

DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

# CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA 91125

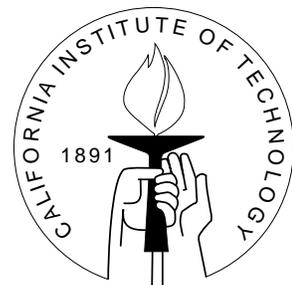
## TESTING MODELS WITH MULTIPLE EQUILIBRIA BY QUANTILE METHODS

Federico Echenique

California Institute of Technology

Ivana Komunjer

University of California at San Diego



**SOCIAL SCIENCE WORKING PAPER 1244**

December 2005

# Testing Models with Multiple Equilibria by Quantile Methods

Federico Echenique

Ivana Komunjer

## Abstract

We derive an econometric test for the presence of monotone comparative statics in models with multiple equilibria. The test applies to many economic models, such as single-person decision, industrial organization, macroeconomic and game-theory models. These models have complementarities between exogenous and endogenous variables. We show that, as a result, extreme (large and small) conditional quantiles of the endogenous variable are increasing in the exogenous variable. We develop a likelihood-ratio test based on estimates for the conditional quantiles of the endogenous variable, which is an asymptotic extension of Bartholomew's (1959a,b) "chi-bar squared" test. Our assumptions are weak; we remain agnostic about the cardinality, location and probabilities of the equilibrium set, and make no restrictions on the equilibrium-selection procedure.

JEL classification numbers: C1,C5

Key words: Econometrics of Games, Monotone Comparative Statics, Quantile Regression

# Testing Models with Multiple Equilibria by Quantile Methods \*

Federico Echenique

Ivana Komunjer

## 1 Introduction

Comparative statics predictions—how exogenous variables affect endogenous variables—are important to establish in economic models. Often, the models possess multiple equilibria, which poses severe challenges for establishing such comparative statics predictions, not to mention testing them econometrically. The standard practice is to impose additional assumptions, so that the model has a unique equilibrium. These additional assumptions are often very strong, and independent of the desired economic explanation.

When there are *complementarities* between exogenous and endogenous variables, despite the possible presence of multiple equilibria, a “monotone comparative statics” prediction holds: There is a smallest and a largest equilibrium, and these change monotonically with exogenous variables. In this paper, we show how monotone comparative

---

\*Acknowledgments: Earlier versions of this paper have been presented at Northwestern University, UC Berkeley, UC Davis, the North American ES summer meeting, the EEE2005 conference, and the 2005 ESSET workshop. Many thanks to Andrés Aradillas-López, Patrick Bajari, Peter Bossaerts, Chris Chambers, Xiaohong Chen, Valentino Dardanoni, Joel Horowitz, Matt Jackson, Michael Jansson, Chuck Manski, Andrea Mattozzi, Rosa Matzkin, Tom Palfrey, Jim Powell, Andrea Rotnitzky, Paul Ruud, Max Stinchcombe, Andrew Sweeting, Elie Tamer, Quang Vuong, Hal White, and Bill Zame for their suggestions and comments.

statics arguments translate into restrictions on the conditional quantiles of the endogenous variable. We construct a likelihood-ratio test of such quantile restrictions. Our test recognizes the existence of (possibly) many equilibria with unknown probabilities of realization.

The monotone comparative statics prediction was first formulated by Milgrom and Roberts (1994) and Villas-Boas (1997), who observed that, in many economic models, there is a smallest and a largest equilibrium, and that these change monotonically with exogenous variables. We illustrate their insight, and how we build on it, using a simple example: we discuss comparative statics in the Solow growth model. Let  $f$  be the production function,  $s$  the savings rate,  $\delta$  the depreciation rate, and  $n$  the rate of population growth in a closed economy. The Solow model determines capital per worker,  $k$ , from the equilibrium condition:  $sf(k) = (n + \delta)k$ . The endogenous variable is then  $k$ . We shall disregard the always-present trivial equilibrium,  $k = 0$ .

If  $f$  exhibits increasing returns, hence is not strictly concave, one typically obtains multiple equilibria—see Figure 1. The literature on growth has emphasized the importance of increasing returns, but growth models avoid multiple equilibria by very specific assumptions on the form of increasing returns (e.g. the “AK model”). One must then wonder if the sole notion of increasing returns is sufficient to generate the literature’s explanations. If it is not sufficient, the literature is fundamentally flawed.

Consider Figure 1. When  $f_1$  is the production function, the equilibria are  $k_1$ ,  $k_2$ , and  $k_3$ . If we change the production function to  $f_2$ , the set of equilibria changes to  $k'_1$ ,  $k'_2$ ,  $k'_3$ ,  $k'_4$  and  $k'_5$ : the largest and the smallest equilibrium have increased. An increase in education, making the economy more productive, might lead to such a change in  $f$ . The Solow growth model then predicts that an exogenous increase in education causes an increase in capital, as long as one looks at the extreme equilibria. This is the monotone comparative statics prediction.

Now consider a researcher testing the hypothesis that an increase in education leads to more capital, using cross-country data on education and income. Assume that all the

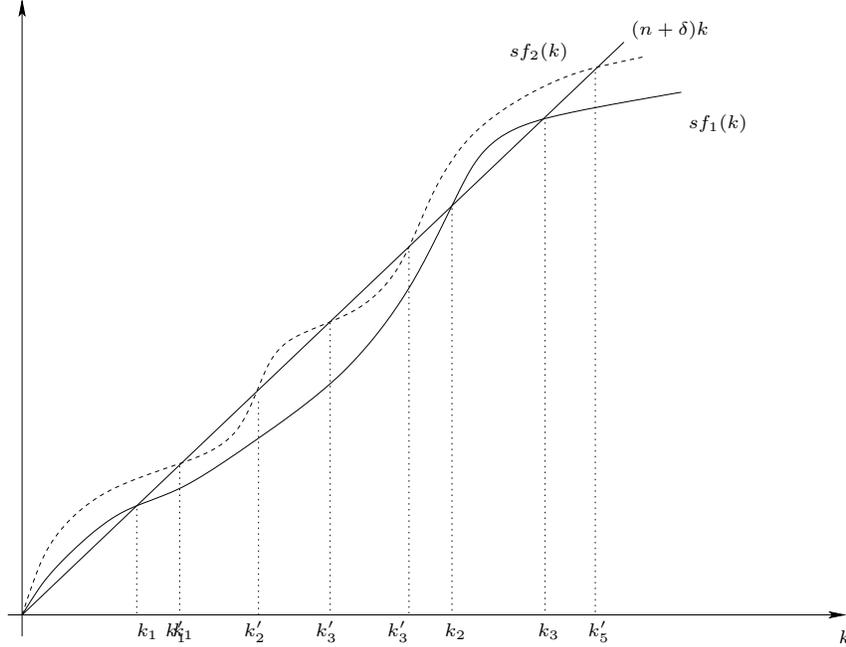


Figure 1: Solow model.

countries in such a data-set can be sorted into two groups: low-education countries with production function  $f_1$ , and high-education countries with  $f_2$ . The researcher assumes that the prediction about extremal equilibria in Figure 1 holds. She, however, ignores the number of equilibria, their true values and the corresponding probabilities. For example, the true equilibrium distribution could be the one in Figure 2.

Figure 2 shows two sets of equilibria, one for each  $f$ . There are Gaussian-like densities around each equilibrium, which reflects that an equilibrium is observed imprecisely. If equilibrium  $k_1$  is selected, we may observe a deviation from  $k_1$  due to, for example, short-run fluctuations around equilibrium, (unobserved) heterogeneities across different countries, or measurement error. We shall reflect our uncertainty about equilibrium selection, and deviations from equilibrium, by using the mixture form depicted in Figure 2.

The case in Figure 2 presents a challenge for our researcher. Note that (a) the largest equilibrium in the low-education group, and the low equilibria in the high-education group have high probabilities; (b) the largest equilibrium in the high-education group has low probability; and (c) three relatively small and likely equilibria appear in the

high-education countries. Were our researcher to compute the mean income in both groups, she should expect it to be smaller in the high education group. Of course, one can find equilibrium distributions for which the conclusion is reversed, and the mean of highly-educated countries is larger. In other words, there are no testable implications for the conditional mean of income.

Note that, in the example of Figure 1, the middle equilibrium ( $k_2$ ) is unstable. Everything we say holds, including the negative statements about testing using mean income, if one ignores unstable equilibria. We would need a more complicated picture, with more than three equilibria, to make our point.

What makes the testing problem *a priori* difficult is that the researcher has little knowledge of what is the true equilibrium distribution: she does not know the number of equilibria in each case, their values, or their corresponding probabilities.

Our solution is to impose the right structure on the tails of probability distributions around each equilibrium such that the effect of  $k_3 < k'_5$  (and  $k_1 < k'_1$ ) is eventually reflected in some large enough quantile of  $k$ . Hence, the researcher should compare some (sufficiently high) quantile of the income in two groups, in order to test the hypothesis of interest. In the problems we consider, then, quantile-based analysis—and not traditional mean-based methods—is the proper econometric methodology. Specifically, we derive a

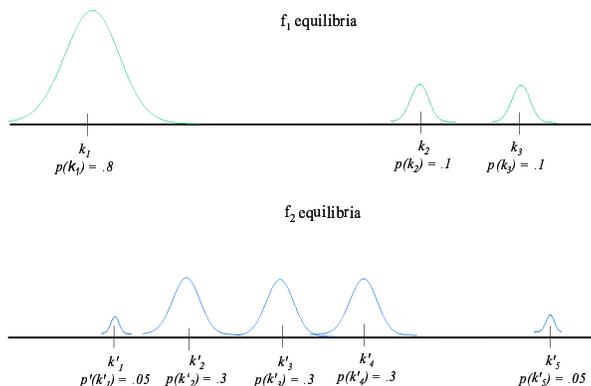


Figure 2: Equilibrium distributions.

likelihood-ratio test for the type of monotone comparative statics exemplified in Figure 1.

It is important to stress that we make few assumptions on the true equilibrium distribution. We only assume that the extremal equilibria have strictly positive probability (alternatively, the test is that the extremal equilibria, of those with positive probability, depend monotonically on the exogenous variables). Similarly, our assumptions on the distribution tails 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.

We now briefly comment on the scope of our contributions. First, our methods apply to many economic models. Some examples are: single-person decision models, IO models (such as Bertrand and Cournot oligopoly, markets with differentiated products), macroeconomic models, growth models, and game-theory models. We provide a more detailed description of these models and how they enter into our setup in Section 2.2. In particular, we note that a more modern version of the Solow model, with optimizing agents, would also fall into our framework.

Second, we show that standard econometric methods, which focus on the mean of the endogenous variable, are *not* suited for testing in the presence of multiple equilibria. Researchers should instead use quantile-based econometric methods. From a practical view-point, our testing methods are readily taken to data. They build on the extensive literature on quantile-based econometric methods which are today well-understood and easy to use in practice.

Finally, the results of the paper give an economic justification for quantile-based econometric methods. Work on these methods has so far been justified by the—reasonable but limited—idea that one should look for richer aspects of the conditional distribution of endogenous variables, beyond the conditional mean.<sup>1</sup> Our results imply that, in models with multiple equilibria, one *must* use quantile-based econometric methods: these models cannot be tested within the classical mean regression framework.

---

<sup>1</sup>Chernozhukov and Hong (2004) is an exception.

**Related Literature** As early as in the work of Bjorn and Vuong (1984, 1985), econometricians recognized the importance of testing economic models that possess multiple equilibria. Proposed solutions to the problem of multiplicity of equilibria have been to assume the probabilities of various equilibrium realizations known (Bjorn and Vuong, 1984); or finitely parametrized (Bjorn and Vuong, 1985; McKelvey and Palfrey, 1995; Bajari, Hong, and Ryan, 2004; Sweeting, 2005). With a model for the equilibrium selection rule in hand, likelihood based estimation and inference is feasible even when multiple equilibria are present. The validity of the results, however, will depend crucially on the correct specification of the equilibrium selection model. Unfortunately, economic theory provides little guidance on how to model equilibrium selection; there is a large body of work on equilibrium selection, but no widely accepted theory.

When no model for the equilibrium selection rule is available, there are alternative approaches to estimation and inference. The first exploits the fact that—despite the multiplicity—some of the model features are uniquely predicted; by focusing attention on those features, one is then able to carry out likelihood-based estimation and inference. Observable implications in models with multiple equilibria were first derived by Jovanovic (1989) who sought conditions for point identification. A number of recent papers in the (quickly expanding) literature on econometrics of games further carry out estimation in such models (Bresnahan and Reiss, 1990, 1991; Berry, 1992; Tamer, 2003). With no simple point-identification conditions available, there is a second approach to estimation and inference. In a variety of economic models studied, the parameters of interest typically satisfy a set of inequality constraints—in addition to the standard equilibrium equality constraints—that one may exploit for set identification (Andrews, Berry, and Jia, 2004; Ciliberto and Tamer, 2004; Kim, 2005).

Mostly, the above papers build on discrete-choice methods and typically try to estimate agents' payoff functions (i.e. estimate the nature of strategic interaction); not only test for the presence of a comparative-statics effect. Understandably, they make more parametric assumptions than we do.

The existing methods are also well-suited for models with few choice variables, as they build on the theory of discrete-choice models. Our methods are the first that apply to models with continuous endogenous variables. Later in the paper we point out some additional differences.

Finally, we note that Athey and Stern (1998) discuss tests for monotone comparative statics. They study firms' choice of organizational form; they do not address equilibrium problems.

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.

## 2 Setup

We shall first introduce our model and some basic notation. We then show that our model is general enough to accommodate many existing economic models, and that the econometric structure allows for some common specifications.

### 2.1 Reduced-form model

Let  $\mathcal{T} \subseteq \mathbf{R}^n$  be a finite set, and  $\leq$  be the usual partial order on  $\mathbf{R}^n$ . A *reduced-form model* is a collection  $((\mathcal{E}_t, p_t), t \in \mathcal{T})$  such that:

1. for all  $t$ ,  $\mathcal{E}_t \subseteq \mathbf{R}_{++}$  is finite and nonempty;
2.  $t < t'$  implies that  $\min \mathcal{E}_t < \min \mathcal{E}_{t'}$  and  $\max \mathcal{E}_t < \max \mathcal{E}_{t'}$ ;
3. for all  $t$ ,  $p_t(\cdot)$  is a probability distribution over  $\mathcal{E}_t$  such that  $p_t(\min \mathcal{E}_t) > 0$  and  $p_t(\max \mathcal{E}_t) > 0$ .

Interpret  $T$  as the exogenous variable,  $\mathcal{T}$  as the set of possible values it can take, and  $\mathcal{E}_t$  as the set of predictions (equilibria) to be tested when  $T$  takes value  $t \in \mathcal{T}$ . The set  $\Theta = \cup_{t \in \mathcal{T}} \mathcal{E}_t$  is the set of all possible equilibrium outcomes as  $t$  varies in  $\mathcal{T}$ . The elements of  $\Theta$  are denoted  $\theta$ .

The probability distribution  $p_t(\cdot)$  reflects some equilibrium-selection procedure. Our results shall build on the comparative statics in item (2) of the definition. We show below how the predictions of a number of economic models satisfy (2). The assumption on  $p_t(\cdot)$  in item (3) is harder to justify; we discuss it in Section 2.4.

We shall use the notation  $\bar{\theta}_t = \max \mathcal{E}_t$  and  $\underline{\theta}_t = \min \mathcal{E}_t$ .

## 2.2 Economic models

Many economic models imply our reduced-form model. If a model's equilibria are the zeroes, or the fixed points, of a function that depends monotonically on  $t$ , then it implies our reduced-form model (Milgrom and Roberts (1994) first made this point, and provided examples).

Following are some examples of well-known models that imply our reduced-form model. We show how the models satisfy the definition of a reduced-form model, except for the requirement that  $\mathcal{E}_t$  is finite. Standard arguments from differential topology imply a generically finite number of equilibria, zeroes or fixed points, in our applications.<sup>2</sup>

### 2.2.1 Individual decision maker.

Consider models of individual decision making, in which the endogenous variable is one-dimensional, and determined as the first-order condition of an optimization problem. Examples are investment-choices by firms and labor-decisions by individuals. Our methods can be used for testing if investment is sensitive to Tobin's  $q$ , as in Hayashi (1982)

---

<sup>2</sup>On the other hand, we believe our methods can be extended to accommodate infinite  $\mathcal{E}_t$ .

and Hayashi and Inoue (1991)), or for testing the effect of wages on labor-participation, as in Juhn (1992).

The most important applications in this category are market-equilibrium models that one solves as a social-planning problem. Thus, many growth and macroeconomic models—such as the ones in Barro and Sala-I-Martin (2003), and Ljungqvist and Sargent (2004)—fall into our framework. For example, a version of the Solow model we discussed in the introduction, but with optimizing agents like in Barro and Sala-I-Martin (2003), and increasing returns, is within our framework.

To see how the models with an individual decision maker reduce to the model in Section 2.1, interpret  $g(\theta, t) = 0$  as a first-order condition, where  $t$  is a parameter and  $\theta$  is the equilibrium. So, if  $g(\theta, t)$  is increasing in  $t$ ,  $A \subseteq \mathbf{R}_+$  is compact,  $g : A \times \mathcal{T} \rightarrow \mathbf{R}$  is continuous,  $g(\inf A, t) > 0$  and  $g(\sup A, t) \leq 0$ , then

$$\mathcal{E}_t = \{\theta \in A : g(\theta, t) = 0\},$$

has the properties described in Section 2.1. Note that the last two assumptions on  $g$  are weak “boundary” conditions; they are satisfied if Inada conditions are satisfied. No additional assumptions on the form of  $g$  are needed. Hence, our methods allow one to test individual decision making models under rather general assumptions on the individual’s objective functions.

### 2.2.2 One-dimensional equilibrium.

Examples of equilibrium models with a one-dimensional endogenous variable are easy to find. For example, in a two-player game, we can compose the two players’ best-response functions and reduce it to our model. As a consequence, duopoly models generally reduce to our reduced-form. Cournot  $n$ -firm oligopoly models also reduce to a one-dimensional equilibrium model by an aggregation procedure (Amir, 1996; Amir and Lambson, 2000). One can thus test, for example, if entry of additional firms to a market causes a decrease in prices (Amir and Lambson, 2000).

Additional examples include some overlapping-generations models, such as the one in Ljungqvist and Sargent (2004), and two-good general equilibrium models (in fact by interpreting  $g(\theta, t) = 0$  as a zero excess-demand condition in the previous subsection).

To see why the one-dimensional equilibrium models reduce to the model in Section 2.1, interpret the fixed-point condition  $g(\theta, t) = \theta$  as an equilibrium condition. So,

$$\mathcal{E}_t = \{\theta \in A : g(\theta, t) = \theta\}, \tag{1}$$

with  $g(\theta, t)$  continuous, increasing in  $t$ , and  $A \subseteq \mathbf{R}_{++}$  compact, fits our reduced-form model (see e.g. Milgrom and Shannon (1994)).

