Invariance of the distribution of the maximum

Fosgerau, Mogens and Lindberg, Per Olov and Mattsson, Lars-Göran and Weibull, Jörgen

Technical University of Denmark and Centre for Transport Studies, KTH Royal Institute of Technology, Sweden, Department of Transport Science and Centre for Transport Studies, KTH Royal Institute of Technology, Department of Transport Science and Centre for Transport Studies, KTH Royal Institute of Technology, Department of Economics, Stockholm School of Economics; Department of Mathematics, KTH Royal Institute of Technology; and Institute for Advanced Study in Toulouse

2015

INVARIANCE OF THE DISTRIBUTION OF THE MAXIMUM

Mogens Fosgerau, Per Olov Lindberg, Lars-Göran Mattsson and Jörgen Weibull*

June 22, 2015

Abstract. Many models in economics involve probabilistic choices where each decision-maker selects the best alternative from a finite set. Viewing the value of each alternative as a random variable, the analyst is then interested in the choice probabilities, that is, the probability for an alternative to give the maximum value. Much analytical power can be gained, both for positive and normative analysis, if the maximum value is statistically independent of which alternative obtains the highest value. This note synthesizes and generalizes previous results on this invariance property. We provide characterizations of the property within a wide class of distributions that comprises the McFadden GEV class, show implications in several directions, and establish connections with copulas. We illustrate the usefulness of the invariance property by way of a few examples.

Keywords: Choice, random utility, extreme value, leader-maximum, invariance, independence.

JEL codes: C10, C25, D01.

This note characterizes, within certain much-used function classes, those multivariate probability distributions that have a certain remarkable and useful invariance property that can be informally described as follows. Consider a population of consumers who face a finite set of alternatives, say, lunch restaurants. At each time a choice is made, the consumer chooses exactly one of the restaurants; he or she is well informed and chooses the one that she finds best. Her choice of restaurant will result in an experienced utility. Let us now compare the experienced utility distributions at the restaurants. The invariance property holds if these distributions are the same for all restaurants. In a statistical sense then, people are thus just as satisfied in one restaurant as in another. The reason why such invariance may even be a possibility is that consumers make voluntary and well-informed choices. The randomness is only in the eyes of the outside observer, who does not know every individual's preferences,

*Fosgerau: Technical University of Denmark and Centre for Transport Studies, KTH Royal Institute of Technology, Sweden. Lindberg and Mattsson: Department of Transport Science and Centre for Transport Studies, KTH Royal Institute of Technology. Weibull: Department of Economics, Stockholm School of Economics; Department of Mathematics, KTH Royal Institute of Technology; and Institute for Advanced Study in Toulouse. This research has been supported by the Danish Strategic Research Council and the Knut and Alice Wallenberg Research Foundation.

only the population distribution of preferences. Imagine that one restaurant improves its food quality or services and/or reduces its prices, and assume that all consumers are informed of this change. Some consumers will then switch to this restaurant from the one they used to go to. The ones who are most eager to do this are those who were least happy at their original restaurant, which they before found to be the best, but no longer so. Since the least happy left, average experienced utility of those who remain in their usual restaurants goes up. If the invariance property holds, the new utility distributions, among clients in the restaurants after one improved its services, will again be identical across restaurants, but now at a higher level.

The invariance property leads to analytical simplification in many applications including discrete choice with an outside option, rent seeking, innovation contests, patent races, and auctions, see Section 3. First, however, Section 1 provides general definitions and some preliminaries. Our main results are presented in Section 2. Section 4 discusses a number of earlier contributions.

1. Definitions and preliminaries

Let $\mathbb{N}$ be the positive integers, $\mathbb{R}$ be the reals and $\mathbb{R}_+$ the non-negative reals. Denote by $\mathcal{F}$ the class of cumulative distribution functions such that $F : \mathbb{R}^n \to [0,1]$ for some integer $n > 1$ and some nonempty and closed (bounded or unbounded) interval $D \subseteq \mathbb{R}$ such that F is twice continuously differentiable on $D^n \subseteq \mathbb{R}^n$ and has positive density f with support D^n. For convenience we will subsequently refer to D^n as "the support". Let $\mathbf{X} = (X_1, ..., X_n)$ be a random vector distributed according to some such F. (We write vectors in bold-face.) Let $\hat{X} = \max_j X_j$ be the maximum of the random vector and let $\hat{F}$ denote its c.d.f.. We write F_j for the partial derivative of F with respect to its j^{th} variable and $F^{(j)}$ for the j^{th} marginal distribution of the multivariate distribution F. Define the selection $\xi \in J = \{1, .., n\}$ by $\xi = \arg\max_j X_j$, where the latter set with probability one is a singleton. Let $p_j = P(\xi = j)$. One may also consider the distribution of the maximum conditional on the selection of a particular alternative $j \in J$: $\hat{F}^{(j)}(t) = P\left(\hat{X} \le t \mid \xi = j\right)$. In a discrete-choice setting, $\hat{F}^{(j)}$ is the distribution of achieved (or experienced) utility, conditional on the choice of alternative $j \in J$, and p_j is its choice probability.

It is relatively straight-forward in this setting to prove the following three equalities:[1]

$$\begin{cases} p_j = \int F_j(s, ..., s)\, ds \quad \forall j \in J \\ \hat{F}^{(j)}(t) = p_j^{-1} \cdot \int_{-\infty}^{t} F_j(s, ..., s)\, ds \quad \forall j \in J, t \in \mathbb{R} \\ \hat{F}(t) = F(t, ..., t) \quad \forall t \in \mathbb{R}. \end{cases} \tag{1}$$

We note that $p_j > 0$ for all $j \in J$, and that the quantities in (1) only depend on how the cumulative distribution function (c.d.f.) $F \in \mathcal{F}$ behaves near the diagonal of its domain. By "invariance" we mean that the conditional distributions $\hat{F}^{(j)}$ are identical across alternatives $j \in J$:

[1] See e.g. Lindberg, Eriksson and Mattsson (1995, Lemma 1). The general case is more complex, see Lindberg (2012).

