Int. J. Game Theory (2006) 34: 91–104
DOI 10.1007/s00182-005-0004-3

ORIGINAL ARTICLE

M. Ali Khan · Kali P. Rath · Yeneng Sun

The Dvoretzky-Wald-Wolfowitz theorem and purification in atomless finite-action games

Published online: 16 March 2006

Abstract In 1951, Dvoretzky, Wald and Wolfowitz (henceforth DWW) showed that corresponding to any mixed strategy into a finite action space, there exists a pure-strategy with an identical integral with respect to a finite set of atomless measures. DWW used their theorem for purification: the elimination of randomness in statistical decision procedures and in zero-sum two-person games. In this short essay, we apply a consequence of their theorem to a finite-action setting of finite games with incomplete and private information, as well as to that of large games. In addition to simplified proofs and conceptual clarifications, the unification of results offered here re-emphasizes the close connection between statistical decision theory and the theory of games.

JEL Classification Numbers: C07 · DO5

A first draft of this paper was completed when Khan and Rath were visiting the *Institute for Mathematical Sciences* at the National University of Singapore in August 2003; they thank the Institute for supporting their visit. A preliminary version was presented at the *Midwest Economic Theory Conference* held at Indiana University, Bloomington in October 2003; the authors are grateful to Eric Balder, Robert Becker, William Thomson and Myrna Wooders for questions and helpful suggestions.

M. Ali Khan
Department of Economics, The Johns Hopkins University, Baltimore, MD 21218, USA

K. P. Rath
Department of Economics and Econometrics, University of Notre Dame, Notre Dame, IN 46556, USA
E-mail: kali.p.rath.1@nd.edu

Y. Sun
Department of Mathematics, National University of Singapore, 2 Science Drive 2, Singapore 117543, Singapore
Department of Economics, National University of Singapore, 1 Arts Link, Singapore 117570, Singapore

Keywords DWW theorem · Atomless · Mixed and pure strategies equilibria · Randomization · Purification

1 Introduction

Dvoretzky, Wald and Wolfowitz (henceforth DWW), in a theorem proved in Dvoretsky et al. (1951b, Theorem 4), and announced in Dvoretsky et al. (1951b, Theorem 1) and in Dvoretsky et al. (1951b, Theorem 2.1), showed that corresponding to any mixed strategy into a finite action space,[1] there exists a pure-strategy such that the two strategies have the same integral with respect to a finite set of atomless measures. They proved their theorem using the Lyapunov theorem for vector measures, and formulated, as a simple consequence of the theorem, a general purification principle that any mixed strategy can be purified to yield the same expected payoffs and distributions. The general principle is then applied to the purification of statistical decision procedures, and of mixed strategies in two-person zero-sum games with finite action sets.[2]

In this paper, we observe that the general purification principle of DWW is applicable to finite games with incomplete information in Radner and Rosenthal (1982) as well as to large non-anonymous games in Schmeidler (1973), both presented in a finite-action setting. The relevance of the DWW insight to the purification problem in finite games with incomplete and diffuse information was already suggested in Radner and Rosenthal (1982, Footnote 3) [3] and in Milgrom and Weber (1985, section 5)[4]. However, the purification result in Milgrom and Weber (1985)does not follow directly from the original result in Dvoretsky et al. (1951a) as claimed therein, but from a new corollary of the DWW Theorem formulated here. In this context, we also clarify different notions of purification implicit in Dvoretsky et al. (1951a), Milgrom and Weber (1985) and Radner and Rosenthal (1982), and prove some results for games with incomplete information, based on what we term *strong purification*. This concept is stronger than all of the purification concepts in the relevant literature. With the alternative mathematical framework in place, it is also straightforward to derive the symmetrization result on equilibria in large anonymous games, as in Khan and Sun (1991) and Mas-Colell (1984).We point out that the derivation of each of the purification results in this paper is not available in the literature, and their directness and simplicity may perhaps be surprising.

To summarize, this essay, in giving a central location to the DWW theorem, re-emphasizes the intimate connection between statistical decision theory and the

[1] We alert the reader to the fact that what we term a mixed strategy here is called a behavioral strategy in Radner and Rosenthal (1982) and Milgrom and Weber (1985), while a mixed strategy there carries a different meaning.

[2] See respectively Dvoretsky et al. (1950, Theorems 5 and 6), Dvoretsky et al. (1951a, Theorems 3.1 and 3.2, section 4, Theorems 5.1 and 5.2) on statistical decision procedures, and Dvoretsky et al. (1950, Theorems 2 and 3), Dvoretsky et al. (1951a, section 9) on two-person zero-sum games. In Dvoretsky et al. (1950, Theorem 4) and Dvoretsky et al. (1951a, section 8), they also consider approximate purification, an issue that is outside the scope of this paper.

[3] Radner and Rosenthal (1982, Footnote 3) noted that their method for the proof of their Theorem 1 is reminiscent of the DWW theorem without giving a proof based on it.

[4] However, the relevance of the DWW insight to the issue of purification in large games has apparently not been noticed in a large and growing literature; see the references in Khan and Sun (2002).

theory of games, an interdisciplinary thrust clearly evident in the classical papers, and one that may possibly suggest fruitful new questions for both subjects.[5]

2 The DWW theorem

Let $\mathbb{R}^n$ denote n-dimensional Euclidean space and $\mathbb{R}^n_+$ its non-negative orthant. For any two measurable spaces $(X, \mathcal{X})$ and $(Y, \mathcal{Y})$, Meas(X, Y) denotes the space of $(\mathcal{X}, \mathcal{Y})$-measurable functions. Where needed, $\mathbb{R}^n$ will be endowed with its Borel σ-algebra. We reproduce the DWW theorem for the reader's convenience.[6]

Theorem [DWW] *Let $(T, \mathcal{T})$ be a measurable space; μ_k, $k = 1, \ldots, m$, finite, atomless measures on $(T, \mathcal{T})$; and $f_i \in Meas(T, \mathbb{R}_+), i = 1, \ldots, n$, such that for all $t \in T$, $\sum_{i=1}^n f_i(t) = 1$. Then there exist measurable functions $f_i^* \in Meas(T, \{0, 1\})$, $i = 1, \ldots, n$, such that $\sum_{i=1}^n f_i^*(t) = 1$ for all $t \in T$ and*

$$\int_T f_i(t) d\mu_k(t) = \int_T f_i^*(t) d\mu_k(t) \quad \text{for all } k = 1, \ldots, m, \quad \text{and} \quad i = 1, \ldots, n.$$

For a finite set A, let $\mathcal{M}(A)$ be the set of probability measures on A, and since it can be embedded in a finite-dimensional Euclidean space, we shall assume it to be endowed with its Borel σ-algebra.