Since our results apply to two-player games, they apply to many models in the literature on the econometrics of games, e.g. Bresnahan and Reiss (1990, 1991); Berry (1992); Tamer (2003); Bajari, Hong, and Ryan (2004) and Ciliberto and Tamer (2004).

### 2.2.3 Multi-dimensional equilibrium.

Equilibrium models with a multidimensional endogenous variable fall into our framework, as long as there are complementarities among the different dimensions of the endogenous variable. Thus, games of strategic complementarities (Topkis, 1979; Vives, 1990; Milgrom and Roberts, 1990) reduce to the model in Section 2.1.

These games include many economic models of interest. For example, price-competition in the—arguably—most common market structure: Bertrand competition with differentiated products (e.g. we could test for some of the effects in Berry, Levinsohn, and Pakes (1995)). They also include macroeconomic models of coordination failures, models of two sided matching, and Hart and Moore’s property-rights model. Topkis (1998) and Vives (1999) contain many more examples.

## 2.3 Econometric model

A *statistical reduced-form model* is a collection

$$((\mathcal{E}_t, p_t, (H_{\theta,t})_{\theta \in \mathcal{E}_t}), t \in \mathcal{T})$$

such that

1.  $((\mathcal{E}_t, p_t), t \in \mathcal{T})$  is a reduced-form model,
2. for all  $t \in \mathcal{T}$  and  $\theta \in \mathcal{E}_t$ ,  $H_{\theta,t}(\cdot)$  is a twice-differentiable distribution function on  $\mathbf{R}_{++}$ , with strictly positive density.

The model should be interpreted as follows: the endogenous variable  $X$  is distributed, given  $T = t$ , as  $F_t(\cdot) \equiv F(\cdot|T = t)$ , where  $F_t(\cdot)$  is a discrete mixture of continuous distributions,

$$F_t(x) = \sum_{\theta \in \mathcal{E}_t} p_t(\theta) H_{\theta,t}(x/\theta) \text{ for any } x > 0. \quad (2)$$

The assumptions on  $H_{\theta,t}(\cdot)$  imply that, for any  $t$  in  $\mathcal{T}$ ,  $F_t(\cdot)$  is twice differentiable on  $\mathbf{R}_{++}$ , and that  $f_t(\cdot) \equiv F'_t(\cdot) > 0$  on  $\mathbf{R}_{++}$ . To alleviate the notation, we shall drop the reference to  $t$  whenever possible.

Given  $\alpha \in (0, 1)$ , let  $q_t(\alpha)$  denote the  $\alpha$ -quantile under  $F_t(\cdot)$ ,

$$q_t(\alpha) \equiv \inf\{y \in \mathbf{R}_{++} : F_t(y) > \alpha\}. \quad (3)$$

Under our assumptions, the quantity in Equation (3) is well defined, and  $q_t(\alpha) = F_t^{-1}(\alpha)$ .

In what follows, we devote particular attention to the distribution tails of the endogenous variable:  $\bar{F}_t(\cdot) \equiv 1 - F_t(\cdot)$ . Similarly, we let  $\bar{H}_{\theta,t}(\cdot) \equiv 1 - H_{\theta,t}(\cdot)$ . Note that given  $t$  in  $\mathcal{T}$  and for any  $\alpha \in (0, 1)$ , we have the following simple relation:

$$q_t(\alpha) = \bar{F}_t^{-1}(1 - \alpha). \quad (4)$$

## 2.4 On the model assumptions

We now comment on the restrictions we have made in our statistical reduced-form model.

### 2.4.1 Mixture assumptions

We do not make any specific assumptions regarding the origin of the mixture distribution in (2). This preserves the generality of our approach. We note however that familiar classes of econometric error models imply the mixture model for  $F_t(\cdot)$ . It is implied, in particular, when conditional on  $T = t$ , the observations come from a multiplicative model

$$X \equiv \theta \cdot \varepsilon, \tag{5}$$

where  $\theta \in \mathcal{E}_t$  is an equilibrium, drawn according to  $p_t(\cdot)$  and  $\varepsilon$  is a multiplicative “error”, drawn from distribution  $H_{\theta,t}(\cdot)$ . This multiplicative structure has the advantage of preserving the non-negativity of the endogenous variable; however, it is not crucial, and one can work with a linear model by performing a logarithmic transformation.

Depending on the application, the “error” variable  $\varepsilon$  in Equation (5) can have different interpretations. For example, in labor-economic studies based on individuals’ reported earnings (Bound, Brown, Duncan, and Rodgers, 1994),  $\varepsilon$  can be thought of as a measurement error; it can arise when  $\theta$  is measured relative to an equivalence scale which is itself measured inaccurately. In game-theoretic models  $\varepsilon$  corresponds to a behavioral error, as in models of Quantal Response Equilibrium (McKelvey and Palfrey, 1995). In the Solow growth model,  $\varepsilon$  may be interpreted as a short-term fluctuation around a long-term steady state  $\theta$ .<sup>3</sup>

In classical error models, it is assumed that  $\varepsilon$  is independent of  $\theta$ , as well as all other variables. In some applications, these assumptions may be justified, but in others they are merely convenient (e.g. see Chen, Hong, and Tamer (2005) for an excellent discussion). In the labor literature, for example, there is well-established evidence against the independence assumption (Bound, Brown, Duncan, and Rodgers, 1994). We allow for arbitrary dependence between the exogenous variable  $T$ , the equilibrium  $\theta$ , and the

---

<sup>3</sup>Note that the “no error” case in which the equilibria  $\theta$  are observed without error, corresponds to having  $H_{\theta,t}(x) = \mathbb{I}(x \leq \theta)$  for any  $x > 0$ , where the function  $\mathbb{I}(\cdot)$  is the indicator function: for any event  $A$ , we have  $\mathbb{I}(A) = 1$  if  $A$  true and  $\mathbb{I}(A) = 0$  otherwise. To avoid complications arising from the non-differentiability of  $F_t(\cdot)$ , we do not formally treat this case.

error  $\varepsilon$ . For example, errors are allowed to be correlated with  $\theta$  and/or to be conditionally heteroskedastic (see Section 4.4).

### 2.4.2 Assumptions on $p$

We have assumed that the largest and smallest equilibria in  $\mathcal{E}_t$  have positive probability under  $p_t(\cdot)$ —this is our only deviation from being agnostic regarding  $p_t(\cdot)$  (one precedent in this respect is Sweeting (2005), who assumes that all equilibria have positive probability). 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  $\mathcal{T} = \{\underline{t}, \bar{t}\}$  with  $\underline{t} < \bar{t}$ . We need that the largest equilibrium in  $\mathcal{E}_{\underline{t}}$ , *of those with positive  $p_{\underline{t}}$ -probability*, be smaller than the largest with positive  $p_{\bar{t}}$ -probability. This is a weaker requirement than the one we have imposed above. It says that the equilibrium-selection mechanism implicit in  $(p_t)_{t \in \mathcal{T}}$  should have the right correlation with respect to changes in  $t$ .

We claim that this correlation can be expected to hold: suppose agents are playing an equilibrium in  $\mathcal{E}_{\underline{t}}$  when the parameter changes to  $\bar{t}$ . Then a broad class of learning dynamics must lead them to play a larger equilibrium (Echenique (2002) contains a proof). Alternatively, the “shift up” of  $g$  in the models in Section 2.2 makes larger equilibria focal.

### 2.4.3 Assumptions on $H$

Our statistical reduced-form model assumes that the  $H_{\theta,t}(\cdot)$  are continuous distributions with support  $\mathbf{R}_{++}$ . We work with unbounded support because most practitioners are probably unwilling to assume a bounded support. Our results, however, are easy to generalize to situations where the support of  $H_{\theta,t}(\cdot)$  is bounded.

We assumed the distributions  $H_{\theta,t}(\cdot)$  to be unknown. In some cases, it might be

preferable to assume  $H_{\theta,t}(\cdot)$  known, at least up to some finite number of parameters; in such cases, the conditional distribution of  $X$  in (2) could in principle be estimated via maximum likelihood methods, provided the equilibrium probabilities  $p_t(\theta)$  are either known or finitely parametrized. However, the presence of unknown equilibrium probabilities  $p_t(\theta)$  in  $F_t(\cdot)$  causes almost all the practical problems of implementation and model estimation with maximum likelihood methods (e.g. see Carroll, Ruppert, and Stefanski (1995)).<sup>4</sup>

## 2.5 On individual heterogeneity

One may be concerned about the effect of individual heterogeneity on our model. Here we show how individual heterogeneity—whether observed or unobserved—may be built into our reduced-form model  $((\mathcal{E}_t, p_t), t \in \mathcal{T})$ .

Suppose that in addition to the exogenous variable  $T$ , and the endogenous variable  $X$ , there is a variable  $Z$  that affects equilibrium sets and probabilities. Let  $\mathcal{Z}$  be the set of its outcomes  $z$ ,  $\mathcal{Z} \subseteq \mathbf{R}^m$ , write  $\mathcal{E}_{t,z}$  for the set of possible equilibria given  $T = t$  and  $Z = z$ , and let  $p_{t,z}(\cdot)$  be the distribution over equilibria in  $\mathcal{E}_{t,z}$ . Suppose now that  $\mathcal{Z}$  is finite and that for each value  $z$  in  $\mathcal{Z}$ , the largest equilibrium in  $\mathcal{E}_{t,z}$  with positive  $p_{t,z}$ -probability is increasing in  $t$ . Let  $\mathcal{E}_t \equiv \cup_{z \in \mathcal{Z}} \mathcal{E}_{t,z}$  and  $p_t(\theta) \equiv \sum_{z \in \mathcal{Z}} p_{t,z}(\theta) p(z)$  where  $p(\cdot)$  is a probability distribution over  $\mathcal{Z}$ . Given that  $\mathcal{Z}$  is finite,  $((\mathcal{E}_t, p_t), t \in \mathcal{T})$  is a reduced-form model and our results will apply. In a sense, what we are doing is to integrate out  $Z$  from the problem and reduce it to the reduced-forms we can analyze.

We require the assumption that the monotone comparative statics result holds conditional on  $Z = z$  so that the largest equilibrium in  $\mathcal{E}_{t,z}$  with positive  $p_{t,z}$ -probability is increasing in  $t$ , for each value of  $z$ . In the applications we can think of, this is a reasonable assumption. For example, in the Solow model, one can let  $Z$  be the population growth,  $n$ . Conditional on  $n$ , then, the monotone comparative statics prediction holds, and an

---

<sup>4</sup>Typically, both the location points of the support of  $\theta$  when  $T = t$  ( $\mathcal{E}_t$ ) and the probabilities attached to them ( $p_t$ ) are estimated using the EM-algorithm.

increase in  $t$  increases the extremal equilibria of  $k$ . In this case, the source of heterogeneity among different countries—their population growth—is observed. Letting  $Z$  be the depreciation rate,  $\delta$ , we obtain the same result; only now the countries are assumed to be heterogeneous in their  $\delta$ 's, which are not necessarily observed by the econometrician.

In the next section, we give a different example of individual heterogeneities in game-theoretic models.

## 2.6 On the source of randomness

In our statistical reduced-form model, conditional on the realization of an equilibrium, there is a random deviation from equilibrium. This is modeled in the mixture form (2) by the  $H$ -distributions. The econometric literature in Industrial Organization includes errors in the payoffs of the players. These errors then imply that equilibrium outcomes are random. Examples of this approach can be found in Bresnahan and Reiss (1990), Bresnahan and Reiss (1991), Tamer (2003), Ciliberto and Tamer (2004), and Bajari, Hong, and Ryan (2004).<sup>5</sup> We also assume randomness in the equilibrium realization, but it takes the form of an unknown equilibrium selection.

In our view, there are no strong reasons to prefer one approach over the other, but the errors-in-payoffs approach presents one question: How should the different information on the realizations of the errors be modeled? Most authors treat the game as a complete-information game, while others (see Aradillas-López (2004) and Bajari and Hong (2005)) allow for incomplete information.<sup>6,7</sup>

---

<sup>5</sup>The literature on testing and estimation with experimental data, on the other hand, makes an assumption similar to ours (McKelvey and Palfrey, 1995).

<sup>6</sup>We note briefly that our reduced-form model allows for incomplete information, as long as we have strategic complementarities (Vives, 1990).

<sup>7</sup>There is a literature on dynamic games of incomplete information. Ericson and Pakes (1995) is a seminal paper; recent contributions are Aguirregabiria and Mira (2004), Pakes, Ostrovsky, and Berry (2004), Pesendorfer and Schmidt-Dengler (2004) and Bajari, Benkard, and Levin (2004). Part of this literature deals with models with multiple equilibria. We conjecture that some of these models can be

That said, our reduced-form model can accommodate for errors in payoffs by treating the errors as unobserved individual heterogeneity (Section 2.5). For example, consider a two-player game described as follows. Player  $i$  chooses a strategy,  $s_i$ , from some set  $S_i$  of real numbers. There is a real exogenous variable  $T$ , and two independent real random variables,  $W_1$ , and  $W_2$  that parameterize the players' payoffs. Given  $T = t$ , and a realization  $(w_1, w_2)$  of  $(W_1, W_2)$ , Player  $i$ 's payoffs are  $u_i(s_1, s_2; t, w_i)$ . Let  $\beta_1(s_2; t, w_1)$  and  $\beta_2(s_1; t, w_2)$  be the players' best-response functions, given values of  $t$  and  $w_i$ . Then the Nash equilibria of this game, are given by the fixed points of the composed best responses,  $\beta(s; t, w_1, w_2) = \beta_1(\beta_2(s; t, w_2); t, w_1)$ . Let  $\mathcal{E}_{t, w_1, w_2}^* = \{s \in S_1 : s = \beta(s; t, w_1, w_2)\}$ . Let  $p_{t, w_1, w_2}(\cdot)$  be the distribution over equilibria in the game specified by the values  $(t, w_1, w_2)$ ; assume that the largest equilibrium has positive probability.

Now, assume that, for  $s_1, s'_1 \in S_1$  with  $s_1 < s'_1$ ,  $u_i(s'_1, s_2; t, w_1) - u_i(s_1, s_2; t, w_1)$  is monotone increasing in  $t$  and  $s_2$ .<sup>8</sup> For all  $s$ , and  $(w_1, w_2)$ ,  $t < t'$  implies  $\beta(s; t, w_1, w_2) < \beta(s; t', w_1, w_2)$  by standard monotone comparative statics arguments (see e.g. Milgrom and Shannon (1994) or Topkis (1998)). Then  $t < t'$  implies that  $\max \mathcal{E}_{t, w_1, w_2}^* < \max \mathcal{E}_{t', w_1, w_2}^*$ , for all  $(w_1, w_2)$ . Finally, let  $p_t(\cdot)$  be the distribution over  $\mathcal{E}_t$  obtained from the distribution of  $(w_1, w_2)$ , and the  $(p_{t, w_1, w_2})$ -distributions.

By setting  $Z \equiv (W_1, W_2)$  and arguing as in Section 2.5 above, we reduce a model with errors in payoffs to our model. Admittedly, the assumption of finite support of  $W_i$  is not standard, but we are unaware of any substantive reasons one should avoid it. An infinite support seems to be, mostly, a convenient modeling choice. Here we use the finite support assumption as a simple way of showing that our model can accommodate errors in payoffs, but one can probably, with more work, allow the errors to have more general supports.

---

analyzed with our techniques.

<sup>8</sup>This model reduces to Tamer's (2003) model in the case where there are two strategies for each player and, using Tamer's notation,  $\beta > 0$  and  $\Delta > 0$ .

### 3 Nature of the problem and results

We first explain our results informally. Consider an example, building on the Solow-model example in the introduction. Let  $\mathcal{T} = \{\underline{t}, \bar{t}\} \subseteq \mathbf{R}, \underline{t} < \bar{t}$ , where  $\underline{t}$  and  $\bar{t}$  denote low- and high-level education, respectively. Denoting the different equilibrium levels of  $k$  by  $\theta$ 's, let  $\mathcal{E}_{\underline{t}} = \{\theta_1, \theta_2, \theta_3\}$  and  $\mathcal{E}_{\bar{t}} = \{\theta'_1, \theta'_2, \theta'_3, \theta'_4, \theta'_5\}$ . The situation is represented in Figure 2.

The problem of obtaining testable implications is to say how the distributions,  $F_{\underline{t}}(\cdot)$  and  $F_{\bar{t}}(\cdot)$ , must differ ( $F_t(\cdot)$  was defined in Equation (2)). All we have to work with is that  $\theta_3 < \theta'_5$  (and  $\theta_1 < \theta'_1$ ), but the probability of the  $\theta'_5$  equilibrium is very low, and there are three very likely equilibria,  $\theta'_2, \theta'_3$  and  $\theta'_4$ , that are smaller than both  $\theta_2$  and  $\theta_3$ .

Note that the mean (and median) under  $F_{\bar{t}}(\cdot)$  is smaller than that under  $F_{\underline{t}}(\cdot)$ . Thus the conditional mean (and median) of  $X$  does not change monotonically in  $t$ . One can change the example so the conditional mean increases instead of decreasing; thus there are no testable implications for the conditional mean of the endogenous variable. One is also more likely to observe a realization under  $F_{\underline{t}}(\cdot)$  that is larger than under  $F_{\bar{t}}(\cdot)$  than vice versa.

Our solution to finding testable implications is to assume the right structure on the distribution tails, so the effect of  $\theta_3 < \theta'_5$  is felt for large enough values of the endogenous variable, irrespective of the values of the corresponding probabilities  $p_{\underline{t}}(\cdot)$  and  $p_{\bar{t}}(\cdot)$ . We show how, for large enough realizations  $x$ , the distribution tails  $\bar{F}_{\underline{t}}(\cdot) \equiv 1 - F_{\underline{t}}(\cdot)$  and  $\bar{F}_{\bar{t}}(\cdot) \equiv 1 - F_{\bar{t}}(\cdot)$  must satisfy:  $\bar{F}_{\underline{t}}(x) < \bar{F}_{\bar{t}}(x)$ .