Definition 1. *A multivariate distribution F has the* **Invariance Property** *if $\hat{F}^{(j)} = \hat{F}$ for all $j \in J$. A random vector $\mathbf{X}$ has the Invariance Property if its distribution has the Invariance Property.*

The following proposition, stated without proof, provides some immediate conclusions regarding the Invariance Property. We use the notational convention that a univariate function applied to a vector is applied to each component of the vector.

Proposition 1. *Consider a random vector $\mathbf{X}$ with c.d.f. $F \in \mathcal{F}$.*
(i) If X_j are i.i.d. then $\mathbf{X}$ has the Invariance Property.
(ii) If $\mathbf{X}$ has the Invariance Property and $T : \mathbb{R} \to \mathbb{R}$ is a strictly increasing function, then $T(\mathbf{X})$ has the Invariance Property.
(iii) $\mathbf{X}$ has the Invariance Property if and only if the selection ξ and the maximum $\hat{X}$ are independent.

We will express results in terms of so-called copulas. These are functions on the unit cubes in Euclidean spaces, defined as follows (see Nelsen, 2006, for an excellent introduction):

Definition 2. *A copula is any function $C : [0,1]^n \to [0,1]$ such that*
(i) $C(\mathbf{x}) = 0$ if $\Pi_{j\in J} x_j = 0$,
(ii) $C(\mathbf{x}) = x_k$ if $\Pi_{j\in J\setminus\{k\}} x_j = 1$,
(iii) If $\mathbf{x}, \mathbf{y} \in [0,1]^n$ and $\mathbf{x} \le \mathbf{y}$, then $V_C([\mathbf{x}, \mathbf{y}]) \ge 0$.

Here $[\mathbf{x}, \mathbf{y}]$ denotes the *box* $\times_{j=1}^{n} [x_j, y_j] \subseteq [0,1]^n$ and $V_C([\mathbf{x}, \mathbf{y}])$ is the *C-volume* of this box, defined as the signed sum of the values $C(\mathbf{v})$ at all vertices $\mathbf{v}$ of $[\mathbf{x}, \mathbf{y}]$, where the sign is positive (negative) if $v_j = x_j$ for an even (odd) number of coordinates $j \in J$. Condition (iii) ensures that the copula assigns non-negative probability mass to any box. By construction, copulas are then c.d.f.s on the unit cube that have uniform marginal distributions.

By Sklar's theorem (e.g. Theorem 2.10.9 in Nelsen, 2006), every multivariate distribution $F : \mathbb{R}^n \to [0,1]$ can be written in terms of its marginal distributions $F^{(j)}$ and a copula C, so that

$$F(\mathbf{x}) = C\left(F^{(1)}(x_1), ..., F^{(n)}(x_n)\right) \quad \forall \mathbf{x} \in \mathbb{R}^n. \tag{2}$$

The copula associated with any c.d.f. thus captures the statistical dependence structure of the multivariate distribution in question.

In order to state our main result we define the class of copulas that are associated with multivariate extreme-value (MEV) distributions, to be called *MEV copulas*. A *multivariate extreme-value (MEV) distribution* is any multivariate distribution H with non-degenerate margins for which there exist an i.i.d. sequence of random vectors, $(\mathbf{X}_m)_{m\in\mathbb{N}}$, all distributed according to some multivariate c.d.f. F, and normalizing vectors $\mathbf{a}_m, \mathbf{b}_m \in \mathbb{R}^n$ for each $m \in \mathbb{N}$ such that all components of all vectors $\mathbf{a}_m$ are positive and, with component-wise maximization, multiplication and division,

$$\lim_{m\to\infty} P\left(\frac{\max_{1\leq l\leq m}\mathbf{X}_l - \mathbf{b}_m}{\mathbf{a}_m} \leq \mathbf{x}\right) = \lim_{m\to\infty}\left[F\left(\mathbf{a}_m\mathbf{x}+\mathbf{b}_m\right)\right]^m = H\left(\mathbf{x}\right)$$

(see Joe, 1997).[2] MEV copulas can be shown to be exactly those copulas that satisfy the following homogeneity property

$$C\left(x_1^\alpha, ..., x_n^\alpha\right) = \left[C\left(x_1, ..., x_n\right)\right]^\alpha \ \forall\mathbf{x}\in(0,1)^n, \ \forall\alpha>0.$$

This follows from

Lemma 1. *A copula C is a copula of an MEV distribution if and only if it is of the form*

$$C\left(\mathbf{x}\right) = \exp\left(-G\left(-\ln\mathbf{x}\right)\right) \tag{3}$$

for some linearly homogenous function G.

(All proofs are given in the Appendix.) We note that if $\mathbf{X} = (X_1, ..., X_n)$ is MEV distributed, then so is any subvector of X. Hence, the copulas of the subvectors also have the form (3).

Lemma 2. *For $I \subset J$, let $\mathbf{X}_I = (X_i)_{i\in I}$ and similarly for $\mathbf{x}_I$. Let $G^{(I)}(\mathbf{x}_I) = \lim_{x_i\to+\infty\ \forall i\notin I} G(\mathbf{x})$. If $\mathbf{X}$ is MEV-distributed with c.d.f. of the form (2) with copula C of the form (3), then $\mathbf{X}_I$ has the c.d.f.*

$$F^{(I)}\left(\mathbf{x}_I\right) = \exp\left(-G^{(I)}\left(-\ln F^{(i)}\left(x_i\right)\right)_{i\in I}\right).$$

2. Results

The main result of this note is the following theorem, which generalizes previous results, see Mattsson, Weibull and Lindberg (2014) and Section 4.

