Affiliation and Dependence in Economic Models*

Luciano I. de Castro†

September 30, 2009

Abstract

Affiliation has been a prominent assumption in the study of economic models with statistical dependence. Despite its large number of applications, especially in auction theory, affiliation has limitations that are important to be aware of. This paper shows that affiliation is a restrictive condition and the intuition usually given for its adoption may be misleading. Also, other usual justifications for affiliation are not compelling. Moreover, some implications of affiliation do not generalize to other definitions of positive dependence. These results show the need to consider alternatives to affiliation. The results of this paper suggest new directions for the study of dependence in economics. The main result classifies economic models of information and proves the existence of a minimally informative random variable that makes types conditionally independent. If this variable is known, then all results that are valid under independence are also valid for these models with statistically dependent types. Complementing this result, we describe a method to study general forms of dependence using grid distributions. Grid distributions are distributions whose densities are constant in squares and they are dense in the set of all distributions. This method allows a comprehensive investigation on the revenue ranking of auctions under general dependence.

JEL Classification Numbers: C62, C72, D44, D82.

Keywords: affiliation, positive dependence, statistical dependence of types, conditional independence, de Finetti's theorem, minimally informative random variable, auctions, pure strategy equilibrium, revenue ranking.

1 Introduction

Asymmetric information is a central theme in modern economics, not only in game theory, but also in industrial organization, general equilibrium, group decision, finance and many other subdisciplines. Most models assume that each agent privately knows a random variable, and

*Earlier versions of some of the results in the first part of this paper appeared in a weaker form in de Castro (2007).

†Department of Economics, University of Illinois at Urbana-Champaign. 1206 South Sixth St., Champaign, IL, 61822. E-mail: decastro.luciano@gmail.com.

these random variables are statistically independent. Although independence is convenient for theoretical manipulations, it is considered a restrictive and unrealistic assumption. Independence is regarded as restrictive because it is satisfied by a "knife-edge" set of distributions, and unrealistic because there are many potential sources of correlation in the real world: media, education, culture or even evolution. Perceiving these limitations early on, economists tried to surpass the mathematical difficulties and include statistical dependence in their models.

The introduction of affiliation was a milestone in the study of dependence in economics. This remarkable contribution was made by Milgrom and Weber (1982a), who borrowed a statistical concept (multivariate total positivity of order 2, MTP_2) and applied it to a general model of symmetric auctions.[1] Affiliation is a generalization of independence—see its definition in section 2—that was introduced through the appealing *positive dependence intuition*: "Roughly, this [affiliation] means that a high value of one bidder's estimate makes high values of the others' estimates more likely" (Milgrom and Weber (1982a, p.1096)). Among many important results, Milgrom and Weber (1982a), were able to show that positive dependence (in the form of affiliation) does not create problem for *pure strategy equilibrium existence*,[2] but it affects in a clear way the *revenue ranking* of auctions, that is, under affiliation, the English and the second-price auction give higher expected revenue than the first-price auction (in self-explanatory symbols: $R_E \geqslant R_2 \geqslant R_1$). These two results suggest an economic interpretation in terms of comparative statics: when the assumption of independence is relaxed in the direction of positive dependence, equilibrium is not a problem and the revenue superiority of the English auction (and second-price auction) increases.[3] From an economic point of view this comparative statics exercise is very interesting, since it clearly indicates what happens to the conclusion of the revenue equivalence theorem (RET) when one of its assumptions is relaxed (from independence to affiliation).[4]

For a quarter of a century, affiliation has been part of the foundations of auction theory and

[1] In two previous papers, Milgrom (1981b) and Milgrom (1981a) presented results that used a particular version of the concept, under the name "monotone likelihood ratio property" (MLRP). It is also clear that Wilson (1969) and Wilson (1977) influenced the development of the affiliation idea. Nevertheless, the concept was fully developed and the term affiliation first appeared in Milgrom and Weber (1982a). See also Milgrom and Weber (1982b). When there is a density function, the property had been previously studied by statisticians under different names. Lehmann (1966)) calls it Positive Likelihood Ratio Dependence (PLRD), Karlin (1968) calls it Total Positivity of order 2 (TP_2) for the case of two variables or Multivariate Total Positivity of Order 2 (MTP_2) for the multivariate case.

[2] Although equilibria in mixed strategies always exist (Jackson and Swinkels (2005)), first-price auctions may fail to possess a pure strategy equilibrium when types are dependent. However, Milgrom and Weber (1982a) proved that affiliation ensures the existence of a symmetric monotonic (increasing) pure strategy equilibrium (SMPSE) for symmetric first-price auctions. Milgrom and Weber (1982a) also proved the existence of equilibrium for second-price auctions with interdependent values. In our setup (private values), the second-price auction always has an equilibrium in weakly dominant pure strategies, which simply consists of bidding the private value.

[3] For private value auctions, which is the focus of this paper, English and second-price auctions are equivalent, which implies $R_E = R_2$. See Milgrom and Weber (1982a).

[4] Besides independence, the RET requires other restrictive conditions, such as symmetry and risk neutrality. The revenue ranking of auctions is undetermined if all those assumptions are relaxed. Thus, the importance of the result is akin to a comparative statistics exercise: holding everything else fixed, what changes if independence is relaxed in the direction of positive dependence?

almost synonymous with dependence in auctions. Affiliation's monotonicity properties (see Theorem 5 of Milgrom and Weber (1982a)) combine well with natural properties of auctions, simplifying the analysis and allowing useful predictions. But the success of affiliation is not restricted to auction theory. Whenever information is important, affiliation may potentially be applied. In fact, researchers in many different areas of economics and finance used the concept to obtain useful results.[5] In sum, few theoretical tools achieved as broad an impact as affiliation.

However, as with any scientific achievement, affiliation has limitations. The purpose of this paper is twofold: to reveal some of affiliation's limitations and to establish new facts that may lead to different approaches in the study of dependence. The assessment of affiliation covers sections 3 and 4 (section 2 introduces the model and gives basic definitions). The main result of this first part is Theorem 4.2, which shows that the revenue ranking implied by affiliation can be reversed, that is, $R_1 > R_2$, even under a strong form of positive dependence (first-order stochastic dominance). In other words, affiliation's and positive dependence's implications may differ. Section 5 collects the results that indicate new directions. The main result of this part is Theorem 5.2, which allows a classification of dependence in cases where it does not matter (independence can be used) and where it may matter. We argue that this classification is fundamental for understanding if dependence is at all significant. Using the ideas described in section 5, experimental or empirical economists can verify whether and to what extent statistical dependence is important and independence is not already a good approximation of the reality. If dependence is important, the results in section 5.2 develop a method for its comprehensive study. In sum, the paper has three clear messages: 1) affiliation may give misleading indications in the study of dependence; 2) independence may be more realistic and useful than it is usually considered; 3) it is possible to study dependence using a general method. Although this outline gives a good overview of this paper, we will describe its content in more detail.

Section 3 discusses whether or not affiliation can be well justified as an assumption in economic models.[6] Section 3.1 shows that affiliation is an extremely narrow condition. Although generality is usually considered as a criterion for judging an assumption (see Stigler (1950)), many common assumptions in economics fail to satisfy this criterion. Thus, affiliation's narrowness may be considered only a (useful) observation. If we have a reasonable intuition for affiliation, its restrictiveness may not be an issue. Section 3.2 reexamines the intuition used by Milgrom and Weber (1982a) to introduce affiliation (the *positive dependence intuition* referred to above). This intuition may be misleading, as there are many different (and weaker) definitions of positive dependence. Even if we accept the intuition and "roughly identify" positive

[5] For instance, Bergin (2001) used affiliation to obtain a generalization of a theorem by Aumann (1976) for the aggregation of information by a set of individuals; Persico (2000) proved a theorem about the usefulness of information for a decision maker under affiliation; and Sobel (2006) also used affiliation to study aggregation of information by groups. This list represents just a very small sample of papers; it would be almost impossible to cite all applications.

[6] This paper uses the word "justification" many times and I want to clarify its intended meaning. Justification is any coherent reason that an economist may have to adopt an assumption. It may be either an intuition for why it should be typically true or just a pragmatic reason such as "it seems to explain reality."

dependence with affiliation, there are problems to reach the conclusion that types are affiliated, as sections 3.3 and 3.4 discuss. Conditional independence is sometimes used as a justification for affiliation, but this justification works only in special cases, as shown in section 3.5. Section 3.6 explains the fact that English auctions are more common than first-price auctions by using the revenue equivalence theorem (RET), instead of the linkage principle, thus showing that the fact that open auctions are more common in the real world should not be considered evidence for affiliation. The rest of section 3 examines a number of alternative justifications for affiliation and show that very few, if any, can count as justifications for a generic use of affiliation in economic models.

On the other hand, sometimes it is argued that even if there is no good reason for the adoption of an assumption, it can be useful to economic theory if its implications are typically true.[7] Therefore, it is useful to consider the robustness of some of the main implications of affiliation, among which we choose equilibrium existence and the revenue ranking of auctions mentioned above. Section 4 shows that some of the implications of affiliation (equilibrium existence and the revenue ranking of auctions) are not preserved if we use other, only slightly less restrictive, notions of positive dependence. This lack of robustness puts in doubt the usefulness of the results that affiliation allow us to prove, since those results may not be typical.[8,9]

These negative results suggest that we may need to consider new approaches. Section 5 presents a series of results that may indicate new directions for the study of dependence in economic models. These results are organized in two parts: a classification of statical dependence using conditional independence (section 5.1) and a general approach that allows the study of dependence in a set of distributions that is dense in the set of all possible distributions (section 5.2).

The main result of section 5.1 is Theorem 5.2, which shows that all economic models fall into three cases: (1) types are independent; (2) types are dependent, but after receiving their private information, players' conditional beliefs about other player's types are independent; (3) types are truly statistically dependent, but there always exists a variable that makes types conditionally independent. In this case, we say that such variable *conditionally splits* the types. This last part of Theorem 5.2 *does not* come from de Finetti's theorem. Theorem 5.2 also states that there is a minimally informative random variable that makes the signals of all individuals

[7] This position was advocated by Friedman (1953).

[8] In section 5.2, we actually show that an important implication of affiliation, namely, the linkage principle, is *typically false* for more general distributions. Thus, Friedman (1953)'s criterion—if the theory produces "good approximation"—may not be satisfied for affiliation either. In a sense, by assuming affiliation economists take the risk of reaching the wrong conclusions.

[9] An opposition that the first part of this paper may bring about is a dismissive attitude like: "we are already know that affiliation is restrictive, as many other assumptions are. So why bother with this assessment?" It should be noted that although the restrictiveness reported here is a new theorem, this result is only an introductory aspect of the assessment made in sections 3 and 4. The assessment is more fundamental and boils down, at the end, to: a) conceptual difficulties of the whole idea (section 3), and b) the potentially misleading implications of the assumption (section 4). Both parts of this paper (the assessment and the proposal) could stand as separate papers, but they complement each other. A negative assessment without a proposal may lead to a "so what?" position. The proposal without the assessment may lead to the question: "Why bother with any new method if we already have a much simpler way (affiliation) to address the question of dependence?" The attempt to avoid both positions explains the current organization of the results.

(conditionally) independent, that is, there is a minimally informative conditional splitter. In general there is no "least" informative conditional splitter in the sense of inclusion by σ-fields, but we introduce a new definition that establishes such an existence. Section 5.1 then discusses how this result can illuminate the study of dependence in economic models. We show that any result that holds under independence also holds true if the conditional splitter is known, even if the types are statistically dependent. Since a conditional splitter always exits and dependence is not an issue if the players know such conditional splitter, any approach to dependence has to rule out first the possibility that this conditional splitter is known by the players. Interestingly, this task may be very difficult to do in empirical works, but it can be done in laboratory settings with a control of players' knowledge. This empirical/experimental question is fundamental in order to determine if (and in which circumstances) dependence is relevant in economic models.

For the cases in which dependence is important, section 5.2 describes a method based in approximating any distribution by grid distributions, which are distributions whose density functions are constant in squares. This set of distributions is dense in the set of all distributions but is sufficiently simple to allow obtaining useful results. This method allows us to have a very good idea of what happens in general in the set of all distributions. In particular, it allows us to obtain an indication of the revenue ranking of auctions both under general dependence and under positive dependence: although the ranking can go in either direction, the "typical" revenue ranking (in a sense explained in section 5.2) is $R_1 > R_2$, exactly the opposite of the ranking given by affiliation.

In many discussions, especially regarding affiliation's implications, we focus primarily on auctions. Because of that, we describe a basic auction setup in section 2. However, our results are not restricted to auction theory and may be of interest to any discipline that uses statistical dependence of signals to describe the private information of economic agents. Section 6 briefly reviews the literature and section 7 discusses potential directions for future work, not only in theory but also in econometrics and experimental economics.

2 Basic model and definitions

As emphasized above, our main results are not restricted to auctions. However, since affiliation's main implications discussed in section 4 refer to auctions, we will describe an auction model below.

There are n bidders, $i = 1, ..., n$. Bidder i receives private information $t_i \in \left[\underline{t}, \overline{t}\right]$ which is the value of the object for himself. The usual notation $t = (t_i, t_{-i})$ = $(t_1, ..., t_n) \in \left[\underline{t}, \overline{t}\right]^n$ is adopted. The (private) values are distributed according to a pdf $f : \left[\underline{t}, \overline{t}\right]^n \to \mathbb{R}_+$ which is symmetric. That is, if $\pi : \{1, ..., n\} \to \{1, ..., n\}$ is a permutation, $f(t_1, ..., t_n) = f\left(t_{\pi(1)}, ..., t_{\pi(n)}\right)$. Let $\overline{f}(x) = \int f(x, t_{-i})\, dt_{-i}$ be a marginal of f. Our main interest is the case where f *is not* the product of its marginals, that is, the case where the types are dependent. We denote by $f(t_{-i} \mid t_i)$ the conditional density $f(t_i, t_{-i}) / \overline{f}(t_i)$.

After knowing his value, bidder i places a bid $b_i \in \mathbb{R}_+$. He receives the object if $b_i > \max_{j \neq i} b_j$. We consider both first and second-price auctions with *private values*. This means that the private information of each bidder (type) is also that bidder's value for the object. As

Milgrom and Weber (1982a) show, second-price and English auctions are equivalent in the case of private values, as we assume here. In a first-price auction, if $b_i > \max_{j\neq i} b_j$, bidder i's utility is $u(t_i - b_i)$ and is $u(0) = 0$ if $b_i < \max_{j\neq i} b_j$. In a second-price auction, bidder i's utility is $u(t_i - \max_{j\neq i} b_j)$ if $b_i > \max_{j\neq i} b_j$ and $u(0) = 0$ if $b_i < \max_{j\neq i} b_j$. For both auctions, ties are randomly broken.

A pure strategy is a function $b : [0,1] \to \mathbb{R}_+$, which specifies the bid $b(t_i)$ for each type t_i. The interim payoff of bidder i, who bids β when his opponent $j \neq i$ follows $b : [0,1] \to \mathbb{R}_+$ is given by

$$\Pi_i(t_i, \beta, b(\cdot)) = u(t_i - \beta) F\left(b^{-1}(\beta) \mid t_i\right) = u(t_i - \beta) \int_{\underline{t}}^{b^{-1}(\beta)} f(t_j \mid t_i)\, dt_j,$$

if it is a first-price auction and

$$\Pi_i(t_i, \beta, b(\cdot)) = \int_{\underline{t}}^{b^{-1}(\beta)} u(t_i - b(t_j)) f(t_j \mid t_i)\, dt_j,$$

if it is a second-price auction.

We focus attention on symmetric monotonic pure strategy equilibrium (SMPSE), which is defined as $b(\cdot)$ such that $\Pi_i(t_i, b(t_i), b(\cdot)) \geqslant \Pi_i(t_i, \beta, b(\cdot))$ for all β and t_i. The usual definition requires this inequality to be true only for almost all t_i. This stronger definition creates no problems and makes some statements simpler, such as those about the differentiability and continuity of the equilibrium bidding function (otherwise, such properties should be qualified by the expression "almost everywhere"). Finally, under our assumptions, the second price auction always has a SMPSE in a weakly dominant strategy, which is $b(t_i) = t_i$.

By reparametrization, we may assume, without loss of generality, $\left[\underline{t}, \overline{t}\right] = [0,1]$. It is also useful to assume $n = 2$, but this is not necessary for most of the results. We also assume risk neutrality, i.e., $u(x) = x$. Thus, unless otherwise stated, the results will be presented under the following setup:

BASIC SETUP: *There are* $n = 2$ *risk neutrals bidders* ($u(x) = x$), *with private values distributed according to a symmetric density function* $f : [0,1]^2 \to \mathbb{R}_+$.

Affiliation is formally defined as follows.[10]

Definition 2.1 *The density function* $f : \left[\underline{t}, \overline{t}\right]^n \to \mathbb{R}_+$ *is affiliated if* $f(t) f(t') \leqslant f(t \wedge t') f(t \vee t')$, *where* $t \wedge t' = (\min\{t_1, t_1'\}, ..., \min\{t_n, t_n'\})$ *and* $t \vee t' = (\max\{t_1, t_1'\}, ..., \max\{t_n, t_n'\})$.

It is useful to introduce the following notation: $\mathcal{D}$ will denote the set of all densities:

$$\mathcal{D} \equiv \{f : [0,1]^n \to \mathbb{R}_+ : \int_{[0,1]^n} f(t) dt = 1\}.$$

The set of all continuous densities will be denoted $\mathcal{C}$ and $\mathcal{A}$ will denote the set of affiliated (continuous or not) densities.

[10] It is possible to define affiliation even if the joint distribution has no density function. See Milgrom and Weber (1982a).

3 Justifications for affiliation

This section evaluates many possible justifications for affiliation. The departure point is the fact proven in subsection 3.1 that affiliation is restrictive. This result motivates the analysis in subsection 3.2 of the positive independent intuition used to introduce affiliation in the first place. In this subsection we also develop the concept of *rough identification*, which is the practice of "roughly" identifying positive dependence and affiliation. We discuss the appropriateness of such methodological shortcut with respect to other parts of the theory in subsections 3.3 and 3.4. Another common justification for affiliation is conditional independence, discussed in subsection 3.5. Subsection 3.6 discusses the "evidence" of affiliation through the predominance of open auctions in the real world. In that subsection, we use the Revenue Equivalence Theorem to explain exactly the same observation. The justifications for affiliation in other sciences do not carry over to economics, as subsection 3.7 explains. Other justifications are considered in subsections 3.8 and 3.9. Subsection 3.10 puts the findings of this section in perspective and prepares for the discussion in section 4.

3.1 Affiliation is restrictive

In this section, we show that affiliation is a restrictive assumption, i.e., the set of affiliated densities is small in the set of all densities. There are two ways to characterize a set as small: topological and measure-theoretic. Although it is possible to show that affiliation is restrictive in the measure-theoretic sense (see de Castro (2007)), here we limit ourselves to the topological result, which is simpler.

Recall that $\mathcal{C}$ denotes the set of continuous density functions $f : [0,1]^n \to \mathbb{R}_+$ and $\mathcal{A}$, the set of affiliated densities (continuous or not). Endow $\mathcal{C}$ with the topology of the uniform convergence, that is, the topology defined by the norm of the sup:

$$\|f\| = \sup_{x\in[0,1]^n} |f(x)|.$$

The following theorem shows that the set of continuous affiliated densities is small in the topological sense.[11] The proofs of this and of all other results are given in the appendix.

Theorem 3.1 *The set of continuous affiliated density functions $\mathcal{C} \cap \mathcal{A}$ is meager.*[12] *More precisely, the set $\mathcal{C} \backslash \mathcal{A}$ is open and dense in $\mathcal{C}$.*

Although the restrictiveness of affiliation seems to be a "folk theorem," it was never stated or formally proven. Thus, Theorem 3.1 (and the measure theoretical result proved in de Castro (2007)) fill this gap in the literature. From an economic point of view, however, the most important aspect of this result may not be its technical contribution, but instead the fact that it

[11] Theorem 3.1 continues to hold if instead of $\mathcal{C}$ we use the set $\mathcal{C}^k$ of k-continuously differentiable functions, with its standard topology. The proof of this fact is essentially the same.

[12] A meager set (or set of first category) is the union of countably many nowhere dense sets, while a set is nowhere dense if its closure has an empty interior. Thus, the theorem says more than that $\mathcal{C} \cap \mathcal{A}$ is meager: $\mathcal{C} \cap \mathcal{A}$ is itself a nowhere dense set, according to the second claim in the theorem.

allows the discussion of a subtle point about the use of affiliation as an assumption of dependence in economic models.

