A TEST FOR MONOTONE COMPARATIVE STATICS

FEDERICO ECHENIQUE AND IVANA KOMUNJER

ABSTRACT. In this paper we design an econometric test for monotone comparative statics (MCS) often found in models with multiple equilibria. Our test exploits the observable implications of the MCS prediction: that the extreme (high and low) conditional quantiles of the dependent variable increase monotonically with the explanatory variable. The main contribution of the paper is to derive a likelihood-ratio test, which to the best of our knowledge, is the first econometric test of MCS proposed in the literature. The test is an asymptotic "chi-bar squared" test for order restrictions on intermediate conditional quantiles. The key features of our approach are: (1) we do not need to estimate the underlying nonparametric model relating the dependent and explanatory variables to the latent disturbances; (2) we make few assumptions on the cardinality, location, or probabilities over equilibria. In particular, one can implement our test without assuming an equilibrium selection rule.

---

1. INTRODUCTION

Comparative statics predictions—or how exogenous variables affect endogenous variables—are important to establish in economic models.[1] Often, the models possess multiple equilibria, and a *monotone comparative statics* (MCS) prediction holds: There is a smallest and a largest equilibrium, and these change monotonically with explanatory variables (see, e.g., Milgrom and Roberts, 1990b; Milgrom and Shannon, 1994; Villas-Boas, 1997). MCS is a feature found in many well-known economic models. Examples are single-person decision models such as models of optimal growth (see, e.g., Barro and Sala-I-Martin, 2003; Ljungqvist and Sargent, 2004) and firms' investment decisions (see, e.g., Hayashi, 1982; Hayashi and Inoue, 1991), oligopoly models à la Berry, Levinsohn, and Pakes (1995), games with network externalities (see, e.g., Ackerberg and Gowrisankaran, 2006), as well as many other models in IO (see Vives, 1999, for survey).

Recent work in econometrics has shown that MCS is a useful property that can aid with model identification. Molinari and Rosen (2008) study the identification power of MCS in an oligopoly pricing game. Lazzati (2012) uses MCS to derive identification results for treatment response models with endogenous social interactions.[2] In both works, the identification analysis is made complicated because of the presence of multiple equilibria (see, e.g., de Paula, 2013, for a survey on multiple equilibria in econometrics of games). Though recognizing the usefulness of MCS in dealing with equilibrium multiplicity, current econometric literature has largely remained silent on the issue of formal tests for MCS.[3] The goal of this paper is to fill this gap.

---

[1]According to Samuelson (1947): "*The usefulness of our theory emerges from the fact that by our analysis we are often able to determine the nature of the changes in our unknown variables resulting from a designated change in one of more parameters.*"

[2]See also Fox and Lazzati (2012) for additional examples.

[3]We note that Athey and Stern (1998) discuss tests for monotone comparative statics, however, only in the context of firms' choice of organizational form. Complementarities play an important

There are two challenges in testing the MCS hypothesis. The first is to obtain testable implications; the second is to construct a formal statistical test and study its properties. In the context of structural models, Echenique and Komunjer (2009) solve the first, but not the second challenge. They obtain testable implications of the MCS property in the form of restrictions on the conditional quantiles of the dependent variable given the explanatory variable.

In this paper, we derive similar restrictions on conditional quantiles in the context of reduced form models with multiple equilibria. Our main contribution is to show how to test those restrictions in a way that is not affected by equilibrium selections. In general, the latter are unknown and have to be treated as nuisance parameters of the problem. Our approach is to first estimate the conditional quantiles nonparametrically, then use those to construct an asymptotic likelihood-ratio test of the order restrictions implied by the MCS. The test relies only on the asymptotic distribution results; it is an extension of the "chi-bar squared" test by Gourieroux, Holly, and Monfort (1982) and Kodde and Palm (1986) to restrictions on conditional quantiles.

It is worth pointing out that we work under quite weak assumptions on the equilibrium selection procedure. All we need to assume is that the latter places nonzero probabilities on the extremal equilibria, and that the equilibrium probabilities do not depend on the unobservables in the model. These assumptions are considerably weaker than assuming the probabilities of various equilibrium realizations known (see, e.g., Bjorn and Vuong, 1984; Jia, 2008); or finitely parameterized (see, e.g., Bjorn and Vuong, 1985; McKelvey and Palfrey, 1995; Ackerberg and Gowrisankaran, 2006; Sweeting, 2009; Bajari, Hong, and Ryan, 2010). They are also weaker than an assumption that only one equilibrium is played in the data (see, e.g. Aguirregabiria and Mira, 2007).

---

role in organizational design, as discussed in, for example Brynjolfsson and Milgrom (2012) and Milgrom and Roberts (1990a). Our results can potentially be of interest to this literature.

Example. Here is an example that illustrates how to use our results. Say that one is interested in testing whether an exogenous change in a policy causes the prices in the market for cars to increase. When there are complementarities between the policy and the agents' choice variables, the effect on prices takes the form of MCS. In the case of car prices, policy changes which increase marginal cost would cause the smallest and largest equilibria to increase. Examples of these policies are environmental regulations (Pakes, Berry, and Levinsohn, 1993) and voluntary export restraints (Berry, Levinsohn, and Pakes, 1999).

Concretely, let Y denote the price and X the policy dummy; further, suppose that an economic theory posits a reduced form model for Y that has the form $Y = g(X)U$, where one observes an "intended equilibrium" $g(X)$, subject to a multiplicative shock U. Here, g is generally unknown. A multiplicative error model is chosen so as to preserve the positivity of the price variable Y. Note, however, that by taking logarithms our entire analysis applies to models with additive errors.

We now explain how this setup would arise in a simplified version of the Berry, Levinsohn, and Pakes (1999) model. For this, assume there are only two firms in the market (American and Japanese) each producing only one car. The Japanese firm is subject to a Voluntary Export Restraint (VER) that is implemented as an implicit tax on firms' production. On the demand side, the consumers choose among the two cars based on a random utility specification. For simplicity, assume that all the observed product characteristics entering the random utilities, except the price, as well as all the observed marginal cost shifters, have been taken out of the model (or, alternatively, think of the analysis as being conditional on those variables). In this setup, Y would be the American and Japanese car prices, X the dummy indicating whether a particular car is subject to the VER, and U would capture
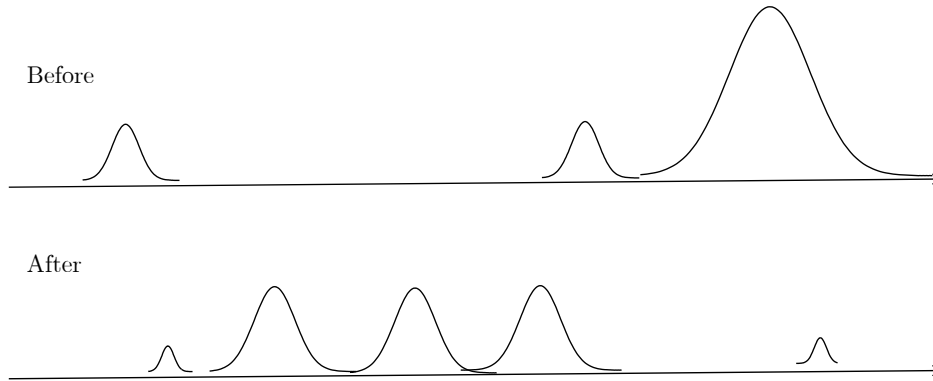
FIGURE 1. Price distribution before the change in policy, and after.

any unobservables in the model (such as unobserved product characteristics and unobserved productivity terms).

Many models which yield predictions for price competition—such as Berry, Levinsohn, and Pakes (1995, 1999), for example—are also likely to have multiple equilibria. We capture this by letting g be a correspondence (a set valued map) instead of a function (single valued map), so there is generally a set $\mathcal{E}_{XU}$ of equilibrium predictions for Y. We assume further that there is an unknown equilibrium selection procedure, which results in a distribution P_X over $\mathcal{E}_{XU}$. This multiple equilibrium model gives rise to a mixture conditional distribution for Y given X.

For example, in the Berry, Levinsohn, and Pakes (1999) model, the realizations of U could vary from one period to the next thus inducing the variation in prices across periods (one can also think of each time period as being a different market in which prices are observed). This results in a distribution of car prices given the VER dummy that due to the presence of multiple equilibria in each period takes a mixture form.

Figure 1 illustrates the effect on the price Y that may result from a change in the policy X. Before the change in the policy, we have three elements in the mixture,

and after we have five. The probabilities under each element are result of some equilibrium selection procedure. The case in Figure 1 presents a challenge as the conditional expectation of Y given X decreases. This means that, in expectation, the car prices *decrease* following the increase in marginal costs. And one can construct examples where following a change in X, the conditional mean of Y increases. Thus— as a result of equilibrium multiplicity—standard practices such as an OLS regression of prices on the policy dummy can be very misleading: the testable implications of the MCS property are not on the conditional mean of prices.

Our solution is to work with restrictions that MCS implies irrespective of the way equilibria are selected. As already said, those restrictions are on the conditional quantiles of Y given X, and we derive them following a reasoning similar to that in Echenique and Komunjer (2009). It is important to stress that we make few assumptions on the true equilibrium distribution. We only assume that P_X puts a positive probability on the extremal equilibria. Similarly, our assumptions on the distribution of the equilibrium deviations U are weak; we need them to belong to a well-known class of distributions in extreme-value theory. This class includes most distributions commonly used in empirical work, such as Gaussian, lognormal and exponential distributions.

Once the appropriate implications of the MCS hypothesis derived, we proceed with a construction of a likelihood-ratio test. In particular, our test is a test for order restrictions on the conditional quantiles of Y given X. We use a two step approach: first, we construct nonparametric estimators for the conditional quantiles of Y given X. The key difficulty here is that the MCS prediction holds only for quantiles that are extreme; hence, we need to use a nonstandard framework to derive their asymptotic distribution (Dekkers and de Haan, 1989; Chernozhukov, 2005; Chernozhukov and Fernandez-Val, 2011).

In the second step, we construct a likelihood-ratio test for order restrictions based on the asymptotic distributions of our conditional quantile estimators. This step presents important challenges as the existing results (Gourieroux, Holly, and Monfort, 1982; Kodde and Palm, 1986) apply only to the conditional means; hence, we need to extend them to our extreme conditional quantile framework. Perhaps an even greater difficulty comes from the presence of numerous nuisance parameters—unknown equilibrium selection probabilities—that we need to eliminate from our test statistic. Unfortunately, the standard approaches of dealing with nuisance parameters fail to work once we exit the usual asymptotic framework. Our solution is to first consider the problem in the exact case (as in Bartholomew, 1959a,b, for example), then extend the obtained solution to our asymptotic framework.

The remainder of the paper is organized as follows: We introduce the model in Section 2, and present the intuition behind our main results in Section 3. In Section 4 we present the basic statistical framework, and develop an approach to estimation in Section 5. Finally, in Section 6 we present our test. Last section concludes. The technical details including the proofs of the results stated in the main text as well as some additional useful results are relegated to an Appendix.

2. Setup

2.1. **Multiple Equilibrium Model.** Consider a familiar nonlinear regression model with a multiplicative error:

$$(1) \qquad\qquad\qquad\qquad Y = g(X)U$$

that relates a dependent variable $Y \in \mathbb{R}_{++}$, an explanatory variable $X \in \mathcal{X}$ with $\mathcal{X}$ finite in $\mathbb{R}$, and a latent disturbance $U \in \mathbb{R}_{++}$.[4] The map $g : \mathcal{X} \to \mathbb{R}_{++}$ in

---

[4]All of our results are easily transposable to additive error models of the form $Y = g(X) + U$ with $Y \in \mathbb{R}$, $U \in \mathbb{R}$, $X \in \mathcal{X}$ and $g : \mathcal{X} \to \mathbb{R}$, by taking logarithms.

Equation (1) is unknown; we assume however that g is positive valued, so that the positivity of the dependent variable is preserved. When the map g is known up to some finite dimensional parameter θ, one can write $g(X, \theta)$ in Equation (1). While the explanatory and dependent variables X and Y are observable, the disturbance U is not; U can be thought of as unaccounted heterogeneity in the model. Finally, note that the random variables Y and U are assumed to be continuous, whereas X is restricted to be discrete.[5] Hereafter, we shall use the lowercase letters y, x and u to denote the realizations of the random variables Y, X and U, respectively.

Underlying the model in Equation (1) is the assumption that, given the explanatory variable X, a unique value of the dependent variable Y can be assigned to each value of the disturbance U. In other words, conditional on X, the mapping from the unobservables to the observables is single valued, and g in Equation (1) is a function. In models that possess multiple equilibria, this letter property is generally violated as more than one value of Y can be associated with each value of U.

In order to adapt our model to multiple equilibria for Y, we shall assume that the map g in Equation (1) is a correspondence $g : \mathbb{R} \rightrightarrows \mathbb{R}_{++}$, which, to each $x \in \mathcal{X}$, assigns the set $\Gamma_x \equiv \{g_i(x), \ldots, g_{N_x}(x) : g_1(x) \leqslant \ldots \leqslant g_{N_x}(x)\}$. The maps g_i which define the image set Γ_x are single valued so every $g_i : \mathbb{R} \to \mathbb{R}_{++}$ is a function. We do not make any assumptions regarding continuity or differentiability of g_i's except that they are Borel-measurable. As a result, there are multiple equilibria for the dependent variable Y in Equation (1) given by $Y_i = g_i(X)U$ with $i = 1, \ldots, N_X$. We then let $\mathcal{E}_{XU} \equiv \{Y_1, \ldots, Y_{N_X}\}$ denote the equilibrium set.[6] Note that all equilibria

---

[5]Recall that X is a policy variable, which in the applications described in the Introduction took the form of a dummy variable. In the general setup, we allow X to be discrete. However, extending our analysis to continuous X's is nontrivial and thus outside the scope of this paper.

[6]Note that while we explicitly allow the cardinality of the equilibrium set N_x to vary with x, we can also accommodate the case in which the latter varies with u provided $\mathrm{Card}(\mathcal{E}_{xu})$ remains bounded by some M_x for every $u \in \mathbb{R}_{++}$.

for Y are ordered in $\mathcal{E}_{XU}$, i.e. $Y_1 \leqslant \ldots \leqslant Y_{N_X}$. We shall work with the following definition.

Definition 1. A *multiple equilibrium model* is a collection $(\mathcal{E}_{XU}, P_X, F_{U|X})$ such that for every $(x, x', u) \in \mathcal{X}^2 \times \mathbb{R}_{++}$ we have:

(i) $\mathcal{E}_{xu} \subseteq \mathbb{R}_{++}$ is finite and nonempty;

(ii) $x < x'$ implies that $\min \mathcal{E}_{xu} < \min \mathcal{E}_{x'u}$ and $\max \mathcal{E}_{xu} < \max \mathcal{E}_{x'u}$;

(iii) P_x is a probability distribution over $\mathcal{E}_{xu}$, $P_x(\min \mathcal{E}_{xu}) > 0$ and $P_x(\max \mathcal{E}_{xu}) > 0$;

(iv) $F_{U|X=x}$ is a twice-differentiable distribution function with positive density on $\mathbb{R}_{++}$.

We assume that $\mathcal{E}_{XU} \subseteq \mathbb{R}_{++}$ is finite, so we accommodate multiple, but finitely-many, equilibria. The assumption is common, and often justified by standard genericity arguments: In parameterized families of economic models, one obtains finitely many equilibria except on sets of measure zero (see, e.g., Mas-Colell, Whinston, and Green, 1995). Our results shall build on the MCS property in item (ii) of Definition 1: an increase in X causes the smallest and largest equilibria in $\mathcal{E}_{XU}$ to increase.

The probability distribution P_X in item (iii) of Definition 1 reflects some equilibrium selection procedure. It is important to note that while the elements of the equilibrium set $\mathcal{E}_{XU}$ vary with U, the probabilities assigned to them by P_X can only depend on X. In other words, the probability π_{Xi} of choosing the ith equilibrium Y_i under P_X ($i = 1, \ldots, N_X$) must not depend on U. In particular, this means that in our setup, the equilibrium selection mechanism can depend on X, as well as on other possibly unobserved variables, provided the latter are independent of U conditional on X.

Our multiple equilibrium model implies the following: given the explanatory variable X, the dependent variable Y is distributed as $F_{Y|X}$, where $F_{Y|X}$ is a discrete

mixture of continuous distributions:

$$(2) \qquad F_{Y|X}(y) = \sum_{i=1}^{N_X} \pi_{Xi} \, F_{U|X}\Big(\frac{y}{g_i(X)}\Big),$$

for any $y \in \mathbb{R}_{++}$, where π_{Xi} $(i = 1, \ldots, N_X)$ is the probability of choosing the ith equilibrium Y_i under P_X. The assumptions on $F_{U|X}$ imply that $F_{Y|X}$ is twice differentiable on $\mathbb{R}_{++}$ with density $f_{Y|X}$ that is positive on $\mathbb{R}_{++}$.

Given $\alpha \in (0,1)$, we let $q_{Y|X}(\alpha)$ denote the α-quantile under $F_{Y|X}$: $q_{Y|X}(\alpha) \equiv \inf\{y \in \mathbb{R}_{++} : F_{Y|X}(y) > \alpha\}$, which under our assumptions also equals $q_{Y|X}(\alpha) = F_{Y|X}^{-1}(\alpha)$. In what follows, we devote particular attention to the distribution tails of the dependent variable: $\bar{F}_{Y|X} \equiv 1 - F_{Y|X}$. Similarly, we let $\bar{F}_{U|X} \equiv 1 - F_{U|X}$. Note that given $\alpha \in (0,1)$, we have the following simple relation:

$$(3) \qquad q_{Y|X}(\alpha) = \bar{F}_{Y|X}^{-1}(1 - \alpha).$$

2.2. **On the Model Assumptions.** We now comment on the restrictions we have made in our multiple equilibrium model.

2.2.1. *Multiplicative error model.* We have defined the equilibrium set $\mathcal{E}_{XU}$ using the multiplicative error model specification in Equation (1), with g being a correspondence. Alternatively, one can take the mixture in Equation (2) to be one of the primitive assumptions of our multiple equilibrium model. As we shall show in subsequent sections, the mixture property in Equation (2) is instrumental in deriving our results. In particular, the latter do not explicitly use the multiplicative error specification in Equation (1).

This raises the question of the plausibility of the mixture assumption for $F_{Y|X}$. In Echenique and Komunjer (2009) we provide a general result on how such mixtures arise in structural econometric models of the form $r(Y, X) = U$ under fairly weak assumptions on the structural function r.

2.2.2. *Assumptions on P_X.* We have assumed that the largest and smallest equilibria in $\mathcal{E}_{XU}$ have positive probability under P_X—this is our only deviation from being agnostic regarding P_X. We actually need something somewhat weaker, and it will be clear that, without our weaker assumption, no testable implications are possible. We argue here that our assumption is reasonable.