Theorem 1. *Consider any $F \in \mathcal{F}$ such that $F(\mathbf{x}) = C\left(F^{(1)}(x_1), ..., F^{(n)}(x_n)\right) \forall\mathbf{x}\in\mathbb{R}^n$ for some* MEV *copula C that is twice continuously differentiable on the unit cube and has positive partial derivatives. Then F and all the multivariate marginal distributions of F have the Invariance Property if and only if for each $j \in J$ there exists an $\alpha_j > 0$ such that $F^{(j)} = \left[F^{(1)}\right]^{\alpha_j}$.*

[2] More precisely, if $\mathbf{X}_m = (X_{m1}, .., X_{mn})$, then $\max_{1\leq l\leq m}\mathbf{X}_m$ is the vector with j^{th} component $\max_{1\leq l\leq m} X_{lj}$, and $\mathbf{a}_m\mathbf{x}$ is the vector with j^{th} component $a_{mj}x_j$.

The condition that the marginal distributions are positive powers of each other has special significance when the marginal distribution $F^{(1)}$ is standard extreme-value type 1 (Gumbel). Then $F^{(j)}(t) = \exp(-\exp[-t + \ln \alpha_j])$, meaning that the marginal distributions are identical up to a location shift. An MEV distribution with type 1 margins is the basis of the well-known logit family of models, pioneered by McFadden (1974, 1978, 1981).

Theorem 1 characterizes the Invariance Property for those distributions in $\mathcal{F}$ that have twice continuously differentiable MEV copulas. Could it be that invariance of a probability distribution is *equivalent* to the distribution having an MEV copula and marginal distributions that are powers of each other? That conjecture is false because, among other things, invariance depends only on the properties of the distribution on the diagonal $L = \{\mathbf{x} \in D^n : x_1 = ... = x_n\}$ of its support:

Proposition 2. *Consider any distributions $F, \tilde{F} \in \mathcal{F}$ with the same support D^n. If F has the Invariance Property and $\tilde{F} = F$ on an open neighborhood of L, then also $\tilde{F}$ has the Invariance Property.*

The next proposition shows that the Invariance Property is preserved under aggregation of components to blocks represented by their maximal member.

Proposition 3. *Suppose that $\mathbf{X} = (X_1, ..., X_n)$ has the Invariance Property, and consider any partition of J into k subsets J_i. For each subset J_i let $Y_i = \max_{j \in J_i}(X_j)$. Then $\mathbf{Y} = (Y_1, ..., Y_k)$ has the Invariance Property.*

3. Applications

This section discusses a number of economics contexts in which the Invariance Property is useful. Throughout we rely on the equivalence of the Invariance Property to the independence of the selection ξ and the maximum $\hat{X}$.

3.1. Discrete choice with outside option. The probabilistic basis of a discrete choice model is a random vector $\mathbf{X}$ of (indirect) utilities, each associated with an alternative $j \in J$. The decision-maker, or, as we will here say, the consumer, chooses the alternative with the highest utility. Such a discrete choice model is also embedded in models of monopolistic competition (Sattinger, 1984; Perloff and Salop, 1985). In many applications, consumers also have an outside option. The utility associated with the outside option acts as a threshold such that the consumer only chooses one of the alternatives in J if the utility of that alternative exceeds that of the outside option. We shall see that such situations are easily treated when the random utilities have the Invariance Property. This seems not to have been observed before.

Suppose then that $\mathbf{X}$ has the Invariance Property, and that alternative j is chosen if $X_j = \hat{X}$ and $\hat{X} > X_0$, where X_0 is the random utility of the outside option, which we take to be statistically independent from $\mathbf{X}$. Let F_0 be the c.d.f. of X_0. We still use ξ to denote the alternative in $J = \{1, ..., n\}$ with maximum utility and p_j to denote the probability that $\xi = j$. The outside option is labelled $j = 0$, and

$\eta \in \{0\} \cup J$ denotes the chosen option among all options. The probability for the outside option is

$$P(\eta = 0) = P\left(\hat{X} \leq X_0\right) = \int F(s, .., s)\, dF_0(s),$$

a probability that can be calculated from the primitives of the model.

It follows from Proposition 1 (iii) that, for any $j \in J$ and $s \in \mathbb{R}$, $P\left(\xi = j \mid \hat{X} > s\right) = p_j$ and thus

$$\begin{aligned} P(\eta = j) &= P\left(\xi = j \wedge \hat{X} > X_0\right) = \int P\left(\xi = j \mid \hat{X} > s\right) \cdot P\left(\hat{X} > s\right) dF_0(s) \\ &= p_j \cdot \left[1 - P(\eta = 0)\right]. \end{aligned}$$

Hence, the presence of an outside option does not affect the probabilities of the inside options, conditional on the outside option not being chosen.

3.2. Rent-seeking, innovation contests and patent races. Baye and Hoppe (2003) establish the strategic equivalence between wide classes of rent-seeking games, innovation contests and patent-races. In their innovation-contest game, n firms compete by employing a finite and positive number of scientists, where each scientist costs $c > 0$ and independently produces an innovation of a random value with c.d.f. H on the unit interval. All firms pay the costs of their scientists, and the firm with the best idea among all firms wins the value of its best idea. The other firms win nothing.

