Discussion Papers of the
Max Planck Institute for
Research on Collective Goods
2020/20

Incomplete-Information Games in Large Populations with Anonymity

Martin F. Hellwig

MAX PLANCK SOCIETY

Discussion Papers of the
Max Planck Institute
for Research on Collective Goods
2020/20

Incomplete-Information Games in Large Populations with Anonymity

Martin F. Hellwig

August 2020

Max Planck Institute for Research on Collective Goods, Kurt-Schumacher-Str. 10, D-53113 Bonn
https://www.coll.mpg.de

Incomplete-Information Games in Large Populations with Anonymity *

Martin F. Hellwig
Max Planck Institute for Research on Collective Goods
Kurt-Schumacher-Str. 10, D - 53113 Bonn, Germany
hellwig@coll.mpg.de

August 17, 2020

Abstract

The paper provides theoretical foundations for models of strategic interdependence under uncertainty that have a continuum of agents and a decomposition of uncertainty into an macro component and an agent-specific micro component, with a law of large numbers for the latter. The decomposition of uncertainty is implied by a condition of *exchangeability* of agents' types, which is imposed equivalently imposed at the level of the prior or at the level of beliefs, i.e., posteriors. Under an additional condition of *anonymity in payoffs*, agents' behaviours are fully determined by their *macro beliefs* about the *cross-section distribution of types* and by the cross-section distribution of other agents' strategies. Any probability distribution over cross-section distributions of types is admissible, but not every macro belief function is compatible with a common prior.

Key Words: Incomplete-information games, large populations, belief functions, common priors, exchangeability, conditional independence, conditional exact law of large numbers.

JEL: C70, D82, D83.

*This paper originated in the challenge posed by a referee's asking for the mathematical foundations of the analysis of common priors in Hellwig (2011),an early version of Section 4.2 of this paper. For helpful discussions and advice, I am grateful to Thomas Mariotti and three referees, as well as Felix Bierbrauer, Christoph Engel, Alia Gizatulina, Christian Hellwig, and especially Yeneng Sun.

1 Introduction

This paper develops theoretical foundations for the study of incomplete-information games with the following properties:

- The payoff for any one agent depends only on the agent's own characteristics and actions and on the cross-section distribution of actions in the population.
- There are many agents, and each agent considers the effect of his own actions on the cross-section distribution of actions to be negligible.
- Uncertainty can be decomposed into an aggregate component and an agent-specific component, and the latter satisfies an exact law of large numbers.

Such games are not covered by the standard approach to studying strategic interdependence with incomplete information, which considers games with finitely many participants where each participants forms beliefs about every other participant's characteristics and actions. This approach foregoes the simplifications that are available if agents care only about the cross-section distribution of the other agents' actions and any one agent is too insignificant to affect the cross-section distribution of actions in the population.

Examples. A few examples illustrate the importance of such games.

Currency attacks and bank runs: In models of currency attacks and bank runs, the payoff to an agent's choice to attack or to run depends on how many agents are also choosing to attack or to run. Any one agent is therefore concerned about the fraction of people in the population that have received bad signals and are likely to speculate against the currency or run on the bank.[1]

Insider trading and market microstructure: Strategic behavior in markets with asymmetric information depends on agents' expectations about

[1] Since Morris and Shin (1998), the literature on currency attacks and bank runs has assumed that each agent privately observes a noisy signal of the fundamental. Given the observation of this signal, the agent forms expectations about the value of the fundamental and about the population share of the set of people who will choose to participate in a currency attack or a bank run. If the chances are good that this population share is high enough for the attack to be successful, the agent will also choose to participate. In addition to Morris and Shin (1998) see C. Hellwig (2002), Rochet and Vives (2004), Goldstein and Pauzner (2005), and Angeletos and Werning (2006).

the relative importance of information trading and liquidity trading. In organized markets in which the identities of traders are not revealed, these expectations concern the distribution of characteristics among the potential traders.[2]

Electoral competition and voting: In voting, the identities of individuals are irrelevant. Only the fractions of the population that vote for or against the given alternatives matter. In models of strategic voting, people form expectations about the distribution of other people's votes. This distribution depends on the distribution of other people's characteristics, i.e., preference parameters or realizations of information variables, and on how these characteristics affect their votes.[3]

Public-good provision and taxation: Models of income taxation usually assume a continuum of agents, with private information about individual productivity and with a law of large numbers for the cross-section productivity distribution.[4] Models of public-good provision usually assume a finite number of agents, with private information about preference parameters; public-good provision levels depend on aggregates of the preference parameters, e.g. the aggregate marginal benefit of an additional unit of the public good. The analysis of public-good provision and income taxation can be integrated in a model with a large population in which aggregate outcomes depend only on cross-section distributions of individual productivity levels and of preference parameters. In a large economy, these distributions are independent of any one agent's characteristics, and in mechanism design approach the distribution of reported characteristics are independent of any one agent's report.[5]

Issues. The notion that any one agent is too insignificant to affect aggregate outcomes is usually formalized by assuming that there is a continuum of agents. Uncertainty is decomposed into an aggregate component and an agent-specific component, and a law of large numbers is assumed for the latter.

This procedure raises several questions. First, what is the relation between these models and the standard Harsanyi/Mertens-Zamir model of strategic interdependence under incomplete information? Second, should we think of the decomposition of uncertainty into aggregate and agent-specific

[2] See, for example Kyle (1985, 1989).
[3] See, e.g. Lindbeck and Weibull (1987), Alesina and Tabellini (1990).
[4] Mirrlees (1971).
[5] See Bierbrauer (2009, 2014), Bierbrauer and Hellwig (2015).

components as being introduced *ad hoc* or can this decomposition itself be derived from some deeper properties of the models? Third, how should we deal with the mathematical difficulties inherent in the notion of a continuum of agents with agent-specific uncertainty?

The standard model of incomplete information relies on the notion of *types* that was introduced by Harsanyi (1967/68) and formalized by Mertens and Zamir (1985). If A is the set of agents, then, for each $a \in A$, there is a set T_a of possible types of agent a and, for each type $t_a \in T_a$, a belief $b_a(t_a)$ of agent a with the type t_a. The belief $b_a(t_a)$ is a probability distribution over vectors $(t_{a'})_{a' \neq a}$ of types of the other agents. Heuristically, we may think of agents using the information provided by the observation of their own types to form probabilistic beliefs about the other agents' types.

In one version of this approach, which was promoted by Harsanyi himself, agents' beliefs are treated as conditional probabilities under a common prior. Agents's types are determined by a move of nature, whose "strategy" is embodied in some commonly known prior probability distribution over type constellations. The game of incomplete information is thus treated as a game of *imperfect* information in which agents know the "strategy" but not the actual choices of nature.

The common-prior approach to modelling incomplete information has the advantage that it provides a unified framework for modelling uncertainty, information and beliefs. In most of this paper, I will therefore assume that there is a common prior. In the concluding section, however, I will argue that the condition of anonymity in beliefs can be applied to the belief $b_a(t_a)$ even if this belief is not derived from a common prior.[6]

Anonymity. The main contribution of the analysis will be to show that, in models with a continuum of agents, the properties listed above are implied by conditions of *anonymity,* which ensure that agents' names play no role. Most importantly, a condition of *anonymity in beliefs* ensures that each agent a with type t_a and probabilistic belief $b_a(t_a)$ thinks about the types $t_{a'}$ of agents $a' \neq a$ as the realizations of *essentially pairwise exchangeable random variables* so that their joint distribution under the belief $b_a(t_a)$ is unchanged if the names of any two of them are interchanged.

A second condition, *anonymity in payoffs*, postulates that an agent's payoff from any action depends on the other agents' actions *only* through the cross-section distribution of these actions. Which of the other agents is

[6] For a controversy about the common-prior approach, see Gul (1998), Aumann (1998).

taking which action makes no difference as long as the cross-section distribution of actions is the same.

If both anonymity conditions hold, in payoffs and in beliefs, the agent's expected payoff from any action depends on his or her expectations about other agents *only* through the agent's probabilistic expectations about the cross-section distribution of the other agents' types that are induced by the belief $b_a(t_a)$ and through the cross-section distribution of the other agents' strategies.

Anonymity in beliefs has the important implication that, from the perspective of agent a with belief $b_a(t_a)$, the other agents' types are essentially conditionally independent and identically distributed random variables. The conditioning variable, relative to which the other agents' types are conditionally independent, can be identified with the cross-section distribution of types. Moreover, with a continuum of agents, an exact law of large numbers implies that the cross-section distribution of types and the conditional probability distribution of any one agent's type coincide.[7] The decomposition of uncertainty into an aggregate component and an agent-specific component is thus an automatic by-product of exchangeability: The cross-section distribution of types represents the macro component, the different agents' individual types the micro-components.

The Mathematical Conundrum. With a continuum of agents, the formalization of the space of agents, of the underlying probability space, and of the random variables that determine agents' types requires some care. With finitely many agents, the vectors $(t_{a'})_{a' \in A}$ of types of all agents and $(t_{a'})_{a' \in A \setminus \{a\}}$ of types of agents other than a (about which agent a forms beliefs) can simply be treated as elements of the finite-dimensional product spaces $\prod\limits_{a' \in A} T_{a'}$ and $\prod\limits_{a' \in A \setminus \{a\}} T_{a'}$. With a continuum of agents, the product spaces $\prod\limits_{a' \in A} T_{a'}$ and $\prod\limits_{a' \in A \setminus \{a\}} T_{a'}$ are unsuitable because, for any given element $\{t_{a'}\}$ of such a space, the mapping $a' \to t_{a'}$ may be non-measurable so that the cross-section distributions of types is not well-defined. In particular, if A is the Lebesgue unit interval and the types $t_{a'}$ of different agents are assumed to be the realizations of (conditionally) independent random

[7] The insight that exchangeability is equivalent to conditional independence relative to some underlying σ-algebra is known as de Finetti's theorem, see de Finetti (1931, 1970/1974).

variables with nontrivial individual uncertainty, the functions $a' \to t_{a'}$ are non-measurable with probability one.[8]

To address this conundrum, Sun (2006) proposed to make the families of measurable sets so large that cross-section distributions are always well defined. If we think of the type $t_{a'}$ of agent a' as the realization of a random variable $\tilde{t}_{a'}(\cdot)$ that is defined on some probability space, the idea is to assume that the family of measurable sets on $A \times \Omega$, the product of the space of agents and the underlying probability space, is large enough so that, for any bounded measurable function f on $A \times \Omega$, integration of the function

$$(a', \omega) \to f(a', \tilde{t}_{a'}(\omega))$$

with respect to agents' names and with respect to states of the world is well defined and, moreover, the integrals have the Fubini property that the order of integration does not make a difference to the result.[9]

Given this formalism, Sun showed that a continuum of essentially pairwise independent and identically distributed random variables satisfies an exact law of large numbers; moreover, if the specified family of measurable sets is sufficiently rich, there is no restriction on the probability distribution of the random variables in question. Qiao et al. (2016) proved analogous results for families of *conditionally* independent random variables; they also established the equivalence between essential pairwise exchangeability and essential pairwise conditional independence with identical conditional probability distributions equal to the sample cross-section distributions.

In this paper, the result of Qiao et al. (2016) cannot be directly applied because it refers to a single underlying probability space. It can be applied to the common prior, but then the question is what that implies for the different agents' beliefs at different types. With beliefs as conditional probability distributions, what is the relation between the property of essential pairwise exchangeability of types at the level of the prior and the property of essential pairwise exchangeability of other agents' types at the level of the conditional distributions?

