Bayesian Nash equilibrium; a statistical test of the hypothesis

Michael Landsberger
Department of Economics
Haifa University, Haifa, Israel
e-mail:lmichael@econ.haifa.ac.il

Boris Tsirelson
School of Mathematical Sciences
Raymond and Beverly Sackler Faculty of Exact Sciences
Tel Aviv University, Tel Aviv, Israel
e-mail: tsirel@math.tau.ac.il

October, 2003

Abstract

Many economic environments are modeled as games under incomplete information. A fundamental concept in these models is the *hypothesis* of a Bayesian Nash equilibrium. However, in spite of the common use of this hypothesis in economic theory and econometric estimation models, no statistical methods have been suggested how to use real life data to test its validity. There are many doubts whether economic agents are sufficiently sophisticated and informed to find their way to such an equilibrium (see, for example, Rubinstein (1998) and Gilboa&Schmeidler (2001)). We show in this paper how a test of the hypothesis can be conducted in the case of a first price auction game with risk averse bidders. We derive a *theoretical* characterization of the distribution of equilibrium bids, and show that, with some modifications, the Kolmogorov conservative test (see Bedford and Meilijson (1997)) can be applied as a statistical tool. However, since our problem is cast as a composite hypothesis in an infinite dimensional space, it is not clear how to verify whether the theoretical characterization is satisfied. In our case, the question is whether a distribution function that satisfies a second order differential inequality can be 'squeezed' into a corridor between two given monotone stepwise functions. The problem motivates the concept of *relative concavity* which we introduce as a *testable* characterization. The latter makes it possible to develope an algorithm that implements the test in a finite number of steps.

[0]We are grateful to E. Dekel, E. Guerre, P. Haile and I. Meilijson for helpful comments on a previous draft of this paper.

1 Introduction

Games under incomplete information constitute a theoretical framework for many economic models, and the notion of a Bayesian Nash equilibrium is a fundamental concept in such games. Nevertheless, many economists have serious doubts whether the equilibrium hypothesis has any grip in reality, in particular in complex environments. Are the players sufficiently sophisticated, informed and able to find their way to such an equilibrium; see, for example, Rubinstein (1998). Obviously, the only way to dispel (or confirm) these misgivings is by confronting the theory with real life data.[1] This is not an easy task because of many reasons, among them lack of data (in many cases) and the fact that there is no single test that fits all games. In principle, confirming (or rejecting) the theory in the case of a Cournot game will still require to design a test for a first price auction, and so on. So, this may be a long way to go but, in our opinion, a necessary one to follow.[2]

We show how a statistical test of the equilibrium hypothesis can be performed in the case of a first price single unit auction game under the IPV paradigm. To this effect, we address the following problems:

- How to obtain from the equilibrium conditions stipulations formulated in terms of observables only (bids). Namely, how to obtain a *theoretical* characterization of the equilibrium distribution of bids.
- How to design an appropriate statistical test and how to implement it; the latter requires the design of a *testable* characterization of the distribution of equilibrium bids and an algorithm that can be used to conduct the test.

The first question was addressed for completeness, but it is the second that we consider as the main achievement of the paper.

Addressing the first question, we derive conditions that are necessary and sufficient for the distribution of equilibrium bids to correspond to a Bayesian Nash equilibrium in a first price auction game under the IPV paradigm, for *some* distribution of valuations and a prescribed concave utility. This *theoretical* characterization is a second order differential inequality. Having done that, we have a composite hypothesis in an infinite dimensional space. We show that a test of the hypothesis can be designed by applying Kolmogorov's conservative test; see Bedford and Meilijson (1997). Roughly speaking, we can define a 'corridor' bounded by two functions that correspond to the empirical distribution of bids, which is the infinite dimensional equivalent of an ordinary confidence interval. To perform the test we must check whether it is possible to 'squeeze' within the corridor a distribution function F that satisfies the second order differential inequality, $F''(b) < \Psi(b, F(b), F'(b))$. At this point, however, the

[1] This way of verifying the validity of theory is not the one recommended by Rubinstein, who places much significance on the a-priori reasonability of assumptions.

[2] Taking the equilibrium hypothesis for granted is common not only in economic theory, but also in econometric analyses. For example, the growing econometric literature about auctions *assumes* equilibrium behavior and uses equilibrium conditions to identify (in the econometric sense) and estimate the underlying distributions of types; see, for example, Guerre, Perrigne and Vuong (2000) and Athey and Haile (2002).

problem is that there seems to be no obvious way how to perform the test, which raises the problem of *implementation.*[3] Obviously, one should design an algorithm to do it, but the differential form does not lend itself to an algorithmic approach. Hence, it follows that while thinking about a characterization that is to be subjected to an empirical test, one must do it in conjunction with an algorithm, or at least a hint how an algorithm can be developed; this has not been a common procedure in economics, so far. Proceeding in this vein leads to the notion of a *testable* characterization that, in contrast with a theoretical characterization, accounts for implementation.

Another obstacle that makes the theoretical characterization problematic for a statistical test is the fact that the set of functions that satisfy the characterization is not closed; consequently, a sequence of functions that satisfy the characterization may converge to a function that violates it. Under such conditions, one can not design a test that is *consistent.*[4] To illustrate these problems and see how to think about possible solutions, we consider, first, the much simpler case where bidders are risk neutral.

Under risk neutrality, the second order differential inequality is reduced to $(1/F^{N-1}(b))'' > 0$, or $-(1/F^{N-1}(b))'' < 0$. This is a non closed set of functions which, as noted above, poses a problem at the level of a statistical test. Fortunately, a closure is around the corner; the set of convex (concave) functions. Setting new coordinates, we can pose the problem as one of 'squeezing' a concave distribution function within a corridor defined by two increasing step-wise functions. More specifically, the question is whether there exists a concave distribution function, H, that satisfies

$$\hat{F}_n(b) - \varepsilon \le H(b) \le \hat{F}_n(b) + \varepsilon$$

where $\varepsilon = C_{\alpha,n}/\sqrt{n}, C_{\alpha,n}$ are (tabulated) critical values of the Kolmogorov statistic $\sqrt{n}\|\hat{F}_n - H\|$, and $\hat{F}_n$ is the empirical distribution of bids. Intuition based on geometric considerations suggests that instead of checking if this is possible by screening (how?) the class of all concave functions, it is enough if we check it for the concave majorant (the lowest concave function) of $\hat{F}_n(b) - \varepsilon$; see Figure 5a. And indeed, thanks to the fact that a concave function is the infimum of linear functions, an envelope, this intuition provides a lead to the design of an algorithm that can implement the test. If concave functions did not have the representation as an envelope, the intuition of checking the concave majorant of $\hat{F}_n(b) - \varepsilon$ would be inappropriate.

The question is whether an approach in this spirit can be developed for the case of risk aversion. Namely, whether the differential characterization, $F''(b) < \Psi(b, F(b), F'(b))$, has a representation as an infimum with respect to some non linear functions, an envelope. We derive such a characterization to which we refer as *relative concavity;* we show that the latter constitutes a *testable* characterization. Then, we proceed by posing the question whether the the lowest distribution function that is everywhere above or equal to $\hat{F}_n(b) - \varepsilon$ satisfies the differential inequality $F''(b) < \Psi(b, F(b), F'(b))$. Having a representation of the characterization as an envelope makes it testable.

[3]The issue of how to implement a statistical test is ISACO

[4]A statistical test is consistent if for a sufficienty large sample the null hypothesis is rejected with probability one, if it is wrong.

To pinpoint the relation between the theoretical characterizations and relative concavity, note that distribution functions that satisfy the theoretical characterization are relatively concave, but the converse need not be true. However, every relatively concave distribution function F, is the limit of some sequence $F_1, F_2, ...$ of functions that satisfy the theoretical characterization. Hence, the set of relatively concave distribution functions is the closure of the set of functions satisfying the second-order differential inequality, $F''(b) < \Psi(b, F(b), F'(b))$. Since the testable characterization is an envelope, it allows to restrict attention to the lowest relatively concave distribution function within the corridor bounded by $\hat{F}_n(b) - \varepsilon$ and $\hat{F}_n(b) + \varepsilon$. Secondly, we know the non linear functions with respect to which the infimum is taken to obtain relative concavity, which indicates how to connect the relevant corners of $\hat{F}_n(b) - \varepsilon$. This sets the scene for the design of an algorithm.

A major weakness of our test is due to the unfortunate fact that economic theory does not offer any coherent and formal alternative to the equilibrium hypothesis. Hence, we must test a composite hypothesis; equilibrium versus non equilibrium. This is a crucial issue which, obviously, affects the power of the test.

Independently of the main role of relative concavity as a testable characterization of equilibrium behavior, it is worthwhile to note that this concept induces a partial order on two-parametric families of functions, and therefore on utilities whose arguments are bids and types (not necessarily in an additive fashion); we refer to this order as 'equilibrium compatibility'. The order stipulates that if a given distribution of bids, F, is equilibrium compatible with our auction game under a utility U_1, and U_1 is less equilibrium compatible than U_2, then F is equilibrium compatible for U_2 as well. This property allows to derive more general results from the statistical test. Since the formula that represents this order resembles the Arrow-Pratt concept of risk aversion, it is interesting to compare the two orders. If U_2 is sufficiently more risk averse than U_1 according to the Arrow-Pratt concept, it often appears that U_2 is also more equilibrium compatible than U_1. However, in general, none of the orders implies the other. Within the CARA family of functions, the two relations are equivalent.

The plan of the paper is as follows: In Section 2 we derive the theoretical characterization of equilibrium distribution of bids. In Section 3 we derive the concept of relative concavity and show that it is necessary for a distribution of bids to correspond to an equilibrium; by construction, relative concavity is an envelope. In Section 4 we elaborate on single crossing conditions for an auction environment under risk aversion, and show that relative concavity in conjunction with single crossing are necessary and (almost) sufficient conditions to assure that a distribution of bids corresponds to an equilibrium. Combining the results obtained in Sections 3 and 4, we compare in Section 5 the concept of *relative concavity* with the theoretical characterization. Using relative concavity, we show in Section 6 how to test the equilibrium hypothesis; we also develope an efficient algorithm that implements the test. Finally, in section 7 we show that relative concavity, in addition to its main role as a testable characterization of equilibrium bids, introduces a new partial order on bi-parametric families of functions, and therefore on utilities whose arguments are, say, bids and types.

2 A Theoretical Characterization of Equilibrium Distribution of Bids

Consider a symmetric first price sealed bid private value auction with N players, having a utility function $U(-b,t)$, where t is a type and b is the bid. With one exception, our assumptions about U are borrowed from Theorem 2 in [9].

$$U(-b,t) \quad \text{is twice continuously differentiable,} \tag{2.1}$$

$$\frac{\partial}{\partial b}U(-b,t) < 0\,, \tag{2.2}$$

$$\frac{\partial^2}{\partial b^2}U(-b,t) < 0\,, \tag{2.3}$$

$$\frac{\partial^2}{\partial t \partial b}\ln U(-b,t) > 0\,. \tag{2.4}$$

We treat t as a valuation (rather than an arbitrary preference parameter running over $[0,1]$):

$$U(-t,t) = 0\,. \tag{2.5}$$

Accordingly, $U(-b,t)$ is considered for $r \le b \le t$ only, where r is the reserve price, and (2.1–2.4) are assumed for $r \le b \le t$. The authors in [9] assumed also $\frac{\partial}{\partial t}U(-b,t) > 0$, which we don't need and, it can be proved that, neither did they. We show that the only device needed to order types is (2.4) which, as we prove, is the single crossing condition for an auction environment. Without it, our as well as their results do not hold. However, once (2.4) was assumed, $\frac{\partial}{\partial t}U(-b,t) > 0$ is redundant.

Remark 1 In the case of a one-dimensional utility function, $U(-b,t) = U(t-b)$, the assumptions become as follows: for $x \ge 0$, $U(x), U'(x), U''(x)$ exist and are continuous, $U'(x) > 0$, $U''(x) < 0$, and $U(0) = 0$. The inequality $(\ln U(x))'' < 0$ follows.

The N types are assumed independent and identically distributed according to a cumulative distribution function G such that $0 < G(r) < 1$, $G(t_{\max}) = 1$ for some $t_{\max}$, and there is a continuous density $G'(t) > 0$ for $t \in [r, t_{\max}]$.

These assumptions (about the auction, the utility function, and the distribution of types), and only these, are valid throughout the paper, except for remarks. Some Remarks may address more special cases, by introducing additional (local) assumptions valid within the remark (see for example Remark 1). Sometimes (see Sect. 4) a remark discusses a more general case by, locally, annuling some of the global assumptions. Other exceptions are Lemma 2 and Appendix A, where $G'(r+)$ is not assumed to exist.