Let X_j be the value of firm j's best idea and let $\mathbf{X} = (X_1, .., X_n)$. Then $\mathbf{X}$ has the joint c.d.f. $F(x_1, .., x_n) = \prod_{j=1}^{n} F^{(j)}(x_j)$, where $F^{(j)}(x_j) = [H(x_j)]^{m_j}$ and m_j is the number of scientists at firm j. Since $F(\mathbf{x}) = \exp\left(-\sum_{j=1}^{n} - \ln F^{(j)}(x_j)\right)$, F has an MEV copula (by Lemma 1), so Theorem 1 applies. Hence, F has the Invariance Property. By Proposition 1 (iii), the identity $\xi \in J$ of the firm that wins the contest and the value $\hat{X}$ of the best idea are statistically independent. Hence, the expected profit to firm j from hiring $m_j \in \mathbb{N}$ scientists can be expressed as

$$\pi_j(m_1, .., m_n) = P(\xi = j) \cdot E(\hat{X}) - c \cdot m_j$$

where, since all scientists have the same probability of producing the best innovation, $P(\xi = j) = m_j / M$ for $M = m_1 + .. + m_n$, and

$$E(\hat{X}) = \int_0^1 P(\hat{X} > s) ds = 1 - \int_0^1 \hat{F}(s) ds = 1 - \int_0^1 [H(s)]^M ds.$$

Hence,

$$\pi_j(m_1, .., m_n) = \left[\frac{1}{M}\left(1 - \int_0^1 [H(s)]^M ds\right) - c\right] \cdot m_j$$

This setting can be generalized in several ways within our framework. The R&D inputs m_j may be any positive real numbers, and the c.d.f. H may be defined on any interval. Moreover, the marginal c.d.f. $F^{(j)}$ of the value of the best idea within any firm j may be of the power form $H^{\varphi_j(m_j)}$ for any positive increasing function φ_j of its R&D input m_j. For example, an S-shaped such function could represent the "synergy" or "critical mass" effect that the marginal return (in terms of innovations or discoveries) from an additional scientist may be highest at some intermediate size of the research unit. Moreover, by using MEV copulas, one may allow for statistical dependence among the values of the best ideas in the different firms. It follows from Lemma 1 that such an MEV copula is defined by a linearly homogenous function G, and one obtains

$$P(\xi = j) = \frac{\varphi_j(m_j) \cdot G_j(\varphi_1(m_1), .., \varphi_n(m_n))}{G(\varphi_1(m_1), .., \varphi_n(m_n))}$$

from Theorem 1 in Mattsson, Weibull and Lindberg (2014), and $E(\hat{X})$ can be derived from the c.d.f. $\hat{F} = H^{G(\varphi_1(m_1),..,\varphi_n(m_n))}$ of $\hat{X}$. We conjecture that the results in Baye and Hoppe (2003) can be generalized, by use of the invariance property, in many important directions for a wide variety of innovation contests, patent races and rent-seeking games.

3.3. Auctions. Following Milgrom and Weber (1982) and Krishna (2002) we briefly consider auctions for an indivisible item in a situation where each bidder $j \in J$ has private information about the item for sale in the form of a random signal X_j. The signal vector $\mathbf{X} = (X_1, ..., X_n)$ has positive dependence among the components, formalized as affiliation (see Krishna, 2002). More precisely, assume that all signals take values in some interval $D \subseteq \mathbb{R}_+$, and let $F : D^n \to [0,1]$ be the joint c.d.f. of the signal vector $\mathbf{X}$ with density $f : D^n \to \mathbb{R}_+$. The signals are (positively) *affiliated* (Milgrom and Weber, 1982) if

$$f(\mathbf{x}) \cdot f(\mathbf{y}) \leq f(\mathbf{x} \vee \mathbf{y}) \cdot f(\mathbf{x} \wedge \mathbf{y})$$

for all $\mathbf{x}, \mathbf{y} \in D^n$.[3] We now add the assumption that the distribution F of $\mathbf{X}$ takes the form $F(\mathbf{x}) = C(\Phi(x_1)^{\alpha_1}, ..., \Phi(x_n)^{\alpha_n})$ for some MEV copula C, univariate c.d.f. Φ, and constants $\alpha_j > 0, j \in J$. To the best of our knowledge, this assumption is new to the auction literature.[4] We proceed to show that it is useful.

A *strategy* for bidder j is any measurable function $\beta_j : D \to \mathbb{R}_+$ that maps bidder j's private signal X_j to a bid $\beta_j(X_j)$. We proceed to calculate the distribution and expectation of the highest and second-highest bids for the case when all bidders use

[3] This property is known in the statistics literature as multivariate total positivity (MTP_2) of the density, see e.g. Karlin and Rinott (1980). See also Krishna (2002).

[4] Joe (1997) collects some results showing the relationship between affiliation and MEV copula. If a density is MTP_2 then all bivariate margins of the c.d.f. are max-infinitely divisible (Thm. 2.3, 2.5 and 2.6). Any MEV copula is max-infinitely divisible. Thus affiliation and MEV copula are related concepts.

the same strictly increasing strategy β.[5] The bidder with the highest signal, $\hat{X}$, then makes the highest bid, $\beta(\hat{X})$, and wins the auction. The distribution of the highest bid is easy to obtain. Writing $\Psi(t)$ for $\Phi\left[\beta^{-1}(t)\right]$, c for $C\left(e^{-\alpha_1},...,e^{-\alpha_n}\right)$, and recalling that C is an MEV copula, we find that the distribution of the maximum bid is

$$\begin{aligned} P\left[\beta\left(\hat{X}\right)\leq t\right] &= C\left(\Psi(t)^{\alpha_1},...,\Psi(t)^{\alpha_n}\right)=C\left(e^{-\alpha_1(-\ln\Psi(t))},...,e^{-\alpha_n(-\ln\Psi(t))}\right)\\ &= c^{-\ln\Psi(t)}=\exp\left[-(\ln c)\cdot\ln\Psi(t)\right]=\Psi(t)^{-\ln c}. \end{aligned}$$

