Rochester Center for Economic Research

Pairwise Comparison Estimation of Censored Transformation Models

Shakeeb Khan and Elie Tamer

Working Paper No. 495
October 2002

UNIVERSITY OF ROCHESTER

Pairwise Comparison Estimation of Censored Transformation Models

Shakeeb Khan[*]
University of Rochester

Elie Tamer[†]
Princeton University

Preliminary and Incomplete Draft
September 2002

Abstract

In this paper a pairwise comparison estimation procedure is proposed for the regression coefficients in a censored transformation model. The main advantage of the new estimator is that it can accommodate covariate dependent censoring without the requirement of smoothing parameters, trimming procedures, or stringent tail behavior restrictions. We also modify the pairwise estimator for other variations of the transformation model and propose estimators for the transformation function itself, as well as regression coefficients in heteroskedastic and panel data models. The estimators are shown to converge at the parametric (root-n) rate, and the results of a small scale simulation study indicate they perform well in finite samples. We illustrate our estimator using the Stanford Heart Transplant data and marriage length data from the CPS fertility supplement.

JEL Classification: C13,C14,C24

Key Words: transformation models, pairwise comparison, maximum rank correlation, duration analysis.

[*]Dept. Econ, University of Rochester, Rochester NY 14627; e-mail: skhan@troi.cc.rochester.edu.
[†]Dept. Econ, Princeton University, Princeton, NJ 08544; e-mail: tamer@princeton.edu.

1 Introduction

The monotonic transformation function in its most basic form is usually expressed as

$$T(y_i) = x_i'\beta_0 + \epsilon_i \quad i = 1, 2, ...n \tag{1.1}$$

where $(y_i, x_i')'$ is a $(k+1)$ dimensional observed random vector, and the random variable ϵ_i is unobserved. The function $T(\cdot)$ is assumed to be monotonic, but otherwise unspecified[1]. The k-dimensional vector β_0 is unknown, and is often the object of interest to be estimated from a random sample of n observations.[2]

The model in equation (1.1) has become increasingly popular in the applied and theoretical econometrics literature. Its popularity stems from two main reasons. First, economic theory rarely provides guidelines on how to specify functional form relationships among variables while (1.1) can accommodate many functional relationships used in practice such as linear, log-linear, or the parametric transformation in Box-Cox models, without suffering from the dimensionality problems encountered when adopting a fully nonparametric approach. The second reason is that (1.1) can be derived from a wide class of duration models which includes the Accelerated Failure Time (AFT) model and the proportional hazard model with unobserved heterogeneity which are both widely popular in the unemployment spell literature. In the proportional hazards model with unobserved heterogeneity, the function $T(\cdot)$ is related to the integrated baseline hazard function- see Ridder(1990) for details.

Several estimators for β_0 have been proposed in the econometrics and statistics literature in the case where ϵ is independent of x. The first was the Maximum Rank Correlation (MRC) estimator proposed in Han(1987)[3]. MRC maximizes the following objective function:

$$H_n(\beta) = \frac{1}{n(n-1)} \sum_{i \neq j} I[y_i > y_j] I[x_i'\beta > x_j'\beta] \tag{1.2}$$

where $I[\cdot]$ denotes the usual indicator function. Consistency of this estimator is based on the condition:

$$P(y_i \geq y_j | x_i, x_j) \geq \frac{1}{2} \quad \text{iff } x_i'\beta \geq x_j'\beta \tag{1.3}$$

[1]The transformation model is sometimes expressed even more generally than in (1.1), where additive separability between ϵ_i and $x_i'\beta_0$ is weakened to monotonicity in each argument.

[2]More recently, the unknown function $T(\cdot)$ has also been a "parameter" of interest to be estimated. While estimation of β_0 will be the initial focus of attention in this paper, we also consider estimation of $T(\cdot)$ later in the paper.

[3]A related rank estimator was proposed in Cuzick(1988).

A similar estimator was proposed in Cavanagh and Sherman(2001). Their Monotone Rank Estimator (MRE) maximized the function:

$$M_n(\beta) = \frac{1}{n(n-1)} \sum_{i \neq j} M(y_i) I[x_i'\beta > x_j'\beta] \tag{1.4}$$

where $M(.)$ is a known monotonic function. Consistency of the MRE is based on the condition:

$$E[y_i|x_i] \quad \text{is monotonic in } x_i'\beta_0 \tag{1.5}$$

which is mildly more general than the condition in (1.3).

Both the MRC and MRE involve non-continuous objective functions which makes their computation relatively difficult. The non-smoothness problem is compounded by the fact that calculation of each objective function involves $O(n \log n)$ operations, as shown in Abrevaya(2001). Nonetheless, algorithms such as Nelder-Meade and Simulated Annealing have been shown to be effective in their computation. Furthermore, they have the advantage of not involving any non-parametric procedures requiring the selection of smoothing parameters, in contrast to the estimators proposed in Powell Stock and Stoker (1989) and Ichimura (1993).

In this paper we propose a pairwise comparison estimator that can accommodate data which is subject to random covariate dependent censoring. The new estimator shares the same advantages of the original MRC- specifically it does not involve any non-parametric procedures, but will be consistent and/or more efficient than the original MRC for a wide class of censored models. Furthermore, the new estimator is numerically equivalent to the MRC for uncensored data, and data exhibiting fixed censoring.

The rest of the paper is organized as follows. The following section describes the model to be estimated and explains the disadvantages of estimating it by the MRC. This then motivates the new estimation procedure which is described in detail, and whose asymptotic properties are provided. Section 3 describes extensions of the new procedure to accommodate doubly censored, heteroskedastic, and panel data, as well as estimate $T(\cdot)$. Section 4 explores the finite sample properties of the new estimators by means of a small scale simulation study. Section 5 applies the new estimator to two data sets, and section 6 concludes by summarizing results and discussing areas for future research. The proofs used in establishing the asymptotic properties of the estimators are left to the appendix.

2 Censored Transformation Model

We consider estimation of the regression coefficients in a transformation model subject to random left censoring. Specifically, as in equation (1.1) we have

$$T(y_i) = x_i'\beta_0 + \epsilon_i \qquad i = 1, 2, ...n \tag{2.1}$$

where ϵ_i is independent of x_i but now the latent dependent variable y_i is no longer always observed. Instead one observes the pair (v_i, d_i) where

$$T(v_i) = \max(x_i'\beta_0 + \epsilon_i, c_i)$$

$$d_i = I[x_i'\beta_0 + \epsilon_i \geq c_i]$$

where c_i is the censoring random variable whose distribution may depend on the covariates x_i but conditional on x_i is assumed to be independent of ϵ_i. It is also assumed that ϵ_i and x_i are independent.

Randomly censored models have received a great deal of attention in the econometrics and statistics literature primarily when the function $T(\cdot)$ is known and strictly increasing function, such as the identity or logarithmic functions. In the latter case the model is often referred to as the Accelerated Failure Time model. Estimators for β_0 when $T(\cdot)$ is known and assumed to be strictly monotonic have been proposed in Buckley and James(1978), Koul, Sousarla and Van Rysin (1980), Ritov(1990), Tsiatsis(1991), Ying Jung and Wei(1995), Yang(1999), Honoré, Khan and Powell(2002) among others. A main disadvantage of these estimators is that they all are based on knowledge of $T(\cdot)$, and some suffer from the additional drawback of assuming that the censoring variable and the observed covariates are statistically independent.

There are few estimators for β_0 in (2.1) when $T(\cdot)$ is unknown and censoring depends on the covariates. We note that the proportional hazards model can be expressed as a transformation model (see, e.g. Ridder(1990)) in which case β_0 could be estimated (even in the presence of covariate dependent censoring) via the partial MLE in Cox(1975). However, this requires that ϵ_i have an extreme value distribution.

It is also well known that the MRC estimator, with y_i replaced with the observed variable v_i can result in a consistent estimator when c_i and x_i are independent. However, even under this strong assumption, it will be rather inefficient, as it "discards" the information in the value of the indicator d_i. Worse still, the MRC is inconsistent in the presence of covariate dependent censoring. This problem can be corrected by weighting observations of v_i by a conditional Kaplan-Meier estimator of the conditional c.d.f. of c_i given x_i as suggested by Cuzick(1988). For example, if the c.d.f. of c_i were known, one could modify the MRC and the MRE as the maximizers of :

$$H_n(\beta) = \frac{1}{n(n-1)} \sum_{i \neq j} (d_i/F(v_i))(d_j/F(v_j)) I[v_i > v_j] I[x_i'\beta > x_j'\beta] \tag{2.2}$$

and

$$M_n(\beta) = \frac{1}{n(n-1)} \sum_{i \neq j} M\left(d_i v_i / F(v_i)\right) I[x_i'\beta > x_j'\beta] \tag{2.3}$$

respectively, where here $F(\cdot)$ denotes the unknown c.d.f; this estimator can be made feasible by replacing the unknown $F(\cdot)$ with the Kaplan Meier estimator, $\hat{F}(\cdot)$. This approach suffers from several drawbacks. For one, computation of a conditional Kaplan-Meier estimator requires the selection of smoothing parameters. Furthermore, it can be very numerically unstable as it divides variables by estimators that are not bounded away from 0. Moreover, weighing by the Kaplan Meier estimator does not allow for fixed censoring.

The assumption of independence between the censoring variable and the covariates is often considered too restrictive. For example it rules out all competing risks models where the researcher only observes the minimum of two dependent variables depending on covariates and having some common covariates. Thus we feel that an estimator for the regression coefficients in a transformation model with covariate dependent censoring that is simple to implement, in the sense that it does not require smoothing parameters or trimming procedures, is something that is lacking in the literature. In this paper we propose an estimator which aims to address this problem.

Our estimator is based on results from the rank regression and pairwise comparison literature in statistics and econometrics- see e.g. Jureckova(1971), Jaeckle(1972) and Powell(1994).