The paper shows that the two properties are actually equivalent in the sense that essential pairwise exchangeability of different agents' types at the level of the prior holds if and only if, under almost every agent's beliefs,

[8] This problem was first identified by Doob (1937, 1953). For early accounts in economics, see Judd (1985) and Feldman and Gilles (1985).

[9] Subsequent work has refined this approach. See in particular Sun and Zhang (2009), Podczeck (2010), and Qiao et al. (2016). Hammond and Sun (2003, 2008) develop a related approach that involves the limits of arbitrarily large finite samples from the given measure space of agents.

the other agents' types are essentially pairwise exchangeable, i.e. satisfy anonymity in beliefs, with probability one. This equivalence property provides a link between the macro-micro decomposition of uncertainty at the level of the prior, which is at the centre of most applied work, and the macro-micro decomposition at the level of beliefs, which are crucial for the analysis of strategic behaviour. In the absence of a common prior, one can still show that anonymity in beliefs induces a decomposition of uncertainty into a macro and a micro component - at the level of beliefs.

Macro Uncertainty and Macro Beliefs Quite generally, anonymity in beliefs implies that the probabilistic belief $b_a(t_a)$ of agent a with type t_a about the other agents' types is fully determined by what I call the agent's *macro belief* $b_a^*(t_a)$ about the cross-section distribution of the other agents' types. Given the macro belief $b_a^*(t_a)$, the underlying belief $b_a(t_a)$ about the overall constellation of other agents' types can be recovered by observing that the agent considers the other agents' types to be conditionally independent and identically distributed with a common conditional probability distribution that coincides with the cross-section distribution.

Similarly, at the level of the prior, with a decomposition of uncertainty in a macro component and a micro component, the macro component is summarized in the probability distribution over cross-section distributions of types that is induced by the common prior. From this probability distribution, the common prior itself can be recovered by using the fact that, conditional on the cross-section distribution of types, the types of different agents are essentially pairwise independent with a common conditional probability distribution equal to the cross-section distribution.

Given that this exercise can always be carried out, the requirement that a probability distribution on cross-section distributions of types can generate a common-prior model of incomplete information imposes no restriction on the scope of admissible macro uncertainty. *Any* probability distribution on cross-section distributions of types can be used as a basis for such a model.

Not every macro belief function, however, is compatible with a common prior. Whereas every probability distribution over cross-section distributions of types can be used to specify a common prior with associated belief and macro belief functions, *not every* macro belief function, and therefore not every belief function that satisfies anonymity in beliefs, is compatible with the existence of a common prior. If the values of the macro belief function, i.e., the probability distributions over cross-section type distributions that are induced by different observations of one's own type are mutually absolutely continuous, I will show that the macro belief function admits the existence of a common prior *if and only if* it satisfies a version of Harsanyi's (1967/68) consistency condition for the existence of a common prior in a certain two-player game of incomplete information.

Plan of the Paper. In the following, Section 2 develops the game theoretic formalism; Section 3 studies the scope for macro uncertainty and macro belief functions. Section 2.1 begins with a general formulation of a strategic game with an atomless continuum of players whose types are the realizations of random variables on a complete probability space. An assumption of anonymity in payoffs specializes the analysis to games in which agents care *only* about the distribution of other agents' actions, not about who does what. The formalism of a Fubini extension (of the product of the space of agents and the probability space) ensures that such distributions are well defined.

Section 2.2 introduces the properties of anonymity in beliefs and exchangeability of types and proves the equivalence result mentioned above. Section 2.3 shows that both properties are also equivalent to properties of conditional independence with identical conditional distributions and give rise to a macro-micro decomposition of uncertainty. Given this decomposition, Section 2.3 shows that all strategically relevant aspects of agents' beliefs are contained in their macro beliefs. The distribution of agents' strategies is then shown to be the key variable for studying strategic interdependence and equilibrium.

Section 3.1 shows that, if the Fubini extension is sufficiently rich, then, from an *ex ante* perspective, the formalism imposes no restriction on the scope of macro uncertainty, i.e., for any probability distribution over cross-section distributions of types, there exists a specification of exchangeable type random variables for different agents that generates the specified probability distribution over cross-section type distributions. Section 3.2 gives necessary and sufficient conditions under which a given macro belief function, i.e. a given function from an agent's own type to the agent's probabilistic beliefs about the cross-section distribution of types, is compatible with a common prior.

The concluding section discusses some open issues.

2 Incomplete-Information Games with a Continuum of Agents

2.1 Agents, Types, Anonymity in Payoffs

Let $(A, \mathcal{A}, \alpha)$ be a complete atomless measure space of agents with $\alpha(A) = 1$. Given this space of agents, I will consider imperfect-information games with the following structure. First, nature chooses a type t_a from a set T_a for each agent a. Then each agent a observes his or her own type and chooses an action s_a from a set S_a. Given the type t_a, the action s_a, and the actions $s_{a'}$ of the other agents, $a' \in A_{-a} := A\backslash\{a\}$, agent a receives the payoff $u_a(t_a, s_a, \{s_{a'}\}_{a' \in A_{-a}})$. The properties of the function u_a will be discussed later.

Because the agent observes t_a before choosing an action, the action is likely to depend on t_a. A *strategy* for the agent is a function $\sigma(\cdot, a)$ from T_a to S_a that indicates how the chosen action depends on t_a.

To model the move of nature, I assume that there is a complete probability space $(\Omega, \mathcal{F}, P)$ and, for each $a' \in A$, a random variable $\tau(\cdot, a')$ such that the type of agent a' is the realization of the random variable $\tau(\cdot, a')$. Given this stochastic specification, a strategy $\sigma(\cdot, a)$ of agent a is a best response to the strategies $\sigma(\cdot, a')$ of agents $a' \in A_{-a}$ if

$$\begin{aligned}
&\int_\Omega u_a(\tau(\omega, a), \sigma(\tau(\omega, a), a), \{\sigma(\tau(\omega, a'), a')\}_{a' \in A_{-a}}) dP(\omega) \\
\geq &\int_\Omega u_a(\tau(\omega, a), \hat{\sigma}(\tau(\omega, a)), \{\sigma(\tau(\omega, a'), a')\}_{a' \in A_{-a}}) dP(\omega) \qquad (2.1)
\end{aligned}$$

for all strategies $\hat{\sigma}(\cdot, a)$ that agent a might choose. A *non-cooperative equilibrium* is a strategy constellation $\{\sigma\cdot, a'\}_{a' \in A}$ such that, for (almost) all $a \in A$, $\sigma(\cdot, a)$ is a best response to $\{\sigma(\cdot, a')\}_{a' \in A\backslash\{a\}}$.

Without loss of generality, I assume that the type space T_a and the action space S_a are the same for all agents, i.e., for some T and S, $T_a = T$ and $S_a = S$ for all $a \in A$.[10] The type space T and the action space S are complete separable metric spaces; they are endowed with the Boral σ-algebras $\mathcal{B}(T)$ and $\mathcal{B}(S)$, and the spaces $\mathcal{M}(T)$ and $\mathcal{M}(S)$ of probability measures on T and S are endowed with the topology of weak convergence and the associated Borel σ-algebras $\mathcal{B}(\mathcal{M}(T))$ and $\mathcal{B}(\mathcal{M}(S))$.

[10] If different agents had different type spaces or action spaces, one could always replace them by the union of type spaces and the union of action spaces, with suitable assumptions about τ and $u_a, a \in A$, ensuring that the "added" types and actions are irrelevant.

The best-response condition (2.1) depends on the other agents' names. In the dependence of the payoff u_a on $\{s_{a'}\}_{a' \in A_{-a}}$, it can make a difference whether action $s' \in S$ is taken by agent a' and action s'' by agent a'' or the other way around. The following assumption eliminates this possibility.

Anonymity in Payoffs: There exists a continuous function $u_a^* : T \times S \times \mathcal{M}(S)$ such that

$$u_a(t_a, s_a, \{s_{a'}\}_{a' \in A_{-a}}) = u_a^*(t_a, s_a, D(\{s_{a'}\}_{a' \in A_{-a}})) \qquad (2.2)$$

for all $a \in A$, for all $t_a \in T_a, s_a \in S_a$, and all constellations $\{s_{a'}\}_{a' \in A_{-a}}$ of actions of agents $a' \in A_{-a}$ for which the cross-section distribution $D(\{s_{a'}\}_{a' \in A_{-a}})$ is well defined.

Anonymity in payoffs implies that agent a's payoff depends on the other agents' actions only through the cross-section distribution $D(\{s_{a'}\}_{a' \in A_{-a}})$. This property implies, in particular, that the agent's payoff is unchanged under any permutation of the other agents' names.[11]

The cross-section distribution $D(\{s_{a'}\}_{a' \in A_{-a}})$ is a measure on S such that, for any set $B \in \mathcal{B}(S)$,

$$D(B | \{s_{a'}\}_{a' \in A_{-a}}) = \alpha_{-a}(\{a' \in A_{-a} | s_{a'} \in B\}).$$

For this measure to be well defined, the map $a' \to s_{a'}$ must be measurable. For

$$\{s_{a'}\}_{a' \in A_{-a}} = \{\sigma(\tau(\omega, a'), a')\}_{a' \in A_{-a}},$$

this requirement is satisfied if the mappings $a' \to \tau(\omega, a')$ and $(t, a') \to \sigma(t, a')$ are measurable.

Since τ is part of the exogenous data, measurability of τ will be imposed by assumption. Following Sun (2006) and Qiao et al. (2016), I will assume that τ is measurable with respect to a Fubini extension of the product σ-algebra $\mathcal{F} \otimes \mathcal{A}$. To make this assumption precise, I first define the concept of a Fubini extension.

[11] If the measure space $(A, \mathcal{A}, \alpha)$ is homogeneous, for example, if $(A, \mathcal{A}, \alpha)$ is a hyperfinite Loeb space, invariance to measurable permutations of agents' names is in fact *equivalent* to the requirement that the other agents' actions affect the agent's payoff only through their cross-section distribution; see Khan and Sun (1999, Section 4). Notice that the Lebesgue unit interval is not homogeneous, so the requirement that the other agents' actions affect the agent's payoff only through their cross-section distribution is actually stronger than the requirement of invariance to measurable permutations of agents' names. The reason is that the Lebesgue σ-algebra is based on neighbourhood structures, and therefore the set of measurable permutations of names is smaller than required for equivalence.