Note that the tail  $\bar{F}_{\underline{t}}(\cdot)$  of  $F_{\underline{t}}(\cdot)$  is related to that of  $H_{\theta, \underline{t}}(\cdot)$ , denoted  $\bar{H}_{\theta, \underline{t}}(\cdot) \equiv 1 - H_{\theta, \underline{t}}(\cdot)$ , via

$$\bar{F}_{\underline{t}}(x) = p_{\underline{t}}(\theta_1)\bar{H}_{\theta_1, \underline{t}}(x/\theta_1) + p_{\underline{t}}(\theta_2)\bar{H}_{\theta_2, \underline{t}}(x/\theta_2) + p_{\underline{t}}(\theta_3)\bar{H}_{\theta_3, \underline{t}}(x/\theta_3).$$

Then,

$$\bar{F}_{\underline{t}}(x) = \bar{H}_{\theta_3, \underline{t}}(x/\theta_3) \left[ p_{\underline{t}}(\theta_3) + p_{\underline{t}}(\theta_1) \frac{\bar{H}_{\theta_1, \underline{t}}(x/\theta_1)}{\bar{H}_{\theta_3, \underline{t}}(x/\theta_3)} + p_{\underline{t}}(\theta_2) \frac{\bar{H}_{\theta_2, \underline{t}}(x/\theta_2)}{\bar{H}_{\theta_3, \underline{t}}(x/\theta_3)} \right] \quad (6)$$

Assume that the tails of  $H_{\theta,\underline{t}}(\cdot)$  satisfy the following properties: for the largest equilibrium  $\theta_3$  in  $\mathcal{E}_{\underline{t}}$  we have

$$\lim_{x \rightarrow \infty} \frac{\bar{H}_{\theta_3,\underline{t}}(\lambda x)}{\bar{H}_{\theta_3,\underline{t}}(x)} = 0, \quad (7)$$

whenever  $\lambda > 1$ ; and for any other  $\theta$ , with  $\theta < \theta_3$ , suppose that

$$\frac{\bar{H}_{\theta,\underline{t}}(x)}{\bar{H}_{\theta_3,\underline{t}}(x)} \quad (8)$$

is bounded as  $x$  goes to  $\infty$ .

Property (7) requires that the tail of the distribution  $H_{\theta_3,\underline{t}}(\cdot)$  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. Property (8), on the other hand, ensures that  $\bar{H}_{\theta_3,\underline{t}}(x)$  does not decrease towards 0 faster than  $\bar{H}_{\theta,\underline{t}}(x)$ , for  $\theta < \theta_3$ . This property is trivially satisfied when  $H_{\theta,t}(\cdot)$  is independent of  $\theta$ , for example.

The two properties imply that the term in brackets in Equation 6 converges to  $p_{\underline{t}}(\theta_3)$  as  $x$  grows. Indeed,

$$\frac{\bar{H}_{\theta_1,\underline{t}}(x/\theta_1)}{\bar{H}_{\theta_3,\underline{t}}(x/\theta_3)} = \left[ \frac{\bar{H}_{\theta_1,\underline{t}}(x/\theta_1)}{\bar{H}_{\theta_3,\underline{t}}(x/\theta_1)} \right] \left[ \frac{\bar{H}_{\theta_3,\underline{t}}(x/\theta_1)}{\bar{H}_{\theta_3,\underline{t}}(x/\theta_3)} \right], \quad (9)$$

so that letting  $y = x/\theta_1$  and  $\lambda = \theta_3/\theta_1 > 1$  and using (7) we have  $B = \bar{H}_{\theta_3,\underline{t}}(\lambda y)/\bar{H}_{\theta_3,\underline{t}}(y) \rightarrow 0$ . Since by (8) the  $A$  term is bounded, the ratio in (9) converges to 0.

Thus,

$$\bar{F}_{\underline{t}}(x) \sim p_{\underline{t}}(\theta_3) \bar{H}_{\theta_3,\underline{t}}(x/\theta_3),$$

as  $x$  goes to  $\infty$  and we establish that the behavior of  $F_{\underline{t}}(x)$  for large  $x$  is driven by  $\theta_3$ .

Under analogous assumptions on  $H_{\theta,\bar{t}}(\cdot)$ ,  $\bar{F}_{\bar{t}}(x)$  behaves like  $p_{\bar{t}}(\theta'_5) \bar{H}_{\theta'_5,\bar{t}}(x/\theta'_5)$ . Thus,

$$\frac{\bar{F}_{\underline{t}}(x)}{\bar{F}_{\bar{t}}(x)} \sim \left[ \frac{p_{\underline{t}}(\theta_3)}{p_{\bar{t}}(\theta'_5)} \right] \left[ \frac{\bar{H}_{\theta_3,\underline{t}}(x/\theta_3)}{\bar{H}_{\theta_3,\underline{t}}(x/\theta'_5)} \right] \left[ \frac{\bar{H}_{\theta_3,\underline{t}}(x/\theta'_5)}{\bar{H}_{\theta_3,\bar{t}}(x/\theta'_5)} \right] \left[ \frac{\bar{H}_{\theta_3,\bar{t}}(x/\theta'_5)}{\bar{H}_{\theta'_5,\bar{t}}(x/\theta'_5)} \right]. \quad (10)$$

Since  $\theta_3 < \theta'_5$ , the assumptions on  $H_{\theta,\underline{t}}(\cdot)$  and  $H_{\theta,\bar{t}}(\cdot)$  imply that the  $A$  term goes to 0 and the  $C$  term is bounded as  $x$  grows. If, in addition, we assume that

$$\frac{\bar{H}_{\theta_3,\underline{t}}(x)}{\bar{H}_{\theta_3,\bar{t}}(x)} \quad (11)$$

is bounded as  $x$  goes to  $\infty$ , then the  $B$  term is bounded. So  $\bar{F}_{\underline{t}}(x)/\bar{F}_{\bar{t}}(x)$  converges to 0 irrespective of the values of  $p_{\underline{t}}(\cdot)$  and  $p_{\bar{t}}(\cdot)$ . Hence, for large enough  $x$ , the tail of  $F_{\bar{t}}(x)$  is thicker than that of  $F_{\underline{t}}(x)$ ; this is the essence of our testable implication.

To summarize, Statements (7), (8) (and their analogues for  $\bar{t}$ ) and (11) together ensure that the ratio of  $\bar{F}_{\underline{t}}(\cdot)$  to  $\bar{F}_{\bar{t}}(\cdot)$  goes to zero. This is our testable implication:  $\bar{F}_{\underline{t}}(x) < \bar{F}_{\bar{t}}(x)$  for  $x$  large enough. As a result, large enough population quantiles must be larger under  $F_{\bar{t}}(\cdot)$  than under  $F_{\underline{t}}(\cdot)$ . 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 only we omit the dependence of  $H_{\theta,t}(\cdot)$  on  $t$  and  $\theta$ .

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

$$\lim_{x \rightarrow c} \frac{\bar{H}(\lambda x)}{\bar{H}(x)} = \lambda^{\rho}. \quad (12)$$

The notion of regular variation was first introduced by Karamata (1930)); see e.g. 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  $\infty$ . Most of the distributions employed in economics, such as the Gaussian, exponential and lognormal distributions, are in  $\mathcal{R}_{-\infty}$  at  $\infty$ . The distributions in  $\mathcal{R}_{-\infty}$  at  $\infty$  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  $x^{-\alpha}$ .<sup>9</sup>

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

$$\lim_{x \rightarrow \infty} \frac{\bar{H}(\lambda x)}{\bar{H}(x)} = \begin{cases} 0 & \text{if } \lambda > 1 \\ \infty & \text{if } \lambda < 1. \end{cases} \quad (13)$$

The discussion in Section 3 should suggest that Statement (13) is a useful property. Now, the property in Statement (13) does not control the rate at which  $\bar{H}(\lambda x)/\bar{H}(x)$  converges. By using a subclass of ( $-\infty$ )-varying distribution tails, called  $\Gamma$  (de Haan, 1970), we can exercise this control.

**Definition 2.** *A distribution tail  $\bar{H}(\cdot)$  belongs to the class  $\Gamma$ ,  $\bar{H} \in \Gamma$ , if there exists a function  $a : \mathbf{R}_{++} \rightarrow \mathbf{R}_{++}$  such that for  $\lambda > 0$ ,*

$$\lim_{x \rightarrow \infty} \frac{\bar{H}(x + \lambda a(x))}{\bar{H}(x)} = \exp(-\lambda); \quad (14)$$

*$a(\cdot)$  is called the auxiliary function of  $\bar{H}(\cdot)$ .*

When  $\bar{H} \in \Gamma$ , one can show that  $a(\cdot)$  can be chosen as  $a(\cdot) \equiv \bar{H}(\cdot)/h(\cdot)$  (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  $\Gamma$  are: exponential, two-parameter Gamma, Gaussian, lognormal, and Weibull (see e.g. Embrechts, Kluppelberg, and Mikosch (1997)).

The tail properties in Equations (12) and (14) translate into similar properties for the inverse function  $\bar{H}^{-1} : (0, 1) \rightarrow \mathbf{R}_{++}$  (see Lemma 8) and the class of regularly varying functions is closed under inversion. The inverses of functions in  $\Gamma$ , however, do not belong to  $\Gamma$  but form a class called  $\Pi$  (de Haan, 1970, 1974).

---

<sup>9</sup>This implies that all the moments of a random variable with a ( $-\infty$ )-varying distribution tail are finite. Examples of distributions with  $\rho$ -varying tails,  $\rho > -\infty$ , which *do not* have finite moments are: (1) a stable law with index  $\alpha$ ,  $0 < \alpha < 2$ , for which  $\rho = -\alpha$ ; (2) a Cauchy distribution, for which  $\rho = -1$ . Hence the use of those distributions is not permitted in our framework.

**Definition 3.** A function  $\bar{H}^{-1} : (0, 1) \rightarrow \mathbf{R}_{++}$  belongs to the class  $\Pi$ ,  $\bar{H}^{-1} \in \Pi$ , if there exist functions  $b : \mathbf{R}_{++} \rightarrow \mathbf{R}_{++}$  and  $a : \mathbf{R}_{++} \rightarrow \mathbf{R}_{++}$  such that, for  $\mu \in (0, 1)$ ,

$$\lim_{y \downarrow 0} \frac{\bar{H}^{-1}(\mu y) - b(y)}{a(y)} = -\ln \mu. \quad (15)$$

When  $\bar{H}(\cdot)$  belongs to  $\Gamma$  with auxiliary function  $\tilde{a}(\cdot)$ , Equation (15) 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 statistical reduced-form model  $((\mathcal{E}_t, p_t, (H_{\theta,t})_{\theta \in \mathcal{E}_t}), t \in \mathcal{T})$  and impose structure on the tails  $\bar{H}_{\bar{\theta}_t, t}(\cdot)$ .

**Assumption S1:** Say that a statistical reduced-form model  $((\mathcal{E}_t, p_t, (H_{\theta,t})_{\theta \in \mathcal{E}_t}), t \in \mathcal{T})$  satisfies assumption S1 if, for every  $t \in \mathcal{T}$ ,  $\bar{H}_{\bar{\theta}_t, t}(\cdot)$  is in  $\mathcal{R}_{-\infty}$  at  $\infty$ , and

$$\text{for every } \theta \in \Theta \text{ with } \theta < \bar{\theta}_t, \frac{\bar{H}_{\theta, t}(x)}{\bar{H}_{\bar{\theta}_t, t}(x)} \text{ is bounded as } x \text{ goes to } \infty. \quad (16)$$

We now show how the properties of the tails  $\bar{H}_{\bar{\theta}_t, t}(\cdot)$  translate into properties of the tail  $\bar{F}_t(\cdot)$  of the conditional distribution of the endogenous variable—recall that  $F_t(\cdot)$  was defined by Equation (2).

**Lemma 1.** If  $((\mathcal{E}_t, p_t, (H_{\theta,t})_{\theta \in \mathcal{E}_t}), t \in \mathcal{T})$  satisfies S1, then for every  $t \in \mathcal{T}$ :

- (i)  $\bar{F}_t(\cdot)$  is in  $\mathcal{R}_{-\infty}$  at  $\infty$ , and  $\bar{F}_t(x) \sim p_t(\bar{\theta}_t) \bar{H}_{\bar{\theta}_t, t}(x/\bar{\theta}_t)$  as  $x \rightarrow \infty$ ;
- (ii)  $\bar{H}_{\bar{\theta}_t, t}^{-1}(\cdot)$  and  $\bar{F}_t^{-1}(\cdot)$  are in  $\mathcal{R}_0$  at 0, and  $\bar{F}_t^{-1}(y) \sim \bar{\theta}_t \bar{H}_{\bar{\theta}_t, t}^{-1}(y)$  as  $y \downarrow 0$ .

Thus, the limit behavior of the distribution tail  $\bar{F}_t(\cdot)$  is determined by the largest equilibrium in  $\mathcal{E}_t$  (i.e.  $\bar{\theta}_t$ ) and its probability,  $p_t(\bar{\theta}_t)$ . In the limit, the conditional quantiles of  $X$  are proportional to the quantiles under  $H_{\bar{\theta}_t, t}(\cdot)$ , and the constant of proportionality equals  $\bar{\theta}_t$ .

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

**Assumption S2:** Say that a statistical reduced-form model  $((\mathcal{E}_t, p_t, (H_{\theta,t})_{\theta \in \mathcal{E}_t}), t \in \mathcal{T})$  satisfies S2 if it satisfies S1 and, in addition, for every  $\theta \in \Theta$ , and

$$\text{for every } (t, t') \in \mathcal{T}^2 \text{ with } t < t', \frac{\bar{H}_{\bar{\theta}_t, t}(x)}{\bar{H}_{\bar{\theta}_{t'}, t'}(x)} \text{ is bounded as } x \text{ goes to } \infty. \quad (17)$$

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

**Theorem 2.** *If  $((\mathcal{E}_t, p_t, (H_{\theta,t})_{\theta \in \mathcal{E}_t}), t \in \mathcal{T})$  satisfies S2, then for any  $(t, t') \in \mathcal{T}^2$ :*

(i) *there is  $\bar{x} \in \mathbf{R}_{++}$  such that  $t < t'$  implies  $\bar{F}_t(x) < \bar{F}_{t'}(x)$  for all  $x \geq \bar{x}$ ;*

(ii) *there is  $\bar{\alpha} \in (0, 1)$  such that  $t < t'$  implies  $q_t(\alpha) < q_{t'}(\alpha)$  for all  $\alpha \in [\bar{\alpha}, 1)$ .*

The idea in (i) of Theorem 2 is that, if the distribution  $H_{\bar{\theta}_t, t}(\cdot)$  is not too heavy-tailed, the effect of  $t$  on the largest equilibrium in  $\mathcal{E}_t$  will eventually be noticed in the tail of  $F_t(\cdot)$ . In a sense, there is a race between the potentially damaging effect of other  $\theta_t \in \mathcal{E}_t$ , and the effect of the largest equilibrium  $\bar{\theta}_t$ . Since  $p_t(\cdot)$  is arbitrary,  $p_t(\cdot)$  can work in favor of the other  $\theta_t \in \mathcal{E}_t$ , as in Figure 2. But the  $(-\infty)$ -varying condition on  $\bar{H}_{\bar{\theta}_t, t}(\cdot)$  and Property (17) together guarantee that the largest  $\theta_t$  wins the race. Hence, for large values of  $x$ , the conditional distributions  $F_t(\cdot)$  of the endogenous variable have tails that increase monotonically with  $t$ , a property akin to monotonicity in first-order stochastic dominance. Equivalently, (ii) in Theorem 2 has consequences for the quantiles of  $X$  conditional on  $T = t$ . In the limit, the conditional quantiles of the endogenous variable are monotone increasing in  $t$ .

### 4.3 Further Model Implications

Theorem 2 suggests one can use estimates of conditional quantiles under  $F_t(\cdot)$  for testing, but there are several difficulties. First, the theorem does not determine  $\bar{x}$  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  $H_{\theta_t, t}(\cdot)$  while maintaining

our agnosticism about the  $p_t(\cdot)$  distributions. Third, given the asymptotic distributions of estimates for quantiles under  $F_t(\cdot)$ , we need to derive a test that is not influenced by the  $p_t(\cdot)$  distributions nor the non-extremal values in  $\mathcal{E}_t$ , for which our model makes no predictions.

In order to deal with the asymptotics, we need to impose further structure on the “largest” distribution tail  $\bar{H}_{\bar{\theta}_t,t}(\cdot)$ : in addition to being  $(-\infty)$ -varying,  $\bar{H}_{\bar{\theta}_t,t}(\cdot)$  is now assumed to belong to the class  $\Gamma$ .

**Assumption S3:** *Say that a statistical reduced-form model  $((\mathcal{E}_t, p_t, (H_{\theta,t})_{\theta \in \mathcal{E}_t}), t \in \mathcal{T})$  satisfies S3 if it satisfies S1 and, in addition, for every  $t \in \mathcal{T}$  we have  $\bar{H}_{\bar{\theta}_t,t} \in \Gamma$  with auxiliary function  $a_{\bar{\theta}_t,t}(\cdot)$ .*

This allows us to show the following results on the tails of conditional distributions  $F_t(\cdot)$  of the endogenous variable.

**Lemma 3.** *If  $((\mathcal{E}_t, p_t, (H_{\theta,t})_{\theta \in \mathcal{E}_t}), t \in \mathcal{T})$  satisfies S3, then for every  $t \in \mathcal{T}$ :*

- (i)  $\bar{F}_t \in \Gamma$  with auxiliary function  $a_t(\cdot)$ ,  $a_t(x) = \bar{\theta}_t a_{\bar{\theta}_t,t}(x/\bar{\theta}_t)$  for all  $x > 0$ ;
- (ii)  $\bar{H}_{\bar{\theta}_t,t}^{-1}(\cdot)$  and  $\bar{F}_t^{-1}(\cdot)$  are in  $\Pi$  with auxiliary functions  $a_{\bar{\theta}_t,t} \circ \bar{H}_{\bar{\theta}_t,t}^{-1}(\cdot)$  and  $a_t \circ \bar{F}_t^{-1}(\cdot)$  in  $\mathcal{R}_0$  at 0, and  $a_t(\bar{F}_t^{-1}(y)) \sim \bar{\theta}_t a_{\bar{\theta}_t,t}(\bar{H}_{\bar{\theta}_t,t}^{-1}(y))$  as  $y \downarrow 0$ .

Lemma 3 presents two results: First, that the  $\Gamma$  (resp.  $\Pi$ ) properties of  $\bar{H}_{\bar{\theta}_t,t}(\cdot)$  (resp.  $\bar{H}_{\bar{\theta}_t,t}^{-1}(\cdot)$ ) continue to hold for any  $\bar{F}_t(\cdot)$  (resp.  $\bar{F}_t^{-1}(\cdot)$ ) provided Property (16) is satisfied. Hence, we will only need to make assumptions on the behavior of  $\bar{H}_{\bar{\theta}_t,t}(\cdot)$  in Equation (2) in order to fully characterize the behavior of  $\bar{F}_t(\cdot)$  as  $x$  gets large. Note that this result is particularly important if we want to preserve our agnosticism about the probabilities  $p_t(\cdot)$  over equilibria  $\theta_t$ .