The inverse of the c.d.f. of the highest bid $\beta\left(\hat{X}\right)$ is the mapping $s\rightarrow\Psi^{-1}\left(s^{-\frac{1}{\ln c}}\right)$. Hence, the expected maximum bid is[6]

$$E\left[\beta\left(\hat{X}\right)\right]=\int_0^1\Psi^{-1}\left(s^{-\frac{1}{\ln c}}\right)ds=\int_0^1\beta\left(\Phi^{-1}\left(s^{-\frac{1}{\ln c}}\right)\right)ds.$$

Writing c_j for $C\left(e^{-\alpha_1},..,1,..,e^{-\alpha_n}\right)$, where $e^{-\alpha_j}$ in position j is replaced by 1, it is also straightforward to derive the distribution of the second highest bid, $\hat{X}_{(2)}$ (the actual payment in a second-price auction):

$$\begin{aligned} P\left[\beta\left(\hat{X}_{(2)}\right)\leq t\right] &= P\left[\hat{X}\leq\beta^{-1}(t)\right]+\\ &\quad+\sum_{j=1}^{n}\left[C\left(\Psi(t)^{\alpha_1},..,1,..,\Psi(t)^{\alpha_n}\right)-C\left(\Psi(t)^{\alpha_1},...,\Psi(t)^{\alpha_n}\right)\right]\\ &= \sum_{j=1}^{n}\Psi(t)^{-\ln c_j}-(n-1)\Psi(t)^{-\ln c}. \end{aligned}$$

The expectation of the second highest bid is then

$$\begin{aligned} E\left[\beta\left(\hat{X}_{(2)}\right)\right] &= \int_0^\infty\left(1-\sum_{j=1}^{n}\Psi(t)^{-\ln c_j}+(n-1)\Psi(t)^{-\ln c}\right)dt\\ &= \sum_{j=1}^{n}\int_0^\infty\left[1-\Psi(t)^{-\ln c_j}\right]dt\ -\ (n-1)\int_0^\infty\left[1-\Psi(t)^{-\ln c}\right]dt\\ &= \sum_{j=1}^{n}\int_0^1\beta\left(\Phi^{-1}\left(s^{-\frac{1}{\ln c_j}}\right)\right)ds-(n-1)E\left[\beta\left(\hat{X}\right)\right]. \end{aligned}$$

We thus have explicit expressions for the distribution and expectation of the highest and second highest bids. By the Invariance Property, the probability distributions of

[5] Such monotonicity is known to hold in symmetric equilibria in sealed-bid first-price and second-price auctions, see e.g. Krishna (2002).

[6] These equations follow from the fact that the expectation of any positive random variable X with c.d.f. F can be written as the integral of $1-F(x)$ over $x\in\mathbb{R}_+$, or, when F has an inverse, of $F^{-1}(y)$ over $y\in[0,1]$.

the highest and second-highest bids are independent of which bidder actually won. We finally note that these derivations were made without any symmetry assumption concerning the distribution of the signal vector X. In particular, asymmetric statistical dependency among bidders' signals is permitted.

4. Previous literature

Researchers in economics and in probability theory have returned to the topic of invariance a number of times. This section presents a summary and highlights how our results are related to earlier results.

Strauss (1979, Theorem 5) claimed that if the components of a random vector $\mathbf{X}$ are of the additive form $X_j = \nu_j + Z_j$, where the ν_j:s are scalars and $\mathbf{Z}$ is a random vector with c.d.f. F, then the Invariance Property holds for all vectors $\boldsymbol{\nu}$ if and only if $F(\mathbf{z}) = \phi\left[G\left(e^{-\mathbf{z}}\right)\right]$ for some positive homogenous function G (with positive degree) and some function $\phi : \mathbb{R}_+ \to [0,1]$ such that F is a c.d.f.[7] It is straightforward to show that the subset of such functions G that are linearly homogenous corresponds exactly to the set of G functions that generate MEV copulas (see Lemma 1). Strauss (1979, Theorem 4) further showed that the multinomial logit model is the only additive random utility model with independent error terms Z_j for which the Invariance Property holds. Anas and Feng (1988) later showed that the multinomial logit model has the weaker Invariance Property $E[\hat{X}_j] = E[\hat{X}]$ for all $j \in J$. This follows, however, directly from Strauss (1979).

de Palma and Kilani (2007) also considered the additive random utility case with independent error terms and reproved the above result by Strauss (1979), which follows from his Theorem 4. They also proved the claim that for error terms that are i.i.d. with finite expectation, the above weaker Invariance Property, $E[\hat{X}_j] = E[\hat{X}]$ for all $j \in J$, is equivalent to the error terms being extreme-value type 1 (Gumbel) distributed. The if-part is the same as the previous claim by Anas and Feng (1988).

Also Train and Wilson (2008) re-derived the Invariance Property for multinomial logit models. They used this property for maximum likelihood estimation in a class of combined stated- and revealed-preference experiments, where each respondent considered the same alternatives in both experiments and where the attributes of the alternatives in the stated-preference experiments were varied on the basis of the respondent's revealed-preference choice.