To fix ideas, let $\mathcal{X} = \{\underline{x}, \overline{x}\} \subseteq \mathbb{R}$ with $\underline{x} < \overline{x}$. We need that for every $u \in \mathbb{R}_{++}$, the largest equilibrium in $\mathcal{E}_{\underline{x}u}$, *of those with positive $P_{\underline{x}}$ probability*, be smaller than the largest equilibrium in $\mathcal{E}_{\overline{x}u}$ with positive $P_{\overline{x}}$ probability. This is a weaker requirement than the one we have imposed above. It says that the equilibrium selection mechanism implicit in P_X should have the right correlation with respect to changes in X.

We claim that this correlation can be expected to hold: suppose agents are playing an equilibrium in $\mathcal{E}_{\underline{x}u}$ when the explanatory variable changes to $\overline{x}$. Then a broad class of learning dynamics must lead them to play a larger equilibrium; Echenique (2002) presents a formal statement and proof.

2.2.3. *Assumptions on $F_{U|X}$.* Our multiple equilibrium model assumes that $F_{U|X}$ is a continuous distribution with support $\mathbb{R}_{++}$. It is worth pointing out that we let $F_{U|X}$ be unknown. In some cases, it might be preferable to assume $F_{U|X}$ known, at least up to some finite-dimensional parameter; in such cases, the conditional distribution of Y in Equation (2) could in principle be estimated via maximum likelihood methods, provided the equilibrium selection probabilities P_X are either known or finitely parameterized. However, the presence of unknown equilibrium probabilities P_X in $F_{Y|X}$ causes almost all the practical problems of implementation and model estimation with maximum likelihood methods.[7]

---

[7]For example, if the estimation is carried out by using the EM-algorithm, both the location of different equilibria and the probabilities attached to them need to be estimated (see, e.g., Carroll, Ruppert, and Stefanski, 1995).
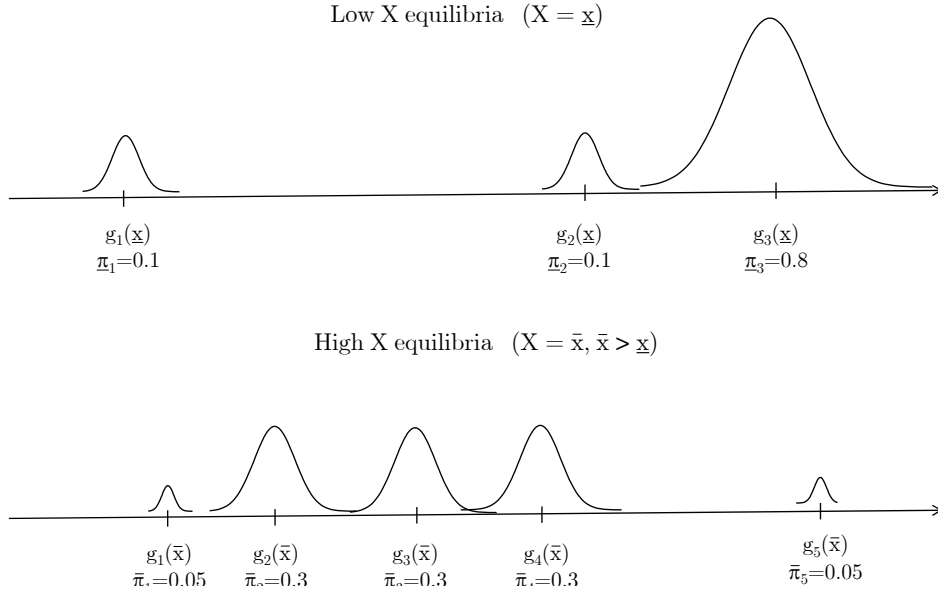
FIGURE 2. Equilibrium distributions.

3. NATURE OF THE PROBLEM AND RESULTS

We first explain our results informally. Consider again our example in which $\mathcal{X} = \{\underline{x}, \overline{x}\} \subseteq \mathbb{R}$, $\underline{x} < \overline{x}$ where $\underline{x}$ and $\overline{x}$ denote low- and high-level of the explanatory variable. In addition, letting $\underline{y}_i$ and $\overline{y}_j$ denote the equilibrium levels when $(X = \underline{x}, U = u)$ and $(X = \overline{x}, U = u)$, respectively, assume that $\mathcal{E}_{\underline{x}u} = \{\underline{y}_1, \underline{y}_2, \underline{y}_3\}$ and $\mathcal{E}_{\overline{x}u} = \{\overline{y}_1, \overline{y}_2, \overline{y}_3, \overline{y}_4, \overline{y}_5\}$, where $\underline{y}_i = g_i(\underline{x})u$ and $\overline{y}_i = g_i(\overline{x})u$. The situation is represented in Figure 2.

The problem of obtaining testable implications is to say how the distributions $F_{Y|X=\underline{x}}$ and $F_{Y|X=\overline{x}}$ must differ ($F_{Y|X}$ was defined in Equation (2)). All we have to work with is that $\underline{y}_3 < \overline{y}_5$ (and $\underline{y}_1 < \overline{y}_1$), but the probability of the $\overline{y}_5$ equilibrium is very low, that of $\underline{y}_3$ is very high, and there are three equally likely equilibria with high sum of probabilities, $\overline{y}_2, \overline{y}_3$ and $\overline{y}_4$, that are smaller than $\underline{y}_3$.

Note that the mean (and median) of the dependent variable under $F_{Y|X=\overline{x}}$ is smaller than that under $F_{Y|X=\underline{x}}$. Thus the conditional mean (and median) of Y does not

change monotonically in X. One can change the example so the conditional mean increases instead of decreasing; thus the MCS property in item (ii) of Definition 1 produces no testable implications for the conditional mean of the dependent variable. One is also more likely to observe a realization under $F_{Y|X=\underline{x}}$ that is larger than under $F_{Y|X=\overline{x}}$ than vice versa.

Our solution to finding testable implications is to assume the right structure on the distribution tails, so the effect of $\underline{y}_3 < \overline{y}_5$ is felt for large enough values of the dependent variable, irrespective of the values of the corresponding probabilities $P_{\underline{x}}$ and $P_{\overline{x}}$. We show how, for large enough realizations y of Y, the distribution tails $\bar{F}_{Y|X=\underline{x}} \equiv 1 - F_{Y|X=\underline{x}}$ and $\bar{F}_{Y|X=\overline{x}} \equiv 1 - F_{Y|X=\overline{x}}$ must satisfy $\bar{F}_{Y|X=\underline{x}}(y) < \bar{F}_{Y|X=\overline{x}}(y)$.

To further simplify the notation, let $\underline{\pi}_i$ (resp. $\overline{\pi}_j$) denote the probabilities assigned to the elements of $\mathcal{E}_{\underline{x}u}$ (resp. $\mathcal{E}_{\overline{x}u}$) under $P_{\underline{x}}$ (resp. $P_{\overline{x}}$). Note that the tail $\bar{F}_{Y|X}$ is related to $\bar{F}_{U|X} \equiv 1 - F_{U|X}$ via:

$$(4) \quad \bar{F}_{Y|X=\underline{x}}(y) = \underline{\pi}_3 \bar{F}_{U|X=\underline{x}}\big(y/g_3(\underline{x})\big) + \underline{\pi}_2 \bar{F}_{U|X=\underline{x}}\big(y/g_2(\underline{x})\big) + \underline{\pi}_1 \bar{F}_{U|X=\underline{x}}\big(y/g_1(\underline{x})\big),$$

Assume that the tails of $F_{U|X=\underline{x}}$ satisfy the following property:

$$(5) \qquad\qquad \lim_{u \to \infty} \frac{\bar{F}_{U|X=\underline{x}}(\lambda u)}{\bar{F}_{U|X=\underline{x}}(u)} = 0,$$

whenever $\lambda > 1$. Property (5) requires that the tail of the distribution $F_{U|X}$ is not too heavy. As we explain below, it is a well-known condition in the statistics of extreme values, and it is satisfied by most distributions familiar to practitioners.

Now,

$$\frac{\bar{F}_{U|X=\underline{x}}\big(y/g_2(\underline{x})\big)}{\bar{F}_{U|X=\underline{x}}\big(y/g_3(\underline{x})\big)} = \frac{\bar{F}_{U|X=\underline{x}}(\lambda z)}{\bar{F}_{U|X=\underline{x}}(z)},$$

where we have let $z \equiv y/g_3(\underline{x})$ and $\lambda \equiv g_3(\underline{x})/g_2(\underline{x}) > 1$, and similarly with g_1 in place of g_2. So, dividing by $\bar{F}_{U|X=\underline{x}}\big(y/g_3(\underline{x})\big)$ throughout Equation (4), and using

Property (5), we obtain that:

$$\bar{F}_{Y|X=\underline{x}}(y) \sim \underline{\pi}_3 \bar{F}_{U|X=\underline{x}}\big(y/g_3(\underline{x})\big) \text{ as } y \text{ goes to } \infty.$$

In other words, the behavior of $\bar{F}_{Y|X=\underline{x}}(y)$ for large y is driven solely by the largest equilibrium $\underline{y}_3$. Under analogous assumptions on the tails of $F_{U|X=\overline{x}}$, it is easy to show that $\bar{F}_{Y|X=\overline{x}}(y)$ behaves like $\overline{\pi}_5 \bar{F}_{U|X=\overline{x}}\big(y/g_5(\overline{x})\big)$. Thus,

$$(6) \qquad \frac{\bar{F}_{Y|X=\underline{x}}(y)}{\bar{F}_{Y|X=\overline{x}}(y)} \sim \underbrace{\left[\frac{\underline{\pi}_3}{\overline{\pi}_5}\right]}_{A} \underbrace{\left[\frac{\bar{F}_{U|X=\underline{x}}\big(y/g_3(\underline{x})\big)}{\bar{F}_{U|X=\underline{x}}\big(y/g_5(\overline{x})\big)}\right]}_{B} \underbrace{\left[\frac{\bar{F}_{U|X=\underline{x}}\big(y/g_5(\overline{x})\big)}{\bar{F}_{U|X=\overline{x}}\big(y/g_5(\overline{x})\big)}\right]}_{C}.$$

From item (iii) in Definition 1, we know that the term A is bounded. Since $\underline{y}_3 < \overline{y}_5$, our assumption on $F_{U|X=\underline{x}}$ in Equation (5) implies that the B term goes to 0 as y grows. If, in addition, we assume that:

$$(7) \qquad \frac{\bar{F}_{U|X=\underline{x}}(y)}{\bar{F}_{U|X=\overline{x}}(y)} \text{ is bounded as } y \text{ goes to } \infty,$$

then the C term is bounded. So $\bar{F}_{Y|X=\underline{x}}(y)/\bar{F}_{Y|X=\overline{x}}(y)$ converges to 0 irrespective of the probabilities under $P_{\underline{x}}$ and $P_{\overline{x}}$. Hence, for large enough y, the tail of $F_{Y|X=\overline{x}}(y)$ is thicker than that of $F_{Y|X=\underline{x}}(y)$; this is the essence of our testable implication.

To summarize, Statements (5) and (7) together ensure that the ratio of $\bar{F}_{Y|X=\underline{x}}$ to $\bar{F}_{Y|X=\overline{x}}$ goes to zero. This is our testable implication: $\bar{F}_{Y|X=\underline{x}}(y)/\bar{F}_{Y|X=\overline{x}}(y)$ for y large enough. As a result, large enough population quantiles must be larger under $F_{Y|X=\overline{x}}$ than under $F_{Y|X=\underline{x}}$. In the next section we show how this result generalizes.

4. Econometric Framework

A useful statistical framework to formalize the basic ideas in Section 3 is that of regularly-varying functions. We first give some preliminary definitions, and results on regularly-varying functions. We then exploit this theory to develop statistical tests for the models in Section 2.

4.1. **Regular Variation Theory.** In this subsection, H denotes a distribution function with positive density h on $\mathbb{R}_{++}$ and distribution tail $\bar{H} \equiv 1 - H$. We shall focus on the behavior of $\bar{H}$ in $+\infty$, knowing that analogous results can be obtained at zero.

Definition 2. A distribution tail $\bar{H} : \mathbb{R}_{++} \to (0, 1)$ is *regularly varying* at c, $0 \leqslant c \leqslant \infty$, with index ρ, $-\infty \leqslant \rho < \infty$, denoted $\bar{H} \in \mathcal{R}_\rho$ at c, if for $\lambda > 0$:

$$(8) \qquad \lim_{x \to c} \frac{\bar{H}(\lambda x)}{\bar{H}(x)} = \lambda^\rho.$$

The notion of regular variation was first introduced by Karamata (1930) (see Resnick, 1987, for an exposition). When c is understood we shall often abuse notation and write $\bar{H} \in \mathcal{R}_\rho$.

We focus on regular variation at $c = \infty$ with index $\rho = -\infty$, denoted by $\mathcal{R}_{-\infty}$ at ∞. Most of the distributions employed in economics, such as the Gaussian, exponential and lognormal distributions, are in $\mathcal{R}_{-\infty}$ at ∞. The distributions in $\mathcal{R}_{-\infty}$ at ∞ are also called "$(-\infty)$-varying" or "rapidly varying." They are moderately heavy-tailed, or light-tailed, meaning that their tails decrease to zero faster than any power law $x^{-\alpha}$.[8]

Note that the special case of $\bar{H}(\cdot)$ being in $\mathcal{R}_{-\infty}$ at ∞ is defined by

$$(9) \qquad \lim_{x \to \infty} \frac{\bar{H}(\lambda x)}{\bar{H}(x)} = \begin{cases} 0 & \text{if} \quad \lambda > 1 \\ \infty & \text{if} \quad \lambda < 1. \end{cases}$$

The discussion in Section 3 should suggest that Statement (9) is a useful property. Now, the property in Statement (9) does not control the rate at which $\bar{H}(\lambda x)/\bar{H}(x)$

---

[8]This implies that all the moments of a random variable with a $(-\infty)$-varying distribution tail are finite. Examples of distributions with ρ-varying tails, $\rho > -\infty$, which *do not* have finite moments are: (1) a stable law with index α, $0 < \alpha < 2$, for which $\rho = -\alpha$; (2) a Cauchy distribution, for which $\rho = -1$. Hence the use of those distributions is not permitted in our framework.

converges. By using a subclass of $(-\infty)$-varying distribution tails, called Γ (see, e.g., de Haan, 1970), we can exercise this control.

Definition 3. A distribution tail $\bar{H}$ belongs to the *class* Γ, $\bar{H} \in \Gamma$, if there exists a function $a : \mathbb{R}_{++} \to \mathbb{R}_{++}$ such that for $\lambda > 0$,

$$(10) \qquad \lim_{x \to \infty} \frac{\bar{H}(x + \lambda a(x))}{\bar{H}(x)} = \exp(-\lambda);$$

a is called the auxiliary function of $\bar{H}$.

When $\bar{H} \in \Gamma$, one can show that a can be chosen as $a \equiv \bar{H}/h$ (we shall often make this choice).

That $\Gamma \subseteq \mathcal{R}_{-\infty}$ is a direct consequence of Theorem 1.5.1 in de Haan (1970). Examples of distributions whose tails are in Γ are: exponential, two-parameter Gamma, Gaussian, lognormal, and Weibull (see, e.g., Embrechts, Kluppelberg, and Mikosch, 1997).

The tail properties in Equations (8) and (10) translate into similar properties for the inverse function $\bar{H}^{-1} : (0,1) \to \mathbb{R}_{++}$ (see Lemma 5) and the class of regularly varying functions is closed under inversion. The inverses of functions in Γ, however, do not belong to Γ but form a class called Π (see, e.g., de Haan, 1970, 1974).

Definition 4. A function $\bar{H}^{-1} : (0,1) \to \mathbb{R}_{++}$ belongs to the *class* Π, $\bar{H}^{-1} \in \Pi$, if there exist functions $b : \mathbb{R}_{++} \to \mathbb{R}_{++}$ and $a : \mathbb{R}_{++} \to \mathbb{R}_{++}$ such that, for $\mu \in (0,1)$,

$$(11) \qquad \lim_{y \downarrow 0} \frac{\bar{H}^{-1}(\mu y) - b(y)}{a(y)} = -\ln \mu.$$

When $\bar{H}$ belongs to Γ with auxiliary function $\tilde{a}$, Equation (11) holds with $b(y) \equiv \bar{H}^{-1}(y)$ and $a(y) \equiv \tilde{a}(\bar{H}^{-1}(y))$.

4.2. **Testable Implications: General Result.** We now return to our multiple equilibrium model $(\mathcal{E}_{XU}, P_X, F_{U|X})$ and impose structure on the distribution tails $\bar{F}_{U|X}$ of the disturbances.

Assumption S1. Say that a multiple equilibrium model $(\mathcal{E}_{XU}, P_X, F_{U|X})$ satisfies assumption S1 if, for every $x \in \mathcal{X}$, $\bar{F}_{U|X=x}$ is in $\mathcal{R}_{-\infty}$ at ∞.

We now show how the properties of the tails $\bar{F}_{U|X}$ translate into properties of the tail of the conditional distribution of the dependent variable $\bar{F}_{Y|X}$ in Equation (2). Recall that π_{XN_X} denotes the probability of choosing the largest equilibrium $Y_{N_X} = g_{N_X}(X)U$ under P_X.

Lemma 1. If $(\mathcal{E}_{XU}, P_X, F_{U|X})$ satisfies S1, then for every $x \in \mathcal{X}$:

(i) $\bar{F}_{Y|X=x}$ is in $\mathcal{R}_{-\infty}$ at ∞, and $\bar{F}_{Y|X=x}(y) \sim \pi_{xN_x} \bar{F}_{U|X=x}\big(y/g_{N_x}(x)\big)$ as $y \to \infty$;

(ii) $\bar{F}^{-1}_{U|X=x}$ and $\bar{F}^{-1}_{Y|X=x}$ are in $\mathcal{R}_0$ at 0, and $\bar{F}^{-1}_{Y|X=x}(v) \sim g_{N_x}(x)\bar{F}^{-1}_{U|X=x}(v)$ as $v \downarrow 0$.

Thus, the limit behavior of the distribution tail $\bar{F}_{Y|X}$ is determined by the largest equilibrium in $\mathcal{E}_{XU}$ and its probability. In the limit, the conditional quantiles of Y are proportional to the quantiles under $F_{U|X}$, and the constant of proportionality equals $g_{N_X}(X)$.

In order to generalize the argument in Section 3 we need to strengthen our assumptions:

Assumption S2. Say that a multiple equilibrium model $(\mathcal{E}_{XU}, P_X, F_{U|X})$ satisfies S2 if it satisfies S1 and, in addition, for every $(x, x') \in \mathcal{X}^2$ such that $x < x'$, we have:

(12) $$\frac{\bar{F}_{U|X=x}(u)}{\bar{F}_{U|X=x'}(u)} \text{ is bounded as } u \text{ goes to } \infty.$$

Using the above assumptions together with Lemma 1 allows us to derive our first main result :

Theorem 1. If $(\mathcal{E}_{XU}, P_X, F_{U|X})$ satisfies S2, then for any $(x, x') \in \mathcal{X}^2$ there is $\bar{y} \in \mathbb{R}_{++}$ such that $x < x'$ implies $\bar{F}_{Y|X=x}(y) < \bar{F}_{Y|X=x'}(y)$ for all $y \geqslant \bar{y}$. Equivalently, there is $\bar{\alpha} \in (0, 1)$ such that $x < x'$ implies $q_{Y|X=x}(\alpha) < q_{Y|X=x'}(\alpha)$ for all $\alpha \in [\bar{\alpha}, 1)$.