The second result of Lemma 3 is to show how  $a_t \circ \bar{F}_t^{-1}(\cdot)$  relates to  $a_{\bar{\theta}_t,t} \circ \bar{H}_{\bar{\theta}_t,t}^{-1}(\cdot)$ . We shall prove that these expressions are involved in the formulation of the central limit theorem for empirical conditional quantiles under  $F_t(\cdot)$ . In other words, the results of Lemma 3 are essential for understanding the asymptotic properties of the estimators

for conditional quantiles of  $X$ , and hence for constructing an econometric test of the implication derived in Theorem 2.

Before discussing in more details the questions of estimation and inference we give a simple example, illustrating how a commonly-used set of assumptions falls into our framework.

#### 4.4 An Illustrative Example

All assumptions we have made about the distributions  $H_{\theta,t}(\cdot)$  are satisfied, for example, in the multiplicative error model (5) with log-normal and conditionally heteroskedastic errors. Let  $\mu_{\theta,t} : \Theta \times \mathcal{T} \rightarrow \mathbf{R}$  and  $\sigma_{\theta,t} : \Theta \times \mathcal{T} \rightarrow \mathbf{R}_{++}$ , and say that, conditional on  $T = t$  and  $\theta = \theta_t$ ,  $\ln \varepsilon_t$  is normally distributed with mean  $\mu_{\theta_t,t}$  and variance  $\sigma_{\theta_t,t}^2$ ,

$$\ln \varepsilon_t | t, \theta_t \sim \mathcal{N}(\mu_{\theta_t,t}, \sigma_{\theta_t,t}^2).$$

A simple application of L'Hôpital's rule shows that for any  $(t, \theta_t) \in \mathcal{T} \times \mathcal{E}_t$ , the ratio  $\bar{H}_{\bar{\theta}_t,t}(\lambda x) / \bar{H}_{\bar{\theta}_t,t}(x)$  converges to 0 as  $x$  goes to  $\infty$ , i.e. that  $\bar{H}_{\bar{\theta}_t,t}(\cdot)$  is in  $\mathcal{R}_{-\infty}$  at  $\infty$ . If, in addition, for a given realization  $t$  of the exogenous variable  $T$ ,  $\sigma_{\theta_t,t}^2$  increases with  $\theta$ , then a similar argument shows that the ratio in Equation (16) converges to 0, hence is bounded as  $x$  goes to  $\infty$ . Therefore a log-normal error whose conditional variance increases with the equilibrium level satisfies our assumption *S1*.

Similarly, if for any given value of  $\theta \in \Theta$ , the variance  $\sigma_{\theta,t}^2$  increases with  $t$ , then the ratio in Equation (16) converges to 0 and it therefore bounded as  $x$  goes to  $\infty$ . In other words a log-normal error whose conditional variance is increasing in both  $t$  and  $\theta$  satisfies our assumption *S2*. A simple model for the conditional variance that satisfies this requirement is, for example, that of linear heteroskedasticity:  $\sigma_{\theta,t} = \kappa_t \theta$ , where  $\kappa_t > 0$  is such that  $\kappa_t < \kappa_{t'}$  as long as  $t < t'$ . Note that no restrictions are imposed on the conditional mean  $\mu_{\theta,t}$  of the log-error.<sup>10</sup>

---

<sup>10</sup>In particular we do not need to assume  $\mu_{\theta_t,t} + \sigma_{\theta_t,t}^2/2 = 0$  which would guarantee that  $E(\ln \varepsilon_t | t, \theta_t) = 1$ . The latter restriction is a type of identification condition we would impose, were we to estimate say

It is a well-known result that the log-normal family belongs to the class  $\Gamma$  of distributions and that their auxiliary functions are of the form  $a_{\theta,t}(x) = \sigma_{\theta,t}^2 x / (\ln x - \mu_{\theta,t})$  (see e.g. Resnick (1987), Embrechts, Kluppelberg, and Mikosch (1997)); hence, any log-normal error model that satisfies  $S1$  will also satisfy our assumption  $S3$ .

## 5 Estimation

### 5.1 Notation and Setup

Consider a random sample  $(X_{t,1}, \dots, X_{t,N_t})$  of size  $N_t$  from a distribution function  $F_t(\cdot)$ , with  $t \in \mathcal{T}$ . Let  $(x_{t,1}, \dots, x_{t,N_t})$  denote the realizations of  $(X_{t,1}, \dots, X_{t,N_t})$  and write  $\hat{F}_{t,N_t}(\cdot)$  to be the empirical distribution function,  $\hat{F}_{t,N_t}(y) \equiv N_t^{-1} \sum_{k=1}^{N_t} \mathbb{1}(x_{t,k} \leq y)$  for  $y > 0$ . For a given  $\alpha$ ,  $0 < \alpha < 1$ , the empirical quantile under  $F_t(\cdot)$ ,  $t \in \mathcal{T}$ , is then given by

$$\hat{q}_{t,N_t}(\alpha) \equiv \inf\{y \in \mathbf{R}_{++} : \hat{F}_{t,N_t}(y) > \alpha\}. \quad (18)$$

Under standard regularity conditions, the estimator in Equation (18) is consistent for the true  $\alpha$ -quantile under  $F_t(\cdot)$ , defined in Equation (3). Consistency of  $\hat{q}_{t,N_t}(\alpha)$  can be extended to cases where the sample  $(X_{t,1}, \dots, X_{t,N_t})$  is weakly dependent, provided additional moment assumptions (Pollard, 1991; Portnoy, 1991; Koenker and Zhao, 1996; Komunjer, 2005; Chernozhukov, 2005); for the sake of simplicity, we focus on the independent case.

To alleviate the notation, we drop the reference to  $t$  when doing so introduces no ambiguities. Hence we use the notation  $N$ ,  $\hat{F}_N(\cdot)$  and  $\hat{q}_N(\alpha)$  to denote the size of the sample  $(X_1, \dots, X_N)$ , the corresponding empirical distribution function and the  $\alpha$ -quantile estimator in Equation (18).

As pointed out previously, the main object of interest are  $\alpha$ -quantiles with probabilities  $\alpha$  close to unity. How close  $\alpha$  is to 1 is determined by the sample size  $N$ ; hence we 

---

the support of  $\theta$  as  $T = t$  (i.e.  $\mathcal{E}_t$ ) or the corresponding probabilities  $p_t$ ; this however is not our objective here.

let this probability be a function of the sample size, and we denote it by  $\alpha_N$ . Knowing how  $\alpha$  varies with  $N$  will then enable us to answer the question: for a given sample size  $N$  how large  $\alpha$  needs to be for the ordering in Theorem 2(ii) to hold.

## 5.2 Central Limit Theory for Intermediate Empirical Quantiles

We now derive the asymptotic distribution of  $\hat{q}_N(\alpha_N)$  in Equation (18) when  $\lim_{N \rightarrow \infty} \alpha_N = 1$  and when  $(1 - \alpha_N)N$  has a positive limit as  $N$  goes to infinity. In particular, we consider the case where  $\lim_{N \rightarrow \infty} (1 - \alpha_N)N = \infty$ . This last condition describes how fast  $\alpha$  has to go to unity relative to the sample size  $N$ ; knowing that  $N^{-1} = o(1 - \alpha_N)$  we can use an appropriate limit theory result to derive an asymptotic distribution of the  $\alpha$ -quantile estimator  $\hat{q}_N(\alpha_N)$  in Equation (18).

We shall need the following lemma.

**Lemma 4.** *Consider a random sample  $(X_1, \dots, X_N)$  of size  $N$  from  $F(\cdot)$  and let  $\hat{q}_N(\alpha_N)$  be the corresponding empirical  $\alpha_N$ -quantile. If the distribution tail  $\bar{F} \in \Gamma$  with auxiliary function  $a(\cdot)$  and with density  $f(\cdot)$  which is eventually non-increasing, then, provided  $\lim_{N \rightarrow \infty} \alpha_N = 1$  and  $\lim_{N \rightarrow \infty} (1 - \alpha_N)N = \infty$  we have:*

$$\sqrt{N(1 - \alpha_N)} \frac{\hat{q}_N(\alpha_N) - q(\alpha_N)}{a(q(\alpha_N))} \xrightarrow{d} \mathcal{N}$$

and

$$\frac{\hat{q}_N(\beta_N) - \hat{q}_N(\alpha_N)}{a(q(\alpha_N))} \xrightarrow{p} \ln \rho$$

where  $\mathcal{N}$  is a standard Gaussian random variable and  $\beta_N$  is such that  $\alpha_N < \beta_N < 1$  and  $(1 - \alpha_N)/(1 - \beta_N) \rightarrow \rho$  with  $\rho > 1$ .

Lemma 4 presents two limit results. The first was proved by Falk (1989). The second is new.

The first result in the lemma shows the asymptotic behavior of intermediate empirical quantiles when  $\alpha_N$  depends on the sample size  $N$ . It is an extension of the well-known result for central  $\alpha$ -quantiles with  $\alpha \in (0, 1)$  fixed (Mosteller, 1946; Smirnov, 1952; Siddiqui,

1960; Bahadur, 1966; Bassett and Koenker, 1978; Powell, 1984, 1986), to the case where  $\alpha$  increases with the sample size  $N$ . Dekkers and de Haan (1989) and Chernozhukov (2005) prove this extension under an additional assumption on the tail behavior of  $F(\cdot)$ . 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 4 is important because it gives us a consistent estimator of the variance of the empirical quantile. Recall that Theorem 2 says that the conditional quantiles of  $X$  must be increasing in  $t$ . 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  $t$ . Our second result in Lemma 4 allows us to solve the problem.

The second limit result of Lemma 4 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)). The result by Dekkers and de Haan (1989) requires that  $d\bar{F}^{-1}(y)/dy$  is in  $\mathcal{R}_{-1}$  at 0, an assumption that we need to avoid because it would imply a restriction on the equilibrium-selection probabilities,  $p_t(\cdot)$ . By focusing on consistency, and not on the asymptotic distribution of quantile spacings, we obtain a result only assuming that  $\bar{F}(\cdot)$  is in  $\Gamma$  and that  $f(\cdot)$  is eventually non-increasing. Consistency, in turn, is sufficient for our testing procedure.

We should note that the assumption that  $f(\cdot)$  be eventually non-increasing imposes no restriction on the equilibrium-selection probabilities  $p_t(\cdot)$ , and follows from requiring the density of  $H_{\theta,t}(\cdot)$  to be eventually non-increasing.

### 5.3 Estimates for Conditional Quantiles under $((\mathcal{E}_t, p_t, (H_{\theta,t})_{\theta \in \mathcal{E}_t}), t \in \mathcal{T})$

We now assume a collection of random samples for different values of  $T$ . Concretely, consider a statistical reduced-form model  $((\mathcal{E}_t, p_t, (H_{\theta,t})_{\theta \in \mathcal{E}_t}), t \in \mathcal{T})$ , and assume we observe realizations from  $k \geq 2$  random samples  $(X_{t_1,1}, \dots, X_{t_1,N_{t_1}})$  to  $(X_{t_k,1}, \dots, X_{t_k,N_{t_k}})$  of sizes

$N_{t_1}$  to  $N_{t_k}$ , respectively. The samples are assumed independent and drawn from the  $k$  distributions  $F_{t_1}(\cdot)$  to  $F_{t_k}(\cdot)$ , respectively, with  $(t_1, \dots, t_k) \in \mathcal{T}^k$ .

In order to use the results of Lemma 4 we need to impose the following assumption on the tails of the distributions  $H_{\theta,t}(\cdot)$ :

**Assumption S4:** *Say that a statistical reduced-form model  $((\mathcal{E}_t, p_t, (H_{\theta,t})_{\theta \in \mathcal{E}_t}), t \in \mathcal{T})$  satisfies S4 if it satisfies S3 and, in addition, for every  $(t, \theta_t) \in \mathcal{T} \times \mathcal{E}_t$  the densities  $h_{\theta_t,t}(\cdot)$  are eventually non-increasing.*

The limit results of Lemma 4 then yield the following result:

**Theorem 5.** *Assume  $((\mathcal{E}_t, p_t, (H_{\theta,t})_{\theta \in \mathcal{E}_t}), t \in \mathcal{T})$  satisfies S4, and let  $(X_{t_j,1}, \dots, X_{t_j,N_{t_j}})$ ,  $1 \leq j \leq k$ , be  $k$  independent random samples of size  $N_{t_j}$  from  $F_{t_j}(\cdot)$  with  $t_j \in \mathcal{T}$ . If for all  $j$ ,  $1 \leq j \leq k$ ,  $0 < \alpha_{N_{t_j}} < \beta_{N_{t_j}} < 1$ ,  $\lim_{N_{t_j} \rightarrow \infty} \alpha_{N_{t_j}} = 1$ ,  $\lim_{N_{t_j} \rightarrow \infty} (1 - \alpha_{N_{t_j}}) N_{t_j} = \infty$  and  $\lim_{N_{t_j} \rightarrow \infty} (1 - \alpha_{N_{t_j}}) / (1 - \beta_{N_{t_j}}) = \rho_{t_j}$  with  $\rho_{t_j} > 1$ , then*

$$\sqrt{(1 - \alpha_{N_{t_j}}) N_{t_j} (\ln \rho_{t_j})} \left[ \frac{\hat{q}_{t_j, N_{t_j}}(\alpha_{N_{t_j}}) - q_{t_j}(\alpha_{N_{t_j}})}{\hat{q}_{t_j, N_{t_j}}(\alpha_{N_{t_j}}) - \hat{q}_{t_j, N_{t_j}}(\beta_{N_{t_j}})} \right] \xrightarrow{d} \mathcal{N}_j,$$

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

For any given  $k \geq 2$ , the results of Theorem 5 allow us to determine the asymptotic behavior of estimates for conditional quantiles under  $F_t(\cdot)$ . With conditional quantile estimators in hand, we can then test the implications in Theorem 2.

For the purpose of testing, we make an assumption on the rate of growth of the different samples. The assumption ensures that the  $(1 - \alpha_{N_{t_j}}) N_{t_j}$  grow at the same speed, and that we consider the same  $\alpha_N$ -quantile, for all  $k$  samples. We can then formulate our results in the standard asymptotic framework, i.e. as  $N \rightarrow \infty$ . Concretely, assume that the sample sizes  $(N_{t_1}, \dots, N_{t_k})$  and the corresponding probabilities  $(\alpha_{N_{t_1}}, \dots, \alpha_{N_{t_k}})$  are such that there exists  $\alpha_N$ ,  $0 < \alpha_N < 1$ ,  $\lim_{N \rightarrow \infty} (1 - \alpha_N) N = \infty$ , and  $c_{t_j}$ ,  $1 \leq j \leq k$ , such that

$$\alpha_{N_{t_j}} = \alpha_N \text{ and } N_{t_j} = c_{t_j} N, \text{ with } c_{t_j} > 0 \text{ for every } j, 1 \leq j \leq k. \quad (19)$$

## 6 Testing

### 6.1 Test Hypotheses

From Theorem 2, the implication of our statistical reduced-form model  $((\mathcal{E}_t, p_t, (H_{\theta,t})_{\theta \in \mathcal{E}_t}), t \in \mathcal{T})$  is that  $t_1 < \dots < t_k$  implies  $q_{t_1}(\alpha_N) < \dots < q_{t_k}(\alpha_N)$  as  $\alpha_N \rightarrow 1$ . Hence, we are interested in testing whether an increase in the exogenous variable results in an increase in conditional quantiles of the endogenous variable. The opposite case of interest is the one in which an increase in  $t$  produces no effect on the conditional quantiles of  $X$ , so that  $t_1 < \dots < t_k$  and  $q_{t_1}(\alpha_N) = \dots = q_{t_k}(\alpha_N)$ . Those two cases define our null and alternative hypotheses.

More formally, for given values  $t_1 < \dots < t_k$  we test the null hypothesis  $H_0 : q_{t_1}(\alpha_N) = \dots = q_{t_k}(\alpha_N)$ , as  $\alpha_N \rightarrow 1$ , against an ordered alternative  $H_1 : q_{t_1}(\alpha_N) \leq \dots \leq q_{t_k}(\alpha_N)$ , as  $\alpha_N \rightarrow 1$ , with strict inequality for at least one value of  $j$ ,  $1 \leq j \leq k$ .

Our test statistic is a function of estimates for conditional quantiles under  $F_t(\cdot)$ ; from Theorems 2 and 5 we know that the latter satisfy the following property:

**Corollary 6.** *Assume  $((\mathcal{E}_t, p_t, (H_{\theta,t})_{\theta \in \mathcal{E}_t}), t \in \mathcal{T})$  satisfies S2 and S4. If for  $N \in \mathbb{N}$ ,  $(X_{t_j,1}, \dots, X_{t_j, c_{t_j} N}), 1 \leq j \leq k$ , are  $k$  independent random samples of size  $c_{t_j} N$  from  $F_{t_j}(\cdot)$ ,  $t_j \in \mathcal{T}$ , then, as  $N \rightarrow \infty$ ,  $t_1 < \dots < t_k$  implies  $\mu_{t_1} < \dots < \mu_{t_k}$ , with*

$$\mu_{t_j} \equiv q_{t_j}(\alpha_N), \frac{\hat{q}_{t_j, c_{t_j} N}(\alpha_N) - \mu_{t_j}}{\hat{\sigma}_{t_j}} \xrightarrow{d} \mathcal{N}_j, \text{ and } \hat{\sigma}_{t_j} \equiv \frac{\hat{q}_{t_j, c_{t_j} N}(\alpha_N) - \hat{q}_{t_j, c_{t_j} N}(\beta_N)}{(\ln \rho) \sqrt{c_{t_j} (1 - \alpha_N) N}},$$

*provided  $0 < \alpha_N < \beta_N < 1$ ,  $\lim_{N \rightarrow \infty} \alpha_N = 1$ ,  $\lim_{N \rightarrow \infty} (1 - \alpha_N) N = \infty$  and  $\lim_{N \rightarrow \infty} (1 - \alpha_N) / (1 - \beta_N) = \rho$ , with  $\rho > 1$ , and where  $\mathcal{N}_1, \dots, \mathcal{N}_k$  are  $k$  independent standard normal random variables.*

### 6.2 Exact test for Order Restrictions

Assume for the moment that all the distribution results from Corollary 6 are exact rather than being asymptotic, i.e. assume that for some probability  $\alpha_N$  close to 1 and for large

enough  $N$ ,  $(\hat{q}_{t_1, c_{t_1} N}(\alpha_N), \dots, \hat{q}_{t_k, c_{t_k} N}(\alpha_N))$  is a sample from  $k$  independent and normally-distributed random variables with means  $(\mu_{t_1}, \dots, \mu_{t_k})$  and variances  $(\hat{\sigma}_{t_1}^2, \dots, \hat{\sigma}_{t_k}^2)$ . Our null and alternative hypotheses are then equivalent to  $H_0 : \mu_{t_1} = \dots = \mu_{t_k}$  and  $H_1 : \mu_{t_1} \leq \dots \leq \mu_{t_k}$  with at least one strict inequality. Note that having observed  $\hat{q}_{t_j, N_{t_j}}(\alpha_N)$  and  $\hat{q}_{t_j, N_{t_j}}(\beta_N)$ , the variances  $\hat{\sigma}_{t_j}^2$ ,  $1 \leq j \leq k$ , are known. So the implications of our reduced-form model in Section 2.3 can be restated in terms of the means  $(\mu_{t_1}, \dots, \mu_{t_k})$  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 (e.g. Bartholomew (1959a,b); Barlow, Bartholomew, Bremner, and Brunk (1972) and Robertson and Wegman (1978)). We shall review Bartholomew's results, as they are instrumental in the extension to the our case of asymptotic normality.