To motivate an estimator in terms of the rank regression and pairwise comparison literature for the problem at hand, we define the vector $\mathbf{y}_i = (v_i, d_i)'$. To construct a rank regression estimator, analogous to Han(1987), we wish to construct a function:

$$f_{ij} \equiv f(\mathbf{y}_i, \mathbf{y}_j)$$

which satisfies the property

$$E[I[f_{ij} \geq 0]|x_i, x_j] \geq E[I[f_{ji} \geq 0]|x_i, x_j] \text{ iff } x_i'\beta_0 \geq x_j'\beta_0 \tag{2.4}$$

Alternatively, in terms of the pairwise comparison literature, we define the vector $\mathbf{z}_i = (v_i, d_i, x_i')'$, and we wish to construct the function

$$e_{ij}(\beta) \equiv e(\beta, \mathbf{z}_i, \mathbf{z}_j)$$

which satisfies

$$e_{ij}(\beta_0) - e_{ji}(\beta_0) \text{ is symmetric around 0, conditional on } x_i, x_j \tag{2.5}$$

For the uncensored transformation model, Han(1987) sets $f_{ij} = y_i - y_j$. For the problem at hand with covariate dependent censoring, we propose an alternative form for $f_{ij}(\beta)$ that satisfies (2.4), and a resulting rank regression estimator. This will also suggest a form for $e_{ij}(\beta)$, and place the estimator within the class of pairwise comparison estimators.

We first define the random variables

$$y_{1i} = v_i \tag{2.6}$$
$$y_{0i} = d_i v_i + (1 - d_i) \cdot (-\infty) \tag{2.7}$$

from which we define f_{ij}, and consequently $I[f_{ij} \geq 0]$ as

$$f_{ij} = y_{1i} - y_{0j} \tag{2.8}$$
$$I[f_{ij} \geq 0] = (1 - d_j) + d_j(v_i - v_j) \tag{2.9}$$

We wish to show that (2.4) holds for the censored transformation model. Another way to motivate our estimator is to notice that by the definition of y_1 and y_0, we have

$$y_0 \leq y \leq y_1$$

and hence that

$$y_0 \leq T^{-1}(x\beta_0 + \epsilon) \leq y_1$$

which implies that

$$x_i\beta \geq x_j\beta \quad \Rightarrow \Pr(y_{1i} \geq y_{0j}) \quad \geq \frac{1}{2} \tag{2.10}$$

Our result is based on the following conditions:

I1 Letting S_X denote the support of x_i, and let $\mathcal{X}_{uc}$ denote the set

$$\{x \in S_X : P(d_i = 1 | x_i = x) > 0\}$$

Then $\mathcal{X}_{uc}$ has positive measure.

I2 The random variable ϵ_i is distributed independently of the random vector $(c_i, x_i')'$.

I3 S_X is not contained in any proper linear subspace of R^k. Furthermore, the first component of x_i has everywhere positive Lebesgue density, conditional on the other components.

We have the following identification result, whose proof is left to the appendix.

Lemma 2.1 *Under Assumptions I1-I3, (2.4) holds.*

It is this result which motivates our estimator. Before describing it in detail, we note that the object of interest β_0 is only identified up to scale as the function $T(\cdot)$ is unknown. Following convention, we set the first component of the vector β_0 to 1, express $\beta_0 = (1, \theta_0')'$ and consider estimation of θ_0. We let $x^{(1)}$ denote the first component of x_i and x_i^{-1} denote its remaining components. Following standard notation, for any $\theta \in \Theta$, we let β denote $(1, \theta')'$.

Our censoring robust rank estimator, which we refer to hereafter as CRMRC, is of the form:

$$\hat{\theta} = \arg\max_{\theta\in\Theta} \frac{1}{n(n-1)} \sum_{i\neq j} I[f_{ij} \geq 0] I[x_i'\beta \geq x_j'\beta] \tag{2.11}$$

$$= \arg\max_{\theta\in\Theta} \frac{1}{n(n-1)} \sum_{i\neq j} (d_j I[v_i \geq v_j] + (1 - d_j)) I[x_i'\beta \geq x_j'\beta] \tag{2.12}$$

where Θ denotes the parameter space.

Remark 2.1 *The above estimator is numerically equivalent to maximizing the objective function:*

$$\frac{1}{n(n-1)} \sum_{i \neq j} I[y_{0i} \geq y_{1j}] I[x_i'\beta_0 \geq x_j'\beta_0] = \frac{1}{n(n-1)} \sum_{i \neq j} d_i I[v_i \geq v_j] I[x_i'\beta_0 \geq x_j'\beta_0]$$

Interestingly, our estimator has an inherent "asymmetry" in the objective function, where we include one censoring indicator, but not the other. It appears that this asymmetry is what permits us to accommodate covariate dependent censoring.

Remark 2.2 *To interpret the above as a pairwise comparison estimator, we can define*

$$e_{ij}(\beta) = sgn\{I[f_{ij}(\beta_0) > 0] - I[(x_i - x_j)'\beta_0 > 0]\}$$

and the conditional symmetry of $e_{ij}(\beta) - e_{ji}(\beta_0)$ *follows from Lemma 2.1*

We first establish consistency of the CRMRC. For this we require the additional condition that the parameter space is compact:

I4 Θ is a compact subset of $\mathbf{R}^{k-1}$.

The following theorem, whose proof is left to the appendix, establishes the consistency of the CRMRC.

Theorem 2.1 *Under Assumptions* **I1**-**I4**,

$$\hat{\theta} \xrightarrow{p} \theta_0$$

Remark 2.3 *We note that the consistency of the proposed estimator follows from Lemma 2.1. Consequently, as is the case with the MRC and MRE, the estimator is applicable to models even more general than (1.1). Specifically, additive separability between* $x_i'\beta_0$ *and* ϵ_i *is often not required if we have monotonicity in each of the two arguments.*

We now establish the limiting distribution theory of the CRMRC. The arguments are completely analogous to those used in Sherman(1993) for establishing the asymptotic distribution of the MRC. Our results are based on a set of assumptions analogous to those found in Sherman(1993), and we deliberately choose notation to match his as closely as possible.

Recalling that $\mathbf{z}_i$ denotes the vector $(d_i, v_i, x_i')'$, we define

$$\begin{aligned}\tau(\mathbf{z},\theta) &= E[(d_i I[v \geq v_i] + (1-d_i))I[x'\beta \geq x_i'\beta]] \\ &+ E[(dI[v_i \geq v] + (1-d))]I[x_i'\beta \geq x'\beta]]\end{aligned}$$

Finally, we let $\mathcal{N}$ denote a neighborhood of θ_0.

A1 θ_0 lies in the interior of Θ, a compact subset of R^{k-1}.

A2 For each $\mathbf{z}$, the function $\tau(\mathbf{z},\cdot)$ is twice differentiable in a neighborhood of θ_0. Furthermore, the vector of second derivatives of $\tau(\mathbf{z},\cdot)$ satisfies the following Lipschitz condition:

$$\|\nabla_2\tau(\mathbf{z},\theta) - \nabla_2\tau(\mathbf{z},\theta_0)\| \leq M(\mathbf{z})\|\theta - \theta_0\|$$

where ∇_2 denotes the second derivative operator and $M(\cdot)$ denotes an integrable function of z.

A3 $E[\|\nabla_1\tau(\mathbf{z}_i,\theta_0)\|^2]$ and $E[\|\nabla_2\tau(\mathbf{z}_i,\theta_0)\|]$ are finite.

A4 $E[\nabla_2\tau(\mathbf{z}_i,\theta_0)]$ is non-singular.

We now state the main theorem, characterizing the asymptotic distribution of the CRMRC; its proof is left to the appendix.

Theorem 2.2 *Under Assumptions I1-I4, A1-A4,*

$$\sqrt{n}(\hat{\theta} - \theta_0) \Rightarrow N(0, V^{-1}\Delta V^{-1}) \tag{2.13}$$

where $V = E[\nabla_2\tau(\mathbf{z}_i,\theta_0)]/2$ *and* $\Delta = E[\nabla_1\tau(\mathbf{z}_i,\theta_0)\nabla_1\tau(\mathbf{z}_i,\theta_0)']$.

We conclude this section with a brief discussion on conducting inference with the CRMRC. The asymptotic variance matrix can be estimated in a similar fashion to the estimator in Sherman(1993). As is the case with that estimator, the selection of smoothing parameters will be required. Unfortunately, it has not been formally established that the bootstrap is asymptotically valid in this setting, or else inference could be conducted without the selection of smoothing parameters.

Also, the CRMRC can be used to construct model specification tests by comparing its value to those of existing estimators. For example, the CRMRC may be compared to the MRE or MRC to test for the presence of covariate dependent censoring. We can compare the CRMRC to the relative coefficients obtained from Cox's partial likelihood estimator (PLE) to test for the presence of unobserved heterogeneity. Also, we can compare the CRMRC to relative coefficients obtained from the Tsiatsis(1990) and/or Ying(1995) estimators, to test for particular functional forms of the transformation.

3 Extensions of the CRMRC

In this section we propose two extensions of the CRMRC to accommodate doubly censored and heteroskedastic data.

3.1 Doubly Censored Data

Many data sets are subject to double (i.e. left and right) random censoring. Examples are when the dependent variable is duration until an event occurs, and individuals are regularly and frequently surveyed or tested for an interval of time. If the occurrence of the event (e.g. unemployment, cancerous tumor) is detected on the first survey/test, the duration is left censored, and if no such events have occurred by the last survey/test, the duration is right censored.

In the monotonic transformation framework, the doubly censored regression model can be expressed as follows. (1.1) still holds, but the econometrician does not always observe the dependent variable $y_i \equiv T^{-1}(x_i'\beta_0 + \epsilon_i)$. Instead one observes the doubly censored sample, which we can express as the pair (v_i, d_i) where

$$d_i = I[c_{1i} < x_i'\beta_0 + \epsilon_i \leq c_{2i}] + 2 \cdot I[x_i'\beta_0 + \epsilon_i \leq c_{1i}] + 3 \cdot I[c_{2i} > x_i'\beta_0 + \epsilon_i]$$

$$v_i = I[d_i = 1] \cdot (x_i'\beta_0 + \epsilon_i) + I[d_i = 2]c_{1i} + I[d_i = 3]c_{2i}$$

where $I[\cdot]$ denotes the usual indicator function, c_{1i}, c_{2i} denote left and right censoring variables, whose distributions may depend on the covariates x_i and who satisfy $P(c_{1i} < c_{2i}) = 1$.

