \( n^{1/2} \hat{\lambda} \overset{d}{\to} N (0, \Omega_{10}^{-1} \Omega_{10}^{-1}) \).

Proof of Proposition 2. We first we show that (\( C \)) holds for \( \hat{h} = \left[ \hat{g} \left( x_{1i}^T \theta_{10} \right), \hat{g}' \left( x_{1i}^T \theta_{10} \right) \right] \), and \( \hat{\theta}_{22} = \arg \min_{\theta_{22} \in \Theta} \left\| m_{2i} \left( \theta_{10}, \theta_{22}, \hat{h} \right) \right\| \). Let \( \hat{g} \left( x_{1i}^T \theta_{10} \right) = \hat{g} \) and \( g_0 \left( x_{1i}^T \theta_{10} \right) = g_0 \), and similarly for \( \hat{g}' \) and \( g_0' \). Note that using the same results of Xue & Zhu (2006) it can be shown that for \( \beta \geq 2 \)

\[
E \left[ |\hat{g} - g_0|^\beta \right] = O \left( h^{2\beta} \right) + O \left( n^{-\beta/2} h^{1-\beta} \right) \quad (14)
\]

\[
E \left[ |\hat{g}' - g_0'|^\beta \right] = O \left( h_1^{\beta} \right) + O \left( n^{-\beta/2} h_1^{1-2\beta} \right),
\]

uniformly in \( X_1 \) and thus \( \sup_{X_1} \left| \hat{h} - h_0 \right| \overset{p}{\to} 0 \). By T, LNN, CS and (14)

\[
\left\| \hat{\theta}_2 - \theta_{20} \right\| \leq \left\| \left( \sum_{i=1}^n x_{2i} x_{2i}^T \right)^{-1} \right\| \left\| \sum_{i=1}^n x_{2i} \varepsilon_i \right\| + \left\| \sum_{i=1}^n x_{2i} (\hat{g} - g_0) \right\| \\
= o_p (1) + O_p (1) \left( O \left( h^2 \right) + O \left( nh^{-1} \right) \right) \to 0.
\]
Next we verify $(M)$. Note that by $T$

$$
\max_i \| m_{1i} \left( \theta_{10}, \tilde{\theta}_2, \tilde{\theta}_3 \right) \| \leq C \max_i \| g'_0 x_{1i} \| + C \max_i \left\| (\hat{g}' - g'_0) x_{1i} \| + \\
\| \tilde{\theta}_2 - \theta_2 \| \max_i \| g'_0 x_{1i} \| + C \max_i \| g'_0 x_{1i} (\hat{g} - g) \| + C \max_i \left\| (\hat{g}' - g'_0) x_{1i} (\hat{g} - g_0) \right\| + C \| \tilde{\theta}_2 - \theta_2 \| \max_i \| x_{1i} x_{2i} (\hat{g} - g_0) \| = \sum_{j=1}^{6} T_j.
$$

By Borel-Cantelli lemma both $T_1$ and $T_3$ are $o_{a.s.} \left( n^{1/2} \right)$ since $E \left( T_j^2 \right) < \infty \ (j = 1, 3)$. By $M$ for any $\epsilon > 0$

$$
Pr \left( n^{1/2} \| T_2 \| > \epsilon \right) \leq \sum_{i=1}^{n} E \left[ \left\| (\hat{g}' - g'_0) x_{1i} \| \right\|^2 / (n \epsilon^2) \right] = O \left( h_{1}^{2} \right) + O \left( (nh_{1}^{-3})^{-1} \right) \rightarrow 0
$$

$$
Pr \left( n^{1/2} \| T_4 \| > \epsilon \right) \leq \sum_{i=1}^{n} E \left[ (\hat{g} - g_0) x_{1i} g'_0 \| \right)^2 / (n \epsilon^2) = O \left( h^4 \right) + O \left( (nh)^{-1} \right) \rightarrow 0
$$

and similarly for $T_6$. Finally $M$ and $CS$ show that

$$
Pr \left( n^{1/2} \| T_5 \| > \epsilon \right) = O \left( h^8 + n^{-2} h^{-3} \right)^{1/2} O \left( h_{1}^{4} + n^{-2} h_{1}^{-7} \right)^{1/2} \rightarrow 0.
$$