Generality is usually listed as a criterion for judging an assumption or theory—see, for instance, Stigler (1950, Section VIII-A, p. 392) or Fudenberg (2006, p. 695). However, some economists may disagree about the importance of such a criterion, since many economic theories are based on narrow assumptions. Following this methodological position, an economist may consider that the restrictiveness of affiliation established by Theorem 3.1 is not relevant for an assessment of affiliation. Fortunately, Theorem 3.1 is not subjected to this philosophical dismissal. If generality is important, then Theorem 3.1 implies that affiliation is not a good assumption; on the other hand, if generality is not important, we could stick to the classic setting and assume only independence, which is simpler. Thus, regardless of our methodological position about the relevance of generality/restrictiveness, Theorem 3.1 *suggests* that affiliation may not provide the best model for information economics.

Of course, this is still only a preliminary observation. The gap between independence and affiliation, despite being small, may contain the economically relevant cases. This possibility is actually suggested by the intuition given for affiliation. The next subsection revisits such an intuition.

3.2 The intuition for affiliation may be misleading

Affiliation was introduced through the *positive dependence intuition*: "a high value of one bidder's estimate makes high values of the others' estimates more likely" (Milgrom and Weber (1982a, p. 1096)). This intuition is very appealing, because positive dependence describes a circumstance likely to happen in the real world. In fact, many authors introduce affiliation through this intuition or some of its variations.

Affiliation captures this intuition, as we illustrate in Figure 1, below. Affiliation requires that the product of weights at points (x', y') and (x, y) (where both values are high or both are low) be greater than the product of weights at (x, y') and (x', y) (where they are high and low, alternatively). In other words, the distribution puts more weight on the points in the diagonal than outside it.

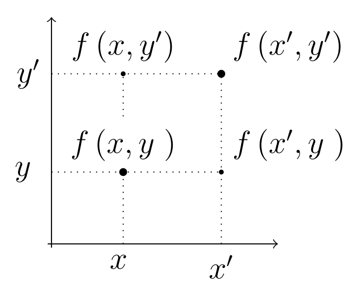


Figure 1 — The pdf f is affiliated if $x \leqslant x'$ and $y \leqslant y'$ imply $f(x, y')\, f(x', y) \leqslant f(x', y')\, f(x, y)$.

However, as long as we are interested in *positive dependence*, as this intuition suggests, affiliation is not the only definition available. In the statistical literature many concepts have been proposed to correspond to the notion of positive dependence. For simplicity, let us consider only the bivariate case and assume that the two real random variables X and Y have joint distribution F and strictly positive density function f. The following concepts are formalizations of the notion of positive dependence for X and Y:[13]

Property I — X and Y are positively correlated (PC) if $cov(X, Y) \geqslant 0$.

Property II — X and Y are said to be positively quadrant dependent (PQD) if for all non-decreasing functions g and h, $cov(g(X), h(Y)) \geqslant 0$.

Property III — The real random variables X and Y are said to be associated (As) if for all non-decreasing functions g and h, $cov(g(X, Y), h(X, Y)) \geqslant 0$.

Property IV — Y is said to be left-tail decreasing in X (denoted LTD($Y|X$)) if for all y, the function $x \mapsto \Pr[Y \leqslant y | X \leqslant x]$ is non-increasing in x. X and Y satisfy Property IV if LTD($Y|X$) and LTD($X|Y$).

Property V — Y is said to be positively regression dependent on X (denoted PRD($Y|X$)) if $\Pr[Y \leqslant y | X = x] = F(y|x)$ is non-increasing in x for all y. X and Y satisfy Property V if PRD($Y|X$) and PRD($X|Y$).[14]

Property VI — Y is said to be Inverse Hazard Rate Decreasing in X (denoted IHRD($Y|X$)) if $\frac{F(y|x)}{f(y|x)}$ is non-increasing in x for all y, where $f(y|x)$ is the pdf of Y conditional to X. X and Y satisfy Property VI if IHRD($Y|X$) and IHRD($X|Y$).

Since there are many alternative definitions of positive dependence, a natural question is: "How do such definitions compare with affiliation?" The following theorem provides the answer.

Theorem 3.2 *Let affiliation be Property VII. Then,*

$$(VII) \Rightarrow (VI) \Rightarrow (V) \Rightarrow (IV) \Rightarrow (III) \Rightarrow (II) \Rightarrow (I),$$

and all implications are strict.

For this theorem, we used only seven concepts for simplicity. Yanagimoto (1972) defines more than thirty concepts of positive dependence and, again, affiliation is the most restrictive of all but one.

One can say that the main contribution of this section is *not* the mathematical result presented as Theorem 3.2, but the observations that: 1) positive dependence was our primary

[13] Most of the concepts can be properly generalized to multivariate distributions. See, for example, Lehmann (1966) and Esary, Proschan, and Walkup (1967). The hypothesis of strictly positive density function is made only for simplicity.

[14] This property is also known as monotonicity in the first-order stochastic dominance sense.

target in the study of dependence in auctions; 2) affiliation *is not* positive dependence but just one among many possible definitions—and it is, in fact, one of the most restrictive.

This observation is important for an assessment of the assumption. If we believe that *positive dependence* corresponds to the set of economically relevant cases, then affiliation may not be the correct assumption or, in other words, the received intuition may be misleading. Accepting the intuition, we may believe that we are covering exactly the important cases, when we are not. The contribution here is to warn of this potential gap. In fact, we will show in the subsections below that the gap between the intuition and the actual assumption is complicated for other assumptions of the model. For this discussion, it will be useful to define *rough identification* as the identification of positive dependence with affiliation, based on the intuition that "roughly, this [affiliation] means that a high value of one bidder's estimate makes high values of the others' estimates more likely."

3.3 Rough identification and multidimensionality of information

We argued above that the rough identification may be misleading. In this subsection, we show that even if the rough identification is considered acceptable to the estimates that bidders make, it does not necessarily follow that affiliation is justified.

The starting point of our argument is a quote from Milgrom and Weber (1982a, p. 1093-4), where they explain why their model is more realistic:

> "(...) consider the situation in an auction for mineral rights on a tract of land where the value of the rights depends on the unknown amount of recoverable ore, its quality, its ease of recovery, and the prices that will prevail for the processed mineral."

Thus, according to Milgrom and Weber (1982a) the typical situation is that of multidimensional information. In other words, it is likely that the expected value of an object is a function of various variables (quality, price, etc.), which is to say that private information is a multidimensional signal.[15] After making the estimates for each variable, the bidders use some model to reach an estimate of the value of the object. Now, the (reasonable) idea that the bidders' estimates are positively correlated is translated by *rough identification* into the assumption that the estimates of each variable are affiliated across bidders. The question is, does this imply that the values of the object are affiliated? It is useful to rephrase this question in more formal terms: let $X_i^1, ..., X_i^m$ denote the estimates of a bidders for the m relevant variables and let $\tau_i \equiv v_i\left(X_i^1, ..., X_i^m\right)$ be bidder i's estimation for the value of the object. The rough identification leads us to accept that $X_1^k, ..., X_N^k$ are affiliated for $k = 1, ..., m$. Would this imply that the τ_i are affiliated? Unfortunately, the answer is no, as the following example shows.[16]

[15]Of course, this observation cannot be taken from a naive point of view. In many cases, multidimensional random variables can be reduced to unidimensional ones. Nevertheless, we will illustrate that such reduction is not free of consequences.

[16]In this discussion and in the example, we have worked with private values, although Milgrom and Weber (1982a) emphasized common values. However, the negative result presented in the example is even easier to obtain with common values.

Example 3.3 *Auction of an Oil Lease Consider the auction of a tract between two bidders. Buyer i has a private estimation of the oil quality in the field, (q_i), and the amount of recoverable ore (s_i). Estimates of these two variables are drawn from independent distributions, but q_1 and q_2 are affiliated, as well s_1 and s_2. The value of the field is calculated as $\tau_i = p(q_i)\, s_i - c(s_i)$, where $p(.)$ denotes the price of the oil according to its quality and $c(.)$ is the cost of oil extraction, depending, obviously, on the size. For simplicity, we will give numerical examples with discrete values: the size can be S (small), M (medium) or B (big). The quality can be L (low) or H (high). There are two bidders and their signals obey the distributions below that are easily checked to be affiliated.*

big	$1/18$	$1/12$	$1/6$
medium	$1/12$	$1/6$	$1/12$
small	$1/6$	$1/9$	$1/18$
2 ↑ / 1→	*small*	*medium*	*big*

Table 1 — Joint Distribution of Bidder's Estimates for the size of the reserve.

High	$1/3$	$1/3$
Low	$1/6$	$1/6$
2 ↑ / 1→	*Low*	*High*

Table 2 — Joint Distribution for the Estimates of the Quality of the Oil

We represent the small (S) size as 1, the medium (M) size as 2 and the big (B) as 3, take a non-decreasing function of costs, $c(1) = c(2) = 1$ and $c(3) = 3$. Let $p(L) = 1$ and $p(H) = 3$. With this, the possible values of the hole are $\tau_i = 0, 1$ or 3. Such specification leads us to the following distribution of types τ_i :

3	$5/36$	$5/36$	$1/6$
1	$23/216$	$5/36$	$7/72$
0	$2/27$	$5/72$	$5/72$
$\tau_2 \uparrow / \tau_1 \rightarrow$	0	1	3

Table 3 — Distribution of Values

It is easy to see that this distribution is not affiliated: for example, using the four probabilities in the right down corner, we have $\frac{5}{36} \cdot \frac{5}{72} > \frac{7}{72} \cdot \frac{5}{72}$.[17]

The above example suggests that there are serious problems when affiliation has to be applied to multidimensional settings.[18] The point in the example is actually valid in more

[17]We can go further. Suppose that the small size (S) is $\frac{1}{42}$, the medium size (M) is $\frac{5}{6}$ and the big size (B) is $\frac{6}{7}$. We take an increasing function of costs, $c\left(\frac{1}{42}\right) = 0$, $c\left(\frac{5}{6}\right) = \frac{29}{6}$ and $c\left(\frac{6}{7}\right) = 5$. Let $p(L) = 6$ and $p(H) = 7$. With these values, the possible values of the petroleum field are $\tau_i = \frac{1}{7}, \frac{1}{6}$ or 1 and the distribution showed in Table 3 remains the same, just substituting 0, 1 and 3 with $\frac{1}{7}, \frac{1}{6}$ and 1. Then, if bidder 1 has common value utility $u_1 = \frac{\tau_1+\tau_2}{2}$, as usual, the expected utility turns out to be non-monotonic. Indeed, $E\left[\frac{\tau_1+\tau_2}{2} | \tau_1 = \frac{1}{7}\right] = 0.3332 > 0.3310 = E\left[\frac{\tau_1+\tau_2}{2} | \tau_1 = \frac{1}{6}\right]$.

[18]Previous examples of these problems were provided by Reny and Perry (1999) and Reny and Zamir (2002). Note that our point here is different from these papers since we are pointing out a fundamental problem for Milgrom and Weber (1982a)'s basic single object symmetric model, not for its extensions, as those papers do.

general terms: it can be shown that unless the function v_i that defines τ_i depends only on a single variable, the τ_i will not be affiliated.[19] We conclude that even if we are ready to accept the rough identification, we will not be able to justify affiliation. In other words, the justification for Milgrom and Weber (1982a)'s basic model with affiliation is based on a case where it holds only under stronger assumptions.[20]

Let us summarize our findings. We began this section by assuming that the rough identification was true, which implied that bidders' estimates of the many relevant variables were affiliated. We have seen that affiliation would pass to the value of the object only if the bidders' estimates were unidimensional. This assumption may be reasonable in auctions of non-durable goods, such as fish, if the bidders do not need to take in account different aspects of the product.

Nevertheless, our faith in the theory could lead us to expect that, under the assumptions of 1) rough identification and 2) bidders' estimates are just one real variable, the theory could finally be well grounded in. Unfortunately, this is yet not true, as we discuss next.

3.4 Rough identification and sufficient statistics

To describe the remaining difficulties, let us quote again Milgrom and Weber (1982a):

> "To represent a bidder's information by a single real-valued signal is to make two substantive assumptions. Not only must his signal be a sufficient statistic for all of the information he possesses concerning the value of the object to him, it must also adequately summarize his information concerning the signals received by the other bidders. The derivation of such a statistic from several separate pieces of information is in general a difficult task. It is in the light of these difficulties that

[19]It is worth mentioning that this problem is related only to affiliation and not to positive dependence. Indeed, Jogdeo (1977) shows that if the estimates are associated (Property III above) and the functions v_i are concordant (monotonic in the same direction) then the τ_i will also be associated.

[20]Consider also the following reasoning, which closely relates the use of many variables to the need of positive correlation (and, hence, affiliation, under the rough identification):

> "What are we to use in place of the independence assumption? When the bidder's costs or valuation depend on some common random factors, so that all the bidders are estimating the same variables, their estimates will be positively correlated even if their estimation *errors* are independent. Positive correlation has been especially prominent in models of auctions for oil and gas drilling rights, where the rights being acquired are, to a first approximation, of equal value to each of the bidders, and the main uncertainties concern such common factors as the quantities of recoverable hydrocarbons, the cost of recovery, the costs of transporting the product to market (perhaps through as yet undeveloped pipelines over the Arctic Slope), future world energy prices, and so on. The common uncertainties found in these auctions also play a large role in the sale of items like wine or art which are purchased at least partly for their savings or investment value, as the parties estimate what it would cost to purchase the same vintage in the future or what the eventual resale price for the painting will be. So there is good reason to believe that positive correlations among value estimates will often be present.
>
> The actual equilibrium analysis of auctions relies on a stronger notion than positive correlation. The appropriate concept, known as affiliation, was introduced by Milgrom and Weber (1982)." Milgrom (1989, p. 13-4)

we choose to view each X_i as a "value estimate," which may be correlated with the "estimates" of others but is the only piece of information available to bidder i." Footnote 14, p. 1097

Now, the reasonable positive dependence intuition and the rough identification, that were previously justified for the estimates of the bidders, have to apply, indeed, for sufficient statistics. But sufficient statistics can be much more complex than simple estimates, as the quotation above points out. It is not clear why the positive dependence intuition should be considered reasonable for sufficient statistics. This point is, of course, related to the previous one, about multidimensional of estimates, but it is qualitatively different, since sufficient statistics summarize not only the bidder's own estimates, but also "his information concerning the signals received by the other bidders". Therefore, affiliation's justification seems even harder than the previous subsection suggested.[21]

This discussion suggests that we need a theory to derive and characterize the sufficient statistics from the pieces of information possessed by the bidders. In the absence of such a theory, we cannot be sure of what we are really assuming when we think of sufficient statistics as affiliated values. Of course, it is possible that in many settings the assumption is reasonable, but it is necessary to classify in which situations this is (approximately) true.

3.5 Conditional independence

A standard way to justify affiliation is to appeal to conditional independence. In fact, affiliation was originally motivated using conditional probabilities (see Milgrom and Weber (1982a, p. 1094). Conditional independence models assume that the signals of bidders are conditionally independent, given a variable v (the intrinsic value of the object, for instance). Since symmetry is the same as exchangeability, which is the main assumption of de Finetti's Theorem, some auction specialists seem to believe that de Finetti's Theorem implies that conditional independence holds in symmetric auctions *without loss of generality*. De Finetti's theorem states the following:

De Finetti's Theorem. *Consider a sequence of random variables X_1, X_2,..., and assume that they are exchangeable, that is, assume that the distribution of $(X_1, ..., X_n)$ is equal to the distribution of $\left(X_{\pi(1)}, ..., X_{\pi(n)}\right)$, for any n and any permutation $\pi : \mathbb{N} \to \mathbb{N}$. Then, there is a random variable Q such that all X_1, X_2,..., are conditionally independent (and identically distributed) given Q.*[22]

[21] More technically, one has to remember that a general theory of beliefs (about the beliefs) should take in account models of higher order beliefs (or even the universal type space of Mertens and Zamir (1985)). In this framework, it is not well known what affiliation implies or requires.

[22] De Finneti proved this theorem for the case where the X_i are Bernoulli variables. Hewitt and Savage (1955) extended it to the general setting. The statement above is somewhat vague. A precise statement is as follows: Let X_1, X_2,..., be an exchangeable sequence of random variables with values in a set S. Then there exists a probability measure μ on the set of probability measures $\Delta(S)$ such that for all measurable sets $A_1, ..., A_n$,

$$\Pr(X_1 \in A_1, ..., X_n \in A_n) = \int_{\Delta(S)} Q(A_1) \cdots Q(A_n) \mu(dQ).$$

Unfortunately, however, de Finneti's theorem *is not valid* for standard models of auction theory, even assuming symmetry. The reason is that standard auction models consider a finite number of players and, hence, a finite number of random variables. De Finneti's theorem is valid only for an (infinite) sequence of random variables.[23] The following example illustrates the problem:

Example 3.4 *Consider two random variables, X_1 and X_2, taking values in $\{0,1\}$, with joint distribution given by: $P(X_1 = 0, X_2 = 1) = P(X_1 = 1, X_2 = 0) = \frac{1}{2} - \varepsilon$ and $P(X_1 = 0, X_2 = 0) = P(X_1 = 1, X_2 = 1) = \varepsilon$. It is easy to see that X_1 and X_2 are symmetric (exchangeable). In the appendix, we show that the conclusion of de Finetti's Theorem cannot hold if $\varepsilon < 1/4$.*[24]

Thus, de Finneti's Theorem *does not* imply that conditional independence is a generic condition in symmetric auctions. There is, however, another way to justify conditional independence, as we discuss in section 5.1. However, even if we are ready to assume conditional independence, this is not yet sufficient for affiliation. To see this, assume that the pdf of the signals conditional to v, $f(t_1, ..., t_n|v)$, is C^2 (twice continuously differentiable) and has full support. It can be proven that the signals are affiliated if

$$\frac{\partial^2 \log f(t_1, ..., t_n|v)}{\partial t_i \partial t_j} \geqslant 0,$$

and

$$\frac{\partial^2 \log f(t_1, ..., t_n|v)}{\partial t_i \partial v} \geqslant 0, \tag{1}$$

for all i, j (see Topkis (1978, p. 310)). It is important to note that conditional independence implies only that

$$\frac{\partial^2 \log f(t_1, ..., t_n|v)}{\partial t_i \partial t_j} = 0.$$

Thus, conditional independence is not sufficient for affiliation. To obtain affiliation, one needs to assume (1) above, i.e., that t_i and v are affiliated. In other words, to obtain affiliation from conditional independence, one has to assume affiliation itself. Thus, conditional independence does not give an economic justification for affiliation.

The fact that we are not able to find a justification in the general model of conditional independence does not imply that it does not exist, at least in special cases. See the results and discussion in section 5.1.

[23] One can assume that there are an infinite number of potential players in the auction, but for some reason only a finite number of them actually participate. Then, one can apply de Finetti's theorem. However, this will be of course *with* a loss of generality.

[24] See also Proposition 5.7 below. Example 3.4 generalizes an example given by Diaconis and Freedman (1980). They prove an approximation version of de Finneti's theorem for a finite set of random variables. See a discussion of their paper after Proposition 5.7.

There is a particular conditional independence model where affiliation can be reasonably justified. Assume that the signals t_i are a common value plus an individual error, that is, $t_i = v + \varepsilon_i$, where the ε_i are independent and identically distributed. Now, we almost have the result that the signals t_1, ..., t_n are affiliated: it is still necessary to assume an additional condition. Let g be the pdf of the ε_i, $i = 1$, ..., n. Then, $t_1, ..., t_n$ are affiliated if and only if g is a strongly unimodal function.[25,26]

3.6 The explanation of open auctions predominance

It is sometimes argued that affiliation is a sound assumption because it successfully explains why English auctions are more common in reality than first-price auctions. The explanation comes from Milgrom and Weber (1982a)'s famous result that English auctions yield (weakly) higher expected revenue than first-price auctions under affiliation. This result is usually contrasted with the Revenue Equivalence Theorem (RET), which says that all standard auctions (under some assumptions) give exactly the same revenue. This subsection argues that such an explanation cannot be considered as evidence of affiliation and, therefore, it shall not be regarded as an affiliation's justification. The main point is the fact that this explanation is not the only plausible one. Indeed, this subsection offers a new explanation for the English auction predominance based exactly in the RET theorem.