For the double censored regression model estimators have been proposed by Zhang and Li(1996), Ren and Gu(1997) to name a few. Both of these require a linear regression specification and the censoring variables to be independent of the covariates.

With $T(\cdot)$ unknown, once can again perform MRC using v_i as the dependent variable if x_i is independent of (c_{1i}, c_{2i}). However in the doubly censored case the efficiency loss can be very severe for ignoring the value of d_i.

To estimate β_0 in the general model with $T(\cdot)$ unknown and covariate dependent censoring, we first define y_{1i}, y_{0i} as

$$y_{1i} = I[d_i < 3]v_i I[d_i = 3] \cdot +\infty \tag{3.1}$$

$$y_{0i} = I[d_i \neq 2]v_i + I[d_i = 2] \cdot -\infty \tag{3.2}$$

and accordingly we may define $f_{ij}, I[f_{ij} \geq 0]$ as:

$$f_{ij} = y_{1i} - y_{0j}$$

$$I[f_{ij} \geq 0] = I[d_i = 3] + I[d_j = 2] - (I[d_i = 3] * I[d_j = 2]) + (I[d_i = 1] + I[d_i = 2]) * (I[d_j =$$

Letting d_{1i}, d_{2i}, d_{3i} denote $I[d_i = 1], I[d_i = 2], I[d_i = 3]$, respectively, we can express the CRMRC for doubly censored data as:

$$\begin{aligned} \hat{\theta} &= \arg\max_{\theta \in \Theta} \frac{1}{n(n-1)} \sum_{i \neq j} ((d_{1i} + d_{2i}) \cdot (d_{1j} + d_{3j}) I[v_i \geq v_j] \\ &+ (d_{3i} + d_{2j} - d_{3i} d_{2j})) I[x_i'\beta \geq x_j'\beta] \end{aligned} \tag{3.3}$$

The following theorem, whose proof is left to the appendix, establishes the asymptotic distribution of the CRMRC in the doubly censored model. Asymptotic distribution theory is based on the on Assumptions AD1-AD4 below. We first need to introduce some further notation for the doubly censored case. Now $\mathbf{z}_i$ denotes the vector $(d_{1i}, d_{2i}, d_{3i}, v_i, x_i')'$, we define

$$\begin{aligned} \tau_d(\mathbf{z}, \theta) &= E[(d_1 d_{3i} I[v \geq v_i]) I[x'\beta \geq x_i'\beta]] \\ &+ E[(d_{1i} d_3 I[v_i \geq v]] I[x_i'\beta \geq x'\beta]] \end{aligned}$$

Finally, we let $\mathcal{N}$ denote a neighborhood of θ_0.

AD1 θ_0 lies in the interior of Θ, a compact subset of R^{k-1}.

AD2 For each $\mathbf{z}$, the function $\tau_d(\mathbf{z}, \cdot)$ is twice differentiable in a neighborhood of θ_0. Furthermore, the vector of second derivatives of $\tau_d(\mathbf{z}, \cdot)$ satisfies the following Lipschitz condition:

$$\|\nabla_2 \tau_d(\mathbf{z}, \theta) - \nabla_2 \tau_d(\mathbf{z}, \theta_0)\| \leq M(\mathbf{z})\|\theta - \theta_0\|$$

where ∇_2 denotes the second derivative operator and $M(\cdot)$ denotes an integrable function of z.

AD3 $E[\|\nabla_1 \tau_d(\mathbf{z}_i, \theta_0)\|^2]$ and $E[\|\nabla_2 \tau_d(\mathbf{z}_i, \theta_0)\|]$ are finite.

AD4 $E[\nabla_2 \tau_d(\mathbf{z}_i, \theta_0)]$ is non-singular.

Theorem 3.1 *Under Assumptions AD1-AD4,*

$$\sqrt{n}(\hat{\theta} - \theta_0) \Rightarrow N(0, V_d^{-1} \Delta_d V_d^{-1}) \tag{3.4}$$

where $V_d = E[\nabla_2 \tau_d(\mathbf{z}_i, \theta_0)]/2$ *and* $\Delta_d = E[\nabla_1 \tau_d(\mathbf{z}_i, \theta_0) \nabla_1 \tau_d(\mathbf{z}_i, \theta_0)']$.

3.2 Estimating the Transformation Function (incomplete)

In this section we consider estimation of the transformation function $T(\cdot)$. For the uncensored model, Cuzick(1988) showed that the "infinite dimensional parameter $T(\cdot)$ could be estimated at the parametric (root-n) rate by his proposed rank regression estimator. The estimator was then modified to accommodate random, though covariate term independent censoring. Other estimators for the transformation function have been proposed in Horowitz(1996), Gorgens and Horowitz(1999), Ye and Duan(1997) and Chen(2002). An attractive feature of the rank estimators in Cuzick(1988), Chen(2002) is that they did not require smoothing parameters. Specifically, for the uncensored model, Chen(2002) proposed maximizing the following rank based objective function with respect to γ:

$$C_n(\gamma) = \sum_{i \neq j} (I[y_i \geq y_1] - I[y_j \geq 0]) I[z_i - z_j \geq \gamma] \tag{3.5}$$

where y_1 denotes the point in the domain of $T(\cdot)$ the function is to be estimated, $z_i = x_i'\beta_0$, where the assumption of a known β_0 does not affect the rate of convergence of the estimator of $\gamma_0 \equiv T(y_1)$ since estimators of β_0 converging at the root-n rate exist.

To accommodate censoring, Cuzick(1988), Chen(2002) divided the terms in the above objective function by the estimated survivor function of the censoring variable, which could be obtained using the Kaplan Meier estimator. There are certain drawbacks with this approach which we attempt to address here. One is that the procedure breaks down with fixed censoring. More importantly, it does not allow for covariate dependent censoring. While covariate dependent censoring might be accommodated with a conditional Kaplan Meier estimator, this would require the selection of a smoothing parameter, which the rank estimator aimed to avoid, as well as be very numerically unstable.

Here we propose an alternative approach to accommodate random, covariate dependent censoring, when estimating the transformation function. We note that we can assume β_0 is known, as we have already provided an estimator which converges as the parametric rate, and here we let $z_i = x_i'\beta_0$. We also note that the transformation function is only identified up to location and scale, so normalizations need to be adopted. As a scale normalization, we set the first component of $\beta_0 = 1$ as before. Here we adopt the usual location normalization by assuming some point y_0, which we set w.l.o.g. to 0, satisfies $T(0) = 0$. We propose an estimator for $T(y_1)$ for some point y_1 in the context of left censoring. To do so, we define the following variables:

$$d_{1ij} = I[y_{0i} \geq y_1] - I[y_{1j} \geq y_1] = d_i I[v_i \geq y_1] - I[v_j \geq 0] \tag{3.6}$$

$$d_{2ij} = I[y_{1i} \geq y_1] - I[y_{0j} \geq 0] = I[v_i \geq y_1] - d_j I[v_j \geq 0] \tag{3.7}$$

To accommodate covariate dependent censoring, we propose the following estimator:

$$\hat{\kappa}_1 = \arg\max_{\kappa} \sum_{i \neq j} (I[d_{1ij} = 1] - I[d_{2ij} = -1]) I[z_i - z_j \geq \kappa] \tag{3.8}$$

The following theorem characterizes the limiting distribution of these two rank estimators of the transformation function:

Theorem 3.2

3.3 Heteroskedastic Models

One of the assumptions that the estimation procedures introduced in this paper have been based on is that the disturbance term ϵ_i be distributed independently of the covariates x_i. This assumption may be overly restrictive in the sense that it rules out any form of conditional heteroskedasticity. In this section we relax the independence assumption by assuming only one of the quantiles of ϵ_i, say the median, is independent of the covariates. Khan(2000) proposed a two step rank estimator for a heteroskedastic transformation model, but did not allow for random censoring. To permit random, covariate dependent censoring, we now make the assumption that the random variables c_i, ϵ_i are statistically independent given x_i.

We illustrate here identification for the univariate censoring case. Similar arguments can be used to attain point identification results for the double censoring case.

Point identification is characterized by the following lemma, whose proof is left to the appendix:

Lemma 3.1 *Define the set* $\mathcal{X}$ *such that*

$$\mathcal{X} = \{x : \Pr(T(c) - x\beta \leq 0|x) = 1\}$$

Assume further that $\Pr_x(\mathcal{X}) > 0$*. Moreover, the random variable* c *is such that* $\epsilon \perp c|x$*. Finally, define the random variables* $y_{1i} = v_i$ *and* $y_{0i} = d_i v_i + (1 - d_i) \cdot -\infty$*. Then we have that*

$$Med(T(y_0)|x) = Med(T(y)|x) = Med(T(y_1)|x) = x\beta$$

if and only if $x \in \mathcal{X}$*.*

The above identification result, along with the invariance of medians, suggests an (infeasible) rank estimator based on the conditional medians of y_{0i} and y_{1i}. Letting $m_0(x_i), m_1(x_i)$ denote these conditional median functions, we would estimate β_0 by maximizing the function

$$Q(\beta) = \frac{1}{n(n-1)} \sum_{i \neq j} I[m_1(x_i) \geq m_0(x_j)] I[x_i'\beta \geq x_j'\beta] \tag{3.9}$$

To construct a feasible estimation procedure, we replace the unknown median functions in the above estimator with their nonparametric estimators. To construct these estimators, we adopt the local polynomial approach introduced in Chaudhuri(1991). For a detailed description of the estimator, see Chaudhuri(1991). Here, we simply let $\hat{m}_0^{\delta_n,p}(x_i)$, $\hat{m}_1^{\delta_n,p}(x_i)$ denote the local polynomial estimators where the superscripts denote the bandwidth sequence (δ_n), and order of polynomial (p) used. Conditions on δ_n and p are stated in the theorem below characterizing the limiting distribution of our estimator of β_0. To avoid the technical difficulty of dealing with a smoothing parameter inside an indicator function, we define our heteroskedasticity robust estimator of β_0, denoted here as $\hat{\beta}_{ht}$ as follows:

$$\hat{\beta}_{ht} = \arg\max_{\beta\in\mathcal{B}} \frac{1}{n(n-1)} \sum_{i\neq j} K_{h_n}(\hat{m}_1^{\delta_n,p}(x_i) - \hat{m}_0^{\delta_n,p}(x_j)) I[x_i'\beta \geq x_j'\beta] \tag{3.10}$$

where $K_{h_n}(\cdot) \equiv K(\cdot/h_n)/h_n$, with $K(\cdot)$ denoting a smooth approximating function to an indicator function (i.e. a cumulative distribution function), and h_n denotes a sequence of positive constants, converging to 0, such that in the limit we have an indicator function. This smoothing technique was introduced in the seminal work of Horowitz(1992).