Our Theorem 1 states that, for any multivariate distribution with MEV copula, the Invariance Property holds for all multivariate marginal distributions if and only if the univariate marginal distributions are positive powers of each other. A special case of this result was given by Resnick and Roy (1990b) who showed that MEV distributions with Gumbel marginals have the Invariance Property. Mattsson, Weibull and Lindberg (2014) generalized this to the "if" claim in the present Theorem 1. Resnick and Roy (1990a) proved a special case of the "only if" claim in our Theorem 1,

[7] An incomplete proof of this claim was given in Robertson and Strauss (1981). Lindberg, Eriksson and Mattsson (1995) completed the proof.

namely that the Invariance Property for two independent non-negative random variables implies that their c.d.f.s are positive powers of each other. Mattsson, Weibull and Lindberg (2014) generalized this to an arbitrary finite number of independent random variables.

5. Appendix

Lemma 1 is essentially in Joe (1997) but his proof is incomplete.

Proof of Lemma 1. Necessity: Note with Joe (1997, p. 173-174) that an MEV copula C must satisfy $C(\mathbf{x}^{\alpha}) = C(\mathbf{x})^{\alpha}$ for all $\mathbf{x} \in (0,1)^n$ and $\alpha > 0$. Then $G(\mathbf{y}) \equiv -\ln C(e^{-\mathbf{y}})$ is linearly homogenous.

Sufficiency: Suppose C has the form (**2**). Let F be a c.d.f. with copula C and marginals that are unit exponential on the negative half-axis, i.e. $F_i(x) = e^x$ on $(-\infty, 0]$. Then $F(\mathbf{x}) = C(e^{\mathbf{x}})$. Let $\mathbf{X}_j = (X_{j1}, ..., X_{jn})$ be i.i.d. with c.d.f. F, and let $M_i^k = \max_{j\leq k} X_{ji}$. Let $\mathbf{Y}^k = \left(kM_1^k, ..., kM_n^k\right)$, then $\mathbf{Y}^k$ has the c.d.f.

$$\begin{aligned} P\left(\mathbf{Y}^k \leq \mathbf{x}\right) &= P\left[\left(kM_1^k, ..., kM_n^k\right) \leq (x_1, ..., x_n)\right] \\ &= P\left[\left(\max_{j\leq k} X_{j1}, ..., \max_{j\leq k} X_{jn}\right) \leq \left(\frac{x_1}{k}, ..., \frac{x_n}{k}\right)\right] \\ &= \prod_{j\leq k} P\left[(X_{j1}, ..., X_{jn}) \leq \left(\frac{x_1}{k}, ..., \frac{x_n}{k}\right)\right] \\ &= C\left(e^{x_1/k}, ..., e^{x_n/k}\right)^k = C\left(e^{\mathbf{x}}\right). \end{aligned}$$

Thus all $\mathbf{Y}^k$ have the same distribution, and hence they converge in distribution to $\mathbf{Y}^1$, say, with the same c.d.f. $C(e^{\mathbf{x}})$. Therefore $C(e^{\mathbf{x}})$ is the c.d.f. of an MEV distribution, and C is an MEV copula. ∎

Proof of Theorem 1. Sufficiency: This follows from Theorem 1 in Mattsson, Weibull and Lindberg (2014), combined with the observation in Lemma 2 that the multivariate marginal distributions of MEV copulas are also MEV copulas, and that twice differentiability with positive partials is inherited.

Necessity: Consider the c.d.f. $P(X_1 \leq x_1, X_2 \leq x_2) = F(x_1, x_2, \infty, ..., \infty) = C\left(F^{(1)}(x_1), F^{(2)}(x_2), 1, ..., 1\right)$. By assumption, this c.d.f. has the Invariance Property. As noted in Lemma 2, the copula for (X_1, X_2) inherits the MEV property from C. We may thus ignore the last but two dimensions of F and C, and assume that $|J| = 2$ at no loss of generality. It remains to show that $F^{(2)} = \left[F^{(1)}\right]^{\alpha}$ for some $\alpha > 0$.

First, by Lemma 1, we may write

$$F(\mathbf{x}) = \exp\left(-G\left(-\ln F^{(1)}(x_1), -\ln F^{(2)}(x_2)\right)\right), \tag{4}$$

where G is linearly homogenous and satisfies the properties necessary for F to be a c.d.f. Since C by hypothesis is twice continuously differentiable with positive partials,

G inherits these properties (except at the origin). To see this, note that $G(y) \equiv -\ln C\left(e^{-y_1}, e^{-y_2}\right)$.

Second, by Proposition 1, the Invariance Property remains under any strictly increasing transformation of the components. Hence, it is no loss of generality to apply such a transformation so that $F^{(1)}$ is a Gumbel distribution, $F^{(1)}(x_1) = \exp\left(-e^{-x_1}\right)$ on $\mathbb{R}$. By (1) we then have, for all $t \in \mathbb{R}$,

$$\begin{aligned}
\hat{F}^{(1)}(t) &= \frac{1}{p_1}\int_{-\infty}^{t} F(s,s)\, G_1\left(e^{-s}, -\ln F^{(2)}(s)\right) e^{-s} ds = \\
\hat{F}^{(2)}(t) &= \frac{1}{p_2}\int_{-\infty}^{t} F(s,s)\, G_2\left(e^{-s}, -\ln F^{(2)}(s)\right) \frac{f^{(2)}(s)}{F^{(2)}(s)} ds,
\end{aligned}$$

where $f^{(2)} > 0$ is the density of $F^{(2)}$ and p_1 and p_2 are the associated choice probabilities. Differentiation with respect to t gives

$$\frac{1}{p_1} F(t,t)\, G_1\left(e^{-t}, -\ln F^{(2)}(t)\right) \cdot e^{-t} = \frac{1}{p_2} F(t,t)\, G_2\left(e^{-t}, -\ln F^{(2)}(t)\right) \cdot \frac{f^{(2)}(t)}{F^{(2)}(t)}.$$

or (since $F(t,t) > 0$ and $G_i(\mathbf{x}) > 0$ for all positive $\mathbf{x} \in \mathbb{R}^2$):

$$f^{(2)}(t) = \frac{p_2 e^{-t} G_1\left(e^{-t}, -\ln F^{(2)}(t)\right)}{p_1 G_2\left(e^{-t}, -\ln F^{(2)}(t)\right)} \cdot F^{(2)}(t) \;\; \forall t \in \mathbb{R}. \tag{5}$$