Next we show that $(\Omega_{10})$ holds. Note that

$$
\hat{\Omega}_{10} \left( \theta_{10}, \tilde{\theta}_2, \tilde{\theta}_3 \right) = \sum_{i=1}^{n} \left[ g'_0 x_{1i} J_{0}^{\left(j\right)} \right]^{\tau} [ g'_0 x_{1i} J_{0}^{\left(j\right)} ] \tilde{\epsilon}_i / n + \sum_{j=1}^{n} R_i R'_i / n + \sum_{i=1}^{n} [ g'_0 x_{1i} J_{0}^{\left(j\right)} ]^{\tau} \tilde{\epsilon}_i R^\tau / n + \sum_{i=1}^{n} R_i [ g'_0 x_{1i} J_{0}^{\left(j\right)} ] \tilde{\epsilon}_i / n = \sum_{j=1}^{4} T_j,
$$

where

$$
R_i = \left( \hat{g}' - g'_0 \right) \left[ g'_0 x_{1i} J_{0}^{\left(j\right)} \right]^{\tau} \tilde{\epsilon}_i + (g_0 - \hat{g}) \left[ g'_0 x_{1i} J_{0}^{\left(j\right)} \right]^{\tau} + \left[ g'_0 x_{1i} J_{0}^{\left(j\right)} \right]^{\tau} \tilde{\epsilon}_i (\hat{g}' - g'_0)
$$

$$
= \sum_{j=1}^{5} R_{ij}.
$$

By LLN $T_1 \overset{p}{\rightarrow} \Omega_{10}$. Next note that $\sum_{i=1}^{n} R_i \| / n \leq \sum_{i=1}^{n} \sum_{j=1}^{5} C \| \| / n$, and that by (14) $E \| R_{i1} \|^2 = O \left( h_1^2 \right) + O \left( n^{-1} h_1^{-3} \right) \rightarrow 0$, $E \| R_{i2} \|^2 = O \left( h^4 \right) + O \left( (nh)^{-1} \right) \rightarrow 0$, 18
\[ E \| R_{i4} \|^2 = O \left( h^8 + n^{-2}h^{-3} \right)^{1/2} O \left( h_1^4 + n^{-2}h_1^{-7} \right)^{1/2} \rightarrow 0 \quad \text{and} \quad E \| R_{i5} \|^2 = O \left( h_1^2 \right) + O \left( n^{-1}h_1^{-3} \right) \quad \text{while by consistency of} \ \hat{\theta}_2 \ \text{and CMT} \ \sum_{i=1}^n \| R_{i3} \|^2 / n \ \overset{p}{\to} \ 0. \ \text{Thus LLN} \ \sum_{i=1}^n \| R_i \|^2 / n \ \overset{p}{\to} \ 0 \quad \text{and hence by CS} \ T_1 \ \overset{p}{\to} \ 0. \ \text{Similarly by CS} \]

\[ \| T_j \| \leq \left( \frac{n}{\left( \sum_{i=1}^n \| R_i \|^2 / n \right)} \right)^{1/2} \left( \frac{n}{\left( \sum_{i=1}^n \left\| g_0^i x_{i1}^\tau J_{g_0}^{\tau} \right\|^2 / n \right)} \right)^{1/2} \rightarrow 0 \]

for \( j = 3, 4 \). Finally we verify (N). Let \( \bar{x}_{J_{g_0}^{(j)}} = J_{g_0}^{\tau} \left[ x_{1i} - E \left( x_{1i} | x_{1i}^\tau \theta_{10} \right) \right] \) and \( \bar{x}_{2i} = x_{2i} - E \left( x_{2i} | x_{1i}^\tau \theta_{10} \right) \).