Theorem 2 in [9] ensures existence and uniqueness of a symmetric equilibrium bid function $B(t)$, $t \in [r, t_{\max}]$, the function being continuously differentiable and satisfying the differential equation

$$B'(t)\frac{\partial}{\partial b}\Big|_{b=B(t)} \ln U(-b,t) + (N-1)\frac{d}{dt}\ln G(t) = 0 \tag{2.6}$$

with the boundary condition $B(r) = r$. Thus, the distribution of types generates a distribution of bids, F:

$$F(B(t)) = G(t)\ . \tag{2.7}$$

The equilibrium distribution of bids is in a sense more regular than that of types. For example, the density of bids is differentiable even when the distribution of types is not (see Theorem 1). From an economic point of view, it is natural to attribute these properties to the competitive environment induced by the bidding process. However, there is one result that deviates from this intuition. The density of bids tends to infinity near the reserve price, irrespective of risk aversion (see Example 1 and Fig. 1 at the end of the section). The singularity of the distribution of bids is an unexpected implication of continuity of the density of types at the same point (Lemma 1). Moreover, they are equivalent (Lemma 2). From an economic point of view, this result is somewhat at variance with what one might expect, given that risk aversion makes bidders to bid closer to their valuations. However, types close (from above) to the reserve price use the utility function within a small region, where it is almost linear.

The singularity at the reserve price poses a problem when we prove Theorem 1 and therefore, the following Lemmas are needed.

Lemma 1 *The function F is twice continuously differentiable on $(r, b_{\max}]$, and*

$$\begin{aligned} &\lim_{b\to r+}\left((b-r)^{1/2}\frac{d}{db}\ln F(b)\right) = \\ = \quad & -2\lim_{b\to r+}\left((b-r)^{3/2}\frac{d^2}{db^2}\ln F(b)\right) = \left(\frac{1}{2(N-1)}\frac{G'(r+)}{G(r)}\right)^{1/2}. \end{aligned}$$

Lemma 2 *Assume that $G'(t)$ does not converge to any finite positive limit for $t \to r+$ (while other assumptions remain in force). Then $(b-r)^{3/2}(d^2/db^2)\ln F(b)$ does not converge to any finite negative limit for $b \to r+$.*

The proofs, and some additional asymptotic relations near the reserve price, are relegated to Appendix A.

The following Theorem provides a formal theoretical characterization of the equilibrium distribution of bids.

Theorem 1 *A distribution F of bids is generated by some distribution of types if and only if F satisfies the following conditions:*

(a) $0 < F(r) < 1$; $F(b_{\max}) = 1$ *for some* $b_{\max}$; *F is twice continuously differentiable on* $(r, b_{\max}]$; *and* $F'(b) > 0$ *for* $b \in (r, b_{\max}]$;

(b) $(b-r)^{3/2}F''(b)$ *has a finite negative limit for* $b \to r+$;

(c)

$$\frac{d^2}{db^2}\ln F(b) < -\frac{1}{N-1}\frac{\partial^2}{\partial b^2}\ln U(-b,t) \tag{2.8}$$

whenever $b \in (r, b_{\max})$ *and* $t \in (r, +\infty)$ *are related by*

$$\frac{d}{db} \ln F(b) = -\frac{1}{N-1} \frac{\partial}{\partial b} \ln U(-b,t) \, . \tag{2.9}$$

Proof The "only if" part: assume that F is generated by some G; we have to prove (a), (b), (c).

Lemma 1 ensures that $F''(b)$ exists and is continuous on $(r, b_{\max}]$, and satisfies Condition (b); the rest of Condition (a) is obvious.

It follows from (2.7) that

$$B'(t) \frac{d}{db}\Big|_{b=B(t)} \ln F(b) = \frac{d}{dt} \ln G(t) \, ; \tag{2.10}$$

combine it with (2.6) and cancel $B'(t)$:

$$\frac{\partial}{\partial b}\Big|_{b=B(t)} \ln U(-b,t) + (N-1) \frac{d}{db}\Big|_{b=B(t)} \ln F(b) = 0 \, , \tag{2.11}$$

which means that $b = B(t)$ implies (2.9). On the other hand, if $b \neq B(t)$, then $b = B(t_1)$ for some $t_1 \neq t$. The pair (b, t_1) satisfies (2.9), therefore (b,t) cannot satisfy (2.9), since the right-hand side of (2.9) is strictly monotone in t due to (2.4). So, (2.9) holds if and only if $b = B(t)$.

We have to prove (2.8) for $b = B(t)$. Insert $b = B(t)$ into (2.9), differentiate in t, and replace $B(t)$ with b again:

$$\begin{aligned} &\frac{db}{dt} \cdot \frac{d^2}{db^2} \ln F(b) = \\ &= -\frac{1}{N-1} \Big(\frac{db}{dt} \cdot \frac{\partial^2}{\partial b^2} \ln U(-b,t) + \frac{\partial^2}{\partial t \partial b} \ln U(-b,t) \Big) \, ; \end{aligned} \tag{2.12}$$

$$\begin{aligned} &\frac{db}{dt} \cdot \Big(\frac{d^2}{db^2} \ln F(b) + \frac{1}{N-1} \cdot \frac{\partial^2}{\partial b^2} \ln U(-b,t) \Big) = \\ &= -\frac{1}{N-1} \frac{\partial^2}{\partial t \partial b} \ln U(-b,t) \, . \end{aligned} \tag{2.13}$$

Here $db/dt = B'(t) > 0$. The right-hand side is negative by (2.4), so, (2.8) follows.

The "if" part: assume (a), (b), (c); we have to prove that F is generated by some G. For a fixed b consider the right-hand side of (2.9) as a function of t. The function decreases strictly due to (2.4). For $t \to b$, the function tends to $+\infty$ due to (2.5). For $t \to +\infty$, the function tends to 0, since

$$-\frac{\partial}{\partial b} \ln U(-b,t) \le \frac{1}{t-b} \tag{2.14}$$

due to (2.3) and (2.5). It follows that for any $b \in (r, b_{\max})$ there is one and only one t satisfying (2.9). Smoothness of F and U ensures that t is continuously differentiable in b. Differentiating (2.9) in b while treating t as the function of b, we get the same as (2.12), (2.13) but multiplied by dt/db. The right-hand side of (2.13) is negative by (2.4); its left-hand side multiplied by dt/db is negative by (2.8); therefore $dt/db > 0$; we see that t is a strictly increasing function of $b \in (r, b_{\max})$.

Let $b \to r$, then $F'(b) \to \infty$ due to (b), therefore $1/(t-b) \to \infty$ due to (2.9) and (2.14); so, $t \to r$ for $b \to r$. Let $b \to b_{\max}$, then $F'(b)$ has a non-zero limit due to (a), therefore $t - b$ remains bounded due to (2.9) and (2.14); so, there is $t_{\max} \in (r, +\infty)$ such that $t \to t_{\max}$ for $b \to b_{\max}$. The derivative dt/db has a finite non-zero limit for $b \to b_{\max}$, since (2.13) remains non-singular. (In contrast, there is a singularity for $b \to r$, see (b).) The inverse function maps $[r, t_{\max}]$ onto $[r, b_{\max}]$.

So, (2.9) determines a strictly increasing, continuously differentiable on $(r, t_{\max}]$ function $b = B(t)$ such that $B(r) = r$ and $B(t_{\max}) = b_{\max}$; we'll see that B is the equilibrium bid function. Define G by (2.7). Clearly, G satisfies our assumptions: $0 < G(r) < 1$, $G(t_{\max}) = 1$, and $G'(t) > 0$ is continuous on $(r, t_{\max}]$. However, it is not yet clear, what happens to $G'(r+)$.

As before, (2.7) gives (2.10). Also, (2.9) for $b = B(t)$ gives (2.11). Combining (2.10) and (2.11) we get (2.6), which means that the function B satisfies the differential equation of the equilibrium. The initial condition $B(r) = r$ is also satisfied. Therefore B is the equilibrium bid function. Lemma 2 ensures that $G'(r+)$ exists. □

Note that the set of functions that satisfy (a), (b) and (c) in Theorem 1 is not closed: a sequence of functions that satisfy the conditions may converge to a function that violates them. A closure of the set is established when a *testable* characterization (relative concavity) is derived.

Remark 2 Considering the case, $U(-b, t) = U(t - b)$, Condition (c) of Theorem 1 can be written as follows:

$$\frac{F''(b)}{F(b)} - \Big(\frac{F'(b)}{F(b)}\Big)^2 < -\frac{1}{N-1}\Big(\frac{U''(t-b)}{U(t-b)} - \Big(\frac{U'(t-b)}{U(t-b)}\Big)^2\Big) \tag{2.15}$$

whenever $b \in (r, b_{\max})$ and $t \in (r, +\infty)$ are related by

$$\frac{F'(b)}{F(b)} = \frac{1}{N-1}\frac{U'(t-b)}{U(t-b)} \, . \tag{2.16}$$

Remark 3 Under risk neutrality, $U(t - b) = t - b$, we have

$$(\ln F(b))'' < \frac{1}{N-1}\frac{1}{(t-b)^2} \tag{2.17}$$

whenever

$$(\ln F(b))' = \frac{1}{N-1}\frac{1}{t-b} \, . \tag{2.18}$$

Eliminating t we get what may be referred to as a second-order differential inequality:

$$(\ln F(b))'' < (N-1)(\ln F(b))'^2 \, . \tag{2.19}$$

In fact, the condition "(2.18) implies (2.17)" is an implicit differential inequality, since (2.18) determines t as a function of b. Similarly, Condition (c) of Theorem 1 is an implicit second-order differential inequality; unfortunately, it does not have a form like (2.19). We may rewrite (2.19) as $F''(b)F(b) - F'^2(b) < (N-1)F'^2(b)$; $F''(b)F(b) < NF'^2(b)$. Using the general identity $(F^{-(N-1)}(b))'' = -(N-1)F^{-(N+1)}(b) \cdot (F''(b)F(b) - NF'^2(b))$ we get

$$(F^{-(N-1)}(b))'' > 0\,. \tag{2.20}$$

Consequently, under risk neutrality, the required condition is that the function $1/F^{N-1}(b)$ is convex. Choosing appropriate coordinates (transforming $1/F^{N-1}(b)$), we can pose the problem in terms of a concave distribution function.

To illustrate the convexity property (2.20), and the singularity at the reserve price, we solve explicitly and graph the following example:

Example 1 Assume risk neutrality, and let valuations be uniformly distributed on $(0,1)$, and $N = 2$. In the absence of any reserve price the equilibrium strategy is $B(t) = t/2$, generating bids distributed uniformly on $(0,1/2)$, in contradiction to Condition (b) of Theorem 1. We got a contradiction because the assumptions under which the theorem was proven stipulate a reserve price r such that $0 < G(r) < 1$; we ignored this condition. To conform with the assumption that underlie the Theorem, we assume $0 < r < 1$. Then, the optimal strategy can be calculated by the well-known integral formula:

$$B(t) = \frac{1}{2}\Big(t + \frac{r^2}{t}\Big)\,. \tag{2.21}$$

It generates the following distribution of bids, shown in Fig. 1:

$$F(b) = b + \sqrt{b^2 - r^2}, \qquad F'(b) = 1 + \frac{b}{\sqrt{b^2 - r^2}} \quad \text{for } b \in (r,\ (r^2+1)/2)\,. \tag{2.22}$$

Now, $(b-r)^{3/2}F''(b) = -r^2(b+r)^{-3/2} \to -(1/2)\sqrt{r/2}$ for $b \to r+$, which conforms to Condition (b). The function $1/F(b)$, shown in Fig. 2 (a), is convex, which conforms to Condition (c) via (2.20).

3 Equilibrium Bid Distributions and Envelopes

Having established the theoretical characterization of equilibrium conditions in terms of observables (distribution of bids) is merely the first, and by no means the major, step on the way to show how to conduct a statistical test of the equilibrium hypothesis. The first question that must be resolved is finding, or designing, an appropriate statistical test. Next, unlike when testing statistical hypotheses in lower dimensions where, typically, there are no problems of implementation of a test, in our case the implementation of the test is a major problem. We suggest in Sect. 6 a statistical test

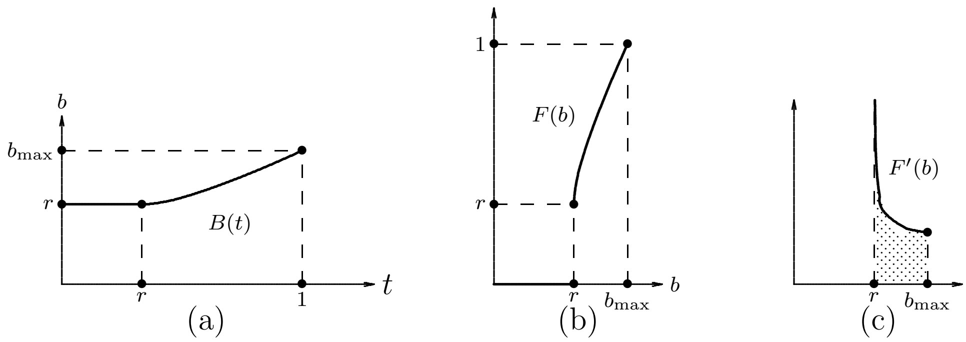


Figure 1: (a); Optimal strategy, (b); cumulative distribution function of equilibrium bids, (c); the density of its absolutely continuous part.