We next state the limiting distribution theory for $\hat{\beta}_{ht}$. Our limiting distribution theory for this estimator is based on the following assumptions:

Assumptions on the Median Functions

Q1 For any value $x^{(d)}$ in the support of $x_i^{(d)}$, $m_j(\cdot) \quad j = 0, 1$ is k times differentiable in $x_i^{(c)}$. Letting $\nabla_k m_j(x^{(c)}, x^{(d)})$ denote the vector of k^{th} order derivatives of $m_j(\cdot)$ in $x_i^{(c)}$, we assume the following Lipschitz condition:

$$\|\nabla_k m_j(x_1^{(c)}, x^{(d)}) - \nabla_k m_j(x_2^{(c)}, x^{(d)})\| \leq \mathcal{K}\|x_1^{(c)} - x_2^{(c)}\|^\gamma$$

for all values $x_1^{(c)}, x_2^{(c)}$ in the support of $x_i^{(c)}$, where $\|\cdot\|$ denotes the Euclidean norm, $\gamma \in (0, 1]$, and $\mathcal{K}$ is some positive constant. In the theorems to follow, we will let $p = k + \gamma$ denote the *order of smoothness* of the quantile function.

Assumptions on the Trimming Function

T The trimming function $\tau : \Re^d \mapsto \Re^+$ is continuous, bounded, and bounded away from zero on its support, denoted by $\mathcal{X}_t$, a compact subset of $\Re^d$.

Assumptions on the Regressors

B1 The sequence of $d + 2$ dimensional vectors (v_i, d_i, x_i) are independent and identically distributed.

B2 The regressor vector x_i has support which is a subset of $\Re^d$.

We order the components of x_i so it can be written as $x_i = ({x_i}^{(d)}, {x_i}^{(c)})'$. Let d_c denote $\dim({x_i}^{(c)})$. Assume that $1 \leq d_c \leq d$ and that the support ${x_i}^{(c)}$ is a convex subset of

$\Re^{d_c}$ and has nonempty interior. Assume that the support of ${x_i}^{(d)}$ is a finite number of points lying in $\Re^{d-d_c}$. We will let $f_X(x)$ denote the product of the conditional (Lebesgue) density of $x_i^{(c)}$ given $x_i^{(d)}$ (denoted by $f_{X^{(c)}|X^{(d)}=x^{(d)}}(x^{(c)})$) and the marginal probability mass function of $X^{(d)}$ (denoted by $f_{X^{(d)}}(x^{(d)})$).

B3 $f_{X^{(c)}|X^{(d)}}(x^{(c)})$ is continuous and bounded on the support of $x_i^{(c)}$.

B4 Assume that $\mathcal{X}_t = \mathcal{X}_{t(d-1)} \times \mathcal{X}_{td}$ where $\mathcal{X}_{t(d-1)}$ and $\mathcal{X}_{td}$ are compact subsets with non-empty interiors of the supports of the first $d-1$ components, and the d^{th} component of x_i, respectively. For each $x \in \mathcal{X}_t$, denote its first $d-1$ components by $x_{(d-1)}$. $\mathcal{X}_t$ will be assumed to have the following properties:

B4.1 $\mathcal{X}_t$ is not contained in any proper linear subspace of $\Re^d$.

B4.2 $f_X(x) \geq \epsilon_0 > 0 \quad \forall x \in \mathcal{X}_t$, for some constant ϵ_0.

Assumptions on the Median Residual Terms

D1 Let $u_{1i} = y_{1i} - m_1(x_i)$; in a neighborhood of 0, u_{1i} has a conditional (Lebesgue) density, denoted by $f_{u1|X_i=x}(\cdot)$ which is continuous, and bounded away from 0 and infinity for all values of $x \in \mathcal{X}_t$. As a function of x, $f_{u|X_i=x}$ is Lipschitz continuous for all values of u_{1i} in a neighborhood of 0. Define u_{0i} analogously and assume it has analogous properties.

Furthermore, we require conditions on the smoothness of the median functions. Let

$$\begin{aligned}\tau_{q1}(x,\theta) &= \int I[x \in \mathcal{X}]I[u \in \mathcal{X}]\tau(x)I[m_1(x) \geq m_0(u)]I[x'\beta(\theta) > u'\beta(\theta)]dF_X(u) \\ &+ \int I[x \in \mathcal{X}]I[u \in \mathcal{X}]\tau_q(u)I[m_1(u) \geq m_0(x)]I[u'\beta(\theta) > x'\beta(\theta)]dF_X(u)\end{aligned}$$

and let

$$\tau_{q2}(x,\theta) = \int I[x \in \mathcal{X}]I[u \in \mathcal{X}]I[x'\beta(\theta) > u'\beta(\theta)]dF_X(u)$$

let $\mathcal{N}$ be a neighborhood of the $d-1$ dimensional vector θ_0. Then we impose the following additional assumptions:

E1 For each x in the support of x_i, $\tau_{q1}(x,\cdot)$ is differentiable of order 2, with Lipschitz continuous second derivative on $\mathcal{N}$.

E2 $E[\nabla_2 \tau_{q1}(\cdot, \theta_0)]$ is negative definite

E3 For each x in the support of x_i, $\tau_{q2}(x, \cdot)$ is continuously differentiable on $\mathcal{N}$.

E4 $E[\|\nabla_1 \tau_{q2}(\cdot, \theta_0)\|^2] < \infty$

Finally, we impose conditions on the second stage smoothed indicator function and bandwidth:

SI1 The function $K(\cdot)$ is positive, strictly increasing, twice differentiable with bounded first and second derivatives, and satisfies the following:

SI1.1 $\lim_{x\to+\infty} K(x) = 1$, $\lim_{x\to-\infty} K(x) = 0$

SI1.2 $\int_{-\infty}^{\infty} K'(x)dx = 1$

SI2 $h_n > 0$ and $h_n \to 0$.

The following theorem establishes that these additional assumptions, along with a stronger smoothness condition on the quantile function and further restrictions on the bandwidth sequence, are sufficient for root-n consistency and asymptotic normality of the proposed estimator:

Theorem 3.3 *Assume that $p > 3d_c/2$, and that in the first stage, k is set to int(p) and the bandwidth sequences satisfy $\sqrt{n}\delta_n^p \to 0$, $\log n\sqrt{n^{-1}\delta_n^{-3d_c}} \to 0$ and*

$$\sqrt{n}h_n^{-2}(\delta_n^{2p} + \log n \cdot n^{-1}\delta_n^{-d_c}) \to 0$$

. Define

$$\begin{aligned}\delta(y_{1i}, y_{0i}, x_i) &= \tau(x_i) f^{-1}_{u_{1i}|x_i}(0) f'_{m_0}(m_1(x_i))(I[y_{1i} \le m_1(x_i)] - 0.5)\nabla_1 \tau_{q2}(x_i, \theta_0) \\ &+ \tau(x_i) f^{-1}_{u_{0i}|x_i}(0) f'_{m_1}(m_0(x_i))(I[y_{0i} \le m_0(x_i)] - 0.5)\nabla_1 \tau_{q2}(x_i, \theta_0)\end{aligned}$$

where $f'_{m1}(\cdot), f'_{m0}(\cdot)$ denote derivatives of density functions of the median functions; then under Assumptions A,B,Q,T,E,SI

$$\sqrt{n}(\hat{\theta} - \theta_0) \Rightarrow N(0, V_q^{-1}\Delta_q V_q^{-1}) \tag{3.11}$$

where $\Delta_q = E[\delta_q(y_i, x_i)\delta_q(y_i, x_i)']$ and $V_q = \frac{1}{2}E[\nabla_2 \tau_{q1}(x_i, \theta_0)]$.

3.4 Panel Data (incomplete)

We note here how the new rank estimator can be modified to accommodate fixed effects in longitudinal panel data sets. Transformation models with fixed effects have been considered in Lee(1997), Abrevaya(2000,2001). None of these were able to incorporate covariate dependent random censoring. The transformation model with fixed effects is usually expressed as:

$$T(y_i^{(t)}) = \alpha_i + x_i^{(t)'}\beta_0 + \epsilon_i^{(t)} \quad i = 1, 2, ...N \quad t = 1, 2, ...T \tag{3.12}$$

where here α_i denotes the individual effect; following usual panel data asymptotics, we assume N is arbitrarily large, and T is fixed at a small number; w.l.o.g., we set $T = 2$. To estimate β_0 in an uncensored model, Abrevaya proposed the "leap frog" estimator, which maximized the objective function:

$$LF(\beta) = \frac{1}{n}\sum_{i=1}^{n} I[y_i^{(1)} < y_j^{(1)}]I[y_i^{(2)} > y_j^{(2)}]I[\Delta x_i'\beta > \Delta x_j'\beta] \tag{3.13}$$

where superscripts denote time periods, and Δ denotes the time difference operator. Now we assume the econometrician does not observe y_{it}, but instead the pair (v_{it}, d_{it}) where $v_{it} = \max(y_{it}, c_{it})$ and d_{it} is a censoring indicator for person i in period t.

Define y_{0i}, y_{1i} as before, and letting superscripts denote time periods, we propose maximizing the following objective function to accommodate random censoring in the panel data model:

$$\begin{aligned} LFCR(\beta) &= \frac{1}{n(n-1)}\sum_{i\neq j} I[y_{1i}^{(1)} < y_{0j}^{(1)}]I[y_{0i}^{(2)} > y_{1j}^{(2)}]I[\Delta x_i'\beta > \Delta x_j'\beta] \qquad (3.14) \\ &= \frac{1}{n(n-1)}\sum_{i\neq j} \Big((1-d_j^{(1)}) + d_j^{(1)} I[v_i^{(1)} \leq v_j^{(1)}]\Big)\cdot d_i^{(2)} I[v_i^{(2)} \geq v_j^{(2)}]I[\Delta x_i'\beta > \Delta x_j'\beta] \end{aligned}$$

The following theorem characterizes the limiting distribution theory of the panel data estimator:

Theorem 3.4

4 Monte Carlo Results

In this section we explore the finite sample properties of the new estimators introduced in this paper by reporting results obtained from a small scale simulation study.