Let us begin by observing the fact that when auctions began to be used thousands of years ago, when most people were illiterate, the task of writing a bid in a sealed envelope would be highly non trivial. Thus, the fact that open auctions were more common in the past is an immediate consequence of technological restrictions at that time. This simple observation explains the predominance of open auctions in the past. Why is that they are still more prevalent *today*, where the cost differences are not that large? The RET suggests that all auction mechanisms give the same revenue. In this case, sellers do not have any incentive to switch from open auctions to sealed-bid ones. Therefore, the predominance of open auctions can be sustained.

It is worth clarifying that we do not claim that this explanation is better than the one provided by affiliation. Our point is weaker: the predominance of open auctions is not necessarily a justification or an evidence of affiliation. On the contrary, we have just seen how we can use the RET to explain exactly the same thing! It is also interesting to observe that it is difficult to find empirical works supporting the thesis that English auctions are revenue superior.[27] A justification of affiliation through the revenue superiority of English auctions would need first to establish this superiority in empirical findings.

[25]The term is borrowed from Lehmann (1986). A function is strongly unimodal if $\log g$ is concave. A proof of the affirmation can be found in Lehmann (1986, Example 1, p. 509), or obtained directly from the previous discussion.

[26]Even if g is strongly unimodal, so that $t_1, ..., t_n$ are affiliated, it is not true in general that t_1, ..., t_n, ε_1, ..., ε_n, v are affiliated.

[27]For a review of empirical works in auctions, see Laffont (1997).

3.7 The use of affiliation in other sciences

As we commented in the introduction, affiliation is used—under other names—in other sciences. Thus, a natural question would be: "How can the use of affiliation be justified in other sciences but not in economics?"

Affiliation is used in statistics, as Positive Likelihood Ratio Dependence (PLRD), the name given by Lehmann (1966) when he introduced the concept, or in reliability theory, as Total Positivity of order 2 (TP_2) for the case of two variables, or Multivariate Total Positivity of Order 2 (MTP_2) for n variables, after Karlin (1968). TP_2 is used when there are good reasons for adopting special distributions in some problems, and those distributions happen to satisfy the TP_2 condition. An example of this can be seen in the historical notes of Barlow and Proschan (1965, Chapter 1) about reliability theory. It is natural to assume that the failure rates of components or systems follow specific probabilistic distributions (exponentials, for instance), and such special distributions have the TP_2 property. Thus, the corresponding theory of total positive distributions can be advantageously used. Another example of this is the use of copulas.[28] If we assume that the distribution is in a family of copulas that have the MTP property, then the use of affiliation's properties and implications is advantageous and justified by the choice of the set of distribution functions.

In the case of economic models, especially auction theory, the random variables (types) represent information gathered by the bidders. There are some situations where we can assume special forms of distributions, but in general there is no justification for such assumptions. In fact, specific distributions are rarely assumed in the theory.[29] Thus, the compelling justification that is presented for applications in reliability theory or statistics is not valid for economics.

3.8 Monotonic comparative statics

Monotone comparative statistics play a special role in economics. Results related to this technique are considered central to modern economics. See, for instance, Milgrom and Roberts (1990) and Milgrom and Shannon (1994)). Thus, if one is interested in this kind of results, affiliation may seem a natural assumption, not because of its soundness, but because of the results that it delivers. In sum, this justification is based on the results, not in the foundations of the assumption.

However, even if one really wants to focus only on monotone comparative statics results, it is not completely clear that affiliation is the right assumption. Reny and Zamir (2004) show that affiliation is not sufficient to imply the standard single-crossing condition in asymmetric auctions (although it is sufficient to imply a weaker condition that turns out to be sufficient for equilibrium when the signals are unidimensional). More important, they show that an auction with multidimensional affiliated signals may fail to have equilibrium in monotone strategies. Also, McAdams (2003) gives an example with three bidders and affiliated types where a non-monotonic equilibrium can exist. Perry and Reny (1999) show that the linkage principle can

[28] See, for instance, Li, Paarsch, and Hubbard (2007).

[29] McAfee and Vincent (1992) make a similar observation, when they note the "lack of any a priori guidance about the appropriate distribution" (p. 512).

fail in multi-unit auctions. In other words, affiliation is not in general sufficient to monotonic equilibria or monotone comparative statics results out of some standard cases. On the other hand, independence is in general sufficient for monotonic results. Thus, one can think that the advantage and importance of monotonic methods may provide a stronger support for the use of independence than for the use of affiliation.[30]

3.9 Other Justifications

Another kind of justification for affiliation may arise when we have some reason to believe or accept that the bidders' types have a specific distribution. If such distribution has the affiliation property, then the use of affiliation is justified. This is the case, for example, when we assume that the distributions are in some family of affiliated copulas. As we discuss in section 3.7, the problem with this justification is that there is very little support for using specific distributions as a model of players' beliefs.

In some cases, the assumption can be justified only by a pragmatic reason: "use affiliation because it works." Although this justification may not appear compelling at least at first glance, it is generally accepted when the focus is on the results, not on the appealing properties of the assumption. In a sense, the assumption is justified by methodological advantages, not by its soundness. This kind of justification is, in fact, very common, especially in applied works. Note however that such kind of justification would also apply to independence.

3.10 Discussion

This section showed that most justifications for the use of affiliation in economics are not well founded. However, it should be noted that the results and discussions in this section do not imply that affiliation is not satisfied in the real world; only empirical tests can assert this. Such tests are necessary, but they seem largely absent in the current experimental or empirical literature.[31]

Even if affiliation is not valid in the real world, the above analysis does not constitute yet a definitive reason to dismiss it. Rather, it is important to verify whether its implications are typically true or not. As Friedman (1953) argues, the most important criterion for judging an assumption is whether the resulting theory "yields sufficiently accurate predictions" (p. 14). This methodological position motivates the next section, which analyzes affiliation's implications.

---

[30] Again, it is possible to conceive a combination of criteria that would justify affiliation. For instance, the economist may want monotonicity *and* generality. And in some cases, affiliation does deliver monotonic results and, in some aspects, is more general. But in this case, other positive dependence conditions, such as those presented in section 3.2, should also be considered. For instance, van Zandt and Vives (2007)'s theorem discussed in section 4 shows that Property V above (first-order stochastic dominance) is sometimes sufficient to monotonic equilibrium existence.

[31] See Laffont (1997) for a survey of empirical literature on auctions. We are aware of only two independent working papers proposing tests of affiliation: de Castro and Paarsch (2008) and Jun, Pinkse, and Wan (2008). Both papers were motivated by an earlier version of this paper. See also section 7 for a final discussion of this topic.

4 Affiliation's implications may fail even under strong positive dependence

Affiliation has been used in the proof of many results. These results can be classified in two groups: facts that are already true for the independent case (affiliation allows a generalization) and predictions that are qualitatively different from the case of independence. In this section, we will focus on one implication for each of these groups.

The first one is the existence of symmetric monotonic pure strategy equilibrium (SMPSE) for first price auctions, generalized from independence to affiliation. The second one is the revenue ranking of auctions: under affiliation, the English and the second-price auction give expected revenue at least as high as the first price auction (a fact that we denote by $R_2 \geqslant R_1$). This last result is in contrast with the case of independence, where the Revenue Equivalence Theorem (RET) implies the equality of the expected revenues $(R_2 = R_1)$.[32,33] Both implications were obtained by Milgrom and Weber (1982a) and I chose them because of their importance. The purpose of this section is to verify whether these implications (existence of SMPSE and $R_2 \geqslant R_1$) are true in a more general setting.

4.1 Equilibrium existence

Is the existence of SMPSE true under other definitions of positive dependence (see section 3.2)? Theorem 4.1 below shows that the following property is sufficient:[34]

Property VI′ — The joint (symmetric) distribution of X and Y satisfy Property VI′ if for all x, x' and y in $[0, 1]$,

$$x \geqslant y \geqslant x' \Rightarrow \frac{F(y|x')}{f(y|x')} \geqslant \frac{F(y|y)}{f(y|y)} \geqslant \frac{F(y|x)}{f(y|x)}.$$

It is easy to see that Property VI implies Property VI′ (under symmetry and full support). Thus, the question becomes whether or not it is possible to generalize the existence of SMPSE for Property V or even further.

If we define $\Pi(x, y) = (x - b(y)) F(y|x)$, where $b(\cdot)$ is a candidate for symmetric equilibrium,[35] then equilibrium existence is equivalent to $\Pi(x, x) \geqslant \Pi(x, y)$. Since $b(\cdot)$ is monotonic, one may conjecture that the monotonicity of $F(y|x)$ — as Property V assumes — may be sufficient for equilibrium existence, through some single crossing arguments (see Athey (2001)). Since Property V is still a strong property of positive dependence, this conjecture

[32] Since affiliation contains independence as a special case, the results can be *qualitatively* different, but must have an overlap.

[33] Both the revenue ranking under affiliation and the RET requires symmetry, risk neutrality and the same payoff by the lowest type of bidders.

[34] Motivated by an earlier version of this paper, Monteiro and Moreira (2006) obtained other generalizations of equilibrium existence for non-affiliated variables. Their results are not directly related to positive dependence properties.

[35] This candidate is increasing and unique, as we can show using standard arguments. See Maskin and Riley (1984) or de Castro (2008).

may be considered reasonable. In fact, the reader may think that the following recent result by van Zandt and Vives (2007) actually *proves* that first-order stochastic dominance is sufficient for equilibrium existence in auctions:

Theorem (van Zandt and Vives, 2007): *Assume that for each player i:*

1. *the utility function is supermodular in the own player's action a_i, has increasing differences in (a_i, a_{-i}), and has increasing differences in (a_i, t); and*
2. *the beliefs mapping $p_i : T_i \to \mathcal{M}_i$ is increasing in the first-order stochastic dominance partial order.*

Then there exist a greatest and a least Bayesian Nash equilibrium, and each one is in monotone strategies.

Despite these compelling reasons, the conjecture that Property V is sufficient for equilibrium existence in auctions is actually false; the following theorem clarifies that SMPSE existence does not generalize beyond Property VI.[36]

Theorem 4.1 *If $f : [0,1]^2 \to \mathbb{R}$ satisfies Property VI′, there is a SMPSE. Nevertheless, Property V is not sufficient for the existence of SMPSE.*

This theorem shows that the rationale for existence of SMPSE existence for Property V do not survive a formalization of the result. Simply, Property V is not strong enough to control the equilibrium inequality $\Pi(x,x) \geqslant \Pi(x,y)$ for every pair of points (x,y).

4.2 Revenue ranking

The next implication—$R_2 \geqslant R_1$—is also an inequality, but it is an inequality over expected values, not specific realizations. For some realizations of the variables, the second-price auction can give less revenue than the first-price auction, but for the inequality $R_2 \geqslant R_1$ to be true is sufficient that the opposite happens on *average*. Since this is a statement about average cases, one could expect that the revenue ranking $R_2 \geqslant R_1$ would be stable across the cases of positive dependence.

There is yet another way of reaching the same conclusion: it is the intuition for the revenue ranking $R_2 \geqslant R_1$, which is a contribution of Klemperer (2004, p.48-9):

> [In a first-price auction, a] player with value $v + dv$ who makes the same bid as a player with a value of v will pay the same price as a player with a value of v when she wins, but because of affiliation she will expect to win a bit less often [than in the case of independence]. That is, her higher signal makes her think her

[36] van Zandt and Vives (2007)'s main result does not apply because even simple auctions with 2 players and private-values do not satisfy one of their assumptions (increasing differences). In fact, if $t_i > a'_j > a'_i > a_j > a_i$ then $(t_i - a'_i)1_{[a'_i > a'_j]} - (t_i - a'_i)1_{[a'_i > a_j]} = -(t_i - a'_i) < 0$ while $(t_i - a_i)1_{[a_i > a'_j]} - (t_i - a_i)1_{[a_i > a_j]} = 0$, to the contrary of the increasing differences requirement.

competitors are also likely to have higher signals, which is bad for her expected profits.

But things are even worse in a second-price afilliated private-values auction for the buyer. Not only does her probability of winning diminish, as in the first-price auction, but her costs per victory are higher. This is because affiliation implies that contingent on her winning the auction, the higher her value the higher expected second-highest value which is the price she has to pay. Because the person with the highest value will win in either type of auction they are both equally efficient, and therefore the higher consumer surplus in first-price auction implies higher seller revenue in the second-price auction.

This intuition appeals mainly to the notion of positive dependence. Thus, the intuition should lead us to believe that the revenue ranking is still valid under other definitions of positive dependence. Despite these intuitive arguments, however, the following theorem shows that the implication $R_2 \geqslant R_1$ is not robust for weaker definitions of positive dependence.

Theorem 4.2 *If f satisfies Property VI' (see definition above), then the second-price auction gives greater revenue than the first-price auction $(R_2 \geqslant R_1)$. Specifically, the revenue difference is given by*

$$n \int_0^1 \int_0^x b^{'}(y) \left[\frac{F(y|y)}{f(y|y)} - \frac{F(y|x)}{f(y|x)} \right] f(y|x)\, dy \cdot f(x)\, dx$$

where n is the number of players and $b(\cdot)$ is the first-price equilibrium bidding function, or by

$$n \int_0^1 \int_0^x \left[\int_0^y L(\alpha|y)\, d\alpha \right] \cdot \left[1 - \frac{F(y|x)}{f(y|x)} \cdot \frac{f(y|y)}{F(y|y)} \right] \cdot f(y|x)\, dy \cdot f(x)\, dx, \tag{2}$$

where $L(\alpha|t) = \exp\left[- \int_\alpha^t \frac{f(s|s)}{F(s|s)} ds \right]$.

More importantly, Property V is not *sufficient for this revenue ranking.*

Interestingly, the empirical literature has tested affiliation's implication that the English auction gives higher revenue than the first-price auction, but there is no clear confirmation of this prediction.[37]

4.3 Discussion

The results of this section are essentially negative: affiliation's implications are not robust. This puts in doubt the comparative statics conclusion mentioned in the introduction. Namely, Milgrom and Weber (1982a)'s results and intuition suggest that when the assumption of independence is relaxed in the direction of positive dependence, equilibrium is not a problem and the revenue superiority of the English auction (and second-price auction) increases. Section 3

[37] See Laffont (1997).

shows that there is a gap between affiliation and positive dependence and Theorems 4.1 and 4.2 show that this gap is important for the two implications analyzed, even in the narrow setup considered here (private values auctions with risk neutrality and symmetry).[38] In sum, it may be the case that affiliation is providing us with misleading indications.

It should be emphasized that there is no mathematical problem with earlier contributions. However, as economists we see in our results more than just their mathematical content.[39] This methodological position is motivated by the hope that our results tell us something about economic phenomena. The results in this section are relevant to clarify the limitations of this hope with respect to affiliation.[40]

Of course, a skeptical reader may not view economic results in this way. Even in this case, the results of this section still provide technical contributions: a counterexample for equilibrium existence in auctions if just Property V (first-order stochastic dominance) is satisfied (Theorem 4.1); a new proof of the revenue ranking of auctions ($R_2 \geqslant R_1$), a generalization of this ranking to Property VI and a respective counterexample if only Property V is satisfied (Theorem 4.2).

5 Alternative approach to the study of dependence

In this section, we consider an alternative approach to the study of dependence, which is divided into two parts. Each part contains results that are useful and informative by themselves. We collect them here because together they offer a general approach to the the dependence problem.

The first part (section 5.1 below) shows that there always exists a *minimally informative* random variable that makes any set of random variables conditionally independent. Theorem 5.2 also provides a classification of the cases of statistical dependence where the dependence matters and where it does not matter (all the results that hold under independence can be used). As we argue in that subsection, this result suggests that independence may be actually less restrictive that appears at first sight. However, of course there are cases where dependence

[38]The fact that we worked with the restrictive symmetric risk neutral private values setup is not motivated by the difficulty in making these observations in a broader context. On the contrary, these results would be *easier* to obtain in more general settings. Our assumptions are restrictive exactly to show that affiliation's implications are not robust even in a particular case.

[39]For a useful illustration of this difference, consider again the revenue equivalence theorem (RET) and the revenue ranking of auctions under affiliation (RRA), i.e., the fact that the second-price auction gives expected revenue than the second-price auction ($R_2 \geqslant R_1$). From a mathematical point of view, RRA *requires less* (affiliation instead of independence) than RET, but also *says less* ($R_2 \geqslant R_1$ instead of $R_2 = R_1$). Usually, an economist does not compare RET and RRA in this way (requires less, says less). What is understood by the contribution of RRA with respect to the RET is a kind of "comparative statics" result, in the sense that positive dependence (affiliation) points out in the direction of increasing revenue advantage by the second-price auction, as discussed above.

[40]A qualification is necessary, though: here we only proved that other (yet strong) definitions of positive dependence are not sufficient for affiliation's implication. In principle, it would still be possible that affiliation's implications were *typically* true in the class of all distributions with positive dependence. However, the results presented in section 5.2 indicate that this weaker claim is also not true; affiliation's implications are *typically false* in the universe of all positive dependent distributions.

is truly important. For those cases, the second part proposes the use of grid distributions. This method is of great generality, but still tractable. The main results regarding this approach are described here, but they are developed more extensively in a separate paper (de Castro, 2008).

5.1 Conditional Independence

Consider the following definition:

Definition 5.1 (Conditional splitter) *Let $(\Omega, \Sigma, \Pr)$ be a probabilistic space, such that Ω is a Polish (complete separable metrizable) space. Given σ-fields $\mathcal{F}^i \subset \Sigma$, $i = 1, ..., n$, we say that $\mathcal{F}$ conditionally splits (or $\mathcal{F}$ is a conditional splitter of) $\mathcal{F}^1$, ..., $\mathcal{F}^n$ if $\mathcal{F}^1$, ..., $\mathcal{F}^n$ are conditionally independent given $\mathcal{F}$. We say that a variable Z conditionally splits variables $X^1, ..., X^n$ if the σ-field generated by Z, denoted $\sigma(Z)$, conditionally splits $\sigma(X^1), ..., \sigma(X^n)$.*[41]

It is useful to observe that the information content of a variable Z must refer to the σ-field $\sigma(Z)$ associated with it, and not with the variable's value. For instance, it is clear that the variable $Y = 2Z$ contains exactly the same information as Z does and, as natural, $\sigma(Y) = \sigma(Z)$ but $Y \neq Z$. The following theorem, whose statement and proof are slightly informal (a more formal statement and proof are given in appendix B), is the main result of this section.

Theorem 5.2 *Consider a game of asymmetric information with n players, such that each player $i = 1, ..., n$ receives a random variable (type) $t_i \in T_i$ and there is a joint distribution on all types. Then one of the following alternatives happens:*

1. *the types are statistically independent;*
2. *the types are statistically dependent, but when each player receives his type, it becomes common knowledge that the (conditional) beliefs are independent;*
3. *the types are statistically dependent, but there is a conditional splitter of the types.*[42] *Moreover:*
   - (a) *any conditional splitter contains strictly more information that it is common knowledge for the players;*[43]

[41] The σ-field generated by a r.v. $Z : \Omega \to \mathbb{R}$ is defined by $\sigma(Z) = Z^{-1}(\mathcal{B})$, where $\mathcal{B}$ is the Borel σ-field in $\mathbb{R}$.

[42] Thus, if players become aware of the outcome of this conditional splitter, case 3 is converted into case 2.

[43] Note that this statement requires proof, as it could in principle be the case that a random variable with information completely different from the common knowledge information is a conditional splitter. Note also a subtle aspect of this statement: if Z is a conditional splitter and C is a variable representing the common knowledge information, statement 3(a) says only that $\sigma(Z) \supset \sigma(C)$. It may happen that there exists a variable Y, with $\sigma(Y) \not\subset \sigma(C)$ such that if the players are informed of Y, the types will be conditionally independent. This only means that $\sigma(Y, C)$ is a conditional splitter and this does not contradict statement 3(a), since obviously $\sigma(Y, C) \supset \sigma(C)$.

(b) there exists a conditional splitter that is minimally informative (in the sense of inclusion by σ-fields);[44]

(c) if the support of types is a finite set, there is an algorithm to find all the minimally informative conditional splitters Z;[45]

(d) in general, there is not a least informative conditional splitter in the sense of inclusion by σ-fields;[46]

(e) nevertheless, in the finite case there is a least informative conditional splitter in the sense of proximity of conditional beliefs.[47]

Case 1 in Theorem 5.2 is familiar and requires no comments. Cases 2 and 3 seem to require less and deliver more than de Finetti's theorem. While de Finetti's theorem requires exchangeability and an infinite number of random variables (as we showed in example 3.4), case 3 in Theorem 5.2 covers the case of a finite number of random variables that are not necessarily exchangeable and states the existence of a variable that makes all types conditionally independent. More important, at first glance, this existence seems to be in contradiction with example 3.4. The contradiction is, of course, only apparent. The difference between the two settings is that Theorem 5.2 does not assume nor does it deliver symmetric distributions, while de Finetti's theorem requires the conditional distribution to be symmetric (exchangeable) but also delivers identical (i.i.d.) distributions. Therefore, Theorem 5.2 is incomparable to de Finetti's theorem.[48] The following example illustrates this matter.