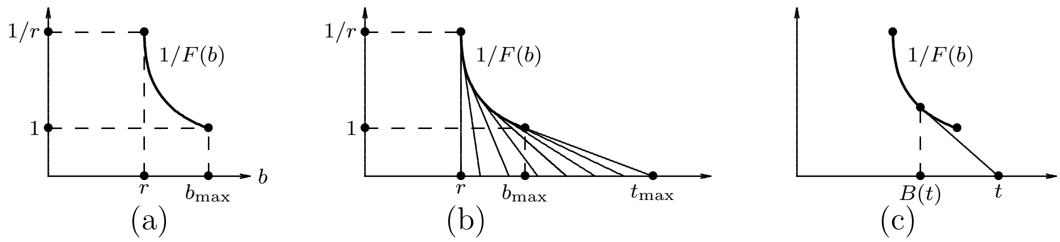


Figure 2: (a); The convex function $1/F(b)$, (b) as a supremum of linear functions, (c); a graphical representation of optimality.

that defines a 'corridor' bounded by the functions $\hat{F}_n(b)-\varepsilon$ and $\hat{F}_n(b)+\varepsilon$, where $\hat{F}_n(b)$ is the empirical distribution, and ε is a number determined by the desired significance level of the test, see Sect. 6 and Figure 6(a) there. The corridor is the infinite dimensional equivalent of an ordinary confidence interval. The test stipulates that if F is an equilibrium bid function (satisfies the second order differential inequality,

$$F''(b) < \Psi(b, F(b), F'(b)) \tag{3.1}$$

implied by (a), (b) and (c) in Theorem 1), it should be possible to 'squeeze' it within the corridor bounded by the functions $\hat{F}_n(b)-\varepsilon$ and $\hat{F}_n(b)+\varepsilon$.

At this point, we face the problem that there seems to be no established procedure of how such a test can be implemented. To illustrate the difficulties that emerge, and indicate avenues that may lead to a solution, we consider the case of risk neutrality.

Remark 4 We have shown in Remark 3 that under risk neutralit a distribution function $H(b)$ corresponds to an equilibrium if it is concave and fits within the corridor bounded by the functions $\hat{F}_n(b)-\varepsilon$ and $\hat{F}_n(b)+\varepsilon$. The question is how to verify this condition; should we screen (how?) all concave functions and check whether we can find one that satisfies the condition. It is not clear how to implement such a search, and fortunately there is no need to think about it thanks to the representation of a concave function as an infimum of linear functions. This property reduces the problem to one of finding whether the concave majorant of $\hat{F}_n(b)-\varepsilon$ is within the corridor; the majorant can be found by connecting the relevant corners of $\hat{F}_n(b)-\varepsilon$ by straight lines. Such a formulation allows algorithmic considerations, which makes concavity a *testable* characterization.

Trying to mimick this procedure under risk aversion requires to think whether it is possible to replace the characterization obtained in Theorem 1 by one that has a representation as an infimum with respect to some non linear functions. To proceed in this vein, we take a different route than the one that led us to Theorem 1, keeping in mind that we need a characterization that is *testable*.

Optimality of a bid b_1 for a type t_1 means that for any b_2

$$U(-b_1, t_1)F^{N-1}(b_1) \geq U(-b_2, t_1)F^{N-1}(b_2) \,, \tag{3.2}$$

that is,

$$F(b_2) \leq F(b_1) \cdot (U(-b_1, t_1)/U(-b_2, t_1))^{1/(N-1)} \,. \tag{3.3}$$

Equality is reached for $b_2 = b_1$ (and only there); thus,

$$F(b_2) = \min(F(b_1) \cdot (U(-b_1, t_1)/U(-b_2, t_1))^{1/(N-1)}) \,, \tag{3.4}$$

where the minimum is taken over all pairs (t_1, b_1) such that $b_1 = B(t_1)$. Considering (3.4), we observe a factor, $F(b_1) \cdot U^{1/(N-1)}(-b_1, t_1)$, that does not contain b_2, and another factor, $U^{-1/(N-1)}(-b_2, t_1)$, that does not contain F. Introduce

$$\varphi_t(b) = U^{-1/(N-1)}(-b, t) \,, \tag{3.5}$$

$$C_t = F(B(t)) \cdot U^{1/(N-1)}(-B(t), t) \,, \tag{3.6}$$

then (3.4) takes the form $F(b) = \min_t(C_t\varphi_t(b))$. Equality is reached if and only if $b = B(t)$. Hence, we arrive at a necessary condition:

$$F(b) = \min_t(C_t\varphi_t(b)) \quad \text{for some } C_t \geq 0 \,. \tag{3.7}$$

Distribution functions that satisfy (3.7)are *relatively concave.* Note that (3.7) is an *envelope* condition; F must be the envelope of a one-parametric subfamily $(C_t\varphi_t)_t$ of the two-parametric family $(C\varphi_t)_{C,t}$ of non-linear functions. The indices after the parenthesis are parameters of the functions, while the argument b in $\varphi_t(b)$ was suppressed. The question is whether (3.7) is also a sufficient condition. Namely, if F satisfies (3.7), was it generated by the equilibrium strategy B under some distribution G of types.

Before addressing this question in the next section, we take a small detour and show that (3.7) is sufficient under risk neutrality. Namely, we show, by construction, how to recover C_t, B and then, the distribution of valuations, G. In a nutshell, (3.7) is sufficient under risk neutrality since the latter satisfies single crossing conditions. As we show in the next section, to make (3.7) sufficient under risk aversion, single crossing conditions must be explicitly assumed.

Remark 5 Under risk neutrality (3.5) becomes $\varphi_t(b) = (t - b)^{-1/(N-1)}$ and (3.7) becomes $F(b) = \min_t C_t(t - b)^{-1/(N-1)}$, which can be rewritten as $1/F^{N-1}(b) =$

$\max_t \tilde{C}_t(t-b)$ for some $\tilde{C}_t$. It means that $1/F^{N-1}(b)$ is the maximum for a family of linear (in b) functions, therefore it is a convex function, which explains (2.20), $(F^{-(N-1)}(b))'' > 0$. Any convex function is the maximum of some linear functions, that correspond to tangent lines. The tangent line to the graph of $1/F^{N-1}(\cdot)$, touching it at $(b, 1/F^{N-1}(b))$, is the graph of the linear function $b \mapsto \tilde{C}_t(t-b)$, where

$$\tilde{C}_t = -(1/F^{N-1}(b))' , \tag{3.8}$$

$$t = b - \frac{1/F^{N-1}(b)}{(1/F^{N-1}(b))'} ; \tag{3.9}$$

the latter is equivalent to (2.18), which explains (2.9) for the special case of risk neutrality.

Hence, $\tilde{C}_t(t-b)$ can be obtained from (3.8) and the strategy $B(t)$ can be obtained from (3.9); the functional dependence between t and b defines, implicitly, the strategy. Equipped with $B(t)$ and $F(t)$, we can obtain the distribution of valuations by the relation, $F(B(t)) = G(t)$.

4 Single Crossing in an Auction Environment

The single crossing property is applied when each type t is described by its indifference curves $f(x,y,t) = \text{const}$ on the plane of some relevant variables x, y. The slope of a contour that corresponds to a type evaluated at a point (x,y) of the plane, is obtainable via the total differential:

$$\begin{aligned} 0 &= df(x,y,t) = f_1(x,y,t)\,dx + f_2(x,y,t)\,dy , \\ & f_1 = \frac{\partial f}{\partial x} , \quad f_2 = \frac{\partial f}{\partial y} , \\ \frac{dy}{dx} &= -\frac{f_1(x,y,t)}{f_2(x,y,t)} . \end{aligned}$$

Types may be ordered by their slopes; in general, each (x,y) determines its own order on the set of types. The single crossing condition stipulates that the order does not depend on (x,y). Usually, one assumes that types are parametrized according to the order, and formulates the single crossing condition in the form

$$\frac{\partial}{\partial t}\Big(-\frac{f_1(x,y,t)}{f_2(x,y,t)} \Big) > 0 . \tag{4.1}$$

In our case, a type t may be described by the indifference curves $pU(-b,t) = \text{const}$ on the plane (p,b), where p is the winning probability and b stands for a bid. The slope along the contour (of constant expected utility) is given by:

$$\frac{db}{dp} = -\frac{\frac{\partial}{\partial p}(pU(-b,t))}{\frac{\partial}{\partial b}(pU(-b,t))} = -\frac{1}{p\frac{\partial}{\partial b}\ln pU(-b,t)} , \tag{4.2}$$

which should be increasing with t; that is, $(\partial/\partial b)\ln U(-b,t)$ should increase, which is just Condition (2.4). Some reformulations: $(\partial/\partial b)\ln U(-b,t_1) < (\partial/\partial b)\ln U(-b,t_2)$

for $b < t_1 < t_2$; $(\partial/\partial b)\ln(U(-b,t_2)/U(-b,t_1)) > 0$ for $b < t_1 < t_2$; $U(-b_1,t_2)/U(-b_1,t_1) < U(-b_2,t_2)/U(-b_2,t_1)$ for $b_1 < b_2 < t_1 < t_2$; and finally,

$$\frac{U(-b_2,t_2)}{U(-b_1,t_2)} > \frac{U(-b_2,t_1)}{U(-b_1,t_1)} \quad \text{for } b_1 < b_2 < t_1 < t_2\,, \tag{4.3}$$

which makes explicit the meaning of the order of types, implicit in (2.4).

Maskin and Riley [9] did not indicate whether they were aware that their assumption (2.4) is the single crossing condition for an auction environment. However, they used it to prove existence and uniqueness of equilibrium strategy under risk aversion. We show in this section that when single crossing is not satisfied, not only sufficiency of (3.7) does not hold, but other 'irregularities' emerge; for example, the monotone relation between types and bids may be violated.

Both (2.4) and (4.3) are simpler than the original formulation of single crossing, (4.1), since the latter involves a two-parametric family of curves. (The formal reason for the simplification is the disappearance of p immediately after (4.2).) The family consists of indifference curves $pU(-b,t) = C = \text{const}$ on the plane (p,b), for all types t. In (4.2), b was treated as a function of p, but now we prefer to treat p as a function of b, namely, $p = C/U(-b,t)$. Using (3.5) we may write the family as $(C\varphi_t^{N-1})_{C,t}$. The single crossing condition constraints intersections between indifference curves of types. Namely, if $t_1 < t_2$ and $C_1\varphi_{t_1}^{N-1}(b) = C_2\varphi_{t_2}^{N-1}(b)$, then $(C_1\varphi_{t_1}^{N-1}(b))' > (C_2\varphi_{t_2}^{N-1}(b))'$; derivatives in b are meant; the order of types is inverted since dp/db is the inverse to db/dp. The constraint excludes multiple crossing, which explains the term "single crossing". It is invariant under monotone transformations of p (or b, or both). In particular, the transformation $\tilde{p} = p^{1/(N-1)}$ turns curves $p = C\varphi_t^{N-1}(b)$ into $\tilde{p} = \tilde{C}\varphi_t(b)$ with $\tilde{C} = C^{1/(N-1)}$. Renaming the parameter, we conclude that the family $(C\varphi_t)_{C,t}$ satisfies the single crossing condition. If $C_1\varphi_{t_1}(b_{12}) = C_2\varphi_{t_2}(b_{12})$ then $C_1\varphi_{t_1}(b) < C_2\varphi_{t_2}(b)$ for $b < b_{12}$, and $C_1\varphi_{t_1}(b) > C_2\varphi_{t_2}(b)$ for $b > b_{12}$. (Compare it with (4.3).)

To show that without explicit resort to single crossing, (3.7) is not sufficient under risk aversion, we show that it is not compatible with (2.8) and (2.9) which are necessary and sufficient. To this effect, assume that F satisfies (3.7); for each b there is a t such that $F(b) = C_t\varphi_t(b)$. The function $C_t\varphi_t(b_1) - F(b_1)$ of b_1 reaches its minimal value 0 at $b_1 = b$, where necessary conditions of first and second order are satistied:

$$F'(b) = C_t\varphi_t'(b) \quad \text{and} \quad F''(b) \le C_t\varphi_t''(b) \quad \text{whenever} \quad F(b) = C_t\varphi_t(b)\,. \tag{4.4}$$

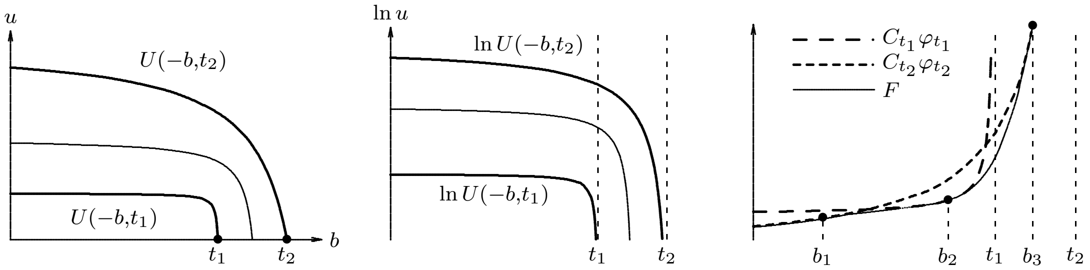
We may also rewrite (3.7) as

$$\ln F(b) = \min_t(\ln C_t + \ln\varphi_t(b)) \quad \text{for some } C_t \tag{4.5}$$

and get, similarly to (4.4),

$$\frac{d}{db}\ln F(b) \;=\; \frac{d}{db}\ln\varphi_t(b)\,, \tag{4.6}$$

$$\frac{d^2}{db^2}\ln F(b) \;\le\; \frac{d^2}{db^2}\ln\varphi_t(b) \tag{4.7}$$



(a) (b) (c)

Figure 3: A utility function, (a), and its logarithm, (b), violating (2.4); consequences for $C_t\varphi_t$ and F, (c): double crossing, and non-monotone dependence between bids and types (b_1 and b_3 correspond to t_2, while b_2 corresponds to t_1).

whenever $\ln F(b) = \ln C_t + \ln \varphi_t(b)$ (that is, $F(b) = C_t\varphi_t(b)$). Taking into account that $\ln \varphi_t(b) = -(1/(N-1)) \ln U(-b, t)$ we observe that (4.6) and (4.7) conform with (2.9) and (2.8) respectively, (though "$<$" in (2.8) is a little stronger than "$\le$" in (4.7)). However, (2.8) and (2.9) appeared in the combination "(2.8) whenever (2.9)", while (4.6) and (4.7) appear in a different combination: "both (4.6) and (4.7) whenever $F(b) = C_t\varphi_t(b)$". Hence, {(2.8), (2.9)} and {(4.6), (4.7)} are not equivalent, and since {(2.8), (2.9)} were proven to be necessary and sufficient, {(4.6), (4.7)} are not.