\[ \sum_{i=1}^n m_{1i} \left( \theta_{10}, \hat{\theta}_2, h_0 \right) / n^{1/2} = \sum_{i=1}^n \hat{g}_0 J_{g_0}^{\tau} x_{1i} \left[ 1 - x_{2i}^\tau \Omega_{22}^{-1} \bar{x}_{2i} \right] \varepsilon_i / n^{1/2} + o_p \left( 1 \right) \]

where \( \Omega_{22} = E \left( \bar{x}_{2i} \bar{x}_{2i}^\tau \right) \). Note that

\[ \sum_{i=1}^n \sum_{i=1}^n m_{1i} \left( \theta_{10}, \hat{\theta}_2, h \right) / n^{1/2} = \sum_{i=1}^n \hat{g}_0 \bar{x}_{J_{g_0}^{(j)}} x_{1i} \left[ 1 - \bar{x}_{2i}^\tau \Omega_{22}^{-1} \bar{x}_{2i} \right] \varepsilon_i / n^{1/2} + \]

\[ \sum_{i=1}^n \left( \hat{g}_0 - g_0 \right) J_{g_0}^{\tau} x_{1i} \varepsilon_i / n^{1/2} + \sum_{i=1}^n \left( \hat{g}_0 - g_0 \right) \left( g_0 - \hat{g} \right) J_{g_0}^{\tau} x_{1i} / n^{1/2} + \]

\[ \sum_{i=1}^n g_0 \left( g_0 - \hat{g} \right) J_{g_0}^{\tau} x_{1i} / n^{1/2} + \sum_{i=1}^n g_0 \varepsilon_i E \left( J_{g_0}^{\tau} x_{1i} | x_{1i}^\tau \theta_{10} \right) / n^{1/2} - \]

\[ \sum_{i=1}^n g_0 \varepsilon_i x_{J_{g_0}^{(j)}} E \left( x_{2i}^\tau | x_{1i}^\tau \theta_{10} \right) / n^{1/2} = \sum_{j=1}^6 T_j. \]

\( T_1 \ \overset{d}{\to} \ N \left( 0, E \left[ q_i \left( \theta_{10} \right) q_i \left( \theta_{10} \right)^\tau \right] \right) \) by CLT. Next again by Lemma 3 \( E \| T_2 \|^2 = O \left( h_1^2 \right) + O \left( n^{-1}h_1^{-3} \right), \ E \| T_3 \|^2 = O \left( h^8 + n^{-2}h^{-3} \right)^{1/2} O \left( h_1^4 + n^{-2}h_1^{-7} \right)^{1/2}, \ E \| T_4 \|^2 = O \left( h^4 \right) + O \left( (nh)^{-1} \right) \) and thus \( T_j \ \overset{p}{\to} \ 0 \ (j = 2, 3, 4) \). Finally using similar arguments as those used by Xue & Zhu (2006)

\[ T_5 = \sum_{i=1}^n \left[ \sum_{j=1}^n \hat{w}_j \left( \theta_{10}, t \right) g_0 + g'_0 \right] \varepsilon_i E \left( J_{g_0}^{\tau} x_{1i} | x_{1i}^\tau \theta_{10} \right) / n^{1/2} + \]

\[ \sum_{i=1}^n \sum_{j=1}^n \hat{w}_j \left( \theta_{10}, t \right) g_0 \varepsilon_i E \left( J_{g_0}^{\tau} x_{1i} | x_{1i}^\tau \theta_{10} \right) / n^{1/2} = \sum_{j=1}^2 T_{5j}. \]
and

\[ E \|T_{51}\|^2 = \sum_{i=1}^{n} \left( E \left[ \sum_{j=1}^{n} -\tilde{w}_j (\theta_{10}, t) g_0 \left( x_{ij}^\tau \theta_{10} \right) + g_0' \left( x_{ij}^\tau \theta_{10} \right) \right] \right)^{1/2} \times \]

\[ \left( E \left[ \tilde{\varepsilon}_i E \left( J^\tau_{\theta_0^*} x_{1i} | x_{1i}^\tau \theta_{10} \right) \right] ^{4} \right)^{1/2} / n = O (h_1^2) \to 0 \]

\[ E \|T_{52}\|^2 = \sum_{i=1}^{n} E \left[ \left( \sum_{j=1}^{n} \tilde{w}_j (\theta_{10}, t) \right)^{4} \right] ^{1/4} (E \|g_0 \tilde{\varepsilon}_i\|^{4})^{1/4} \times \]

\[ \left( E \left[ E \left( J^\tau_{\theta_0^*} x_{1i} | x_{1i}^\tau \theta_{10} \right) \right] ^{4} \right)^{1/4} / n = O \left( n^{-3/2} h_1^{-7/2} \right) \to 0 , \]

and similarly for \( T_6 \). Thus all of the conditions of Theorem 1 are met, hence the conclusion. ■