We introduce the following notation:  $\hat{q} \equiv (\hat{q}_{t_1, c_{t_1} N}(\alpha_N), \dots, \hat{q}_{t_k, c_{t_k} N}(\alpha_N))'$ ,  $\mu \equiv (\mu_{t_1}, \dots, \mu_{t_k})'$  and  $\hat{\Sigma} \equiv \text{diag}(\hat{\sigma}_{t_1}^2, \dots, \hat{\sigma}_{t_k}^2)$ . Hence, for a given value of  $N$ , the  $k$ -vector  $\hat{q}$  is multivariate normal with mean  $\mu$  and diagonal covariance matrix  $\hat{\Sigma}$ . Letting  $A$  be a  $(k-1) \times k$ -matrix

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

we can write the null and the alternative hypotheses as

$$H_0 : \{A\mu = 0\} \text{ against } H_1 : \{A\mu \preceq 0 \text{ and } A\mu \neq 0\}, \quad (20)$$

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

The likelihood-ratio test of the above hypothesis is based on the statistic

$$\hat{\xi}_{LR} \equiv -2 \ln \frac{\max_{A\mu=0} L(\hat{q}|\mu, \hat{\Sigma})}{\max_{A\mu \preceq 0} L(\hat{q}|\mu, \hat{\Sigma})}, \quad (21)$$

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

$$L(\hat{q}|\mu, \hat{\Sigma}) = \frac{1}{(2\pi)^{k/2} (\det \hat{\Sigma})^{1/2}} \exp \left[ -(\hat{q} - \mu)' \hat{\Sigma}^{-1} (\hat{q} - \mu) \right]. \quad (22)$$

Combining Equations (21) and (22) then yields

$$\hat{\xi}_{LR} = \min_{A\mu=0} (\hat{q} - \mu)' \hat{\Sigma}^{-1} (\hat{q} - \mu) - \min_{A\mu \leq 0} (\hat{q} - \mu)' \hat{\Sigma}^{-1} (\hat{q} - \mu). \quad (23)$$

Barlow, Bartholomew, Bremner, and Brunk (1972) show that the test statistic in Equation (23)—similar to the  $\chi^2$  statistic used to test for  $H_0$  against the most general form of alternative  $H_2 : \mu_{t_1} \neq \dots \neq \mu_{t_k}$ —is a weighted average of  $\chi^2$  distributions with  $d$  degrees of freedom ( $\chi_d^2$ ) with  $0 \leq d \leq k - 1$ , and is denoted by  $\bar{\chi}_k^2$ . ( $\chi_0^2$  denotes a point mass at 0.)

The  $\bar{\chi}_k^2$  distribution of the likelihood-ratio test statistic  $\hat{\xi}_{LR}$  depends on the number of quantiles being compared  $k$ , as well as their variances  $\hat{\sigma}_{t_j}^2$  through the probability weights attached to each distribution  $\chi_d^2$ . For example, when  $k = 2$  and 3, the distribution of  $\hat{\xi}_{LR}$  is given by

$$\hat{\xi}_{LR} \stackrel{d}{=} \frac{1}{2} \chi^2(0) + \frac{1}{2} \chi^2(1), \text{ for } k = 2, \quad (24)$$

$$\hat{\xi}_{LR} \stackrel{d}{=} \frac{\hat{\alpha}}{2\pi} \chi^2(0) + \frac{1}{2} \chi^2(1) + \left[ \frac{1}{2} - \frac{\hat{\alpha}}{2\pi} \right] \chi^2(2), \text{ for } k = 3, \quad (25)$$

where the constant  $\hat{\alpha}$ ,  $-\pi < \hat{\alpha} < \pi$ , is defined as  $\hat{\alpha} \equiv \arccos \left[ \hat{\sigma}_{t_2}^2 / \sqrt{(\hat{\sigma}_{t_1}^2 + \hat{\sigma}_{t_2}^2)(\hat{\sigma}_{t_2}^2 + \hat{\sigma}_{t_3}^2)} \right]$ .

In the special case where the variances  $\hat{\sigma}_{t_j}^2$  are equal Bartholomew (1959b) computes the  $\bar{\chi}_k^2$  critical values for a number of values for  $k$  ( $2 \leq k \leq 12$ ). When the variances are different, exact critical values for  $\bar{\chi}_k^2$  are hard to obtain if  $k \geq 5$ , though there is no difficulty in obtaining their numerical values for any  $k$  (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 case obtained in Section 5, 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  $\text{Id}_k$ .

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

**Assumption S5:** *Say that a statistical reduced-form model  $((\mathcal{E}_t, p_t, (H_{\theta,t})_{\theta \in \mathcal{E}_t}), t \in \mathcal{T})$  satisfies S5 if it satisfies S4 and, in addition,  $H_{\theta,t}(\cdot) = H(\cdot)$  for every  $(t, \theta) \in \mathcal{T} \times \Theta$ .*

In the error model (5), for example, the assumption S5 is satisfied if the error  $\varepsilon$  is independent of the exogenous variable  $T$  and the equilibrium level  $\theta$ . When the distributions  $H_{\theta,t}(\cdot)$  do not depend on  $t$  or  $\theta$ , the same is true for the quantities involved in the previous limit results. In particular, under S5 we have that  $a_{\bar{\theta}_t,t} \circ \bar{H}_{\bar{\theta}_t,t}^{-1}(\cdot) = a \circ \bar{H}^{-1}(\cdot)$  for all  $(t, \theta_t) \in \mathcal{T} \times \mathcal{E}_t$ ; so using Lemma 3 (ii) gives us

$$\frac{\sigma_{t_j}}{\sigma_{t_i}} = \frac{a_{t_i}(q_{t_j}(\alpha_N))/\sqrt{c_{t_j}(1-\alpha_N)N}}{a_{t_i}(q_{t_i}(\alpha_N))/\sqrt{c_{t_i}(1-\alpha_N)N}} \sim \sqrt{\frac{c_{t_i} \bar{\theta}_{t_i}}{c_{t_j} \bar{\theta}_{t_j}}}, \text{ as } \alpha_N \rightarrow 1. \quad (26)$$

In other words, the scaling constants  $\sigma_{t_j}$  that control how fast the empirical quantiles  $\hat{q}_{t_j, c_{t_j} N}(\alpha_N)$  converge to the true quantiles  $q_{t_j}(\alpha_N)$ , are all of the same size. In that case, we have the following result.

**Theorem 7.** *Assume  $((\mathcal{E}_t, p_t, (H_{\theta,t})_{\theta \in \mathcal{E}_t}), t \in \mathcal{T})$  satisfies S5. If for  $N \in \mathbb{N}$ ,  $(X_{t_j,1}, \dots, X_{t_j, c_{t_j} N})$ ,  $1 \leq j \leq k$ , are  $k$  independent samples of  $c_{t_j} N$  independent draws from  $F_{t_j}(\cdot)$ ,  $t_j \in \mathcal{T}$ , then as  $N \rightarrow \infty$ , the likelihood-ratio statistic  $\hat{\xi}_{LR}$  is asymptotically distributed as  $\bar{\chi}_k^2$ , 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 (24) – (25).

The conclusion of Theorem 7 remains valid if, instead of being independent of  $\theta$  and  $t$ , the distributions  $H_{\theta,t}(\cdot)$  are such that for any  $(t, t') \in \mathcal{T}^2$ , the ratios  $h_{\bar{\theta}_t,t}(\bar{H}_{\bar{\theta}_t,t}^{-1}(y))/h_{\bar{\theta}_{t'},t'}(\bar{H}_{\bar{\theta}_{t'},t'}^{-1}(y))$  have some finite and strictly positive limit as  $y$  goes to zero.

## 7 Conclusion

In this paper we design an econometric test for monotone comparative statics predictions suited for testing models with multiple equilibria. Many economic models possess multiple equilibria, which has traditionally posed a severe challenge for deriving testable implications. 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 methods without assuming an equilibrium-selection rule.

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

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 any asymptotic extension of the “chi-bar” squared test.

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 exogenous variable—would offer an interesting alternative way of testing the monotonicity

prediction.

## References

- AGUIRREGABIRIA, V., AND P. MIRA (2004): “Sequential Estimation of Dynamic Discrete Games,” Mimeo, Boston University.
- AMIR, R. (1996): “Cournot Oligopoly and the Theory of Supermodular Games,” *Games and Economic Behavior*, 15, 132–148.
- AMIR, R., AND V. E. LAMBSON (2000): “On the Effects of Entry in Cournot Markets,” *Review of Economic Studies*, 67(2), 235–254.
- ANDREWS, D. W. K., S. BERRY, AND P. JIA (2004): “Confidence Regions for Parameters in Discrete Games with Multiple Equilibria, with an Application to Discount Chain Store Location,” Mimeo, Yale University.
- ARADILLAS-LÓPEZ, A. (2004): “Semiparametric Estimation of a Simultaneous Game with Incomplete Information,” Mimeo, UC Berkeley.
- 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., AND H. HONG (2005): “Semiparametric Estimation of Dynamic Games of Incomplete Information,” Mimeo, Duke University.
- BAJARI, P., H. HONG, AND S. RYAN (2004): “Identification and Estimation of Discrete Games of Complete Information,” Mimeo, Duke University.
- BAJARI, P. L., C. L. BENKARD, AND J. LEVIN (2004): “Estimating Dynamic Models of Imperfect Competition,” NBER Working Paper 10450.

- 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., J. LEVINSOHN, AND A. PAKES (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63(4), 841–890.
- BERRY, S. T. (1992): "Estimation of a Model of Entry in the Airline Industry," *Econometrica*, 60(4), 889–917.
- BJORN, P. A., AND Q. H. VUONG (1984): "Simultaneous 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 Social Science Working Paper 537.
- BOUND, J., C. BROWN, G. DUNCAN, AND W. RODGERS (1994): "Evidence on the Validity of Cross-Sectional and Longitudinal Labor Market Data," *Journal of Labor Economics*, 12, 345–368.
- BRESNAHAN, T. F., AND P. C. REISS (1990): "Entry in Monopoly Markets," *Review of Economic Studies*, 57(4), 531–553.
- (1991): "Empirical Models of Discrete Games," *Journal of Econometrics*, 48, 57–81.

- CARROLL, R. J., D. RUPPERT, AND L. A. STEFANSKI (1995): *Measurement Error in Nonlinear Models*. Chapman and Hall.
- CHEN, X., H. HONG, AND E. TAMER (2005): “Measurement Error Models with Auxiliary Data,” *Review of Economic Studies*, 72, 343–366.
- CHERNOZHUKOV, V. (2005): “Extremal Quantile Regression,” *Annals of Statistics*, 33, forthcoming.
- CHERNOZHUKOV, V., AND H. HONG (2004): “Likelihood Estimation and Inference in a Class of Nonregular Econometric Models,” *Econometrica*, 72(5), 1445–1480.
- CILIBERTO, F., AND E. TAMER (2004): “Market Structure and Multiple Equilibria in the Airline Industry,” Mimeo, Northeastern University and University of Virginia.
- 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.
- 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.
- EMBRECHTS, P., C. KLUPPELBERG, AND T. MIKOSCH (1997): *Modelling Extremal Events for Insurance and Finance*. Springer-Verlag.
- ERICSON, R., AND A. PAKES (1995): “Markov-Perfect Industry Dynamics: A Framework for Empirical Work,” *Review of Economic Studies*, 62(1), 53–82.

- FALK, M. (1989): “A Note on Uniform Asymptotic Normality of Intermediate Order Statistics,” *Annals of the Institute of Statistical Mathematics*, 41, 19–29.
- 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.
- JOVANOVIC, B. (1989): “Observable Implications of Models with Multiple Equilibria,” *Econometrica*, 57(6), 1431–1437.
- JUHN, C. (1992): “Decline of Male Labor Market Participation: The Role of Declining Market Opportunities,” *Quarterly Journal of Economics*, 107(1), 79–121.
- KARAMATA, J. (1930): “Sur un mode de croissance régulière des fonctions,” *Mathematica*, 5, 38–53.
- KIM, K. I. (2005): “Semiparametric Estimation of Signaling Games,” Mimeo, UCLA.
- 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.
- LJUNGQVIST, L., AND T. J. SARGENT (2004): *Recursive Macroeconomic Theory*. MIT Press (Second Edition).

- 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 (1990): “Rationalizability, Learning and Equilibrium in Games with Strategic Complementarities,” *Econometrica*, 58(6), 1255–1277.
- (1994): “Comparing Equilibria,” *American Economic Review*, 84(3), 441–459.
- MILGROM, P., AND C. SHANNON (1994): “Monotone Comparative Statics,” *Econometrica*, 62(1), 157–180.
- MOSTELLER, F. (1946): “On some Useful ”Inefficient” Statistics,” *Annals of Mathematical Statistics*, 17, 377–408.
- PAKES, A., M. OSTROVSKY, AND S. BERRY (2004): “Simple Estimators for the Parameters of Discrete Dynamic Games,” Mimeo, Harvard University.
- PESENDORFER, M., AND P. SCHMIDT-DENGLER (2004): “Identification and Estimation of Dynamic Games,” NBER Working Paper 9726.
- POLLARD, D. (1991): “Asymptotics for Least Absolute 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.
- 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. (2005): “Coordination Games, Multiple Equilibria and the Timing of Radio Commercials,” Mimeo, Northwestern University.
- TAMER, E. (2003): “Incomplete Simultaneous Discrete Response Model with Multiple Equilibria,” *Review of Economic Studies*, 70(1), 147–165.
- TOPKIS, D. M. (1979): “Equilibrium Points in Nonzero-Sum n-Person Submodular Games,” *SIAM Journal of Control and Optimization*, 17(6), 773–787.
- (1998): *Supermodularity and Complementarity*. Princeton University Press, Princeton, New Jersey.
- VILLAS-BOAS, J. M. (1997): “Comparative Statics of Fixed Points,” *Journal of Economic Theory*, 73(1), 183–198.
- VIVES, X. (1990): “Nash Equilibrium with Strategic Complementarities,” *Journal of Mathematical Economics*, 19(3), 305–321.
- (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.

## Appendix A Preliminary Lemmas.

Lemmas 8, 9 and 10 are simple translations of results in de Haan (1970) to our problem. Lemmas 11, and 12 present more substantial preliminary results we shall need in the proof of Lemma 4.

In the sequel,  $\bar{H}(\cdot)$  is a distribution tail  $\bar{H} : \mathbf{R}_{++} \rightarrow (0, 1)$  and  $\bar{H}^{-1}(\cdot)$  the corresponding quantile function  $\bar{H}^{-1} : (0, 1) \rightarrow \mathbf{R}_{++}$ .

**Lemma 8.** *If  $\bar{H} \in \mathcal{R}_{-\infty}$  at  $\infty$ , then  $\bar{H}^{-1} \in \mathcal{R}_0$  at 0.*

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

**Lemma 9.** *If  $\bar{H}^{-1} \in \mathcal{R}_0$  at 0, then for all sequences  $\{a_N\}$  and  $\{a'_N\}$  of positive numbers with  $\lim_{N \rightarrow \infty} a_N = \lim_{N \rightarrow \infty} a'_N = 0$  and  $\lim_{N \rightarrow \infty} a_N/a'_N = c$  ( $0 < c < \infty$ ), we have*

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

*Proof.* Let  $U(x) \equiv \bar{H}^{-1}(1/x)$  for all  $x > 1$  so  $U \in \mathcal{R}_0$  at  $\infty$ . 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  $\lim_{N \rightarrow \infty} \alpha_N = \lim_{N \rightarrow \infty} \alpha'_N = \infty$  and  $\lim_{N \rightarrow \infty} \alpha_N/\alpha'_N = 1/c$  ( $0 < 1/c < \infty$ ). By applying Corollary 1.2.1 (2) in de Haan (1970) we then have  $\lim_{N \rightarrow \infty} \bar{H}^{-1}(a_N)/\bar{H}^{-1}(a'_N) = \lim_{N \rightarrow \infty} U(\alpha_N)/U(\alpha'_N) = (1/c)^0 = 1$ . ■

**Lemma 10.** *Consider a distribution tail  $\bar{H} \in \Gamma$  with auxiliary function  $a(\cdot)$ . Let  $H(\cdot)$  be twice differentiable on  $\mathbf{R}_{++}$  with a density  $h(\cdot)$  that is eventually non-increasing. Let  $W(x) \equiv \bar{H}^{-1}(\exp(-x))$ , for  $x > 0$ . Then  $W(\cdot)$  is twice continuously differentiable on  $\mathbf{R}_{++}$  with  $W'(x) = a[\bar{H}^{-1}(\exp(-x))]$ , for  $x > 0$ , and for any real interval  $[a, b]$  we have:*

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

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

*Proof.* First we prove (i). Note that  $a(W(x)) = \exp(-x)/h(W(x)) = W'(x)$ . From Lemma 3 we know that  $a \circ \bar{H}^{-1} \in \mathcal{R}_0$  at 0, so

$$\lim_{x \rightarrow \infty} \frac{W'(x+s)}{W'(x)} = \lim_{x \rightarrow \infty} \frac{a(\bar{H}^{-1}(\exp(-x-s)))}{a(\bar{H}^{-1}(\exp(-x)))} = 1, \text{ for } s > 0. \quad (27)$$

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  $a(W(x)) = W'(x)$  implies that

$$\frac{W''(x+s)}{W'(x)} = \left[ \frac{W'(x+s)}{W'(x)} \right] a'(W(x+s)). \quad (28)$$

The bracketed term on the right-hand side of Equation (28) converges to 1 uniformly on  $[a, b]$  by item (i) of the Lemma. We shall prove that  $a'(W(x+s)) \rightarrow 0$  as  $x \rightarrow \infty$  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

$$\lim_{x \rightarrow \infty} \frac{\bar{H}(x)h'(x)}{[h(x)]^2} = -1, \text{ i.e. } \lim_{x \rightarrow \infty} a'(x) = 0. \quad (29)$$

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

$$a'(y(x)) = \sup_{y \in [W(x+a), W(x+b)]} a'(y). \quad (30)$$

Now,  $y(x) \rightarrow \infty$  as  $x \rightarrow \infty$  because  $W(\cdot)$  is monotone increasing. Then the right-hand-side of Equation (30) converges to 0 as  $x \rightarrow \infty$ , because  $a'(\cdot)$  converges to 0 (29). This proves the needed uniform convergence of  $a'(W(x+s))$  in Equation (28). ■

**Lemma 11.** *Let  $\{c_N\}$  be a sequence of strictly positive real numbers such that  $\lim_{N \rightarrow \infty} c_N = \infty$ , and consider  $f : \mathbf{R} \rightarrow \mathbf{R}$ . Let  $\{X_N\}$  and  $\{Y_N\}$  be two independent stochastic processes. If*

1.  $X_N \xrightarrow{d} X$ , as  $N \rightarrow \infty$ , for some  $X$  with continuous distribution  $F(\cdot)$ ,
2.  $Y_N \xrightarrow{as} \infty$ , as  $N \rightarrow \infty$ ,
3. for each  $K > 0$ ,  $\lim_{y \rightarrow \infty} f(x + y) = 1$  uniformly in  $x \in [-K, K]$ .