Since F with $F^{(1)}(x_1) \equiv \exp\left(-e^{-x_1}\right)$ and $F^{(2)} = \left[F^{(1)}\right]^{\alpha}$ for any $\alpha > 0$ has the Invariance Property (by the established sufficiency claim of this theorem), $F^{(2)}$ satisfies this equation, where p_1 and $p_2 = 1 - p_1$ are the associated choice probabilities. Suppose that $\bar{F}^{(2)}$ is a solution to (5) for $p_1 = \bar{p}_1$ and $p_2 = \bar{p}_2$, where (i) $\bar{p}_1$ and $\bar{p}_2 = 1 - \bar{p}_1$ are the choice probabilities associated with $\bar{F}$, and (ii) $\bar{F}^{(2)} \neq \left[F^{(1)}\right]^{\alpha}$ for all $\alpha > 0$. We will show that no such solution $\bar{F}^{(2)}$ exists. For this purpose we first show that, for any $\bar{p}_1 \in (0,1)$ there exists some $\alpha > 0$ such that $\bar{p}_1 = p_1(\alpha)$, where $p_1(\alpha)$ and $p_2(\alpha) = 1 - p_1(\alpha)$ are the choice probabilities under the c.d.f. F with $F^{(1)}(x_1) \equiv \exp\left(-e^{-x_1}\right)$ and $F^{(2)} = \left[F^{(1)}\right]^{\alpha}$. This can be established as follows. By Theorem 1 in McFadden (1978),

$$p_1(\alpha) = \frac{G_1(1,\alpha)}{G(1,\alpha)} \quad \text{and} \quad p_2(\alpha) = \frac{\alpha G_2(1,\alpha)}{G(1,\alpha)}$$

(set $V_1 = 0$ and $V_2 = \ln\alpha$ in McFadden's equation (12)). By homogeneity of each G_i one obtains

$$\frac{1 - p_1(\alpha)}{p_1(\alpha)} = \alpha \cdot \frac{G_2\left(1/(1+\alpha), \alpha/(1+\alpha)\right)}{G_1\left(1/(1+\alpha), \alpha/(1+\alpha)\right)}$$

Moreover, since each G_i is continuous and positive on $\Delta = \left\{\mathbf{y} \in \mathbb{R}^2_+ : y_1 + y_2 = 1\right\}$ also the ratio G_2/G_1 is continuous and positive on Δ. Since Δ is compact, there exists

$a, b > 0$ such that $a \leq G_2(\mathbf{y})/G_1(\mathbf{y}) \leq b$ for all $\mathbf{y} \in \Delta$. Thus, for any $\bar{p}_1 \in (0,1)$ there exist an $\bar{\alpha} > 0$ such that $p_1(\bar{\alpha}) = \bar{p}_1$.

According to the Picard-Lindelöf Theorem (see e.g. Theorem 3.1 in Hale (1969)), an ordinary differential equation such as (5) has a unique (local) solution through any given point $(t_0, x_0) \in \mathbb{R} \times (0,1)$, for $x_0 = F^{(2)}(t_0)$ if

$$\Psi(t, x) = \frac{p_2 e^{-t} G_1(e^{-t}, -\ln x)}{p_1 G_2(e^{-t}, -\ln x)} \cdot x$$

defines a continuous function on $\mathbb{R} \times (0,1)$ that is locally Lipschitz-continuous in x. Our function Ψ is continuously differentiable in x and thus also locally Lipschitz continuous in x. By uniqueness, the solution $F^{(2)} = \left[F^{(1)}\right]^{\bar{\alpha}}$ and the hypothesized alternative solution, $\bar{F}^{(2)}$ cannot intersect at any point (where their values are in $(0,1)$). Hence, $\bar{F}^{(2)}(t)$ either lies above or below our solution for all $t \in \mathbb{R}$. Assume it always lies above; $F^{(2)} < \bar{F}^{(2)}$ on $\mathbb{R}$ (the opposite case can be treated in the same way). Define the random vector $\mathbf{Z} = (Z_1, Z_2)$ such that its c.d.f. is the above copula C. Let $F^{-(i)}$ be the inverse of $F^{(i)}$, for $i = 1, 2$, and let $\bar{F}^{-(2)}$ be the inverse of $\bar{F}^{(2)}$.[8] Clearly $\bar{F}^{-(2)} < F^{-(2)}$ on $(0,1)$.