**Proof of Proposition 3.** We first show that (C) holds. The results of Chen et al. (2002) show that

\[ \partial \sum_{i=1}^{n} x_{2i} \left[ dN_i (t) - Y_i (t) d\Lambda_\varepsilon \left( \bar{g} (t) + x_{1i}^\tau \theta_{10} + x_{2i}^\tau \theta_2 \right) \right] \right] / \theta_2 \overset{P}{\to} \Omega_2 \]

uniformly in \( t \in (0, \tau] \) and \( \theta_2 \in \Gamma (\theta_{20}, \delta) \) where \( \Gamma (\theta_{20}, \delta) \) is an open ball centred at \( \theta_{20} \) with radius \( \delta \) and

\[ \Omega_2 = - \int_0^\tau E \left[ (x_{2i} - \mu_2 (t)) x_{2i}^\tau \lambda_\varepsilon (g (t) + x_{1i}^\tau \theta) Y_i (t) \right] dg (t) / n \]

which can be shown to be negative semi-definite. Thus \( \sum_{i=1}^{n} m_{2i} (\theta_{10}, \theta_2, \bar{g}) = 0 \) is quasi-convex with probability approaching 1 and therefore \( \theta_2 \) is consistent. Given the consistency of \( \theta_2 \), the consistency of \( \bar{g} \) follows using the same arguments of Chen et al. (2002) since

\[ \int_0^\tau \sum_{i=1}^{n} \left[ dN_i (t) - Y_i (t) d\Lambda_\varepsilon \left( g (t) + x_{1i}^\tau \theta \right) \right] = \int_0^\tau \sum_{i=1}^{n} \left[ dN_i (t) - Y_i (t) d\Lambda_\varepsilon (g (t) + x_{1i}^\tau \theta_0) \right] + o_p (1) . \]
Next we investigate (M). By the consistency of $\hat{\theta}_2$

\[
\max_i \left\| m_{1i} \left( \theta_{10}, \hat{\theta}_2, \hat{h} \right) \right\| \leq \max_i \left\| \int_0^\infty x_{1i} \left[ dN_i (t) - Y_i (t) \right] d\Lambda_\varepsilon (g_0 (t) + x_i^T \theta_0) \right\| + \frac{o_p (1)}{\alpha_{\alpha, s}} (n^{1/2}) + X_0 O_p (1) o_p (1)
\]

by the Borel-Cantelli Lemma and M respectively.

Next we show that (\Omega) holds. Note that

\[
\left\| \hat{\Omega}_1 \left( \theta_{10}, \hat{\theta}_2, \hat{h} \right) - \hat{\Omega} (\theta_0, h) \right\| \leq \left\| \Lambda_\varepsilon \left( g \left( t, \theta_{10}, \hat{\theta}_2 \right) + x_{1i}^T \theta_{10} + x_{2i}^T \hat{\theta}_2 \right) - \Lambda_\varepsilon (g_0 (t) + x_i^T \theta_0) \right\| \times \left\| \sum_{i=1}^n x_{1i} x_{i}^T / n \right\| = o_p (1) O_p (1)
\]

by consistency of $\hat{g}$, $\hat{\theta}_2$ and CMT. Thus $\hat{\Omega} (\theta_0, h) \xrightarrow{p} \Omega_0$ by T and LLN. Finally (N) holds for $\sum_{i=1}^n m_{i} \left( \theta_{10}, \hat{\theta}_2, \hat{h} \right) / n^{1/2}$ since (Lu & Liang 2006)

\[
\sum_{i=1}^n m_{1i} \left( \theta_{10}, \hat{\theta}_2, \hat{h} \right) / n^{1/2} = \sum_{i=1}^n m_{1i} \left( \theta_{10}, \theta_{20}, h \right) / n^{1/2} - \sum_{i=1}^n x_{1i} \partial \Lambda_\varepsilon \left( g \left( \theta_{10}, \hat{\theta}_2 \right) + x_i^T \theta_0 \right) / \partial \theta_2 \left( \hat{\theta}_2 - \theta_{20} \right) n^{1/2} \]