We first turn attention to the basic CRMRC. Our base design involves two regressors, and an additive of error term, which we express, in the absence of censoring as:

$$T(y_i) = \alpha_0 + x_{1i}\beta_0 + x_{2i} + \epsilon_i$$

where x_{1i}, x_{2i} are distributed as a chi-squared with one degree of freedom, and standard normal, respectively; α_0, β_0 were each set to 1. We considered 2 functional forms for $T(\cdot)$, error distribution pairs

1. $T^{-1}(x) = x$; $\epsilon_i \sim$mixture of two normals, centered around -1,2, respectively.
2. $T^{-1}(x) = x^3$; $\epsilon_i \sim$standard normal.

We simulated three types of censoring: 1) covariate dependent left censoring, where the censoring variable was distributed as $0.5 * z_i + x_{1i} - x_{2i} + 1$; 2) double covariate independent censoring, where the left censoring variable was distributed $0.5 * z_i$ were z_i was standard normal and the right censoring variable is distributed as the left censoring variable plus a chi-squared random variable +1.5; 3) double covariate dependent censoring where the left censoring variable was the same as in 1) and the relationship between the two censoring variables was the same as in 2).

Tables I-IV We report results for 3 estimators: 1)CRMRC 2) the MRC 3) the MRE with $M(\cdot)$ set to the identity function, For each estimator and each design the summary statistics mean bias, median bias. root mean squared error (RMSE) and median absolute deviation (MAD) are reported for 100,200, and 400 observations, with 401 replications. As there is only one parameter to compute, each estimator was evaluated by means of a grid search of 500 evenly spaced points over the interval [-2,2]. The simulation results are in accordance with the theory. For covariate dependent left censoring, the results clearly establish the benefits of the CRMRC. It performs quite well with bias and RMSE values shrinking at the parametric rate. In complete contrast, the MRC and MRE perform very poorly for both functional forms, with RMSE values in most cases not reducing, and sometimes even increasing with the sample size.

For double covariate independent censoring, all estimators have RMSE's shrinking at the parametric rate, but the efficiency gains of the CRMRC are very apparent for both functional form error distribution pairs. For covariate dependent censoring, the results are similar to the one sided censoring case- only the CRMRC exhibits root-n consistency and the others are clearly inconsistent.

Tables V-VI report results for panel data models. Here the regressors in the first period were defined as above, and in the second period, they were defined as the average of the regressor values in the first period and regressor values from an independent draw from the same distribution. The fixed effects were set as a linear combination of all regressor values in both periods plus a standard normal. The error terms in each period were i.i.d standard normal, and we considered a cubic transformation. For covariate independent censoring, the censoring variable was set to $0.5 * z_i$ in each period, where z_i again represents a standard normal distribution. For covariate dependent censoring, we set the censoring variable in each period to be the same (stochastic) function of the regressors in that time period as was used for the left censoring cross-sectional designs. Results are reported for 2 estimators: the CRMRC, and the Leap-frog estimator in Abrevaya(1999)(referred to here as LF) , noting that the latter may be theoretically inconsistent in both (covariate dependent and covariate independent) cases.

The results indicate that the CRMRC performs very well in both designs, the RMSE shrinking at the parametric rate. In contrast LF performs very poorly for the covariate dependent censoring design, with biases and RMSE values staying large for all sample sizes. LF performs better at the covariate independent design, but its bias stays at 15% as the sample size increases from 200 to 400, suggesting consistency is suspect here as well.

We next turn attention to estimation of the function $T(\cdot)$. We consider the same base design, with the same two functional forms, but now only consider left censoring designs, one with covariate dependence and the other with covariate independence. We report results for two estimators: the CRMRC and Chen(2002) rank estimator, referred to here as CRNK. Tables VI-X report mean bias and RMSE for both estimators for a grid of 11 values of γ_0 at 100 and 400 observations. Again, the simulation results agree with the theory. Both estimators perform well for the covariate independent censoring case, with the CRNK doing slightly better at 400 observations, but only the CRMRC performs adequately in the covariate dependent censoring case.

In summary, the results from our simulation indicate that the CRMRC estimators introduced in this paper perform adequately well in finite samples, so it can be applied in empirical settings, which we turn to in the following section.

5 Empirical Illustrations

In this section we further explore the finite sample properties of the new estimators proposed in this paper by ways of two empirical illustrations.

5.1 Stanford Heart Transplant Data

We consider the well studied Stanford heart transplant data set published in Miller and Halpern (1982), of which an earlier subset of these data is available in the text by Kalbfleisch and Prentice (1980. Summarized in this data set are the survival times of 184 patients who received heart transplants at the Stanford University Medical Center, as well as an indicator variable which equals one if the patient was dead (uncensored) at the time the data were collected, the age of the patient (in years) at the time of the transplant, a tissue-mismatch score variable, and a waiting time variable. We estimate the following model of the survival times,

$$T(v_i) = \min\{\alpha_0 + \beta_0 x_i + \gamma_0 z_i + \rho_0 w_i + \varepsilon_i, c_i\}, \tag{5.1}$$

where the dependent variable v_i is the observed survival time (in days), x_i is age of patient i, z_i is the tissue mismatch score, and w_i is the waiting time variable.

For this model, covariate dependent censoring seems quite plausible. Larger censoring times correspond to earlier transplants; if transplants for younger or older patients were not typically performed in the earlier years, this would induce a dependence between censoring and the covariate age.

We drop all the incomplete observations to obtain a total of 69 patients that have complete records for the mismatch and waiting time variables. We standardize the coefficient on age to one and provide estimates using the CRMRC and MRC. Table 1 summarizes our results. In addition to providing point estimates, we estimate standard errors by the mean absolute deviation of the boostrapped c.d.f, divided by 0.67, as was done in Honoré, Khan and Powell(2002).

5.2 Marriage length in the CPS

We further illustrate our estimator by studying the effects of age at first marriage and other covariates on first marriage length. For couples who are still married for the first time at

Table 1: **Stanford Heart Data Estimation Results**

Regressor	Parameter	Median Absolute Deviation/.67
CRMRC		
Waiting till Transplant	-1.78	1.46
Mismatch	-.74	.86
MRC		
Waiting till Transplant	-.52	1.07
Mismatch	-.66	.76

Table 2: **Descriptive Statistics for CPS Marriage and Fertility Data**

Variable	Mean	Standard Deviation	Min	Max
Age at First Marriage	22.5	5.9	14	78.5
Age	64.8	10.11	50	99
Race	.85	.31	0	1
Educ	12.2	3.1	1	19

the date of the interview, their marriage length variable is right censored. Moreover, it can be argued that divorce is correlated with age at first marriage which makes the censoring point (time of divorce) correlated with age. We draw a random sample of 1000 observations from the 1985 marriage and fertility June CPS where we restrict our choice to individuals who have been married at least once and who are 50 years of age or older at the time of the interview. Table 2 provides descriptive statistics of the data. Moreover, the average first marriage length for divorcees is 33 years with a standard deviation of 16 years. The amount of censoring is 52% which means that almost half of our sample of ever married couples have been divorced at least once. Using age at first marriage and race as regressors, and standardizing the coefficient on age at first marriage to one, we compute the CRMRC and MRC estimators. Race coefficient values of 28.12 and 35.12 were obtained using CRMRC and MRC, respectively. [Bootstrapped confidence bands to come]

6 Conclusions

This paper introduced new estimation procedures for several censored transformation models. With the exception of the heteroskedasticity-robust variation, the new procedures have the attractive properties of requiring no smoothing parameters. All estimators were robust to censoring that depends on the covariates. The estimators are shown to converge at the parametric rate with asymptotic normal distributions. A simulation study indicated it performed well in finite samples, and also illustrated how erroneous existing rank estimators can be if the censoring variable depends on covariates. Two empirical illustrations applied the new estimator to a Stanford heart transplant data set and a data set involving marriage duration. In both cases, the new estimators gave different results than an estimator which did not permit covariate dependent censoring and/or required known transformation functions.

The results in this paper suggest areas for future research. For one, it would be useful to formally establish identification for the transformation function, and the coefficients in the panel data model. Also, it would be useful to explore under what conditions identification can be achieved if the censoring variable is not distributed independently of the error term.

References

[1] Abrevaya, J. (1999), "Leapfrog estimation of a fixed-effects model with unknown transformation of the dependent variable." *Journal of Econometrics*, 93, 203-228.

[2] Abrevaya, J. (1999), "Large computation of the maximum rank correlation estimator." *Economics letters*, 62, 279-288.

[3] Buckley, J. and I. James (1979), "Linear Regression with Censored Data," *Biometrika*, 66, 429-436.

[4] Cavanagh, C. and R.P. Sherman (1998), "Rank Estimators for Monotonic Index Models", *Journal of Econometrics*, 84, 351-381

[5] Chaudhuri, P. (1991a), "Global Nonparametric Estimation of Conditional Quantiles and their Derivatives", *Journal of Multivariate Analysis*, 39, 246-269

[6] Chaudhuri, P. (1991b), "Nonparametric Quantile Regression", *Annals of Statistics*, 19, 760-777

[7] Chaudhuri, P., K. Doksum, and A. Samarov (1997), "On Average Derivative Quantile Regression", *Annals of Statistics*, 25, 715-744

[8] Cox D.R. (1972), "Regression Models and Life Tables," *Journal of the Royal Statistical Society Series B*, 34,187-220.

[9] Cox D.R. (1975), "Partial Likelihood" *Biometrika*, 62, 269-276

[10] Cuzick, J. (1988), "Rank Regression", *Annals of Statistics*, 16, 1369-1389.

[11] Dabrowska, D. (1992), "Nonparametric Quantile Regression with Censored Data", *Sankhya, Ser. A*, 54, 252-259

[12] Fan, J. and I. Gijbels (1996), *Local Polynomial Modelling and its Applications'*, New York: Chapman and Hall.