Assuming single crossing, Lemma 3 stipulates that (3.7) guarranties (weak) monotonicity of strategies and reconciles between {(2.8), (2.9)} and {(4.6), (4.7)}.

Lemma 3 *Let F satisfy* (3.7). *Then*

(a) *the equality $F(b) = C_t\varphi_t(b)$ establishes an increasing correspondence between b and t.*

(b) *If F is continuously differentiable, then* (4.6) *holds if and only if $F(b) = C_t\varphi_t(b)$.*

(c) *If F is twice continuously differentiable, then* (4.6) *implies* (4.7).

Proof (a) Assume the contrary: there exist $b_1 < b_2 < t_1 < t_2$ such that $F(b_1) = C_{t_2}\varphi_{t_2}(b_1)$ and $F(b_2) = C_{t_1}\varphi_{t_1}(b_2)$. We have $C_{t_1}\varphi_{t_1}(b_1) \ge C_{t_2}\varphi_{t_2}(b_1)$ and $C_{t_1}\varphi_{t_1}(b_2) \le C_{t_2}\varphi_{t_2}(b_2)$. Take $b \in [b_1, b_2]$ such that $C_{t_1}\varphi_{t_1}(b) = C_{t_2}\varphi_{t_2}(b)$. If $b_1 < b$ then single crossing requires $C_{t_1}\varphi_{t_1}(b_1) < C_{t_2}\varphi_{t_2}(b_1)$, which is impossible. Therefore $b = b_1 < b_2$; single crossing requires $C_{t_1}\varphi_{t_1}(b_2) > C_{t_2}\varphi_{t_2}(b_2)$, which is also impossible.

(b) We know that (4.6) holds whenever $F(b) = C_t\varphi_t(b)$. Assume that there are b, t satisfying (4.6) while $F(b) \ne C_t\varphi_t(b)$. By (3.7), $F(b) = C_{\tilde t}\varphi_{\tilde t}(b)$ for some $\tilde t$. Therefore, (4.6) is satisfied for $b, \tilde t$, but also for b, t. It contradicts the strict monotonicity in t of the right-hand side of (4.6), ensured by (2.4) and (3.5).

(c) We know that (4.7) holds whenever $F(b) = C_t\varphi_t(b)$. It remains to use (b). □

Consequenty, by Lemma 3, the necessary condition (3.7) is close to be sufficient assuming single crossing. However, some obstacles persist. First, (3.7) does not exclude the possibility that F coincides with a single $C_t\varphi_t$ on some interval, in which case all b in the interval correspond to a single t. Second, (3.7) does not exclude the possibility that F' jumps down at some b, in which case the single b corresponds to all t of some interval. (Thus, the "increasing correspondence" in Lemma 3 (a) is not an increasing function but, in general, a many-to-many correspondence.) Third, the "$\le$" sign in (4.7) is weaker than the "$<$" sign in (2.8). We show in Appendix D that these technical issues can be resolved by introducing an appropriate smoothing procedure.

The following remark illustrates the claim made before that without single crossing monotonicity of strategies may fail.

Remark 6 Consider the utility function shown in Fig. 3 (a); it satisfies (2.1)–(2.3) and (2.5) but violates (2.4), since (2.4) requires $(\partial/\partial b)\ln U(-b,t)$ to increase in t, which is not the case, see Fig. 3 (b). Some choice of C_t leads to the situation shown in Fig. 3 (c): graphs of $C_{t_1}\varphi_{t_1}$ and $C_{t_2}\varphi_{t_2}$ cross more than once. Note that each function $C_t\varphi_t$ is increasing, and tends to infinity at t. The steeper function, $C_{t_1}\varphi_{t_1}$, in Fig. 3 (c) tends to ∞ at t_1 and the flatter one, $C_{t_2}\varphi_{t_2}$, does so at t_2 (the latter does not show clearly because of technical reasons). The relation between b and t, defined by the equality $F(b) = C_t\varphi_t(b)$, is not monotone (neither increasing nor decreasing), since it holds for (b_1,t_2), (b_2,t_1), and (b_3,t_2).

5 Relative Concavity and the Theoretical Characterization

Having explored various aspects of *relative concavity*, we summarize its properties and relation with the theoretical characterization. Starting with the two-parametric family of all linear functions (straight lines) and taking minima, we get concave functions. Similarly, we may start with another two-parametric family of non-linear functions (curves) satisfying the single crossing condition, and take minima. Doing so, we get functions to which we refer as relatively concave. Having no intention to develop here an abstract theory of relative concavity, we consider only families of the form $(C\varphi_t)_{C,t}$, where $C \in [0,\infty)$ and $t \in (r,\infty)$ are the two parameters, $\varphi_t(b) = (1/U(-b,t))^{1/(N-1)}$ for $b \in [r,t)$, $\varphi_t(b) = +\infty$ for $b \in [t,\infty)$, and U is a given function satisfying (2.1)–(2.5). Consequently, a function F is called relatively concave with respect to U (and N, which is omitted), if it satisfies (3.7).

Relating the theoretical characterization, (a), (b), (c) of Theorem 1, to relative concavity, it follows that distribution functions that satisfy these conditions are relatively concave (follows immediately from Theorem 1 and (3.4)); the converse does not hold. Every relatively concave distribution function F is the limit of some sequence $F_1, F_2, \dots$ of functions satisfying Conditions (a), (b), (c) of Theorem 1. It means basically, that every relatively concave function can be approximated by smooth, strictly relatively concave functions. (Other requirements of (a), (b) can be satisfied by small local changes.) The distribution functions which satisfy the theoretical characterization are inconsistent; namely, functions that satisfy the characterization converge to a function that does not satisfy it. On the other hand, distribution functions which are relatively concave, are consistent; namely, if $F_1, F_2, \dots$ are relatively concave and

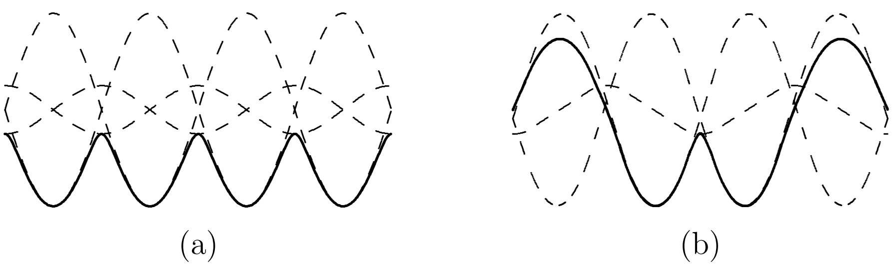


Figure 4: Minimum of solutions of the differential equation satisfies the differential inequality (a), but the converse is generally wrong (b).

$F_n \to F$, then F is also relatively concave. Such functions describe the closure of the non-closed set described by Theorem 1.

Condition (3.7) is closely related to the differential characterization stated in the Theorem. Namely, functions $C\varphi_t$ are solutions of the differential equation $F''(b) = \Psi(b, F(b), F'(b))$ where Ψ is the same as in (3.1). Heuristically, one could write a "proportion"

$$\frac{\text{differential equation}}{\text{differential inequality}} = \frac{\text{family of functions}}{\text{their envelopes}},$$

which is instructive to discuss in much simpler cases. For example, take $\Psi_0(b, F(b), F'(b)) = 0$ identically, then the differential equation becomes $F''(b) = 0$, its solutions being linear functions $F(b) = c_1 b + c_2$. The differential inequality becomes $F''(b) < 0$, its solutions being concave functions. A function is a minimum (envelope) of linear functions if and only if it is concave. However, a concave function F need not satisfy $F''(b) < 0$ for two reasons: possible nonexistence of $F''(b)$, or $F''(b)$ may vanish. The set of all concave functions is the closure of the set of F satisfying $F''(b) < 0$.

Another example: take $\Psi_1(b, F(b), F'(b)) = -F(b)$, then the differential equation becomes $F''(b) = -F(b)$, its solutions being harmonic (sine) functions $F(b) = c_1 \cos b + c_2 \sin b$. A minimum (envelope) of harmonic functions satisfies the differential inequality $F''(b) < -F(b)$ (see Fig. 4(a)) with the two reservations (possible nonexistence of $F''(b)$ or vanishing of $F''(b) + F(b)$). However, F satisfying $F''(b) < -F(b)$ need not be a minimum of harmonic functions, see Fig. 4(b). It means that necessity of (3.7) (in contrast to sufficiency) is not a general property of second order differential equations and inequalities.

Finally, we may ask what got 'lost' on the conceptual level when we replace the theoretical characterization by relative concavity. Remembering Theorem 1, the following Corollary provides the answer:

Corollary 1 *For any distribution function F, the following two conditions are equivalent.*

(a) *There exist distributions $F_1, F_2, \dots$ of bids, generated by some distributions $G_1, G_2, \dots$ of types, such that $F_k(b) \to F(b)$ for all $b \in (r, \infty)$.*

(b) *The function F is relatively concave on (r, ∞).*

Before completing this section we elaborate briefly on some technical properties of relative concavity that illustrate, from a different point of view than so far, the need for having it established.

Remember that in the risk neutral case the second order differential inequality becomes ordinary concavity (see the transition from (2.19) to (2.20)). This result leads to the interesting question whether a similar transition can be found for the risk averse case by rectifying the family of functions $(C\varphi_t)_{C,t}$. More specifically, is there a nonlinear transformation of the (b, y)-plane that transforms each curve $y = C\varphi_t(b)$ into a straight line? Should this be possible, we would not need the new concept of relative concavity.

It turns out that the existence of such a transformation is closely related to Desargues's theorem of projective geometry, illustrated in Fig. 5 (a), see for example [4], p. 6. Namely, the answer to our question depends on an answer to an entirely different, and very nice, mathematical problem that goes back to the 17-th century. As opposed to the family of straight lines, an arbitrary two-parametric family of curves, in general in general does not satisfy Desargues's theorem, and therefore cannot be rectified. However, our family, $(C\varphi_t)_{C,t}$ is somewhat special. We do not elaborate here on this issue, and the interested reader is referred to Appendix B where it is shown that our family of function can not be rectified. Namely, there is no coordinate system under which the non linear functions on the plane that emerge under risk aversion, can be linearized, which implies that risk aversion introduces a new structure that can not be handled by relying on methods applied under risk neutrality. Consequently, the new concept of relative concavity is essential.

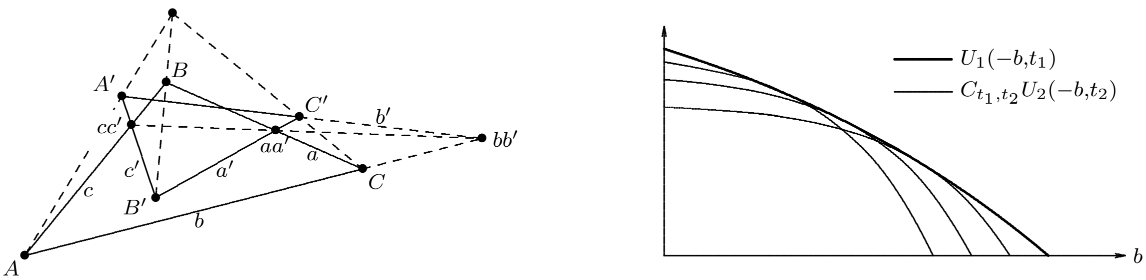
6 The Statistical Test

Given a sample $b_1, \dots, b_n$ from an unknown continuous distribution F, we may estimate F by the empirical distribution function

$$\hat{F}_n(b) = \frac{\#\{k : b_k \le b\}}{n} \,. \tag{6.1}$$

It is well-known that $\|\hat{F}_n - F\| = \sup_b |\hat{F}_n(b) - F(b)|$ converges to 0 in probability, when $n \to \infty$. Moreover,

$$\Pr(\|\hat{F}_n - F\| > \lambda/\sqrt{n}) \xrightarrow[n\to\infty]{} 1 - L(\lambda) \,, \tag{6.2}$$



(a) (b)

Figure 5: (a) Desargues's two-triangle theorem. If two triangles have corresponding verticles joined by concurrent lines, then the intersections of corresponding sides are collinear. That is, if the lines AA', BB', CC' all pass through one point, then the intersection points aa', bb', cc' all lie on one line. (b) Comparing two utility functions via (7.1).

$$L(\lambda) = 1 - 2\sum_{k=1}^{\infty}(-1)^{k+1}\exp(-2k^2\lambda^2) = \\ = \frac{\sqrt{2\pi}}{\lambda}\sum_{k=1}^{\infty}\exp(-(2k-1)^2\pi^2/(8\lambda^2))\,, \tag{6.3}$$

which is a well-known result of A. Kolmogorov (see for example [11] p. 142–143), widely used for testing statistical hypotheses. Quantiles (critical values) $C_{\alpha,n}$ of $\sqrt{n}\|\hat{F}_n - F\|$ for finite n do not depend on F; they are tabulated, and $C_{\alpha,n} \to C_\alpha$ for $n \to \infty$, C_α being defined by $1 - L(C_\alpha) = \alpha$.