Fubini Extension: Given two complete probability spaces $(\Omega, \mathcal{F}, P)$, and $(I, \mathcal{I}, \lambda)$, the probability space $(\Omega \times I, \mathcal{W}, Q)$ is a *Fubini extension* of the product space $(\Omega \times I, \mathcal{F} \otimes \mathcal{I}, P \otimes \lambda)$ if $\mathcal{F} \otimes \mathcal{I} \subset \mathcal{W}$, $Q|\mathcal{F} \otimes \mathcal{I} = P \otimes \lambda$, and, for any real-valued Q-integrable function f on $(\Omega \times I, W)$, (i) the sections $f(\cdot, i)$ and $f(\omega, \cdot)$ are integrable, respectively, on $(\Omega, \mathcal{F}, P)$ for λ-almost all $i \in I$, and on $(I, \mathcal{I}, \lambda)$ for P-almost all $\omega \in \Omega$, and (ii) the functions

$$i \to \int_\Omega f(\omega, i) dP(\omega) \quad \text{and} \quad \omega \to \int_I f(\omega, i) d\lambda(i) \tag{2.3}$$

are integrable, respectively, on $(I, \mathcal{I}, \lambda)$ and $(\Omega, \mathcal{F}, P)$ with

$$\int_{\Omega \times I} f(\omega, i) dQ \;=\; \int_\Omega \left[\int_I f(\omega, i) d\lambda(i) \right] dP(\omega) = \int_I \left[\int_\Omega f(\omega, i) dP(\omega) \right] d\lambda(i) \tag{2.4}$$

Remark 2.1 *Given two complete probability spaces $(\Omega, \mathcal{F}, P)$, and $(I, \mathcal{I}, \lambda)$ and a Fubini extension $(\Omega \times I, \mathcal{W}, Q)$ of the product space $(\Omega \times I, \mathcal{F} \otimes \mathcal{I}, P \otimes \lambda)$, let f be a $\mathcal{W}$-measurable function from $\Omega \times I$ to a complete separable metric space X, with Borel σ-algebra $\mathcal{B}(X)$. Then, for P-almost all $\omega \in \Omega$, the cross-section distribution $D(\{f(\omega, i)\}_{i \in I}) = \lambda \circ f(\omega, \cdot)^{-1}$ is well defined and the mapping*

$$\omega \to D(\{f(\omega, i)\}_{i \in I})$$

from $(\Omega, \mathcal{F})$ into the space $\mathcal{M}(X)$ of probability measures on X is measurable, where $\mathcal{M}(X)$ is endowed with the Borel σ-algebra $\mathcal{B}(\mathcal{M}(X))$ that is induced by the topology of weak convergence on $\mathcal{M}(X)$.

To reflect the fact that the probability space $(\Omega \times I, W, Q)$ has $(\Omega, \mathcal{F}, P)$, and $(I, \mathcal{I}, \lambda)$ as its marginal spaces, as required by the Fubini property, I write $W = \mathcal{F} \boxtimes \mathcal{I}$ and $Q = P \boxtimes \lambda$, so the notation $(\Omega \times I, \mathcal{F} \boxtimes \mathcal{I}, P \boxtimes \lambda)$ indicates that I refer to a Fubini extension of the product $(\Omega \times I, \mathcal{F} \otimes \mathcal{I}, P \otimes \lambda)$.

In the following, I will sometimes identify the space $(I, \mathcal{I}, \lambda)$ in the definition of a Fubini extension with the space $(A, \mathcal{A}, \alpha)$ of all agents and sometimes with the space $(A_{-a}, \mathcal{A}_{-a}, \alpha_{-a})$ of all agents other than a, where $\mathcal{A}_{-a}$ is the σ-algebra of sets in $\mathcal{A}$ that do not contain a and $\alpha_{-a} := \alpha|\mathcal{A}_{-a}$ is the restriction of the measure α to $\mathcal{A}_{-a}$. One easily checks that, if $(\Omega \times A, \mathcal{F} \boxtimes \mathcal{A}, P \boxtimes \alpha)$ is a Fubini extension of the product $(\Omega \times A, \mathcal{F} \otimes \mathcal{A}, P \otimes \alpha)$, then, for any $a \in A$, $(\Omega \times A_{-a}, \mathcal{F} \boxtimes \mathcal{A}_{-a}, P \boxtimes \alpha_{-a})$ is a Fubini extension of the product $(\Omega \times A_{-a}, \mathcal{F} \otimes \mathcal{A}_{-a}, P \otimes \alpha_{-a})$, where $\mathcal{F} \boxtimes \mathcal{A}_{-a}$ is the family of sets $X \subset \Omega \times A_{-a}$ such that $X = Y \backslash (\Omega \times \{a\})$ for some $Y \in \mathcal{F} \boxtimes \mathcal{A}$ for some $X = Y \backslash (\Omega \times \{a\})$ for some $Y \in \mathcal{F} \boxtimes \mathcal{A}$ and $P \boxtimes \alpha_{-a}$ is the restriction of $P \boxtimes \alpha$ to $\mathcal{F} \boxtimes \mathcal{A}_{-a}$.

Measurability of τ : The function τ is a measurable mapping from a Fubini extension $(\Omega \times A, \mathcal{F} \boxtimes \mathcal{A}, P \boxtimes \alpha)$ of the product probability space $(\Omega \times A, \mathcal{F} \otimes \mathcal{A}, P \otimes \alpha)$ to the type space T.

From Remark 2.1, one immediately obtains:[12]

Remark 2.2 *For P-almost all $\omega \in \Omega$ and any $a \in A$, the cross-section distribution of types of agents other than a, $D(\{\tau(\omega, a')\}_{a' \in A_{-a}}) = \alpha_{-a} \circ \tau(\omega, \cdot)^{-1}$, is well defined and the function*

$$\omega \to D(\{\tau(\omega, a')\}_{a' \in A})$$

is a measurable mapping from $(\Omega, \mathcal{F})$ into $\mathcal{M}(T)$.

Remark 2.3 *Assume that the mapping $(t, a') \to \sigma(t, a')$ from $T \times A_{-a}$ into S is measurable. Then for P-almost all $\omega \in \Omega$ and any $a \in A$, the cross-section distribution of actions of agents other than a,*

$$D(\{\sigma(\tau(\omega, a'), a'\}_{a' \in A_{-a}}) = \alpha_{-a} \circ \sigma(\tau(\omega, \cdot), \cdot)^{-1},$$

is well defined and the function

$$\omega \to D(\{\sigma(\tau(\omega, a'), a')\}_{a' \in A_{-a}})$$

is a measurable mapping from $(\Omega, \mathcal{F})$ into $\mathcal{M}(S)$.

Given the assumption of anonymity in payoffs, Remark 2.3 implies that, if the mapping $(t, a') \to \sigma(t, a')$ is measurable, the best-response condition (2.1) can be rewritten in the form

$$\begin{aligned} &\int_\Omega u_a^*(\tau(\omega, a), \sigma(\tau(\omega, a), a), D(\{\sigma(\tau(\omega, a'), a')\}_{a' \in A_{-a}}))dP(\omega) \\ \geq &\int_\Omega u_a^*(\tau(\omega, a), \hat{\sigma}(\tau(\omega, a), a), D(\{\sigma(\tau(\omega, a'), a')\}_{a' \in A_{-a}}))dP(\omega). \end{aligned} \tag{2.5}$$

[12]Here and elsewhere in the paper, it is useful to recall that, if Q is a measure on a space X and f is a measurable function from X to some other space Y, then $Q \circ f^{-1}$, a measure on Y, indicates the distribution of $f(x)$ that is induced when x is distributed as Q.

In this formulation, agents' names still matter. To be sure, agent a only cares about the distribution $D(\{\sigma(\tau(\omega, a'), a')\}_{a' \in A_{-a}})$ of the other agents' actions, but this distribution depends on the interplay between the type constellation $\{t_{a'}\}_{a' \in A_{-a}}$ and the strategy constellations $\{\sigma(\cdot, a')\}_{a' \in A_{-a}}$ of the other agents.

If the other agents all use the same (measurable) strategy $\sigma^* : T \to S$, this interplay takes a very simple form and one has

$$D(\{\sigma(\tau(\omega, a'), a')\}_{a' \in A_{-a}}) = D(\{\tau(\omega, a')\}_{a' \in A_{-a}}) \circ (\sigma^*)^{-1}, \tag{2.6}$$

so that the agent is only concerned about the cross-section distribution $D(\{\tau(\omega, a')\}_{a' \in A_{-a}})$ of the other agents' types and does not care about which agent has which type. However, the assumption that all other agents to use the same strategy is special. After all, strategy choices are endogenous. With enough symmetry in the exogenous data, equilibrium strategy choices may in fact be symmetric, but that would be a very special case. Even if the payoff function u_a^* was the same for all a, the assumptions that I have imposed so far are not sufficient for this conclusion.

2.2 Anonymity in Beliefs and Exchangeability of Types

If the other agents choose different strategies, asymmetries in the beliefs that agent a has about the types $t_{a'}, t_{a''}$ of agents a' and a'' may affect the agent's behaviour. To see the role of beliefs, note, that, if a regular conditional distribution $b_a(\tau(\cdot, a))$ for $\{\tau(\cdot, a')\}_{a' \in A_{-a}}$ given $\tau(\cdot, a)$ exists, one can rewrite (2.5) in the form

$$\begin{aligned}
&\int_\Omega u_a^*(\tau(\omega, a), \sigma(\tau(\omega, a), a), D(\{\sigma(t_{a'}, a')\}_{a' \in A_{-a}}))db_a(\{t_{a'}\}_{a' \in A_{-a}}|\tau(\omega, a))dP(\omega) \\
\geq &\int_\Omega u_a^*(\tau(\omega, a), \hat{\sigma}(\tau(\omega, a), a), D(\{\sigma(t_{a'}, a')\}_{a' \in A_{-a}}))db_a(\{t_{a'}\}_{a' \in A_{-a}}|\tau(\omega, a))dP(\omega).
\end{aligned} \tag{2.7}$$

Trivially, the strategy $\sigma(\cdot, a)$ satisfies the best-response condition (2.7) if and only if, for P-almost all $\omega \in \Omega$, the action $\sigma(\tau(\omega, a), a)$ maximizes the conditional expectation

$$\int_\Omega u_a^*(\tau(\omega, a), s_a, D(\{\sigma(t_{a'}, a')\}_{a' \in A_{-a}}))db_a(\{t_{a'}\}_{a' \in A_{-a}}|\tau(\omega, a)) \tag{2.8}$$

over the set S. Maximization of (2.8) however, depends on the belief $b_a(\cdot|\tau(\omega, a))$. If this belief treats the types $t_{a'}, t_{a''}$ of agents a' and a'' asymmetrically, the agent's best response to the other agents' strategies will reflect this asymmetry.

To eliminate the impact of the other agents' names on agent a's beliefs, I use a version of de Finetti's notion of *exchangeability.* The basic idea is that agent a regards the random variables $\tau(\cdot, a'), a' \in A$, as being symmetric in the sense that their joint distribution is unchanged by a permutation of the agents' names. Whereas de Finetti assumed mutual exchangeability, Hammond and Sun (2006, 2008) and Qiao et al. (2016) showed that, with a large family of random variables, mutual exchangeability is essentially equivalent to pairwise exchangeability. The word "essential" refers to the fact that the properties hold for all but a negligible set of random variables in the family.

Exchangeability Given two complete probability spaces $(\Omega, \mathcal{F}, P)$, and $(I, \mathcal{I}, \lambda)$ and a family $f(\cdot, i), i \in I$, of $\mathcal{F}$-measurable functions from Ω to a complete separable metric space X with Borel σ-algebra $\mathcal{B}(X)$, the random variables $f(\cdot, i)$, $i \in I$, are *essentially pairwise exchangeable* if there exists a Borel probability measure π on $(X \times X, \mathcal{B}(X) \times \mathcal{B}(X))$ such that, λ-almost all $i_1 \in I$,

$$P(f(\cdot, i_1)^{-1}(B_1) \cap f(\cdot, i_2)^{-1}(B_2)) = \pi(B_1 \times B_2)$$

for λ-almost all $i_2 \in I$ and for all $B_1, B_2 \in \mathcal{B}(X)$.

In the present context, there are two ways to think about exchangeability. First, given the role of beliefs in the objective function (2.8), one can think about exchangeability as a property of beliefs. Second, in a common-prior model, one can also think about exchangeability as a property of the initial move of nature.

In the following, I will consider both notions and study the relation between them, as well as their strategic implications. To avoid confusion, I will use the term *exchangeability of types* for the property at the level of the mapping τ and the term *anonymity in beliefs* for the property at the level of the belief $b_a(t_a)$ of agent a with type t_a.