We shall now present a corollary of the DWW Theorem that is used to derive a variety of purification results for atomless games. It also has independent interest. Note that Assertions (1) and (2) in the corollary are stated separately in Dvoretsky et al. (1951a, section 3) as Theorems 3.1 and 3.2 for atomless distribution functions; we put them together for general atomless probability measures. Assertions (2) and (3) are quoted in Milgrom and Weber (1985, last paragraph on page 629) as a result in Dvoretsky et al. (1951a). However, Assertion (3) is not proved in Dvoretsky et al. (1951a) as is claimed in Milgrom and Weber (1985); we prove it here and put it together with Assertions (1) and (2). The proof of the corollary we present here adapts the proof in Dvoretsky et al. (1951a, section 3) along with the additional trick of applying the DWW theorem to a specially chosen function $v_{\ell+n+1}(a, t) = 1_{A \setminus F(t)}$ for proving Assertion (3).

Corollary 1 *Let $(T, \mathcal{T})$ be a measurable space; μ_k, $k = 1, \ldots, m$, atomless probability measures on $(T, \mathcal{T})$; A a finite set represented as $\{a_1, \ldots, a_n\}$; v_j, $j = 1, \ldots, \ell$, elements of $Meas(A \times T, \mathbb{R})$ that are integrable with respect to each of the measures μ_k; and $g \in Meas(T, \mathcal{M}(A))$. Let $g(t; S)$ represent the value of the probability measure $g(t)$ at $S \subseteq A$ and $g(t; da)$ the integration operator with respect to it. Then there exists $g^* \in Meas(T, A)$ such that for all $k = 1, \ldots, m$,*

1. *For all $j = 1, \ldots, \ell$, $\int_T \int_A v_j(a, t) g(t; da) d\mu_k(t) = \int_T v_j(g^*(t), t) d\mu_k(t)$;*
2. *For all $B \subseteq A$, $\int_T g(t; B) d\mu_k(t) = \mu_k g^{*-1}(B)$;*

[5] In this connection, we refrain from the consideration of sequential statistical procedures, and, as noted above, from issues relating to approximate purification, see Dvoretsky et al. (1950); Dvoretsky et al. (1951a) and Wald and Wolfowitz (1951).

[6] See Dvoretsky et al. (1951a, Theorem 2.1) and the proof of Theorem 4 in Dvoretsky et al. (1951b).

3. $g^*(t) \in \{a_i \in A : g(t; \{a_i\}) > 0\} \equiv$ *supp* $g(t)$ *for* μ_k*-almost all* $t \in T$.

Proof We shall supplement the ℓ given functions v_j by $n+1$ additional functions. Towards this end, let v_j be the indicator function $1_{\{a_{j-\ell}\}}$ on A for $j = \ell+1, \ldots, \ell+n$ and $v_{\ell+n+1}(a, t) = 1_{A \setminus F(t)}$, where $F(t) = \{a \in A : g(t; \{a\}) > 0\}$, $A \setminus F(t)$ denotes set-theoretic subtraction, and for a set B, 1_B is the indicator function of the set B. Certainly v_j is bounded and in Meas$(A \times T, \mathbb{R})$ for each $j = \ell+1, \ldots, \ell+n+1$. For notational convenience, let $\ell + n + 1 = q$.

We first establish that there exists $g^* \in$ Meas(T, A) such that

$$\int_T \int_A v_j(a, t) g(t; da) d\mu_k(t) = \int_T v_j(g^*(t), t) d\mu_k(t) \quad \text{for all } k = 1, \ldots, m \quad \text{and} \quad j = 1, \ldots, q. \tag{1}$$

Towards this end, define for any $W \in \mathcal{T}$, $\nu_{ijk}(W) = \int_W v_j(a_i, t) d\mu_k(t)$ for all $i = 1, \ldots, n$, $j = 1, \ldots, q$, and $k = 1, \ldots, m$. Certainly ν_{ijk} are atomless finite (signed) measures. Define for each $i = 1, \ldots, n$, $f_i(t) = g(t)(\{a_i\})$. Certainly $\sum_{i=1}^n f_i(t) = 1$. On applying the DWW Theorem to ν_{ijk} and to f_i, we are guaranteed the existence of functions $f_i^* \in$ Meas$(T, \{0, 1\})$, $i = 1, \ldots, n$ such that $\sum_{i=1}^n f_i^*(t) = 1$ for all $t \in T$, and

$$\int_T f_i(t) d\nu_{ijk}(t) = \int_T f_i^*(t) d\nu_{ijk}(t) \quad \text{for all } i = 1, \ldots, n; \quad j = 1, \ldots, q \quad \text{and} \quad k = 1, \ldots, m. \tag{2}$$

On substituting the values of ν_{ijk} in (2), we obtain

$$\int_T v_j(a_i, t) f_i(t) d\mu_k(t) = \int_T v_j(a_i, t) f_i^*(t) d\mu_k(t) \quad \text{for all } i = 1, \ldots, n; \ j = 1, \ldots, q \quad \text{and} \quad k = 1, \ldots, m,$$

which, in turn, implies that for any $j = 1, \ldots, q$, and $k = 1, \ldots, m$,

$$\int_T \sum_{i=1}^n v_j(a_i, t) f_i(t) d\mu_k(t) = \int_T \sum_{i=1}^n v_j(a_i, t) f_i^*(t) d\mu_k(t). \tag{3}$$