The idea of Theorem 1 is that, if the distribution $F_{U|X}$ is not too heavy-tailed, the effect of X on the largest equilibrium in $\mathcal{E}_{XU}$ will eventually be noticed in the tail of $F_{Y|X}$. In a sense, there is a race between the potentially damaging effect of other equilibria in $\mathcal{E}_{XU}$, and the effect of the largest equilibrium Y_{N_X}. Since P_X is arbitrary, P_X can work in favor of the other equilibria in $\mathcal{E}_{XU}$, as in Figure 2. But the $(-\infty)$-varying condition on $\bar{F}_{U|X}$ and Property (12) together guarantee that the largest equilibrium wins the race. Hence, for large values of y, the conditional distributions $F_{Y|X=x}(y)$ of the dependent variable have tails that increase monotonically with x, a property akin to monotonicity in first-order stochastic dominance. Equivalently, Theorem 1 has consequences for the quantiles of Y conditional on X. In the limit, the conditional quantiles of the dependent variable given X are monotone increasing in X.

4.3. **Further Model Implications.** Theorem 1 suggests one can use estimates of conditional quantiles under $F_{Y|X}$ for testing, but there are several difficulties. First, the theorem does not determine $\bar{y}$ or $\bar{\alpha}$; it does not identify the quantiles for which we have testable implications. Second, we need to know the (asymptotic) distribution of the conditional quantile estimates—the key is to derive the latter by imposing structure on the distributions $F_{U|X}$ while maintaining our agnosticism about the P_X distributions. Third, given the asymptotic distributions of estimates for quantiles under $F_{Y|X}$, we need to derive a test that is not influenced by the P_X distributions nor the non-extremal values in $\mathcal{E}_{XU}$, for which our model makes no predictions.

In order to deal with the asymptotics, we need to impose further structure on the distribution tail $\bar{F}_{U|X}$: in addition to being $(-\infty)$-varying, $\bar{F}_{U|X}$ is now assumed to belong to the class Γ.

Assumption S3. Say that a multiple equilibrium model $(\mathcal{E}_{XU}, P_X, F_{U|X})$ satisfies S3 if it satisfies S1 and, in addition, for every $x \in \mathcal{X}$ we have $\bar{F}_{U|X=x} \in \Gamma$ with auxiliary function a_x^U.

This allows us to show the following results on the tails of conditional distributions $F_{Y|X}$ of the dependent variable.

Lemma 2. *If $(\mathcal{E}_{XU}, P_X, F_{U|X})$ satisfies S3, then for every $x \in \mathcal{X}$:*
(i) $\bar{F}_{Y|X=x} \in \Gamma$ with auxiliary function $a_x^Y(y) = g_{N_x}(x)a_x^U\big(y/g_{N_x}(x)\big)$ for all $y > 0$;
(ii) $\bar{F}_{U|X=x}^{-1}$ and $\bar{F}_{Y|X=x}^{-1}$ are in Π with auxiliary functions $a_x^U \circ \bar{F}_{U|X=x}^{-1}$ and $a_x^Y \circ \bar{F}_{Y|X=x}^{-1}$ in $\mathcal{R}_0$ at 0, and $a_x^Y\big(\bar{F}_{Y|X=x}^{-1}(v)\big) \sim g_{N_x}(x)a_x^U\big(\bar{F}_{U|X=x}^{-1}(v)\big)$ as $v \downarrow 0$.

Lemma 2 presents two results: First, that the Γ (resp. Π) properties of $\bar{F}_{U|X}$ (resp. $\bar{F}_{U|X}^{-1}$) continue to hold for $\bar{F}_{Y|X}$ (resp. $\bar{F}_{Y|X}^{-1}$). Hence, we will only need to make assumptions on the behavior of $\bar{F}_{U|X}$ in Equation (2) in order to fully characterize the behavior of $\bar{F}_{Y|X}(y)$ as y gets large. Note that this result is particularly important if we want to preserve our agnosticism about the probabilities P_X over equilibria in $\mathcal{E}_{XU}$.

The second result of Lemma 2 is to show how $a_X^Y \circ \bar{F}_{Y|X}^{-1}$ relates to $a_X^U \circ \bar{F}_{U|X}^{-1}$. We shall prove that these expressions are involved in the formulation of the central limit theorem for empirical conditional quantiles under $F_{Y|X}$. In other words, the results of Lemma 2 are essential for understanding the asymptotic properties of the estimators for conditional quantiles of Y given X, and hence for constructing an econometric test of the implication derived in Theorem 1.

5. ESTIMATION

5.1. **Notation and Setup.** Fix $x \in \mathcal{X}$ and assume that the econometrician observes some large number T_x of realizations of the dependent variable Y obtained when the

explanatory variable X takes the value x. More formally, let $(Y_{x,1}, \ldots, Y_{x,T_x})$ be a random sample of size T_x from a distribution function $F_{Y|X=x}$. Let $(y_{x,1}, \ldots, y_{x,T_x})$ denote the realizations of $(Y_{x,1}, \ldots, Y_{x,T_x})$ and write $\hat{F}_{Y|X=x}$ to be the empirical distribution function, $\hat{F}_{Y|X=x}(y) \equiv T_x^{-1} \sum_{t=1}^{T_x} \mathbb{1}(y_{x,t} \leqslant y)$ for $y > 0$. For a given α, $0 < \alpha < 1$, the empirical quantile under $F_{Y|X=x}$ is then given by:

$$(13) \qquad \hat{q}_{Y|X=x}(\alpha) \equiv \inf\{y \in \mathbb{R}_{++} : \hat{F}_{Y|X=x}(y) > \alpha\}.$$

Under standard regularity conditions, the estimator in Equation (13) is consistent for the true α-quantile under $F_{Y|X=x}$. Consistency of $\hat{q}_{Y|X=x}(\alpha)$ can be extended to cases where $(Y_{x,1}, \ldots, Y_{x,T_x})$ is a weakly dependent time-series, provided additional assumptions hold (Pollard, 1991; Portnoy, 1991; Koenker and Zhao, 1996; Komunjer, 2005; Chernozhukov, 2005; Chernozhukov and Fernandez-Val, 2011); for the sake of simplicity, we focus on the independent case.

To alleviate the notation, we drop the reference to x when doing so introduces no ambiguities. Hence we use the notation $(Y_1, \ldots, Y_T)$, T, $\hat{F}$ and $\hat{q}(\alpha)$ to denote the random sample under $F_{Y|X=x}$, its size, the corresponding empirical distribution function and the α-quantile estimator in Equation (13).

As pointed out previously, the main object of interest are α-quantiles with probabilities α close to unity. How close α is to 1 is determined by the sample size T; hence we let this probability be a function of the sample size, and we denote it by α_T. Knowing how α varies with T will then enable us to answer the question: for a given sample size T how large α needs to be for the ordering in Theorem 1(ii) to hold.

5.2. **Central Limit Theory for Intermediate Empirical Quantiles.** We now derive the asymptotic distribution of $\hat{q}(\alpha_T)$ in Equation (13) when $\lim_{T \to \infty} \alpha_T = 1$ and when $(1 - \alpha_T)T$ has a positive limit as T goes to infinity. In particular, we

consider the case where $\lim_{T\to\infty}(1 - \alpha_T)T = \infty$. This last condition describes how fast α has to go to unity relative to the sample size T; knowing that $T^{-1} = o(1 - \alpha_T)$ we can use an appropriate limit theory result to derive an asymptotic distribution of the α-quantile estimator $\hat{q}(\alpha_T)$ in Equation (13).

We shall need the following lemma.

Lemma 3. *Consider a random sample $(Y_1, \ldots, Y_T)$ of size T from F and let $\hat{q}(\alpha_T)$ be the corresponding empirical α_T-quantile. If the distribution tail $\bar{F} \in \Gamma$ with auxiliary function a and with density f which is eventually non-increasing, then, provided $\lim_{T\to\infty} \alpha_T = 1$ and $\lim_{T\to\infty}(1 - \alpha_T)T = \infty$ we have:*

$$\sqrt{T(1 - \alpha_T)}\, \frac{\hat{q}(\alpha_T) - q(\alpha_T)}{a\big(q(\alpha_T)\big)} \xrightarrow{d} \mathcal{N} \quad and \quad \frac{\hat{q}(\beta_T) - \hat{q}(\alpha_T)}{a\big(q(\alpha_T)\big)} \xrightarrow{p} \ln \rho \, ,$$

where $\mathcal{N}$ is a standard Gaussian random variable and β_T is such that $\alpha_T < \beta_T < 1$ and $(1 - \alpha_T)/(1 - \beta_T) \to \rho$ with $\rho > 1$.

Lemma 3 presents two limit results. The first was proven by Falk (1989). The second is new.

The first result in the lemma shows the asymptotic behavior of intermediate empirical quantiles when α_T depends on the sample size T. It is an extension of the well-known result for central α-quantiles with $\alpha \in (0, 1)$ fixed (see, e.g., Mosteller, 1946; Smirnov, 1952; Siddiqui, 1960; Bahadur, 1966; Bassett and Koenker, 1978; Powell, 1984, 1986), to the case where α increases with the sample size T. Dekkers and de Haan (1989), Chernozhukov (2005), and Chernozhukov and Fernandez-Val (2011) prove this extension under an additional assumption on the tail behavior of F. While it is not new, we include a proof of the first result to make the paper self-contained, and because it requires little beyond what we need to prove the second result.

The second limit result of Lemma 3 is important because it gives us a consistent estimator of the variance of the empirical quantile. Recall that Theorem 1 says that the conditional quantiles of Y given X must be increasing in X. With consistent estimators of quantiles in hand, a test seems easy to derive. The problem, though, is that we do not know how the asymptotic variances of the quantile estimators change with X. Our second result in Lemma 3 allows us to solve the problem.

The second limit result of Lemma 3 extends a result on the asymptotic distribution of the quantile spacings derived by Dekkers and de Haan (1989) for the case $\rho = 2$ (see also Chernozhukov, 2005; Chernozhukov and Fernandez-Val, 2011). The result by Dekkers and de Haan (1989) requires that $d\bar{F}^{-1}(v)/dv$ be in $\mathcal{R}_{-1}$ at 0, an assumption that we need to avoid because it implies a restriction on the equilibrium selection P_X. By focusing on consistency, and not on the asymptotic distribution of quantile spacings, we obtain a result only assuming that $\bar{F}$ in Γ and that f if eventually non-increasing. Consistency, in turn, is sufficient for our testing procedure.

We should note that the assumption that f be eventually non-increasing imposes no restriction on the equilibrium selection probabilities π_{Xi} in Equation (2), and follows from requiring the density of $F_{U|X}$ to be eventually non-increasing.

5.3. **Estimates for Conditional Quantiles under** $(\mathcal{E}_{XU}, P_X, F_{U|X})$. We now assume a collection of random samples for different values $x \in \mathcal{X}$ of the explanatory variable X. Concretely, consider a multiple equilibrium model $(\mathcal{E}_{XU}, P_X, F_{U|X})$, and assume we observe realizations from $k \geqslant 2$ random samples $(Y_{x_1,1}, \ldots, Y_{x_1,T_{x_1}})$ to $(Y_{x_k,1}, \ldots, Y_{x_k,T_{x_k}})$ of sizes T_{x_1} to T_{x_k}, respectively. To ease the notation, for any $j = 1, \ldots, k$, we let $(Y_{j,1}, \ldots, Y_{j,T_j})$ denote $(Y_{x_j,1}, \ldots, Y_{x_j,T_{x_j}})$; in other words, we replace the subscript x_j with j whenever doing so does not introduce any ambiguity. The k samples are assumed independent and drawn from the distributions $F_{Y|X=x_1}$ to $F_{Y|X=x_k}$, respectively, with $(x_1, \ldots, x_k) \in \mathcal{X}^k$.

In order to use the results of Lemma 3 we need to impose the following assumption on the tails of $F_{U|X}$:

Assumption S4. Say that a multiple equilibrium model $(\mathcal{E}_{XU}, P_X, F_{U|X})$ satisfies S4 if it satisfies S3 and, in addition, the densities $h_{U|X}$ are eventually non-increasing.

The limit results of Lemma 3 then yield the following result:

Theorem 2. *Assume* $(\mathcal{E}_{XU}, P_X, F_{U|X})$ *satisfies S4, and consider* k *independent random samples* $(Y_{j,1}, \ldots, Y_{j,T_j}), j = 1, \ldots, k,$ *each of size* $T_j \geqslant 1$ *and drawn from* $F_{Y|X=x_j}$ *with* $x_j \in \mathcal{X}$. *If for every* j *we have:* $0 < \alpha_{T_j} < \beta_{T_j} < 1$, $\lim_{T_j \to \infty} \alpha_{T_j} = 1$, $\lim_{T_j \to \infty} (1 - \alpha_{T_j}) T_j = \infty$ *and* $\lim_{T_j \to \infty} (1 - \alpha_{T_j})/(1 - \beta_{T_j}) = \rho_j$ *with* $\rho_j > 1$, *then as* $T \to \infty$:

$$\frac{\hat{q}_{Y|X=x_j}(\alpha_{T_j}) - \mu_j}{\sigma_j} \overset{d}{\to} \mathcal{N}_j \quad \text{with } \mu_j \equiv q_{Y|X=x_j}(\alpha_{T_j}) \text{ and } \sigma_j \equiv \frac{a_{x_j}^Y(\mu_j)}{\sqrt{(1 - \alpha_{T_j})T_j}},$$

where $a_{x_j}^Y$ *is the auxiliary function of* $F_{Y|X=x_j}$, $a_{x_j}^Y = \bar{F}_{Y|X=x_j}/f_{Y|X=x_j}$, *and* $\mathcal{N}_1, \ldots, \mathcal{N}_k$ *are* k *independent standard normal random variables. Moreover, the scaling constants* σ_j *can be consistently estimated via:*

$$\frac{\hat{\sigma}_j}{\sigma_j} \equiv \sigma_j^{-1} \frac{\hat{q}_{Y|X=x_j}(\alpha_{T_j}) - \hat{q}_{Y|X=x_j}(\beta_{T_j})}{\ln \rho_j \sqrt{(1 - \alpha_{T_j})T_j}} \overset{p}{\to} 1.$$

For any given $k \geqslant 2$, the results of Theorem 2 allow us to determine the asymptotic behavior of estimates for conditional quantiles under $F_{Y|X}$. With conditional quantile estimators in hand, we can then test the implications in Theorem 1.

For the purpose of testing, we make an assumption on the rate of growth of the different samples. The assumption ensures that the $(1 - \alpha_{T_j})T_j$ grow at the same speed, and that we consider the same α_T-quantile, for all k samples. We can then formulate our results in the standard asymptotic framework, i.e. as $T \to \infty$. Concretely, assume that the sample sizes $(T_1, \ldots, T_k)$ and the corresponding

probabilities $(\alpha_{T_1}, \ldots, \alpha_{T_k})$ are such that for every j there exist α_T and c_j that satisfy:

$$(14) \quad \alpha_{T_j} = \alpha_T \text{ and } T_j = c_j T, \text{ with } 0 < \alpha_T < 1, \lim_{T \to \infty} (1 - \alpha_T)T = \infty, \text{ and } c_j > 0.$$

6. TESTING

6.1. **Test Hypotheses.** From Theorem 1, the observable restriction of our multiple equilibrium model $(\mathcal{E}_{XU}, P_X, F_{U|X})$ is that $x_1 < \ldots < x_k$ implies $q_{Y|X=x_1}(\alpha_T) < \ldots < q_{Y|X=x_k}(\alpha_T)$ as $\alpha_T \to 1$. Hence, we are interested in testing weather an increase in the explanatory variable results in an increase in the conditional quantiles of the dependent variable. The opposite case of interest is the one in which an increase in X produces no effect on the conditional quantiles of Y given X, so that we have $x_1 < \ldots < x_k$ and $q_{Y|X=x_1}(\alpha_T) = \ldots = q_{Y|X=x_k}(\alpha_T)$. Those two cases define our alternative and null hypotheses, respectively.

More formally, for given values $x_1 < \ldots < x_k$ in $\mathcal{X}^k$ we test the null hypothesis $H_0 : q_{Y|X=x_1}(\alpha_T) = \ldots = q_{Y|X=x_k}(\alpha_T)$, as $\alpha_T \to 1$, against an ordered alternative $H_1 : q_{Y|X=x_1}(\alpha_T) \leqslant \ldots \leqslant q_{Y|X=x_k}(\alpha_T)$, as $\alpha_T \to 1$, with strict inequality for at least one value of j, $1 \leqslant j \leqslant k$.

Our test statistic is a function of estimates for conditional quantiles under $F_{Y|X}$; from Theorems 1 and 2 we know that the latter satisfy the following property:

Corollary 3. *Assume $(\mathcal{E}_{XU}, P_X, F_{U|X})$ satisfies S2 and S4, and let $(Y_{j,1}, \ldots, Y_{j,c_j T})$, $j = 1, \ldots, k$, be k independent random samples of size $c_j T$ (with $T \geqslant 1$) drawn from $F_{Y|X=x_j}$, $x_j \in \mathcal{X}$. If $0 < \alpha_T < \beta_T < 1$, $\lim_{T \to \infty} \alpha_T = 1$, $\lim_{T \to \infty} (1 - \alpha_T)T = \infty$ and $\lim_{T \to \infty} (1 - \alpha_T)/(1 - \beta_T) = \rho$, with $\rho > 1$, then as $T \to \infty$:*

$$x_1 < \ldots < x_k \text{ implies } \mu_1 < \ldots < \mu_k$$

where $\mu_j \equiv q_{Y|X=x_j}(\alpha_T)$, and

$$\frac{\hat{q}_{Y|X=x_j}(\alpha_T) - \mu_j}{\hat{\sigma}_j} \xrightarrow{d} \mathcal{N}_j \ \text{with} \ \hat{\sigma}_j \equiv \frac{\hat{q}_{Y|X=x_j}(\alpha_T) - \hat{q}_{Y|X=x_j}(\beta_T)}{\ln \rho \sqrt{c_j(1 - \alpha_T)T}}$$

where $\mathcal{N}_1, \dots, \mathcal{N}_k$ are k independent standard normal random variables.

Note that the asymptotic distribution result in Corollary 3 exploits the sample size growth assumptions made in Equation (14). It follows by applying Slutsky's Theorem to the results derived in Theorem 2.