Technically, the belief $b_a(t_a)$ is a measure on the range R_a of the mapping $\omega \to \boldsymbol{\tau}^a(\omega) := \{\tau(\omega, a')\}_{a' \in A_{-a}}$ where R_a is endowed with the σ-algebra $\mathcal{R}_a$, the coarsest σ-algebra under which the mapping $\omega \to \boldsymbol{\tau}^a(\omega)$ from $(\Omega, \mathcal{F})$ to R_a is measurable.

Anonymity in Beliefs For any $a \in A$ and $t_a \in T$, the measure $b_a(t_a)$ on $(R_a, \mathcal{R}_a)$ satisfies *anonymity in beliefs* if, under this measure, the types $t_{a'}$ of agents $a' \neq a$ are essentially pairwise exchangeable.

Exchangeability of Types Given the measurable mapping τ from the Fubini extension $(\Omega \times A, \mathcal{F} \boxtimes \mathcal{A})$ of the product space $(\Omega \times A, \mathcal{F} \otimes \mathcal{A})$ to the type space T, the random variables $\tau(\cdot, a), a \in A$, are essentially pairwise exchangeable.

If we think about beliefs as conditional distributions and about anonymity in beliefs as a property that holds almost surely, rather than just for some type t_a, then, from an *ex ante* perspective, anonymity in beliefs can be interpreted as a form of *conditional* exchangeability of the random variables $\tau(\cdot, a')$, $a' \in A_{-a}$. The following result shows that this conditional exchangeability is in fact equivalent to exchangeability.

Proposition 2.4 *Given a measurable mapping τ from the Fubini extension $(\Omega \times A, \mathcal{F} \boxtimes \mathcal{A}, P \boxtimes \alpha)$ of the product probability space $(\Omega \times A, \mathcal{F} \otimes \mathcal{A}, P \otimes \alpha)$ to the type space T and a function b from $T \times A$ to the space of probability measures on $(R_a, \mathcal{R}_a)$ such that, for α-almost all $a \in A$, $b(\tau(\cdot, a), a)$ is a regular conditional distribution for $\{\tau(\cdot, a')\}_{a' \in \mathcal{A}_{-a}}$ given $\tau(\cdot, a)$, the following statements are equivalent:*

(a) For α-almost every $a \in A$, for P-almost all $\omega \in \Omega$, the probability measure $b_a(\tau(\omega, a))$ satisfies anonymity in beliefs.

(b) The random variables $\tau(\cdot, a), a \in A$, are essentially pairwise exchangeable.

2.3 The Macro-Micro Decomposition of Uncertainty

Anonymity in beliefs and exchangeability of types have two important implications: First, they ensure that agents' best responses to the other agents' strategies do not depend on the other agents' names in the sense of who is doing what. Second, they provide for a decomposition of uncertainty into a macro and a micro component, with a law of large numbers holding for the latter.

For models with a continuum of random variables, Hammond and Sun (2003, 2008) have shown that the property of essential pairwise exchangeability is equivalent to the property of essential pairwise conditional independence relative to some countably generated σ-algebra, with identical conditional distributions. Moreover, Qiao et al. (2016) have shown that, with measurability relative to a Fubini extension, the conditioning σ-algebra can be identified with the algebra generated by the cross-section distributions of the random variables in question, and by a conditional law of large numbers, the conditional probability distribution of any one of the random variables and the cross-section sample distribution coincide. The following definition and propositions adapt their analysis to the present setting.

Essential Conditional Pairwise Independence Given two complete probability spaces $(\Omega, \mathcal{F}, P)$, and $(I, \mathcal{I}, \lambda)$, a countably generated sub-σ-algebra $\mathcal{C}$ of $\mathcal{F}$, and a family $f(\cdot, i), i \in I$, of $\mathcal{F}$-measurable functions from Ω to a complete separable metric space X with Borel σ-algebra $\mathcal{B}(X)$, the random variables $f(\cdot, i)$, $i \in I$, are *essentially pairwise conditionally independent given* $\mathcal{C}$ if, for λ-almost all $i_1 \in I$, the random variables $f(\cdot, i_1)$ and $f(\cdot, i_2)$ are conditionally independent given $\mathcal{C}$, for λ-almost all $i_2 \in I$.

Proposition 2.5 *Given a measurable mapping τ from the Fubini extension $(\Omega \times A, \mathcal{F} \boxtimes \mathcal{A}, P \boxtimes \alpha)$ of the product probability space $(\Omega \times A, \mathcal{F} \otimes \mathcal{A}, P \otimes \alpha)$ to the type space T, the following statements are equivalent:*

(a) The random variables $\tau(\cdot, a), a \in A$, are essentially pairwise exchangeable.

(b) The random variables $\tau(\cdot, a), a \in A$, are essentially pairwise conditionally independent given the sub-σ-algebra $\mathcal{D}$ of $\mathcal{F}$ that is generated by the mapping

$$\omega \to D(\{\tau(\omega, a)\}_{a \in A}), \tag{2.9}$$

and, moreover, for α-almost every $a \in A$, the mapping (2.9) is a regular conditional distribution for $\tau(\cdot, a)$ given $\mathcal{D}$.

Proposition 2.6 *Given a measurable mapping τ from the Fubini extension $(\Omega \times A, \mathcal{F} \boxtimes \mathcal{A}, P \boxtimes \alpha)$ of the product probability space $(\Omega \times A, \mathcal{F} \otimes \mathcal{A}, P \otimes \alpha)$ to the type space T and a function b from $T \times A$ to the space of probability measures on $(R_a, \mathcal{R}_a)$ such that, for a-almost all $a \in A$, $b(\tau(\cdot, a), a)$ is a regular conditional distribution for $\{\tau(\cdot, a')\}_{a' \in \mathcal{A}_{-a}}$ given $\tau(\cdot, a)$, the following statements are equivalent:*

(a) For α-almost every $a \in A$, for P-almost all $\omega \in \Omega$, the probability measure $b_a(\tau(\omega, a))$ satisfies anonymity in beliefs.

(b) For α-almost every $a \in A$, for P-almost all $\omega \in \Omega$, under the probability measure $b_a(\tau(\omega, a))$, the types $t_{a'}$ of agents $a' \neq a$ are essentially pairwise conditionally independent given the sub-σ-algebra $\hat{\mathcal{D}}$ of $\mathcal{R}_a$ that is generated by the mapping

$$\{t_{a'}\}_{a' \in A_{-a}} \to D(\{t_{a'}\}_{a' \in A_{-a}}), \tag{2.10}$$

and, moreover, for α_{-a}-almost every $a' \in A_{-a}$, the mapping (2.10) is a regular conditional distribution for $t_{a'}$ given $\hat{\mathcal{D}}$.

Propositions 2.5 and 2.6 have two components. One component asserts the equivalence of exchangeability of types or anonymity in beliefs with essential pairwise conditional independence (with identical conditional distributions). The other component asserts a conditional law of large numbers whereby the cross-section distribution of types is almost surely equal to the common conditional distribution of types given the σ-algebra that is generated by the cross-section distributions.

Conditional independence and the validity of the exact law of large numbers over the continuum of agents provide for a decomposition of uncertainty into macro and micro components. The macro component concerns the cross-section distribution of types, the micro component the type of each individual agent. Conditional independence and the law of large numbers ensure that, conditonal on the cross-section distribution of types, each agent's individual type has a probability distribution that is equal to the cross-section distribution.

The formulation of Proposition 2.6 relies on the specification of beliefs in terms of regular conditional distributions under a common prior. This formulation may create an impression that anonymity in beliefs and the macro-micro decomposition should really be treated as properties of the common prior. Such an impression would however be mistaken. The following result provides an analogue to Proposition 2.6 that does not refer to a common prior.

Proposition 2.7 *For any $a \in A$ and $t_a \in T$, the measure $b_a(t_a)$ satisfies anonymity in beliefs if and only if there exists a countably generated sub-σ-algebra $\mathcal{C}_a \subset \mathcal{R}_a$ and a $\mathcal{C}_a$-measurable mapping μ from R_a to the space of measures on $(R_a, \mathcal{R}_a)$ such that, under the measure $b_a(t_a)$, the types $t_{a'}$ of agents $a' \neq a$ are essentially pairwise conditionally independent and identically distributed with the common regular conditional distribution $\mu(\cdot)$. If the measure $b_a(t_a) \circ \boldsymbol{\tau}^a(\cdot)$ on $(\Omega, \mathcal{F})$ is absolutely continuous with respect to P, the sub-σ-algebra $\mathcal{C}_a$ coincides with the σ-algebra $\hat{\mathcal{D}}$ of $\mathcal{R}_a$ that is generated by the mapping*

$$\{t_{a'}\}_{a' \in A_{-a}} \to D(\{t_{a'}\}_{a' \in A_{-a}}), \tag{2.11}$$

and the regular conditional distribution $\mu(\cdot)$ coincides with the mapping (2.11).

The first half of Proposition 2.7 relies on Hammond and Sun (2008), the second half on Qiao et al. (2016). The assumption that $b_a(t_a) \circ \boldsymbol{\tau}^a(\cdot)$ on $(\Omega, \mathcal{F})$ be absolutely continuous with respect to P ensures that the projection mapping $(\{t_{a'}\}_{a' \in A_{-a}}, \hat{a}) \to t_{\hat{a}}$ is measurable with respect to a Fubini extension of the product σ-algebra $\mathcal{R}_a \otimes \mathcal{A}_{-a}$, as required for the exact law of large numbers inherent in the second half of the proposition.

2.4 Macro Beliefs and Strategic Behaviour

The fact that the macro-micro decomposition of uncertainty is obtained at the level of beliefs as well as the common prior is important because strategic behaviour depends on beliefs. From (2.7) and (2.8) above, we know that, with anonymity in payoffs, an agent's strategic behaviour depends on the agent's probabilistic beliefs about the cross-section distribution of the other agents' actions. The following result, which encompasses both Propositions 2.6 and 2.7, shows that this cross-section distribution of the other agents' actions can be expressed in terms of the cross-section distribution of the other agents' types and the cross-section distribution of the other agents' strategies.

Proposition 2.8 *Given $a \in A$ and $t_a \in T$, assume that, under the measure $b_a(t_a)$ on $(R_a, \mathcal{R}_a)$, the types $t_{a'}$ of agents $a' \neq a$ are essentially pairwise conditionally independent given the sub-σ-algebra $\hat{\mathcal{D}}$ of $\mathcal{R}_a$ that is generated by the mapping*

$$\{t_{a'}\}_{a' \in A_{-a}} \to D(\{t_{a'}\}_{a' \in A_{-a}}); \tag{2.12}$$

assume also that, for α_{-a}-almost every $a' \in A_{-a}$, the mapping (2.10) is a regular conditional distribution for $t_{a'}$ given $\hat{\mathcal{D}}$. If the mapping $(t, a') \to \sigma(t, a')$ from $T \times A_{-a}$ into S is measurable, then

$$D(\{\sigma(t_{a'}, a')\}_{a' \in A_{-a}}) = \int_{\hat{a} \in A_{-a}} D(\{t_{a'}\}_{a' \in A_{-a}}) \circ \sigma(\cdot, \hat{a})^{-1} d\alpha(\hat{a}) \tag{2.13}$$

for $b_a(t_a)$-almost all $\{t_{a'}\}_{a' \in A_{-a}} \in R_a$.