For each $i = 1, \ldots, n$, let $T_i = \{t \in T : f_i^*(t) = 1\}$ and $g^* : T \longrightarrow A$, $g^*(t) = a_i$ for all $t \in T_i$. It is clear from the properties of f_i^* that $\{T_i\}$ is a measurable decomposition of T, and that therefore g^* is well-defined. Then, equation (3) implies that for any $j = 1, \ldots, q$ and $k = 1, \ldots, m$,

$$\begin{aligned}\int_T \int_A v_j(a,t) g(t;da) d\mu_k(t) &= \int_T \sum_{i=1}^n v_j(a_i,t) g(t;\{a_i\})) d\mu_k(t) \\ &= \int_T \sum_{i=1}^n v_j(a_i,t) f_i(t) d\mu_k(t) \\ &= \int_T \sum_{i=1}^n v_j(a_i,t) f_i^*(t) d\mu_k(t) \\ &= \sum_{i=1}^n \int_{T_i} v_j(a_i,t) d\mu_k(t) \\ &= \sum_{i=1}^n \int_{T_i} v_j(g^*(t),t) d\mu_k(t) \\ &= \int_T v_j(g^*(t),t) d\mu_k(t),\end{aligned}$$

and the proof of the claim is complete.

Now, on choosing $j = 1, \ldots, \ell$, in Equation (1), we obtain the first assertion of the corollary.

Next, to show that the distribution induced by $g^*(t)$ is the same as that induced by g on A for any of the m measures, we apply equation (1) to the function $v_j,\ j = \ell + 1, \ldots, \ell + n$, to obtain that for any $k = 1, \ldots, m$,

$$\begin{aligned}\int_T g(t;\{a_i\}) d\mu_k(t) &= \int_T \int_A 1_{\{a_i\}}(a) g(t;da) d\mu_k(t) \\ &= \int_T 1_{\{a_i\}}(g^*(t)) d\mu_k(t) = \mu_k g^{*-1}(\{a_i\})\end{aligned}$$

holds for all $i = 1, \ldots, n$. This implies Assertion (2).

Finally, on choosing $j = \ell + n + 1$, in equation (1), we obtain for any $k = 1, \ldots, m$,

$$\int_T \int_A 1_{A \setminus F(t)}(a) g(t;da) d\mu_k(t) = \int_T 1_{A \setminus F(t)}(g^*(t)) d\mu_k(t).$$

Since the first element of the equality is zero by construction, so is the second element, and this implies that $1_{A \setminus F(t)}(g^*(t))$ equals zero for μ_k-almost all $t \in T$. Thus, for μ_k-almost all $t \in T,\ g^*(t) \in \text{supp } g(t)$. □

3 Finite games with incomplete information: conditional independence

In this section, we consider finite games with incomplete information as formulated by Milgrom and Weber (1985). A *game with incomplete information* Γ_{MW} consists of a finite set of ℓ players and an information space available to them. Each player i is endowed with a finite action set A_i (denote the product space $\Pi_{j=1}^{\ell} A_j$ by A), a measurable space $(T_i, \mathcal{T}_i)$ representing his possible types, and a payoff function $u_i : A \times T_0 \times T_i \longrightarrow \mathbb{R}$. Let the measurable space $(T_0, \mathcal{T}_0)$, $T_0 = \{t_{01}, \ldots, t_{0m}\}$, represent the space of common states that affect the payoffs of all the players. The product measurable space $(T, \mathcal{T}) \equiv (\Pi_{j=0}^{\ell} T_j, \Pi_{j=0}^{\ell} \mathcal{T}_j)$ equipped with a probability measure η constitutes the information space of the game. Assume that for any $a \in A$, $u_i(a, t_0, t_i)$ is integrable on $(T, \mathcal{T}, \eta)$.[7]

For each $t_{0k} \in T_0$, $k = 1, \ldots, m$, let $\eta(\cdot; t_{0k})$ denote the conditional probability[8] on the space $(\Pi_{j=1}^{\ell} T_j, \Pi_{j=1}^{\ell} \mathcal{T}_j)$. Following Milgrom and Weber (1985), we shall assume that for each $i = 1, \ldots, \ell$, the marginal $\eta_i(\cdot; t_{0k})$ of $\eta(\cdot; t_{0k})$ on the space $(T_i, \mathcal{T}_i)$ is atomless and that $\eta(\cdot; t_{0k}) = \Pi_{i=1}^{\ell} \eta_i(\cdot; t_{0k})$. The latter condition is simply a formalization of the intuitive statement that conditional on T_0, the players' types are independent. It is also abbreviated to as conditional independence of probability measures. We shall denote the measure $\eta_i(\cdot; t_{0k})$ by μ_{ik}. The marginal of η on $(T_0, \mathcal{T}_0)$ is denoted by η_0, and $\eta_0(t_{0k})$ by μ_{0k} for each $k = 1, \ldots, m$.

For any player i, a *mixed strategy* is an element of $\mathrm{Meas}(T_i, \mathcal{M}(A_i))$, and a *pure strategy* is an element of $\mathrm{Meas}(T_i, A_i)$. A pure strategy $f_i \in \mathrm{Meas}(T_i, A_i)$ can also be viewed as a mixed strategy g_i by taking $g_i(t_i)$ to be the Dirac measure $\delta_{f_i(t_i)}$ at $f_i(t_i)$ for each $t_i \in T_i$. A *mixed (pure) strategy profile* is a collection $g = \{g_i\}_{i=1}^{\ell}$ of mixed (pure) strategies that specify a mixed (pure) strategy for each player. For a player $i = 1, \ldots, \ell$, we shall use the following (conventional) notation: $A_{-i} = \Pi_{1 \le j \le \ell, j \ne i} A_j$, $T_{-i} = \Pi_{1 \le j \le \ell, j \ne i} T_j$, $a = (a_i, a_{-i})$ for $a \in A$, $(t_1, \ldots, t_\ell) = (t_i, t_{-i})$ for $(t_1, \ldots, t_\ell) \in \Pi_{j=1}^{\ell} \mathcal{T}_j$, and $g = (g_i, g_{-i})$ for a strategy profile g.[9]