6.2. **Exact Test for Order Restrictions.** Assume for the moment that all the distribution results from Corollary 3 are exact rather than being asymptotic, i.e. assume that for some probability α_T close to 1 and for large enough T, $(\hat{q}_{Y|X=x_1}(\alpha_T), \dots, \hat{q}_{Y|X=x_k}(\alpha_T))$ is a sample from k independent and normally-distributed random variables with means $(\mu_1, \dots, \mu_k)$ and variances $(\hat{\sigma}_1^2, \dots, \hat{\sigma}_k^2)$. Our null and alternative hypotheses are then equivalent to $H_0 : \mu_1 = \dots = \mu_k$ and $H_1 : \mu_1 \leqslant \dots \leqslant \mu_k$ with at least one strict inequality. Note that having observed $\hat{q}_{Y|X=x_j}(\alpha_T)$ and $\hat{q}_{Y|X=x_j}(\beta_T)$, the variances $\hat{\sigma}_j^2$ are known. So the implications of our multiple equilibrium model $(\mathcal{E}_{XU}, P_X, F_{U|X})$ can be restated in terms of the means $(\mu_1, \dots, \mu_k)$ of k independent Gaussian random variables with known variances.

A likelihood-ratio (LR) test of H_0 against H_1 is now available from the existing literature (see, e.g., Bartholomew, 1959a,b; Barlow, Bartholomew, Bremner, and Brunk, 1972; Robertson and Wegman, 1978). We shall review Barholomew's results, as they are instrumental in showing how the extension of exact results works in the asymptotic case.

We introduce the following notation: $\hat{q} \equiv (\hat{q}_{Y|X=x_1}(\alpha_T), \dots, \hat{q}_{Y|X=x_k}(\alpha_T))'$, $\mu \equiv (\mu_1, \dots, \mu_k)'$ and $\hat{\Sigma} \equiv \text{diag}(\hat{\sigma}_1^2, \dots, \hat{\sigma}_k^2)$. Hence, for a given value of T, the k-vector $\hat{q}$ is multivariate normal with mean μ and diagonal covariance matrix $\hat{\Sigma}$. Letting A be

a $(k-1) \times k$-matrix defined as:

$$
A \equiv \begin{bmatrix} 1 & -1 & & (0) \\ & \ddots & \ddots & \\ (0) & & 1 & -1 \end{bmatrix},
$$

we can write the null and the alternative hypotheses as:

$$
(15) \qquad H_0 : \{A\mu = 0\} \text{ against } H_1 : \{A\mu \preccurlyeq 0 \text{ and } A\mu \neq 0\} \,,
$$

where the inequalities $\preccurlyeq$ and $\succcurlyeq$ are understood as component wise.

The test in Equation (15) is based on the likelihood-ratio statistic:

$$
(16) \qquad \hat{\xi}_{LR} \equiv -2\ln \frac{\max\limits_{A\mu=0} L(\hat{q}|\mu, \hat{\Sigma})}{\max\limits_{A\mu \preccurlyeq 0} L(\hat{q}|\mu, \hat{\Sigma})},
$$

where $L(\hat{q}|\mu, \hat{\Sigma})$ is the likelihood function:

$$
(17) \qquad L(\hat{q}|\mu, \hat{\Sigma}) = \frac{1}{(2\pi)^{k/2}(\det \hat{\Sigma})^{1/2}} \exp\left[-(\hat{q} - \mu)'\hat{\Sigma}^{-1}(\hat{q} - \mu) \right].
$$

Combining Equations (16) and (17) then yields:

$$
(18) \qquad \hat{\xi}_{LR} = \min_{A\mu=0}(\hat{q} - \mu)'\hat{\Sigma}^{-1}(\hat{q} - \mu) - \min_{A\mu \preccurlyeq 0}(\hat{q} - \mu)'\hat{\Sigma}^{-1}(\hat{q} - \mu).
$$

Barlow, Bartholomew, Bremner, and Brunk (1972) show that the test statistic in Equation (18)—similar to the χ^2 statistic used to test H_0 against the most general form of alternative $H_2 : \mu_1 \neq \ldots \neq \mu_k$—is a weighted average of χ^2 distributions with d degrees of freedom (χ_d^2) with $0 \leqslant d \leqslant k-1$, and is denoted $\overline{\chi}_k^2$ (χ_0^2 denotes a point mass at 0).

The $\overline{\chi}_k^2$ distribution of the likelihood-ratio test statistic $\hat{\xi}_{LR}$ depends on the number of quantiles being compared k, as well as their variances $\hat{\sigma}_j^2$ through the probability

weights attached to each distribution χ_d^2. For example, when $k = 2$ and 3, the distribution of $\hat{\xi}_{LR}$ is given by:

$$
(19) \qquad\qquad\qquad\qquad \hat{\xi}_{LR} \overset{d}{=} \frac{1}{2}\chi^2(0) + \frac{1}{2}\chi^2(1), \text{ for } k \ = \ 2,
$$

$$
(20) \qquad \hat{\xi}_{LR} \overset{d}{=} \frac{\hat{\alpha}}{2\pi}\chi^2(0) + \frac{1}{2}\chi^2(1) + \left[\frac{1}{2} - \frac{\hat{\alpha}}{2\pi}\right]\chi^2(2), \text{ for } k \ = \ 3,
$$

and $\hat{\alpha} \equiv \arccos\left[\hat{\sigma}_2^2 / \sqrt{(\hat{\sigma}_1^2 + \hat{\sigma}_2^2)(\hat{\sigma}_2^2 + \hat{\sigma}_3^2)}\right]$ is a constant, $-\pi < \hat{\alpha} < \pi$.

In the special case where the variances $\hat{\sigma}_j^2$ are equal, Bartholomew (1959b) computes the $\overline{\chi}_k^2$ critical values for a number of values for k $(2 \leqslant k \leqslant 12)$. When the variances are different, exact critical values for $\overline{\chi}_k^2$ are hard to compute analytically if $k \geqslant 5$, though there is no difficulty in obtaining their numerical values for any k (see, e.g., Barlow, Bartholomew, Bremner, and Brunk, 1972). Stochastic upper and lower bounds for the distribution of $\hat{\xi}_{LR}$ have been obtained by Robertson and Wright (1982) and Dardanoni and Forcina (1998).

6.3. **Asymptotic Test.** We shall now derive a test for the implication obtained in Corollary 3, where normality is only asymptotic. Using the notation of Section 6.2, the k-vector $\hat{\Sigma}^{-1/2}(\hat{q} - \mu)$ is asymptotically multivariate normal with mean vector 0_k and identity covariance matrix Id_k.

Note that the standard way of dealing with asymptotically valid order restriction tests (see, e.g., Gourieroux, Holly, and Monfort, 1982; Kodde and Palm, 1986) does not apply here, as the components of the scaling matrix $\hat{\Sigma}$ are not all proportional to $T^{-1/2}$. In order to make sure that $\hat{\Sigma}$ does not become ill-scaled as T gets large— that some of the variance terms $\hat{\sigma}_j$ become infinitely large compared to others—we assume the following:

Assumption S5. Say that a multiple equilibrium model $(\mathcal{E}_{XU}, P_X, F_{U|X})$ satisfies S5 if it satisfies S4 and, in addition, U is independent of X.

When the distribution $F_{U|X}$ does not depend on X, the same is true for the quantities involved in the previously derived limit results. In particular, under S5 we have that $a_x^U \circ \bar{F}_{U|X=x}^{-1} = a^U \circ \bar{F}_U^{-1}$ for all $x \in \mathcal{X}$, so:

$$
\begin{aligned}
\frac{\sigma_j}{\sigma_i} &= \frac{a_{x_j}^Y(q_{Y|X=x_j}(\alpha_T))/\sqrt{c_j(1-\alpha_T)T}}{a_{x_i}^Y(q_{Y|X=x_i}(\alpha_T))/\sqrt{c_i(1-\alpha_T)T}} \\
&= \sqrt{\frac{c_i}{c_j}} \frac{a_{x_j}^Y \circ \bar{F}_{Y|X=x_j}^{-1}(\alpha_T)}{a_{x_i}^Y \circ \bar{F}_{Y|X=x_i}^{-1}(\alpha_T)} \\
(21) \qquad\qquad &\sim \sqrt{\frac{c_i}{c_j}} \frac{g_{N_{x_i}}(x_i)}{g_{N_{x_j}}(x_j)} \quad \text{as } \alpha_T \to 1,
\end{aligned}
$$

where the last equality uses the asymptotic proportionality of $a_X^Y \circ \bar{F}_{Y|X}$ and $a^U \circ \bar{F}_U$ that was established in Lemma 2 (ii).

Now, consider again the limit results derived in Corollary 3. The asymptotic equivalence result established in Equation (21) guarantees that the scaling constants σ_j that control how fast the empirical quantiles $\hat{q}_{Y|X=x_j}(\alpha_T)$ converge to the true quantiles $q_{Y|X=x_j}(\alpha_T)$, are all of the same size. In that case, we have have the following result.

Theorem 4. *Assume* $(\mathcal{E}_{XU}, P_X, F_{U|X})$ *satisfies S5. If for* $T \geqslant 1$, $(Y_{j,1}, \ldots, Y_{j,c_jT})$, $j = 1, \ldots, k$, *are* k *independent random samples of size* c_jT *drawn from* $F_{Y|X=x_j}$, $x_j \in \mathcal{X}$, *then as* $T \to \infty$, *the likelihood-ratio statistic* $\hat{\xi}_{LR}$ *is asymptotically distributed as* $\bar{\chi}^2{}_k$, *with weights that are consistently estimated by weights obtained in the exact Gaussian case.*

For example, when $k = 2$ and 3, the asymptotic distribution of $\hat{\xi}_{LR}$ is that derived in Equations (19)-(20).

It is worth pointing out that the conclusion of Theorem 4 remains valid if, instead of being independent of X, the distribution $F_{U|X}$ is such that for any $(x, x') \in \mathcal{X}^2$

we have:

$$(22) \quad \lim_{v \downarrow 0} \frac{f_{U|X=x}(\bar{F}_{U|X=x}^{-1}(v))}{f_{U|X=x'}(\bar{F}_{U|X=x'}^{-1}(v))} \text{ exists, is strictly positive and independent of } (x, x').$$

The requirement in Equation (22) is weaker than that of independence, since it only restricts the behavior of the auxiliary function a_X^U evaluated at the tail quantiles $\bar{F}_{U|X}^{-1}$. In particular, if the auxiliary function a_X^U is constant, as in the case of exponentially distributed random variables, the requirement in (22) holds without imposing independence of U and X.

6.4. **Limitations.** Finally, we briefly discuss some important limitations of our approach. First, note that since we only use the MCS prediction to derive our test, the latter is based solely on the extremal conditional quantiles of Y given X. Our approach is not equilibria other than the largest and the smallest one, thus our test is not based on the observations that would correspond to those equilibria. While this allows us to cover a variety of models in which MCS holds, it comes at a cost in terms of the power of our test. Indeed, if one knows that the MCS prediction needs to hold at equilibria other than the largest and the smallest one (e.g. all the stable equilibria), then incorporating this additional information in the testing procedure will likely result in improved power properties since the test will now be based on a larger set of observations.

Second, the implementation of our test requires a choice for the probability α_T that satisfies the conditions in Equation (14). Since there are a number of possible choices for α_T, a data-dependent rule for determination of an "optimal" α_T would be useful in practice. This, however, is likely to be quite a difficult problem, akin to developing rigorous rules for bandwidth selection in nonparametric methods. We therefore leave this issue for future research.

7. CONCLUSION

In this paper we design an econometric test for monotone comparative statics predictions suited for testing models with multiple equilibria. Our approach may be characterized as nonparametric as we do not make assumptions on the cardinality, location or probabilities over equilibria. In particular, one can implement our test without assuming an equilibrium selection rule.

First, we show how monotone comparative statics predictions translate into observable implications on the distribution of the dependent variable. In particular, we show that high enough conditional quantiles of the dependent variable increase when the explanatory variable increases.

Second, we construct a likelihood-ratio test for equality of high conditional quantiles against an ordered alternative, as predicted by the monotone comparative statics arguments. The test is an asymptotic extension of the "chi-bar squared" test. Even though the focus of this paper is on quantiles with probabilities close to one, all of our results—when properly transposed—continue to hold for probabilities close to zero.

Finally, we point out some extensions: our likelihood-ratio test can be accommodated to test other hypotheses of interest, such as the unrestricted order among conditional quantiles. Provided that quantile probabilities increase towards one at the same speed as the sample size—which would satisfy the requirement of "large enough" quantile in our paper—this would give rise to other limit distributions. It would be interesting to compare our existing test with one based on such extreme conditional quantiles. In order to carry out our likelihood-ratio test, we needed to eliminate the nuisance parameters—quantile variances—by replacing them with their probability limits. An alternative approach is to use the asymptotic distribution results of the quantile spacings and derive a better approximation to standardized

quantiles in the small sample. Finally, a regression-based approach—in which the conditional quantile is modeled as a function of the explanatory variable—would offer an interesting alternative way of testing the monotonicity prediction.

Appendix A. Proofs of results stated in the text

Proof of Lemma 1. Fix $x \in \mathcal{X}$ and assume $\bar{F}_{U|X=x} \in \mathcal{R}_{-\infty}$ at ∞. Let $R_x : \mathbb{R}_{++} \to \mathbb{R}_{++}$ be given by

$$(23) \qquad R_x(y) \equiv \pi_{xN_x} \frac{\bar{F}_{U|X=x}\big(y/g_{N_x}(x)\big)}{\bar{F}_{Y|X=x}(y)}.$$

Note that R_x is well defined, as from item (iv) in Definition 1 and Equation (2) we know $\bar{F}_{Y|X=x}(y) > 0$, for all $y > 0$. Moreover,

$$(24) \qquad R_x(y) = \frac{\pi_{xN_x} \bar{F}_{U|X=x}\big(y/g_{N_x}(x)\big)}{\pi_{xN_x} \bar{F}_{U|X=x}\big(y/g_{N_x}(x)\big) \left[1 + \displaystyle\sum_{1 \leqslant i < N_x} \frac{\pi_{xi} \bar{F}_{U|X=x}\big(y/g_i(x)\big)}{\pi_{xN_x} \bar{F}_{U|X=x}\big(y/g_{N_x}(x)\big)} \right]}.$$

Given that $\bar{F}_{U|X=x}$ is $(-\infty)$-varying at ∞, we have

$$(25) \qquad \lim_{y\to\infty} \frac{\bar{F}_{U|X=x}\big(y/g_i(x)\big)}{\bar{F}_{U|X=x}\big(y/g_{N_x}(x)\big)} = \lim_{z\to\infty} \frac{\bar{F}_{U|X=x}\big(z g_{N_x}(x)/g_i(x)\big)}{\bar{F}_{U|X=x}(z)} = 0,$$

with $z = y/g_{N_x}(x)$. Moreover, from item (iii) in Definition 1 we know that π_{xi}/π_{xN_x} is bounded, so

$$(26) \qquad \lim_{y\to\infty} R_x(y) = 1,$$

and

$$(27) \qquad \bar{F}_{Y|X=x}(y) \sim \pi_{xN_x} \bar{F}_{U|X=x}\big(y/g_{N_x}(x)\big) \text{ as } y \to \infty.$$

¿From Equation (27), we have

$$\lim_{y\to\infty} \frac{\bar{F}_{Y|X=x}(\lambda y)}{\bar{F}_{Y|X=x}(y)} = \lim_{y\to\infty} \frac{\bar{F}_{U|X=x}\big(\lambda y/g_{N_x}(x)\big)}{\bar{F}_{U|X=x}\big(y/g_{N_x}(x)\big)} = \lim_{z\to\infty} \frac{\bar{F}_{U|X=x}(\lambda z)}{\bar{F}_{U|X=x}(z)},$$

so $\bar{F}_{Y|X=x} \in \mathcal{R}_{-\infty}$ at ∞, which together with Equation (27) shows that item (i) holds.

We shall now prove item (ii). Using Lemma 5, $\bar{F}_{U|X=x}^{-1}$ is 0-varying at 0: for $\lambda > 0$,

$$(28) \qquad \lim_{v \downarrow 0} \frac{\bar{F}_{U|X=x}^{-1}(\lambda v)}{\bar{F}_{U|X=x}^{-1}(v)} = 1.$$

On the other hand, $\lim_{y \to \infty} \bar{F}_{Y|X=x}(y) = 0$ and Equation (26) together imply that $\lim_{y \to \infty} \bar{F}_{Y|X=x}(y) R_x(y) / \pi_{xN_x} = 0$, and

$$\lim_{y \to \infty} \frac{\bar{F}_{Y|X=x}(y) R_x(y) / \pi_{xN_x}}{\bar{F}_{Y|X=x}(y)} = \frac{1}{\pi_{xN_x}}.$$

That $\bar{F}_{U|X=x}^{-1}$ is 0-varying at 0 then implies, by Lemma 6,

$$(29) \qquad \lim_{y \to \infty} \frac{\bar{F}_{U|X=x}^{-1}\big(\bar{F}_{Y|X=x}(y) R_x(y) / \pi_{xN_x}\big)}{\bar{F}_{U|X=x}^{-1}\big(\bar{F}_{Y|X=x}(y)\big)} = [\pi_{xN_x}]^0 = 1.$$

Now, using the definition of $R_x(y)$ in Equation (23), we have

$$(30) \qquad \frac{\bar{F}_{U|X=x}^{-1}\big(\bar{F}_{Y|X=x}(y) R_x(y) / \pi_{xN_x}\big)}{\bar{F}_{U|X=x}^{-1}\big(\bar{F}_{Y|X=x}(y)\big)} = \frac{y/g_{xN_x}(x)}{\bar{F}_{U|X=x}^{-1}\big(\bar{F}_{Y|X=x}(y)\big)},$$

so Equation (29) implies that $y/g_{xN_x}(x) \sim \bar{F}_{U|X=x}^{-1}\big(\bar{F}_{Y|X=x}(y)\big)$ as y goes to ∞. Letting $v = \bar{F}_{Y|X=x}(y)$ we then have

$$(31) \qquad \bar{F}_{Y|X=x}^{-1}(v) \sim g_{xN_x}(x) \bar{F}_{U|X=x}^{-1}(v) \text{ as } v \downarrow 0.$$

Equations (28) and (31) give

$$\lim_{v \downarrow 0} \frac{\bar{F}_{Y|X=x}^{-1}(\lambda v)}{\bar{F}_{Y|X=x}^{-1}(v)} = 1 \text{ for } \lambda > 0,$$

so $\bar{F}_{Y|X=x}^{-1}$ is 0-varying at 0 which together with Equations (28) and (31) shows (ii), and thus completes the proof of Lemma 1. $\qquad\qquad\square$

Proof of Theorem 1. The proof of Theorem 1 follows from Lemma 1 easily by the argument used in Section 3. We include it here for completeness. Consider $(x_1, x_2) \in \mathcal{X}^2$ such that $x_1 < x_2$. From Lemma 1(i),

$$\frac{\bar{F}_{Y|X=x_1}(y)}{\bar{F}_{Y|X=x_2}(y)} \sim \frac{\pi_{x_1 N_{x_1}} \bar{F}_{U|X=x_1}\big(y/g_{N_{x_1}}(x_1)\big)}{\pi_{x_2 N_{x_2}} \bar{F}_{U|X=x_2}\big(y/g_{N_{x_2}}(x_2)\big)}, \text{ as } y \to \infty.$$