A simple hypothesis F may be tested by means of the Kolmogorov statistic $\sqrt{n}\|\hat{F}_n - F\|$; if it exceeds its critical value $C_{\alpha,n}$, the hypothesis F is rejected. The probability of type I error (rejection of the true hypothesis) is equal to α.

A composite hypothesis, formalized by a set $\mathcal{F}$ of allowed distribution functions F, also may be tested; namely, the hypothesis $\mathcal{F}$ is rejected, if $\sqrt{n}\|\hat{F}_n - F\|$ exceeds $C_{\alpha,n}$ for all $F \in \mathcal{F}$. The statistic used is $\sqrt{n}\operatorname{dist}(\hat{F}_n, \mathcal{F}) = \sqrt{n}\inf_{F\in\mathcal{F}}\|\hat{F}_n - F\|$. Such a procedure is known as the conservative Kolmogorov test, see for example [3]. The probability of type I error does not exceed α. It depends on $\mathcal{F}$ and F. An extreme example: if $\mathcal{F}$ is dense among all distributions, then $\operatorname{dist}(\hat{F}_n, \mathcal{F}) = 0$ always, and the hypothesis is never rejected.

Every F that does not belong to $\mathcal{F}$ may be considered as an alternative. Assume that $\mathcal{F}$ is not a closed set; it means that some alternatives $F \notin \mathcal{F}$ belong to the closure $\overline{\mathcal{F}}$ of $\mathcal{F}$, that is, satisfy $\operatorname{dist}(F, \mathcal{F}) = 0$. The test is ineffective against such alternatives, since the type II error (non-rejecting of the wrong hypothesis) is highly probable: $\operatorname{dist}(\hat{F}_n, \mathcal{F}) \le \operatorname{dist}(\hat{F}_n, F) + \operatorname{dist}(F, \mathcal{F}) = \operatorname{dist}(\hat{F}_n, F)$, therefore $\Pr(\operatorname{dist}(\hat{F}_n, \mathcal{F}) > C_{\alpha,n}/\sqrt{n}) \le \Pr(\operatorname{dist}(\hat{F}_n, F) > C_{\alpha,n}/\sqrt{n}) = \alpha$.

Is it a drawback of the conservative Kolmogorov test? For some bizarre hypotheses $\mathcal{F}$ it really is. An example: the hypothesis claims that the unknown distribution is concentrated on rational numbers only. Here, $\mathcal{F}$ is dense among all distributions, therefore the test is utterly ineffective. In contrast, the evident test “reject the hypothesis if the sample contains at least one irrational number” rejects any continuous

F almost surely, irrespective of n.

Return to equilibrium bids, assuming that N and U are known, while G is unknown. The corresponding set of distributions, described by Theorem 1, is not closed (its closure is described by Corollary 1). Does it mean that the conservative Kolmogorov test is inappropriate here? No, it does not, for the following reason. The convergence $F_k \to F$ of Conjecture 1 is in fact stronger than claimed there, it is a convergence in variation: $\int |F'_k(b) - F'(b)|\, db \to 0$ for $k \to \infty$, which can be shown using the fact that all F_k and F are relatively concave. The convergence in variation (called also norm convergence) implies convergence of probabilities for every event. Any test rejects the hypothesis when a definite event R occurs; we have $\Pr(R|F) = \lim \Pr(R|F_k) \le \alpha$, which shows that no test can be effective against $F \in \overline{\mathcal{F}}$. The mathematical reason is that the set $\mathcal{F}$ has the same closure in different topologies (the weak topology and the norm topology). Given that no statistical test is able to discriminate $\mathcal{F}$ and $\overline{\mathcal{F}}$, from now on our hypothesis is the closed set $\mathcal{F} = \mathcal{F}_{N,U}$ of relatively concave distribution functions.

The conservative Kolmogorov test is consistent against all alternatives, and powerful against $1/\sqrt{n}$ local alternatives. Here, the only relevant property of $\mathcal{F}_{N.U}$ is its closeness: if $F \notin \mathcal{F}_{N,U}$ then $\operatorname{dist}(F, \mathcal{F}_{N,U}) > 0$. If n, α, β satisfy

$$\sqrt{n} \operatorname{dist}(F, \mathcal{F}_{N,U}) \ge C_{\alpha,n} + C_{\beta,n} \tag{6.4}$$

then

$$\Pr(\sqrt{n} \operatorname{dist}(\hat{F}_n, \mathcal{F}_{N,U}) \le C_{\alpha,n}) \le \beta \tag{6.5}$$

due to the triangle inequality: $\operatorname{dist}(F, \mathcal{F}_{N,U}) \le \operatorname{dist}(\hat{F}_n, \mathcal{F}_{N,U}) + \|\hat{F}_n - F\|$, therefore $\Pr(\sqrt{n} \operatorname{dist}(\hat{F}_n, \mathcal{F}_{N,U}) \le C_{\alpha,n}) \le \Pr(\sqrt{n} \operatorname{dist}(F, \mathcal{F}_{N,U}) - \sqrt{n}\|\hat{F}_n - F\| \le C_{\alpha,n}) \le \Pr(\sqrt{n}\|\hat{F}_n - F\| \ge C_{\beta,n}) = \beta$. Now, if $n \to \infty$ while $F \notin \mathcal{F}_{N,U}$ is fixed, then for each $\beta > 0$ the inequality (6.4) holds eventually, implying (6.5). It is the consistency:

$$\Pr(\sqrt{n} \operatorname{dist}(\hat{F}_n, \mathcal{F}_{N,U}) \le C_{\alpha,n}) \xrightarrow[n\to\infty]{} 0\,, \tag{6.6}$$

the left-hand side being the probability of type II error (non-rejecting the wrong hypothesis) under the fixed alternative F. Consider now $1/\sqrt{n}$ local alternatives F_n, that is, $\sqrt{n} \operatorname{dist}(F_n, \mathcal{F}_{N,U}) \to C$, $0 < C < \infty$. Let β satisfy $C > C_\alpha + C_\beta$, then $\sqrt{n} \operatorname{dist}(F_n, \mathcal{F}_{N,U}) \ge C_{\alpha,n} + C_{\beta,n}$ for n large enough, therefore (6.5) holds under F_n, which shows that the test is powerful against $1/\sqrt{n}$ local alternatives.

So, the hypothesis $\mathcal{F}_{N,U}$ is rejected if

$$\sqrt{n} \operatorname{dist}(\hat{F}_n, \mathcal{F}_{N,U}) > C_{\alpha,n}\,, \tag{6.7}$$

in other words, if there is no relatively concave function F on $[r, \infty)$ such that for all $b \in [r, \infty)$

$$\hat{F}_n(b) - \varepsilon \le F(b) \le \hat{F}_n(b) + \varepsilon \tag{6.8}$$

where $\varepsilon = C_{\alpha,n}/\sqrt{n}$. How to determine, whether such F exists, or not?

Consider a simpler case: the usual concavity (relative to linear functions). Geometrical intuition tells us (see Fig. 6(a)) that we may check only a single function F,

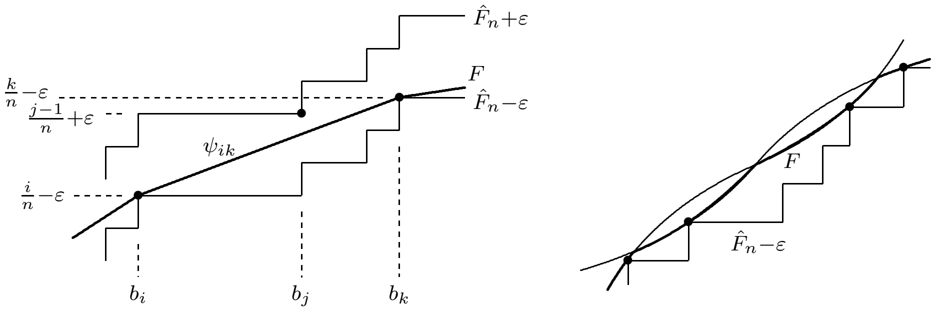


Figure 6: how to check (6.8) for the usual concavity (a); a bizarre majorant may appear (b) in the absence of single crossing.

namely the concave majorant of $\hat{F}_n - \varepsilon$ (that is, the minimal concave F such that $F \ge \hat{F}_n - \varepsilon$), and the majorant is piecewise linear, each linear segment extending from a node $(b_i, \hat{F}_n(b_i+) - \varepsilon) = (b_i, i/n - \varepsilon)$ to another node $(b_k, \hat{F}_n(b_k+) - \varepsilon) = (b_k, k/n - \varepsilon)$. Any pair (i,k) such that $1 \le i < k \le n$ determines a linear function $\psi_{ij}(b)$ interpolating the two nodes,

$$\psi_{ik}(b_i) = \frac{i}{n} - \varepsilon \, , \qquad \psi_{ik}(b_k) = \frac{k}{n} - \varepsilon \, , \tag{6.9}$$

but only those pairs (i,k) are relevant that satisfy $\hat{F}_n(b) - \varepsilon \le \psi_{ik}(b)$ for all b. The inequality $\hat{F}_n(b) - \varepsilon \le F(b)$ is satisfied trivially, the other inequality $F(b) \le \hat{F}_n(b) + \varepsilon$ may be checked at nodes:

$$\psi_{ik}(b_j) \le \frac{j-1}{n} + \varepsilon \tag{6.10}$$

for all i,j,k such that $1 \le i < j < k \le n$. (If it holds for all relevant pairs (i,k) then it holds also for other pairs.) So, for the usual concavity (6.7) could be checked in a finite number of operations.

Fortunately, all said above remains true for relative concavity. Of course, functions ψ_{ik} are no more linear; rather, they belong to the given two-parametric family $\{C\varphi_t\}$:

$$\psi_{ik}(b) = C_{ik}\varphi_{t_{ik}}(b) \, , \tag{6.11}$$

where C_{ik} and t_{ik} are determined by (6.9). It is implied that $i/n > \varepsilon$. Single crossing is crucial, it excludes situations like the one shown in Fig. 6(b).

So, the test (6.7) is not too difficult to carry out due to the specific structure of $\mathcal{F}_{N,U}$. The conservative Kolmogorov test may be used to build a conservative confidence set, which is well-known (see for example [3]). There are several possibilities, all based on the two requirements:

$$F \text{ is relatively concave w.r.t. } U \, , \tag{6.12}$$

$$\hat{F}_n - \varepsilon \le F \le \hat{F}_n + \varepsilon \, ; \tag{6.13}$$

of course, (6.12) is treated via (3.7) and (3.5), and $\varepsilon = C_{\alpha,n}/\sqrt{n}$ in (6.13).

If we consider U known while F unknown, we may use (6.12), (6.13) for defining a set of bid distributions F. The set contains the (unknown) true bid distribution with the probability at least $1-\alpha$, which means that the set is a (conservative) confidence set.

If we consider both U and F unknown, we may use (6.12), (6.13) for defining a confidence set of pairs (U, F).

If we consider F unknown while U nuisance (that is, unknown and uninteresting), we may define a confidence set of F by requiring (6.12), (6.13) to hold for at least one U. However, we consider in detail only the following, most interesting case.

Let U be unknown while F nuisance. We define a confidence set $\hat{\mathcal{U}}_n$ by requiring (6.12), (6.13) to hold for at least one F. Fortunately, the set $\hat{\mathcal{U}}_n$ is defined by a finite number of inequalities: (6.10) must hold whenever $1 \le i < j < k \le n$ and $i/n > \varepsilon$; as before, ψ_{ik} is defined by (6.9) and (6.11). More explicitly, t_{ik} is the solution of the equation $\varphi_{t_{ik}}(b_i)/\varphi_{t_{ik}}(b_k) = (i/n-\varepsilon)/(k/n-\varepsilon)$, that is,

$$\frac{U(-b_k, t_{ik})}{U(-b_i, t_{ik})} = \left(\frac{i/n-\varepsilon}{k/n-\varepsilon}\right)^{N-1} ; \tag{6.14}$$

the solution exists and is unique, since the left-hand side increases in t due to (2.4), see (4.3), and it increases from 0 to 1 when t increases from b_k to $+\infty$. Further, $C_{ik}^{N-1} = (i/n-\varepsilon)^{N-1}U(-b_i, t_{ik}) = (k/n-\varepsilon)^{N-1}U(-b_k, t_{ik})$, and (6.10) takes the form

$$\left(\frac{i}{n}-\varepsilon\right)^{N-1} U(-b_i, t_{ik}) = \left(\frac{k}{n}-\varepsilon\right)^{N-1} U(-b_k, t_{ik}) \le \left(\frac{j-1}{n}+\varepsilon\right)^{N-1} U(-b_j, t_{ik}) \, . \tag{6.15}$$

That is the empirical restriction on the utility function U.