Assume that the players play the mixed strategy profile $g = \{g_i\}_{i=1}^{\ell}$. Then, the resulting expected payoff for player i can be written as

$$U_i(g) = U_i(g_1, \ldots, g_\ell) = \int_T \int_A u_i(a, t_i, t_0) g_1(t_1; da_1) \cdots g_\ell(t_\ell; da_\ell) d\eta$$

$$= \sum_{k=1}^{m} \mu_{0k} \int_{T_i} \int_{A_i} v_{ik}^{g}(a_i, t_i) g_i(t_i; da_i) d\mu_{ik}(t_i), \quad (4)$$

where $v_{ik}^{g}(a_i, t_i)$ (which depends on the mixed strategy profile g) equals

$$\int_{t_{-i} \in T_{-i}} \int_{a_{-i} \in A_{-i}} u_i(a_i, a_{-i}, t_i, t_{0k}) \prod_{j \ne i} g_j(t_j; da_j) d \prod_{j \ne i} \mu_{jk}(t_j), \quad (5)$$

[7] A boundedness condition on the payoffs is assumed in Milgrom and Weber (1985, p. 623).

[8] It always exists since T_0 is finite.

[9] From now on, without any ambiguity, we shall abbreviate $\Pi_{1 \le j \le \ell, j \ne i}$ to $\Pi_{j \ne i}$.

and the second equality in equation (4) is obtained from the assumption of conditional independence. The fact that the functions $v^g_{ik}(a_i, t_i)$ are in $\text{Meas}(A_i \times T_i, \mathbb{R})$ is a consequence of Fubini's theorem.[10] For each $j = 1, \ldots, \ell$, denote the measure $\int_{T_j} g_j(t_j, \cdot)d\mu_{jk}(t_j)$ on A_j by $\gamma^{g_j}_{jk}$. Then we can rewrite equation (5) as

$$v^g_{ik}(a_i, t_i) = \int\limits_{a_{-i} \in A_{-i}} u_i(a_i, a_{-i}, t_i, t_{0k})d \prod_{j \neq i} \gamma^{g_j}_{jk}(a_{-i}). \tag{6}$$

Thus, the ith player's expected payoff depends on the actions of the other players only through the conditional distribution of their strategies induced on their action spaces.

In the following, we define different types of equivalent strategy profiles, and this furnishes different types of purification concepts.

Definition 1 *Let $g = \{g_i\}^\ell_{i=1}$ and $g^* = \{g^*_i\}^\ell_{i=1}$ be two mixed strategy profiles.*

1. *The strategy profile g is said to be* payoff equivalent *to g^* if $U_i(g) = U_i(g^*)$ for all players $i = 1, \ldots, \ell$.*
2. *The strategy profiles g and g^* are said to be* strongly payoff equivalent *if (i) they are payoff equivalent; (ii) for any player i and any given mixed strategy $g'_i \in Meas(T_i, \mathcal{M}(A_i))$, $U_i(g'_i, g_{-i}) = U_i(g'_i, g^*_{-i})$.*
3. *For a player i, the strategy g_i is said to be* distribution equivalent *to the strategy g^*_i if they have the same conditional distribution on the action space A_i in the sense that for all $k = 1, \ldots, m$, $\int_{T_i} g_i(t_i; \cdot)d\mu_{ik}(t_i) = \int_{T_i} g^*_i(t_i; \cdot)d\mu_{ik}(t_i)$. The strategy profile g is said to be* distribution equivalent *to the strategy profile g^* if g_i is distribution equivalent to g^*_i for all players $i = 1, \ldots, \ell$.*
4. *For a player i, the pure strategy g^*_i is said to be* strongly distribution equivalent *to the strategy g_i if g^*_i is distribution equivalent to g_i, and for each $k = 1, \ldots, m$, $g^*_i(t_i) \in supp\ g_i(t_i)$ for μ_{ik}-almost all $t_i \in T_i$. The pure strategy profile g^* is said to be* strongly distribution equivalent *to the strategy profile g if g^*_i is strongly distribution equivalent to g_i for all players $i = 1, \ldots, \ell$.*
5. *A pure strategy profile g^* is said to be a* strong purification *of the strategy profile g if g^* is both strongly payoff equivalent and strongly distribution equivalent to g.*

Item (2)(ii) above says that the expected payoff of player i from an arbitrary mixed strategy is the same irrespective of whether his opponents play g_{-i} or g^*_{-i}. It is thus clear that if two strategy profiles are strongly payoff equivalent and one is an equilibrium of the game Γ_{MW}, then the other is also an equilibrium.

The following two examples establish, in the simple context of a single player games, that payoff equivalence and distribution equivalence are independent concepts. The first example shows that strong distribution equivalence does not imply payoff equivalence.

Example 1 Let $(T, \mathcal{T}, \lambda)$ be the unit Lebesgue interval. Let the action set $A = \{0, 1\}$. Let ν be the uniform distribution on A, which is to say $\nu(\{0\}) = 1/2 =$

[10] For the details of this theorem, see, for example, Ash (1972).

$\nu(\{1\})$, and let $g \in \text{Meas}(T, \mathcal{M}(A))$ such that $g(t) = \nu$ for all $t \in T$. Let $g^* : T \longrightarrow A$ be such that

$$g^*(t) = \begin{cases} 1 & \text{for } 0 \leq t < \frac{1}{2}, \\ 0 & \text{for } \frac{1}{2} \leq t \leq 1. \end{cases}$$

It is clear that $g^*(t) \in \text{supp}\, g(t)$ for all $t \in T$, and the induced distribution λg^{*-1} on A is identical to ν. Hence, g^* is strongly distribution equivalent to the pure strategy g.