Now note that

$$(32) \quad \frac{\bar{F}_{U|X=x_1}\big(y/g_{N_{x_1}}(x_1)\big)}{\bar{F}_{U|X=x_2}\big(y/g_{N_{x_2}}(x_2)\big)} = \frac{\bar{F}_{U|X=x_1}\big(y/g_{N_{x_1}}(x_1)\big)}{\bar{F}_{U|X=x_1}\big(y/g_{N_{x_2}}(x_2)\big)} \frac{\bar{F}_{U|X=x_1}\big(y/g_{N_{x_2}}(x_2)\big)}{\bar{F}_{U|X=x_2}\big(y/g_{N_{x_2}}(x_2)\big)}.$$

From item (ii) in Definition 1 we have $g_{N_{x_2}}(x_2) > g_{N_{x_1}}(x_1)$, and by assumption S1 $\bar{F}_{U|X=x_1} \in \mathcal{R}_{-\infty}$ at ∞, so

$$\lim_{y \to \infty} \frac{\bar{F}_{U|X=x_1}\big(y/g_{N_{x_1}}(x_1)\big)}{\bar{F}_{U|X=x_1}\big(y/g_{N_{x_2}}(x_2)\big)} = \lim_{z \to \infty} \frac{\bar{F}_{U|X=x_1}\big(z g_{N_{x_2}}(x_2)/g_{N_{x_1}}(x_1)\big)}{\bar{F}_{U|X=x_1}(z)} = 0,$$

where $z \equiv y/g_{N_{x_2}}(x_2)$. So the first term on the right-hand side of Equation (32) goes to 0 as y gets large. From Property (12) in Assumption S2 and given $x_1 < x_2$, we know that the second term of the right-hand side of Equation (32) is bounded as y increases. Finally, we know that $\pi_{x_1 N_{x_1}}/\pi_{x_2 N_{x_2}} < \infty$ since from item (iii) in Definition 1 $\pi_{x_2 N_{x_2}} > 0$. Combining the facts above,

$$\lim_{y \to \infty} \frac{\bar{F}_{Y|X=x_1}(y)}{\bar{F}_{Y|X=x_2}(y)} = 0,$$

so there is $\bar{y}_1 > 0$ such that, if $y \geqslant \bar{y}_1$ then $\bar{F}_{Y|X=x_1}(y) < \bar{F}_{Y|X=x_2}(y)$. Since $\mathcal{X}$ is finite, there is $\bar{y}$ such that if $y \geqslant \bar{y}$ then $\bar{F}_{Y|X=x}(y) < \bar{F}_{Y|X=x'}(y)$ for all $(x, x') \in \mathcal{X}^2$ with $x < x'$. Note that for any $x \in \mathcal{X}$ and $v \in (0, 1)$, $\bar{F}_{Y|X=x}^{-1}(v) = q_{Y|X=x}(1 - v)$. From the above we know that, for any $(x_1, x_2) \in \mathcal{X}^2$ such that $x_1 < x_2$, there is $\underline{v}_1 \in (0, 1)$ such that if $v \leqslant \underline{v}_1$ then $q_{Y|X=x_1}(1 - v) < q_{Y|X=x_2}(1 - v)$. Equivalently, letting $\bar{\alpha}_1 \equiv 1 - \underline{v}_1$, for $\alpha \in [\bar{\alpha}_1, 1)$ we have $q_{Y|X=x_1}(\alpha) < q_{Y|X=x_2}(\alpha)$. $\mathcal{X}$ being finite guarantees that the result holds for any $(x, x') \in \mathcal{X}^2$ by the same reasoning as above. $\square$

Proof of Lemma 2. Fix $x \in \mathcal{X}$ and assume $\bar{F}_{U|X=x}$ is in Γ with auxiliary function a_x^U; for R_x defined in Equation (23) we have:

$$\frac{R_x\big(g_{N_x}(x)y + g_{N_x}(x)\lambda a_x^U(y)\big)}{R_x\big(g_{N_x}(x)y\big)} =$$

(33)
$$\left[\frac{\pi_{xN_x}\bar{F}_{U|X=x}\big(y + \lambda a_x^U(y)\big)}{\bar{F}_{Y|X=x}\big(g_{N_x}(x)y + g_{N_x}(x)\lambda a_x^U(y)\big)}\right] \left[\frac{\bar{F}_{Y|X=x}\big(g_{N_x}(x)y\big)}{\pi_{xN_x}\bar{F}_{U|X=x}(y)}\right].$$

From Equation (26), the left-hand side in Equation (33) converges to 1 as $y \to \infty$. On the other hand,

$$\lim_{y \to \infty} \frac{\pi_{xN_x}\bar{F}_{U|X=x}\big(y + \lambda a_x^U(y)\big)}{\pi_{xN_x}\bar{F}_{U|X=x}(y)} = \exp(-\lambda),$$

since a_x^U is the auxiliary function of $\bar{F}_{U|X=x}$. Then we have:

$$
\begin{aligned}
\exp(\lambda) &= \lim_{y \to \infty} \frac{\bar{F}_{Y|X=x}\big(g_{N_x}(x)y\big)}{\bar{F}_{Y|X=x}\big(g_{N_x}(x)y + g_{N_x}(x)\lambda a_x^U(y)\big)} \\
&= \lim_{z \to \infty} \frac{\bar{F}_{Y|X=x}(z)}{\bar{F}_{Y|X=x}\big(z + g_{N_x}(x)\lambda a_x^U(z/g_{N_x}(x))\big)},
\end{aligned}
$$

using the change of variable $z \equiv g_{N_x}(x)y$. Hence $\bar{F}_{Y|X=x} \in \Gamma$:

(34)
$$\lim_{y \to \infty} \frac{\bar{F}_{Y|X=x}\big(y + \lambda a_x^Y(y)\big)}{\bar{F}_{Y|X=x}(y)} = \exp(-\lambda),$$

with auxiliary function a_x^Y defined as $a_x^Y(y) \equiv g_{N_x}(x)a_x^U(y/g_{N_x}(x))$ for all $y > 0$, which shows item (i).

In order to show item (ii) we exploit the fact that for any sequence $\{\varphi_s\}_{s>0}$ of monotone increasing functions $\varphi_s : \mathbb{R}^+ \to (0,1)$, $\lim_{s\to\infty}\varphi_s(x) = \varphi(x)$ for all continuity points $x > 0$ of φ, implies $\lim_{s\to\infty}\varphi_s^{-1}(z) = \varphi^{-1}(z)$ for all continuity points $z \in (0,1)$ of φ^{-1} (see, e.g., Lemma 1.9 in de Haan, 1974). Let then

$$\varphi_s(y) \equiv 1 - \frac{\bar{F}_{U|X=x}(s + y\, a_x^U(s))}{\bar{F}_{U|X=x}(s)} \quad \text{for all } y > 0.$$

That $\bar{F}_{U|X=x} \in \Gamma$ implies $\lim_{s\to\infty} \varphi_s(y) = 1 - \exp(-y)$ for all $y > 0$. Letting $\varphi(y) \equiv 1 - \exp(-y)$, we then have for $t \in (0,1)$:

$$\lim_{s\to\infty} \frac{\bar{F}_{U|X=x}^{-1}\big((1-t)\bar{F}_{U|X=x}(s)\big) - s}{a_x^U(s)} = \lim_{s\to\infty} \varphi_s^{-1}(t) = \varphi^{-1}(t) = -\ln(1-t).$$

Letting $v \equiv \bar{F}_{U|X=x}(s)$ and $\mu \equiv 1 - t$ gives:

$$(35) \qquad \lim_{v\downarrow 0} \frac{\bar{F}_{U|X=x}^{-1}(v\mu) - \bar{F}_{U|X=x}^{-1}(v)}{a_x^U\big(\bar{F}_{U|X=x}^{-1}(v)\big)} = -\ln\mu \text{ for } \mu \in (0,1).$$

Thus $\bar{F}_{U|X=x}^{-1} \in \Pi$ as in Definition 4 with auxiliary function $a_x^U \circ \bar{F}_{U|X=x}^{-1}$.

Moreover, for any $\lambda > 0$, letting $\mu \equiv \lambda$ and $\nu \equiv \lambda^{-1}$ we have:

$$\frac{a_x^U\big(\bar{F}_{U|X=x}^{-1}(\lambda v)\big)}{a_x^U\big(\bar{F}_{U|X=x}^{-1}(v)\big)} =$$

$$(36) \qquad -\left[\frac{\bar{F}_{U|X=x}^{-1}(\mu v) - \bar{F}_{U|X=x}^{-1}(v)}{a_x^U\big(\bar{F}_{U|X=x}^{-1}(v)\big)}\right]\left[\frac{a_x^U\big(\bar{F}_{U|X=x}^{-1}(\lambda v)\big)}{\bar{F}_{U|X=x}^{-1}(\lambda\nu v) - \bar{F}_{U|X=x}^{-1}(\lambda v)}\right].$$

Equations (35) and (36) together imply:

$$(37) \qquad \lim_{v\downarrow 0} \frac{a_x^U \circ \bar{F}_{U|X=x}^{-1}(\lambda v)}{a_x^U \circ \bar{F}_{U|X=x}^{-1}(v)} = \frac{\ln\mu}{-\ln\nu} = 1,$$

so $a_x^U \circ \bar{F}_{U|X=x}^{-1} \in \mathcal{R}_0$ at 0. We now study $\bar{F}_{Y|X=x}$: if for any $x \in \mathcal{X}$, we let $\varphi_{x,s}(y) \equiv 1 - \bar{F}_{Y|X=x}\big(s + y\, a_x^Y(s)\big)/\bar{F}_{Y|X=x}(s)$ for all $y > 0$, we have $\lim_{s\to\infty} \varphi_{x,s}(y) = \varphi(y)$. Same reasoning as previously then implies:

$$(38) \qquad \lim_{v\downarrow 0} \frac{\bar{F}_{Y|X=x}^{-1}(v\mu) - \bar{F}_{Y|X=x}^{-1}(v)}{a_x^Y\big(\bar{F}_{Y|X=x}^{-1}(v)\big)} = -\ln\mu \text{ for } \mu \in (0,1).$$

So $\bar{F}_{Y|X=x}^{-1} \in \Pi$ as in Definition 4 with auxiliary function $a_x^Y \circ \bar{F}_{Y|X=x}^{-1}$. Equation (38) and the fact that:

$$\frac{a_x^Y\big(\bar{F}_{Y|X=x}^{-1}(\lambda v)\big)}{a_x^Y\big(\bar{F}_{Y|X=x}^{-1}(v)\big)} = -\left[\frac{\bar{F}_{Y|X=x}^{-1}(\mu v) - \bar{F}_{Y|X=x}^{-1}(v)}{a_x^Y\big(\bar{F}_{Y|X=x}^{-1}(v)\big)}\right]\left[\frac{a_x^Y\big(\bar{F}_{Y|X=x}^{-1}(\lambda v)\big)}{\bar{F}_{Y|X=x}^{-1}(\lambda\nu v) - \bar{F}_{Y|X=x}^{-1}(\lambda v)}\right],$$

with $\lambda > 0$, $\mu \equiv \lambda$ and $\nu \equiv \lambda^{-1}$, then imply that $a_x^Y \circ \bar{F}_{Y|X=x}^{-1} \in \mathcal{R}_0$ at 0.

Given Equation (31) and the definition of a_x^Y it is not surprising to see that $a_x^Y\left(\bar{F}_{Y|X=x}^{-1}(v)\right) \sim g_{N_x}(x) a_x^U\left(\bar{F}_{U|X=x}^{-1}(v)\right)$ as $v \downarrow 0$; however we need a formal proof of that statement. We start by showing that:

$$(39) \qquad \lim_{s\to\infty} F_{Y|X=x}^s(A_s\lambda + b_s) = \exp[-\exp(-\lambda)],$$

with $A_s \equiv a_x^Y(b_s)$ and $b_s \equiv \bar{F}_{Y|X=x}^{-1}(1/s)$. In Equation (34) let $y \equiv \bar{F}_{Y|X=x}^{-1}(1/s)$ so $y \to \infty$ as $s \to \infty$; then

$$\lim_{s\to\infty} \frac{\bar{F}_{Y|X=x}\left(\bar{F}_{Y|X=x}^{-1}(1/s) + \lambda a_x^Y\left(\bar{F}_{Y|X=x}^{-1}(1/s)\right)\right)}{\bar{F}_{Y|X=x}\left(\bar{F}_{Y|X=x}^{-1}(1/s)\right)}$$

$$= \lim_{s\to\infty} s\bar{F}_{Y|X=x}\left(a_x^Y\left(\bar{F}_{Y|X=x}^{-1}(1/s)\right)\lambda + \bar{F}_{Y|X=x}^{-1}(1/s)\right)$$

$$(40) \qquad = \exp(-\lambda).$$

Let $b_s \equiv \bar{F}_{Y|X=x}^{-1}(1/s)$ and $A_s \equiv a_x^Y(b_s)$; the last equality in Equation (40) together with Lemma 2.2.2 in de Haan (1970) then imply Equation (39). We now derive a similar equality involving $F_{U|X=x}$: the last equality in Equation (40) and the tail equivalence property in Equation (27) together imply:

$$\lim_{s\to\infty} s\bar{F}_{U|X=x}\left(\left(A_s/g_{N_x}(x)\right)\lambda + \left(b_s/g_{N_x}(x)\right)\right) = \exp(-\lambda - \ln\pi_{xN_x}).$$

Using again Lemma 2.2.2 in de Haan (1970) then gives:

$$(41) \qquad \lim_{s\to\infty} F_{U|X=x}^s\left(\left(A_s/g_{N_x}(x)\right)\lambda + \left(b_s/g_{N_x}(x)\right)\right) = \exp\left(-\exp(-\lambda - \ln\pi_{xN_x})\right).$$

On the other hand, $\bar{F}_{U|X=x} \in \Gamma$ as in Definition 3 with auxiliary function a_x^U, together with Lemma 2.2.2 in de Haan (1970) imply:

$$(42) \qquad \lim_{s\to\infty} F_{U|X=x}^s(\tilde{A}_s\lambda + \tilde{b}_s) = \exp\left(-\exp(-\lambda)\right),$$

with $\tilde{A}_s \equiv a_x^U(\tilde{b}_s)$ and $\tilde{b}_s \equiv \bar{F}_{U|X=x}^{-1}(1/s)$. Combining Equations (41) and (42) and applying the results of Lemma 2.4.1 in de Haan (1970) on the change of norming

constants (with $A = 1$ and $B = \ln \pi_{xN_x}$), then gives:

$$\frac{\left(A_s/g_{N_x}(x)\right)}{\tilde{A}_s} \to 1 \text{ and } \frac{\left(b_s/g_{N_x}(x)\right) - \tilde{b}_s}{\tilde{A}_s} \to \ln \pi_{xN_x} \text{ as } s \to \infty.$$

So from the first of the above limit results we get:

$$a_x^Y(\bar{F}_{Y|X=x}^{-1}(v)) \sim g_{N_x}(x) a_x^U(\bar{F}_{U|X=x}^{-1}(v)) \text{ as } v \downarrow 0,$$

which completes the proof of item (ii).   □

Proof of Lemma 3. Given a random sample $(Y_1, \ldots, Y_T)$ let $\{Y_{(k)}^{(T)}\}_{k=1}^T$ be the ascending order statistics: $Y_{(1)}^{(T)} \leqslant \ldots \leqslant Y_{(T)}^{(T)}$. Then for any α_T, $0 < \alpha_T < 1$, we have:

$$(43) \qquad\qquad \hat{q}(\alpha_T) = Y_{(m)}^{(T)} \text{ with } m \equiv \lfloor \alpha_T T \rfloor + 1,$$

where $\lfloor \cdot \rfloor$ denotes the greatest integer function, $\lfloor x \rfloor \equiv \max\{n \in \mathbb{N} : n \leqslant x\}$ for $x > 0$. Note that m depends on T. Similarly, for β_T: $\hat{q}(\beta_T) = Y_{(k)}^{(T)}$ where $k \equiv \lfloor \beta_T T \rfloor + 1$.

First we record the following facts, which follow trivially from the definition of m and the hypotheses on α_T in the theorem:

$$(44) \qquad\qquad \lim_{T \to \infty} T - m = \infty,$$

$$(45) \qquad\qquad \lim_{T \to \infty} \frac{T - m}{T} = 0,$$

$$(46) \qquad\qquad \lim_{T \to \infty} \frac{T - m}{(1 - \alpha_T)T} = \lim_{T \to \infty} \frac{T - m + 1}{(1 - \alpha_T)T} = 1.$$

The hypotheses on β_T imply properties (44), (45), and (46) for k, and, in addition, that

$$(47) \qquad\qquad \lim_{T \to \infty} \frac{T - m}{T - k} = \rho.$$

Second, we have:

$$\sqrt{T(1-\alpha_T)} \left[\frac{\hat{q}(\alpha_T) - q(\alpha_T)}{a(q(\alpha_T))} \right]$$

$$= \sqrt{T(1-\alpha_T)} \left[\frac{Y^{(T)}_{(\lfloor \alpha_T T \rfloor + 1)} - \bar{F}^{-1}(1-\alpha_T)}{a(\bar{F}^{-1}(1-\alpha_T))} \right]$$

$$= \sqrt{T-m+1} \left[\frac{Y^{(T)}_{(m)} - \bar{F}^{-1}((T-m)/T)}{a(\bar{F}^{-1}((T-m)/T))} + \frac{\bar{F}^{-1}((T-m)/T) - \bar{F}^{-1}(1-\alpha_T)}{a(\bar{F}^{-1}((T-m)/T))} \right]$$

$$(48) \qquad \cdot \frac{a(\bar{F}^{-1}((T-m)/T))}{a(\bar{F}^{-1}(1-\alpha_T))} \sqrt{\frac{T(1-\alpha_T)}{T-m+1}},$$

and

$$\frac{\hat{q}(\beta_T) - \hat{q}(\alpha_T)}{a(q(\alpha_T))} - \ln \rho$$

$$= \left\{ \left[\frac{Y^{(T)}_{(k)} - Y^{(T)}_{(m)}}{a(Y^{(T)}_{(m)})} - \ln \rho \right] \frac{a(Y^{(T)}_{(m)})}{a(\bar{F}^{-1}((T-m)/T))} \right.$$

$$(49) \qquad \left. + \ln \rho \left[\frac{a(Y^{(T)}_{(m)})}{a(\bar{F}^{-1}((T-m)/T))} - 1 \right] \right\} \frac{a(\bar{F}^{-1}((T-m)/T))}{a(\bar{F}^{-1}(1-\alpha_T))}.$$

The proof of the theorem is done in three steps. We first show (STEP 1) that:

$$(50) \qquad \sqrt{T-m+1} \left[\frac{Y^{(T)}_{(m)} - \bar{F}^{-1}((T-m)/T)}{a(\bar{F}^{-1}((T-m)/T))} \right] \xrightarrow{d} \mathcal{N},$$

$$(51) \qquad \frac{Y^{(T)}_{(k)} - Y^{(T)}_{(m)}}{a(Y^{(T)}_{(m)})} \xrightarrow{P} \ln \rho,$$

where $\mathcal{N}$ is a standard Gaussian random variable. We then show (STEP 2):

$$(52) \qquad \lim_{T \to \infty} \sqrt{T-m+1} \left[\frac{\bar{F}^{-1}((T-m)/T) - \bar{F}^{-1}(1-\alpha_T)}{a(\bar{F}^{-1}((T-m)/T))} \right] = 0$$

$$(53) \qquad \lim_{T \to \infty} \frac{a(\bar{F}^{-1}((T-m)/T))}{a(\bar{F}^{-1}(1-\alpha_T))} = 1.$$

Finally, we show (STEP 3):

$$(54) \qquad \frac{a(Y^{(T)}_{(m)})}{a(\bar{F}^{-1}((T-m)/T))} \xrightarrow{P} 1.$$

The first limit result of Lemma 3 then follows from (48) by (50), (52) and (53) using (46) and Lemma 2.4.1 in de Haan (1970). The second limit result in Lemma 3 follows from (49) by (51) and (54) using (53), (46), and Slutsky's Theorem.