[13] Gorgens, T. and J.L. Horowitz (1999), "Semiparametric Estimation of a Censored Regression Model with an Unknown Transformation of the Dependent Variable", *Journal of Econometrics*, 90, 155-191

[14] Han, A. (1987) "Non Parametric Analysis of a Generalized Regression Model", *Journal of Econometrics*, 35, 303-316

[15] Honoré, B.E., Khan, S. and J.L. Powell (2002) "Quantile Regression under Random Censoring", *Journal of Econometrics*, forthcoming.

[16] Horowitz, J.L. (1996), "Semiparametric Estimation of a Regression Model with an Unknown Transformation of the Dependent Variable", *Econometrica*, 64, 103-137

[17] Kalbfleisch, J.D. and R.L. Prentice (1980), *The Statistical Analysis of Failure Time Data*. New York: Wiley.

[18] Kaplan, E.L. and P. Meier (1958), "Nonparametric Estimation from Incomplete Data," *Journal of the American Statistical Association*, 53, 457-481.

[19] Koul, H., V. Susarla, and J. Van Ryzin (1981), "Regression Analysis with Randomly Right Censored Data," *Annals of Statistics*, 9, 1276-1288.

[20] Lai, T.L. and Z. Ying (1991), "Rank Regression Methods for Left Truncated and Right Censored Data," *Annals of Statistics*, 19, 531-554.

[21] Lin, J.S. and L.J. Wei (1992), "Linear Regression Analysis for Multivariate Failure Time Observations," *Journal of the American Statistical Association*, 87, 1091-1097.

[22] Miller, R. and J. Halpern (1982), "Regression with Censored Data," *Biometrika*, 69, 521-531.

[23] Newey, W.K. and D. McFadden (1994) "Estimation and Hypothesis Testing in Large Samples", in Engle, R.F. and D. McFadden (eds.) , *Handbook of Econometrics, Vol. 4*, Amsterdam: North-Holland.

[24] Pakes, A. and D. Pollard (1989), "Simulation and the Asymptotics of Optimization Estimators", *Econometrica*, 57, 1027-1057

[25] Powell, J.L. (1994) "Estimation of Semiparametric Models", in Engle, R.F. and D. McFadden (eds.), *Handbook of Econometrics, Vol. 4*, Amsterdam: North-Holland.

[26] Powell, J.L., J.H. Stock, and T.M. Stoker (1989) "Semiparametric Estimation of Index Coefficients", *Econometrica*, 57, 1404-1430.

[27] Ridder, G. (1990) "The Non-parametric Identification of Generalized Accelerated Failure-time Models", *Review of Economic Studies*, 57, 167-182

[28] Ritov, Y. (1990), "Estimation in a Linear Regression Model with Censored Data," *Annals of Statistics*, 18, 303-328.

[29] Serfling, R.J. (1980) *Approximation Theorems of Mathematical Statistics*, New York: Wiley.

[30] Sherman, R.P. (1993), "The Limiting Distribution of the Maximum Rank Correlation Estimator", *Econometrica*, 61, 123-137

[31] Sherman, R.P. (1994a), "U-Processes in the Analysis of a Generalized Semiparametric Regression Estimator", *Econometric Theory*, 10, 372-395

[32] Sherman, R.P. (1994b), "Maximal Inequalities for Degenerate U-Processes with Applications to Optimization Estimators ", *Annals of Statistics*, 22, 439-459

[33] Tsiatis, A.A. (1990), "Estimating Regression Parameters Using Linear Rank Tests for Censored Data," *Annals of Statistics*, 18, 354-372.

[34] van den Berg, G.J. (2001), "Duration Models: Specification, Identification and Multiple Durations", in Heckman, J.J. and E. Leamer, eds., *Handbook of Econometrics, Vol. 5*, Amsterdam: North-Holland.

[35] Wei, L.J., Z. Ying, and D.Y. Lin (1990), "Linear Regression Analysis of Censored Survival Data Based on Rank Tests." *Biometrika*, 19, 417-442.

[36] Wang, J.-G. (1987), "A Note on the Uniform Consistency of the Kaplan-Meier Estimator" *Annals of Statistics*, 15, 1313-1316

[37] Yang, S. (1999), "Censored Median Regression Using Weighted Empirical Survival and Hazard Functions", *Journal of the American Statistical Association*, 94, 137-145

[38] Ye, J. and N. Duan (1997) "Nonparametric $n^{-1/2}$-consistent Estimation for the General Transformation Models", *Annals of Statistics*, 25, 2682-2717

[39] Ying, Z., S.H. Jung, and L.J. Wei (1995), "Survival Analysis with Median Regression Models," *Journal of the American Statistical Association*, 90, 178-184

A Appendix

A.1 Proof of Lemma 2.1

Recall we observe the vector $\mathbf{z}_i \equiv (v_i, d_i, x_i')'$ which we assume to be generated from the model:

$$T(v_i) = \max(x_i'\beta_0 + \epsilon_i, c_i)$$

$$d_i = I[x_i'\beta_0 + \epsilon_i \leq c_i]$$

To prove the lemma, we define two random variables which are functions of $\mathbf{z}_i$ and hence are observable. We define:

$$\begin{aligned} y_{1i} &= v_i \\ y_{0i} &= I[d_i = 1]v_i + I[d_i = 0] \cdot -\infty \end{aligned}$$

Note that establishing the conclusion of the lemma is equivalent to establishing that

$$P(y_{1i} \geq y_{0j}|x_i, x_j) \geq P(y_{1j} \geq y_{0i}|x_i, x_j) \tag{A.1}$$

whenever $z_i \equiv x_i'\beta_0 \geq z_j \equiv x_j'\beta_0$. To do so, we can decompose the left hand side of the above equation as follows:

$$P(y_{1i} \geq y_{0j}, c_i \geq c_j|x_i, x_j) + P(y_{1i} \geq y_{0j}, c_i \leq c_j|x_i, x_j) \tag{A.2}$$

and similarly decompose the right hand side of (A.1). We first compare

$$P(y_{1i} \geq y_{0j}, c_i \geq c_j | x_i, x_j) \tag{A.3}$$

to

$$P(y_{1j} \geq y_{0i}, c_i \geq c_j | x_i, x_j) \tag{A.4}$$

Focusing initially on (A.3), we decompose the event into the disjoint union of three cases: $(d_j = 0), (d_i = 1, d_j = 1), (d_i = 0, d_j = 1)$. Conditioning on c_i, c_j, and suppressing the event $c_i \geq c_j$ and the fact we are conditioning on c_i, c_j, x_i, x_j, we have, by using the monotonicity of the transformation function, (A.3) is:

$$\begin{aligned} & P(\epsilon_j \leq c_j - z_j) && \text{(A.5)} \\ + \quad & P(\epsilon_i - \epsilon_j \geq z_j - z_i, \epsilon_i \geq c_i - z_i, \epsilon_j \geq c_j - z_j) && \text{(A.6)} \\ + \quad & P(c_i \geq z_j + \epsilon_j, \epsilon_i \leq c_i - z_i, \epsilon_j \geq c_j - z_j) && \text{(A.7)} \end{aligned}$$

We denote (A.5) by $F(c_j - z_j)$ where $F(\cdot)$ denotes the c.d.f. of ϵ_j. We next decompose (A.6) as

$$P(\epsilon_i - \epsilon_j \geq z_j - z_i, \epsilon_i \geq c_i - z_i, \epsilon_j \geq c_j - z_j, \epsilon_j + z_j - z_i \geq c_i - z_i) + \tag{A.8}$$

$$P(\epsilon_i - \epsilon_j \geq z_j - z_i, \epsilon_i \geq c_i - z_i, \epsilon_j \geq c_j - z_j, \epsilon_j + z_j - z_i < c_i - z_i) \tag{A.9}$$

Noting that $c_i \geq c_j$, we can express the sum of these two terms as

$$P(\epsilon_i \geq c_i - z_i, \epsilon_j \geq c_j - z_j, \epsilon_j \leq c_i - z_j) + P(\epsilon_i \geq \epsilon_j + z_j - z_i, \epsilon_j \geq c_i - z_j) \tag{A.10}$$

which we can express as:

$$(F(c_i - z_j) - F(c_j - z_j))\,(1 - F(c_i - z_i)) + (1 - F(c_i - z_j)) - \int_{c_i - z_j}^{\infty} F(e + z_j - z_i)dF(e) \tag{A.11}$$

Similarly we can express (A.7) as

$$(F(c_i - z_j) - F(c_j - z_j)) \cdot F(c_i - z_i) \tag{A.12}$$

Therefore by summing the three pieces in (A.5),(A.6), (A.7), and averaging over the censoring variables, we get (A.3) can expressed as:

$$\int \left\{ 1 - \int_{c_i - z_j}^{\infty} F(e + z_j - z_i)dF(e) \right\} I[c_i \geq c_j] dF_{c|x}(c_i|x_i) dF_{c|x}(c_j|x_j) \tag{A.13}$$

where $F_{c|x}(\cdot)$ denotes the conditional c.d.f. of the censoring variable. We turn attention now to (A.4), which can be decomposed into two disjoint cases $(d_i = 0), (d_i = 1, d_j = 1)$ since the case $(d_i = 1, d_j = 0)$ cannot occur when $c_i \geq c_j$. Using analogous arguments we can express (A.4) as

$$\int \left\{ 1 - \int_{c_i - z_i}^{\infty} F(e + z_i - z_j) dF(e) \right\} I[c_i \geq c_j] dF_{c|x}(c_i|x_i) dF_{c|x}(c_j|x_j) \tag{A.14}$$

Thus the difference between (A.3) and (A.4) is

$$\int \left\{ \int_{c_i - z_i}^{\infty} F(e + z_i - z_j) dF(e) - \int_{c_i - z_j}^{\infty} F(e + z_j - z_i) dF(e) \right\} I[c_i \geq c_j] dF_{c|x}(c_i|x_i) dF_{c|x}(c_j|x_j) \tag{A.15}$$

we note the above expression is non-negative whenever $z_i \geq z_j$ as the differences between the two terms, each involving non-negative integrands, is the area of integration, which is larger for the first term whenever $z_i \geq z_j$, and the difference between $F(e + z_i - z_j)$ and $F(e + z_j - z_i)$, which is also positive whenever $z_i \geq z_j$. This shows (A.1) for the case when $c_i \geq c_j$. For the case $c_i < c_j$ we proceed similarly and find that the difference between the left hand side and right hand side in (A.1) can be expressed as

$$\int \left\{ \int_{c_j - z_i}^{\infty} F(e + z_i - z_j) dF(e) - \int_{c_i - z_j} F(e + z_j - z_i) dF(e) \right\} I[c_i < c_j] dF_{c|x}(c_i|x_i) dF_{c|x}(c_j|x_j) \tag{A.16}$$

Again, the above expression is non-negative whenever $z_i \geq z_j$, and this is also for two reasons. The area of integration as well as the integrand is larger for the first term in the above difference whenever $z_i \geq z_j$. Since we have shown (A.1) to be true for both cases $c_i \geq c_j, c_i < c_j$, this completes the proof. ■

A.2 Proof of Theorem 2.1

To show consistency it suffices to show 4 conditions (see e.g. Newey and MacFadden(1994), Theorem 2.1.): compactness, uniform convergence, continuity, identification.