Next, let $u : A \times T \longrightarrow \mathbb{R}$ be such that for all $t \in T$,

$$u(a,t) = \begin{cases} 0 & \text{for } a = 1, \\ g^*(t) & \text{for } a = 0. \end{cases}$$

We can now compute the two expressions:

$$\begin{aligned} U(g) &= \int_T \int_A u(a,t) g(t; da) d\lambda(t) = \frac{1}{2} \int_T [u(0,t) + u(1,t)] d\lambda(t) \\ &= \frac{1}{2} \int_T g^*(t) d\lambda(t) = \frac{1}{4}, \\ U(g^*) &= \int_T \int_A u(a,t) \delta_{g^*(t)}(da) d\lambda(t) = \int_T u(g^*(t), t) d\lambda(t) \\ &= \int_0^{1/2} u(1,t) d\lambda(t) + \int_{1/2}^1 u(0,t) d\lambda(t) = 0. \end{aligned}$$

Thus, g^* is not payoff equivalent (thus not strongly payoff equivalent) to g.

The second example shows that (strong) payoff equivalence does not imply distribution equivalence; and furthermore, (strong) payoff equivalence and distribution equivalence do not imply strong distribution equivalence.

Example 2 Let $(T, \mathcal{T}, \lambda)$ be the unit Lebesgue interval. Let the action set $A = \{0, 1\}$. Let $u : A \times T \longrightarrow \mathbb{R}$ be the constant payoff function with value 1, and ν the uniform distribution on A ($\nu(\{0\}) = 1/2 = \nu(\{1\})$), and let $g \in \text{Meas}(T, \mathcal{M}(A))$ be such that

$$g(t) = \begin{cases} \nu & \text{for } 0 \leq t < \frac{1}{2}, \\ \delta_{\{1\}} & \text{for } \frac{1}{2} \leq t \leq 1. \end{cases}$$

Define $g', g^* : T \to A$ such that $g'(t) = 0$ for all $t \in T$, and

$$g^*(t) = \begin{cases} 1 & \text{for } 0 \leq t < \frac{3}{4}, \\ 0 & \text{for } \frac{3}{4} \leq t \leq 1. \end{cases}$$

Then, g' is payoff equivalent to g but not distribution equivalent to g; g^* is both payoff and distribution equivalent to g but not strongly distribution equivalent to g.

Various notions of purification have been used in the literature. Both payoff equivalence and distribution equivalence are used for the purification of statistical decision procedures in Dvoretsky et al. (1951a, p. 2).[11] Payoff equivalence is used for purification of mixed strategies in two-person zero-sum games in Dvoretsky et al. (1950, Theorems 2 and 3, p. 257) and Dvoretsky et al. (1951a, p. 20), and of mixed strategy equilibria in finite games with incomplete information in Radner and Rosenthal (1982, p. 403).

Milgrom and Weber (1985) defined a notion of purification in section 5 that is weaker than the strong distribution equivalence in Definition 1; they also claimed in the last paragraph on page 629 and in the proof of their Theorem 4 that DWW proved in Dvoretsky et al. (1951a) that any (possibly non-equilibrium) mixed strategy profile has a strongly distribution equivalent pure strategy profile. We can make two observations in this connection. The first point is that this latter result is not contained in Dvoretsky et al. (1951a) as claimed in Milgrom and Weber (1985) (though it follows from our Corollary 1 here).[12] The second point concerns payoff equivalence. Since expected payoffs provide a starting point for discussing players' behavior in a game with incomplete information, a reasonable purification concept for a general mixed strategy profile ought at least to lead to equivalent expected payoffs; however, Example 1 shows that this (minimal) requirement does not follow from strong distribution equivalence.

The following theorem provides a result in terms of strong purification.

Theorem 1 *In the game Γ_{MW}, every mixed strategy profile has a strong purification.*

Proof Let $g = (g_1, \ldots, g_\ell)$ be a mixed strategy profile for the game Γ_{MW}. Fix any player $i = 1, \ldots, \ell$. For each $k = 1, \ldots, m$, compute μ_{ik} and v^g_{ik}. Now apply Corollary 1 to the collection

$$\{(T_i, \mathcal{T}_i), \{\mu_{ik}\}_{k=1}^m, A_i, \{v^g_{ik}\}_{k=1}^m, g_i\}$$

to obtain a pure strategy $g^*_i \in \text{Meas}(T_i, A_i)$ such that for all $k = 1, \ldots, m$,

(i) $\int_{T_i} \int_{A_i} v^g_{ik}(a_i, t_i) g_i(t_i; da_i) d\mu_{ik}(t_i) = \int_{T_i} v^g_{ik}(g^*_i(t_i), t_i) d\mu_{ik}(t_i)$;

(ii) for all $B \subseteq A_i$, $\int_{T_i} g_i(t_i; B) d\mu_{ik}(t_i) = \mu_{ik} g^{*-1}_i(B) = \gamma^{g_i}_{ik}$;

(iii) $g^*_i(t_i) \in \text{supp}\, g_i(t_i)$ for μ_{ik}-almost all $t_i \in T_i$.

Let $g^* = (g^*_1, \ldots, g^*_\ell)$. Then (ii) and (iii) above imply that the pure strategy profile g^* is strongly distribution equivalent to the strategy profile g. All that remains to be shown is the strong payoff equivalence of g^* and g.

Towards this end, consider any mixed strategy $g'_i \in \text{Meas}(T_i, \mathcal{M}(A_i))$. Denote (g'_i, g_{-i}) by g' and (g'_i, g^*_{-i}) by g'^*. By equations (4) and (6), the expected payoffs of player i with g' and g'^* are respectively given by

$$U_i(g') = \sum_{k=1}^m \mu_{0k} \int_{T_i} \int_{A_i} v^{g'}_{ik}(a_i, t_i) g'_i(t_i; da_i) d\mu_{ik}(t_i), \tag{7}$$

[11] The notions of payoff equivalence and distribution equivalence here are called respectively *equivalence* and *strong equivalence* in the bottom of page 2 in Dvoretsky et al. (1951a).