Example 5.3 *(Example 3.4 continued.)*

Consider the same distribution described in Example 3.4 and fix $\varepsilon = \frac{1}{6} < \frac{1}{4}$. As stated in Example 3.4, there is no variable Z such that $X|Z$ and $Y|Z$ are conditionally independent *and* symmetric. However, consider a variable $Z \in \{0, 1\}$ such that the joint distribution with X

[44] By "minimally informative" we mean the following: if Y is another variable that makes the types conditionally independent, and Y contains less information than Z (i.e., $\sigma(Y) \subset \sigma(Z)$), then Y contains as much information as Z ($\sigma(Y) = \sigma(Z)$). Note that the existence of minimally informative variables is not trivial: conditional independence is not preserved under the intersection of σ-fields.

[45] Naturally, our algorithm produces σ-fields, not exactly specific variables.

[46] By *least informative conditional splitter* we mean the following: if Y is another conditional splitter, then $\sigma(Y) \supset \sigma(Z)$. This least informative conditional splitter would contain strictly more information that the common knowledge, unless we are in case 2 instead of case 3 of this Theorem.

[47] This statement is inaccurate. We represent conditional expectations by the associated Markov transitions and observe that they can be seen as functions in L^2. Therefore, we can show that there is a unique conditional splitter that is the closest one (in the standard L^2 norm) to the conditional expectation given the common knowledge information. A formal description of our notion is only possible after a number of technical definitions, which we prefer to postpone to the appendix B. There are good reasons for using this definition, but, of course, it can be disputed. Our objective here is just to show that it is possible to give a reasonable definition of "least informative" that allows to obtain existence. Since this is not a central point for this paper's objective, we will not discuss this further.

[48] Theorem 5.7 below provides necessary and sufficient condition for the conclusion of de Finetti's theorem with a finite number of variables in a setting that covers examples 3.4 and 5.3.

and Y is given by:

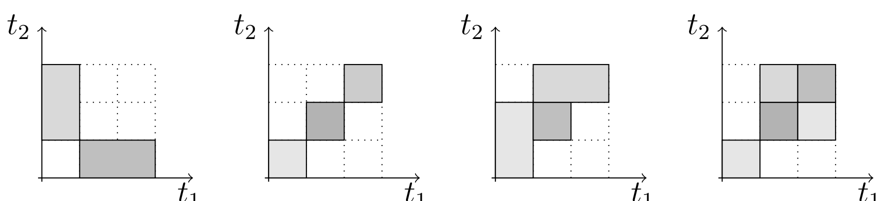
$$\begin{bmatrix} a_{000} & a_{010} & a_{001} & a_{011} \\ \\ a_{100} & a_{110} & a_{101} & a_{111} \end{bmatrix} = \begin{bmatrix} \frac{3}{48} & \frac{1}{48} & \frac{5}{48} & \frac{15}{48} \\ \\ \frac{15}{48} & \frac{5}{48} & \frac{1}{48} & \frac{3}{48} \end{bmatrix},$$

where $a_{ijk} = \Pr(X = i, Y = j, Z = k)$, for $(i, j, k) \in \{0, 1\}^3$. If we write $a_{ij|k}$ for $\Pr(X = i, Y = j|Z = k)$, for $(i, j, k) \in \{0, 1\}^3$, we obtain:

$$\begin{bmatrix} a_{00|0} & a_{01|0} \\ \\ a_{10|0} & a_{11|0} \end{bmatrix} = \begin{bmatrix} \frac{3}{24} & \frac{1}{24} \\ \\ \frac{15}{24} & \frac{5}{24} \end{bmatrix}; \begin{bmatrix} a_{00|1} & a_{01|1} \\ \\ a_{10|1} & a_{11|1} \end{bmatrix} = \begin{bmatrix} \frac{5}{24} & \frac{15}{24} \\ \\ \frac{1}{24} & \frac{3}{24} \end{bmatrix}. \tag{3}$$

Note that $\Pr(X = 0|Z = 0) = \frac{1}{6}$ and that $\Pr(Y = 0|Z = 0) = \frac{3}{4}$. Therefore, $\Pr(X = 0, Y = 0|Z = 0) = \frac{3}{24} = \Pr(X = 0|Z = 0) \Pr(Y = 0|Z = 0)$. A similar verification can be done to all other cases, showing that $X|Z$ is independent of $Y|Z$ and, consequently, Z makes X and Y conditionally independent (is a conditional splitter). Note, however, that the claim in example 3.4 is not violated, since the conditional distributions in (3) are not symmetric.

Figure 2 below illustrates cases 2 and 3 in Theorem 5.2. There are two players with continuous types t_1 and t_2 whose support is the union of the rectangles showed in Figure 2. The random variable Z indicates the rectangle that contains the realization of t_1 and t_2; the types are conditionally independent given the rectangle (which occurs, for instance, if the distribution is uniform in each rectangle). The darker the rectangle, the bigger the probability of that rectangle. Note that the information contained in the variable Z is common knowledge in case 2, and it is not completely informative (does not imply the knowledge of the other player's type) in case 3. Note also that even in case 3, it is possible that some realization of Z is common knowledge, as shown in Figure 2 (d).



(a) Case 2 (b) Case 2 (c) Case 3 (d) Case 3

Figure 2 — Examples of cases 2 and 3 in Theorem 5.2.

Although case 2 seems special, it is not possible to say, from a theoretical point of view, whether it is typical or not in economic applications. An example will illustrate this claim. Suppose that an econometrician observes data on wine auctions. Analyzing the data, the econometrician observes that the bids (and therefore the values, assuming symmetry and affiliation) are extremely positively correlated. However, the vintage, the producer (and in some cases, the previous prices) of that wine are common knowledge to market participants. If the econometrician also has this common knowledge information and controls (conditions) on it, this large correlation will reduce and maybe even disappear. If it disappears, this corresponds exactly to case 2.

The proof of Theorem 5.2 also gives conditions that characterize the three cases in its statement (see lemma B.16). Although the proof is too long to describe here, it is useful to state separately a step in the proof that may be of interest by itself.

Proposition 5.4 *Let $\mathcal{F}^1, ..., \mathcal{F}^n$ be the sub-σ-fields of Σ. There exists a σ-field $\mathcal{Z} \subset \Sigma$ such that $\mathcal{F}^1, ..., \mathcal{F}^n$ are conditionally independent given $\mathcal{Z}$. More specifically, given random variables X^1, ..., X^n, then there exists a random variable Z that makes them conditionally independent.*

While Proposition 5.4 only states the existence of a random variable that makes the types conditionally independent, Theorem 5.2 deals also with *minimally informative* variables. Thus, Proposition 5.4 can be proven with a complete informative variable Z, i.e., a variable that contains all the information about X^1, ..., X^n. That this variable exists may not seem completely obvious. In particular, since $X = (X^1, ..., X^n)$ is n-dimensional, one may be confused by the fact that all information in a model with multidimensional information can be summarized by a one-dimensional variable (Z). As it turns out, the information contained in any vector of random variable (with values in $\mathbb{R}^n$) can be summarized by a single dimensional random variable.[49] The following (coding) argument can be instructive.

Assume, without loss of generality, that X^i is in $[0, 1]$. For each $i = 1, 2, ...$, write X^i as $0.X_{i1}X_{i2}X_{i3}...$. From this, define Z as the r.v. in $[0, 1]$ whose realization is given by:

$$Z = 0.X_{11}X_{21}...X_{n1}X_{12}X_{22}...X_{n2}X_{13}X_{23}...$$

Recall that two random variables X and Y are conditional independent given Z if and only if the additional knowledge of Y does not improve the assessment of X once one knows V, that is, $\Pr(X|Z) = \Pr(X|Y, Z)$. It is easy to see that Z contains all the information that $(X^1, X^2, ..., X^n)$ contains, that is, $\Pr(X^i|Z) = \Pr(X^i|X^j, Z)$.

Unfortunately, however, this "coding argument" *is not* a formal proof because it considers conditioning with respect to null events ($Z = z$). As it is well-known, a number of paradoxes may arise from this kind of procedure. See, for example, Billingsley (1995, Exercise 33.1, p. 441) and the following quote: "There are pathological examples showing that the interpretation of conditional probabilities in terms of an observer with partial information breaks down in certain cases." (Billingsley (1995, p. 437)) This "proof" is useful, however, from an intuitive point of view, because it appeals directly to the notion of information. Consider the following quote from Billingsley (1995, p. 58-9): "The heuristic equating of σ-field and information is helpful even though it sometimes breaks down, and of course proofs are indifferent to whatever illusions and vagaries brought them into existence."

[49]This observation is interesting by itself, because it suggests that the gap between the results obtained for one-dimensional and multidimensional information models *is not* due to information complexity, but rather to techniques employed to obtain those results, such as techniques based on order, monotonicity, calculus, etc.

The confusion about the use of de Finetti's theorem to deliver conditional independence (discussed after example 3.4) shows that the statement and proof of Proposition 5.4 is already helpful.[50] The result itself, however, is not totally useful from a practical point of view, because it is based on a fully informative random variable. This fact highlights the minimal informativeness established by Theorem 5.2. This result is of economic importance, because the less information is sufficient to make bidders conditionally independent, the easier it is for bidders to acquire it. And if conditional independence turns out to be common knowledge, as in case 2, then dependence does not matter: we can apply any result that is valid for independent variables.

Lemma 5.5 *Any result that holds with independent types also holds with the statistically dependent variables in case 2.*[51]

Proof. Although the types in case 2 are statistically dependent, they are independent given the common knowledge information observed in the interim stage. Following the arguments in Harsanyi (1967-8), this interim stage is exactly the realistic stage from where we construct the ex ante stage (and impose the common prior assumption). For the given realization of types in the interim stage, it is not necessary that the ex ante stage be the original one: it can just be the event whose occurrence is common knowledge. In that event, the variables are independent. The result follows. ∎

An immediate corollary of Lemma 5.5 is:

Corollary 5.6 *The Revenue Equivalence Theorem (RET) holds for dependent types in case 2 (provided the other assumptions of this theorem also hold).*

In particular, the variables illustrated in Figure 2 (b) are truly affiliated, but the RET holds for this case. Thus, it is important for the study of dependence in economics to be able to distinguish when we are in case 2. For econometric applications, this seems a rather difficult task, because one has to know what is common knowledge among the participants, and this piece of information may be unobservable to the econometrician. In the wine auctions example given above, it is possible that more than just the vintage is common knowledge; for example, some brochure could have been distributed or some explanation (information) about the object may have been given at the the time of the auction, which the econometrician is not aware and/or cannot use as a control.

Even if the situation corresponds to case 3 instead of case 2, it is possible that Z can be learned with the repetition of the game. Alternatively, the information contained in Z may be available for acquisition (from a consultant or a spy, for instance). If Z becomes common knowledge, case 3 will reduce to case 2. Thus statistical dependence is only relevant in case

[50] Although the result contained in Proposition 5.4 is certainly known by specialists, we were unable to find a good reference for it. See remark B.28 in appendix B for a review of results related to Theorem 5.2 and Proposition 5.4.

[51] This statement is slightly informal, but its content should be clear: the conclusions hold conditionally to each piece of common knowledge information. For instance, we can have monotone pure strategy equilibria in each common knowledge part, but this does not imply that the equilibria will be overall monotonic.

3 with a Z that cannot be (is not) learned. Even if a conditional splitter cannot be learned, it may happen that after conditioning to the common knowledge, the remaining correlation is so small that independence already gives a good approximation, since revenue and equilibria vary continuously with the distribution.[52] In sum, independence can already give a good approximation in many cases of interest. However, if this is true or not is an empirical/experimental question, not a theoretical one. The following summarizes the discussion in this section:

1. There is always a random variable (r.v.) Z that makes the types of the bidders conditionally independent.
2. If this r.v. is common knowledge, then the appropriate framework is that of independent types models.
3. If this r.v. can be learned with time, then the game will converge to the previous situation.
4. Only if the r.v. cannot become common knowledge, then a model with dependence is truly necessary.

   Therefore:

5. It is important to test whether or not independence is a good approximation in economically relevant environments.

5.1.1 De Finetti's Theorem for a finite number of random variables

It was noted earlier that Theorem 5.2 ensures the existence of conditional independence but it does not deliver de Finetti's Theorem implication that the conditional probabilities are *identical*, that is, symmetric. Although the implication given in Theorem 5.2 is sufficient for our purposes, in some applications one may be interested in the stronger implication, which we will call *symmetric conditional independence*. This terminology means that the variables are not only conditionally independent given Z, but also symmetrically (identically) distributed. The full exploration of this question is beyond the scope of this paper, but we present here a theorem in a simple setting in which the question can be totally clarified.

Theorem 5.7 *Let X and Y be symmetric (exchangeable) binary random variables that are statistically dependent. Then there is a binary random variable Z that makes X and Y symmetrically conditionally independent if and only if X and Y are positively correlated.*

Theorem 5.7 generalizes the claim made in Example 3.4. It should be noted that Diaconis (1977) and Diaconis and Freedman (1980) have previously studied versions of de Finetti's theorem for a finite set of variables, but these papers do not present any result comparable

[52]This statement can be formalized by using the sup-norm in the space of continuous densities. This observation is sufficiently simple and intuitive, so that we do not pursue its formalization. See, however, items 3 and 4 in Theorem 5.8 below for related results.

to Theorem 5.7. They are interested in how the conclusion of de Finetti's theorem is *asymptotically* true for a finite set of exchangeable variables when the size of the set increases. In contrast, we are interested in a fixed (and small) set of random variables and look for a (necessary and sufficient) condition for de Finetti's conclusion to hold *exactly*.

5.2 Grid Distributions

If we are in case 3 of Theorem 5.2, then we have to deal with general cases of dependence. The third alternative is designed to be a comprehensible yet tractable approach to general dependence in economic models. In fact, this approach can also be used to study asymmetries in auctions. This approach uses grid distributions to approximate any distribution of types. Grid distributions consist of all distributions whose densities are constant in squares. Figure 3 below represents a grid distribution whose density is constant in four squares that form a partition of its support. This example corresponds to the simple case of two players with values v_1 and v_2, with two intervals for each player.

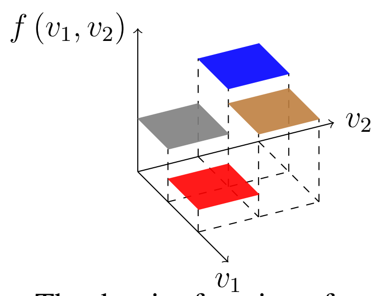


Figure 3 — The density function of a grid distribution.

Many results concerning this class of distribution functions are presented in de Castro (2008). For describing some of those results, let $\mathcal{D}$ be the set of all probability density functions (p.d.f.'s) $f : [0,1]^n \to \mathbb{R}$, representing the joint distribution of types of n players in a game of incomplete information (more specifically, an auction). Also, let $\mathcal{D}^k$ represent the set of grid distributions $f \in \mathcal{D}$, where the interval $[0,1]$ is divided in k sub-intervals and let $\mathcal{D}^\infty \equiv \cup_{k\in\mathbb{N}}\mathcal{D}^k$ be the set of all grid distributions (with arbitrary number of intervals).

Theorem 5.8 *Consider first-price auctions with n players, with private values in $[0,1]$.*[53] *Then:*

1. *The set of grid distributions $\mathcal{D}^\infty$ is dense in the set of all p.d.f.'s $\mathcal{D}$. More specifically:*
2. *Let $T^k : \mathcal{D} \to \mathcal{D}^k$ denote the projection of $\mathcal{D}$ over $\mathcal{D}^k$. Then, $T^k(f) \to f$ in the strong sense.*[54]

[53] The first four claims are not restricted to first-price auctions.

[54] By strong, we mean in the L^1-norm. If f is continuous, the convergence can be established also in the supnorm.

3. *If $f \in \mathcal{D}$ is continuous and symmetric, and there exists k_0 such that all auction with distribution $T^k(f)$ have a monotonic pure strategy equilibrium (MPSE) for all $k \geqslant k_0$, then f also has a MPSE. Moreover, the equilibria for f is the limit of the equilibria of $T^k(f)$ as $k \to \infty$.*

4. *Conversely, if f has a MPSE, then for each $\epsilon > 0$, there exists k_ϵ such that $T^k(f)$ has a MPS ϵ-equilibrium for all $k \geqslant k_\epsilon$ and which converge to f's equilibrium.*

5. *If $f \in \mathcal{D}^\infty$ is symmetric, there is a unique candidate for a symmetric MPSE (SMPSE); moreover a closed form expression of this bidding function is available;*[55]

6. *If $f \in \mathcal{D}^\infty$ is symmetric and there are only two risk neutral bidders ($n = 2$), there exists an algorithm that decides in finite time whether there is or there is not a symmetric monotonic pure strategy equilibrium for this auction.*[56] *For $f \in \mathcal{D}^k$, the algorithm requires less than $3(k^2 + k)$ comparisons. The algorithm is* exact*, in the sense that there is an equilibrium if and only if the algorithm detects its modulo only errors in elementary operations.*[57]

7. *If $f \in \mathcal{D}^\infty$ is symmetric or asymmetric, there are $n > 2$ bidders with any risk attitude, there is still an algorithm to verify if a MPSE exists or not, although it will not be* exact *as before.*

A more detailed statement of the above results and their proofs can be found in de Castro (2008), together with some other facts about grid distributions. The facts that grid distributions are dense (item 1 in Theorem 5.8 above) and there are sufficient continuity properties (items 3 and 4) show that the study of auctions using grid distributions is meaningful. This method allows the use of simulations to investigate what is the typical situation in the set of all distributions. In particular, this method allows us to verify that there are many cases of equilibrium existence in which the revenue ranking is exactly the opposite of that implied by affiliation. Typically the first-price auction gives higher expected revenue than the second-price auction, even for the cases with positive dependence.

Figure 4 illustrates the general distribution of the relative revenue difference $(R_2^f - R_1^f)/R_2^f$, where R_1^f (R_2^f) denotes the expected revenue of the first (second) price auction when the distribution of types is given by the density function $f \in \mathcal{D}^4$ and there are two risk neutral bidders. As the reader can see, the peak occurs at –5%, while the linkage principle would imply a positive number.[58] The following remarks are in order:

[55] Since grid distributions are not continuous, standard techniques for dealing with the solution of the respective differential equations need adaptation. See de Castro (2008) for details.

[56] The existence of such an algorithm is not trivial, since there are infinitely many types and the equilibrium condition should in principle be checked in infinite many points, which is not feasible.

[57] By elementary operations we mean sums, multiplications, divisions, comparisons and square and third degree roots.

[58] The histogram refers to symmetric functions in $f \in \mathcal{D}^4$ generated at random, and considers only $f \in \mathcal{D}^4$ for which a SMPSE exists for the first-price auction with $n = 2$ players. The algorithm mentioned in point 6 above was implemented to obtain the results.

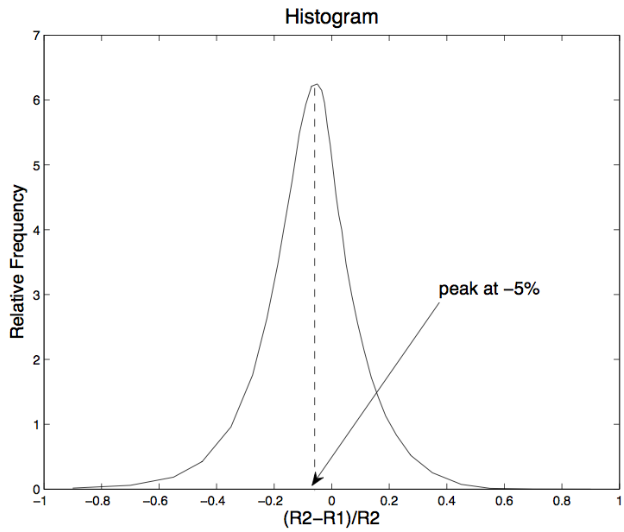


Figure 4: Histogram of $\frac{R_2-R_1}{R_2}$, for $k=4$, with symmetry and risk neutrality. The peak of the relative frequency occurs at -5%.