Let the random vector $\mathbf{Y} = (Y_1, Y_2) = (F^{-(1)}(Z_1), F^{-(2)}(Z_2))$. Then $\mathbf{Y}$ has the c.d.f.

$$F_{\mathbf{Y}}(\mathbf{y}) = P\left[Z_1 \leq F^{(1)}(y_1) \wedge Z_2 \leq F^{(2)}(y_2)\right] = C(F^{(1)}(y_1), F^{(2)}(y_2)).$$

Then $p_2(\bar{\alpha}) = P(Y_2 \geq Y_1) = P\left[F^{-(2)}(Z_2) \geq F^{-(1)}(Z_1)\right] = \bar{p}_2$ by the choice of $\bar{\alpha}$. Similarly, the random vector $\bar{\mathbf{Y}} = (Y_1, \bar{Y}_2) = (F^{-(1)}(Z_1), \bar{F}^{-(2)}(Z_2))$ has the c.d.f. $\bar{F}(y_1, y_2) = C(F^{(1)}(y_1), \bar{F}^{(2)}(y_2))$. But since $\bar{F}^{-(2)} < F^{-(2)}$, and $\mathbf{Z}$ has positive density everywhere,

$$p_2(\bar{\alpha}) = P\left[F^{-(2)}(Z_2) \geq F^{-(1)}(Z_1)\right] > P\left[\bar{F}^{-(2)}(Z_2) \geq F^{-(1)}(Z_1)\right] = P(\bar{Y}_2 \geq Y_1) = \bar{p}_2,$$

a contradiction. ∎

Proof of Proposition 3. Let $\mathbf{X}$ have the invariance property, i.e., $\hat{F}^{(j)} = \hat{F}$ for all j. It is sufficient to establish the proposition for the case when two alternatives, say 1 and 2, are merged via the maximum operation. Define $Y = X_1 \vee X_2$. It is sufficient to show that $P(Y \leq t \mid \xi \in \{1,2\}) = \hat{F}(t)$. But

$$\begin{aligned} P(Y \leq t \mid \xi \in \{1,2\}) &= \frac{P([Y \leq t] \wedge [\xi \in \{1,2\}])}{P(\xi \in \{1,2\})} \\ &= \frac{P([Y \leq t] \wedge [\xi = 1]) + P([Y \leq t] \wedge [\xi = 2])}{p_1 + p_2} \\ &= \frac{P(Y \leq t \mid \xi = 1) P(\xi = 1) + P(Y \leq t \mid \xi = 2) P(\xi = 2)}{p_1 + p_2} \\ &= \frac{p_1}{p_1 + p_2} \hat{F}^1(t) + \frac{p_2}{p_1 + p_2} \hat{F}^2(t) = \hat{F}(t). \ \blacksquare \end{aligned}$$

[8] These inverse functions are well-defined since F has positive density on $\mathbb{R}^2$.

References

[1] Anas, A., and C.M. Feng (1988): "Invariance of expected utility in logit models", *Economics Letters* 27 (1), 41-45.

[2] Baye, M.R. and H.C. Hoppe (2003): "The strategic equivalence of rent-seeking, innovation, and patent-race games", *Games and Economic Behavior* 44 (2), 217-226.

[3] de Palma, A., and K. Kilani (2007): "Invariance of conditional maximum utility", *Journal of Economic Theory* 132 (1), 137-146.

[4] Hale, J.K. (1969): *Ordinary Differential Equations*. New York: Wiley and Sons.

[5] Joe, H. (1997): *Multivariate Models and Dependence Concepts*. London: Chapman & Hall.

[6] Karlin, S., and Y. Rinott (1980): "Classes of orderings of measures and related correlation inequalities I: Multivariate totally positive distributions", *Journal of Multivariate Analysis* 10, 467-498.

[7] Krishna, V. (2002): *Auction Theory*. Cambridge (USA): Academic Press.

[8] Lindberg, P.O. (2012): *Contributions to Probabilistic Discrete Choice*. Ph.D. Thesis, Department of Transport Science, KTH Royal Institute of Technology, Stockholm. (http://kth.diva-portal.org/smash/get/diva2:528102/FULLTEXT01.pdf)

[9] Lindberg, P.O., E. A. Eriksson, and L.-G. Mattsson (1995): "Invariance of achieved utility in random utility models", *Environment and Planning A* 27 (1), 121-142.

[10] Mattsson, L.-G., J. Weibull, and P.O. Lindberg (2014): "Extreme values, invariance and choice probabilities", *Transportation Research Part B* 59, 81-95.

[11] McFadden, D. (1974): "Conditional logit analysis of qualitative choice behavior", in Zarembka (ed.), *Frontiers in Econometrics*. New York: Academic Press.

[12] McFadden, D. (1978): "Modelling the choice of residential location", Chapter 25 in Karlqvist, Lundqvist, Snickars, Weibull (eds.), *Spatial Interaction Theory and Planning Models*. Amsterdam: North-Holland

[13] McFadden, D. (1981): "Econometric models of probabilistic choice", in Manski, McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Applications*. Cambridge (USA): MIT Press.

[14] Milgrom, P., and R. Weber (1982): "A theory of auctions and competitive bidding", *Econometrica* 50, 1089-1122.

[15] Nelsen, R. B. (2006): *An Introduction to Copulas, Vol. 2.* Berlin: Springer Series in Statistics.

[16] Perloff, J. M., and S. C. Salop (1985): "Equilibrium with product differentiation", *Review of Economic Studies* 52(1), 107–120.

[17] Resnick, S. and R. Roy (1990a): "Multivariate extremal processes, leader processes and dynamic choice models", *Advances in Applied Probability* 22 (2), 309-331.

[18] Resnick, S., and R. Roy (1990b): "Leader and maximum independence for a class of discrete choice models", *Economics Letters* 33 (3), 259-263.

[19] Robertson, C. A. and D. J. Strauss (1981): "A characterization theorem for random utility variables", *Journal of Mathematical Psychology* 23 (2): 184–89.

[20] Sattinger, M. (1984): "Value of an additional firm in monopolistic competition", *Review of Economic Studies* 51(2), 321–332.

[21] Strauss, D. (1979): "Some results on random utility models", *Journal of Mathematical Psychology* 20 (1): 35–52.

[22] Train, K., and W. W. Wilson (2008): "Estimation on stated-preference experiments constructed from revealed-preference choices", *Transportation Research Part B* 42 (3), 191-203.