STEP 1: This step takes a key idea from the proof of Theorem 3.1 in Dekkers and de Haan (1989). Let $A_1, \ldots, A_T$ be independent and identically distributed standard exponential random variables. Let $A_{(1)}^{(T)} \leqslant \ldots \leqslant A_{(T)}^{(T)}$ be the ascending order statistics of $(A_1, \ldots, A_T)$. Then, by using the probability integral transform, we have $\{Y_{(m)}^{(T)}\}_{m=1}^T \stackrel{d}{=} \{\bar{F}^{-1}(\exp(-A_{(m)}^{(T)}))\}_{m=1}^T$.

Now, let $W(x) \equiv \bar{F}^{-1}(\exp(-x))$ for $x > 0$; we have $Y_{(m)}^{(T)} \stackrel{d}{=} W(A_{(m)}^{(T)})$ and $W(\ln(T/(T-m))) = \bar{F}^{-1}((T-m)/T)$. Moreover,

$$a\left(W(x)\right) = \frac{\exp(-x)}{f(W(x))} = W'(x).$$

Let $\eta_T \equiv \ln(T/(T-m))$; then, $a\left(\bar{F}^{-1}\left((T-m)/T\right)\right) = W'(\eta_T)$. So the expression in Statement (50) can be written as:

$$
\begin{aligned}
\sqrt{T-m+1} &\left[\frac{Y_{(m)}^{(T)} - \bar{F}^{-1}((T-m)/T)}{a(\bar{F}^{-1}((T-m)/T))} \right] \\
&\stackrel{d}{=} \sqrt{T-m+1} \left[\frac{W(A_{(m)}^{(T)}) - W(\eta_T)}{W'(\eta_T)} \right] \\
&\stackrel{d}{=} \sqrt{T-m+1} \int_0^{Z_T/\sqrt{T-m+1}} \frac{W'(\eta_T + s)}{W'(\eta_T)} ds,
\end{aligned}
$$

(55)

where $Z_T \equiv \sqrt{T-m+1}[A_{(m)}^{(T)} - \ln(T/(T-m))]$.

Then, by Lemma 10, $Z_T \stackrel{d}{\to} \mathcal{N}_1$ as $T \to \infty$. But Lemma 7(i) and Statement (41) imply that the integrand on the right-hand side of (55) converges uniformly to 1 on compact intervals, as $T \to \infty$. So Lemma 8 implies Statement (50).

The proof of Statement (51) is similar. We have:

$$
\frac{Y_{(k)}^{(T)} - Y_{(m)}^{(T)}}{a(Y_{(m)}^{(T)})} - \ln \rho
$$

$$
\stackrel{d}{=} \frac{W(A_{(k)}^{(T)}) - W(A_{(m)}^{(T)})}{W'(A_{(m)}^{(T)})} - \ln \rho
$$

$$
\stackrel{d}{=} \left[\sqrt{\frac{T-m}{\rho-1}} \int_0^{V_T/\sqrt{\frac{T-m}{\rho-1}}} \frac{W'(A_{(m)}^{(T)} + \ln \rho + s)}{W'(A_{(m)}^{(T)})} ds \right] \sqrt{\frac{\rho-1}{T-m}}
$$

(56)
$$
+ \int_0^{\ln \rho} \left[\frac{W'(A_{(m)}^{(T)} + s)}{W'(A_{(m)}^{(T)})} - 1 \right] ds,
$$

where $V_T \equiv \sqrt{(T-m)/(\rho-1)}[A_{(k)}^{(T)} - A_{(m)}^{(T)} - \ln \rho]$. Note that $\{V_T\}$ and $\{A_{(m)}^{(T)}\}$ are independent (Renyi, 1953) and that $A_{(m)}^{(T)} \stackrel{as}{\to} \infty$ (see, e.g. Theorem 4 in Watts (1980)). By Lemma 10, we have $V_T \stackrel{d}{\to} \mathcal{N}_2$ as $T \to \infty$, and the integrand in the first term of Equation (56) converges uniformly to 1 on compact intervals. Hence, using Lemma 8 and Statement (39) the first term in brackets in Equation (56) converges in distribution. It is multiplied by $[(\rho-1)/(T-m)]^{-1/2}$, which goes to zero. So the first summand of expression (56) converges in probability to 0 (it converges in distribution to the constant 0, so it converges in probability). On the other hand, the second summand in expression (56) converges to 0 a.s.: Note that $A_{(m)}^{(T)} \stackrel{as}{\to} \infty$ a.s. (by, e.g. Theorem 4 in Watts (1980)) and the integrand converges to 0 uniformly on compact intervals (Lemma 7 (i)), so the integral converges to 0 for a full measure of realizations of $\{A_{(m)}^{(T)}\}$. This establishes Statement (51).

STEP 2: We now prove (52) and (53). Using the notation in Step 1:

$$
\begin{aligned}
\sqrt{T-m+1}\; &\left[\frac{\bar{F}^{-1}((T-m)/T)-\bar{F}^{-1}(1-\alpha_T)}{a(\bar{F}^{-1}((T-m)/T))}\right]\\
&= \sqrt{T-m+1}\left[\frac{W(\eta_T)-W(\ln(1/(1-\alpha_T)))}{W'(\eta_T)}\right]\\
&= \sqrt{T-m+1}\int_{-\ln(1-\alpha_T)-\eta_T}^{0}\frac{W'(\eta_T+s)}{W'(\eta_T)}ds\\
&\sim \sqrt{T-m+1}\left[0-\ln\frac{T-m}{(1-\alpha_T)T}\right]\;\text{as } T\to\infty.
\end{aligned}
$$

(57)

The equivalence in Statement (57) follows by exchanging the limit and the integral, using the uniform convergence established in Lemma 7(i), and the fact that Statement (46) implies:

$$
\lim_{T\to\infty}\left[-\ln(1-\alpha_T)-\eta_T\right]=\lim_{T\to\infty}\ln\frac{T-m}{(1-\alpha_T)T}=0.
$$

Using $|\ln\{(T-m)/[(1-\alpha_T)T]\}|\leqslant|(T-m)/[(1-\alpha_T)T]-1|$, we then get:

$$
\begin{aligned}
\sqrt{T-m+1}\left|\ln\frac{T-m}{(1-\alpha_T)T}\right| &\leqslant \sqrt{T-m+1}\left|\frac{T-m}{(1-\alpha_T)T}-1\right|\\
&= \sqrt{T-m+1}\left|\frac{\alpha_T T-\lfloor\alpha_T T\rfloor-1}{(1-\alpha_T)T}\right|\\
&\leqslant 2\frac{\sqrt{T-m+1}}{(1-\alpha_T)T}\to 0,\;\text{as } T\to\infty,
\end{aligned}
$$

where the convergence to 0 follows from Statement (46). By Statement (57), then, this proves (52). To prove (53), note that Lemma 2(ii) implies that $a\circ\bar{F}^{-1}\in\mathcal{R}_0$ at 0. So Statements (45) and (46), and Lemma 6 give (53).

STEP 3: The proof of (54), in turn is similar to that of (50) in Step 1. We have:

$$\frac{a(Y^{(T)}_{(m)})}{a(\bar{F}^{-1}((T-m)/T))} - 1$$

$$\overset{d}{=} \frac{W'(A^{(T)}_{(m)}) - W'(\eta_T)}{W'(\eta_T)}$$

(58)
$$\overset{d}{=} \left[\sqrt{T-m} \int_0^{Z_T/\sqrt{T-m+1}} \frac{W''(\eta_T + s)}{W'(\eta_T)} ds \right] \frac{1}{\sqrt{T-m}},$$

where as previously $Z_T = \sqrt{T-m+1}[A^{(T)}_{(m)} - \ln(T/(T-m))]$. Then, by Lemma 10, $Z_T \overset{d}{\to} \mathcal{N}_1$ as $T \to \infty$. But Lemma 7(ii) implies that the integrand on the right-hand side of (58) converges uniformly to 0 on compact intervals. So Lemma 9 and Statement (39) imply Statement (54). $\square$

Proof of Theorem 2. If $(\mathcal{E}_{XU}, P_X, F_{U|X})$ satisfies S4 then it satisfies S3; hence we can use Lemma 2(i) to show that for any $x_j \in \mathcal{X}$, $1 \leqslant j \leqslant k$, the conditional distribution tails $\bar{F}_{Y|X=x_j} \in \Gamma$. Moreover, from Equation (2) we know that for any $1 \leqslant j \leqslant k$,

$$f_{Y|X=x_j}(y) = \sum_{i=1}^{N_{x_j}} \frac{\pi_{x_j i}}{g_i(x_j)} f_{U|X=x_j}\left(\frac{y}{g_i(x_j)}\right) \text{ for any } y > 0.$$

Under S4 the densities $f_{U|X=x_j}$ are all eventually non-decreasing; hence the same holds for $f_{Y|X=x_j}$. If for each $1 \leqslant j \leqslant k$, we have $0 < \alpha_{T_j} < \beta_{T_j} < 1$, $\lim_{T_j \to \infty} \alpha_{T_j} = 1$, $\lim_{T_j \to \infty} (1 - \alpha_{T_j}) T_j = \infty$ and $\lim_{T_j \to \infty} (1 - \alpha_{T_j})/(1 - \beta_{T_j}) = \rho_j$ with $\rho_j > 1$, then the results of Lemma 3 apply for all $1 \leqslant j \leqslant k$, i.e.

$$\sqrt{T_j(1 - \alpha_{T_j})} \frac{\hat{q}_{Y|X=x_j}(\alpha_{T_j}) - q_{Y|X=x_j}(\alpha_{T_j})}{a_{x_j}^Y(q_{Y|X=x_j}(\alpha_{T_j}))} \overset{d}{\to} \mathcal{N}_j,$$

and

$$\frac{\hat{q}_{Y|X=x_j}(\beta_{T_j}) - \hat{q}_{Y|X=x_j}(\alpha_{T_j})}{a_{x_j}^Y(q_{Y|X=x_j}(\alpha_{T_j}))} \overset{p}{\to} \ln \rho_j,$$

where $a_X^Y \equiv \bar{F}_{Y|X}/f_{Y|X}$ is the auxiliary function of $F_{Y|X}$ and $\mathcal{N}_j$, $1 \leqslant j \leqslant k$, are k independent standard normal random variables. The conclusion Theorem 2 follows by letting $\mu_j \equiv q_{Y|X=x_j}(\alpha_{T_j})$ and $\sigma_j \equiv a_{x_j}^Y(\mu_j)/\sqrt{T_j(1-\alpha_{T_j})}$, and using the independence of different samples $(Y_{j,1}, \ldots, Y_{j,T_j})$. $\qquad\qquad\qquad\square$

Proof of Theorem 4. The proof is done in five steps:

STEP1: we work with the first minimization problem in Equation (18):

$$(59) \qquad \min_{\mu}(\mu - \hat{q})'\hat{\Sigma}^{-1}(\mu - \hat{q}),$$

$$\text{subject to } A\mu = 0.$$

Let $\mathcal{L} : \mathbb{R}^{2k-1} \to \mathbb{R}$ be the corresponding Lagrangian $\mathcal{L}(\mu, \lambda) = (\mu - \hat{q})'\hat{\Sigma}^{-1}(\mu - \hat{q}) + \lambda'A\mu$, where λ denotes the $(k-1)$-vector of Lagrange multipliers (dual variables) associated with the constraint $A\mu = 0$. A is full rank and the (Lagrange) dual function $g : \mathbb{R}^{k-1} \to \mathbb{R} \cup \{-\infty\}$ is $g(\lambda) \equiv \inf_\mu \mathcal{L}(\mu, \lambda) = -\frac{1}{4}\lambda'A\hat{\Sigma}A'\lambda + \lambda'A\hat{q}$. The dual problem is then:

$$(60) \qquad \max_{\lambda} -\frac{1}{4}\lambda'A\hat{\Sigma}A'\lambda + \lambda'A\hat{q},$$

with λ unconstrained. The solutions to the dual and primal problems (60) and (59) are:

$$(61) \qquad \lambda_0 = 2(A\hat{\Sigma}A')^{-1}A\hat{q},$$

$$(62) \qquad \mu_0 = \hat{q} - \frac{1}{2}\hat{\Sigma}A'\lambda_0 = \hat{q} - \hat{\Sigma}A'(A\hat{\Sigma}A')^{-1}A\hat{q},$$

and we have:

$$(63) \quad (\mu_0 - \hat{q})'\hat{\Sigma}^{-1}(\mu_0 - \hat{q}) = \hat{q}'A'(A\hat{\Sigma}A')^{-1}A\hat{q} = -\frac{1}{4}\lambda_0'A\hat{\Sigma}A'\lambda_0 + \lambda_0'A\hat{q}.$$

Similarly, we consider the dual of the second minimization problem in (18):

(64) $$\min_{\mu}(\mu - \hat{q})'\hat{\Sigma}^{-1}(\mu - \hat{q}),$$

subject to $A\mu \preccurlyeq 0$.

The dual is:

(65) $$\max_{\lambda} -\frac{1}{4}\lambda'A\hat{\Sigma}A'\lambda + \lambda'A\hat{q},$$

subject to $\lambda \succcurlyeq 0$.

Letting λ_1 and μ_1 denote the solutions to the dual and primal problems (65) and (64) we again have:

(66) $$(\mu_1 - \hat{q})'\hat{\Sigma}^{-1}(\mu_1 - \hat{q}) = -\frac{1}{4}\lambda_1'A\hat{\Sigma}A'\lambda_1 + \lambda_1'A\hat{q}.$$

STEP 2: using Equations (63) and (66) the likelihood-ratio statistic in (18) then equals:

$$
\begin{aligned}
\hat{\xi}_{LR} &= \max_{\lambda}\left(-\frac{1}{4}\lambda'A\hat{\Sigma}A'\lambda + \lambda'A\hat{q}\right) - \max_{\lambda:\lambda\succcurlyeq 0}\left(-\frac{1}{4}\lambda'A\hat{\Sigma}A'\lambda + \lambda'A\hat{q}\right) \\
&= \max_{\lambda}\left[\hat{q}'\hat{\Sigma}^{-1}\hat{q} - \left(\frac{1}{2}\hat{\Sigma}A'\lambda - \hat{q}\right)'\hat{\Sigma}^{-1}\left(\frac{1}{2}\hat{\Sigma}A'\lambda - \hat{q}\right)\right] \\
&\quad - \max_{\lambda:\lambda\succcurlyeq 0}\left[\hat{q}'\hat{\Sigma}^{-1}\hat{q} - \left(\frac{1}{2}\hat{\Sigma}A'\lambda - \hat{q}\right)'\hat{\Sigma}^{-1}\left(\frac{1}{2}\hat{\Sigma}A'\lambda - \hat{q}\right)\right] \\
&= \min_{\lambda:\lambda\succcurlyeq 0}\left[\left(\frac{1}{2}\hat{\Sigma}A'\lambda - \hat{q}\right)'\hat{\Sigma}^{-1}\left(\frac{1}{2}\hat{\Sigma}A'\lambda - \hat{q}\right)\right] \\
&\quad - \min_{\lambda}\left[\left(\frac{1}{2}\hat{\Sigma}A'\lambda - \hat{q}\right)'\hat{\Sigma}^{-1}\left(\frac{1}{2}\hat{\Sigma}A'\lambda - \hat{q}\right)\right] \\
&= \min_{\lambda:\lambda\succcurlyeq 0}\left[\left(\frac{1}{2}\hat{\Sigma}A'\lambda_0 - \frac{1}{2}\hat{\Sigma}A'\lambda\right)'\hat{\Sigma}^{-1}\left(\frac{1}{2}\hat{\Sigma}A'\lambda_0 - \frac{1}{2}\hat{\Sigma}A'\lambda\right)\right],
\end{aligned}
$$

where the last equality follows by a simple geometric argument. Combining the above with Equations (61)-(62) then gives:

$$\hat{\xi}_{LR} = \min_{\lambda:\lambda \succcurlyeq 0} \left\| \hat{\Sigma}^{-1/2} (\hat{q} - \mu_0) - \frac{1}{2} \hat{\Sigma}^{1/2} A' \lambda \right\|^2,$$

where $\|X\|^2 \equiv X'X$ for any $X \in \mathbb{R}^k$. Letting $\hat{R} \equiv (A\hat{\Sigma}A')^{-1} A\hat{\Sigma}^{1/2}$ and $\nu \equiv \frac{1}{2}\hat{\Sigma}^{1/2} A'\lambda$ (so that $\lambda = 2\hat{R}\nu$) we then have:

$$(67) \qquad\qquad \hat{\xi}_{LR} = \min_{\nu:\hat{R}\nu \succcurlyeq 0} \left\| \hat{\Sigma}^{-1/2} (\hat{q} - \mu_0) - \nu \right\|^2.$$

STEP3: we consider the dual of the minimization problem in Equation (67):

$$(68) \qquad\qquad \max_{\beta:\beta \succcurlyeq 0} \left[-\frac{1}{4}\beta' \hat{R}\hat{R}'\beta - \beta' \hat{R}\hat{\Sigma}^{-1/2}(\hat{q} - \mu_0) \right],$$

where β is a $(k-1)$-vector of Lagrange multipliers. Note that

$$-\frac{1}{4}\beta' \hat{R}\hat{R}'\beta - \beta' \hat{R}\hat{\Sigma}^{-1/2}(\hat{q} - \mu_0)$$
$$= \left\| (A\hat{\Sigma}A')^{-1/2} A(\hat{q} - \mu_0) \right\|^2 - \left\| \tfrac{1}{2}(A\hat{\Sigma}A')^{-1/2}\beta + (A\hat{\Sigma}A')^{-1/2} A(\hat{q} - \mu_0) \right\|^2,$$

so the quantity in Equation (68) is equivalent to:

$$(69) \quad \left\| (A\hat{\Sigma}A')^{-1/2} A(\hat{q} - \mu_0) \right\|^2 - \min_{\beta:\beta \succcurlyeq 0} \left\| \frac{1}{2}(A\hat{\Sigma}A')^{-1/2}\beta + (A\hat{\Sigma}A')^{-1/2} A(\hat{q} - \mu_0) \right\|^2.$$