- The histogram moves to the left (the peak becomes more negative) when k increases.
- The result is qualitatively the same if only positive dependent distributions are considered.
- The result is also qualitatively the same if instead of $n=2$, we consider more players.
- Although we have run 10^7 simulations, the results are already stable for 10^5.
- The histogram is similar (the peak is also negative) if we consider the absolute difference $R_2^f - R_1^f$ instead of the relative $(R_2^f - R_1^f)/R_2^f$.
- The functions $f \in \mathcal{D}^k$ were picked at random (generated by a uniform distribution). Of course, with other distributions over $\mathcal{D}^k$, the histogram may be different. It should be stressed that this is not a problem of the method proposed here; we used the uniform distribution over $\mathcal{D}^k$ for lack of guidance about what kind of distribution would be the more realistic one. If the empirical/experimental literature provides guidance as to what are the typical kinds of distribution, the simulations could be easily adapted.[59]
- The objective of the exercise shown in Figure 4 is to compare the revenue ranking of first-price and second-price auctions under general (positive) dependence with affiliation's

[59]The determination of what are the typical forms of dependence in the real world is a very important contribution that experimental or empirical economists could make.

revenue ranking in the simplest of the settings, but the method can be applied more generally.[60]

An alternative method to approximate general distributions is to use discrete distributions. Nevertheless, the use of discrete values precludes us from using the convenient tools of differential calculus. This advantage is important because differential calculus is much more tractable than finite differences methods. In particular, differential calculus allows for a characterization of equilibrium strategies. Grid distribution combine the benefits of the simplicity of finite values and the tools previously developed.

A possible opposition to the results in this section is that it refers more to simulations than direct proofs, as usual in economic theory. However, the use of grid distributions does not exclude the possibility of proving general theorems. Indeed, there is a one-to-one correspondence between grid distributions in $\mathcal{D}^k$ and sets of k^n random variables in $[0,1]$. In particular, distributions in $\mathcal{D}^\infty$ can be modeled by sequences of independent random variables. Thus, many tools of modern Probability can be used to prove general theorems about the properties of grid distributions. By continuity, such kind of results will be informative about the general case. de Castro (2008) illustrates this idea by *proving* that the set of distributions with pure strategy equilibrium is a small set (it essentially has measure zero) in the set of all distributions. It should be noted also that this author would not be able to anticipate this fact without the simulations that the grid method makes possible.

Another observation is that the simulations are not done for particular parametric examples, as it is usually the case with simulations. The simulations considered here allow the consideration of the generic picture (a dense set), not particular examples. In this sense, they are much more informative and useful than usual simulation exercises.

Another conceivable opposition is that this method is restricted to the analysis of MPSE and it would be desirable to consider mixed equilibria, which are more general. However, the restriction to MPSE is a characteristic of most of the received auction theory literature. Although it is certainly desirable to know what happens with mixed strategy equilibria, the difficulties related to this task are already considerable for games far simpler than auctions. In other words, our focus on MPSE follows the standard practice in the received literature. It seems reasonable to pursue a complete understanding of MPSE's properties before passing to the much more complex case of mixed strategies.

6 Related literature

Few papers have pointed out restrictions or limitations to the implications of affiliation. Perry and Reny (1999) presented an example of a multi-unit auction where the linkage principle fails and the revenue ranking is reversed, even under affiliation. This result shows that revenue ranking is not robust when the number of objects increases from one to many. In contrast, one

[60] However, the algorithm for equilibrium existence is *exact* only in the special case described in item 6 of Theorem 5.8 (see also item 7). We were able to verify, however, that the general algorithm is accurate. See de Castro (2008) for more details.

of our results shows that the revenue ranking is not robust even if we maintain the number of objects but allow for other kinds of dependences. Klemperer (2003) argues that, in real auctions, affiliation is not as important as asymmetry and collusion and illustrates his arguments with examples of the 3G auctions conducted in Europe in 2000–2001.

Nevertheless, much more has been written in accordance with the conclusions of affiliation. McMillan (1994, p.152) says that the auction theorists working as consultants to the FCC in spectrum auctions advocated for the adoption of the open auction using the linkage principle as an argument: "Theory says, then, that the government can increase its revenue by publicizing any available information that affects the licensee's assessed value." The disadvantages of the open format in the presence of risk aversion and collusion were judged "to be outweighed by the bidders' ability to learn from other bids in the auction" (p. 152). Milgrom (1989) emphasizes affiliation as the explanation for the predominance of the English auction over the first-price auction.

On the other hand, the experimental and empirical literature show an amazing lack of studies about whether affiliation holds or not. The available studies investigated only some of the implications of affiliation. See Kagel (1995) for a survey about the experimental literature and Laffont (1997) for a survey of empirical results in auctions. See also footnote 31.

7 Final remarks

This paper began with an assessment of affiliation because this assumption is widely used as a substitute for independence not only in auction theory but also in economics of information in general. This (theoretical) criticism of this paper serves the purpose of motivating the need for new results and new approaches to the study of dependence in economic models, such as those presented in section 5. Although this criticism suggests that affiliation is unlikely to hold in typical situations, it cannot state that it does not hold in general; only empirical studies can assert this.

The criticism is based in a series of results (theorems and examples), but it does not contain a clear and simple thought experiment to show that affiliation is not realistic, as Allais (1953) and Ellsberg (1961) do in their famous criticisms of expected utility. The reason for this absence is that such an example cannot exist in a simple form. This fact comes from Theorem 5.2 and the discussion following it, which show that it is not possible to rule out independence (and, therefore, affiliation) without a control of what is common knowledge for the participants. In other words, it is virtually impossible to assert what kind of dependence is typical from a purely theoretical point of view. This impossibility raises the question of testing affiliation and dependence in more general terms. Again, Theorem 5.2 implies that this test requires special care with characteristics that are unobservable but may be (commonly) known by market participants.

Experimental studies could shed light on the actual distribution of values across individuals, controlling for the common knowledge. It would be very helpful to develop methods to determine the values that people attribute to objects in an auction and whether those values are correlated or not. With respect to econometrics, an obvious need is to develop methods to test the affiliation of bidders' values, controlling for the common knowledge (if this is possible).[61] It would also be useful to develop techniques to describe the kind of dependence of the bids in real auctions. It would be very helpful to learn whether the kind of dependence is different across different markets and how these differences can be characterized. For instance, is there less correlation in Internet auctions, where the participants are consumers with almost no interaction, than in auctions where the participants are firms or professionals acting in the same industry? Yet another direction of research would be the development of econometric techniques to deal with dependence out of affiliation.[62]

In sum, there is much yet to be done to fully understand dependence in economics.

A Proofs

A.1 Proof of Theorem 3.1 .

First, we prove that $C\backslash A$ is open. If $f \in C\backslash A$, then

$$f(x) f(x') > f(x \wedge x') f(x \vee x'),$$

for some $x, x' \in [0,1]^n$. Fix such x and x' and define $K = f(x) + f(x') + f(x \wedge x') + f(x \vee x') > 0$. Choose $\varepsilon > 0$ such that $2\varepsilon K < f(x) f(x') - f(x \wedge x') f(x \vee x')$ and let $B_\varepsilon(f)$ be the open ball with radius ε and centre in f. Thus, if $g \in B_\varepsilon(f)$, $\|f - g\| < \varepsilon$, which implies $g(x) > f(x) - \varepsilon$, $g(x') > f(x') - \varepsilon$, $g(x \wedge x') < f(x \wedge x') + \varepsilon$, $g(x \vee x') < f(x \vee x') + \varepsilon$, so that

$$\begin{aligned}
& g(x) g(x') - g(x \wedge x') g(x \vee x') \\
> \; & [f(x) - \varepsilon][f(x') - \varepsilon] - [f(x \wedge x') + \varepsilon][f(x \vee x') + \varepsilon] \\
= \; & f(x) f(x') - f(x \wedge x') f(x \vee x') - \varepsilon [f(x) + f(x') + f(x \wedge x') + f(x \vee x')] \\
= \; & f(x) f(x') - f(x \wedge x') f(x \vee x') - \varepsilon K \\
> \; & \varepsilon K > 0,
\end{aligned}$$

which implies that $B_\varepsilon(f) \subset C\backslash A$, as we wanted to show.

Now, let us show that $C\backslash A$ is dense, that is, given $f \in C$ and $\varepsilon > 0$, there exists $g \in B_\varepsilon(f) \cap C\backslash A$. Since $f \in C$, it is uniformly continuous (because $[0,1]^n$ is compact), that is, given $\eta > 0$, there exists $\delta > 0$ such that $\|x - x'\|_{\mathbb{R}^n} < 2\delta$ implies $|f(x) - f(x')| < \eta$. Take $\eta = \varepsilon/4$ and the corresponding δ.

Choose $a \in (4\delta, 1 - 4\delta)$ and define x (x') by specifying that their first $\left\lfloor \frac{n}{2} \right\rfloor$ coordinates are equal to $a - \delta$ $(a + \delta)$ and the last ones to be equal to $a + \delta$ $(a - \delta)$. Thus, $x \wedge x' = (a - \delta, ..., a - \delta)$ and $x \vee x'$ = $(a + \delta, ..., a + \delta)$. Let x_0 denote the vector $(a, ..., a)$. For $y = x$, x', $x \wedge x'$ or $x \vee x'$, we have: $|f(y) - f(x_0)| < \eta$. Let $\xi : (-1, 1)^n \to \mathbb{R}$ be a smooth function

[61] See footnote 31.

[62] Grid distributions can be useful for this task. See de Castro and Paarsch (2008).

that vanishes outside $\left(-\frac{\delta}{2}, \frac{\delta}{2}\right)^n$ and equals 1 in $\left(-\frac{\delta}{4}, \frac{\delta}{4}\right)^n$. Define the function g by

$$\begin{aligned} g(y) &= f(y) + 2\eta\xi(y-x) + 2\eta\xi(y-x') \\ &\quad -2\eta\xi(y - x\wedge x') - 2\eta\xi(y - x\vee x'). \end{aligned}$$

Observe that $\|g-f\| = 2\eta = \varepsilon/2$, that is, $g \in B_\varepsilon(f)$. In fact, $g \in B_\varepsilon(f) \cap C\backslash A$, because

$$\begin{aligned} g(x) &= f(x) + 2\eta > f(x_0) + \eta; \\ g(x') &= f(x) + 2\eta > f(x_0) + \eta; \\ g(x\wedge x') &= f(x\wedge x') - 2\eta < f(x_0) - \eta; \\ g(x\vee x') &= f(x\vee x') - 2\eta < f(x_0) - \eta, \end{aligned}$$

which implies

$$\begin{aligned} & g(x)\,g(x') - g(x\wedge x')\,g(x\vee x') \\ > & \left[f(x_0)+\eta\right]^2 - \left[f(x_0)-\eta\right]^2 \\ = & 4\eta > 0, \end{aligned}$$

as we wanted to show. ∎

A.2 Proof of Theorem 3.2.

The proof of Theorem 3.2 is divided in two parts: the implications and the counterexamples.

A.2.1 Implications

It is obvious that $(III) \Rightarrow (II) \Rightarrow (I)$. The implication $(IV) \Rightarrow (III)$ is Theorem 4.3. of Esary, Proschan, and Walkup (1967). The implication $(V) \Rightarrow (IV)$ is proved by Tong (1980, p. 80). The implication $(VII) \Rightarrow (VI)$ is Lemma 1 of Milgrom and Weber (1982a). Thus, we need only to prove $(VI) \Rightarrow (V)$.

For this, assume that $H(y|x) \equiv \frac{f(y|x)}{F(y|x)}$ is non-decreasing in x for all y. Then, $H(y|x) = \partial_y\left[\ln F(y|x)\right]$ and we have

$$1 - \ln\left[F(y|x)\right] = \int_y^\infty H(s|x)\,ds \geqslant \int_y^\infty H(s|x')\,ds = 1 - \ln\left[F(y|x')\right],$$

if $x \geqslant x'$. Then, $\ln\left[F(y|x)\right] \leqslant \ln\left[F(y|x')\right]$, which implies that $F(y|x)$ is non-increasing in x for all y, as required by Property V.

A.2.2 Counterexamples

The counterexamples for each passage are given by Tong (1980, Chapter 5), except those involving Property (VI): $(V) \nRightarrow (VI)$, $(VI) \nRightarrow (VII)$. For the counterexample of $(V) \nRightarrow (VI)$, consider the following symmetric and continuous pdf defined on $[0,1]^2$:

$$f(x,y)=\frac{d}{1+4(y-x)^2}$$

where $d=[\arctan(2)-\ln(5)/4]^{-1}$ is the suitable constant for f to be a pdf We have the marginal given by

$$f(y)=\frac{k}{2}[\arctan 2(1-y)+\arctan 2(y)]$$

so that we have, for $(x,y)\in[0,1]^2$:

$$f(x|y)=2\left[1+4(y-x)^2\right]^{-1}[\arctan 2(1-y)+\arctan 2(y)]^{-1},$$

$$F(x|y)=\frac{[\arctan 2(x-y)+\arctan 2(y)]}{\arctan 2(1-y)+\arctan 2(y)}$$

and

$$\frac{F(x|y)}{f(x|y)}=2\left[1+4(y-x)^2\right][\arctan(2x-2y)+\arctan(2y)].$$

Observe that for $y'=0.91>y=0.9$ and $x=0.1$,

$$\frac{F(x|y')}{f(x|y')}=0.366863>0.366686=\frac{F(x|y)}{f(x|y)},$$

which violates Property (VI). On the other hand,

$$\begin{aligned}\partial_y[F(x|y)] &= \frac{\frac{2}{1+4y^2}-\frac{2}{1+4(x-y)^2}}{\arctan(2-2y)+\arctan(2y)}\\ &\quad -\frac{[\arctan(2x-2y)+\arctan(2y)]\left[\frac{2}{1+4y^2}-\frac{2}{1+4(1-y)^2}\right]}{[\arctan(2-2y)+\arctan(2y)]^2}\end{aligned}$$

In the considered range, the above expression is non-positive, so that Property (V) is satisfied. Then, $(V)\nRightarrow(VI)$.

Now we will establish that $(VI)\nRightarrow(VII)$. Fix an $\varepsilon<1/2$ and consider the symmetric density function over $[0,1]^2$:

$$f(x,y)=\begin{cases}k_1, & \text{if } x+y\leqslant 2-\varepsilon\\ k_2, & \text{otherwise}\end{cases}$$

where $k_1>1>k_2=2\left[1-k_1\left(1-\varepsilon^2/2\right)\right]/\varepsilon^2>0$ and $\varepsilon\in(0,1/2)$. For instance, we could choose $\varepsilon=1/3$, $k_1=19/18$ and $k_2=1/18$. The conditional density function is given by

$$f(y|x)=\begin{cases}1, & \text{if } x\leqslant 1-\varepsilon\\ \frac{k_1}{k_2(x+\varepsilon-1)+k_1(2-\varepsilon-x)}, & \text{if } x>1-\varepsilon \text{ and if } y\leqslant 2-\varepsilon-x\\ \frac{k_2}{k_2(x+\varepsilon-1)+k_1(2-\varepsilon-x)}, & \text{otherwise}\end{cases}$$

and the conditional c.d.f. is given by:

$$F(y|x) = \begin{cases} 1, & \text{if } x \leqslant 1-\varepsilon \\ \frac{k_1 y}{k_2(x+\varepsilon-1)+k_1(2-\varepsilon-x)}, & \text{if } x > 1-\varepsilon \text{ and if } y \leqslant 2-\varepsilon-x \\ \frac{k_2(y+x+\varepsilon-2)+k_1(2-\varepsilon-x)}{k_2(x+\varepsilon-1)+k_1(2-\varepsilon-x)}, & \text{otherwise} \end{cases}$$

and

$$\frac{F(y|x)}{f(y|x)} = \begin{cases} 1, & \text{if } x \leqslant 1-\varepsilon \\ y, & \text{if } x > 1-\varepsilon \text{ and if } y \leqslant 2-\varepsilon-x \\ y+x+\varepsilon-2+k_1/k_2(2-\varepsilon-x), & \text{otherwise} \end{cases}$$

Since $1-k_1/k_2 < 0$, the above expression is non-increasing in x for all y, so that Property (VI) is satisfied. On the other hand, it is obvious that Property (VII) does not hold:

$$f(0.5, 0.5) f\left(1-\frac{\varepsilon}{2}, 1-\frac{\varepsilon}{2}\right) = k_2 k_1 < k_1^2 = f\left(0.5, 1-\frac{\varepsilon}{2}\right) f\left(0.5, 1-\frac{\varepsilon}{2}\right).$$

This shows that $(VI) \nRightarrow (VII)$.

A.3 Proof of Theorem 4.1.

The equilibrium existence follows from Milgrom and Weber (1982a)'s proof. For the counterexample, consider the pdf defined in the proof of Theorem 3.2:

$$f(x,y) = \frac{d}{1+4(y-x)^2},$$

where $d = [\arctan(2) - \ln(5)/4]^{-1}$. In the proof of Theorem 3.2, we established that this pdf satisfies Property V but not Property VI and that:

$$F(x|y) = \frac{[\arctan 2(x-y) + \arctan 2(y)]}{\arctan 2(1-y) + \arctan 2(y)}.$$

If there is a SMPSE, it has to be given by the following expression:[63]

$$b(y) = y - \int_0^y \exp\left[-\frac{1}{2}\int_z^y \frac{1}{\arctan 2w} dw\right] dz$$

cannot be an equilibrium, that is, to verify the existence of x and y such that

$$(y - b(y)) F(y|y) < (y - b(x)) F(x|y).$$

[63] This is the solution of the standard differential equation. For a careful proof of this intuitive fact, see de Castro (2008).

This simplifies to the condition:

$$\frac{\int_0^y \exp\left[-\frac{1}{2}\int_z^y \frac{1}{\arctan 2w} dw\right] dz}{y - x + \int_0^x \exp\left[-\frac{1}{2}\int_z^x \frac{1}{\arctan 2w} dw\right] dz} < \frac{\arctan 2\left(x-y\right)}{\arctan 2y} + 1.$$

Let $y = 0.5$ and $x = 1$. Mathematica gives $\int_0^y \exp\left[-\frac{1}{2}\int_z^y \frac{1}{\arctan 2w} dw\right] dz = 0.391128$ and $\int_0^x \exp\left[-\frac{1}{2}\int_z^x \frac{1}{\arctan 2w} dw\right] dz = 0.745072$. Thus, we have:

$$\frac{0.391128}{-0.5 + 0.745072} = 1.59597 < 2 = \frac{\arctan 2\left(x-y\right)}{\arctan 2y} + 1,$$

which concludes the verification for the counterexample of SMPSE existence.

A.4 Proof of Theorem 4.2.