Thus, with anonymity in beliefs, the cross-section distribution of actions of agents other than a depends only on the cross-section distribution of types and the constellation $\{\sigma(\cdot, \hat{a})\}_{\hat{a} \in A_{-a}}$ of the other agents' strategies. In fact, equation (2.13) shows that the strategies $\sigma(\cdot, \hat{a}), \hat{a} \in A_{-a}$, affect the distribution $D(\{\sigma(t_{a'}, a')\}_{a' \in A_{-a}})$ only through their cross-section distribution $\Sigma^a := \alpha_{-a} \circ (\boldsymbol{\sigma}^a)^{-1}$. so one can write

$$D(\{\sigma(t_{a'}, a')\}_{a' \in A_{-a}}) = \int_{S^T} D(\{t_{a'}\}_{a' \in A_{-a}}) \circ (\sigma^*)^{-1} d\Sigma^a(\sigma^*), \tag{2.14}$$

which is unchanged under any permutation of the other agents' names. In (2.14), Σ^a is a measure on the space of measurable functions from T to S, and σ^*, the variable of integration, is a generic element of this function space.

With anonymity in beliefs as well as payoffs, expression (2.8) for agent a's conditionally expected payoff from action s_a with type $\tau(\omega, a)$ and belief $b_a(\tau(\omega, a))$ can then be written as

$$\int_{R_a} u_a^*(\tau(\omega, a), s_a, \int_{S^T} D(\{t_{a'}\}_{a' \in A_{-a}}) \circ (\sigma^*)^{-1} d\Sigma^a(\sigma^*))\, db_a(\{t_{a'}\}_{a' \in A_{-a}} | \tau(\omega, a)). \tag{2.15}$$

In this expression, the agent's belief $b_a(\tau(\omega, a))$ concerns only the cross-section type distribution $D(\{t_{a'}\}_{a' \in A_{-a}})$. One can therefore rewrite (2.15) in the form

$$\int_{\mathcal{M}(T)} u_a^*(\tau(\omega, a), s_a, \int_{S^T} \delta \circ (\sigma^*)^{-1} d\Sigma(\sigma^*))\, db_a^*(\delta | \tau(\omega, a)), \tag{2.16}$$

where, for any $t_a \in T$,

$$b_a^*(t_a) := b_a(t_a) \circ D(\cdot)^{-1}. \tag{2.17}$$

is the probability distribution for $D(\{t_{a'}\}_{a' \in A_{-a}})$ that is induced by $b_a(t_a)$.

I will refer to $b_a^*(t_a)$ as the *macro belief* of agent a with type t_a. Because the measure α assigns zero weight to the individual agent a, we also have

$$D(\{t_{a'}\}_{a' \in A_{-a}}) = D(\{t_{a'}\}_{a' \in A})$$

for all $\{t_{a'}\}_{a' \in A}$ and all a. Under exchangeability of types, therefore, for α-almost all $a \in A$ and P-almost all $\omega \in \Omega$, the macro belief $b_a^*(\tau(\omega, a))$ is independent of a. The fact that the measure α assigns zero weight to the individual agent a also implies that the distribution $\Sigma^a := \alpha_{-a} \circ (\boldsymbol{\sigma}^a)^{-1}$ of strategies pursued by agents other than a is the same for all a.

Thus, for $a \in A$ and P-almost all $\omega \in \Omega$, one can therefore rewrite (2.16) in the form

$$\int_{\mathcal{M}(T)} u_a^*(\tau(\omega, a), s_a, \Delta(\delta, \Sigma))\, db^*(\delta | \tau(\omega, a)), \tag{2.18}$$

where

$$\Delta(\delta, \Sigma) := \int_{S^T} \delta \circ (\sigma^*)^{-1} d\Sigma(\sigma^*)) \tag{2.19}$$

is the cross-section distribution of actions that is induced by the cross-section distribution of types δ and the cross-section distribution of actions Σ.

Anonymity in payoffs and anonymity in beliefs have thus been used to transform the objective function in (2.1), where agents rely on beliefs and expectations about the types and actions of every single other agent, into a form were agents rely only on beliefs and expectations about cross-section distributions of types and cross-section distributions of strategies of the other agents. Whereas the other agents' names play a substantive role in (2.8), and even more so in (2.1), they do not even appear in (2.16) or (2.18).

Expression (2.18) also indicates that, with exchangeability of types, the cross-section distribution of strategies is the key endogenous variable in any analysis of strategic behaviour and strategic interdependence. In this formulation, it is natural to think about (Bayes-Nash) equilibrium in terms of distributions.

Equilibrium Strategy Distribution A measure Σ on the space of measurable functions $\sigma^* : T \to S$ is an *equilibrium strategy distribution* if there exist strategies $\boldsymbol{\sigma}^a : T \to S, a \in A$, such that, (i) for α-almost every $a \in A$ and P-almost every $\omega \in \Omega$, the action $\boldsymbol{\sigma}^a(\tau(\omega, a))$ maximizes the objective (2.18) over S, and (ii) $\Sigma := \alpha \circ (\boldsymbol{\sigma}^a)^{-1}$.

I will *not* discuss under what conditions an equilibrium strategy distribution exists. Some of the issues that arise are routine, e.g., in addition to continuity of u_a^*, one needs a compactness condition on S or a boundary condition in u_a^* to ensure that, for any $\tau(\omega, a) \in T$ and $\Delta \in \mathcal{M}(S)$, there exists $s_a \in S$ that maximizes (2.18). One also needs an additional measurability condition on u_a^* to ensure that the action $\boldsymbol{\sigma}^a(\tau(\omega, a))$ that maximizes (2.18) over S can be taken to be measurable in $\tau(\omega, a)$ and a.

An important issue is nonroutine, however: Because an equilibrium strategy distribution is a measure on a space of functions, the topology on this space and the structural properties of utility functions and of the macro belief function must be specified in such a way that the continuity and compactness conditions for a fixed-point argument are satisfied. Milgrom and Weber (1985) have provided such conditions for models with finitely many agents (without anonymity). I conjecture that their arguments can be applied in the present as well.[13]

3 The Scope of Macro Uncertainty

3.1 Cross-Section Type Distributions

I now turn to the question whether anonymity in beliefs and exchangeability of types impose any additional implicit restrictions on the mapping τ. The answer to this question depends on the specification of the Fubini extension $\mathcal{F} \boxtimes \mathcal{A}$ of the product σ-algebra $\mathcal{F} \otimes \mathcal{A}$. For example, if $\mathcal{F} \boxtimes \mathcal{A}$ is equal to the product σ-algebra $\mathcal{F} \otimes \mathcal{A}$ itself, then, as was shown by Sun (2006) and Hammond and Sun (2008), exchangeability and (conditional) independence are incompatible with the measurability assumption on τ *except* for the case where $\tau(\omega, a)$ is the same for α-almost all a, for P-almost all ω. In this case, the type distribution $D(\{\tau(\omega, a)\}_{a \in A})$ would almost surely be degenerate, macro uncertainty would only concern the value of the type, which is common to (almost) all agents, and, conditional on the common value of the type, there would be no further individual uncertainty.[14]

As discussed by Sun (2006) and Qiao et al. (2016), this degeneracy is avoided if the Fubini extension $\mathcal{F} \boxtimes \mathcal{A}$ of the product σ-algebra $\mathcal{F} \otimes \mathcal{A}$ is *rich*. This requirement excludes the product σ-algebra $\mathcal{F} \otimes \mathcal{I}$.

Richness of the Fubini Extension A Fubini extension $(\Omega \times I, \mathcal{F} \boxtimes \mathcal{I}, P \boxtimes \lambda)$ of a product probability space $(\Omega \times I, \mathcal{F} \otimes \mathcal{I}, P \otimes \lambda)$ is said to be *rich* if there exists a measurable function h from $(\Omega \times I, \mathcal{F} \boxtimes \mathcal{I}, P \boxtimes \lambda)$ to the unit interval such that (i) the random variables $h(\cdot, i)$, $i \in I$, are essentially pairwise independent, i.e., for λ-almost all $i_1 \in I$, the random variables $h(\cdot, i_1)$ and $h(\cdot, i_2)$ are independent for λ-almost all $i_2 \in I$, and, moreover, (ii) for λ-almost every $i \in I$, the random variable $h(\cdot, i)$ has a uniform distribution.

[13] I also conjecture that, with an atomless measure space of agents, there is no need to allow for (mixed) behaviour strategies rather than pure strategies.

[14] Proposition 2.1 of Sun (2006) shows that, if h is a measurable function from the product space $(\Omega \times I, \mathcal{F} \otimes \mathcal{I}, P \otimes \lambda)$ to the unit interval and if the random variables $h(\cdot, i)$, $i \in I$, are essentially pairwise independent, the random variables $h(\cdot, i)$, $i \in I$, must be essentially trivial, i.e., for λ-almost all $i \in I$, $h(\cdot, i)$ must be constant. Proposition 4 in Hammond and Sun (2008) provides a version of this result with essential pairwise conditional independence.

Conditions for the existence of a rich Fubini extension are given in Sun (2006), Sun and Zhang (2009), and Podczeck (2010). In particular, Sun (2006) shows that a rich Fubini extension exists if $(I, \mathcal{I}, \lambda)$ is a hyperfinite Loeb space. Sun and Zhang (2009) show that, whereas a rich Fubini extension fails to exist if I is the unit interval with the Lebesgue σ-algebra, an extended Lebesgue unit interval, with a larger σ-algebra, does permit the construction of a rich Fubini extension of the product $(\Omega \times I, \mathcal{F} \otimes \mathcal{I},\, P \otimes \lambda)$.

The following proposition shows that, if the Fubini extension $\mathcal{F} \boxtimes \mathcal{A}$ is rich, there is *no restriction on macro uncertainty*, i.e. uncertainty about cross-section type distributions, and the only restriction on micro uncertainty comes from the principle that, for a given cross-section distribution of types, the conditional probability distribution of the random variable $\tau(\cdot, a)$ is equal to the cross-section distribution of types.

Proposition 3.1 *Let $\tilde{\delta}$ be any $\mathcal{M}(T)$-valued random variable on $(\Omega, \mathcal{F}, P)$ and let $\mathcal{D}$ be the sub-σ-algebra of $\mathcal{F}$ that is generated by $\tilde{\delta}$. If the Fubini extension $(\Omega \times A, \mathcal{F} \boxtimes \mathcal{A}, P \boxtimes \alpha)$ is rich, there exists an $\mathcal{F} \boxtimes \mathcal{A}$-measurable mapping τ from $\Omega \times A$ to T such that the following statements hold:*

(a) Exchangeability of types holds, i.e., the random variables $\tau(\cdot, a'), a' \in A$, are essentially pairwise exchangeable.

(b) Conditionally on $\mathcal{D}$, the random variables $\tau(\cdot, a'), a' \in A$, are essentially pairwise independent and, for α-almost every $a' \in A$, the mapping $\omega \to \tilde{\delta}(\omega)$ is a regular conditional distribution for $\tau(\cdot, a')$ given $\mathcal{D}$.

(c) For P-almost all $\omega \in \Omega$,

$$\tilde{\delta}(\omega) = D(\{\tau(\omega, a')\}_{a' \in A}), \tag{3.1}$$

i.e., conditionally on $\mathcal{D}$, an exact law of large numbers holds.

(d) For α-almost every $a \in A$, there exists a function b_a from T to the space of measures on $(R_a, \mathcal{R}_a)$ such that $b_a(\tau(\cdot, a))$ is a regular conditional distribution for $\{\tau(\cdot, a')\}_{a' \in A_{-a}}$ given $\tau(\cdot, a)$ and, moreover, for P-almost every $\omega \in \Omega$, $b_a(\tau(\cdot, a))$ satisfies anonymity in beliefs.