We first turn attention to the proof of identification. Let $Q(\beta)$ denote the limiting objective function. We need to show that this is uniquely maximized at β_0. Let $b \neq \beta_0$. We can express $Q(\beta_0) - Q(b)$ as

$$E_X[P(y_{1i} \geq y_{0j}|x_i, x_j) \left(I[x_i'\beta_0 \geq x_j'\beta_0] - I[x_i'\beta \geq x_j'\beta] \right)] \tag{A.17}$$

which we can write as

$$E[I[y_{1i} \geq y_{0j}] \left(I[x_i'\beta_0 \geq x_j'\beta_0, x_i'b < x_j'b] - I[x_i'\beta_0 < x_j'\beta_0, x_i'b \geq x_j'b] \right)] \tag{A.18}$$

which we can rearrange to express as:

$$E[(I[y_{1i} \geq y_{0j}] - I[y_{1j} \geq y_{0i}])I[x_i'\beta_0 \geq x_j'\beta_0, x_i'b < x_j'b]] \tag{A.19}$$

By the previous lemma, the above expectation is non-negative, and only equal to 0 when $b = \beta_0$ by Assumption **I3**. This establishes that the limiting objective function is uniquely maximized at β_0, proving identification. Turning attention to the other three items, we note that compactness holds by Assumption, uniform convergence follows from uniform laws of large numbers for U-statistics with bounded kernel functions (see, e.g. Sherman(1994), and continuity follows from the smoothness of the density of $x_i'\beta_0$ which follows from **I3**. This establishes consistency. ■

A.3 Proof of Theorem 2.2

We note that virtually identical arguments as in Sherman(1993) can be used, as the objective functions of the MRC and the CRMRC are very similar. The only component of the proof there that does not immediately carry over to the problem at hand is establishing the Euclidean property of the class of functions in the objective function. For the problem at hand, we consider the class of functions:

$$\mathcal{F} = \{f(\cdot, \cdot, \theta) : \theta \in \Theta\} \tag{A.20}$$

where for each $(\mathbf{z}_1, \mathbf{z}_2) \in S \times S$, $\theta \in \Theta$, we can define

$$f(\mathbf{z}_1, \mathbf{z}_2, \theta) = I[y_{11} > y_{02}]I[x_1'\beta > x_2'\beta] \tag{A.21}$$

where with our notation, recall β is a function of θ. Alternatively, we can define,

$$f(\mathbf{z}_1, \mathbf{z}_2, \theta) = I[y_{01} > y_{12}]I[x_1'\beta > x_2'\beta] = d_1 I[v_1 > v_2]I[x_1'\beta > x_2'\beta] \tag{A.22}$$

It is easier to establish the Euclidean property (with respect to the constant envelope 1) for the above definition of $f(\cdot, \cdot, \theta)$. Note the class of functions

$$f_2(\mathbf{z}_1, \mathbf{z}_2, \theta) = I[v_1 > v_2]I[x_1'\beta > x_2'\beta] \tag{A.23}$$

is Euclidean for envelope 1 from identical subgraph set arguments used in Sherman(1993). The class of functions:

$$f_2(\mathbf{z}_1, \mathbf{z}_2, \theta) = d_1 \tag{A.24}$$

is trivially Euclidean for envelope 1 as it does not depend on θ. The Euclidean property of $f = f_1 \cdot f_2$ follows from Lemma 2.14(ii) in Pakes and Pollard(1989).

A.4 Proof of Lemma 3.1

(only if) Consider the following

$$\begin{aligned}
\Pr(T(y_1) - x\beta \leq 0|x) &= \Pr(y_1 \leq T^{-1}(x\beta)|x) \\
&= \Pr(y_1 \leq T^{-1}(x\beta), d = 1|x) + \Pr(y_1 \leq T^{-1}(x\beta), d = 0|x) \\
&= \Pr(y \leq T^{-1}(x\beta), d = 1|x) + \Pr(c \leq T^{-1}(x\beta), d = 0|x) \\
&= \Pr(\epsilon \leq 0, \epsilon \geq T(c) - x\beta|x) + \Pr(T(c) \leq x\beta, \epsilon \leq T(c) - x\beta|x) \\
&= \Pr(\epsilon \leq 0|x) - \Pr(\epsilon \leq 0, \epsilon \leq T(c) - x\beta|x) + \Pr(T(c) \leq x\beta, \epsilon \leq T(c) - x\beta|x) \\
&= \Pr(\epsilon \leq T(c) - x\beta|x)
\end{aligned}$$

small where the last equality follows from the hypothesis that $x \in \mathcal{X}$.

$$\begin{aligned}
\Pr(T(y_0) = x\beta \leq 0|x) &= \Pr(y_0 \leq T^{-1}(x\beta)|x) \\
&= \Pr(y \leq T^{-1}(x\beta), d = 1|x) + \Pr(d = 0|x) \\
&= \Pr(\epsilon \leq 0, \epsilon \geq T(c) - x\beta|x) + \Pr(\epsilon \leq T(c) - x\beta|x) \\
&= \Pr(\epsilon \leq 0|x) - \Pr(\epsilon \leq 0, \epsilon \leq T(c) - x\beta|x) + \Pr(\epsilon \leq T(c) - x\beta|x) \\
&= \Pr(\epsilon \leq T(c) - x\beta|x)
\end{aligned}$$

where the last equality follows from the hypothesis. As we can see that for $x \in \mathcal{X}$, we have

$$\begin{aligned}
\Pr(T(y_1) - x\beta \leq 0|x) &= \Pr(T(y_1) - x\beta \leq 0|x) \\
&= \Pr(T(y) - x\beta \leq 0|x) \\
&= \frac{1}{2}
\end{aligned}$$

which implies that the medians are the same.

(if) Now we have

$$\begin{aligned}
\Pr(\epsilon \leq 0|x) &= \Pr(T(y_1) - x\beta \leq 0|x) \\
&= \Pr(T(y_1) - x\beta \leq 0, d = 1|x) + \Pr(T(y_1) - x\beta \leq 0, d = 0|x) \\
&= \Pr(T(y_1) - x\beta \leq 0, \epsilon \geq T(c) - x\beta|x) + \Pr(T(y_1) - x\beta \leq 0, \epsilon \leq T(c) - x\beta|x) \\
&= \Pr(T(c) - x\beta \leq \epsilon \leq 0|x) + \Pr(\epsilon \leq T(c) - x\beta \leq 0|x) \\
&= \Pr(\epsilon \leq 0; T(c) - x\beta \leq 0|x) \\
&= \Pr(\epsilon \leq 0) \Pr(T(c) - x\beta \leq 0|x) \\
&\Rightarrow \Pr(T(c) - x_\eta \leq 0|x) = 1
\end{aligned}$$

The last equality follows from the hypothesis that $\epsilon \perp c|x$ which is the maintained assumption. ■

A.5 Proof of Theorem 3.3

The asymptotic properties follow from arguments that are very similar to those used in Khan(2001), so we only provide a sketch of the steps involved. First we expand the kernel function of the estimated median functions around the kernel of the true median functions in (3.10), yielding the sum of the three components

$$\Gamma_n(\beta) \equiv \frac{1}{n(n-1)} \sum_{i \neq j} K_{h_n}(m_{1i} - m_{0j}) I[x_i'\beta \geq x_j'\beta] \tag{A.25}$$

$$H_n(\beta) \equiv \frac{1}{n(n-1)} \sum_{i \neq j} K_{h_n}'(m_{1i} - m_{0j}) h_n^{-1}((\hat{m}_{1i} - m_{1i}) - (\hat{m}_{0j} - m_{0j}) I[x_i'\beta \geq x_j'\beta] \tag{A.26}$$

$$R_n(\beta) \frac{1}{n(n-1)} \sum_{i \neq j} K_{h_n}''(m_{1i}^* - m_{0j}^*) h_n^{-2} (\hat{m}_{1i} - m_{1i} - \hat{m}_{0j} + m_{0j})^2 I[x_i'\beta \geq x_j'\beta] \tag{A.27}$$

where we have adopted the shorthand notation $\hat{m}_{1i}, m_{1i}$ denotes $\hat{m}_1^{\delta_n,p}(x_i), m_1(x_i)$ respectively, and * denotes intermediate values.

First we deal with (A.26). It follows by uniform rates of convergence for median function estimators over compact sets, (see, e.g. Chaudhuri(1991)) where these rates depend on p, δ_n, Assumptions SI1,SI2, and the rates imposed on δ_n, h_n stated in the theorem $R_n(\beta)$ is $o_p(1/n)$ uniformly over β within an $O_p(1/\sqrt{n})$ neighborhood of β_0.