Now, let:

$$(70) \qquad\qquad \hat{Z} \equiv -(A\hat{\Sigma}A')^{-1/2} A(\hat{q} - \mu_0) \text{ and } \gamma \equiv \frac{1}{2}(A\hat{\Sigma}A')^{-1/2}\beta$$

(so $\beta = 2(A\hat{\Sigma}A')^{1/2}\gamma$); combining Equations (67)-(69) then yields:

$$(71) \qquad\qquad \hat{\xi}_{LR} = \|\hat{Z}\|^2 - \min_{\gamma:(A\hat{\Sigma}A')^{1/2}\gamma \succcurlyeq 0} \|\hat{Z} - \gamma\|^2.$$

Let $P_{\hat{\mathcal{C}}}\hat{Z}$ denote the orthogonal projection of $\hat{Z}$ on the cone $\hat{\mathcal{C}}$, defined as: $\hat{\mathcal{C}} \equiv \{\gamma \in \mathbb{R}^{k-1} : (A\hat{\Sigma}A')^{1/2}\gamma \succcurlyeq 0\}$. The LR statistic in Equation (71) then equals:

$$(72) \qquad\qquad \hat{\xi}_{LR} = \|P_{\hat{\mathcal{C}}}\hat{Z}\|^2.$$

STEP 4: under the null hypothesis H_0 we have $A\mu = 0$ (in addition to $A\mu_0 = 0$) so that the quantity in Equation (70) can be written as $\hat{Z} = BV$, with $B \equiv -(A\hat{\Sigma}A')^{-1/2}A\hat{\Sigma}^{1/2}$ and $V \equiv \hat{\Sigma}^{-1/2}(\hat{q} - \mu)$. Under conditions of Corollary 3, the k-vector V converges in distribution to $V \xrightarrow{d} \mathcal{N}(0_k, \mathrm{Id}_k)$ as $T \to \infty$, and the $(k-1) \times k$-matrix B is such that $BB' = \mathrm{Id}_{k-1}$; hence as $T \to \infty$, we have $\hat{Z} \xrightarrow{d} Z \equiv \mathcal{N}(0_{k-1}, \mathrm{Id}_{k-1})$, under the null hypothesis H_0. Now, for every j, $1 \leqslant j \leqslant k$, let:

$$\sigma_j \equiv \frac{a_{x_j}^Y(q_{Y|X=x_j}(\alpha_T))}{\sqrt{c_j T(1 - \alpha_T)}},$$

and consider the matrix $\sigma_1^{-2}(A\hat{\Sigma}A')$; its entries are:

$$\sigma_1^{-2}(A\hat{\Sigma}A') = \begin{bmatrix} \frac{\hat{\sigma}_1^2}{\sigma_1^2} + \frac{\sigma_2^2}{\sigma_1^2}\frac{\hat{\sigma}_2^2}{\sigma_2^2} & -\frac{\sigma_2^2}{\sigma_1^2}\frac{\hat{\sigma}_2^2}{\sigma_2^2} & & (0) \\ -\frac{\sigma_2^2}{\sigma_1^2}\frac{\hat{\sigma}_2^2}{\sigma_2^2} & \frac{\sigma_2^2}{\sigma_1^2}\frac{\hat{\sigma}_2^2}{\sigma_2^2} + \frac{\sigma_3^2}{\sigma_1^2}\frac{\hat{\sigma}_3^2}{\sigma_3^2} & -\frac{\sigma_3^2}{\sigma_1^2}\frac{\hat{\sigma}_3^2}{\sigma_3^2} & \\ & & \ddots & \\ (0) & & -\frac{\sigma_{k-1}^2}{\sigma_1^2}\frac{\hat{\sigma}_{k-1}^2}{\sigma_{k-1}^2} & \frac{\sigma_{k-1}^2}{\sigma_1^2}\frac{\hat{\sigma}_{k-1}^2}{\sigma_{k-1}^2} + \frac{\sigma_k^2}{\sigma_1^2}\frac{\hat{\sigma}_k^2}{\sigma_k^2} \end{bmatrix}$$

¿From Lemma 3 and Theorem 2 we know that for every j, $1 \leqslant j \leqslant k$, $\sigma_j^{-2}\hat{\sigma}_j^2 \xrightarrow{p} 1$. Moreover, using Lemma 2(ii), and the fact that $F_{U|X}$ does not depend on X so we can write it as $F_{U|X} = F_U$ with auxiliary function a^U, we have:

$$\frac{\sigma_j}{\sigma_1} \sim \sqrt{\frac{c_1}{c_j}} \frac{g_{N_{x_j}}(x_j)}{g_{N_{x_1}}(x_1)} \frac{a^U(\bar{F}_U^{-1}(\alpha_T))}{a^U(\bar{F}_U^{-1}(\alpha_T))} = \sqrt{\frac{c_1}{c_j}} \frac{g_{N_{x_j}}(x_j)}{g_{N_{x_1}}(x_1)},$$

so as $T \to \infty$ we have $\sigma_1^{-2}(A\hat{\Sigma}A') \xrightarrow{p} \Omega$ with a symmetric $(k-1) \times (k-1)$-matrix Ω given by:

(73)
$$\Omega \equiv \begin{bmatrix} 1 + \frac{c_1 g_{N_{x_2}}(x_2)^2}{c_2 g_{N_{x_1}}(x_1)^2} & -\frac{c_1 g_{N_{x_2}}(x_2)^2}{c_2 g_{N_{x_1}}(x_1)^2} & & (0) \\ -\frac{c_1 g_{N_{x_2}}(x_2)^2}{c_2 g_{N_{x_1}}(x_1)^2} & \frac{c_1 g_{N_{x_2}}(x_2)^2}{c_2 g_{N_{x_1}}(x_1)^2} + \frac{c_{t1} g_{N_{x_3}}(x_3)^2}{c_3 g_{N_{x_1}}(x_1)^2} & -\frac{c_{t1} g_{N_{x_3}}(x_3)^2}{c_3 g_{N_{x_1}}(x_1)^2} & \\ & \ddots & & \ddots & \\ (0) & & -\frac{c_1 g_{N_{x_{k-1}}}(x_{k-1})^2}{c_{k-1} g_{N_{x_1}}(x_1)^2} & \frac{c_1 g_{N_{x_{k-1}}}(x_{k-1})^2}{c_{k-1} g_{N_{x_1}}(x_1)^2} + \frac{c_1 g_{N_{x_k}}(x_k)^2}{c_k g_{N_{x_1}}(x_1)^2} \end{bmatrix}$$

Hence, using the fact that $\hat{\mathcal{C}}$ equals $\hat{\mathcal{C}} = \{\gamma \in \mathbb{R}^{k-1} : \sigma_1^{-2}(A\hat{\Sigma}A')^{1/2}\gamma \succcurlyeq 0\}$, we have

that the minimand in Equation (71) converges in probability to a well defined limit:

$$(74) \qquad \hat{\xi}_{LR} \xrightarrow{p} \xi_{LR} \equiv \|Z\|^2 - \min_{\gamma:\Omega^{1/2}\gamma\succcurlyeq0} \|Z - \gamma\|^2 = \|P_{\mathcal{C}}Z\|^2,$$

where $Z \overset{d}{=} \mathcal{N}(0_{k-1}, \mathrm{Id}_{k-1})$, and $P_{\mathcal{C}}Z$ denotes the orthogonal projection of Z on the

cone $\mathcal{C} \equiv \{\gamma \in \mathbb{R}^{k-1} : \Omega^{1/2}\gamma \succcurlyeq 0\}$ with Ω as defined in Equation (73).

STEP 5: In order to determine the distribution of ξ_{LR} in Equation (74) we use

the following lemma:

Lemma 4 (Gourieroux, Holly, and Monfort (1982)). *Let Z be a standard normal*

random vector of dimension $k - 1 \geqslant 1$, i.e. $Z \overset{d}{=} \mathcal{N}(0_{k-1}, \mathrm{Id}_{k-1})$ and let C be a

nonsingular symmetric $(k - 1) \times (k - 1)$-matrix whose columns are denoted $C_j, j =$

$1, \ldots, k - 1$. To each vector $C_j, j = 1, \ldots, k - 1$, we associate a vector $C_j^{\perp} \in \mathbb{R}^{k-1}$

such that: $C_j^{\perp}$ is orthogonal to any $C_i, i \neq j$, and $C_j' C_j^{\perp} < 0$. For each subset S of

the set $\{1, \ldots, k - 1\}$ we define the cone:

$$\mathcal{C}_S \equiv \{y \in \mathbb{R}^{k-1} : y = \sum_{i=1}^{k-1} \alpha_i A_i, \text{ with } \alpha_i \leqslant 0, i = 1, \ldots, k - 1, \text{ and}$$

$$A_i = C_i \text{ when } i \notin S \text{ and } A_i = C_i^{\perp} \text{ when } i \in S\}.$$

Consider the orthogonal projection of Z on the cone $\mathcal{C}_{(1,\ldots,k-1)}$, denoted $P_{\mathcal{C}_{(1,\ldots,k-1)}}Z$.

Then the distribution of $\|P_{\mathcal{C}_{(1,\ldots,k-1)}}Z\|^2$ is a mixture of chi-square distributions:

$$\|P_{\mathcal{C}_{(1,\ldots,k-1)}}Z\|^2 \overset{d}{=} \sum_{d=0}^{k-1} \omega(d)\chi^2(d) \text{ with } \omega(d) = \sum_{S:\dim S=d} \Pr\{P_{\mathcal{C}_{(1,\ldots,k-1)}}Z \in C_S\},$$

where the sequence of weights $\omega(d), d = 0, \ldots, k - 1$ satisfies $\omega(d) \geqslant 0$ and

$\sum_{d=0}^{k-1} \omega(d) = 1$ and $\chi^2(0)$ denotes the point mass distribution at zero.

Apply Lemma 4 to the $(k - 1) \times (k - 1)$-matrix $\Omega^{1/2}$ by letting $C \equiv \Omega^{1/2}$. Using

the notation from Lemma 4, we then have that $\mathcal{C} = \mathcal{C}_{\{1,\ldots,k-1\}}$. Combining Lemma 4

with Equation (74) then yields the result of Theorem 4. Note that the entries of Ω can be consistently estimated using $\hat{\sigma}_1^{-2}(A\hat{\Sigma}A')$; hence the probability weights $\omega(d)$ can be consistently estimated by $\hat{\omega}(d)$, where $\hat{\omega}(d)$ are the weights obtained in the exact Gaussian case. $\square$

Appendix B. Auxiliary Lemmas

Lemmas 5, 6 and 7 are simple translations of results in de Haan (1970) to our problem. Lemmas 8, 9 and 10 present more substantial preliminary results we shall need in the proof of Lemma 3. In the sequel, $\bar{H}$ is a distribution tail $\bar{H} : \mathbb{R}_{++} \to (0,1)$ and $\bar{H}^{-1}$ the corresponding quantile function $\bar{H}^{-1} : (0,1) \to \mathbb{R}_{++}$.

Lemma 5. *If $\bar{H} \in \mathcal{R}_{-\infty}$ at ∞, then $\bar{H}^{-1} \in \mathcal{R}_0$ at 0.*

Proof of Lemma 5. Let $U(x) \equiv \bar{H}(x)$ for all $x > 0$; U is non-increasing. If U is $-\infty$-varying at ∞, then by Corollary 1.2.1 (5) in de Haan (1970), the function $x \mapsto \inf\{y|U(y) \leqslant 1/x\}$ is 0-varying at ∞. It is easy to verify that this function is $x \mapsto \bar{H}^{-1}(1/x)$. Then for $\lambda > 0$, $\lim_{y\downarrow 0} \bar{H}^{-1}(\lambda y)/\bar{H}^{-1}(y) = \lim_{s\to\infty} \bar{H}^{-1}(\lambda/s)/\bar{H}^{-1}(1/s) = 1$ where $s \equiv 1/y$. Thus $\bar{H}^{-1}$ is 0-varying at 0. $\square$

Lemma 6. *If $\bar{H}^{-1} \in \mathcal{R}_0$ at 0, then for all sequences $\{a_N\}$ and $\{a'_N\}$ of positive numbers with $\lim_{N\to\infty} a_N = \lim_{N\to\infty} a'_N = 0$ and $\lim_{N\to\infty} a_N/a'_N = c$ (with $0 < c < \infty$), we have*

$$\lim_{N\to\infty} \frac{\bar{H}^{-1}(a_N)}{\bar{H}^{-1}(a'_N)} = 1.$$

Proof of Lemma 6. Let $U(x) \equiv \bar{H}^{-1}(1/x)$ for all $x > 1$ so $U \in \mathcal{R}_0$ at ∞. Let $\{\alpha_N\}$ and $\{\alpha'_N\}$ be sequences of positive numbers with $\alpha_N \equiv 1/a_N$ and $\alpha'_N \equiv 1/a'_N$ so that $\lim_{N\to\infty} \alpha_N = \lim_{N\to\infty} \alpha'_N = \infty$ and $\lim_{N\to\infty} \alpha_N/\alpha'_N = 1/c$ ($0 < 1/c < \infty$). By applying Corollary 1.2.1 (2) in de Haan (1970) we then have $\lim_{N\to\infty} \bar{H}^{-1}(a_N)/\bar{H}^{-1}(a'_N) = \lim_{N\to\infty} U(\alpha_N)/U(\alpha'_N) = (1/c)^0 = 1$. $\square$

Lemma 7. *Consider a distribution tail $\bar{H} \in \Gamma$ with auxiliary function a. Let H be twice differentiable on $\mathbb{R}_{++}$ with a density h that is eventually non-increasing. Let $W(x) \equiv \bar{H}^{-1}(\exp(-x))$, for $x > 0$. Then W is twice continuously differentiable on $\mathbb{R}_{++}$ with $W'(x) = a[\bar{H}^{-1}(\exp(-x))]$, for $x > 0$, and for any real interval $[a, b]$ we have:*

(i) $\lim_{x\to\infty} W'(x + s)/W'(x) = 1$, *uniformly for s in $[a, b]$;*

(ii) $\lim_{x\to\infty} W''(x + s)/W'(x) = 0$, *uniformly for s in $[a, b]$.*

Proof of Lemma 7. First we prove (i). Note that $a\left(W(x)\right) = \exp(-x)/h(W(x)) = W'(x)$. From Lemma 2 we know that $a \circ \bar{H}^{-1} \in \mathcal{R}_0$ at 0, so

$$(75) \qquad \lim_{x\to\infty} \frac{W'(x + s)}{W'(x)} = \lim_{x\to\infty} \frac{a(\bar{H}^{-1}(\exp(-x - s)))}{a(\bar{H}^{-1}(\exp(-x)))} = 1, \text{ for } s > 0.$$

By Corollary 1.2.1 in de Haan (1970), the convergence is uniform on intervals $[a, b]$ with $a > 0$. This implies that the convergence is uniform on arbitrary intervals $[a, b]$ by the change of variables $y = x - |a| - \eta$, for some $\eta > 0$ (and for $x > |a| + \eta$) by the resulting uniform convergence on $[\eta, b + |a| + \eta]$.

We now prove (ii). First note that $a\left(W(x)\right) = W'(x)$ implies that

$$(76) \qquad \frac{W''(x + s)}{W'(x)} = \left[\frac{W'(x + s)}{W'(x)}\right] a'(W(x + s)).$$

The bracketed term on the right-hand side of Equation (76) converges to 1 uniformly on $[a, b]$ by item (i) of the Lemma. We shall prove that $a'(W(x + s)) \to 0$ as $x \to \infty$ uniformly on $[a, b]$; combined, these two properties establish (ii).

Now $a(x) = \bar{H}(x)/h(x)$, so $a'(x) = -1 - \bar{H}(x)h'(x)/[h(x)]^2$. Then, $\bar{H} \in \Gamma$ implies, by Theorem 2.7.4 in de Haan (1970) (or Proposition 1.18 in Resnick (1987)), that

$$(77) \qquad \lim_{x\to\infty} \frac{\bar{H}(x)h'(x)}{[h(x)]^2} = -1, \text{ i.e. } \lim_{x\to\infty} a'(x) = 0.$$

Fix $x > 0$ large enough so that $x + a > 0$. The range of $a'(W(x + s))$ when $s \in [a, b]$ is the same as the range of $a'(y)$ when $y \in [W(x + a), W(x + b)]$, as W is monotone

increasing. Since a' is continuous, we can let $y(x)$ be such that

$$(78) \qquad a'(y(x)) = \sup_{y \in [W(x+a), W(x+b)]} a'(y).$$

Now, $y(x) \to \infty$ as $x \to \infty$ because W is monotone increasing. Then the right-hand-side of Equation (78) converges to 0 as $x \to \infty$, because a' converges to 0 (77). This proves the needed uniform convergence of $a'(W(x+s))$ in Equation (76). □

Lemma 8. *Let $\{c_T\}$ be a sequence of strictly positive real numbers such that $\lim_{T\to\infty} c_T = \infty$, and consider $f : \mathbb{R} \to \mathbb{R}$. Let $\{X_T\}$ and $\{Y_T\}$ be two independent stochastic processes. If*

(1) *$X_T \xrightarrow{d} X$, as $T \to \infty$, for some X with continuous distribution F,*

(2) *$Y_T \xrightarrow{as} \infty$, as $T \to \infty$,*

(3) *for each $K > 0$, $\lim_{y \to \infty} f(x+y) = 1$ uniformly in $x \in [-K, K]$.*

Then

$$c_T \int_0^{X_T/c_T} f(x + Y_T)\, dx \xrightarrow{d} X, \text{ as } T \to \infty.$$

Proof of Lemma 8. Fix a realization $\{y_T\}$ of $\{Y_T\}$ such that $\lim_{T\to\infty} y_T = \infty$; the almost sure convergence in item 2 ensures that $\{y_T\}$ with $\lim_{T\to\infty} y_T = \infty$ have full measure. Let $z \in \mathbb{R}_+$ and denote by B_T the event

$$\left\{ c_T \int_0^{X_T/c_T} f(x + y_T)\, dx \leqslant z \right\}.$$

Let $\varepsilon > 0$. We shall prove that there is a T^* such that $T \geqslant T^*$ implies that $|P(B_T) - F(z)| < \varepsilon$; here P denotes the probability measure on the space on which $\{X_T\}$ is defined.

Fix $\delta > 0$ such that $F(z/(1-\delta)) - F(z/(1+\delta)) < \varepsilon/4$. Let $K \in \mathbb{R}$ be large enough that $K > z/(1-\delta)$, $F(-K) < \varepsilon/4$ and $1 - F(K) < \varepsilon/4$. Since $X_T \xrightarrow{d} X$, there is T_1

such that $n \geqslant T_1$ implies

(79) $$P\{|X_T| > K\} < \varepsilon/2$$

(80) $$F(z/(1+\delta)) - \varepsilon/4 < P\{X_T \leqslant z/(1+\delta)\}$$

(81) $$P\{X_T \leqslant z/(1-\delta)\} < F(z/(1-\delta)) + \varepsilon/4$$

Let $B_T^K = B_T \cap \{|X_T| \leqslant K\}$. Then, by Statement (79), $T \geqslant T_1$ implies that $P(B_T) - P(B_T^K) \leqslant P\{|X_T| > K\} < \varepsilon/2$.