An example. Restrict ourselves to the CARA utility function, $U_\alpha(-b, t) = (1/\alpha)(1-e^{-\alpha(t-b)})$. Denote for convenience $x_i = (i/n-\varepsilon)^{N-1}$, $y_j = ((j-1)/n+\varepsilon)^{N-1}$, $x_k = (k/n-\varepsilon)^{N-1}$, then

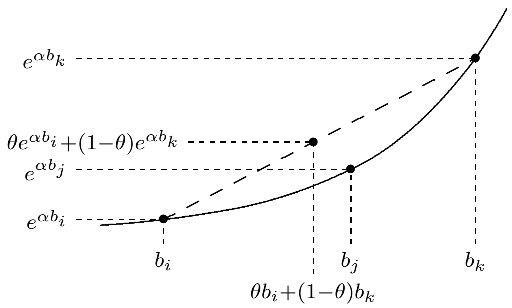
$$t_{ik} = \frac{1}{\alpha} \ln \frac{x_k e^{\alpha b_k} - x_i e^{\alpha b_i}}{x_k - x_i} \, , \tag{6.16}$$

$$C_{ik}^{N-1} = \frac{1}{\alpha} x_i x_k \frac{e^{\alpha b_k} - e^{\alpha b_i}}{x_k e^{\alpha b_k} - x_i e^{\alpha b_i}} \, , \tag{6.17}$$

and (6.15) can be transformed into

$$e^{\alpha b_j} \le \theta e^{\alpha b_i} + (1-\theta) e^{\alpha b_k} \tag{6.18}$$

where $\theta = (x_i(x_k-y_j))/(y_j(x_k-x_i))$, that is, $1-\theta = (x_k(y_j-x_i))/(y_j(x_k-x_i))$, see Fig. 7 (a). It may happen that $b_j \le \theta b_i + (1-\theta)b_k$, in which case (6.18) is satisfied for all $\alpha \in [0,\infty)$ due to convexity of the exponential function. Otherwise, if $b_j > \theta b_i + (1-\theta)b_k$, there is a value $\alpha_{ijk}^{\min} \in (0,\infty)$ such that (6.18) is satisfied for all $\alpha \in [\alpha_{ijk}^{\min}, \infty)$ but violated for all $\alpha \in [0, \alpha_{ijk}^{\min})$. So, we define $\alpha_{ijk}^{\min}$ as the root α of the equation $e^{\alpha b_j} = \theta e^{\alpha b_i} + (1-\theta)e^{\alpha b_k}$, provided that $b_j > \theta b_i + (1-\theta)b_k$,



(a)

Figure 7: The empirical restriction (6.18) on the parameter α of the CARA utility function.

otherwise $\alpha_{ijk}^{\min} = 0$. Then we put $\alpha_{\min} = \max_{i<j<k} \alpha_{ijk}^{\min}$. The empirical restriction on α is $\alpha \in [\alpha_{\min}, \infty)$.

In general, the confidence set $\hat{\mathcal{U}}_n$ is monotone w.r.t. the partial order (7.1). That is, if U_1 is less equilibrium compatible than U_2, then $U_1 \in \hat{\mathcal{U}}_n$ implies $U_2 \in \hat{\mathcal{U}}_n$.

7 Equilibrium Compatibility

The concept of relative concavity gives raise to another partial order for two-parametric families, and therefore for utility functions; we shall use the term 'equilibrium compatible' in reference to utility functions that obey this order. Hence, U_1 is less equilibrium compatible than U_2, if any function relatively concave w.r.t. U_1 is, necessarily, relatively concave w.r.t. U_2. That is, if a function f can be represented as $f(b) = \min_t(C_{t,1}U_1^{-1/(N-1)}(-b,t))$, then it can be represented as $f(b) = \min_t(C_{t,2}U_2^{-1/(N-1)}(-b,t))$. The following equivalent definition eliminates the seeming dependence on N: U_1 can be represented as

$$U_1(-b,t_1) = \max_{t_2} C_{t_1,t_2} U_2(-b,t_2) \,, \tag{7.1}$$

see Fig. 5 (b). The condition may be expressed by means of logarithmic derivatives (similarly to (2.8), (2.9)):

$$\frac{\partial^2}{\partial b^2} \ln U_1(-b,t_1) \geq \frac{\partial^2}{\partial b^2} \ln U_2(-b,t_2) \tag{7.2}$$

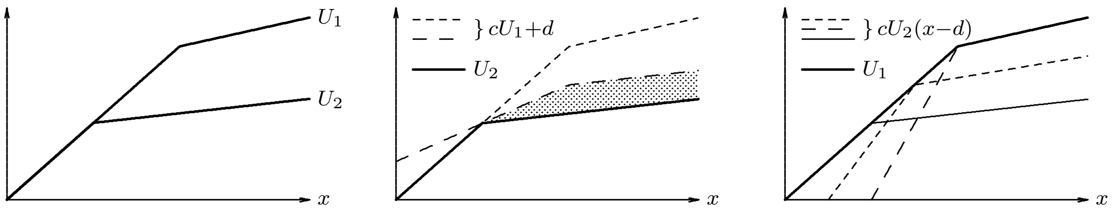
whenever

$$\frac{\partial}{\partial b} \ln U_1(-b,t_1) = \frac{\partial}{\partial b} \ln U_2(-b,t_2) \,. \tag{7.3}$$

For single-variable utility functions, $U_k(-b,t) = U_k(t-b)$, it means

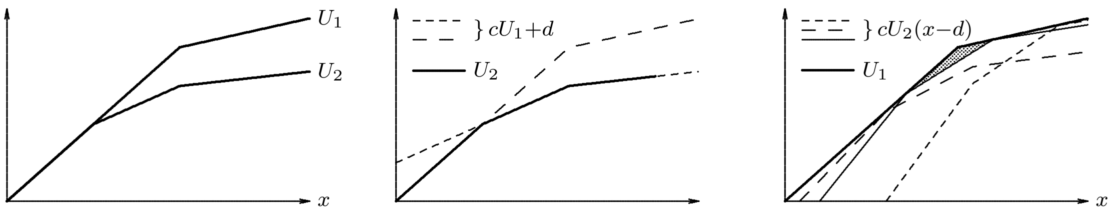
$$\frac{U_1''(x_1)}{U_1(x_1)} \geq \frac{U_2''(x_2)}{U_2(x_2)} \quad \text{whenever} \quad \frac{U_1'(x_1)}{U_1(x_1)} = \frac{U_2'(x_2)}{U_2(x_2)} \,. \tag{7.4}$$

An example: CARA utility functions $U_1 = U_{\alpha_1}$, $U_2 = U_{\alpha_2}$ with parameters $\alpha_1 < \alpha_2$. Here $U_k''(x) = -\alpha_k U'(x)$, which makes (7.4) evident.



(a) (b) (c)

Figure 8: (a) Two utility functions. (b) U_2 is not more risk averse than U_1. (c) Nevertheless, U_2 is more equilibrium compatible than U_1.



(a) (b) (c)

Figure 9: (a) Two utility functions. (b) U_2 is more risk averse than U_1. (c) Nevertheless, U_2 is not more equilibrium compatible than U_1.

The reader must have noticed the structural similarity between (7.4) and the definition of risk aversion. In both definitions we have ratios of derivatives of different order of utilities. Not surprisingly, (7.4) is more complicated but so is the concept (equilibrium compatibility versus risk aversion).

The practical importance of this order is that if two utilities, say, U_1 and U_2 are ordered so that U_1 is less equilibrium compatible than U_2, then if we know that a given distribution of bids is compatible with equilibrium assuming that the utility function is U_1, then it is also compatible with equilibrium if the utility is U_2.

A natural question is the relation between the two orders. It appears that no one of them implies the other. Appropriate counterexamples are shown in Figures 8, 9. The functions used are neither smooth nor strictly concave, they are piecewise linear, which is of no importance, since the same holds for smooth strictly concave functions that are close enough to the functions shown. The relation "U_2 is more risk averse than U_1" is violated for U_1, U_2 of Fig. 8 and satisfied for U_1, U_2 of Fig. 9, which can be checked by plotting the function J such that $U_2(x) = J(U_1(x))$; it is concave for the case of Fig. 9 but not Fig. 8. However, both relations can be expressed in terms of relative concavity. Namely, U_2 is more risk averse than U_1 if and only if U_2 is relatively concave w.r.t. the family $(cU_1(x) + d)$ with two parameters c, d. In Fig. 8 (b), the relatively concave majorant of U_2 differs from U_2 by the shaded region. In Fig. 9 (b), U_2 is the minimum of just two functions from the family. On the other hand, U_2 is more equilibrium compatible than U_1 if and only if U_1 is relatively convex w.r.t. the family $(cU_2(x - d))$ with two parameters c, d. That is the case for Fig. 8 (c) but not Fig. 9 (c); here the convex minorant of U_1 differs from U_1 by the shaded triangle. For more on the relation between the two orders, see Appendix C.

8 Testing, Using Experimental or Real Life Data; Pros and Cons

We showed in this paper how to tests the equilibrium hypothesis in the case of a first price auction, using *real life* or *field* data. To the best of our knowledge nothing in this vein has been suggested in the literature. However, there is a large literature in which various postulates of auction theory, including equilibrium, have been challenged using experimental data; see Kagel and Roth (1995), ch.7. Researchers that used experimental data generated many interesting results over the years, while tests based on field data have not been conducted yet. Nevertheless, a comparison of the advantages of each method is useful in order to understand the merits of each.

An almost inherent weakness of any tests based on field data is that we do not have data on valuations, or signals. Since any statistical test must be formulated in terms of observables only, we suggested to test the equilibrium hypothesis by deriving hypotheses about the distribution of bids. This is obviously an indirect method which must entail a price in terms of efficiency. There can not be a test that can be applied not using data on valuations, and come up with efficiency that could have been obtained if valuations were known. This conclusion leads to an important *potential* advantage of testing the hypothesis using exprimental data. In such experiments, valuations and bids are known to whoever conducted the experiment.

Unfortunately, the experimental literature reveals a serious shortcoming of this methodology as it has been applied. All papers that we were aware of were based on models that assumed uniform distributions of signals. We guess that the only reason for using, almost as a rule, uniform distributions, was simplicity. This is understandable, but a heavy price was paid; whatever was achieved applies *only* to this, very special, distribution.

We believe that even the most 'fanatic' proponents of the equilibrium hypothesis are not so naive as to claim that agents always play equilibrium; namely, that they are willing and able to do so. It is a sure guess that such a naive approach stands no chance to be verified. At the best we can hope that equilibrium is played in 'most' or 'many' situations. Hence, proving or disproving a claim using uniform distributions does not have implications that go beyond this, very special, case. Of course, one can claim that this situation does not indicate a weakness of the methodology, since there are no conceptual problems with conducting the experimental studies using more general distributions, although the fact that the huge experimental literature sticked to this distribution only, serves as an indication that experimentalists find a more general approach hard to pursue.

As shown in Section 7, our test can be conducted assuming *any* concave utility. But then, at least initially, our test may not be easy to comprehend, although once understood, can be implemented routinely.

Appendix

A Asymptotics near the Reserve Price

In the appendix, G is assumed to be continuously differentiable on $(r, t_{\max}]$, not necessarily on $[r, t_{\max}]$. When needed, existence of $G'(r+)$ will be stipulated explicitly.

Lemmas 4, 5 will be used in the proof of Lemmas 1, 2. As usual, $X \sim Y$ means $X/Y \to 1$; $o(1)$ means something that tends to 0; and $o(X)$ means $X \cdot o(1)$.