[12] It is claimed in Dvoretsky et al. (1951a, p. 6) that equations (4.5) to (4.8) can be obtained by applying equations (3.3) and (3.5) to the loss functions $W(F_i, d_{m_1 \cdots m_\ell}, x)$. While (4.5), (4.6) and (4.8) can be obtained in this way, (4.7) cannot. The authors thank Zhixiang Zhang for this observation.

$$U_i(g'^*) = \sum_{k=1}^{m} \mu_{0k} \int_{T_i} \int_{A_i} v_{ik}^{g'^*}(a_i, t_i) g_i'(t_i; da_i) d\mu_{ik}(t_i), \tag{8}$$

where $v_{ik}^{g'^*}(a_i, t_i) = \int_{a_{-i} \in A_{-i}} u_i(a_i, a_{-i}, t_i, t_{0k}) d\Pi_{j \neq i} \gamma_{jk}^{g_j^*}(a_{-i})$. By (ii) above and the fact that $g'_{-i} = g_{-i}$ and $g'^*_{-i} = g^*_{-i}$, we have $\gamma_{jk}^{g'^*j} = \gamma_{jk}^{g_j^*} = \gamma_{jk}^{g_j} = \gamma_{jk}^{g'_j}$ for all $j = 1, \ldots, \ell$ with $j \neq i$ and $k = 1, \ldots, m$. Hence $v_{ik}^{g'^*}(\cdot, \cdot) = v_{ik}^{g'}(\cdot, \cdot)$ for all $k = 1, \ldots, m$. By equations (7) and (8), we have $U_i(g') = U_i(g'^*)$, and also

$$U_i(g^*) = \sum_{k=1}^{m} \mu_{0k} \int_{T_i} v_{ik}^{g}(g_i^*(t_i), t_i) d\mu_{ik}(t_i). \tag{9}$$

Hence, $U_i(g^*) = U_i(g)$ by equation (4) and (i) above. Therefore, g_i^* is strongly payoff equivalent to g_i. Since player i was chosen arbitrarily, the proof is finished. □

4 Finite games with private information: mutual independence

Next, we turn to finite games with private information as formulated by Radner and Rosenthal (1982). We shall reformulate it as a special case of the game Γ_{MW} considered in the last section. This allows a synthetic treatment of finite games with private information that is independent or conditionally independent. A *game with private information* Γ_{RR} consists of a finite set of ℓ players, each of whom is endowed with a finite action set A_i (as above, A denotes the product space $\Pi_{j=1}^{\ell} A_j$), an information space constituted by a pair of ℓ measurable spaces $(T_i, \mathcal{T}_i)$ and $(S_i, \mathcal{S}_i)$ together with a probability measure μ on the product measurable space $(S, \mathcal{S}) \equiv (\Pi_{j=1}^{\ell}(T_j \times S_j), \Pi_{j=1}^{\ell}(\mathcal{T}_j \otimes \mathcal{S}_j)$, and finally, a payoff function $u_i : A \times S_i \longrightarrow \mathbb{R}$.

For any point $s = (t_1, s_1, \ldots, t_\ell, s_\ell) \in S$, and for any $i = 1, \ldots, \ell$, let (ζ_i, σ_i) be the coordinate projections, which is to say that $\zeta_i(s) = t_i$ and $\sigma_i(s) = s_i$. We shall assume that for every player i, the distribution $\eta_i = \mu\zeta_i^{-1}$ of ζ_i is atomless, and that the random variables $\{\zeta_j : j \neq i\}$ together with the random variable $\xi_i \equiv (\zeta_i, \sigma_i)$ form a mutually independent set of random variables. Assume that for any $a \in A$, $u_i(a, \sigma_i(\cdot))$ is integrable on $(S, \mathcal{S}, \mu)$.

Mixed (pure) strategies as well as mixed (pure) strategy profiles for the game Γ_{RR} can be defined as in section 3. For a given mixed strategy profile $g = \{g_i\}_{i=1}^{\ell}$, the resulting expected payoff for player i is given by

$$\begin{aligned} U_i(g) &= U_i(g_1, \ldots, g_\ell) \\ &= \int_S \int_{A_\ell} \cdots \int_{A_1} u_i(a, \sigma_i(s)) g_1(\zeta_1(s); da_1) \cdots g_\ell(\zeta_\ell(s); da_\ell) d\mu(s). \end{aligned} \tag{10}$$

Denote the measure $\int_S g_j(\zeta_j(s), \cdot) d\mu(s) = \int_{T_i} g_j(t_j, \cdot) d\eta_j(t_j)$ by $\gamma_j^{g_j}$. The assumption of independence implies that

$$U_i(g) = \int_{A_{-i}} \int_S \int_{A_i} u_i(a, \sigma_i(s)) g_i(\zeta_i(s); da_i) d\mu(s) d \prod_{j \neq i} \gamma_j^{g_j}(a_{-i}). \tag{11}$$

The definition of equivalent strategy profiles in the game Γ_{MW} in Definition 1 is still valid here for the game Γ_{RR}.

Next, choose $v_i \in \text{Meas}(A \times T_i, \mathbb{R})$ such that $v_i(a, \zeta_i(\cdot))$ is the conditional expectation $E(u_i(a, \sigma_i(\cdot))|\zeta_i)$.[13] Then, the property of conditional expectation implies that for any $a_{-i} \in A_{-i}$,

$$\int_S \int_{A_i} u_i(a, \sigma_i(s)) g_i(\zeta_i(s); da_i) d\mu(s) = \int_S \int_{A_i} v_i(a, \zeta_i(s)) g_i(\zeta_i(s); da_i) d\mu(s)$$
$$= \int_{T_i} \int_{A_i} v_i(a, t_i) g_i(t_i; da_i) d\eta_i(t_i). \quad (12)$$

By equations (11) and (12), we obtain that

$$U_i(g) = \int_{T_i} \int_A v_i(a, t_i) g_i(t_i; da_i) d\eta_i(t_i) d \prod_{j \neq i} \gamma_j^{g_j}(a_{-i})$$
$$= \int_{\Pi_{j=1}^{\ell} T_j} \int_A v_i(a, t_i) g_1(t_1; da_1) \cdots g_\ell(t_\ell; da_\ell) d \prod_{j=1}^{\ell} \eta_j. \quad (13)$$

We shall now define a special case of the game Γ_{MW} in section 3 by taking T_0 to be a singleton $\{t_{01}\}$. The payoff function for player i is v_i instead of the u_i there in the definition of the game Γ_{MW} in section 3. The measure η on $(T, \mathcal{T}) \equiv (\Pi_{j=0}^{\ell} T_j, \Pi_{j=0}^{\ell} \mathcal{T}_j)$ is the product measure $\delta_{t_{01}} \times \Pi_{j=1}^{\ell} \eta_j$. The remaining elements of the model are the same as those in section 3.