Then

$$c_N \int_0^{X_N/c_N} f(x + Y_N) dx \xrightarrow{d} X, \text{ as } N \rightarrow \infty.$$

*Proof.* Fix a realization  $\{y_N\}$  of  $\{Y_N\}$  such that  $\lim_{N \rightarrow \infty} y_N = \infty$ ; the almost sure convergence in item 2 ensures that  $\{y_N\}$  with  $\lim_{N \rightarrow \infty} y_N = \infty$  have full measure. Let  $z \in \mathbf{R}_+$  and denote by  $B_N$  the event

$$\left\{ c_N \int_0^{X_N/c_N} f(x + y_N) dx \leq z \right\}.$$

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

Fix  $\delta > 0$  such that  $F(z/(1 - \delta)) - F(z/(1 + \delta)) < \varepsilon/4$ . Let  $K \in \mathbf{R}$  be large enough that  $K > z/(1 - \delta)$ ,  $F(-K) < \varepsilon/4$  and  $1 - F(K) < \varepsilon/4$ . Since  $X_N \xrightarrow{d} X$ , there is  $N_1$  such that  $n \geq N_1$  implies

$$P\{|X_N| > K\} < \varepsilon/2 \tag{31}$$

$$F(z/(1 + \delta)) - \varepsilon/4 < P\{X_N \leq z/(1 + \delta)\} \tag{32}$$

$$P\{X_N \leq z/(1 - \delta)\} < F(z/(1 - \delta)) + \varepsilon/4 \tag{33}$$

Let  $B_N^K = B_N \cap \{|X_N| \leq K\}$ . Then, by Statement (31),  $N \geq N_1$  implies that  $P(B_N) - P(B_N^K) \leq P\{|X_N| > K\} < \varepsilon/2$ .

The convergence in item 3 is uniform on  $[-K, K]$ , so there is  $N^*$  such that  $N^* \geq N_1$  and such that  $N \geq N^*$  implies that, for all  $\tilde{x} \in [-K, K]$ ,  $(1 - \delta) < f(\tilde{x} + y_N) < (1 + \delta)$ .

Then,  $N \geq N^*$  implies

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

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

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

if  $\tilde{x} < 0$ .

Then  $P\{X_N(1 + \delta) \leq z, X_N \geq 0\} \leq P(B_N^K) \cap \{X_N \geq 0\} \leq P\{X_N(1 - \delta) \leq z, X_N \geq 0\}$ .

And since  $z \geq 0$ ,  $P\{X_N(1 - \delta) \leq z, X_N < 0\} = P(B_N^K) \cap \{X_N < 0\} = P\{X_N(1 + \delta) \leq z, X_N < 0\}$ .

Hence,  $P\{X_N(1 + \delta) \leq z\} \leq P(B_N^K) \leq P\{X_N(1 - \delta) \leq z\}$ .

Now,

$$\begin{aligned} |F(z) - P\{X_N(1 + \delta) \leq z\}| &\leq |F(z) - F(z/(1 + \delta))| \\ &\quad + |F(z/(1 + \delta)) - P\{X_N(1 + \delta) \leq z\}| \\ &\leq \varepsilon/4 + \varepsilon/4, \end{aligned}$$

by the definition of  $\delta$  and Statement (32). And similarly for  $P\{X_N(1 - \delta) \leq z\}$ . So

$$|F(z) - P(B_N^K)| < \varepsilon/2.$$

Finally, then,  $N \geq N^*$  implies that

$$\begin{aligned} |F(x) - P(B_N)| &\leq |F(z) - P(B_N^K)| + |P(B_N) - P(B_N^K)| \\ &< \varepsilon/2 + \varepsilon/2. \end{aligned}$$

The argument for  $z < 0$  is analogous. The proof follows because  $\{X_N\}$  and  $\{Y_N\}$  are independent. ■

**Lemma 12.** *Let  $\{c_N\}$  be a sequence of strictly positive real numbers such that  $\lim_{N \rightarrow \infty} c_N = \infty$ , and consider  $f : \mathbf{R} \rightarrow \mathbf{R}$ . Let  $\{X_N\}$  be a stochastic process and  $\{y_N\}$  a sequence of strictly positive real numbers. If*

1.  $X_N \xrightarrow{d} X$ , as  $N \rightarrow \infty$ , for some  $X$  with continuous distribution  $F(\cdot)$ ,

2.  $\lim_{N \rightarrow \infty} y_N = \infty$ ,

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

Then

$$c_N \int_0^{X_N/c_N} f(x + y_N) dx \xrightarrow{P} 0, \text{ as } N \rightarrow \infty.$$

*Proof.* Let  $\eta > 0$  and denote by  $B_N$  the event

$$\left\{ \left| c_N \int_0^{X_N/c_N} f(x + y_N) dx \right| \leq \eta \right\}.$$

We shall prove that  $P(B_N) \rightarrow 1$ .

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  $f(\cdot)$  on  $[-K, K]$ , there is  $N^*$  such that  $N \geq N^*$  implies that, for all  $\tilde{x} \in [-K, K]$ ,  $|f(\tilde{x} + y_N)| < \eta/K$ . Then, for all  $N \geq N^*$  and  $\tilde{x} \in [-K, K]$ ,

$$\left| c_N \int_0^{\tilde{x}/c_N} f(x + y_N) dx \right| \leq |\tilde{x}| \eta/K \leq \eta,$$

as  $|\tilde{x}| \leq K$ . So  $N \geq N^*$  implies that  $P(B_N) \geq P\{X_N \leq K\} > 1 - \varepsilon$ , by the definition of  $K$ . ■

**Lemma 13.** Let  $A_1, \dots, A_N$  be a random sample from  $F_A(x) = 1 - \exp(-x)$  with  $x > 0$ , and let  $A_{(1)}^{(N)} \leq \dots \leq A_{(N)}^{(N)}$  be the ascending order statistics of  $(A_1, \dots, A_N)$ . Consider orders  $(m, k) \in \mathbb{N}^2$  such that  $m < k \leq N$ . If  $m \rightarrow \infty$ ,  $k \rightarrow \infty$  and  $N \rightarrow \infty$  in a way that  $(N - m) \rightarrow \infty$ ,  $(N - m)/N \rightarrow 0$ ,  $(N - k) \rightarrow \infty$ ,  $(N - k)/N \rightarrow 0$  and  $(N - m)/(N - k) \rightarrow \rho$  where  $\rho > 1$ , then

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

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

*Proof.* Using Renyi (1953) representation, we know that  $\{A_{(N-k+1)}^{(N)} - A_{(N-k)}^{(N)}\}_{k=1}^N \stackrel{d}{=} \{Z_k/k\}_{k=1}^N$  where  $A_{(0)}^{(N)} \equiv 0$  and where  $Z_1, \dots, Z_N$  are independent and identically distributed standard exponential random variables. Then for any  $m$ ,  $1 \leq m \leq N$ , and any

$k, m < k \leq N$ , we have

$$A_{(m)}^{(N)} \stackrel{d}{=} \sum_{j=N-m+1}^N \frac{Z_j}{j} \text{ and } A_{(k)}^{(N)} - A_{(m)}^{(N)} \stackrel{d}{=} \sum_{l=N-k+1}^{N-m} \frac{Z_l}{l}, \quad (34)$$

which are independent. When  $m \rightarrow \infty, k \rightarrow \infty$  and  $N \rightarrow \infty$  in a manner that  $(N-m) \rightarrow \infty, (N-m)/N \rightarrow 0, (N-k) \rightarrow \infty, (N-k)/N \rightarrow 0$  and  $(N-m)/(N-k) \rightarrow \rho$  with  $\rho > 1$ , we can apply the central limit theorem in Liapunov's form to the sums of random variables in Equation (34) (see Theorem 4 in Renyi (1953), for example) to get

$$\frac{A_{(m)}^{(N)} - M_1}{S_1} \xrightarrow{d} \mathcal{N}_1 \text{ and } \frac{A_{(k)}^{(N)} - A_{(m)}^{(N)} - M_2}{S_2} \xrightarrow{d} \mathcal{N}_2, \quad (35)$$

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

$$\begin{aligned} M_1 &\equiv \sum_{j=N-m+1}^N \frac{1}{j} = \sum_{l=1}^N \frac{1}{l} - \sum_{n=1}^{N-m} \frac{1}{n} \\ &= \ln N + \gamma + O(N^{-1}) - \ln(N-m) - \gamma + O((N-m)^{-1}) \\ &= \ln \frac{N}{N-m} + O((N-m)^{-1}), \end{aligned} \quad (36)$$

$$\begin{aligned} S_1^2 &\equiv \sum_{j=N-m+1}^N \frac{1}{j^2} \\ &= \frac{1}{N-m+1} - \frac{1}{N} + \frac{\theta}{(N-m)(N-m+1)} \\ &= \frac{1}{N-m+1} + o((N-m)^{-1}), \end{aligned} \quad (37)$$

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

$$\begin{aligned} M_2 &\equiv \sum_{j=N-k+1}^{N-m} \frac{1}{j} = \sum_{l=1}^{N-m} \frac{1}{l} - \sum_{n=1}^{N-k} \frac{1}{n} \\ &= \ln \frac{N-m}{N-k} + O((N-m)^{-1}) \\ &= \ln \rho + O((N-m)^{-1}), \end{aligned} \quad (38)$$

$$\begin{aligned} S_2^2 &\equiv \sum_{j=N-k+1}^{N-m} \frac{1}{j^2} \\ &= \frac{1}{N-k+1} - \frac{1}{N-m} + \frac{\phi}{(N-k)(N-k+1)} \\ &= \frac{\rho-1}{N-m} + o((N-m)^{-1}), \end{aligned} \quad (39)$$

where  $0 < \phi < 1$  and  $\rho > 1$ . Combining Equations (35) – (39) then yields the result. ■

## Appendix B Proofs of results stated in the text.

*Proof of Lemma 1.* Fix  $t \in \mathcal{T}$  and assume  $\bar{H}_{\bar{\theta}_t, t} \in \mathcal{R}_{-\infty}$  at  $\infty$ . Let  $U_t : \mathbf{R}_{++} \rightarrow \mathbf{R}_{++}$  be given by

$$U_t(x) \equiv p_t(\bar{\theta}_t) \frac{\bar{H}_{\bar{\theta}_t, t}(x/\bar{\theta}_t)}{\bar{F}_t(x)}. \quad (40)$$

Note that  $U_t(\cdot)$  is well defined, as for any  $\theta_t \in \mathcal{E}_t$  we have  $\bar{H}_{\theta_t, t}(x) > 0$  and therefore  $\bar{F}_t(x) > 0$ , for all  $x > 0$ . The definition of  $F_t(\cdot)$  in Equation (2) implies that

$$U_t(x) = \frac{p_t(\bar{\theta}_t) \bar{H}_{\bar{\theta}_t, t}(x/\bar{\theta}_t)}{p_t(\bar{\theta}_t) \bar{H}_{\bar{\theta}_t, t}(x/\bar{\theta}_t) \left[ 1 + \sum_{\theta \in \mathcal{E}_t \setminus \{\bar{\theta}_t\}} \frac{p_t(\theta) \bar{H}_{\theta, t}(x/\theta)}{p_t(\bar{\theta}_t) \bar{H}_{\bar{\theta}_t, t}(x/\bar{\theta}_t)} \right]}. \quad (41)$$

Now note that

$$\frac{p_t(\theta) \bar{H}_{\theta, t}(x/\theta)}{p_t(\bar{\theta}_t) \bar{H}_{\bar{\theta}_t, t}(x/\bar{\theta}_t)} = \frac{p_t(\theta)}{p_t(\bar{\theta}_t)} \frac{\bar{H}_{\theta, t}(x/\theta)}{\bar{H}_{\bar{\theta}_t, t}(x/\bar{\theta}_t)} \frac{\bar{H}_{\bar{\theta}_t, t}(x/\bar{\theta}_t)}{\bar{H}_{\bar{\theta}_t, t}(x/\bar{\theta}_t)}, \quad (42)$$

and that for any  $\theta \in \mathcal{E}_t \setminus \{\bar{\theta}_t\}$ ,  $\theta < \bar{\theta}_t$ . Given that  $\bar{H}_{\bar{\theta}_t, t}(\cdot)$  is  $(-\infty)$ -varying at  $\infty$ , we have

$$\lim_{x \rightarrow \infty} \frac{\bar{H}_{\bar{\theta}_t, t}(x/\theta)}{\bar{H}_{\bar{\theta}_t, t}(x/\bar{\theta}_t)} = \lim_{s \rightarrow \infty} \frac{\bar{H}_{\bar{\theta}_t, t}(s\bar{\theta}_t/\theta)}{\bar{H}_{\bar{\theta}_t, t}(s)} = 0, \quad (43)$$

(with  $s = x/\bar{\theta}_t$ ), so the last term on the right hand side of Equation (42) goes to 0 as  $x$  gets large. Property (16) implies that the middle term  $\bar{H}_{\theta, t}(x/\theta)/\bar{H}_{\bar{\theta}_t, t}(x/\bar{\theta}_t)$  is bounded.

So

$$\lim_{x \rightarrow \infty} U_t(x) = 1, \quad (44)$$

and

$$\bar{F}_t(x) \sim p_t(\bar{\theta}_t) \bar{H}_{\bar{\theta}_t, t}(x/\bar{\theta}_t), \text{ as } x \rightarrow \infty. \quad (45)$$

From Equation (45)

$$\lim_{x \rightarrow \infty} \frac{\bar{F}_t(\lambda x)}{\bar{F}_t(x)} = \lim_{x \rightarrow \infty} \frac{\bar{H}_{\bar{\theta}_t, t}(\lambda x/\bar{\theta}_t)}{\bar{H}_{\bar{\theta}_t, t}(x/\bar{\theta}_t)} = \lim_{s \rightarrow \infty} \frac{\bar{H}_{\bar{\theta}_t, t}(\lambda s)}{\bar{H}_{\bar{\theta}_t, t}(s)},$$

so  $\bar{F}_t \in \mathcal{R}_{-\infty}$  at  $\infty$ , which together with Equation (45) shows that item (i) holds.

We shall now prove item (ii). Using Lemma 8,  $\bar{H}_{\bar{\theta}_t, t}^{-1}(\cdot)$  is 0-varying at 0: for  $\lambda > 0$ ,

$$\lim_{y \downarrow 0} \frac{\bar{H}_{\bar{\theta}_t, t}^{-1}(\lambda y)}{\bar{H}_{\bar{\theta}_t, t}^{-1}(y)} = 1. \quad (46)$$

On the other hand,  $\lim_{x \rightarrow \infty} \bar{F}_t(x) = 0$  and Equation (44) imply that  $\lim_{x \rightarrow \infty} \bar{F}_t(x)U_t(x)/p_t(\bar{\theta}_t) = 0$ , and

$$\lim_{x \rightarrow \infty} \frac{\bar{F}_t(x)U_t(x)/p_t(\bar{\theta}_t)}{\bar{F}_t(x)} = \frac{1}{p_t(\bar{\theta}_t)}.$$

That  $\bar{H}_{\bar{\theta}_t, t}^{-1}(\cdot)$  is 0-varying at 0 then implies, by Lemma 9,

$$\lim_{x \rightarrow \infty} \frac{\bar{H}_{\bar{\theta}_t, t}^{-1}(\bar{F}_t(x)U_t(x)/p_t(\bar{\theta}_t))}{\bar{H}_{\bar{\theta}_t, t}^{-1}(\bar{F}_t(x))} = [p_t(\bar{\theta}_t)]^0 = 1. \quad (47)$$

Now, using Equation (40)

$$\frac{\bar{H}_{\bar{\theta}_t, t}^{-1}(\bar{F}_t(x)U_t(x)/p_t(\bar{\theta}_t))}{\bar{H}_{\bar{\theta}_t, t}^{-1}(\bar{F}_t(x))} = \frac{x/\bar{\theta}_t}{\bar{H}_{\bar{\theta}_t, t}^{-1}(\bar{F}_t(x))}, \quad (48)$$

so Equation (47) implies that  $x/\bar{\theta}_t \sim \bar{H}_{\bar{\theta}_t, t}^{-1}(\bar{F}_t(x))$  as  $x$  goes to  $\infty$ . Letting  $y = \bar{F}_t(x)$  we then have

$$\bar{F}_t^{-1}(y) \sim \bar{\theta}_t \bar{H}_{\bar{\theta}_t, t}^{-1}(y) \text{ as } y \downarrow 0. \quad (49)$$

Equations (46) and (49) give

$$\lim_{y \downarrow 0} \frac{\bar{F}_t^{-1}(\lambda y)}{\bar{F}_t^{-1}(y)} = 1 \text{ for } \lambda > 0,$$

so  $\bar{F}_t^{-1}(\cdot)$  is 0-varying at 0 which together with Equations (46) and (49) shows (ii). ■

*Proof of Theorem 2.* The proof of Theorem 2 follows from Lemma 1 easily by the argument used in Section 3. We include it here for completeness. (i) Consider  $(t_1, t_2) \in \mathcal{T}^2$  such that  $t_1 < t_2$ . From Lemma 1(i),

$$\frac{\bar{F}_{t_1}(x)}{\bar{F}_{t_2}(x)} \sim \frac{p_{t_1}(\bar{\theta}_{t_1})\bar{H}_{\bar{\theta}_{t_1}, t_1}(x/\bar{\theta}_{t_1})}{p_{t_2}(\bar{\theta}_{t_2})\bar{H}_{\bar{\theta}_{t_2}, t_2}(x/\bar{\theta}_{t_2})}, \text{ as } x \rightarrow \infty.$$