\[
= \sum_{i=1}^n \int_0^\infty \left\{ x_{1i} - \hat{\mu}_1 (t) - \hat{\Omega}_{12} \left( \hat{\theta} \right) \hat{\Omega}_{22} \left( \hat{\theta} \right)^{-1} (x_{2i} - \hat{\mu}_2 (t)) \right\} dM_i (t) / n^{1/2} + o_p (1)
\]

and the result follows by CLT. Thus all the conditions of Theorem 1 are met hence the conclusion.
7 Tables

Table 1. Empirical size (in %) for single index model at 5% and 10% significance level

<table>
<thead>
<tr>
<th></th>
<th>50</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_{EL}$</td>
<td>7.913</td>
<td>12.713</td>
</tr>
<tr>
<td>$D^{a2}_{EL}$</td>
<td>7.452</td>
<td>11.783</td>
</tr>
<tr>
<td>$LM_{EL}$</td>
<td>8.781</td>
<td>13.506</td>
</tr>
<tr>
<td>$D_{ET}$</td>
<td>7.337</td>
<td>12.121</td>
</tr>
<tr>
<td>$D^{a2}_{ET}$</td>
<td>7.393</td>
<td>12.056</td>
</tr>
<tr>
<td>$LM_{ET}$</td>
<td>7.664</td>
<td>12.171</td>
</tr>
<tr>
<td>$D_{EU}$</td>
<td>8.823</td>
<td>13.229</td>
</tr>
<tr>
<td>$D^{a2}_{EU}$</td>
<td>8.202</td>
<td>12.842</td>
</tr>
<tr>
<td>$LM_{EU}$</td>
<td>7.813</td>
<td>11.902</td>
</tr>
</tbody>
</table>

Table 2. Empirical size (in %) for partially linear model at 5% and 10% significance level

<table>
<thead>
<tr>
<th></th>
<th>50</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_{EL}$</td>
<td>8.859</td>
<td>13.389</td>
</tr>
<tr>
<td>$D^{a2}_{EL}$</td>
<td>8.344</td>
<td>12.516</td>
</tr>
<tr>
<td>$LM_{EL}$</td>
<td>9.833</td>
<td>14.75</td>
</tr>
<tr>
<td>$D_{ET}$</td>
<td>8.063</td>
<td>12.095</td>
</tr>
<tr>
<td>$D^{a2}_{ET}$</td>
<td>8.354</td>
<td>12.531</td>
</tr>
<tr>
<td>$LM_{ET}$</td>
<td>8.579</td>
<td>13.298</td>
</tr>
<tr>
<td>$D_{EU}$</td>
<td>9.666</td>
<td>15.079</td>
</tr>
<tr>
<td>$D^{a2}_{EU}$</td>
<td>9.184</td>
<td>14.235</td>
</tr>
<tr>
<td>$LM_{EU}$</td>
<td>8.472</td>
<td>13.301</td>
</tr>
<tr>
<td>$W$</td>
<td>10.254</td>
<td>16.201</td>
</tr>
</tbody>
</table>
Table 3. Empirical size (in %) for transformation model at 5% and 10% nominal level