For a mixed strategy profile $g = \{g_i\}_{i=1}^{\ell}$, the expected payoff for player i is

$$V_i(g) = \int_T \int_A v_i(a, t_i) g_1(t_1; da_1) \cdots g_\ell(t_\ell; da_\ell) d\eta, \quad (14)$$

which equals $U_i(g)$ by equation (13). Thus, we obtain a special case of the game Γ_{MW} that has the same expected payoffs as the game Γ_{RR}. The following theorem follows obviously from Theorem 1.

Theorem 2 *In the game Γ_{RR}, every mixed strategy profile has a strong purification.*

In comparison with Radner and Rosenthal (1982, Theorem 1), the above theorem provides a strong purification result not only for an equilibrium but also for an arbitrary mixed strategy profile. On the other hand, a consequence of the DWW theorem that is available in Dvoretsky et al. (1951a) [i.e., Parts (1) and (2) of Corollary 1] already implies that any mixed strategy equilibrium is strongly payoff equivalent to a pure strategy equilibrium; one can simply ignore the argument used to establish (iii) in the first paragraph of the proof of Theorem 1.

[13] It is easy to see such a function v_i exists since A is finite.

5 Large non-anonymous games

Next, we turn to games with many players as in Schmeidler (1973). Let $(T, \mathcal{T}, \lambda)$ be an atomless probability space formalizing the space of players and $A = \{a_1, \ldots, a_n\}$ be the space of actions. A payoff function u is a continuous function on $A \times \mathcal{M}(A)$. Let $\mathcal{U}$ be the space of payoff functions endowed with its Borel σ-algebra generated by the sup-norm topology. A *large non-anonymous game* $\mathcal{G}$ is an element of $\text{Meas}(T, \mathcal{U})$. We shall also denote $\mathcal{G}(t)$ by u_t, and since one can always rescale the payoffs, we assume without any loss of generality that there is $M > 0$ such that for all $t \in T$, $\|u_t\| \leq M$.

A *mixed strategy profile* g (respectively a *pure strategy profile* g^*) is an element of $\text{Meas}(T, \mathcal{M}(A))$ $(\text{Meas}(T, A))$. Let the distribution induced on A by the mixed strategy profile g be denoted by γ^g, where for all $B \subseteq A$, $\gamma^g(B) = \int_T g(t; B)d\lambda(t)$. For any mixed strategy profile g, the expected payoff of player t is given by $\sum_{i=1}^n u_t(a_i, \gamma^g)g(t; \{a_i\}) = \int_A u_t(a, \gamma^g)g(t; da)$. The average payoff $U(g)$ is $\int_T \int_A u_t(a, \gamma^g)g(t; da)d\lambda$.

An *equilibrium in mixed strategies* g is an element of $\text{Meas}(T, \mathcal{M}(A))$ such that for λ-almost all $t \in T$, player t maximizes her expected payoff. Thus, for λ-almost all $t \in T$,

$$\sum_{i=1}^n u_t(a_i, \gamma^g)g(t; a_i) \geq \sum_{i=1}^n u_t(a_i, \gamma^g)p_i \quad \text{for all } p \in \left\{ p \in \mathbb{R}_+^n : \sum_{i=1}^n p_i = 1 \right\},$$

which implies that if $g(t; \{a_i\}) > 0$, then $u_t(a_i, \gamma^g) \geq u_t(a, \gamma^g)$ for all $a \in A$; this means that $u_t(a_i, \gamma^g)$ and $\int_A u_t(a, \gamma^g)g(t; da)$ take the maximum value of $u_t(\cdot, \gamma^g)$ on A.

An *equilibrium in pure strategies* is simply a pure strategy profile g^* such that for λ-almost all $t \in T$, player t maximizes her payoff $u_t(\cdot, \gamma^{g^*})$. It is clear that for any pure strategy profile g^*, γ^{g^*} is simply λg^{*-1}. An equilibrium in pure strategies g^* is a *purification* of an equilibrium g in mixed strategies, if $\gamma^g = \gamma^{g^*}$. The proof of the following theorem only uses a consequence of the DWW theorem that is already available in Dvoretsky et al. (1951a), which is to say, Parts (1) and (2) of Corollary 1 above.

Theorem 3 *Any mixed strategy equilibrium g for the game $\mathcal{G}$ has a purification.*

Proof We also write $u_t(a, \gamma^g)$ as $u(a, t, \gamma^g)$. Apply parts (1) and (2) of Corollary 1 to the collection $\{(T, \mathcal{T}), \lambda, A, u(\cdot, \cdot, \gamma^g), g\}$ to obtain a pure strategy profile $g^* \in \text{Meas}(T, A)$ such that $\lambda g^{*-1} = \int_T g(t; \cdot)d\lambda = \gamma^g$ and

$$\int_T \int_A u_t(a, \gamma^g)g(t; da)d\lambda = \int_T u_t(g^*(t), \gamma^g)d\lambda. \tag{15}$$

For λ-almost all $t \in T$, since $\int_A u_t(a, \gamma^g)g(t; da)$ takes the maximum value of $u_t(\cdot, \gamma^g)$ on A, we have $\int_A u_t(a, \gamma^g)g(t; da) \geq u_t(g^*(t), \gamma^g)$. equation (15) implies that $\int_A u_t(a, \gamma^g)g(t; da) = u_t(g^*(t), \gamma^g)$ for λ-almost all $t \in T$. Hence, for λ-almost all $t \in T$, $u_t(g^*(t), \gamma^g) = u_t(g^*(t), \lambda g^{*-1})$ takes the maximum value of $u_t(\cdot, \gamma^g)$ on A, which means that g^* is an equilibrium in pure strategies and hence also a purification of g. □

6 Large anonymous games

Finally, we turn to anonymous games considered in Mas-Colell (1984) and Khan and Sun (1991). A large anonymous game is a probability measure μ in $\mathcal{M}(\mathcal{U})$, where $\mathcal{U}$ is the space of payoff functions as specified in the previous section. The game is said to be *dispersed* if μ is atomless.