Lemma 4 *The following asymptotic relations hold when $b, t \to r$ and $t > b > r$:*

$$\frac{\partial}{\partial b} \ln U(-b,t) \sim \frac{-1}{t-b}\,, \tag{A.1}$$

$$\frac{\partial^2}{\partial b^2} \ln U(-b,t) \sim \frac{-1}{(t-b)^2}\,, \tag{A.2}$$

$$\frac{\partial^2}{\partial t \partial b} \ln U(-b,t) \sim \frac{1}{(t-b)^2}\,. \tag{A.3}$$

Proof. $(\partial/\partial b)U(-b,t) = -U_1(-b,t)$ and $(\partial/\partial t)U(-b,t) = U_2(-b,t)$, functions U_1, U_2 being continuous and strictly positive on the closed region $r \le b \le t < \infty$. Differentiating (2.5) gives $U_1(-t,t) = U_2(-t,t)$. We assume that $U_1(-r,r) = U_2(-r,r) = 1$ without loss of generality, since multiplying U by any positive constant leaves (A.1)–(A.3) invariant. Note that

$$U(-b,t) \sim t-b \quad \text{for } b,t \to r \text{ and } t > b > r\,, \tag{A.4}$$

since $U(-b,t) = U(-b,t) - U(-t,t) = U_1(-r,r) \cdot (t-b) + o(t-b)$. We have

$$\begin{aligned}
&U(-b,t)\frac{\partial}{\partial b} \ln U(-b,t) = -U_1(-b,t) \to -1\,,\\
&U^2(-b,t)\frac{\partial^2}{\partial b^2} \ln U(-b,t) = -U_1^2(-b,t) + U(-b,t)\frac{\partial^2}{\partial b^2}U(-b,t) \to -1\,,\\
&U^2(-b,t)\frac{\partial^2}{\partial t \partial b} \ln U(-b,t) =\\
&= U_1(-b,t)U_2(-b,t) - U(-b,t)\frac{\partial^2}{\partial t \partial b}U(-b,t) \to 1\,;
\end{aligned}$$

it remains to substitute (A.4). □

Lemma 5 *The following asymptotic relations hold when $b, t \to r$ and $t > b > r$:*

$$t - B(t) \sim t - r\,, \tag{A.5}$$

$$B'(t) \sim (N-1)(t-r)\frac{d}{dt} \ln G(t)\,, \tag{A.6}$$

$$B(t) - r \sim (N-1)(t-r)\Big(\ln G(t) - \frac{1}{t-r}\int_r^t \ln G(s)\,ds \Big)\,, \tag{A.7}$$

$$\frac{d}{db}\Big|_{b=B(t)} \ln F(b) \sim \frac{1}{N-1}\frac{1}{t-r}\,, \tag{A.8}$$

$$\Big(\frac{d}{dt}\ln G(t)\Big)\cdot\Big(1 + o(1) - (N-1)(t-r)^2 \frac{d^2}{db^2}\Big|_{b=B(t)} \ln F(b)\Big) \sim$$

$$\sim \frac{1}{N-1}\frac{1}{t-r}\,. \tag{A.9}$$

Proof. Combine the differential equation (2.6) and (A.1):

$$\frac{B'(t)}{t-B(t)} \sim (N-1)\frac{d}{dt}\ln G(t)\,; \tag{A.10}$$

integrability of the right-hand side near t implies the same for the other side. We have

$$\int_r^t \frac{B'(s)}{s-B(s)}\,ds \ge \int_r^t \frac{B'(s)}{t-B(s)}\,ds = -\ln(t-B(s))\Big|_{s=r}^{s=t} = \ln\frac{t-r}{t-B(t)} \ge 0\,;$$

when $t \to r$, the integral tends to 0, therefore the logarithm tends to 0, which means (A.5). Substituting (A.5) into (A.10) gives (A.6). Integrate it, using integration by parts:

$$\begin{aligned} B(t) - r = \int_r^t B'(s)\,ds &\sim \int_r^t (N-1)(s-r)\,d\ln G(s) = \\ = \ (N-1)(s-r)\ln G(s)\Big|_{s=r}^{s=t} &- (N-1)\int_r^t \ln G(s)\,ds = \\ = \ (N-1)(t-r)\Big(\ln G(t) &- \frac{1}{t-r}\int_r^t \ln G(s)\,ds \Big)\,, \end{aligned}$$

which is (A.7).

We know from (2.11) that

$$\frac{d}{db}\Big|_{b=B(t)} \ln F(b) = \frac{-1}{N-1}\frac{\partial}{\partial b}\Big|_{b=B(t)} \ln U(-b,t)\,; \tag{A.11}$$

it is not a vicious circle, since (2.11) was deduced from (2.1)–(2.7) only. Use (A.1) and (A.5):

$$\frac{d}{db}\Big|_{b=B(t)} \ln F(b) \sim \frac{-1}{N-1}\cdot\frac{-1}{t-B(t)} \sim \frac{1}{N-1}\frac{1}{t-r}\,,$$

which is (A.8).

Differentiate (A.11) in t (differentiability of the left-hand side follows from that of the other side) and use (A.2), (A.3), (A.5):

$$(N-1)B'(t)\frac{d^2}{db^2}\Big|_{b=B(t)} \ln F(b) =$$

$$
\begin{aligned}
&= -B'(t)\frac{\partial^2}{\partial b^2}\Big|_{b=B(t)} \ln U(-b,t) - \frac{\partial^2}{\partial t\partial b}\Big|_{b=B(t)} \ln U(-b,t) = \\
&= -B'(t)\frac{-1}{(t-B(t))^2}\cdot(1+o(1)) - \frac{1}{(t-B(t))^2}\cdot(1+o(1)) = \\
&= B'(t)\frac{1}{(t-r)^2}(1+o(1)) - \frac{1}{(t-r)^2}(1+o(1))\,.
\end{aligned}
$$

Further,

$$
\begin{aligned}
& B'(t)\cdot\Big(\frac{1}{(t-r)^2}(1+o(1)) - (N-1)\frac{d^2}{db^2}\Big|_{b=B(t)} \ln F(b)\Big) = \\
&= \frac{1}{(t-r)^2}(1+o(1))\,;
\end{aligned}
$$

substituting (A.6) gives

$$
\begin{aligned}
&(N-1)(t-r)\Big(\frac{d}{dt}\ln G(t)\Big)\cdot\Big(\frac{1}{(t-r)^2}(1+o(1)) - (N-1)\frac{d^2}{db^2}\Big|_{b=B(t)} \ln F(b)\Big) = \\
&= \frac{1}{(t-r)^2}(1+o(1))\,,
\end{aligned}
$$

thus, (A.9) is proved. □

Proof of Lemma 1 The right-hand side of (A.11) is continuously differentiable in t, therefore the same holds for the other side. Taking into account that $B'(t)$ exists and is continuous and strictly positive, we conclude that $\ln F(b)$ and $F(b)$ are twice continuously differentiable in b. Denote $C = G'(r+)/G(r)$, then $\ln G(t) = \ln G(r) + C(t-r) + o(t-r)$, and

$$
\frac{1}{t-r}\int_r^t \ln G(s)\,ds = \ln G(r) + \frac{1}{2}C(t-r) + o(t-r)\,.
$$

Substitute it into (A.7): $B(t) - r \sim (N-1)(t-r)\cdot\frac{1}{2}C(t-r)$, therefore

$$
t - r \sim \Big(\frac{2}{(N-1)C}(B(t)-r)\Big)^{1/2}\,. \tag{A.12}
$$

We use (A.9):

$$
\begin{aligned}
& C\Big(1 + o(1) - (N-1)(t-r)^2\frac{d^2}{db^2}\Big|_{b=B(t)} \ln F(b)\Big) \sim \frac{1}{N-1}\frac{1}{t-r}\,; \\
& -(N-1)(t-r)^2\frac{d^2}{db^2}\Big|_{b=B(t)} \ln F(b) = \\
&= \frac{1}{C}\frac{1}{N-1}\frac{1}{t-r}(1+o(1)) - 1 - o(1) \sim \frac{1}{C}\frac{1}{N-1}\frac{1}{t-r}\,; \\
& \frac{d^2}{db^2}\Big|_{b=B(t)} \ln F(b) \sim \frac{-1}{(N-1)^2 C(t-r)^3}\,;
\end{aligned}
$$

substitute (A.12):

$$\frac{d^2}{db^2}\Big|_{b=B(t)} \ln F(b) \sim \frac{-1}{(N-1)^2 C}\Big(\frac{2}{(N-1)C}(B(t)-r)\Big)^{-3/2} =$$
$$= \quad -2^{-3/2}(N-1)^{-1/2}C^{1/2}(B(t)-r)^{-3/2}\,,$$

therefore

$$\frac{d^2}{db^2} \ln F(b) \sim -\frac{1}{2}\Big(\frac{C}{2(N-1)}\Big)^{1/2}(b-r)^{-3/2}\,,$$

which is the second statement of the lemma. Integration gives

$$\frac{d}{db} \ln F(b) \sim \Big(\frac{C}{2(N-1)}\Big)^{1/2}(b-r)^{-1/2}\,,$$

which is the first statement of the lemma. □

Proof of Lemma 2 Assume existence of the limit

$$\lim_{b\to r+}\Big((b-r)^{3/2}\frac{d^2}{db^2}\ln F(b)\Big) = -C \in (-\infty, 0)\,; \tag{A.13}$$

we have to prove existence of the limit $G'(r+) \in (0,\infty)$. Integrating the relation $(d^2/db^2)\ln F(b) \sim -C(b-r)^{-3/2}$ gives $(d/db)\ln F(b) \sim 2C(b-r)^{-1/2}$. Substitute it into (A.8):

$$2C(B(t)-r)^{-1/2} \sim \frac{1}{N-1}\frac{1}{t-r}\,. \tag{A.14}$$

We combine (A.13) and (A.9):

$$\Big(\frac{d}{dt}\ln G(t)\Big)\cdot\Big(1+o(1)+(N-1)(t-r)^2\cdot C(B(t)-r)^{-3/2}\Big) \sim \frac{1}{N-1}\frac{1}{t-r}\,;$$

use (A.14):

$$\Big(\frac{d}{dt}\ln G(t)\Big)\cdot\Big(1+o(1)+(N-1)(t-r)^2 C\Big(\frac{1}{2C(N-1)(t-r)}\Big)^3\Big) \sim$$
$$\sim \frac{1}{N-1}\frac{1}{t-r}\,;$$
$$\Big(\frac{d}{dt}\ln G(t)\Big)\cdot\Big(1+o(1)+\frac{1}{8(N-1)^2C^2(t-r)}\Big) \sim \frac{1}{N-1}\frac{1}{t-r}\,;$$
$$\frac{d}{dt}\ln G(t) \sim \frac{8(N-1)^2C^2(t-r)}{(N-1)(t-r)} = 8(N-1)C^2\,,$$

which means that $G'(r+)/G(r) = 8(N-1)C^2$. □

B Transforming Curves into Straight Lines

A profound mathematical theory of two-parametric families of curves on a plane, pioneered by Eli Cartan in 1920-th (see [1], § 6H), is presented by V. Arnold ([1], § 6). A necessary and sufficient condition is given for such a family to be rectifiable, that is, for existence of a smooth transformation of the plane, turning the given family into the family of straight lines. The criterion is formulated in terms of the second-order differential equation describing the family.

We specialize the criterion for our case: the family of curves $y = C\varphi_t(b)$ on the (b, y)-plane, with two parameters C, t; here $\varphi_t(b) = U^{-a}(-b, t)$, $a = 1/(N-1)$, and U is a utility function satisfying (2.1–2.5). The corresponding second-order differential equation $y'' = \Phi(b, y, y')$ is obtained by eliminating C, t from

$$y = CU^{-a}(-b,t)\,, \tag{B.1}$$

$$y' = CaU^{-a-1}(-b,t)U'(-b,t)\,, \tag{B.2}$$

$$y'' = Ca(a+1)U^{-a-2}(-b,t)U'^2(-b,t) - CaU^{-a-1}(-b,t)U''(-b,t)\,; \tag{B.3}$$

(here U', U'' mean derivatives of U in its first argument, $(-b)$). We have

$$\frac{y'}{y} = a\frac{U'(-b,t)}{U(-b,t)}\,, \tag{B.4}$$

$$\frac{y''}{y} = a(a+1)\frac{U'^2(-b,t)}{U^2(-b,t)} - a\frac{U''(-b,t)}{U(-b,t)} = \frac{a+1}{a}\Big(\frac{y'}{y}\Big)^2 - a\frac{U''(-b,t)}{U(-b,t)}\,. \tag{B.5}$$

Given b, y, y', we should determine t from (B.4) and substitute it into (B.5).

Example 2 The CARA case, $U_\alpha(-b,t) = (1/\alpha)(1 - e^{-\alpha(t-b)})$, is especially simple. Here, $U'' = -\alpha U'$, thus (B.5) turns into

$$\frac{y''}{y} = \frac{a+1}{a}\Big(\frac{y'}{y}\Big)^2 + \alpha\frac{y'}{y}\,, \tag{B.6}$$

and the differential equation is obtained without determining t.

For rectifiability, it is necessary (see [1], § 6E) that the right-hand side $\Phi(b, y, y')$ of the differential equation $y'' = \Phi(b, y, y')$ is a polynomial in y' of degree ≤ 3. For CARA, according to (B.6), it appears to be a quadratic polynomial. In general, according to (B.5), the condition is

$$\frac{U''(-b,t)}{U(-b,t)} = P_3\Big(\frac{U'(-b,t)}{U(-b,t)}, b\Big)\,, \tag{B.7}$$

where P_3 is a polynomial of degree ≤ 3 in its first argument (its dependence on b need not be polynomial). Taking into account that $(\ln U)' = U'/U$ and $(\ln U)'' = (U''/U) - (U'/U)^2$, we may reformulate (B.7) as

$$\frac{\partial^2}{\partial b^2}\ln U(-b,t) = P_3\Big(\frac{\partial}{\partial b}\ln U(-b,t), b\Big) \tag{B.8}$$

for another polynomial P_3 of degree ≤ 3 in its first argument.