The dominant strategy for each bidder in the second-price auction is to bid his value: $b^2\left(t\right) = t$. Then, the expected payment by a bidder in the second-price auction, P^2, is given by:

$$\begin{aligned} P^2 &= \int_{\left[\underline{t},\bar{t}\right]} \int_{\left[\underline{t},x\right]} y f\left(y|x\right) dy \cdot f\left(x\right) dx = \\ &= \int_{\left[\underline{t},\bar{t}\right]} \int_{\left[\underline{t},x\right]} \left[y - b\left(y\right)\right] f\left(y|x\right) dy \cdot f\left(x\right) dx + \int_{\left[\underline{t},\bar{t}\right]} \int_{\left[\underline{t},x\right]} b\left(y\right) f\left(y|x\right) dy \cdot f\left(x\right) dx, \end{aligned}$$

where $b\left(\cdot\right)$ gives the equilibrium strategy for symmetric first-price auctions. Thus, the first integral can be substituted by $\int_{\left[\underline{t},\bar{t}\right]} \int_{\left[\underline{t},x\right]} b'\left(y\right) \frac{F(y|y)}{f(y|y)} f\left(y|x\right) dy \cdot f\left(x\right) dx$, from the first-order condition: $b'\left(y\right) = \left[y - b\left(y\right)\right] \frac{f(y|y)}{F(y|y)}$. The last integral can be integrated by parts, to:

$$\begin{aligned} & \int_{\left[\underline{t},\bar{t}\right]} \int_{\left[\underline{t},x\right]} b\left(y\right) f\left(y|x\right) dy \cdot f\left(x\right) dx \\ = & \int_{\left[\underline{t},\bar{t}\right]} \left[b\left(x\right) F\left(x|x\right) - \int_{\left[\underline{t},x\right]} b'\left(y\right) F\left(y|x\right) dy\right] \cdot f\left(x\right) dx \\ = & \int_{\left[\underline{t},\bar{t}\right]} b\left(x\right) F\left(x|x\right) \cdot f\left(x\right) dx - \int_{\left[\underline{t},\bar{t}\right]} \int_{\left[\underline{t},x\right]} b'\left(y\right) F\left(y|x\right) dy \cdot f\left(x\right) dx \end{aligned}$$

In the last line, the first integral is just the expected payment for the first-price auction, P^1. Thus, we have

$$
\begin{aligned}
D &= P^2 - P^1 \\
&= \int_{[\underline{t},\bar{t}]} \int_{[\underline{t},x]} b^{'}(y) \frac{F(y|y)}{f(y|y)} f(y|x)\, dy \cdot f(x)\, dx \\
&\quad - \int_{[\underline{t},\bar{t}]} \int_{[\underline{t},x]} b^{'}(y)\, F(y|x)\, dy \cdot f(x)\, dx \\
&= \int_{[\underline{t},\bar{t}]} \int_{[\underline{t},x]} b^{'}(y) \left[\frac{F(y|y)}{f(y|y)} f(y|x) - F(y|x) \right] dy \cdot f(x)\, dx \\
&= \int_{[\underline{t},\bar{t}]} \int_{[\underline{t},x]} b^{'}(y) \left[\frac{F(y|y)}{f(y|y)} - \frac{F(y|x)}{f(y|x)} \right] f(y|x)\, dy \cdot f(x)\, dx
\end{aligned}
$$

Remember that $b(t) = \int_{[\underline{t},t]} \alpha dL(\alpha|t) = t - \int_{[\underline{t},t]} L(\alpha|t)\, d\alpha$, where $L(\alpha|t) = \exp\left[-\int_{\alpha}^{t} \frac{f(s|s)}{F(s|s)} ds \right]$. So, we have

$$
\begin{aligned}
b^{'}(y) &= 1 - L(y|y) - \int_{[\underline{t},y]} \partial_y L(\alpha|y)\, d\alpha \\
&= \frac{f(y|y)}{F(y|y)} \int_{[\underline{t},y]} L(\alpha|y)\, d\alpha.
\end{aligned}
$$

We conclude that

$$
\begin{aligned}
D &= \int_{[\underline{t},\bar{t}]} \int_{[\underline{t},x]} \frac{f(y|y)}{F(y|y)} \int_{[\underline{t},y]} L(\alpha|y)\, d\alpha \left[\frac{F(y|y)}{f(y|y)} - \frac{F(y|x)}{f(y|x)} \right] f(y|x)\, dy \cdot f(x)\, dx \\
&= \int_{[\underline{t},\bar{t}]} \int_{[\underline{t},x]} \left[\int_{[\underline{t},y]} L(\alpha|y)\, d\alpha \right] \cdot \left[1 - \frac{F(y|x)}{f(y|x)} \cdot \frac{f(y|y)}{F(y|y)} \right] \cdot f(y|x)\, dy \cdot f(x)\, dx,
\end{aligned}
$$

which is the desired expression if we multiply by the number n of players.

For the counterexample, consider the grid distribution $f : [0,1]^2 \to \mathbb{R}_+, f \in \mathcal{D}^4$ (see section 5.2), defined by:

$$
f(x,y) = a_{mp} \text{ if } (x,y) \in \left(\frac{m-1}{k}, \frac{m}{k} \right] \times \left(\frac{p-1}{k}, \frac{p}{k} \right],
$$

for $m, p \in \{1,2,3,4\}$, where

$$
\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} = \begin{bmatrix} 2.7468 & 0.0803 & 0.1195 & 0.0696 \\ 0.0803 & 0.3200 & 0.5271 & 0.1224 \\ 0.1195 & 0.5271 & 1.7814 & 0.5650 \\ 0.0696 & 0.1224 & 0.5650 & 1.2705 \end{bmatrix}.
$$

The definition of f at the zero measure set of points $\{(x,y) = \left(\frac{m}{k}, \frac{p}{k}\right) : m = 0 \text{ or } p = 0\}$ is arbitrary. This distribution satisfies Property V (but not Property VI). Moreover, the first-price auction with this distribution has a SMPSE and a higher expected revenue than the correspondent second-price auction ($R_2 < R_1$). These claims can be verified with the methodology developed by de Castro (2008) and discussed in section 5.2.

A.5 Proof of Example 3.4.

Although this is implied by Theorem 5.7, we give here a direct proof for this example. Let p denote the probability of Heads and let μ be a distribution over coins. Then: $\Pr(\text{Heads,Heads}) = \varepsilon = \int (p)^2 \mu(dp)$, and $\Pr(\text{Tails,Tails}) = \varepsilon = \int (1-p)^2 \mu(dp) = \int (1)\mu(dp) + \int (-2p)\mu(dp) + \int (p)^2 \mu(dp) = 1 - 2E[p] + \varepsilon$. Then, $1 - 2E[p] = 0$, or $E[p] = 1/2$. This implies: Var[p]$= \int (p - E[p])^2 \mu(dp) = \int (p^2 - p + \frac{1}{4})\mu(dp) = \int (p)^2 \mu(dp) - \frac{1}{4} = \varepsilon - \frac{1}{4}$. Since Var[p] is non-negative, $\varepsilon \geqslant \frac{1}{4}$. ∎

B Conditional Independence

This appendix develops the concepts and tools necessary to the proofs of results in section 5.1. This appendix is organized in subsections for clarity and didactic reasons. We begin with some notation.

B.1 Preliminaries

Let $(\Omega, \Sigma, \Pr)$ be a probability space, where Ω is a Polish (separable complete metrizable) space and Σ, its Borel σ-field, unless otherwise specified. Let $\mathcal{B}$ denote the Borel σ-field in $\mathbb{R}$ (the set of real numbers). Given a σ-field $\mathcal{F} \subset \Sigma$, a random variable Y is $\mathcal{F}$-measurable if $Y^{-1}(B) \equiv \{\omega : Y(\omega) \in B\} \in \mathcal{F}$ for every $B \in \mathcal{B}$. Let $\sigma(Y)$ denote the smallest σ-field with respect to which Y is measurable. If $\mathcal{C}$ is a class of sets, $\sigma(\mathcal{C})$ denotes the σ-field generated by $\mathcal{C}$, that is, the smallest σ-field containing $\mathcal{C}$. Let Σ° denote the set of Σ-measurable null sets, that is, $\Sigma^\circ \equiv \{A \in \Sigma : \Pr(A) = 0\}$. The completion of $\mathcal{F} \subset \Sigma$ is $\bar{\mathcal{F}}$ given by $\sigma(\mathcal{F} \cup \Sigma^\circ)$. As usual, $F \Delta G$ denotes $(F \cap G^c) \cup (F^c \cap G)$. We have the following:

Lemma B.1 *$F \in \bar{\mathcal{F}}$ if and only if there is $G \in \mathcal{F}$ such that* $\Pr(F \Delta G) = 0$.

Proof. Define $\tilde{\mathcal{F}} \equiv \{F \in \Sigma : \exists G \in \mathcal{F} \text{ such that } \Pr(F \Delta G) = 0\}$. It is clear that $\tilde{\mathcal{F}} \supset \mathcal{F} \cup \Sigma^\circ$. The fact that $F^c \Delta G^c = F \Delta G$ implies that $\tilde{\mathcal{F}}$ is closed to complementation. If $F_1, ..., F_n, ... \in \tilde{\mathcal{F}}$, with corresponding $G_1, ..., G_n, ... \in \mathcal{F}$ such that $\Pr(F_n \Delta G_n) = 0$, let $F = \cup_{n \in \mathbb{N}} F_n$ and $G = \cup_{n \in \mathbb{N}} G_n$. Then, $F^c \cap G = (\cap_n F_n^c) \cap \cup_n G_n \subset \cup_{n \in \mathbb{N}} (F_n^c \Delta G_n)$ and similarly for $F \cap G^c$. Thus, $F \Delta G \subset \cup_{n \in \mathbb{N}} (F_n \Delta G_n)$. Therefore, countable additivity implies that $\tilde{\mathcal{F}}$ is closed to countable unions, which shows that $\tilde{\mathcal{F}}$ is a σ-field. Since it contains $\mathcal{F} \cup \Sigma^\circ$, $\tilde{\mathcal{F}} \supset \bar{\mathcal{F}}$. On the other hand, assume that $F \in \tilde{\mathcal{F}}$ and let $G \in \mathcal{F}$ be such that $F \Delta G \in \Sigma^\circ$. Then $G \cap F^c \in \Sigma^\circ \subset \bar{F}$ and $F \cup G = G \cup (F \Delta G) \in \bar{F}$. Therefore, $F = (F \cup G) \setminus (G \cap F^c) \in \bar{F}$, which shows that $\tilde{\mathcal{F}} \subset \bar{\mathcal{F}}$ and concludes the proof. ∎

Definition B.2 *We say that $\mathcal{F}$ and $\mathcal{G}$ are equivalent σ-fields if $\bar{\mathcal{F}} = \bar{\mathcal{G}}$.*

We will consider random variables $X^i : \Omega \to \mathbb{R}$, for $i = 1, ..., n$ and Z.[64] The vector $(X^1, ..., X^n)$ will be denoted by X. We will denote the σ-fields $\sigma(X^i)$, $\sigma(X_1, ..., X_n)$ and $\sigma(Z)$ by $\mathcal{X}^i$, $\mathcal{X}$ and $\mathcal{Z}$, respectively.

Definition B.3 *Given a σ-field $\mathcal{F} \subset \Sigma$, a regular conditional probability given $\mathcal{F}$ is a function $Q : \Omega \times \Sigma \to \mathbb{R}_+$ satifying:*

(a) $\omega \mapsto Q(\omega, A)$ is $\mathcal{F}$-measurable, for any $A \in \Sigma$;

(b) for every $B \in \mathcal{F}$ and $A \in \Sigma$,

$$\int_B Q(\omega, A)\, d\Pr(\omega) = \Pr(A \cap B);$$

(c) for each ω, $Q(\omega, \cdot)$ is a (countably additive) probability measure on (Ω, Σ);[65]

In this case, the (regular) conditional probability $Q(\omega, A)$ will be denoted by $\Pr(A|\mathcal{F})_\omega$. Following the usual practice, ω will be sometimes omitted in $\Pr(\cdot|\mathcal{F})_\omega$.

Although conditional probabilities always exist, sometimes there does not exist a *regular* conditional probability. See Billingsley (1995, Exercise 33.11, p. 443). However, regular conditional probabilities always exist if Ω is a Polish (complete, separable, metrizable) space and Σ is its Borel field, as we assume here. (See Billingsley (1995, Theorem 33.3, p. 439)). Thus, we will consider only regular conditional probabilities in what follows, and refer to them simply as conditional probabilities.

Definition B.4 *We say that the sub-σ-fields $\mathcal{F}^1, ..., \mathcal{F}^n$ are conditionally independent given $\mathcal{F} \subset \Sigma$ if for any $A^i \in \mathcal{F}^i \subset \Sigma$, for $i = 1, ..., n$, we have:*

$$\Pr(\cap_i A^i|\mathcal{F}) = \Pi_i \Pr(A^i|\mathcal{F}). \tag{4}$$

We denote this by $\perp\!\!\!\perp (\mathcal{F}^1, ..., \mathcal{F}^n)|\mathcal{F}$ or, if $n = 2$, by $\mathcal{F}^1 \perp\!\!\!\perp \mathcal{F}^2|\mathcal{F}$ and we say that $\mathcal{F}$ conditionally splits (or is a conditional splitter of) $(\mathcal{F}^1, ..., \mathcal{F}^n)$. We say that the random variables (r.v.) $X^1, ..., X^n$ are conditionally independent given a r.v. Z if $\perp\!\!\!\perp (\mathcal{X}^1, ..., \mathcal{X}^n)|\mathcal{Z}$ and we also denote this by $\perp\!\!\!\perp (X^1, ..., X^n)|Z$. In this case, we also say that Z conditionally splits (or is a conditional splitter of) $(X^1, ..., X^n)$.

The following non-trivial results will be needed:

Lemma B.5 (Mutual Conditional Independence) *The σ-fields $\mathcal{F}_t$, $t \in T$ are conditionally independent given $\mathcal{F}$ (denoted $\perp\!\!\!\perp \mathcal{F}_t|\mathcal{F}$) if and only if $\mathcal{F}_S \perp\!\!\!\perp \mathcal{F}_{T\setminus S}|\mathcal{F}$ for all sets $S \subset T$, where $\mathcal{F}_S \equiv \vee_{t\in S}\mathcal{F}_t$ is the σ-field generated by $\cup_{t\in S}\mathcal{F}_t$.*

[64] Our theory can be generalized for random variables taking values in more general spaces, but this seems sufficient for our purposes and avoid unnecessary complications.

[65] If only the first two conditions are satisfied, we have just a conditional probability (not regular).

Proof. It is sufficient to adapt the proof of Kallenberg (2002, Lemma 3.8 (ii), p. 51) to conditional independence. ∎

Lemma B.6 (Doob) *For any σ-fields $\mathcal{F}, \mathcal{G}$ and $\mathcal{H}$, $\mathcal{F} \perp\!\!\!\perp \mathcal{G}|\mathcal{H}$ if and only if*

$$\Pr[H|\mathcal{F}, \mathcal{G}] = \Pr[H|\mathcal{G}] \; a.s., \forall H \in \mathcal{H}.$$

Proof. See Kallenberg (2002, Proposition 6.6, p. 110). ∎

B.2 Proof of Proposition 5.4

To prove Proposition 5.4, we need the following lemma, which is in fact the first half of that proposition.

Lemma B.7 *Let $\mathcal{F}^1, ..., \mathcal{F}^n$ be the sub-σ-fields of Σ. There exist a σ-field $\mathcal{Z} \subset \Sigma$ such that $\mathcal{F}^1, ..., \mathcal{F}^n$ are conditionally independent given $\mathcal{Z}$.*

Proof. For each set $S \subset N \equiv \{1, ..., n\}$, let $\mathcal{Z}^S \equiv \vee_{i \in S} \mathcal{F}^i$, that is, $\mathcal{Z}^S$ denotes the σ-field generated by $\cup_{i \in S} \mathcal{F}^i$. For simplicity, we write $\mathcal{Z}$ Instead of $\mathcal{Z}^N$. By lemma B.5, it is sufficient to prove that $\mathcal{Z}^S$ and $\mathcal{Z}^{S^c}$ are conditionally independent given $\mathcal{Z}$ for all $S \subset N$. By lemma B.6, this follows by establishing $\Pr[H|\mathcal{Z}^S, \mathcal{Z}] = \Pr[H|\mathcal{Z}]$, for all $H \in \mathcal{Z}^{S^c}$. But since $\sigma(\mathcal{Z}^S, \mathcal{Z}) = \mathcal{Z}$, the conditional probability is the same in both sides of this equation. ∎

Observe that the existence of $\mathcal{Z}$ given in Lemma B.7 above is yet not sufficient for the last statement in Proposition 5.4 because it is not always true that given a σ-field $\mathcal{Z}$, there exists a r.v. Z such that $\mathcal{Z} = \sigma(Z)$. For this to hold, it is necessary that $\mathcal{Z}$ is countably generated. A σ-field $\mathcal{F}$ is countably generated if there is a countable class of sets $A_1, A_2, ...$ such that $\mathcal{F} = \sigma(A_1, A_2, ...)$. Indeed, we have the following:

Claim B.8 *There exists r.v. Z such that $\mathcal{F} = \sigma(Z)$ if and only if $\mathcal{F}$ is countably generated.*[66]

Proof. If $\mathcal{F} = \sigma(A_1, A_2, ...)$, define: $Z(\omega) = \sum_{k=1}^{\infty} 3^{-k} 1_{A_k}(\omega)$. It is easy to see that $\sigma(Z) = \sigma(A_1, A_2, ...)$. Conversely, assume that $\mathcal{F} = \sigma(Z)$ for some r.v. Z. Since Z is $\mathcal{F}$-measurable, for each $B \in \mathcal{B}$, $Z^{-1}(B) \in \mathcal{F}$. Take a countable class of sets $A_1, A_2, ...$ which generate $\mathcal{B}$ (for instance, the class of intervals $(a, b]$, for $a, b \in \mathbb{Q}$). Since $A_k \in \mathcal{B}$, $Z^{-1}(A_k) \in \mathcal{F}$, which implies $Z^{-1}(\sigma(A_1, A_2, ...)) = Z^{-1}(\mathcal{B}) \subset \mathcal{F}$. But since $\mathcal{F} = \sigma(Z)$ is the smallest σ-field with respect to which Z is measurable, $\mathcal{F} \subset \sigma(Z^{-1}(\mathcal{B})) = Z^{-1}(\mathcal{B})$, concluding the proof of the claim. ∎

Given a countably generated σ-field $\mathcal{F}$ and a sub-σ-field $\mathcal{G} \subset \mathcal{F}$, it is *not* necessarily true that $\mathcal{G}$ is countably generated. To see a counterexample, let $\mathcal{F}$ be the borelianos in $[0, 1]$ and $\mathcal{G}$ the class of all countable or cocountable subsets of $[0, 1]$ (a set is said to be cocountable if its complement is countable). It is not difficult to verify that $\mathcal{G}$ is a sub-σ-field of $\mathcal{F}$, which is

[66] See Billingsley (1995, Exercise 20.1, p. 270). We include the proof for completeness.

not countably generated. Therefore, the rest of the proof of Proposition 5.4 requires the use of equivalent equivalent fields (see definition B.2). From Lemma B.1, we know that two σ-fields $\mathcal{G}, \mathcal{G}' \subset \Sigma$ are equivalent if for every $B \in \mathcal{G}$, there is a $B' \in \mathcal{G}'$ such that $\Pr(B \Delta B') = 0$ and for every $B' \in \mathcal{G}'$, there is a $B \in \mathcal{G}$, such that $\Pr(B \Delta B') = 0$. Consider the following two facts:

Lemma B.9 *Every sub-σ-field $\mathcal{G}$ of Σ is equivalent to a countably generated sub-σ-field $\mathcal{G}'$.*

Proof. Given $E, F \in \Sigma$, let $d(E, F) \equiv \Pr(E \Delta F)$. This defines a pseudo-metric in Σ. Since Σ is the Borel σ-field in Ω, there is a countable set of sets $\{E_n\}$ such that $\Sigma = \sigma(\{E_n\})$. Since the ring generated by $\{E_n\}$ is also countable, then we may assume that $\{E_n\}$ is a ring. By Halmos (1974, Theorem 13.D), for every $m \in \mathbb{N}$ and set $A \in \Sigma$ (in particular, for every $A \in \mathcal{G}$), there is a integer n such that $\Pr(A \Delta E_n) < 1/m$. Let $B_{m,n}$ denote this set. Then, the collection of sets $\{B_{m,n}\} \subset \{E_n\}$ is countable and it is dense in $\mathcal{G}$. Therefore, $\mathcal{G}' \equiv \sigma(\{B_{m,n}\})$ is countably generated and is equivalent to $\mathcal{G}$. ∎

Lemma B.10 *If $\mathcal{F}$ and $\mathcal{G}$ are equivalent σ-fields then $\Pr(\cdot|\mathcal{F}) = \Pr(\cdot|\mathcal{G})$ (a.e.).*

Proof. It can be shown that a function $f : \Omega \to \mathbb{R}$ is $\mathcal{F}$-measurable if and only if there is a $\bar{f} : \Omega \to \mathbb{R}$ which is equal to f a.e. and is $\bar{\mathcal{F}}$-measurable. Thus, if $\omega \mapsto Q(\omega, A)$ is a probability kernel which is $\mathcal{F}$-measurable, it is $\bar{\mathcal{F}}$-measurable (up to a null set). Since the second condition in the definition of conditional expectation is also satisfied for $\bar{\mathcal{F}}$, then $\Pr(\cdot|\mathcal{F}) = \Pr(\cdot|\bar{\mathcal{F}})$. This implies the result. ∎

Conclusion of the proof of proposition 5.4: Lemmas B.7, B.9 and B.10 imply that there exist a random variable Z such that $\perp\!\!\!\perp (X^1, ..., X^n)|Z$, which concludes the proof. ∎

B.3 Theorem 5.2 for the finite support case

In this section, we will assume that each X^i have a finite support $S^i = \{x^i_1, ..., x^i_{k_i}\}$.[67] remain true We will assume the following trivial condition:[68]

Assumption B.11 *For all $x \in S^1 \times \cdots \times S^n$, there is $\omega \in \Omega$ such that $X(\omega) = x$.*

Lemma B.12 *Assume that $\mathcal{F}$ is a σ-field formed by a finite partition $\Pi = \{C_k\}_{k \in K}$ of Ω. For any $A \in \Sigma$,*

$$\Pr(A|\mathcal{F})_\omega = \sum_{k \in K} \Pr(A|C_k)_\omega 1_{C_k}(\omega). \qquad (a.e.)$$

[67] Most results and proofs are exactly the same for the case of countable support, although with some potential complication in the notation. Naturally, the algorithm described at the end of this subsection is restricted to the finite support case.