(e) The belief functions b_a in Statement (d) take the form

$$b_a(t_a) = \int_{\mathcal{M}(T)} \hat{b}_a(\delta)\, d\beta^*(\delta | t_a), \tag{3.2}$$

where $\hat{b}_a(\tilde{\delta}(\cdot))$ is a regular conditional distribution for $\{\tau(\cdot, a')\}_{a' \in A_{-a}}$ given $\tilde{\delta}$ and $\beta^*(\tau(\cdot, a))$ is a regular conditional distribution for $\tilde{\delta}$ given $\tau(\cdot, a)$.

In addition to pulling together the findings of Propositions 2.4 - 2.6, Proposition 3.1 shows that the objects of those propositions, namely the families $\tau(\cdot, a)$, $b_a(\tau(\cdot, a)), a \in A$, of random types and of belief functions can always be well defined. Moreover, for any specification of the macro random variable $\tilde{\delta}$, they can be chosen so that $\tilde{\delta}$ almost surely coincides with the cross-section distribution of $\tau(\cdot, a), a \in A$, as well as the conditional probability distribution given $\tilde{\delta}$ of $\tau(\cdot, a')$, for α-almost all $a' \in A$. Given that, for any $\Phi \in \mathcal{M}(\mathcal{M}(T))$, there is a random variable $\tilde{\delta}$ on $(\Omega, \mathcal{F}, P)$ whose probability distribution is Φ, it follows that, if the Fubini extension $(\Omega \times I, \mathcal{F} \boxtimes \mathcal{I}, P \boxtimes \lambda)$ is rich, any probability distribution over cross-section distributions is admissible.

The existence result in Statement (d) is nontrivial because the σ-algebra $\mathcal{R}_a$ is not, in general, countably generated. The result follows from the law of iterated expectations as spelled out in (3.2). Existence of a regular conditional distribution $\hat{b}_a(\tilde{\delta}(\cdot))$ for $\{\tau(\cdot, a')\}_{a' \in A_{-a}}$ given $\tilde{\delta}$ is obtained from the very construction of the random variables $\tau(\cdot, a'), a' \in A_{-a}$. Existence of a regular conditional distribution $\beta(\tau(\cdot, a))$ for $\tilde{\delta}$ given $\tau(\cdot, a)$ is obtained by standard arguments from the fact that T and $\mathcal{M}(T)$ are complete separable metric spaces and that the σ-algebras $\mathcal{B}(T)$ and $\mathcal{B}(\mathcal{M}(T))$ are countably generated.

3.2 Macro Belief Functions

Whereas Proposition 3.1 implies that every probability distribution over cross-section distributions is admissible, the same cannot be said for macro belief functions. Not every measurable function β from T to $\mathcal{M}(\mathcal{M}(T))$ is compatible with a common prior.

As is well known, in models with finitely many agents, with arbitrary belief functions, the existence of a common prior cannot be taken for granted.[15] In such models, the conditions under which a given set of belief functions is compatible with a common prior are very restrictive, the more so, the more agents there are. In the present setting, with a continuum of agents and belief functions required to satisfy anonymity in beliefs, conditions for compatibility with a common prior are less restrictive than the size of the population might suggest, but even so, there is a problem.

Compatibility with a Common Prior A macro belief function $\beta : T \to \mathcal{M}(\mathcal{M}(T))$ admits a common prior if there exists a mapping $\tau : \Omega \times A \to T$ that is measurable with respect to a rich Fubini extension

[15] Harsanyi (1967/68), Samet (1998 a, b), Feinberg (2000), Rodrigues-Neto (2009), Hellman and Samet (2012), Hellwig (2013).

$\mathcal{F} \boxtimes \mathcal{A}$ of the product σ-algebra $\mathcal{F} \otimes \mathcal{A}$, and, for α-almost every $a \in A$, there exists a regular conditional distribution $b_a(\cdot|\tau(\cdot, a))$ for $\{\tau(\cdot, a'\}_{a' \in \mathcal{A}_{-a}}$ given $\tau(\cdot, a)$ such that for P-almost every $\omega \in \Omega$, $b_a(\cdot|\tau(\omega, a))$ satisfies anonymity in beliefs and the associated macro belief $b_a^*(\tau(\omega, a))$ coincides with $\beta(\tau(\omega, a))$.

Proposition 3.2 *A measurable function $\beta : T \to \mathcal{M}(\mathcal{M}(T))$ admits a common prior if and only if there exist measures $\Psi \in \mathcal{M}(T))$, $\Phi \in \mathcal{M}(\mathcal{M}(T))$, $\Pi \in \mathcal{M}(\mathcal{T} \times \mathcal{M}(T))$ such that*

$$\Pi(B_1 \times B_2) = \int_{B_1} \beta(B_2|t)\, d\Psi(t) \tag{3.3}$$

and

$$\Pi(B_1 \times B_2) = \int_{B_2} \delta(B_1)\, d\Phi(\delta) \tag{3.4}$$

for all $B_1 \in \mathcal{B}(T)$ and $B_2 \in \mathcal{B}(\mathcal{M}(T))$.

To understand this proposition, let $\tau(\cdot, a')$, $a' \in A$, be the family of random types for which β is supposed to be the macro belief function. Let $\tilde{\delta} = D(\{\tau(\cdot, a')\}_{a' \in A})$ be the random variable indicating the cross-section distribution of types, and let $a \in A$ be such that, conditionally on $\tilde{\delta}(\cdot)$, $\tau(\cdot, a)$ is distributed as $\tilde{\delta}(\cdot)$. Let Π be the joint distribution of the pair $(\tau(\cdot, a), \tilde{\delta}(\cdot))$, and let Ψ and Φ be the marginal distributions of $\tau(\cdot, a)$ and $\tilde{\delta}(\cdot)$.

There are two ways to think about Π. First, using the fact that $\beta(\tau(\cdot, a))$ is a regular conditional distribution for $\tilde{\delta}(\cdot)$ given $\tau(\cdot, a)$, one can think about Π as being derived from the marginal distribution Ψ of the type $\tau(\cdot, a)$ of agent a and the macro belief function β. Second, one can think about Π as being derived from the marginal distribution Φ of $\tilde{\delta}(\cdot)$ in combination with the fact that marginal distributions of $\tau(\cdot, a)$ and $\tilde{\delta}(\cdot)$. The first approach yields (3.3), the second (3.4).

Consistency of (3.3) and (3.4) requires that

$$\int_{B_2} \delta(B_1)\, d\Phi(\delta) = \int_{B_1} \beta(B_2|t)\, d\Psi(t) \tag{3.5}$$

for all $B_1 \in \mathcal{B}(T)$ and $B_2 \in \mathcal{B}(\mathcal{M}(T))$. In order to understand what this means, it is useful to note that this condition is formally equivalent to the condition for the existence of a common prior in a *two-player model* in which the type space of player 1 is T, the type space of player 2 is $\mathcal{M}(T)$, the belief function of player 1 is β, and the belief function of player 2 is the identity mapping $\delta \to \delta$. In this two-player model, a common prior Π exists if and only if there exist agent-specific priors Ψ for the type of player 1 and Φ for the type of player 2 such that equation (3.5) holds for all $B_1 \in \mathcal{B}(T)$ and $B_2 \in \mathcal{B}(\mathcal{M}(T))$, in which case Π is given by (3.3) and (3.4).

In the following, I use arguments from the analysis of two-player games to spell out the meaning of the consistency condition (3.5). The aim is to obtain conditions that only refer to the macro belief function β and not also to the measures Ψ and Φ, which are endogenous. I begin with a result showing that Proposition 3.2 can be restated in terms of density functions.[16]

Proposition 3.3 *A measurable function $\beta : T \to \mathcal{M}(\mathcal{M}(T))$ admits a common prior if and only if there exist measures $\Psi \in \mathcal{M}(T))$, $\Phi \in \mathcal{M}(\mathcal{M}(T))$, $\Pi \in \mathcal{M}(\mathcal{T} \times \mathcal{M}(T))$ such that the following statements are true.*

(a) Φ-almost every measure $\delta \in \mathcal{M}(T)$ is absolutely continuous with respect to Ψ and has a density function g_Ψ such that

$$\delta(B_1) = \int_{B_1} g_\Psi(t|\delta)\, d\Psi(t) \tag{3.6}$$

for all $B_1 \in \mathcal{B}(T)$.

(b) For Ψ-almost every $t \in T$, the measure $\beta(t)$ is absolutely continuous with respect to Φ and has a density function f_Φ such that

$$\beta(B_2|t) = \int_{B_2} f_\Phi(\delta|t)\, d\Phi(\delta) \tag{3.7}$$

for all $B_2 \in \mathcal{B}(\mathcal{M}(T))$.

(c) The measure Π is absolutely continuous with respect to the product measure $\Psi \times \Phi$ and has a density function π such that

$$\Pi(B_1 \times B_2) = \int_{B_1} \int_{B_2} \pi(t,\delta)\, d\Phi(\delta) d\Psi(t); \tag{3.8}$$

moreover,

$$\pi(t,\delta) = f_\Phi(\delta|t) = g_\Psi(t|\delta) \tag{3.9}$$

for $\Psi \times \Phi$-almost all $(t,\delta) \in T \times \mathcal{M}(T)$.

[16] This restatement reflects the insight of Samet (1998a) insight that, if a common prior exists, then the marginal distributions of the types of the different participants can be represented as the invariant distributions of Markov processes, with kernels given by compositions of the belief functions. As discussed, e.g., in Doob (1953), Markov kernels are absolutely continuous with respect to the invariant distributions.

The consistency condition (3.9) provides the basis for the following result.

Proposition 3.4 *Let β be a measurable function from T to $\mathcal{M}(\mathcal{M}(T))$ and assume that the measures $\beta(t), t \in T$, are mutually absolutely continuous. If β admits a common prior, then the following statements hold:*

(i) there exists a set $D \in \mathcal{M}(T)$ such that $\beta(D|t) = 1$ for all t and, moreover, the measures $\delta \in D$ are mutually absolutely continuous;

(ii) For any $t_0 \in T$, for $\beta(t_0)$-almost all $\delta_1 \in D$ and δ_1-almost all $t_1 \in T$, there exist density functions $f_1(\cdot|t)$ of the measures $\beta(t), t \in T$, with respect to $\beta(t_1)$, and $g_1(\cdot|\delta)$ of the measures $\delta \in D$ with respect to δ_1 so that the condition

$$\frac{f_1(\delta_2|t_2)}{f_1(\delta_1|t_2)} = \frac{g_1(t_2|\delta_2)}{g_1(t_1|\delta_2)} > 0 \tag{3.10}$$

holds for $\beta(t_1)$-almost all $\delta_2 \in \mathcal{M}(T)$ and δ_1-almost all $t_2 \in T$.

Conversely, if β satisfies (i) and (ii), then β admits a common prior. The common prior is unique. The measures Π, Ψ, Ψ take the form

$$\begin{aligned} \Pi(B_1 \times B_2) &= \lambda(t_1, \delta_1) \int_{B_1} \int_{B_2} \frac{f_1(\delta|t)}{f_1(\delta_1|t)}\, d\beta(\delta|t_0) d\delta_0(t) \\ &= \lambda(t_1, \delta_1) \int_{B_2} \int_{B_1} \frac{g_1(t|\delta)}{g_1(t_1|\delta)}\, d\delta_0(t) d\beta(\delta|t_0), \end{aligned} \tag{3.11}$$

$$\Psi(B_1) = \lambda(t_1, \delta_1) \int_{B_1} \frac{1}{f_1(\delta_0|t)}\, d\delta_0(t) \tag{3.12}$$

$$\Phi(B_2) = \lambda(t_1, \delta_1) \int_{B_2} \frac{1}{g_1(t_0|\delta)}\, d\beta(\delta|t_0), \tag{3.13}$$

for $B_1 \in \mathcal{B}(T)$ and $B_2 \in \mathcal{M}(T)$, where $\lambda(t_1, \delta_1) > 0$ is a scaling factor ensuring that $\Pi(T \times \mathcal{M}(T)) = 1$.