Turning attention to $H_n(\beta)$, with the properties of $K(\cdot)$ in Assumption SI1, we apply the arguments in Lemma A.4 in Khan(2000) that uniformly over β within $o_p(1)$ neighborhoods of β_0, we have

$$H_n(\beta) = (\beta - \beta_0)' \frac{1}{n} \sum_{i=1}^{n} \delta(y_{1i}, y_{0i}, x_i) + o_p(1/n) \tag{A.28}$$

Finally, with regard to $\Gamma_n(\beta)$, we have by the properties of $K(\cdot), h_n$ in Assumption SI1,SI2, using identical arguments as in Lemma A.3 in Khan(2000), that uniformly over β within $o_p(1)$ neighborhoods of β_0, we have

$$\Gamma_n(\beta) = \frac{1}{2}(\beta - \beta_0)' V_q (\beta - \beta_0) + o_p(1/n) \tag{A.29}$$

Combining these three results, the limiting distribution of the estimator follows by applying Lemma A.2 in Khan(2000). ■

TABLE I

Simulation Results for Rank Regression Estimators

One Sided CD Censoring Linear

	β			
	Mean Bias	Med. Bias	RMSE	MAD
100 obs.				
CRMRC	0.0717	-0.0160	0.4282	0.3424
MRC	0.8818	0.9680	0.9049	0.8832
MRE	0.8701	0.9760	0.8969	0.8710
200 obs.				
CRMRC	0.0704	0.0160	0.3433	0.2654
MRC	0.9624	1.0000	0.9664	0.9624
MRE	0.9563	0.9920	0.9613	0.9563
400 obs.				
CRMRC	0.0168	-0.0160	0.2406	0.1843
MRC	0.9905	1.0000	0.9910	0.9905
MRE	0.9879	1.0000	0.9885	0.9879

TABLE II

Simulation Results for Rank Regression Estimators

One Sided CD Censoring Cubic

	β			
	Mean Bias	Med. Bias	RMSE	MAD
100 obs.				
CRMRC	0.0336	0.0080	0.2689	0.2083
MRC	0.7411	0.8280	0.7871	0.7422
MRE	0.4942	0.4760	0.6101	0.5119
200 obs.				
CRMRC	0.0309	0.0040	0.1867	0.1464
MRC	0.7745	0.8640	0.8114	0.7745
MRE	0.4962	0.4720	0.5827	0.5018
400 obs.				
CRMRC	0.0109	0.0040	0.1220	0.0973
MRC	0.8570	0.9200	0.8730	0.8570
MRE	0.5526	0.5240	0.6070	0.5527

TABLE III

Simulation Results for Rank Regression Estimators

Two Sided CI Censoring Cubic

	β			
	Mean Bias	Med. Bias	RMSE	MAD
100 obs.				
CRMRC	0.0443	0.0040	0.3127	0.2452
MRC	0.0688	-0.0400	0.5864	0.4960
MRE	0.0351	-0.0720	0.7114	0.6411
200 obs.				
CRMRC	0.0180	0.0040	0.2074	0.1569
MRC	0.0886	-0.0600	0.5319	0.4417
MRE	0.1036	-0.1120	0.6819	0.6091
400 obs.				
CRMRC	0.0052	0.0000	0.1318	0.1062
MRC	0.0547	-0.0080	0.3972	0.3075
MRE	0.1321	-0.0400	0.6263	0.5428

TABLE IV

Simulation Results for Rank Regression Estimators

Two Sided CD Censoring Cubic

	β			
	Mean Bias	Med. Bias	RMSE	MAD
100 obs.				
CRMRC	0.0917	0.0400	0.3645	0.2782
MRC	0.9880	1.0000	0.9887	0.9880
MRE	0.9895	1.0000	0.9898	0.9895
200 obs.				
CRMRC	0.0944	0.0400	0.2821	0.2097
MRC	0.9984	1.0000	0.9985	0.9984
MRE	0.9999	1.0000	0.9999	0.9999
400 obs.				
CRMRC	0.0812	0.0667	0.1941	0.1507
MRC	0.9999	1.0000	0.9999	0.9999
MRE	1.0000	1.0000	1.0000	1.0000

TABLE V

Simulation Results for Panel Data Estimators

Cubic CI

	β			
	Mean Bias	Med. Bias	RMSE	MAD
100 obs.				
CRMRC	0.0767	0.0133	0.3781	0.2970
LF	0.2237	0.1733	0.4907	0.3817
200 obs.				
CRMRC	-0.0088	-0.0400	0.2483	0.1961
LF	0.1515	0.0667	0.3873	0.2894
400 obs.				
CRMRC	0.0046	-0.0133	0.1865	0.1419
LF	0.1597	0.1200	0.3106	0.2293

TABLE VI

Simulation Results for Panel Data Estimators

Cubic CD

	β			
	Mean Bias	Med. Bias	RMSE	MAD
100 obs.				
CRMRC	0.0664	0.0133	0.3686	0.2878
LF	0.7071	0.8400	0.7790	0.7180
200 obs.				
CRMRC	0.0480	0.0133	0.2700	0.2056
LF	0.7900	0.8933	0.8271	0.7909
400 obs.				
CRMRC	0.0179	-0.0133	0.1666	0.1318
LF	0.8238	0.9200	0.8484	0.8238

TABLE VII

Function Estimation- Linear CI

γ_0:	**-0.950**	**-0.550**	**-0.150**	**0.250**	**0.450**	**0.650**	**0.850**	**1.050**	**1.250**	**1.450**	**1.850**
Mean Bias											
100 obs.											
CRMRC:	-0.039	-0.032	-0.005	-0.010	0.024	0.016	0.019	0.031	0.032	0.046	0.075
CRNK :	-0.016	0.028	0.010	-0.025	0.024	0.019	-0.004	-0.003	0.017	0.020	0.065
400 obs.											
CRMRC:	-0.053	-0.048	-0.030	-0.011	0.013	0.015	0.027	0.031	0.041	0.043	0.041
CRNK :	0.001	-0.010	0.000	-0.007	0.015	0.001	0.001	0.005	-0.001	0.000	-0.006
RMSE											
100 obs.											
CRMRC:	0.266	0.256	0.210	0.161	0.129	0.174	0.228	0.279	0.291	0.311	0.356
CRNK :	0.337	0.272	0.230	0.175	0.150	0.220	0.272	0.317	0.369	0.391	0.424
400 obs.											
CRMRC:	0.147	0.133	0.110	0.071	0.055	0.088	0.103	0.119	0.138	0.148	0.154
CRNK :	0.156	0.139	0.120	0.085	0.062	0.097	0.110	0.130	0.147	0.152	0.174

TABLE VIII

Function Estimation- Linear CD

γ_0:	**-0.950**	**-0.550**	**-0.150**	**0.250**	**0.450**	**0.650**	**0.850**	**1.050**	**1.250**	**1.450**	**1.850**
Mean Bias											
100 obs.											
CRMRC:	-0.322	-0.204	-0.099	-0.043	0.018	0.036	0.052	0.064	0.062	0.072	0.032
CRNK :	-0.819	-0.588	-0.342	-0.185	0.061	0.188	0.261	0.345	0.485	0.576	0.683
400 obs.											
CRMRC:	-0.271	-0.175	-0.109	-0.036	0.019	0.046	0.060	0.072	0.079	0.075	0.072
CRNK :	-0.795	-0.517	-0.307	-0.116	0.042	0.142	0.228	0.308	0.402	0.547	0.675
RMSE											
100 obs.											
CRMRC:	0.544	0.379	0.279	0.168	0.117	0.169	0.231	0.250	0.263	0.287	0.276
CRNK :	1.068	0.844	0.598	0.395	0.272	0.401	0.475	0.569	0.716	0.810	0.857
400 obs.											
CRMRC:	0.320	0.228	0.158	0.088	0.052	0.095	0.116	0.145	0.150	0.162	0.166
CRNK :	0.907	0.622	0.396	0.186	0.094	0.201	0.288	0.378	0.473	0.637	0.770

TABLE IX

Function Estimation- Cubic CI

γ_0:	**-0.983**	**-0.766**	**0.368**	**0.819**	**1.016**	**1.157**	**1.270**	**1.366**	**1.450**	**1.525**	**1.594**
Mean Bias											
100 obs.											
CRMRC:	-0.055	-0.023	0.009	0.025	0.037	0.038	0.051	0.050	0.074	0.071	0.071
CRNK :	-0.017	0.028	0.015	-0.014	0.015	0.002	0.017	0.024	0.043	0.058	0.068
400 obs.											
CRMRC:	-0.056	-0.047	0.016	0.036	0.046	0.039	0.043	0.043	0.039	0.039	0.045
CRNK :	0.007	-0.004	-0.004	0.005	0.005	0.002	0.007	-0.005	-0.003	-0.005	-0.010
RMSE											
100 obs.											
CRMRC:	0.281	0.258	0.185	0.279	0.287	0.306	0.311	0.333	0.351	0.354	0.353
CRNK :	0.338	0.291	0.240	0.324	0.369	0.373	0.386	0.408	0.418	0.426	0.439
400 obs.											
CRMRC:	0.157	0.137	0.090	0.116	0.135	0.143	0.148	0.148	0.145	0.154	0.151
CRNK :	0.155	0.141	0.104	0.133	0.143	0.149	0.157	0.162	0.162	0.167	0.176

TABLE X

Function Estimation- Cubic CD

γ_0:	**-0.983**	**-0.766**	**0.368**	**0.819**	**1.016**	**1.157**	**1.270**	**1.366**	**1.450**	**1.525**	**1.594**
Mean Bias											
100 obs.											
CRMRC:	-0.340	-0.240	0.036	0.076	0.064	0.069	0.067	0.052	0.047	0.034	0.029
CRNK :	-0.830	-0.659	0.193	0.361	0.470	0.534	0.567	0.586	0.643	0.668	0.678
400 obs.											
CRMRC:	-0.279	-0.196	0.045	0.071	0.070	0.073	0.077	0.076	0.080	0.075	0.074
CRNK :	-0.817	-0.578	0.172	0.312	0.387	0.451	0.536	0.570	0.632	0.668	0.713
RMSE											
100 obs.											
CRMRC:	0.552	0.412	0.181	0.257	0.254	0.280	0.283	0.290	0.280	0.274	0.284
CRNK :	1.071	0.916	0.400	0.574	0.702	0.770	0.805	0.803	0.835	0.846	0.844
400 obs.											
CRMRC:	0.327	0.246	0.099	0.142	0.149	0.155	0.158	0.166	0.167	0.163	0.170
CRNK :	0.938	0.688	0.235	0.379	0.464	0.535	0.624	0.663	0.731	0.766	0.804