[68] It is easy to construct examples of spaces that do not satisfy this. For example, let $\Omega = \{a, b\}$, $n = 2$, $X^1(a) = X^2(b) = 0$ and $X^1(b) = X^2(a) = 1$. Then, $(0, 0) \in S^1 \times S^2$ but there is no $\omega \in \Omega$ such that $X(\omega) = (0, 0)$. However, in this case, we can easily add new zero-probability states to the space and change it to a space that satisfy it.

Proof. It is easy to see that the expression on the right above satisfies the two conditions for being a conditional probability. On the other hand, let $J \subset K$ be the set of indices j such that $\Pr(C_j) > 0$. Then $\Pr(\cup_{j \in J^c} C_j) = \sum_{j \in J^c} \Pr(C_j) = 0$ and the sets C_j for $j \notin J$ can be ignored. Since $\Pr(A|\mathcal{F})_\omega$ is $\mathcal{F}$-measurable, it must be constant in each C_j for $j \in J$ and, therefore, it must be $\Pr(A|\mathcal{F})_\omega = \Pr(A|C_k)$ for almost all $\omega \in C_k$. This concludes the proof. ∎

The following definition will be useful in the sequel.

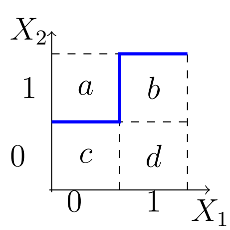
Definition B.13 *A set $C \in \Sigma$ is admissible if $\forall x = (x^1, ..., x^n) \in S^1 \times \cdots \times S^n$,*

$$\Pr(\{X^i = x^i\} \cap C) > 0, \forall i \Rightarrow \Pr(\{X = x\} \cap C) > 0.$$

We say that a partition Π is admissible if for all $C \in \Pi$, C is admissible.

To understand the concept of admissible set, consider the following example.

Example B.14 $\Omega = \{a, b, c, d\}$*;* $\Pi = \{\{a\}, \{b, c, d\}\}$ *and* $\Pr(\omega) > 0, \forall \omega \in \Omega$*. Let the values of $X(\omega)$ be given as in Figure 4. While the set $\{a\} \in \Pi$ is admissible, the set $C = \{b, c, d\}$ is not, because* $\Pr(X_1 = 0, X_2 = 1|C) = 0$ *while* $\Pr(X_1 = 0|C)\Pr(X_2 = 1|C) > 0$*. Note also that $\Pi' = \{\{a, b\}, \{c, d\}\}$ is an admissible partition.*



$\Pi = \{\{a, b\}, \{c, d\}\}$ is admissible, but
$\Pi = \{\{a\}, \{b, c, d\}\}$ is non-admissible.

Figure 4: Admissible and non-admissible partitions.

We have the following:

Proposition B.15 *$\bar{\mathcal{X}}^1, ..., \bar{\mathcal{X}}^n$ are conditionally independent given $\mathcal{F}$ if and only if Π is admissible and*

$$(\Pr(C))^{n-1} = \frac{\Pr(\{X^1 = x^1\} \cap C) \cdots \Pr(\{X^n = x^n\} \cap C)}{\Pr(\{X = x\} \cap C)}, \tag{5}$$

for all $C \in \Pi$ and $x = (x^1, ..., x^n) \in S^1 \times \cdots \times S^n$ s.t. $\Pr(\{X = x\} \cap C) > 0$.

Proof. *Necessity.* Observe that Lemma B.12 gives

$$\Pr(A|\Sigma)_\omega = \sum_{C \in \Pi} \Pr(A|C)_\omega 1_C(\omega) \qquad (a.e.),$$

which implies that the conditional independence must hold for all $C \in \Pi$ for which $\Pr(A \cap C) > 0$. If Π is not admissible, then there exists $x = (x^1, ..., x^n) \in S^1 \times \cdots \times S^n$ and $C \in \Pi$

such that $\Pr(\{X = x\} \cap C) = 0$ and $\Pr(\{X^i = x^i\} \cap C) > 0, \forall i = 1, ..., n$. But then $\Pr(C) > 0$ and $\Pr(\{X = x\}|C) = 0$ while $\prod_{i=1}^n \Pr(\{X^i = x^i\}|C) > 0$. Then $\mathcal{X}^1, ..., \mathcal{X}^n$ is not conditionally independent given Σ, a contradiction.

Since $\{x^i\} \in \mathcal{X}^i$, for $i = 1, ..., n$, necessity of the second condition comes directly from the conditional independence requirement:

$$\begin{aligned}\Pr(X^1 = x^1, ..., X^n = x^n|C) &= \Pr(\{X^1 = x^1\}|C) \cdots \Pr(\{X^n = x^n\}|C) \\ \iff \frac{\Pr(\{X = x\} \cap C)}{\Pr(C)} &= \frac{\Pr(\{X^1 = x^1\} \cap C) \cdots \Pr(\{X^n = x^n\} \cap C)}{(\Pr(C))^n}.\end{aligned}$$

Sufficiency. From Lemma B.12, it is sufficient to check the conditional independence condition for each $C \in \Pi$ satisfying $\Pr(C) > 0$. So, fix an element C of an admissible partition Π. For $i = 1, ..., n$, fix $A^i \in \mathcal{X}^i$ and let $A = A^1 \cap \cdots \cap A^n$. Let $\tilde{S}^i \subset S^i$ be the set of the values of X^i implied by A^i, that is, $\tilde{S}^i \equiv X^i(A^i)$. If for some i and $x^i \in \tilde{S}^i$, $\Pr(\{X^i = x^i\} \cap C) = 0$, we can redefine $\tilde{A}^i = A^i \setminus \{X^i = x^i\}$ and $\tilde{A} = A \setminus \{X^i = x^i\}$ so that $\Pr(\tilde{A}|C) = \Pr(A|C)$ and $\Pr(\tilde{A}^i|C) = \Pr(A^i|C)$. In other words, we can assume without loss of generality that $\Pr(\{X^i = x^i\} \cap C) > 0$ for all i and $x^i \in \tilde{S}^i$. Since Π is admissible, this implies that $\Pr(\{X = x\} \cap C) > 0$ for all $x \in \tilde{S} \equiv \tilde{S}^1 \times \cdots \times \tilde{S}^n$. It is clear that $X(A) \subset \tilde{S}$. Assume that there exists $x \in \tilde{S} \setminus X(A)$. By assumption B.11, there exists $\omega \in \Omega$ such that $X(\omega) = x = (x^1, ..., x^n)$. Without loss of generality, we can assume that $\omega \in A^i$ for all i.[69] Thus, $\omega \in A^1 \cap \cdots \cap A^n = A$. This implies that $x = X(\omega) \in X(A)$, which is a contradiction. This shows that $X(A) = \tilde{S}$, that is, $A = \cup_{x \in \tilde{S}} X^{-1}(x)$. It is also clear that $A^i = \cup_{x^i \in \tilde{S}^i} (X^i)^{-1}(x^i)$.

Since $\Pr(\{X = x\} \cap C) > 0$ for all $x \in \tilde{S}$, the assumption implies that $\Pr(X = x|C) = \prod_{i=1}^n \Pr(X^i = x^i|C)$ for all $x \in \tilde{S}$. Therefore,

$$\begin{aligned}\Pr(A|C) &= \sum_{x \in \tilde{S}} \Pr(X = x|C) = \sum_{x \in \tilde{S}} \prod_{i=1}^n \Pr(X^i = x^i|C) \\ &= \sum_{x^1 \in \tilde{S}^1} \cdots \sum_{x^n \in \tilde{S}^n} \prod_{i=1}^n \Pr(X^i = x^i|C) \\ &= \Pr(A^1|C) \cdots \Pr(A^n|C),\end{aligned}$$

as we wanted to show. ∎

The above result gives an algorithm to find (all) partitions Π that make the variables $X^1, ..., X^n$ conditionally independent.[70] The variable Z will be just the indication of the element of the partition Π that contains the true realization of types. The algorithm can be roughly described as follows:

[69] Since $X^i(\omega) = x^i \in \tilde{S}^i$, if $\Pr(\{\omega\}) > 0$, then $\omega \in A^i \in \mathcal{X}^i$; otherwise, we can put $\tilde{A}^i \equiv A^i \cup \{\omega\} \in \mathcal{X}^i$. Note that this change does not affect the previous assumption that $\Pr(\{X^i = x^i\} \cap C) > 0$.

[70] The algorithm will run in exponential time, but we conjecture that this cannot be significantly improved, unless P = NP.

Input: the finite probabilistic space $(\Omega, \Sigma, \Pr)$ and the partitions $\Pi^1, ..., \Pi^n$ generating $\mathcal{X}^1, ..., \mathcal{X}^n$.

1. Find the common knowledge partition: $\Pi^0 \equiv \wedge_{i=1}^n \Pi^i$, that is, the finest common coarsening partition.

2. Test whether Π^0 is admissible. If Π^0 is not admissible, find a coarser refinement of Π^0 that is admissible and call it Π^0.

3. Put $\Pi := \emptyset$. Let $\Pi^0 = \{C_1, ..., C_K\}$. For $k = 1, ..., K$, do:

   (a) Test whether C_k satisfies (5) for all $x = (x^1, ..., x^n) \in S^1 \times \cdots \times S^n$ such that $\Pr(\{X = x\} \cap C) > 0$. If it satisfies, do $\Pi := \Pi \cup \{C_k\}$. If C_k does not satisfy (5) for some x, do the following:

   (b) Calculate $P(\cdot) = \Pr(\cdot|C_k)$, that is the conditional probability given C_k;

   (c) Obtain $\Sigma_{C_k} = \Sigma \cap C_k$;

   (d) Derive and $\Pi^i_{C_k} \equiv \Pi^i \cap C_k$;

   (e) Call this program for the input (C_k, Σ_{C_k}, P) and partitions $\Pi^1_{C_k}, ..., \Pi^n_{C_k}$. The outcome will be a partition $\Pi_k = \{C_k^1, ..., C_k^m\}$ of C_k. Put $\Pi := \Pi \cup \{C_k^1, ..., C_k^m\}$.

Output: the partition Π.

This algorithm will produce all partitions that make $X^1, ..., X^n$ conditionally independent if step 2 considers all partitions that are coarser refinements of Π^0 and admissible. It will stop at some point because the fully informative partition makes $X^1, ..., X^n$ conditionally independent (see Proposition 5.4).

Gossner, Kalai, and Weber (2009) also have the existence of minimal conditional structure for finite or countable setting treated in this subsection.[71] It should be noted however that the important contribution of this subsection is not existence, since this is known at least since McKean (1963, p. 343, property e) (see also Mouchart and Rolin (1984, Theorem 4.3)). The main contribution of this subsection is the algorithm that it provides. The development of such an algorithm was an open question: see van Putten and van Schuppen (1985) and van Schuppen (1982). Gossner, Kalai, and Weber (2009)'s existence proof is based in the Zorn's Lemma, which is obviously non-constructive. They also provide a characterization of the case where the common knowledge partition makes the variables conditionally independent. However, their definition puts together cases 1 and 2 of of Theorem 5.2. We offer a different characterization of all cases below.

Lemma B.16 *Let S^i denote the finite or countable the support of the X^i. Consider the following two conditions: (i) the common knowledge partition is the (unique) outcome of the above algorithm; (ii) there exist $x^i \in \prod_{i=1}^n S^i$ such that* $\Pr(X = x) = 0$.[72]

[71] I am grateful to Olivier Gossner for pointing out this paper to me. This happened after I had obtained my results.

[72] This lemma is also true without the restriction to finite or countable support. In this case, condition (i) should be that the common knowledge partition makes the variables conditionally independent; and condition (ii) should be that there exist $B^i \subset S^i$ such that $\Pr(B^i) > 0, \forall i$, but $\Pr(B^1 \times \cdots \times B^n) = 0$.

If only condition (i) holds, but not (ii), we are in case 1 of Theorem 5.2. If both conditions hold, we are in case 2. If condition (i) does not hold, we are in case 3.

Proof. It is clear that if condition (i) does not hold, we are in case 3. Under condition (i), the common knowledge partition makes the types conditionally independent. Now, condition (ii) holds, that is, if there exists $x = (x^1, ..., x^n) \in \prod_{i=1}^n S^i$ such that $\Pr(X = x) = 0$, this means that there is i and j such that x^i and x^j cannot be in the same element of common knowledge partition. Therefore, this partition is not trivial, which means that the variables are not independent, that is, they are not in case 1. On the other hand, if condition (ii) does not hold, we have independence (case 1). ∎

This section establishes the existence of minimally informative conditional splitters in the finite support case, provides an algorithm to find all of them and establishes the classification of cases 1, 2 and 3 given in Theorem 5.2. The other statements of Theorem 5.2 for the case of finite support are included in the discussion for the general case, which is done in the next subsection.

B.4 Theorem 5.2 for the general case.

The minimally informative conditional splitters are defined as follows:

Definition B.17 (Minimally informative conditional splitters) *A variable Z that makes $X^1, ..., X^n$ conditional independent is minimally informative, denoted $\perp\!\!\!\perp (X^1, ..., X^n)|_{\min} Z$, if there is no σ-field $\mathcal{F}$ such that: (i)$\perp\!\!\!\perp (X^1, ..., X^n)|\mathcal{F}$, and (ii) $\mathcal{F}$ is equivalent to $\mathcal{Z}$.*

A variable Z that makes $X^1, ..., X^n$ conditional independent is the least informative if for every σ-field $\mathcal{F}$ such that: $\perp\!\!\!\perp (X^1, ..., X^n)|\mathcal{F}$, we have $\mathcal{F} \subset \bar{\mathcal{Z}}$.

Lemma B.16 gives the classification of cases in Theorem 5.2. The following lemma establishes statement 3(a) in Theorem 5.2:

Lemma B.18 *If $\perp\!\!\!\perp (\mathcal{X}^1, ..., \mathcal{X}^n)|\mathcal{F}$, the common knowledge σ-field $\mathcal{K}$ is included in $\mathcal{F}$.*

Proof. It is easy to see that the common knowledge σ-field is given by $\mathcal{K} = \cap_{i=1}^n \mathcal{X}^i$. Let $H \in \mathcal{K}$ and $\perp\!\!\!\perp (\mathcal{X}^1, ..., \mathcal{X}^n)|\mathcal{F}$. We want to prove that $H \in \mathcal{F}$. For each set $S \subset N \equiv \{1, ..., n\}$, let $\mathcal{X}^S \equiv \vee_{i \in S} \mathcal{X}^i$, that is, $\mathcal{X}^S$ denotes the σ-field generated by $\cup_{i \in S} \mathcal{X}^i$. By Lemmas B.5 and B.6 $\perp\!\!\!\perp (\mathcal{X}^1, ..., \mathcal{X}^n)|\mathcal{F}$ is equivalent to $P[H|\mathcal{F}] = P[H|\mathcal{X}^S, \mathcal{F}]$, for all $H \in \mathcal{X}^{S^c}$ and $S \subset \{1, ..., n\}$. Since $H \in \mathcal{K} = \cap_{i=1}^n \mathcal{X}^i \subset \mathcal{X}^{S^c}$, $P[H|\mathcal{F}] = P[H|\mathcal{X}^S, \mathcal{F}]$. Also, $\mathcal{K} \subset \sigma(\mathcal{X}^S, \mathcal{F})$ implies that H is $\sigma(\mathcal{X}^S, \mathcal{F})$-measurable and $P[H|\mathcal{X}^S, \mathcal{F}] = 1_H$. But $P[H|\mathcal{F}] = 1_H$ implies that $H \in \mathcal{F}$, as we wanted to show. ∎

For a proof of statement 3(b), see Mouchart and Rolin (1984, Theorem 4.3) (see also McKean (1963, p. 343, property e)). The algorithm mentioned in item 3(c) for the finite support case is presented in section B.3 above. The following example establishes statement 3(d), that is, it shows that it may not exist a least informative conditionally splitter.

Example B.19 *Let* $\Omega = \{a, b, c\}$ *and* $\Pr(\{\omega\}) = \frac{1}{3}$*, for all* $\omega \in \Omega$*. Consider the following* σ*-fields:*

$$\begin{aligned} \mathcal{F}_1 &= \sigma(\{a, b\}, \{c\}); \\ \mathcal{F}_2 &= \sigma(\{a\}, \{b, c\}). \end{aligned}$$

Note that $\mathcal{F}_1$ *and* $\mathcal{F}_2$ *are not independent because* $\Pr(\{b\}) \neq \Pr(\{a, b\}) \Pr(\{b, c\})$*. However, it is not difficult to verify that* $\mathcal{F}_1 \perp\!\!\!\perp \mathcal{F}_2 | \mathcal{F}_1$ *and* $\mathcal{F}_1 \perp\!\!\!\perp \mathcal{F}_2 | \mathcal{F}_2$ *and these two different* σ*-fields are minimally informative. Thus, it is not possible that there is a least informative conditional splitter.*

The proof of item 3(e) is given in the next subsection.

B.5 Proof of item 3(e) of Theorem 5.2

Although we will prove item 3(e) for the finite case, we will introduce definitions for the general case below, since these are more common.

Definition B.20 *A Markov transition from* $(S, \mathcal{S})$ *to* $(X, \mathcal{X})$ *is a Borel measurable function* $T : S \to \Delta(X)$*, where* $\Delta(X)$ *is the set of all measures in the measurable space* $(X, \mathcal{X})$*, and the topology in* $\Delta(X)$ *(for giving its Borel sets) is its weak* topology, i.e., the* $\sigma(\Delta(X), C_b)$*-topology.*

Definition B.21 *A Markov kernel from* $(S, \mathcal{S})$ *to* $(X, \mathcal{X})$ *is a function* $k : S \times \mathcal{X} \to [0, 1]$ *satisfying the following:*

1. *for each* $s \in S$*, the set function* $k(s, \cdot) : \mathcal{X} \to [0, 1]$ *is a probability measure.*
2. *For each* $A \in \mathcal{X}$*, the mapping* $k(\cdot, A) : S \to [0, 1]$ *is* $\mathcal{S}$*-measurable.*

The reader will note that a conditional probability as defined in definition B.3 is just a Markov kernel as defined by definition B.21. The following two results are informative. They are respectively theorems 19.12 and 19.13 of Aliprantis and Border (2006, p.630).

Lemma B.22 *Let* S *and* X *be separable metrizable spaces. Then for a mapping* $T : S \to \Delta(X)$ *the following statements are equivalent.*

1. T *is a Markov transition, that is,* T *is Borel measurable.*
2. *The function* $k : S \times \mathcal{B}_X \to [0, 1]$*, defined by* $k(s, A) = T_s(A)$*, is a Markov kernel.*

Lemma B.23 *Let* S *and* X *be separable metrizable spaces. Then for a mapping* $k : S \times \mathcal{B}_X \to [0, 1]$ *the following statements are equivalent.*

1. *The function* k *is a Markov kernel.*

2. *The function* $T : S \to \Delta(X)$*, defined by* $s \mapsto T_s(\cdot) = k(s, \cdot)$*, is a Markov transition.*

For now, let Y denote the set $\Delta(\Omega)$, and let $\mathcal{Y}$ denote the weak* topology mentioned in definition B.20. Let $M \subset Y$ denote the set of all independent (product) measures, that is,

$$M = \{\mu \in \Delta(\Omega, \Sigma) : \mu = \mu^1 \times \cdots \times \mu^n, \mu^i \in \Delta(\Omega, \mathcal{X}^i)\}$$

It is easily seen that M is $\mathcal{Y}$-measurable. The following result clarifies the relationship between Markov transitions and conditional probabilities.