In this proposition, condition (3.10) takes the place of the consistency condition (3.9) in Proposition 3.3. Both conditions are variants of Harsanyi's (1967/68) well known necessary condition for the existence of a common prior for a given belief system.

Indeed the underlying argument is the same: If β admits a common prior, there are two ways to evaluate the ratio $\frac{\pi(t_2,\delta_2)}{\pi(t_1,\delta_1)}$ of the joint distribution Π of an agent's type and the cross-section distribution of types. One can write

$$\frac{\pi(t_2, \delta_2)}{\pi(t_1, \delta_1)} = \frac{\pi(t_2, \delta_1)}{\pi(t_1, \delta_1)} \cdot \frac{\pi(t_2, \delta_2)}{\pi(t_2, \delta_1)} = \frac{g_\Psi(t_2|\delta_1)}{g_\Psi(t_1|\delta_1)} \cdot \frac{f_\Phi(\delta_2|t_2)}{f_\Phi(\delta_1|t_2)} \tag{3.14}$$

or, alternatively,

$$\frac{\pi(t_2,\delta_2)}{\pi(t_1,\delta_1)} = \frac{\pi(t_1,\delta_2)}{\pi(t_1,\delta_1)} \cdot \frac{\pi(t_2,\delta_2)}{\pi(t_1,\delta_2)} = \frac{f_\Phi(\delta_2|t_1)}{f_\Phi(\delta_1|t_1)} \cdot \frac{g_\Psi(t_2|\delta_2)}{g_\Psi(t_1|\delta_2)}, \tag{3.15}$$

where in each case the second equation is based on (3.9). For these evaluations of the ratio $\frac{\pi(t_2,\delta_2)}{\pi(t_1,\delta_1)}$ to be compatible with each other, one must have

$$\frac{g_\Psi(t_2|\delta_1)}{g_\Psi(t_1|\delta_1)} \cdot \frac{f_\Phi(\delta_2|t_2)}{f_\Phi(\delta_1|t_2)} = \frac{f_\Phi(\delta_2|t_1)}{f_\Phi(\delta_1|t_1)} \cdot \frac{g_\Psi(t_2|\delta_2)}{g_\Psi(t_1|\delta_2)}. \tag{3.16}$$

Whereas equation (3.16) involves densities with respect to Ψ and Φ, the mutual absolute continuity of Ψ and the measures $\delta \in D$ and the mutual absolute continuity Φ and the measures $\beta(t), t \in T$, imply that (3.16) can be rewritten as

$$\frac{g_0(t_2|\delta_1)}{g_0(t_1|\delta_1)} \cdot \frac{f_0(\delta_2|t_2)}{f_0(\delta_1|t_2)} = \frac{f_0(\delta_2|t_1)}{f_0(\delta_1|t_1)} \cdot \frac{g_0(t_2|\delta_2)}{g_0(t_1|\delta_2)}, \tag{3.17}$$

where $g_0(t_2|\delta_1)$, $f_0(\delta_2|t_2)$, etc. are the corresponding densities with respect to some $\delta_0 \in D$ and $\beta(t_0) \in \mathcal{M}(T)$. Equation (3.17) is exactly Harsanyi's (1967/68) condition, albeit applied to densities in a model with (possibly) a continuum of states, rather than probabilities in a model with a finite number of states. To get from this equation to (3.10), it suffices to set $\delta_0 = \delta_1$ and $t_0 = t_1$ and to note that $g_1(t_2|\delta_1) = g_1(t_1|\delta_1) = 1$ and $f_1(\delta_2|t_1) = f_1(\delta_2|t_1) = 1$ because the density of a measure with respect to itself is identically equal to one.

Whereas Harsanyi's condition is usually discussed as a necessary condition for the existence of a common prior, Proposition 3.4 shows that, under the given conditions, it is also sufficient. This finding hinges on the strict positivity of the densities $f_0(\delta|t)$ and $g_0(t|\delta)$ on the relevant parts of their domains, which in turn is derived from the assumption that the macro beliefs $\beta(t), t \in T$, are mutually absolutely continuous.

The sufficiency part of Proposition 3.4 parallels the finding of Hellwig (2013) that, in an n-player game in which any element of any player's information partition intersects any element of any other player's information partition, for a *strictly positive belief system*, a common prior exists if (and only if) the Harsanyi condition holds for all quadruples that can be obtained from pairs of types for pairs of players (keeping the other players' types fixed). I conjecture that, without the mutual-absolute-continuity assumption, necessary and sufficient conditions along the lines of Rodrigues-Neto (2009) or Hellman and Samet (2012) could still be obtained.

4 Further Considerations

Anonymity in Beliefs in the Absence of a Common Prior. To conclude the paper, I briefly discuss some further issues. First, as mentioned in the introduction, the interpretation of belief functions as regular conditional distributions is controversial. One may therefore ask what becomes of the results of this paper when there is no common prior.

In the absence of a common prior, the belief $b_a(t_a)$ of agent a with type t_a must be taken as a given, without any relation to a prior, common or not. One can still impose the property of anonymity in beliefs, and, by the result of Hammond and Sun (2008), one still finds that, if $b_a(t_a)$ has this property, then, under this belief, relative to some countably generated σ-algebra, the types $t_{a'}, a' \in A_{-a}$, are conditionally independent and identically distributed. To go further and assert a conditional law of large numbers, one needs the formalism of the Fubini extension.

In Hellwig (2019), a previous version of this paper, I actually started from the beliefs $b_a(t_a), t_a \in T, a \in A$, with an assumption that, for some complete probability space $(\Omega_a(t_a), \mathcal{F}_a(t_a), P_a(t_a))$, the belief $b_a(t_a)$ is given as

$$b_a(t_a) = P_a(t_a) \circ \boldsymbol{\tau}^a(\cdot|t_a)^{-1},$$

where

$$\boldsymbol{\tau}^a(\cdot|t_a) = \{\tau_a(\cdot, a'|t_a)\}_{a' \in A_{-a}}$$

and $\tau_a(\cdot, \cdot|t_a)$ is a mapping from $\Omega_a(t_a) \times A_{-a}$ to T that is measurable with respect to a rich Fubini extension of the product space $(\Omega_a(t_a) \times A_{-a}, \mathcal{F}_a(t_a) \otimes \mathcal{A}_{-a}, P_a(t_a) \otimes \alpha_{-a})$. With this formalism, a version of Proposition 2.7 is immediately available to provide for a macro-micro decomposition of uncertainty, with an exact conditional law of large numbers, from the perspective of the belief $b_a(t_a)$.

In this approach, however, where each belief is treated in isolation, with a distinct probability space $(\Omega_a(t_a), \mathcal{F}_a(t_a), P_a(t_a))$ supporting the Fubini formalism, it is difficult to think about a belief *function,* at least if the functions $\boldsymbol{\tau}^a(\cdot|t_a)$, $t_a \in T$, have different ranges. The Fubini formalism requires that the probability spaces $(\Omega_a(t_a), \mathcal{F}_a(t_a), P_a(t_a))$ be complete, so even if $\Omega_a(t_a)$ was taken to be the same for all t_a, the completions $\mathcal{F}_a(t_a), P_a(t_a)$ of the σ-algebras $b_a(t_a) \circ \boldsymbol{\tau}^a(\cdot|t_a)$ need not be. As shown in Hellwig (2019), the difficulty is resolved if one assumes that the measures $b_a(t_a), t_a \in T$, are mutually absolutely continuous; in this case, one can replace any one of the spaces $\Omega_a(t_a)$ with the union of the ranges of the functions $\boldsymbol{\tau}^a(\cdot|t_a)$. Proposition 2.7 then applies to all the beliefs $b_a(t_a), t_a \in T$.

What about Payoff-relevant Aspects of Names? Second, a referee has asked how the formalism would accommodate *labels* that may be strategically relevant even though they are parts of agents' names. The example given by the referee is location, but one can also think of profession, gender, ethnicity, age. Indeed, as shown by Bertrand and Mullainathan (2004), even proper names can be payoff relevant if they are treated as signals of gender, ethnicity, or race.

In addressing the referee's question, one must be careful about the semantics of the word "name". If we think about "names" as abstract identifiers like IP numbers, which in and of themselves are strategically unimportant, then, by definition, labels must be treated as parts of agents' types, rather than their names. Paradoxical though it may seem, in this interpretation of the word "name", whether a person is called Marianne or Sendhil, would be an aspect of the person's type.

Does it make a difference? If we think of labels as being parts of agents' names, the space of agents take the form $A = \hat{A} \times L$, where $\hat{A}$ is a space of strategically irrelevant identifiers and L is the space of labels. For any agent $a \in A$, the value of the label for this agent is given by the projection from A to L, $\ell(a) = proj_L a$. If instead we treat the label ℓ as a part of the agent's type, we must replace the type $\tau(\omega, a)$ of this paper by an extended type $(\tau(\omega, a), \ell(\omega, a))$. The latter formulation has certain advantages.

If labels are treated as parts of agents' names, it may be appropriate to replace the assumption of anonymity in beliefs or exchangeability of types by conditioning on labels, e.g., assuming that for $\alpha \circ \ell^{-1}$-almost all $\ell \in L$, the types $\tau(\cdot, \hat{a}, \ell), \hat{a} \in \hat{A}$, are essentially pairwise exchangeable.

If labels are treated as parts of agents' types, it suffices to apply the exchangeability assumptions of this paper to the extended types $\tau^*(\cdot, a) = (\tau(\cdot, a), \ell(\cdot, a))$, rather than just $\tau(\cdot, a), a \in A$. This assumption would actually encompass the conditional-exchangeability assumption that, for $\alpha \circ \ell^{-1}$-almost all $\ell \in L$, the types $\tau(\cdot, \hat{a}, \ell), \hat{a} \in \hat{A}$, are essentially pairwise exchangeable.

To see this, notice that, with exchangeability of extended types, the pairs $(\tau(\cdot, a), \ell(\cdot, a))$ are essentially pairwise conditionally independent and identically distributed, with conditioninig on the σ-algebra generated by the sample cross-section distributions $D(\{(\tau(\cdot, a'), \ell(\cdot, a'))\}_{a' \in A})$. I claim that, moreover, conditionally on the σ-algebra generated by the cross-section distributions $D(\{(\tau(\cdot, a'), \ell(\cdot, a'))\}_{a' \in A})$ and the labels $\ell(\cdot, a'), a' \in A^*$, the (narrow) types $\tau(\cdot, a), a \in A$, must be essentially conditionally independent and identically distributed, where A^*, with $\alpha(A^*) = 1$, is the set of agents, such that, conditionally on $D(\{(\tau(\cdot, a''), \ell(\cdot, a''))\}_{a'' \in A})$, the pairs

$(\tau(\cdot,a'),\ell(\cdot,a'))$ are conditionally independent and identically distributed. Because, conditionally on $D(\{(\tau(\cdot,a''),\ell(\cdot,a''))\}_{a''\in A})$, the σ-algebras generated by $(\tau(\cdot,a),\ell(\cdot,a))$ and by $(\tau(\cdot,a'),\ell(\cdot,a'))$, $a'\in A^*\backslash\{a\}$, are independent, the conditional distribution for $\tau(\cdot,a)$ given $D(\{(\tau(\cdot,a''),\ell(\cdot,a''))\}_{a''\in A})$ and $\ell(\cdot,a'), a'\in A^*$, is the same as the conditional distribution for $\tau(\cdot,a)$ given $D(\{(\tau(\cdot,a''),\ell(\cdot,a''))\}_{a''\in A})$ and $\ell(\cdot,a)$. Thus, if

$$D(\{(\tau(\cdot,a''),\ell(\cdot,a''))\}_{a''\in A})=\delta \quad\text{and}\quad \ell(\cdot,a)=\ell,$$

the conditional distribution of $\tau(\cdot,a)$ is $\theta(\ell,\delta)$, where $\theta(\cdot,\delta)$ is a regular conditional distribution for t given ℓ when the pair (t,ℓ) has the joint distribution δ. For the relation between labels and types in the narrow sense, one thus gets the same structure as in the case where labels are treated as parts of names and exchangeability is conditioned on labels, but, in addition, one can accommodate macro uncertainty about labels.