Now note that

$$\frac{p_{t_1}(\bar{\theta}_{t_1})\bar{H}_{\bar{\theta}_{t_1}, t_1}(x/\bar{\theta}_{t_1})}{p_{t_2}(\bar{\theta}_{t_2})\bar{H}_{\bar{\theta}_{t_2}, t_2}(x/\bar{\theta}_{t_2})} = \frac{p_{t_1}(\bar{\theta}_{t_1})\bar{H}_{\bar{\theta}_{t_1}, t_1}(x/\bar{\theta}_{t_1})\bar{H}_{\bar{\theta}_{t_1}, t_1}(x/\bar{\theta}_{t_2})\bar{H}_{\bar{\theta}_{t_1}, t_2}(x/\bar{\theta}_{t_2})}{p_{t_2}(\bar{\theta}_{t_2})\bar{H}_{\bar{\theta}_{t_1}, t_1}(x/\bar{\theta}_{t_2})\bar{H}_{\bar{\theta}_{t_1}, t_2}(x/\bar{\theta}_{t_2})\bar{H}_{\bar{\theta}_{t_2}, t_2}(x/\bar{\theta}_{t_2})}. \quad (50)$$

Since  $\bar{\theta}_{t_2} > \bar{\theta}_{t_1}$  and  $\bar{H}_{\bar{\theta}_{t_1}, t_1} \in \mathcal{R}_{-\infty}$  at  $\infty$ , we have  $\lim_{x \rightarrow \infty} \bar{H}_{\bar{\theta}_{t_1}, t_1}(x/\bar{\theta}_{t_1})/\bar{H}_{\bar{\theta}_{t_1}, t_1}(x/\bar{\theta}_{t_2}) = \lim_{y \rightarrow \infty} \bar{H}_{\bar{\theta}_{t_1}, t_1}((\bar{\theta}_{t_2}/\bar{\theta}_{t_1})y)/\bar{H}_{\bar{\theta}_{t_1}, t_1}(y) = 0$  where  $y \equiv x/\bar{\theta}_{t_2}$ . So the second term on the

right-hand side of Equation (50) goes to 0 as  $x$  gets large. From assumption (16),  $\theta < \bar{\theta}_{t_2}$  implies that

$$\frac{\bar{H}_{\theta, t_2}(x/\bar{\theta}_{t_2})}{\bar{H}_{\bar{\theta}_{t_2}, t_2}(x/\bar{\theta}_{t_2})}$$

is bounded. So the fourth term of the right-hand side of Equation (50) is bounded as  $x$  goes to  $\infty$ . From Property (17), and given  $t_1 < t_2$ , we know that

$$\frac{\bar{H}_{\bar{\theta}_{t_1}, t_1}(x/\bar{\theta}_{t_2})}{\bar{H}_{\bar{\theta}_{t_1}, t_2}(x/\bar{\theta}_{t_2})}$$

is bounded as  $x \rightarrow \infty$ . Hence the third term of the right-hand side of Equation (50) is bounded as  $x$  increases. Finally, we know that  $p_{t_1}(\bar{\theta}_{t_1})/p_{t_2}(\bar{\theta}_{t_2}) < \infty$  since  $p_{t_2}(\bar{\theta}_{t_2}) > 0$ , so the first term of the right-hand side of Equation (50) is finite.

Combining the facts above,

$$\lim_{x \rightarrow \infty} \frac{\bar{F}_{t_1}(x)}{\bar{F}_{t_2}(x)} = 0,$$

so there is  $\bar{x}_1 > 0$  such that, if  $x \geq \bar{x}_1$  then  $\bar{F}_{t_1}(x) < \bar{F}_{t_2}(x)$ . Since  $\mathcal{T}$  is finite, there is  $\bar{x}$  such that if  $x \geq \bar{x}$  then  $\bar{F}_t(x) < \bar{F}_{t'}(x)$  for all  $(t, t') \in \mathcal{T}^2$  with  $t < t'$ . (ii) Note that for any  $t \in \mathcal{T}$  and  $y \in (0, 1)$ ,  $\bar{F}_t^{-1}(y) = q_t(1 - y)$ . From (i), for any  $(t_1, t_2) \in \mathcal{T}^2$  such that  $t_1 < t_2$ , there is  $\underline{y}_1 \in (0, 1)$  such that if  $y \leq \underline{y}_1$  then  $q_{t_1}(1 - y) < q_{t_2}(1 - y)$ . Equivalently, letting  $\bar{\alpha}_1 \equiv 1 - \underline{y}_1$ , for  $\alpha \in [\bar{\alpha}_1, 1)$  we have  $q_{t_1}(\alpha) < q_{t_2}(\alpha)$ .  $\mathcal{T}$  being finite guarantees that (ii) holds by the same reasoning as in (i). ■

*Proof of Lemma 3.* Assume  $\bar{H}_{\bar{\theta}_t, t}(\cdot)$  is in  $\Gamma$  with auxiliary function  $a_{\bar{\theta}_t, t}(\cdot)$ ; for  $U_t(\cdot)$  defined in Equation (40) we have

$$\frac{U_t(\bar{\theta}_t x + \bar{\theta}_t \lambda a_{\bar{\theta}_t, t}(x))}{U_t(\bar{\theta}_t x)} = \left[ \frac{p_t(\bar{\theta}_t) \bar{H}_{\bar{\theta}_t, t}(x + \lambda a_{\bar{\theta}_t, t}(x))}{\bar{F}_t(\bar{\theta}_t x + \bar{\theta}_t \lambda a_{\bar{\theta}_t, t}(x))} \right] \left[ \frac{\bar{F}_t(\bar{\theta}_t x)}{p_t(\bar{\theta}_t) \bar{H}_{\bar{\theta}_t, t}(x)} \right] \quad (51)$$

From Equation (44), the left-hand side in Equation (51) converges to 1 as  $x \rightarrow \infty$ . On the other hand,

$$\lim_{x \rightarrow \infty} \frac{p_t(\bar{\theta}_t) \bar{H}_{\bar{\theta}_t, t}(x + \lambda a_{\bar{\theta}_t, t}(x))}{p_t(\bar{\theta}_t) \bar{H}_{\bar{\theta}_t, t}(x)} = \exp(-\lambda),$$

as  $a_{\bar{\theta}_t,t}(\cdot)$  is the auxiliary function of  $\bar{H}_{\bar{\theta}_t,t}(\cdot)$ . Then we have

$$\exp(\lambda) = \lim_{x \rightarrow \infty} \frac{\bar{F}_t(\bar{\theta}_t x)}{\bar{F}_t(\bar{\theta}_t x + \bar{\theta}_t \lambda a_{\bar{\theta}_t,t}(x))} = \lim_{s \rightarrow \infty} \frac{\bar{F}_t(s)}{\bar{F}_t(s + \bar{\theta}_t \lambda a_{\bar{\theta}_t,t}(s/\bar{\theta}_t))},$$

using  $s = \bar{\theta}_t x$ . Hence  $\bar{F}_t \in \Gamma$ :

$$\lim_{s \rightarrow \infty} \frac{\bar{F}_t(s + \lambda a_t(s))}{\bar{F}_t(s)} = \exp(-\lambda), \quad (52)$$

with auxiliary function  $a_t(\cdot)$  defined as  $a_t(x) \equiv \bar{\theta}_t a_{\bar{\theta}_t,t}(x/\bar{\theta}_t)$  for all  $x > 0$ , 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  $\varphi_s : \mathbb{R}^+ \rightarrow (0, 1)$ ,  $\lim_{s \rightarrow \infty} \varphi_s(x) = \varphi(x)$  for all continuity points  $x > 0$  of  $\varphi(\cdot)$ , implies  $\lim_{s \rightarrow \infty} \varphi_s^{-1}(z) = \varphi^{-1}(z)$  for all continuity points  $z \in (0, 1)$  of  $\varphi^{-1}(\cdot)$  (see, e.g., Lemma 1.9 in de Haan, 1974). Let then

$$\varphi_s(x) \equiv 1 - \frac{\bar{H}_{\bar{\theta}_t,t}(s + x a_{\bar{\theta}_t,t}(s))}{\bar{H}_{\bar{\theta}_t,t}(s)} \text{ for all } x > 0.$$

That  $\bar{H}_{\bar{\theta}_t,t} \in \Gamma$  implies  $\lim_{s \rightarrow \infty} \varphi_s(x) = 1 - \exp(-x)$  for all  $x > 0$ . Letting  $\varphi(x) \equiv 1 - \exp(-x)$ , we then have for  $z \in (0, 1)$

$$\lim_{s \rightarrow \infty} \frac{\bar{H}_{\bar{\theta}_t,t}^{-1}(\bar{H}_{\bar{\theta}_t,t}(s)(1-z)) - s}{a_{\bar{\theta}_t,t}(s)} = \lim_{s \rightarrow \infty} \varphi_s^{-1}(z) = \varphi^{-1}(z) = -\ln(1-z).$$

Letting  $y \equiv \bar{H}_{\bar{\theta}_t,t}(s)$  and  $\mu \equiv 1 - z$  gives

$$\lim_{y \downarrow 0} \frac{\bar{H}_{\bar{\theta}_t,t}^{-1}(y\mu) - \bar{H}_{\bar{\theta}_t,t}^{-1}(y)}{a_{\bar{\theta}_t,t}(\bar{H}_{\bar{\theta}_t,t}^{-1}(y))} = -\ln \mu \text{ for } \mu \in (0, 1). \quad (53)$$

Thus  $\bar{H}_{\bar{\theta}_t,t}^{-1} \in \Pi$  as in Definition 3 with auxiliary function  $a_{\bar{\theta}_t,t} \circ \bar{H}_{\bar{\theta}_t,t}^{-1}(\cdot)$ .

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

$$\frac{a_{\bar{\theta}_t,t}(\bar{H}_{\bar{\theta}_t,t}^{-1}(\lambda y))}{a_{\bar{\theta}_t,t}(\bar{H}_{\bar{\theta}_t,t}^{-1}(y))} = - \left[ \frac{\bar{H}_{\bar{\theta}_t,t}^{-1}(\mu y) - \bar{H}_{\bar{\theta}_t,t}^{-1}(y)}{a_{\bar{\theta}_t,t}(\bar{H}_{\bar{\theta}_t,t}^{-1}(y))} \right] \left[ \frac{a_{\bar{\theta}_t,t}(\bar{H}_{\bar{\theta}_t,t}^{-1}(\lambda y))}{\bar{H}_{\bar{\theta}_t,t}^{-1}(\lambda \nu y) - \bar{H}_{\bar{\theta}_t,t}^{-1}(\lambda y)} \right]. \quad (54)$$

Equations (53) and (54) together imply

$$\lim_{y \downarrow 0} \frac{a_{\bar{\theta}_t,t} \circ \bar{H}_{\bar{\theta}_t,t}^{-1}(\lambda y)}{a_{\bar{\theta}_t,t} \circ \bar{H}_{\bar{\theta}_t,t}^{-1}(y)} = \frac{\ln \mu}{-\ln \nu} = 1, \quad (55)$$

so  $a_{\bar{\theta}_t, t} \circ \bar{H}_{\bar{\theta}_t, t}^{-1} \in \mathcal{R}_0$  at 0.

We now study  $\bar{F}_t(\cdot)$ : if for any  $t \in \mathcal{T}$ , we let  $\varphi_{t,s}(x) \equiv 1 - \bar{F}_t(s + xa_t(s))/\bar{F}_t(s)$  for all  $x > 0$ , we have  $\lim_{x \rightarrow \infty} \varphi_{t,s}(x) = \varphi(x)$ . Same reasoning as previously then implies

$$\lim_{y \downarrow 0} \frac{\bar{F}_t^{-1}(y\mu) - \bar{F}_t^{-1}(y)}{a_t(\bar{F}_t^{-1}(y))} = -\ln \mu \text{ for } \mu \in (0, 1). \quad (56)$$

So  $\bar{F}_t^{-1} \in \Pi$  as in Definition 3 with auxiliary function  $a_t \circ \bar{F}_t^{-1}(\cdot)$ . Equation (56) and the equality

$$\frac{a_t(\bar{F}_t^{-1}(\lambda y))}{a_t(\bar{F}_t^{-1}(y))} = - \left[ \frac{\bar{F}_t^{-1}(\mu y) - \bar{F}_t^{-1}(y)}{a_t(\bar{F}_t^{-1}(y))} \right] \left[ \frac{a_t(\bar{F}_t^{-1}(\lambda y))}{\bar{F}_t^{-1}(\lambda \nu y) - \bar{F}_t^{-1}(\lambda y)} \right],$$

with  $\lambda > 0$ ,  $\mu \equiv \lambda$  and  $\nu \equiv \lambda^{-1}$ , then imply that  $a_t \circ \bar{F}_t^{-1} \in \mathcal{R}_0$  at 0.

Given Equation (49) and the definition of  $a_t(\cdot)$ , it is not surprising to see that  $a_t(\bar{F}_t^{-1}(y)) \sim \bar{\theta}_t a_{\bar{\theta}_t, t}(\bar{H}_{\bar{\theta}_t, t}^{-1}(y))$  as  $y \downarrow 0$ ; however we need a formal proof of that statement.

We start by showing that

$$\lim_{x \rightarrow \infty} F_t^x(a_x \lambda + b_x) = \exp[-\exp(-\lambda)], \quad (57)$$

with  $a_x \equiv a_t(b_x)$  and  $b_x \equiv \bar{F}_t^{-1}(1/x)$ . In Equation (52), let  $x \equiv 1/\bar{F}_t(s)$  (or  $s = \bar{F}_t^{-1}(1/x)$ ) so  $x \rightarrow \infty$  as  $s \rightarrow \infty$ ; then

$$\begin{aligned} & \lim_{x \rightarrow \infty} \frac{\bar{F}_t(\bar{F}_t^{-1}(1/x) + \lambda a_t(\bar{F}_t^{-1}(1/x)))}{\bar{F}_t(\bar{F}_t^{-1}(1/x))} \\ &= \lim_{x \rightarrow \infty} x \cdot \bar{F}_t[a_t(\bar{F}_t^{-1}(1/x))\lambda + \bar{F}_t^{-1}(1/x)] \\ &= \exp(-\lambda). \end{aligned} \quad (58)$$

Let  $b_x \equiv \bar{F}_t^{-1}(1/x)$  and  $a_x \equiv a_t(b_x)$ ; the last equality in Equation(58) together with Lemma 2.2.2 in de Haan (1970) then imply Equation (57).

We now derive a similar equality involving  $H(\cdot)$ : the last equality in Equation (58) and the tail equivalence property in Equation (45) imply  $\bar{F}_t(x) \sim p_t(\bar{\theta}_t) \bar{H}_{\bar{\theta}_t, t}(x/\bar{\theta}_t)$ , as  $x \rightarrow \infty$

$$\lim_{x \rightarrow \infty} x \cdot \bar{H}_{\bar{\theta}_t, t}[(a_x/\bar{\theta}_t)\lambda + (b_x/\bar{\theta}_t)] = \exp(-\lambda - \ln p_t(\bar{\theta}_t)).$$

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

$$\lim_{x \rightarrow \infty} H_{\bar{\theta}_t, t}^x((a_x/\bar{\theta}_t)\lambda + (b_x/\bar{\theta}_t)) = \exp[-\exp(-\lambda - \ln p_t(\bar{\theta}_t))], \quad (59)$$

with  $a_x = a_t(b_x)$  and  $b_x = \bar{F}_t^{-1}(1/x)$  as previously. On the other hand,  $\bar{H}_{\bar{\theta}_t, t} \in \Gamma$  as in Definition 2 with auxiliary function  $a_{\bar{\theta}_t, t}(\cdot)$ , together with Lemma 2.2.2 in de Haan (1970) imply

$$\lim_{x \rightarrow \infty} H_{\bar{\theta}_t, t}^x(\tilde{a}_x\lambda + \tilde{b}_x) = \exp[-\exp(-\lambda)], \quad (60)$$

with  $\tilde{a}_x = a_{\bar{\theta}_t, t}(\tilde{b}_x)$  and  $\tilde{b}_x = \bar{H}_{\bar{\theta}_t, t}^{-1}(1/x)$ . Combining Equations (59) and (60) and applying the results of Lemma 2.4.1 in de Haan (1970) on the change of norming constants (with  $A = 1$  and  $B = \ln p_t(\bar{\theta}_t)$ ), then gives

$$\frac{(a_x/\bar{\theta}_t)}{\tilde{a}_x} \rightarrow 1 \text{ and } \frac{(b_x/\bar{\theta}_t) - \tilde{b}_x}{\tilde{a}_x} \rightarrow \ln p_t(\bar{\theta}_t) \text{ as } x \rightarrow \infty.$$

So

$$a_t(\bar{F}_t^{-1}(y)) \sim \bar{\theta}_t a_{\bar{\theta}_t, t}(\bar{H}_{\bar{\theta}_t, t}^{-1}(y)) \text{ as } y \downarrow 0,$$

which completes the proof of item (ii). ■

*Proof of Lemma 4.* Given a random sample  $(X_1, \dots, X_N)$  let  $\{X_{(k)}^{(N)}\}_{k=1}^N$  be the ascending order statistics:  $X_{(1)}^{(N)} \leq \dots \leq X_{(N)}^{(N)}$ . Then for any  $\alpha_N$ ,  $0 < \alpha_N < 1$ , we have

$$\hat{q}_N(\alpha_N) = X_{(m)}^{(N)} \text{ with } m \equiv \lfloor \alpha_N N \rfloor + 1, \quad (61)$$

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

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

$$\lim_{N \rightarrow \infty} N - m = \infty, \quad (62)$$

$$\lim_{N \rightarrow \infty} \frac{N - m}{N} = 0, \quad (63)$$

$$\lim_{N \rightarrow \infty} \frac{N - m}{(1 - \alpha_N)N} = \lim_{N \rightarrow \infty} \frac{N - m + 1}{(1 - \alpha_N)N} = 1. \quad (64)$$

The hypotheses on  $\beta_N$  imply properties (62), (63), and (64) for  $k$ , and, in addition, that

$$\lim_{N \rightarrow \infty} \frac{N - m}{N - k} = \rho. \quad (65)$$

Second, we have

$$\begin{aligned} & \sqrt{N(1 - \alpha_N)} \left[ \frac{\hat{q}_N(\alpha_N) - q(\alpha_N)}{a(q(\alpha_N))} \right] \\ &= \sqrt{N(1 - \alpha_N)} \left[ \frac{X_{(\lfloor \alpha_N N \rfloor + 1)}^{(N)} - \bar{F}^{-1}(1 - \alpha_N)}{a(\bar{F}^{-1}(1 - \alpha_N))} \right] \\ &= \sqrt{N - m + 1} \left[ \frac{X_{(m)}^{(N)} - \bar{F}^{-1}((N - m)/N)}{a(\bar{F}^{-1}((N - m)/N))} + \frac{\bar{F}^{-1}((N - m)/N) - \bar{F}^{-1}(1 - \alpha_N)}{a(\bar{F}^{-1}((N - m)/N))} \right] \times \\ & \hspace{15em} (66) \end{aligned}$$

$$\frac{a(\bar{F}^{-1}((N - m)/N))}{a(\bar{F}^{-1}(1 - \alpha_N))} \sqrt{\frac{N(1 - \alpha_N)}{N - m + 1}},$$

and

$$\begin{aligned} & \frac{\hat{q}_N(\beta_N) - \hat{q}_N(\alpha_N)}{a(q(\alpha_N))} - \ln \rho \\ &= \left\{ \left[ \frac{X_{(k)}^{(N)} - X_{(m)}^{(N)}}{a(X_{(m)}^{(N)})} - \ln \rho \right] \frac{a(X_{(m)}^{(N)})}{a(\bar{F}^{-1}((N - m)/N))} + \right. \\ & \left. \ln \rho \left[ \frac{a(X_{(m)}^{(N)})}{a(\bar{F}^{-1}((N - m)/N))} - 1 \right] \right\} \frac{a(\bar{F}^{-1}((N - m)/N))}{a(\bar{F}^{-1}(1 - \alpha_N))}. \end{aligned} \quad (67)$$

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

$$\sqrt{N - m + 1} \left[ \frac{X_{(m)}^{(N)} - \bar{F}^{-1}((N - m)/N)}{a(\bar{F}^{-1}((N - m)/N))} \right] \xrightarrow{d} \mathcal{N}, \quad (68)$$

$$\frac{X_{(k)}^{(N)} - X_{(m)}^{(N)}}{a(X_{(m)}^{(N)})} \xrightarrow{P} \ln \rho, \quad (69)$$

where  $\mathcal{N}$  is a standard Gaussian random variable.