Lemma B.24 *A Markov transition* $T : \Omega \to \Delta(\Omega) = Y$ *from* $(\Omega, \mathcal{F})$ *to* (Ω, Σ) *represents the conditional probability given* $\mathcal{F}$ *if and only if for every* $B \in \mathcal{F}$ *and* $A \in \Sigma$*,*

$$\int_B T_\omega(A)\, d\Pr(\omega) = \Pr(A \cap B)$$

where $T_\omega(A)$ *represents the probability of the set* $A \in \Sigma$ *under the measure* $T(\omega) \in \Delta(\Omega)$*.*

In this case, $\mathcal{F}$ *conditionally splits* $(X^1, ..., X^n)$ *if and only if* $T(\Omega) \subset M$*.*

Proof. The first part is immediate from definitions B.3, B.20 and B.21 and Lemmas B.22 , B.23. The second part comes from the definition of M above and definition B.4. ∎

Corollary B.25 *The set* C *of Markov transitions* $T : \Omega \to \Delta(\Omega)$ *that conditionally splits* $(X^1, ..., X^n)$ *is a closed convex subset of all Markov transitions.*[73]

Proof. This comes directly from the two characterizing conditions given in Lemma B.24 above. ∎

The following result establishes the fact 3(e) of Theorem 5.2 for Ω finite. It is useful to introduce the notation: $L^p(\mathcal{F}, X)$ for the space $L^p((\Omega, \mathcal{F}, \Pr), X)$, $1 \leqslant p \leqslant \infty$, where X is a Banach space (see Diestel and Uhl (1977)). If $X = \mathbb{R}$, then we will write just $L^p(\mathcal{F})$ instead of $L^p(\mathcal{F}, \mathbb{R})$. Also, in the references below, DU stands for Diestel and Uhl (1977), while DS abbreviates Dunford and Schwartz (1958).

Proposition B.26 *Let* T^0 *denote the Markov transition representing the conditional probability given the common knowledge* σ*-field* $\mathcal{K}$*. If* Ω *is finite, then there is a unique* $T \in C$ *which realizes the minimal distance from* T^0 *to* C*.*

Proof. Let n be the number of points in Ω. Then a Markov transition is a function $T : \Omega \to \Delta(\Omega) \subset \mathbb{R}^n$. Then, C can be seen as a subset of $L^2(\Sigma, \mathbb{R}^n)$.[74] Since C is closed and convex subset of the Hilbert space $L^2(\Sigma, \mathbb{R}^n)$, and $T^0 \notin C$ (because we are considering case 3 of Theorem 5.2), there exists a unique point $\bar{T} \in C$ that realizes $\inf_{T \in C} ||T - T^0||$, where $|| \cdot ||$ denotes the $L^2(\Sigma, \mathbb{R}^n)$-norm (see DS, IV.4.2, p. 248). ∎

[73] We will use this result for Ω finite, so that the topology of $Y = \Delta(\Omega)$ is not important.

[74] Since all $T \in C$ take value in $\Delta(\mathbb{R}^n)$, which is compact, then C is actually a subset of $L^\infty(\Sigma, \mathbb{R}^n) \subset L^2(\Sigma, \mathbb{R}^n)$.

Remark B.27 *Note that Proposition B.26 only holds because C is a subset of a Hilbert space and we used its norm. If we have just a Banach space instead of a Hilbert space, it is not always true that there is a unique point minimizing the distance from C to P^0. For example, consider the set $D = \{(x_1, x_2) \in [0,1]^2 : x_1 = x_2\}$ in $\mathbb{R}^2$ with the sum norm: $||(x_1, x_2)|| = |x_1| + |x_2|$. The distance of the point $x^0 = (1, 0)$ to D is $1 = ||(x_1, x_2) - (1, 0)|| = |x_1 - 1| + |x_2|$ for all $(x_1, x_2) \in D$.*[75] *However, every Hilbert space is reflexive (DS, IV.4.6, p. 250) but $L^2(\Sigma, X)$ is reflexive if and only if X is reflexive (see DU, Corollary IV.1.2, p. 100). Unfortunately, $\Delta(\Omega)$ is reflexive if and only if it is finite dimensional (for instance, if Ω is finite) (DS, IV.13.21, p. 341). These observations show that the ideas in Proposition B.26 do not extend directly to general Ω.*

Remark B.28 (Related works) *Although Proposition 5.4 is known by specialists, we were unable to find a good reference for it. The closest reference is Suppes and Zanotti (1981), who state the existence of a (fully informative) Z when t_i are binary variables. It is interesting to note that Holland and Rosenbaum (1986, p.1525) quote Suppes and Zanotti's theorem and say that their proof "is easily generalized to the discrete case" and that "any distribution on $\mathbb{R}^J$ may be approximated arbitrarily well by a discrete distribution of $\mathbb{R}^J$, and Theorem 1 [Suppes-Zanotti] applies to such discrete approximation." They do not state, discuss or give any references for the result established in Proposition 5.4. We also found no reference for Proposition 5.4 in Probability text-books. Mouchart and Rolin (1984) and van Putten and van Schuppen (1985) prove the first part of Proposition 5.4 (that is, Lemma B.7), for the case of $n = 2$. It is obvious that those authors could have easily stated and proved Lemma B.7, in which case we would just quote them. A similar comment is valid for the existence of random variables (the second part of Proposition 5.4), which does not follow immediately from Lemma B.7.*

The closest references for Theorem 5.2 that we were able to find (unfortunately after obtaining our results) was section 4 of Mouchart and Rolin (1984) (see also section 5 of van Putten and van Schuppen (1985) and Mouchart and Rolin (1985)). They discuss what they call "σ-algebraic realization problem", which is the problem of finding a "minimal conditional independence relation CI$_{min}$" and proved item 3(b) of Theorem 5.2. The fact that there is no least informative conditional splitter (item 3(d) of Theorem 5.2) was known by van Schuppen (1982). However, the classification given in Theorem 5.2 is ours. The algorithm provided in item 3(c) was regarded as an open problem *by van Putten and van Schuppen (1985), which the algorithm provided here solves. The contribution given in item 3(e) of Theorem 5.2 (definition and existence of least informative conditional splitter) is also completely new.*

[75] The norm of Hilbert space is given by an inner product. Naturally, this example fails with the euclidian norm, which comes from an inner product.

B.6 Proof of Theorem 5.7.

Let the distribution of X and Y be given by the table below, that is, $\Pr(X = 1, Y = 0) = b$.

	$Y = 0$	$Y = 1$
$X = 0$	a	b
$X = 1$	b	d

Our purpose is to find a binary variable Z, with joint distribution with X and Y described by the tables below,

$Z = 0$	$Y = 0$	$Y = 1$
$X = 0$	ua	vb
$X = 1$	vb	wd

$Z = 1$	$Y = 0$	$Y = 1$
$X = 0$	$(1-u)a$	$(1-v)b$
$X = 1$	$(1-v)b$	$(1-w)d$

such that the conditional independence conditions are satisfied:

$$\begin{cases} aduw = b^2v^2 \\ ad(1-u)(1-w) = b^2(1-v)^2 \\ u, v, w \in (0,1) \end{cases}$$

Let us define $r \equiv \frac{b^2}{ad}$. From the first equation above, we obtain: $w = r\frac{v^2}{u}$. Thus, $w < 1$ is equivalent to $rv^2 < u$. The second equation simplifies to $(1 - u - w + uw) = r(1 - 2v + v^2)$ and, using the first equation, to $1 - u - w = r(1 - 2v)$. Substituting $w = r\frac{v^2}{u}$ we obtain:

$$\begin{aligned} 1 - u - r\frac{v^2}{u} &= r(1-2v) \iff \\ & u^2 - u\left[1 - r(1-2v)\right] + rv^2 = 0. \end{aligned}$$

Observe that the case $r = 1$ corresponds to independence of X and Y, in which case $u = v = w$ can be anything. So, we assume $r \neq 1$ in what follows. The solution is:

$$u = \frac{[1 - r(1-2v)] \pm \sqrt{[1 - r(1-2v)]^2 - 4rv^2}}{2} \tag{6}$$

The conditions $u, v, w \in (0,1)$ will be satisfied if $u, v \in (0,1)$ and $rv^2 < u$. Let us define the polynomial $P(u) \equiv u^2 - u\left[1 - r(1-2v)\right] + rv^2$ and observe that:

$$\begin{aligned} P(0) &= 0^2 - 0[1 - r(1-y-z)] + rv^2 = rv^2 > 0; \\ P(rv^2) &= (rv^2)^2 - (rv^2)[1 - r(1-2v)] + rv^2 \\ &= rv^2(rv^2 - 1 + r - 2rv + 1) = r^2v^2(1-v)^2 > 0; \\ P(1) &= 1^2 - 1[1 - r(1-2v)] + rv^2 = r(1 - 2v + v^2) = r(1-v)^2 > 0. \end{aligned}$$

Since u must be a root of the polynomial $P(u)$, the conditions $u, v \in (0,1)$ and $rv^2 < u$ are equivalent to $v \in (0,1)$ and:

$$rv^2 < \frac{1-r(1-2v)}{2} < 1 \iff \frac{1-\sqrt{\frac{2-r}{r}}}{2} < v < \min\{\frac{1+\sqrt{\frac{2-r}{r}}}{2}, \frac{1+r}{2r}\};$$

$$[1-r(1-2v)]^2 \geqslant 4rv^2 \iff \begin{cases} v \geqslant \frac{r-1}{2(r-\sqrt{r})} & \text{if } r > 1 \\ v \leqslant \frac{1-r}{2(\sqrt{r}-r)} & \text{if } r < 1 \end{cases}$$

Note that the inequality $rv^2 < u \iff rv^2 < \frac{1-r(1-2v)}{2}$ cannot be satisfied for $r > 2$. For $r \in (1,2]$,

$$\begin{aligned} \frac{r-1}{2(r-\sqrt{r})} &> \frac{1+\sqrt{\frac{2-r}{r}}}{2} \\ \iff (\sqrt{r}-1)(\sqrt{r}+1) &> \frac{1}{\sqrt{r}}(\sqrt{r}+\sqrt{2-r})(r-\sqrt{r}) \\ \iff \sqrt{r}+1 &> \sqrt{r}+\sqrt{2-r}, \end{aligned}$$

which is obviously true for $r \in (1,2]$. Therefore, the conditions cannot be met if $r > 1$.[76] On the other hand, if $r < 1$, then $\sqrt{\frac{2-r}{r}} > 1$ and $\frac{1+r}{2r} > 1$, in which case the inequalities $\frac{1-\sqrt{\frac{2-r}{r}}}{2} < v < \min\{\frac{1+\sqrt{\frac{2-r}{r}}}{2}, \frac{1+r}{2r}\}$ are trivially satisfied for any $v \in (0,1)$. We simplify all conditions to: $v \in (0,1)$ and $v \leqslant \frac{1-r}{2(\sqrt{r}-r)}$, but this last condition is also trivially satisfied for any $v \in (0,1)$, since:

$$\begin{aligned} \frac{1-r}{2(\sqrt{r}-r)} &\geqslant 1 \\ \iff 1-r &\geqslant 2\sqrt{r}-2r \\ \iff (\sqrt{r}-1)^2 &\geqslant 0. \end{aligned}$$

Therefore, if $r < 1$, we can choose any $v \in (0,1)$, obtain u from (6) and put $w = r\frac{v^2}{u}$. This will give the decomposition that we wanted. Note that the condition $r < 1$ is equivalent to $ad > b^2$, which is just positive correlation.

References

AARONSON, J. (1997): *An Introduction to Infinite Ergodic Theory*. American Mathematical Society.

ALIPRANTIS, C., AND K. BORDER (2006): *Infinite dimensional analysis: a hitchhiker's guide*. Springer Verlag.

[76] This gives another proof for the statement in Example 3.4.

ALLAIS, M. (1953): "Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'Ecole Americaine," *Econometrica*, 21(4), 503–546.

ATHEY, S. (2001): "Single Crossing Properties and the Existence of Prue Strategy Equilibria in Games of Incomplete Information," *Econometrica*, 69(4), 861–99.

AUMANN, R. J. (1976): "Agreeing to Disagree," *The Annals of Statistics*, 4(6), 1236–1239.

BARLOW, R. E., AND F. PROSCHAN (1965): *Mathematical theory of reliability*. John Wiley & Sons.

BERGIN, J. (2001): "Common Knowledge with Monotone Statistics," *Econometrica*, 69(5), 1315–1332.

BILLINGSLEY, P. (1995): *Probability and measure*, Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, third edn., A Wiley-Interscience Publication.

DE CASTRO, L. (2007): "Affiliation, Equilibrium Existence and the Revenue Ranking of Auctions," Discussion paper, Working paper, Universidad Carlos III de Madrid.

——— (2008): "Grid Distribution to Study Single Obejct Auctions," Discussion paper, University of Illinois at Urbana-Champaign.

DE CASTRO, L. I., AND H. J. PAARSCH (2008): "Testing Affiliation in Private-Values Models of First-Price Auctions Using Grid Distributions," *University of Illinois at Urbana-Champaign*.

DIACONIS, P. (1977): "Finite forms of de Finetti's theorem on exchangeability," *Synthese*, 36(2), 271–281.

DIACONIS, P., AND D. FREEDMAN (1980): "Finite exchangeable sequences," *Ann. Probab.*, 8(4), 745–764.

DIESTEL, J., AND J. UHL (1977): *Vector Measures*. AMS (Mathematical Surveys no. 15), Providence, R.I.

DUNFORD, N., AND J. SCHWARTZ (1958): *Linear Operators, Part I*. Interscience, New York.

DUPUIS, P., AND R. ELLIS (1997): *A weak convergence approach to the theory of large deviations*. Wiley-Interscience.

ELLSBERG, D. (1961): "Risk, Ambiguity, and the Savage Axioms," *The Quarterly Journal of Economics*, 75(4), 643–669.

ESARY, J., F. PROSCHAN, AND D. WALKUP (1967): "Association of random variables, with applications," *The Annals of Mathematical Statistics*, 38, 1466–1474.

FRIEDMAN, M. (1953): "The methodology of positive economics," *The Philosophy of economics: an anthology*, 2, 180–213.

FUDENBERG, D. (2006): "Advancing Beyond Advances in Behavioral Economics," *Journal of Economic Literature*, 44(3), 694–711.

GOSSNER, O., E. KALAI, AND R. WEBER (2009): "Information Independence and Common Knowledge," *Econometrica*, 77(4), 1317–1328.

HALMOS, P. (1974): *Measure theory*. Springer-Verlag Inc., New York.

HARSANYI, J. C. (1967-8): "Games with incomplete information played by "Bayesian" players, Parts I, II, III," *Management Sci.*, 14, 159–182, 320–334, 486–502.

HEWITT, E., AND L. J. SAVAGE (1955): "Symmetric measures on Cartesian products," *Trans. Amer. Math. Soc.*, 80, 470–501.

HOLLAND, P., AND P. ROSENBAUM (1986): "Conditional association and unidimensionality in monotone latent variable models," *The Annals of Statistics*, 14(4), 1523–1543.

JACKSON, M. O., AND J. M. SWINKELS (2005): "Existence of Equilibrium in Single and Double Private Value Auctions," *Econometrica*, 73(1), 93–139.

JOGDEO, K. (1977): "Association and probability inequalities," *The Annals of Statistics*, 5, 495–504.

JUN, S. J., J. PINKSE, AND Y. WAN (2008): "A Consistent Nonparametric Test of Affiliation in Auction Models," Discussion paper, Penn State University.

KAGEL, J. H. (1995): "Auctions: A Survey of Experimental Research," *In: Kagel, J. H. and Roth, A., The Handbook of Experimental Economics, Princeton University Press,*, pp. 501–85.

KALLENBERG, O. (2002): *Foundations of Modern Probability*. Springer, 2nd edn.

KARLIN, S. (1968): *Total positivity*. Stanford University Press, Stanford, Calif.

KLEMPERER, P. (2003): "Alfred Marshall Lecture: Using and Abusing Economic Theory," *Journal of the European Economic Association*, 1(2-3), 272–300.

KLEMPERER, P. (2004): *Auctions: Theory and Practice*. Princeton University Press.

LAFFONT, J. (1997): "Game theory and empirical economics: The case of auction data," *European Economic Review*, 41(1), 1–35.

LEHMANN, E. (1966): "Some concepts of dependence," *Annals of Mathematical Statistics*, 37, 1137–1153.

——— (1986): *Testing statistical hypotheses*. John Wiley & Sons, 2nd edn.

LI, T., H. PAARSCH, AND T. HUBBARD (2007): "Semiparametric Estimation in Models of First-Price, Sealed-Bid Auctions with Affiliation," Discussion paper, working paper.

MASKIN, E., AND J. RILEY (1984): "Optimal Auctions with Risk Averse Buyers," *Econometrica*, 52(6), 1473–1518.

MCADAMS, D. (2003): "Characterizing Equilibria in Asymmetric First-Price Auctions," Working Paper, MIT.

MCAFEE, R., AND D. VINCENT (1992): "Updating the Reserve Price in Common Value Auctions," *American Economic Review, Papers and Proceedings*, 82(2), 512–518.

MCKEAN, H. P. (1963): "Brownian Motion with a Several-Dimensional Time," *Theory of Probability and its Applications*, 8(4), 335–354.

MCMILLAN, J. (1994): "Selling Spectrum Rights," *Journal of Economic Perspectives*, 8(3), 145–162.

MERTENS, J., AND S. ZAMIR (1985): "Formulation of Bayesian analysis for games with incomplete information," *International Journal of Game Theory*, 14(1), 1–29.

MILGROM, P. (1981a): "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics*, 12(2), 380–391.

——— (1981b): "Rational Expectations, Information Acquisition, and Competitive Bidding," *Econometrica*, 49(4), 921–43.

——— (1989): "Auctions and Bidding: A Primer," *Journal of Economic Perspectives*, 3(3), 3–22.

MILGROM, P., AND J. ROBERTS (1990): "Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities," *Econometrica*, 58, 1255–77.

MILGROM, P., AND C. SHANNON (1994): "Monotone Comparative Statics," *Econometrica*, 62(1), 157–80.

MILGROM, P., AND R. WEBER (1982a): "A Theory of Auctions and Competitive Bidding," *Econometrica*, 50(5), 1089–1122.

MILGROM, P., AND R. J. WEBER (1982b): "The value of information in a sealed-bid auction," *Journal of Mathematical Economics*, 10(1), 105 – 114.

MONTEIRO, P., AND H. MOREIRA (2006): "First-price auctions without affiliation," *Economics Letters*, 91(1), 1–7.

MOUCHART, M., AND J. ROLIN (1984): "A note on conditional independence with statistical applications," *Statistica (Bologna)*, 4, 557–584.

——— (1985): "Letter to the editor. Comment on "Invariance Properties of the Conditional Independence Relation," *Statistica (Bologna)*, 3, 427–430.

PERRY, M., AND P. RENY (1999): "On the Failure of the Linkage Principle in Multi-unit Auctions," *Econometrica*, 67(4), 895–900.

PERSICO, N. (2000): "Information Acquisition in Auctions," *Econometrica*, 68(1), 135–148.

RENY, P. J., AND S. ZAMIR (2004): "On the Existence of Pure Strategy Monotone Equilibria in Asymmetric First-Price Auctions," *Econometrica*, 74(4), 1105–25.

SOBEL, J. (2006): "Information Aggregation and Group Decisions," *Unpublished Paper*.

STIGLER, G. J. (1950): "The Development of Utility Theory. II," *The Journal of Political Economy*, 58(5), 373–396.

SUPPES, P., AND M. ZANOTTI (1981): "When are probabilistic explanations possible?," *Synthese*, 48(2), 191–199.

TONG, Y. (1980): "Probability inequalities in multivariate distributions.," *New York: Academic Press, 1980.*

TOPKIS, D. (1978): "Minimizing a submodular function on a lattice," *Operations Research*, 26, 305–321.

VAN PUTTEN, C., AND J. H. VAN SCHUPPEN (1985): "Invariance Properties of the Conditional Independence Relation," *The Annals of Probability*, 13(3), 934–945.

VAN SCHUPPEN, J. H. (1982): "The strong finite stochastic realization problem: Preliminary results," *Lecture Notes in Control and Information Science*, 44, 179–190.

VAN ZANDT, T., AND X. VIVES (2007): "Monotone equilibria in Bayesian games of strategic complementarities," *Journal of Economic Theory*, 134(1), 339 – 360.

WILSON, R. (1969): "Competitive bidding with disparate information," *Management Science*, 15, 446–448.

WILSON, R. (1977): "A bidding model of perfect competition," *The Review of Economic Studies*, 4, 511–518.

YANAGIMOTO, T. (1972): "Families of positively dependent random variables," *Annals of the Institute of Statistical Mathematics*, 24(1), 559–573.