The convergence in item 3 is uniform on $[-K, K]$, so there is T^* such that $T^* \geqslant T_1$ and such that $T \geqslant T^*$ implies that, for all $\tilde{x} \in [-K, K]$, $(1-\delta) < f(\tilde{x} + y_T) < (1+\delta)$. Then, $T \geqslant T^*$ implies

$$\tilde{x}(1-\delta) < c_T \int_0^{\tilde{x}/c_T} f(x + y_T)\, dx < \tilde{x}(1+\delta),$$

if $\tilde{x} \geq 0$, and

$$\tilde{x}(1+\delta) < c_T \int_0^{\tilde{x}/c_T} f(x + y_T)\, dx < \tilde{x}(1-\delta),$$

if $\tilde{x} < 0$. Then $P\{X_T(1+\delta) \leqslant z, X_T \geqslant 0\} \leqslant P(B_T^K) \cap \{X_T \geqslant 0\} \leqslant P\{X_T(1-\delta) \leqslant z, X_T \geqslant 0\}$. And since $z \geqslant 0$, $P\{X_T(1-\delta) \leqslant z, X_T < 0\} = P(B_T^K) \cap \{X_T < 0\} = P\{X_T(1+\delta) \leqslant z, X_T < 0\}$. Hence, $P\{X_T(1+\delta) \leqslant z\} \leqslant P(B_T^K) \leqslant P\{X_T(1-\delta) \leqslant z\}$.

Tow, $|F(z) - P\{X_T(1+\delta) \leqslant z\}| \leqslant |F(z) - F(z/(1+\delta))| + |F(z/(1+\delta)) - P\{X_T(1+\delta) \leqslant z\}| \leqslant \varepsilon/4 + \varepsilon/4$, by the definition of δ and Statement (80). And similarly for $P\{X_T(1-\delta) \leqslant z\}$. So $|F(z) - P(B_T^K)| < \varepsilon/2$. Finally, then, $T \geqslant T^*$ implies that

$$|F(x) - P(B_T)| \leqslant |F(z) - P(B_T^K)| + |P(B_T) - P(B_T^K)|$$

$$< \varepsilon/2 + \varepsilon/2.$$

The argument for $z < 0$ is analogous. The proof follows because $\{X_T\}$ and $\{Y_T\}$ are independent. $\square$

Lemma 9. *Let $\{c_T\}$ be a sequence of strictly positive real numbers such that $\lim_{T\to\infty} c_T = \infty$, and consider $f : \mathbb{R} \to \mathbb{R}$. Let $\{X_T\}$ be a stochastic process and $\{y_T\}$ a sequence of strictly positive real numbers. If*

(1) *$X_T \overset{d}{\to} X$, as $T \to \infty$, for some X with continuous distribution F,*

(2) *$\lim_{T\to\infty} y_T = \infty$,*

(3) *for each $K > 0$, $\lim_{y\to\infty} f(x + y) = 0$ uniformly in $x \in [-K, K]$.*

Then

$$c_T \int_0^{X_T/c_T} f\left(x + y_T\right) dx \overset{p}{\to} 0, \text{ as } T \to \infty.$$

Proof of Lemma 9. Let $\eta > 0$ and denote by B_T the event

$$\left\{ \left| c_T \int_0^{X_T/c_T} f\left(x + y_T\right) dx \right| \leqslant \eta \right\}.$$

We shall prove that $P(B_T) \to 1$.

Let $\varepsilon > 0$. Let $K > 0$ be large enough that $F(-K) < \varepsilon/2$ and $1 - F(K) < \varepsilon/2$. By the uniform convergence of f on $[-K, K]$, there is T^* such that $T \geqslant T^*$ implies that, for all $\tilde{x} \in [-K, K]$, $|f(\tilde{x} + y_T)| < \eta/K$. Then, for all $T \geqslant T^*$ and $\tilde{x} \in [-K, K]$,

$$\left| c_T \int_0^{\tilde{x}/c_T} f\left(x + y_T\right) dx \right| \leqslant |\tilde{x}|\, \eta/K \leqslant \eta,$$

as $|\tilde{x}| \leqslant K$. So $T \geqslant T^*$ implies that $P(B_T) \geqslant P\left\{X_T \leqslant K\right\} > 1 - \varepsilon$, by the definition of K. $\square$

Lemma 10. *Let $A_1, ..., A_T$ be a random sample from $F_A(x) = 1 - \exp(-x)$ with $x > 0$, and let $A_{(1)}^{(T)} \leqslant ... \leqslant A_{(T)}^{(T)}$ be the ascending order statistics of $(A_1, ..., A_T)$. Consider orders $(m, k) \in \mathbb{T}^2$ such that $m < k \leqslant T$. If $m \to \infty$, $k \to \infty$ and $T \to \infty$*

in a way that $(T - m) \to \infty$, $(T - m)/T \to 0$, $(T - k) \to \infty$, $(T - k)/T \to 0$ *and*
$(T - m)/(T - k) \to \rho$ *where* $\rho > 1$, *then*

$$\sqrt{T - m + 1}\left[A^{(T)}_{(m)} - \ln\frac{T}{T - m}\right] \xrightarrow{d} \mathcal{N}_1 \ \text{and} \ \sqrt{T - m}\left[\frac{A^{(T)}_{(k)} - A^{(T)}_{(m)} - \ln\rho}{\sqrt{\rho - 1}}\right] \xrightarrow{d} \mathcal{N}_2$$

where $\mathcal{N}_1$ *and* $\mathcal{N}_2$ *are two independent standard normal random variables.*

Proof of Lemma 10. Using Renyi's (1953) representation, we know that $\{A^{(T)}_{(T-k+1)} - A^{(T)}_{(T-k)}\}_{k=1}^{T} \overset{d}{=} \{Z_k/k\}_{k=1}^{T}$ where $A^{(T)}_{(0)} \equiv 0$ and where $Z_1, ..., Z_T$ are independent and identically distributed standard exponential random variables. Then for any m, $1 \leqslant m \leqslant T$, and any k, $m < k \leqslant T$, we have

$$(82) \qquad A^{(T)}_{(m)} \overset{d}{=} \sum_{j=T-m+1}^{T} \frac{Z_j}{j} \ \text{and} \ A^{(T)}_{(k)} - A^{(T)}_{(m)} \overset{d}{=} \sum_{l=T-k+1}^{T-m} \frac{Z_l}{l},$$

which are independent. When $m \to \infty$, $k \to \infty$ and $T \to \infty$ in a manner that $(T-m) \to \infty$, $(T-m)/T \to 0$, $(T-k) \to \infty$, $(T-k)/T \to 0$ and $(T-m)/(T-k) \to \rho$ with $\rho > 1$, we can apply the central limit theorem in Liapunov's form to the sums of random variables in Equation (82) (see e.g. Theorem 4 in Renyi (1953)) to get

$$(83) \qquad \frac{A^{(T)}_{(m)} - M_1}{S_1} \xrightarrow{d} \mathcal{N}_1 \ \text{and} \ \frac{A^{(T)}_{(k)} - A^{(T)}_{(m)} - M_2}{S_2} \xrightarrow{d} \mathcal{N}_2,$$

with $\mathcal{N}_1$ and $\mathcal{N}_2$ two independent standard normal random variables where

$$
\begin{aligned}
M_1 &\equiv \sum_{j=T-m+1}^{T} \frac{1}{j} = \sum_{l=1}^{T} \frac{1}{l} - \sum_{n=1}^{T-m} \frac{1}{n} \\
&= \ln T + \gamma + O(T^{-1}) - \ln(T - m) - \gamma + O((T - m)^{-1}) \\
(84) \qquad &= \ln\frac{T}{T - m} + O((T - m)^{-1}),
\end{aligned}
$$

and

$$S_1^2 \equiv \sum_{j=T-m+1}^{T} \frac{1}{j^2}$$

$$= \frac{1}{T-m+1} - \frac{1}{T} + \frac{\theta}{(T-m)(T-m+1)}$$

$$(85) \qquad = \frac{1}{T-m+1} + o((T-m)^{-1}),$$

where γ is the Euler-Mascheroni constant and $0 < \theta < 1$; similarly

$$M_2 \equiv \sum_{j=T-k+1}^{T-m} \frac{1}{j} = \sum_{l=1}^{T-m} \frac{1}{l} - \sum_{n=1}^{T-k} \frac{1}{n}$$

$$= \ln \frac{T-m}{T-k} + O((T-m)^{-1})$$

$$(86) \qquad = \ln \rho + O((T-m)^{-1}),$$

and

$$S_2^2 \equiv \sum_{j=T-k+1}^{T-m} \frac{1}{j^2}$$

$$= \frac{1}{T-k+1} - \frac{1}{T-m} + \frac{\phi}{(T-k)(T-k+1)}$$

$$(87) \qquad = \frac{\rho - 1}{T-m} + o((T-m)^{-1}),$$

where $0 < \phi < 1$ and $\rho > 1$. Combining Equations (83)-(87) then yields the result.

$\square$

References

ACKERBERG, D. A., AND G. GOWRISANKARAN (2006): "Quantifying equilibrium network externalities in the ACH banking industry," *The RAND Journal of Economics*, 37(3), 738–761.

AGUIRREGABIRIA, V., AND P. MIRA (2007): "Sequential Estimation of Dynamic Discrete Games," *Econometrica*, 75(1), 1–53.

ATHEY, S., AND S. STERN (1998): "An Empirical Framework for Testing Theories About Complementarity in Organizational Design," NBER Working Paper 6600.

BAHADUR, R. R. (1966): "A Note on Quantiles in Large Samples," *The Annals of Mathematical Statistics*, 37, 577–580.

BAJARI, P., H. HONG, AND S. P. RYAN (2010): "Identification and Estimation of a Discrete Game of Complete Information," *Econometrica*, 78(5), 1529–1568.

BARLOW, R., D. BARTHOLOMEW, J. BREMNER, AND H. BRUNK (1972): *Statistical Inference under Order Restrictions*. John Wiley and Sons, New York.

BARRO, R. J., AND X. SALA-I-MARTIN (2003): *Economic Growth*. MIT Press.

BARTHOLOMEW, D. J. (1959a): "A Test of Homogeneity for Ordered Alternatives," *Biometrika*, 46, 36–48.

——— (1959b): "A Test of Homogeneity for Ordered Alternatives. II," *Biometrika*, 46, 328–335.

BASSETT, JR., G., AND R. KOENKER (1978): "Asymptotic Theory of Least Absolute Error Regression," *Journal of the American Statistical Association*, 73(363), 618–622.

BERRY, S. T., J. LEVINSOHN, AND A. PAKES (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63(4), 841–890.

——— (1999): "Voluntary Export Restraints on Automobiles: Evaluating a Trade Policy," *American Economic Review*, 89(3), 400–430.

BJORN, P. A., AND Q. H. VUONG (1984): "Simultaenous Equations Models for Dummy Endogenous Variables: A Game Theoretic Formulation with an Application to Labor Force Participation," Caltech Social Science Working Paper 577.

——— (1985): "Econometric Modeling of a Stackelberg Game with an Application to Labor Force Participation," Caltech Working Paper 537.

BRYNJOLFSSON, E., AND P. MILGROM (2012): "Complementarity in organizations," in *The Handbook of Organizational Economics*, ed. by R. Gibbons, and J. Roberts. Princeton Univers. Press.

CARROLL, R. J., D. RUPPERT, AND L. A. STEFANSKI (1995): *Measurement Error in Nonlinear Models*. Chapmann and Hall.

CHERNOZHUKOV, V. (2005): "Extremal Quantile Regression," *Annals of Statistics*, 33, 806–839.

CHERNOZHUKOV, V., AND I. FERNANDEZ-VAL (2011): "Inference for Extremal Conditional Quantile Models, with an Application to Market and Birthweight Risks," *The Review of Economic Studies*, 78(2), 559–589.

DARDANONI, V., AND A. FORCINA (1998): "A Unified Approach to Likelihood Inference on Stochastic Orderings in a Nonparametric Context," *Journal of the American Statistical Association*, 93, 1112–1123.

DE HAAN, L. (1970): "On Regular Variation and its Application to the Weak Convergence of Sample Extremes," University of Amsterdam, Mathematical Centre Tract 32.

——— (1974): "Equivalence Classes of Regularly Varying Functions," *Stochastic Processes and their Applications*, 2, 243–259.

DE PAULA, A. (2013): "Econometric Analysis of Games with Multiple Equilibria," *Annual Review of Economics*, 5(1).

DEKKERS, A. L., AND L. DE HAAN (1989): "On the Estimation of the Extreme-Value Index and Large Quantile Estimation," *Annals of Statistics*, 17, 1795–1832.

ECHENIQUE, F. (2002): "Comparative Statics by Adaptive Dynamics and The Correspondence Principle," *Econometrica*, 70(2), 833–844.

ECHENIQUE, F., AND I. KOMUNJER (2009): "Testing Models with Multiple Equilibria by Quantile Methods," *Econometrica*, 77, 1281–1297.

EMBRECHTS, P., C. KLUPPELBERG, AND T. MIKOSCH (1997): *Modelling Extremal Events for Insurance and Finance.* Springer-Verlag.

FALK, M. (1989): "A Note on Uniform Asymptotic Normality of Intermediate Order Statistics," *Annals of the Institute of Statistical Mathematics*, 41, 19–29.

FOX, J. T., AND N. LAZZATI (2012): "Identification of Potential Games and Demand Models for Bundles," unpublished manuscript, University of Michigan.

GOURIEROUX, C., A. HOLLY, AND A. MONFORT (1982): "Likelihood Ratio Test, Wald Test, and Kuhn-Tucker Test in Linear Models with Inequality Constraints on the Regression Parameters," *Econometrica*, 50, 63–80.

HAYASHI, F. (1982): "Tobin's Marginal q and Average q: A Neoclassical Interpretation," *Econometrica*, 50(1), 213–224.

HAYASHI, F., AND T. INOUE (1991): "The Relation Between Firm Growth and Q with Multiple Capital Goods: Theory and Evidence from Panel Data on Japanese Firms," *Econometrica*, 59(3), 731–753.

JIA, P. (2008): "What Happens When Wal-Mart Comes to Town: An Empirical Analysis of the Discount Retailing Industry," *Econometrica*, 76(6), 1263–1316.

KARAMATA, J. (1930): "Sur un mode de croissance régulière des fonctions," *Mathematica*, 5, 38–53.

KODDE, D. A., AND F. C. PALM (1986): "Wald Criteria for Jointly Testing Equality and Inequality Restrictions," *Econometrica*, 54, 1243–1248.

KOENKER, R., AND Q. ZHAO (1996): "Conditional Quantile Estimation and Inference for ARCH Models," *Econometric Theory*, 12, 793–813.

KOMUNJER, I. (2005): "Quasi-Maximum Likelihood Estimation for Conditional Quantiles," *Journal of Econometrics*, 128, 137–164.

LAZZATI, N. (2012): "Treatment Response with Social Interactions: Partial Identification via Monotone Comparative Statics," Working Paper, University of Michigan.

LJUNGQVIST, L., AND T. J. SARGENT (2004): *Recursive Macroeconomic Theory*. MIT Press (Second Edition).

MAS-COLELL, A., M. D. WHINSTON, AND J. R. GREEN (1995): *Microeconomic Theory*. Oxford University Press.

MCKELVEY, R. D., AND T. R. PALFREY (1995): "Quantal Response Equilibria for Normal Form Games," *Games and Economic Behavior*, 10(1), 6–38.

MILGROM, P., AND J. ROBERTS (1990a): "The economics of modern manufacturing: Technology, strategy, and organization," *The American Economic Review*, 80(3), 511–528.

——— (1990b): "Rationalizability, Learning and Equilibrium in Games with Strategic Complementarities," *Econometrica*, 58(6), 1255–1277.

MILGROM, P., AND C. SHANNON (1994): "Monotone Comparative Statics," *Econometrica*, 62(1), 157–180.

MOLINARI, F., AND A. M. ROSEN (2008): "The Identification Power of Equilibrium in Games: The Supermodular Case," *Journal of Business & Economic Statistics*, 26(3), 297–302.

MOSTELLER, F. (1946): "On some Useful "Inefficient" Statistics," *Annals of Mathematical Statistics*, 17, 377–408.

PAKES, A., S. BERRY, AND J. A. LEVINSOHN (1993): "Applications and Limitations of Some Recent Advances in Empirical Industrial Organization: Price Indexes and the Analysis of Environmental Change," *American Economic Review*, 83(2), 240–246.

Pollard, D. (1991): "Asymptotics for Least Aboslute Deviation Regression Estimators," *Econometric Theory*, 7, 186–199.

Portnoy, S. (1991): "Behavior of Regression Quantiles in Non-Stationary, Dependent Cases," *Journal of Multivariate Analysis*, 38, 100–113.

Powell, J. L. (1984): "Least Absolute Deviations Estimation for the Censored Regression Model," *Journal of Econometrics*, 25, 303–325.

——— (1986): "Censored Regression Quantiles," *Journal of Econometrics*, 32, 143–155.

Renyi, A. (1953): "On the Theory of Order Statistics," *Acta Mathematica Hungarica*, 4, 191–231.

Resnick, S. I. (1987): *Extreme Values, Regular Variation, and Point Processes*. Springer-Verlag.

Robertson, T., and E. J. Wegman (1978): "Likelihood Ratio Tests for Order Restrictions in Exponential Families," *Annals of Statistics*, 6, 485–505.

Robertson, T., and F. T. Wright (1982): "Bounds on Mixtures of Distributions Arising in Order Restricted Inference," *Annals of Statistics*, 10, 302–306.

Samuelson, P. A. (1947): *Foundations of Economic Analysis*. Harvard University Press, Massachusetts.

Siddiqui, M. (1960): "Distribution of Quantiles in Samples Form a Bivariate Population," *Journal of Research of the National Bureau of Standards–B. Mathematics and Mathematical Physics*, 64B, 145–150.

Smirnov, N. (1952): "Limit Distributions for the Terms of a Variational Series," *American Mathematical Society Translations*, 67, 82–143.

Sweeting, A. (2009): "The strategic timing incentives of commercial radio stations: An empirical analysis using multiple equilibria," *RAND Journal of Economics*, 40(4), 710–742.

VILLAS-BOAS, J. M. (1997): "Comparative Statics of Fixed Points," *Journal of Economic Theory*, 73(1), 183–198.

VIVES, X. (1999): *Oligopoly Pricing.* MIT Press, Cambridge, Massachusetts.

WATTS, V. (1980): "The Almost Sure Representation of Intermediate Order Statistics," *Zeitschrift Fur Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 54, 281–285.