We then show (STEP 2)

$$\lim_{N \rightarrow \infty} \sqrt{N - m + 1} \left[ \frac{\bar{F}^{-1}((N - m)/N) - \bar{F}^{-1}(1 - \alpha_N)}{a(\bar{F}^{-1}((N - m)/N))} \right] = 0 \quad (70)$$

$$\lim_{N \rightarrow \infty} \frac{a(\bar{F}^{-1}((N - m)/N))}{a(\bar{F}^{-1}(1 - \alpha_N))} = 1. \quad (71)$$

Finally, we show (STEP 3)

$$\frac{a(X_{(m)}^{(N)})}{a(\bar{F}^{-1}((N-m)/N))} \xrightarrow{p} 1. \quad (72)$$

The first limit result of Lemma 4 then follows from (66) by (68), (70) and (71) using (64) and Lemma 2.4.1 in de Haan (1970).

The second limit result in Lemma 4 follows from (67) by (69) and (72) using (71), (64), 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  $A_1, \dots, A_N$  be independent and identically distributed standard exponential random variables. Let  $A_{(1)}^{(N)} \leq \dots \leq A_{(N)}^{(N)}$  be the ascending order statistics of  $(A_1, \dots, A_N)$ . Then, by using the probability integral transform  $\left\{ X_{(m)}^{(N)} \right\}_{m=1}^N \stackrel{d}{=} \left\{ \bar{F}^{-1}(\exp(-A_{(m)}^{(N)})) \right\}_{m=1}^N$ .

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

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

Let  $\eta_N \equiv \ln(N/(N-m))$ ; then,  $a(\bar{F}^{-1}((N-m)/N)) = W'(\eta_N)$ . So the expression in Statement (68) can be written as

$$\begin{aligned} & \sqrt{N-m+1} \left[ \frac{X_{(m)}^{(N)} - \bar{F}^{-1}((N-m)/N)}{a(\bar{F}^{-1}((N-m)/N))} \right] \\ & \stackrel{d}{=} \sqrt{N-m+1} \left[ \frac{W(A_{(m)}^{(N)}) - W(\eta_N)}{W'(\eta_N)} \right] \\ & \stackrel{d}{=} \sqrt{N-m+1} \int_0^{Z_N/\sqrt{N-m+1}} \frac{W'(\eta_N + s)}{W'(\eta_N)} ds, \end{aligned} \quad (73)$$

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

Then, by Lemma 13,  $Z_N \xrightarrow{d} \mathcal{N}_1$  as  $N \rightarrow \infty$ . But Lemma 10 (i) and Statement (59) imply that the integrand on the right-hand side of (73) converges uniformly to 1 on compact intervals, as  $N \rightarrow \infty$ . So Lemma 11 implies Statement (68).

The proof of Statement (69) is similar. We have

$$\begin{aligned}
& \frac{X_{(k)}^{(N)} - X_{(m)}^{(N)}}{a(X_{(m)}^{(N)})} - \ln \rho \\
& \stackrel{d}{=} \frac{W(A_{(k)}^{(N)}) - W(A_{(m)}^{(N)})}{W'(A_{(m)}^{(N)})} - \ln \rho \\
& \stackrel{d}{=} \left[ \sqrt{\frac{N-m}{\rho-1}} \int_0^{Y_N/\sqrt{\frac{N-m}{\rho-1}}} \frac{W'(A_{(m)}^{(N)} + \ln \rho + s)}{W'(A_{(m)}^{(N)})} ds \right] \sqrt{\frac{\rho-1}{N-m}} \\
& + \int_0^{\ln \rho} \left[ \frac{W'(A_{(m)}^{(N)} + s)}{W'(A_{(m)}^{(N)})} - 1 \right] ds, \tag{74}
\end{aligned}$$

where  $Y_N \equiv \sqrt{\frac{N-m}{\rho-1}} [A_{(k)}^{(N)} - A_{(m)}^{(N)} - \ln \rho]$ . Note that  $\{Y_N\}$  and  $\{A_{(m)}^{(N)}\}$  are independent (Renyi, 1953) and that  $A_{(m)}^{(N)} \xrightarrow{as} \infty$  (see, e.g. Theorem 4 in Watts (1980)). By Lemma 13,  $Y_N \xrightarrow{d} \mathcal{N}_2$  as  $N \rightarrow \infty$ , and the integrand in the first term of Equation (74) converges uniformly to 1 on compact intervals. Hence, using Lemma 11 and Statement (57) the first term in brackets in Equation (74) converges in distribution. It is multiplied by  $(\frac{\rho-1}{N-m})^{-1/2}$ , which goes to zero. So the first summand of expression (74) 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 (74) converges to 0 a.s.: Note that  $A_{(m)}^{(N)} \xrightarrow{as} \infty$  a.s. (by, e.g. Theorem 4 in Watts (1980)) and the integrand converges to 0 uniformly on compact intervals (Lemma 10 (i)), so the integral converges to 0 for a full measure of realizations of  $\{A_{(m)}^{(N)}\}$ . This establishes Statement (69).

STEP 2: We now prove (70) and (71). Using the notation in Step 1:

$$\begin{aligned}
& \sqrt{N-m+1} \left[ \frac{\bar{F}^{-1}((N-m)/N) - \bar{F}^{-1}(1-\alpha_N)}{a(\bar{F}^{-1}((N-m)/N))} \right] \\
& = \sqrt{N-m+1} \left[ \frac{W(\eta_N) - W(\ln(1/(1-\alpha_N)))}{W'(\eta_N)} \right] \\
& = \sqrt{N-m+1} \int_{-\ln(1-\alpha_N)-\eta_N}^0 \frac{W'(\eta_N + s)}{W'(\eta_N)} ds \\
& \sim \sqrt{N-m+1} \left[ 0 - \ln \frac{N-m}{(1-\alpha_N)N} \right] \text{ as } N \rightarrow \infty. \tag{75}
\end{aligned}$$

The equivalence in Statement (75) follows by exchanging the limit and the integral, using the uniform convergence established in Lemma 10 (i), and the fact that Statement (64)

implies

$$\lim_{N \rightarrow \infty} [-\ln(1 - \alpha_N) - \eta_N] = \lim_{N \rightarrow \infty} \ln \frac{N - m}{(1 - \alpha_N)N} = 0.$$

The  $\ln(\cdot)$  function is concave, so  $|\ln \{(N - m)/[(1 - \alpha_N)N]\}| \leq |(N - m)/[(1 - \alpha_N)N] - 1|$ .

Hence,

$$\begin{aligned} \sqrt{N - m + 1} \left| \ln \frac{N - m}{(1 - \alpha_N)N} \right| &\leq \sqrt{N - m + 1} \left| \frac{N - m}{(1 - \alpha_N)N} - 1 \right| \\ &= \sqrt{N - m + 1} \left| \frac{\alpha_N N - \lfloor \alpha_N N \rfloor - 1}{(1 - \alpha_N)N} \right| \\ &\leq 2 \frac{\sqrt{N - m + 1}}{(1 - \alpha_N)N} \rightarrow 0, \text{ as } N \rightarrow \infty, \end{aligned}$$

where the convergence to 0 follows from Statement (64). By Statement (75), then, this proves (70).

To prove (71), note that Lemma 3 (ii) implies that  $a \circ \bar{F}^{-1} \in \mathcal{R}_0$  at 0. So Statements (63) and (64), and Lemma 9, give (71).

STEP 3: The proof of (72), in turn is similar to that of (68) in Step 1. We have

$$\begin{aligned} &\frac{a(X_{(m)}^{(N)})}{a(\bar{F}^{-1}((N - m)/N))} - 1 \\ &\stackrel{d}{=} \frac{W'(A_{(m)}^{(N)}) - W'(\eta_N)}{W'(\eta_N)} \\ &\stackrel{d}{=} \left[ \sqrt{N - m} \int_0^{Z_N/\sqrt{N - m + 1}} \frac{W''(\eta_N + s)}{W'(\eta_N)} ds \right] \frac{1}{\sqrt{N - m}}, \end{aligned} \quad (76)$$

where,  $Z_N = \sqrt{N - m + 1}[A_{(m)}^{(N)} - \ln(N/(N - m))]$ . Then, by Lemma 13,  $Z_N \xrightarrow{d} \mathcal{N}_1$  as  $N \rightarrow \infty$ . But Lemma 10 (ii) implies that the integrand on the right-hand side of (76) converges uniformly to 0 on compact intervals. So Lemma 12 and Statement (57) imply Statement (72). ■

*Proof of Theorem 5.* If  $((\mathcal{E}_t, p_t, (H_{\theta,t})_{\theta \in \mathcal{E}_t}), t \in \mathcal{T})$  satisfies  $S_4$  then it satisfies  $S_3$ ; hence we can use Lemma 3(i) to show that for any  $t_j \in \mathcal{T}$ ,  $1 \leq j \leq k$ , the conditional

distribution tails  $\bar{F}_{t_j}(\cdot) \in \Gamma$ . Moreover, from Equation (2) we know that for any  $j$ ,  $1 \leq j \leq k$ ,

$$f_{t_j}(x) = \sum_{\theta \in \mathcal{E}_{t_j}} \frac{p_{t_j}(\theta)}{\theta} h_{\theta, t_j}(x/\theta) \text{ for any } x > 0.$$

Under  $S_4$  the densities  $h_{\theta, t_j}(\cdot)$  are all eventually non-decreasing; hence the same holds for  $f_{t_j}(\cdot)$ . If for each  $j$ ,  $1 \leq j \leq k$ , we have  $0 < \alpha_{N_{t_j}} < \beta_{N_{t_j}} < 1$ ,  $\lim_{N_{t_j} \rightarrow \infty} \alpha_{N_{t_j}} = 1$ ,  $\lim_{N_{t_j} \rightarrow \infty} (1 - \alpha_{N_{t_j}}) N_{t_j} = \infty$  and  $\lim_{N_{t_j} \rightarrow \infty} (1 - \alpha_{N_{t_j}})/(1 - \beta_{N_{t_j}}) = \rho_{t_j}$  with  $\rho_{t_j} > 1$ , then the results of Lemma 4 apply for all  $j$ ,  $1 \leq j \leq k$ , i.e.

$$\sqrt{N_{t_j}(1 - \alpha_{N_{t_j}})} \left[ a(q_{t_j}(\alpha_{N_{t_j}})) \right]^{-1} \left[ \hat{q}_{t_j, N_{t_j}}(\alpha_{N_{t_j}}) - q_{t_j}(\alpha_{N_{t_j}}) \right] \xrightarrow{d} \mathcal{N}_j,$$

and

$$\frac{\hat{q}_{t_j, N_{t_j}}(\beta_{N_{t_j}}) - \hat{q}_{t_j, N_{t_j}}(\alpha_{N_{t_j}})}{a(q_{t_j}(\alpha_{N_{t_j}}))} \xrightarrow{p} \ln \rho_{t_j},$$

where  $\mathcal{N}_j$ ,  $1 \leq j \leq k$ , are  $k$  independent standard normal random variables. The conclusion Theorem 5 follows by using Slutsky's Theorem and the independence of different samples  $(X_{t_j, 1}, \dots, X_{t_j, N_{t_j}})$ . ■

*Proof of Theorem 7.* The proof is done in five steps:

STEP1: we work with the first minimization problem in Equation (23)

$$\begin{aligned} \min_{\mu} (\mu - \hat{q})' \hat{\Sigma}^{-1} (\mu - \hat{q}), \\ \text{subject to } A\mu = 0. \end{aligned} \tag{77}$$

Let  $\mathcal{L} : \mathbb{R}^{2k-1} \rightarrow \mathbb{R}$  be the corresponding Lagrangian

$$\mathcal{L}(\mu, \lambda) = (\mu - \hat{q})' \hat{\Sigma}^{-1} (\mu - \hat{q}) + \lambda' A\mu,$$

where  $\lambda$  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  $g : \mathbb{R}^{k-1} \rightarrow \mathbb{R} \cup \{-\infty\}$  is

$$\begin{aligned} g(\lambda) &= \inf_{\mu} \mathcal{L}(\mu, \lambda) \\ &= -\frac{1}{4} \lambda' A \hat{\Sigma} A' \lambda + \lambda' A \hat{q}. \end{aligned}$$

The dual problem is then

$$\max_{\lambda} -\frac{1}{4}\lambda' A \hat{\Sigma} A' \lambda + \lambda' A \hat{q}, \quad (78)$$

with  $\lambda$  unconstrained. The solutions to the dual and primal problems (78) and (77) are

$$\lambda_0 = 2(A \hat{\Sigma} A')^{-1} A \hat{q}, \quad (79)$$

$$\begin{aligned} \mu_0 &= \hat{q} - \frac{1}{2} \hat{\Sigma} A' \lambda_0 \\ &= \hat{q} - \hat{\Sigma} A' (A \hat{\Sigma} A')^{-1} A \hat{q}, \end{aligned} \quad (80)$$

and we have

$$\begin{aligned} (\mu_0 - \hat{q})' \hat{\Sigma}^{-1} (\mu_0 - \hat{q}) &= \hat{q}' A' (A \hat{\Sigma} A')^{-1} A \hat{q} \\ &= -\frac{1}{4} \lambda_0' A \hat{\Sigma} A' \lambda_0 + \lambda_0' A \hat{q}. \end{aligned} \quad (81)$$

Similarly, we consider the dual of the second minimization problem in (23)

$$\begin{aligned} \min_{\mu} (\mu - \hat{q})' \hat{\Sigma}^{-1} (\mu - \hat{q}), \quad (82) \\ \text{subject to } A\mu \preceq 0. \end{aligned}$$

The dual is

$$\begin{aligned} \max_{\lambda} -\frac{1}{4}\lambda' A \hat{\Sigma} A' \lambda + \lambda' A \hat{q}, \quad (83) \\ \text{subject to } \lambda \succeq 0. \end{aligned}$$

Letting  $\lambda_1$  and  $\mu_1$  denote the solutions to the dual and primal problems (83) and (82) we again have

$$(\mu_1 - \hat{q})' \hat{\Sigma}^{-1} (\mu_1 - \hat{q}) = -\frac{1}{4} \lambda_1' A \hat{\Sigma} A' \lambda_1 + \lambda_1' A \hat{q}. \quad (84)$$

STEP 2: using Equations (81) and (84) the likelihood-ratio statistic (23) 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 (79) – (80) 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

$$\hat{R} \equiv (A \hat{\Sigma} A')^{-1} A \hat{\Sigma}^{1/2} \text{ and } \nu \equiv \frac{1}{2} \hat{\Sigma}^{1/2} A' \lambda$$

(so that  $\lambda = 2\hat{R}\nu$ ) we then have

$$\hat{\xi}_{LR} = \min_{\nu: \hat{R}\nu \succcurlyeq 0} \left\| \hat{\Sigma}^{-1/2} (\hat{q} - \mu_0) - \nu \right\|^2. \tag{85}$$

STEP3: we consider the dual of the minimization problem in (85):

$$\max_{\beta: \beta \succcurlyeq 0} \left[ -\frac{1}{4} \beta' \hat{R} \hat{R}' \beta - \beta' \hat{R} \hat{\Sigma}^{-1/2} (\hat{q} - \mu_0) \right], \tag{86}$$

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

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





Consider the orthogonal projection of  $Y$  on the cone  $C_{(1,\dots,p)}$ , denoted  $\hat{Y}$ . Then the distribution of  $\|\hat{Y}\|^2$  is a mixture of chi-square distributions

$$\|\hat{Y}\|^2 \sim \sum_{d=0}^p \omega(d) \chi^2(d) \text{ with } \omega(d) = \sum_{S: \dim S=d} P\{\hat{Y} \in C_S\},$$

where the sequence of weights  $\omega(d)$ ,  $d = 0, \dots, p$  satisfies  $\omega(d) \geq 0$  and  $\sum_{d=0}^p \omega(d) = 1$  and  $\chi^2(0)$  denotes the point mass distribution at zero.

Let  $\omega_i, i = 1, \dots, k-1$  denote the columns of the  $(k-1) \times (k-1)$ -matrix  $\Omega^{1/2}$ ; to each  $(k-1)$ -vector  $\omega_i, i = 1, \dots, k-1$ , we associate a vector  $\varphi_i \in \mathbb{R}^{k-1}$  which is orthogonal to any  $\omega_j, j \neq i$ , and is such that  $\omega_i' \varphi_i < 0$  (so  $\omega_i$  and  $\varphi_i$  are not in the same half-space generated by the  $\omega_j$ 's with  $j \neq i$ ). Then, using the notation from Lemma 14, we have that  $C = \left\{ \gamma \in \mathbb{R}^{k-1} : \gamma = \sum_{i=1}^{k-1} \alpha_i \varphi_i, \alpha_i \leq 0, i = 1, \dots, k-1 \right\} = C_{\{1, \dots, k-1\}}$ . Combining Lemma 14 with Equation (92) then yields the result of Theorem 7. Note that the entries of  $\Omega$  can be consistently estimated using  $\hat{\sigma}_{t_1}^{-2}(A\hat{\Sigma}A')$ ; 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. ■