An *equilibrium* τ of the game μ is an element of $\mathcal{M}(A \times \mathcal{U})$ with marginal measures τ_A and $\tau_{\mathcal{U}}$ such that (i) $\tau_{\mathcal{U}}$ is μ, and (ii) $\tau(B_\tau) = \tau(\{(u, a) \in (\mathcal{U} \times A) : u(a, \tau_A) \geq u(x, \tau_A) \text{ for all } x \in A\}) = 1$. An equilibrium τ can be *symmetrized* if there exist $h \in \text{Meas}(\mathcal{U}, A)$ and another equilibrium τ^s such that $\tau_A = \tau^s_A$ and $\tau^s(\text{Graph}_h) = 1$, where $\text{Graph}_h = \{(u, h(u)) \in (\mathcal{U} \times A) : u \in \mathcal{U}\}$. In this case, τ^s is said to be a *symmetric equilibrium.*

Theorem 4 *Every equilibrium of a dispersed large anonymous game can be symmetrized.*

Proof For an equilibrium τ of a dispersed large anonymous game μ, there exists $g \in \text{Meas}(\mathcal{U}, \mathcal{M}(A))$ such that $\tau(B) = \int_{\mathcal{U}} g(u; B_u) d\mu(u)$ where B is the Borel product σ-algebra on $\mathcal{U} \times A$ and B_u the section of B in A.[14] Since $\tau(B_\tau) = 1$, it is easy to see that for μ-almost all $u \in \mathcal{U}$, $g(u; (B_\tau)_u) = 1$ and supp $g(u) \subseteq (B_\tau)_u$. Apply Corollary 1 to the collection $\{(\mathcal{U}, \mathcal{B}(\mathcal{U})), \mu, A, g\}$ to obtain a pure strategy $h \in \text{Meas}(\mathcal{U}, A)$ such that $\mu h^{-1} = \int_{\mathcal{U}} g(u; \cdot) d\mu = \tau_A$ and $h(u) \in \text{supp } g(u) \subseteq (B_\tau)_u$ for μ-almost all $u \in \mathcal{U}$.

Let τ^s be the distribution $\mu(id_{\mathcal{U}}, h)^{-1}$. Then, $\tau^s_{\mathcal{U}} = \mu$ and $\tau^s_A = \mu h^{-1} = \tau_A$, and $B_\tau = B_{\tau^s}$. We can thus obtain that $h(u) \in (B_{\tau^s})_u$ for μ-almost all $u \in \mathcal{U}$. Since $\tau^s(\text{Graph}_h) = 1$, we have $\tau^s(B_{\tau^s}) = 1$, and hence τ^s is a symmetric equilibrium that is a symmetrization of τ.

7 Concluding remarks

The primary motivation of this short essay has been to show that the general purification principle of DWW can be applied synthetically to games based on atomless measure spaces: one-shot, finite player games with incomplete information as well as one-shot large non-anonymous and anonymous games, all in a finite-action setting. In conclusion, we note that the theorems invoked and extended here, along with the strengthened purification concepts to which they have been applied, may have further application to more general settings; in particular, to pure-strategy equilibria in atomless Bayesian games with infinite action sets or over time; see Yannelis and Rustichini (1991), Balder (2002), Khan and Sun (2002) and their references.

References

Ash RB (1972) Real analysis and probability. Academic, New York
Balder EJ (2002) A unifying pair of Cournot–Nash equilibrium existence results. J Econ Theor 102:437–470

[14] For the existence of such a disintegration, see for example, Ash (1972).

Dvoretsky A, Wald A, Wolfowitz J (1950) Elimination of randomization in certain problems of statistics and of the theory of games. Proc Nat Acad Sci USA 36:256–260
Dvoretsky A, Wald A, Wolfowitz J (1951a) Elimination of randomization in certain statistical decision procedures and zero-sum two-person games. Ann Math Stat 22:1–21
Dvoretsky A, Wald A, Wolfowitz J (1951b) Relations among certain ranges of vector measures. Pac J Math 1:59–74
Khan MA, Sun YN (1991) On symmetric Cournot–Nash equilibrium distributions in a finite-action, atomless game. In: Khan MA, Yannelis NC (eds) Equilibrium theory in infinite dimensional spaces. Springer, Berlin Heidelberg New York, pp 325–332
Khan MA, Sun YN (2002) Non-Cooperative games with many players. In: Aumann RJ, Hart S (eds) Handbook of game theory, vol 3. chap 46. Elsevier, Amsterdam, pp 1761–1808
Mas-Colell A (1984) On a theorem of Schmeidler. J Math Econ 13:201–206
Milgrom PR, Weber RJ (1985) Distributional strategies for games with incomplete information. Math Oper Res 10:619–632
Radner R, Rosenthal RW (1982) Private information and pure-strategy equilibria. Math Oper Res 7:401–409
Schmeidler D (1973) Equilibrium points of non-atomic games. J Stat Phys 7:295–300
Wald A, Wolfowitz J (1951) Two methods of randomization in statistics and the theory of games. Ann Math 53:581–586
Yannelis NC, Rustichini A (1991) Equilibrium points of non-cooperative random and Bayesian games. In: Aliprantis CD, Border KC, Luxemberg WAJ (eds) Positive operators, Riesz spaces, and economics. Springer, Berlin Heidelberg New York, pp 23–48