<table>
<thead>
<tr>
<th></th>
<th>cens %</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>200</td>
</tr>
<tr>
<td>( n )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( D_{EL} )</td>
<td>25</td>
<td>7.616</td>
<td>13.542</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>6.384</td>
<td>11.526</td>
</tr>
<tr>
<td>( D_{EL}^{a2} )</td>
<td>25</td>
<td>7.006</td>
<td>12.594</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>6.192</td>
<td>10.604</td>
</tr>
<tr>
<td>( LM_{EL} )</td>
<td>25</td>
<td>8.529</td>
<td>12.188</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>6.830</td>
<td>10.834</td>
</tr>
<tr>
<td>( D_{ET} )</td>
<td>25</td>
<td>7.387</td>
<td>13.298</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>6.001</td>
<td>10.562</td>
</tr>
<tr>
<td>( D_{ET}^{a2} )</td>
<td>25</td>
<td>7.235</td>
<td>13.096</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>6.128</td>
<td>11.277</td>
</tr>
<tr>
<td>( LM_{ET} )</td>
<td>25</td>
<td>7.463</td>
<td>13.882</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>6.703</td>
<td>12.133</td>
</tr>
<tr>
<td>( D_{EU} )</td>
<td>25</td>
<td>7.762</td>
<td>13.542</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>7.038</td>
<td>11.526</td>
</tr>
<tr>
<td>( D_{EU}^{a2} )</td>
<td>25</td>
<td>7.556</td>
<td>12.619</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>7.046</td>
<td>11.696</td>
</tr>
<tr>
<td>( LM_{EU} )</td>
<td>25</td>
<td>7.616</td>
<td>12.871</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>6.384</td>
<td>10.853</td>
</tr>
<tr>
<td>( W )</td>
<td>25</td>
<td>4.878</td>
<td>11.542</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>4.788</td>
<td>11.06</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>200</td>
</tr>
<tr>
<td>-----</td>
<td>----------</td>
<td>-------------</td>
<td>-------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\gamma = 1$</td>
<td></td>
</tr>
<tr>
<td>$D_{EL}$</td>
<td>25</td>
<td>7.387</td>
<td>12.188</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>6.320</td>
<td>10.604</td>
</tr>
<tr>
<td>$D_{EL}^{a2}$</td>
<td>25</td>
<td>6.726</td>
<td>11.586</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>5.882</td>
<td>10.180</td>
</tr>
<tr>
<td>$LM_{EL}$</td>
<td>25</td>
<td>8.018</td>
<td>11.822</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>6.421</td>
<td>9.967</td>
</tr>
<tr>
<td>$D_{ET}$</td>
<td>25</td>
<td>7.018</td>
<td>11.968</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>5.640</td>
<td>9.717</td>
</tr>
<tr>
<td>$D_{ET}^{a2}$</td>
<td>25</td>
<td>6.945</td>
<td>12.441</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>5.699</td>
<td>10.826</td>
</tr>
<tr>
<td>$LM_{ET}$</td>
<td>25</td>
<td>7.314</td>
<td>13.188</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>6.502</td>
<td>11.284</td>
</tr>
<tr>
<td>$D_{EU}$</td>
<td>25</td>
<td>7.141</td>
<td>12.865</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>6.404</td>
<td>11.065</td>
</tr>
<tr>
<td>$D_{EU}^{a2}$</td>
<td>25</td>
<td>6.800</td>
<td>11.231</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>6.270</td>
<td>10.643</td>
</tr>
<tr>
<td>$LM_{EU}$</td>
<td>25</td>
<td>7.235</td>
<td>11.455</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>6.128</td>
<td>10.419</td>
</tr>
<tr>
<td>$W$</td>
<td>25</td>
<td>4.634</td>
<td>11.08</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>4.644</td>
<td>10.728</td>
</tr>
</tbody>
</table>
Table 3. Continued

<table>
<thead>
<tr>
<th>$n$</th>
<th>cens %</th>
<th>$\gamma = 2$</th>
<th>100</th>
<th>200</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$D_{EL}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>25</td>
<td>5.944</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>50</td>
<td>6.004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$D_{EL}^a$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>25</td>
<td>6.255</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>50</td>
<td>5.412</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$LM_{EL}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>25</td>
<td>7.612</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>50</td>
<td>5.971</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$D_{ET}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>25</td>
<td>6.456</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>50</td>
<td>5.358</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$D_{ET}^a$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>25</td>
<td>6.390</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$LM_{ET}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>25</td>
<td>6.948</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>50</td>
<td>6.242</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$D_{EU}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>25</td>
<td>6.784</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>50</td>
<td>5.956</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$D_{EU}^a$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>25</td>
<td>6.460</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>50</td>
<td>6.020</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$LM_{EU}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>25</td>
<td>7.018</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>50</td>
<td>5.944</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$W$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>50</td>
<td>4.412</td>
</tr>
</tbody>
</table>