The necessary condition is not sufficient. In order to reach sufficiency, the dual condition is needed, see [1], § 6G. The duality means that variables and parameters interchange their roles; the equality $y = C\varphi_t(b)$ is now interpreted as a curve on the (t, C)-plane, parametrized by b, y. We have

$$C = yU^a(-b,t)\,, \tag{B.9}$$

$$C' = yaU^{a-1}(-b,t)U'(-b,t)\,, \tag{B.10}$$

$$C'' = ya(a-1)U^{a-2}(-b,t)U'^2(-b,t) + yaU^{a-1}(-b,t)U''(-b,t)\,; \tag{B.11}$$

this time derivatives are taken in t (thus, U' and U'' are not the same as in (B.1–B.7)). Further,

$$\frac{C'}{C} = a\frac{U'(-b,t)}{U(-b,t)}\,, \tag{B.12}$$

$$\frac{C''}{C} = a(a-1)\frac{U'^2(-b,t)}{U^2(-b,t)} + a\frac{U''(-b,t)}{U(-b,t)} = \frac{a-1}{a}\Big(\frac{C'}{C}\Big)^2 + a\frac{U''(-b,t)}{U(-b,t)}\,. \tag{B.13}$$

The dual differential equation is of the form $C'' = \Psi(t, C, C')$, and the dual condition requires $\Psi(t, C, C')$ to be a polynomial in C' of degree ≤ 3. Similarly to (B.7) and (B.8), we get

$$\frac{U''(-b,t)}{U(-b,t)} = Q_3\Big(\frac{U'(-b,t)}{U(-b,t)}, t\Big) \tag{B.14}$$

and finally

$$\frac{\partial^2}{\partial t^2}\ln U(-b,t) = Q_3\Big(\frac{\partial}{\partial t}\ln U(-b,t), t\Big) \tag{B.15}$$

for some polynomial Q_3 of degree ≤ 3 in its first argument. The two conditions (B.8) and (B.15), taken together, are necessary and sufficient for rectifiability.

For a single-variable utility function $U(-b,t) = U(t-b)$, the two condition (B.8) and (B.15) merge into one:

$$(\ln U(x))'' = P_3((\ln U(x))')\,. \tag{B.16}$$

However, for $x \to 0$ we have $U(x) \sim \text{const}/x$, $(\ln U(x))' \sim 1/x$, $(\ln U(x))'' \sim -1/x^2$, thus $P_3(1/x) \sim -1/x^2$ for $x \to 0$, which means that P_3 is in fact a quadratic polynomial, $P_3(z) = -z^2 + (\text{linear terms})$. On the other hand, for $x \to \infty$ we have $(\ln U(x))' \to 0$, $(\ln U(x))'' \to 0$, which means that $P_3(0) = 0$. Also, $(\ln U)'' = (U''/U) - (U'/U)^2 \le -(\ln U)'^2$, which means that $P_3(z) \le -z^2$. So, $P_3(z) = -z^2 - \alpha z$ for some $\alpha \ge 0$. Now, let $z(x) = (\ln U(x))'$, then $z'(x) = P_3(z(x))$, thus

$$\frac{dz}{dx} = P_3(z)\,; \qquad \frac{dz}{P_3(z)} = dx\,; \qquad -\int \frac{dz}{z^2 + \alpha z} = x + \text{const}\,.$$

The case $\alpha = 0$ leads to $U(x) = \text{const}\cdot x$, the risk neutral utility. The case $\alpha > 0$ gives $U(x) = \text{const}\cdot(1 - e^{-\alpha x})$, the CARA utility.

Under risk neutrality, a rectifying transformation follows from (2.20):

$$y = C\varphi_t(b) = C(t-b)^{-1/(N-1)} ,$$
$$\tilde{y} = 1/y^{N-1} = \tilde{C}(t-b) .$$

The CARA case can be reduced to the risk neutral case, which is of some independent interest. The differential equation (2.6) under CARA,

$$\alpha \frac{e^{-\alpha(t-b)}}{1-e^{-\alpha(t-b)}} \frac{db}{dt} = (N-1)\frac{d}{dt}\ln G(t) , \tag{B.17}$$

can be solved by the following change of variables:

$$t = \ln(\alpha\tilde{t}+1)/\alpha , \qquad dt = d\tilde{t}/(\alpha\tilde{t}+1) ,$$
$$b = \ln(\alpha\tilde{b}+1)/\alpha , \qquad db = d\tilde{b}/(\alpha\tilde{b}+1) ,$$
$$e^{-\alpha(t-b)} = \frac{\alpha\tilde{b}+1}{\alpha\tilde{t}+1} , \qquad G(t) = \tilde{G}(\tilde{t}) .$$

Substituting it into (B.17) we get

$$\frac{1}{\tilde{t}-\tilde{b}} \frac{d\tilde{b}}{d\tilde{t}} = (N-1)\frac{d}{d\tilde{t}}\ln \tilde{G}(\tilde{t}) ,$$

which is just (2.6) under risk neutrality.

C Two Orders in Variational Form

Let $U_\alpha(x)$ be a one-parametric family of single-variable utility functions, satisfying the conditions of Remark 1 for each α, and continuously differentiable in the parameter α. The family is said to be of increasing risk aversion, if U_{α_2} is more risk averse than U_{α_1} whenever $\alpha_1 \leq \alpha_2$. Likewise, the family is of increasing equilibrium compatibility, if U_{α_2} is more equilibrium compatible than U_{α_1} whenever $\alpha_1 \leq \alpha_2$.

Define another one-parametric family of functions

$$f_\alpha(x) = -\frac{U''_\alpha(x)}{U'_\alpha(x)} . \tag{C.1}$$

It is well-known that U_{α_2} is more risk averse than U_{α_1} if and only if $f_{\alpha_2}(x) \geq f_{\alpha_1}(x)$ for all x. Thus, (U_α) is of increasing risk aversion if and only if f_α increases in α, which is easy to express in the variational form:

$$0 \leq \frac{\partial}{\partial\alpha} f_\alpha(x) = -\frac{\partial}{\partial\alpha}\frac{U''_\alpha(x)}{U'_\alpha(x)} ; \tag{C.2}$$
$$U''\delta U' - U'\delta U'' \geq 0 , \tag{C.3}$$

where δ means variation,

$$\delta U_\alpha(x) = \Big(\frac{\partial}{\partial\alpha} U_\alpha(x)\Big) d\alpha\,, \quad \delta U'_\alpha(x) = \Big(\frac{\partial}{\partial\alpha} U'_\alpha(x)\Big) d\alpha = \Big(\frac{\partial^2}{\partial\alpha\partial x} U_\alpha(x)\Big) d\alpha \quad \dots \tag{C.4}$$

and, of course, $d\alpha$ is assumed to be positive.

Define one more one-parametric family of functions $g_\alpha(\cdot)$ by

$$-\frac{U''_\alpha(x)}{U_\alpha(x)} = g_\alpha\Big(\frac{U'_\alpha(x)}{U_\alpha(x)}\Big)\,; \tag{C.5}$$

it is correct, since U'_α/U_α is strictly monotone. According to (7.4), U_{α_2} is more equilibrium compatible than U_{α_1} if and only if $g_{\alpha_2}(z) \ge g_{\alpha_1}(z)$ for all z. Thus, (U_α) is of increasing equilibrium compatibility if and only if g_α increases in α, that is, $\delta g \ge 0$. Differentiate (C.5) in α:

$$-\delta\frac{U''}{U} = (\delta g)\Big(\frac{U'}{U}\Big) + g'\Big(\frac{U'}{U}\Big)\cdot\delta\frac{U'}{U}\,;$$

$$(\delta g)\Big(\frac{U'}{U}\Big) = \frac{U''\delta U - U\delta U''}{U^2} + \frac{U'\delta U - U\delta U'}{U^2} g'\Big(\frac{U'}{U}\Big)\,;$$

so, the condition $\delta g \ge 0$ takes the form

$$(U''\delta U - U\delta U'') + (U'\delta U - U\delta U')g'(U'/U) \ge 0\,. \tag{C.6}$$

If U is a CARA utility function, $U(x) = (1/\alpha)(1 - e^{-\alpha x})$, then $U'' = -\alpha U'$, hence $g(z) = \alpha z$ for all z, and (C.6) becomes $(-\alpha U'\delta U - U\delta U'') + \alpha(U'\delta U - U\delta U') \ge 0$, that is,

$$\delta U'' + \alpha\delta U' \le 0\,, \tag{C.7}$$

while (C.3) becomes $-\alpha U'\delta U' - U'\delta U'' \ge 0$, which is again (C.7). It means that the two orders coincide in first order variations around a CARA function. In other words, if U is a CARA function and $U + \delta U$ is some (not just CARA) utility function close to U, then (C.7) is the first order (in δU) condition both for $U + \delta U$ being more risk avers than U, and for $U + \delta U$ being more equilibrium compatible than U.

Example 3 If (U_α) is the CARA family, $U_\alpha(x) = (1/\alpha)(1 - e^{-\alpha x})$, then $U''_\alpha = -\alpha U'_\alpha$ for all α, hence $\delta U'' = -\delta(\alpha U') = -U'd\alpha - \alpha\delta U'$, so (C.7) is satisfied, which means that the CARA family is both of increasing risk aversion (it evidently is) and of increasing equilibrium compatibility (as was noted after (7.4)).

D Smoothing procedures

Start with the simple case where f s concave. Then, for a small $\varepsilon > 0$, we may take

$$f_\varepsilon(x) = \frac{1}{2\varepsilon}\int_{x-\varepsilon}^{x+\varepsilon} f(x_1)\,dx_1 - \varepsilon x^2\,.$$

Consequently, f_ε is strictly concave, continuously differentiable function, and $f_\varepsilon \to f$ for $\varepsilon \to 0$. This approach, however, does not work for relative concavity; here is one that does.

Consider, once again, a smooth, strictly concave function $f(x) = \min_t(tx + C_t)$ determines a relation between x and t by the two following equivalent formulas: $f(x) = tx + C_t$, and $f'(x) = t$. The graph of the relation is an increasing curve on the plane (x, t). If f is just concave, we still have an increasing curve on the plane (x, t), but the curve may contain horizontal or/and vertical segments. The ε-neighborhood of the increasing curve evidently contains smooth increasing curves with no horizontal or vertical segments, or even tangent lines. Such a curve is the graph of a smooth function, $g_\varepsilon(x) = t$, satisfying $g'_\varepsilon(x) > 0$ for all x. Define a function f_ε by the differential equation $f'_\varepsilon(x) = g_\varepsilon(x)$; its evident solution contains an arbitrary constant which, if appropriately chosen, we get f_ε close to f. Thus, f is approximated by a smooth, strictly convex f_ε.

The approach taken above can be generalized readily to relative concavity. A relatively concave function $F(b) = \min_t(C_t\varphi_t(b))$ determines an increasing relation between b and t by the formula $F(b) = C_t\varphi_t(b)$, see Lemma 3 (a). Its graph is an increasing curve on the plane (b, t), possibly containing horizontal or/and vertical segments. We approximate the curve by another curve $t = g_\varepsilon(b)$ with a smooth function g_ε satisfying $g'_\varepsilon(b) > 0$ for all b. Define a function F_ε by the following differential equation (suggested by (4.6)):

$$\frac{d}{db} \ln F_\varepsilon(b) = \left(\frac{d}{db} \ln \varphi_t(b) \right) \Big|_{t=g_\varepsilon(b)} . \tag{D.1}$$

The evident solution contains an arbitrary constant; choosing the constant appropriately, we get F_ε close to F. Thus, F is approximated by a smooth F_ε satisfying Condition (c) of Theorem 1. So, smoothing a relatively concave function is possible, but indirectly, by smoothing the corresponding dependence between b and t. By smoothing a strategy, we indirectly smooth out both distributions (of types and bids), The differential equation (D.1) restores a distribution of bids from an equilibrium strategy.

References

[1] V.I. Arnold, "Geometrical methods in the theory of ordinary differential equations", second edition, Springer-Verlag, New York 1988 (translated from Russian).

[2] S. Athey and P. Haile, "Idetification of standard auction models", *Econometrica*, 70 (2002), 2107-2140.

[3] T. Bedford and I. Meilijson, "A characterization of marginal distributions of (possibly dependent) lifetime variables which right censor each other", *The Annals of Statistics*, 25 (1997), 1622–1645.

[4] H.S.M. Coxeter, "The real projective plane", third edition, Springer-Verlag 1993.

[5] I. Gilboa&D.Schmeidler "A theory of case-based decisions", Cambridge University Press, 2001.

[6] E. Guerre, I. Perrigne and Q. Vuong "Optimal nonparametric estimation of first price auctions", *Econometrica*, 68 (2000), 525-574.

[7] J. Kagel&A. Roth"The Handbook of Experimental Economics", Princeton University Press, (1995).

[8] M. Landsberger and B. Tsirelson, "Augmenting and testing the theory of first price auction (extended version)", Social Science Research Network (electronic), to appear.

[9] E. Maskin and J. Riley, "Optimal auctions with risk averse buyers", *Econometrica*, 52 (1984), 1473-1518.

[10] A. Rubinstein, "Modeling bounded rationality", The MIT Press, 1998.

[11] G.R. Shorack and J.A. Wellner, "Empirical processes with applications to statistics", Wiley 1986.