Turning to strategic behaviour, I first note that any effects of agents' labels on their own behaviours are accommodated by a simple reinterpretation of the notation. In the formalism of this paper, the behaviour of agent a depends on payoff function u_a, the type t_a, and the belief $b_a(t_a)$, as well as the agent's expectations about the strategies chosen by the other agents. To accommodate the effects of a label $\ell(a)$ or $\ell(\omega,a)$ it suffices to reinterpret the triple u_a, t_a, and $b_a(t_a)$ in terms of an extended name $a=(\hat{a}(a),\ell(a))$ and/or an extended type $\tau^*(\omega,a)=(\tau(\omega,a),\ell(\omega,a))$.

Effects of agents' labels on other agents' behaviours presume that labels are observable and that they enter the other agents' payoff functions. For example, the payoff function of agent a in (B.1) might be replaced by

$$u_a(t_a,s_a,\{(s_{a'},\ell_{a'})\}_{a'\in A_{-a}}), \tag{4.18}$$

so that the label $\ell_{a'}$ of agent $a'\neq a$ affects the payoff of agent a directly as well as indirectly, through its effect on the action $s_{a'}=\sigma(\tau^*(\omega,a'),a')$ of agent a'. In this formulation, the condition of anonymity in payoffs might be reformulated so that (4.18) is replaced by

$$u^*_a(t_a,s_a,D(\{(s_{a'},\ell_{a'})\}_{a'\in A_{-a}})), \tag{4.19}$$

where $D(\{(s_{a'},\ell_{a'})\}_{a'\in A_{-a}})$ is now the cross-section distribution of the pairs $(s_{a'},\ell_{a'})$ of actions and labels of the other agents.

If we think about labels as parts of agents' extended types, the specifications (4.18) and (4.19) beg the question why the payoff of agent shouldn't also depend on $t_{a'}$, the part of the extended type of agent a' that is not part of the label $\ell_{a'}$. A straightforward answer would be that agent a cannot observe $t_{a'}$. From this perspective, the important distinction attached to the notion of a *label* is not so much between payoff-irrelevant and payoff-relevant aspects of names, as between observable and unobservable aspects of types.

A Appendix: Proofs

Before turning to the proofs as such, I recall a few basic facts. For a reference, see, e.g., Billingsley (1995), in particular, pp. 41 f.

- Since T is a complete separable metric space, there exists a countable family $\mathcal{P} = \{B^k\}_{k=1}^{\infty}$ of sets in $\mathcal{B}(T)$ that generates $\mathcal{B}(T)$.
- Without loss of generality, the family $\mathcal{P} =\{B^k\}_{k=1}^{\infty}$ may be taken to be a π-system, i.e. a family of sets that is closed under finite intersections.
- A family $\mathcal{Q}$ of subsets of T is said to be a λ-system if it satisfies (i) $T \in \mathcal{Q}$, (ii) if $B \in \mathcal{Q}$, then $T \backslash B \in \mathcal{Q}$, (iii) if $B_1, B_2, ...$ are pairwise disjoint sets in $\mathcal{Q}$, then, $\cup_{n=1}^{\infty} B_n \in \mathcal{Q}$.
- Dynkin's $\pi - \lambda$ Theorem: If $\mathcal{P}$ is a π-system and $\mathcal{Q}$ is a λ-system, then $\mathcal{P} \subset \mathcal{Q}$ implies $\sigma(\mathcal{P}) \subset \mathcal{Q}$.

Proof of Remark 2.1. For any $B \in \mathcal{B}(T)$, let $\chi_B : T \to [0,1]$ be the indicator function of the set B, i.e., let $\chi_B(t) = 1$ for $t \in B$ and $\chi_B(t) = 0$ for $t \notin B$. Since $f : \Omega \times I \to T$ is measurable with respect to the Fubini extension $\mathcal{F} \boxtimes \mathcal{I}$ of the product σ-algebra $\mathcal{F} \otimes \mathcal{I}$, the composition $\chi_B \circ f$ is also measurable with respect to $\mathcal{F} \boxtimes \mathcal{I}$.

Let $\mathcal{P} =\{B^k\}_{k=1}^{\infty}$ be a countable family of subsets of T that is closed under finite intersections and suppose that $\{B^k\}_{k=1}^{\infty}$ generates $\mathcal{B}(T)$. For any k, let Ω_k be the set of $\omega \in \Omega$ for which the section $\chi_B \circ f(\omega, \cdot)$ of the function is $\chi_B \circ f$ integrable on $(I, \mathcal{I}, \lambda)$. By the Fubini property, $P(\Omega_k) = 1$. Because the family $\{B^k\}_{k=1}^{\infty}$ is countable, $P(\cap_{k=1}^{\infty} \Omega_k) = 1$. Thus, for P-almost every $\omega \in \Omega$, for all k, the section $\chi_{B^k} \circ f(\omega, \cdot)$ of the function $\chi_{B^k} \circ f$ is integrable on $(I, \mathcal{I}, \lambda)$.

Let $\mathcal{Q}$ be the family of subsets of T such that, for any $B \in \mathcal{Q}$ and any $\omega \in \cap_{k=1}^{\infty} \Omega_k$, the section $\chi_B \circ f(\omega, \cdot)$ of the function is $\chi_B \circ f$ integrable on $(I, \mathcal{I}, \lambda)$. One easily verifies that $\mathcal{Q}$ is a λ-system. By the argument just given, every set in the π-system $\mathcal{P} =\{B^k\}_{k=1}^{\infty}$ is also contained in $\mathcal{Q}$.

Because $\mathcal{P}$ generates $\mathcal{B}(T)$, Dynkin's $\pi-\lambda$ Theorem implies that every set in $\mathcal{P}$ is also contained in Q. Thus, for P-almost every $\omega \in \Omega$, for all $B \in \mathcal{B}(T)$, the section $\chi_B \circ f(\omega, \cdot)$ of the function $\chi_{B^k} \circ f$ is integrable on $(I, \mathcal{I}, \lambda)$. Moreover, the Fubini property implies that the functions

$$\omega \to \int_I \chi_B \circ f(\omega, i)\, d\lambda(i) \tag{A.1}$$

from $(\Omega, \mathcal{F}, P)$ into $[0, 1]$ are measurable.

For any $\omega \in \cap_{k=1}^{\infty} \Omega_k$ and any $B \in \mathcal{B}(T)$, we have

$$\int_I \chi_B \circ f(\omega, i)\, d\lambda(i) = \lambda \circ f(\omega, \cdot)^{-1}(B) \tag{A.2}$$

By Lemma 1 in Hammond and Sun (2003), it follows that the measurability of the function (A.1) for any $B \in \mathcal{B}(T)$ implies the measurability of the function

$$\omega \to \lambda \circ f(\omega, \cdot)^{-1}$$

from $(\Omega, \mathcal{F})$ into $(\mathcal{M}(T), \mathcal{B}(\mathcal{M}(T))$. The remark follows immediately. ■

Remark 2.2 is a special case of Remark 2.1, with $(\Omega \times I, \mathcal{W}, Q) = (\Omega \times A, \mathcal{F} \boxtimes \mathcal{A}, P \boxtimes \alpha)$ and $f = \tau$. Remark 2.3 is also a special case of Remark 2.1, with $(\Omega \times I, \mathcal{W}, Q) = (\Omega \times A_{-a}, \mathcal{F} \boxtimes \mathcal{A}_{-a}, P \boxtimes \alpha_{-a})$ and f equal to the function $(\omega, a') \to \sigma(\tau(\omega, a'), a')$.

Proof of Proposition 2.4. By Proposition B.3 in Appendix B, it suffices to prove that statement (a) is equivalent to the following statement:

(a*) For α-almost all $a \in A$, the random variables $\tau(\cdot, a'), a' \in A_{-a}$, are essentially conditionally pairwise exchangeable given $\mathcal{C}(a)$, where $\mathcal{C}(a)$ is the sub-σ-algebra of $\mathcal{F}$ that is generated by $\tau(\cdot, a)$.

For this purpose recall that, for any $a \in A$ and any $t_a \in T$, $b(t_a, a)$ is a probability measure on $(R_a, \mathcal{R}_a)$, where R_a is the range of the function $\omega \to \boldsymbol{\tau}^a(\omega) := \{\tau(\omega, a')\}_{a' \in A_{-a}}$ and $\mathcal{R}_a$ is the coarsest σ-algebra under which the mapping $\omega \to \boldsymbol{\tau}^a(\omega)$ from $(\Omega, \mathcal{F})$ to R_a is measurable. For any $a' \in A_{-a}$, trivially, the mapping $\mathbf{t} \to \varphi_{a'}(\mathbf{t}) = t_{a'} = proj_{a'}(\mathbf{t})$, from $(R_a, \mathcal{R}_a)$ to $(T, \mathcal{B}(T))$, is measurable, and so is the mapping $\mathbf{t} \to (\varphi_{a'}(\mathbf{t}), \varphi_{a''}(\mathbf{t})) = (t_{a'}, t_{a''})$, for any $a', a'' \in A_{-a}$, from $(R_a, \mathcal{R}_a)$ to $(T, \mathcal{B}(T)) \times (T, \mathcal{B}(T))$. By the definition of the mapping $\omega \to b(\tau(\omega, a), a)$ as a regular conditional distribution for $\boldsymbol{\tau}^a(\cdot)$ given the sub-σ-algebra $\mathcal{C}(a) \subset \mathcal{F}$ that is generated by $\tau(\cdot, a)$, it follows that, for any a' and $a'' \in A_{-a}$, the mapping $\omega \to b_a(\tau(\omega, a)) \circ (\varphi_{a'}(\cdot), \varphi_{a''}(\cdot))^{-1}$ is a regular conditional distribution for

$(\tau(\cdot, a'), \tau(\cdot, a''))$ given $a', a'' \in A_{-a}$. The equivalence of statement (a) in the proposition and statement (a*) above follows immediately. ■

Proposition 2.5 follows directly from Proposition 3 in Qiao et al. (2016). Proposition 2.6 follows from the argument in the proof of Proposition B.3 in Appendix B in combination with Proposition 3 in Qiao et al. (2016). Alternatively, Proposition 2.6 follows from Proposition 2.7 and Remark **??**.

Proof of Proposition 2.7. Given the definition of $b_a(t_a)$, the first statement follows from Proposition 7 of Hammond and Sun (2008).

To prove the second statement, define a measure Q on $(\Omega, \mathcal{F})$ by setting $Q = b_a(t_a) \circ (\boldsymbol{\tau}^a)^{-1}$. If $b_a(t_a) \circ (\boldsymbol{\tau}^a)^{-1}$ is absolutely continuous with respect to P, then, by the Radon-Nikodym theorem, there exists a density function q on $(\Omega, \mathcal{F})$ such that, for any $F \in \mathcal{F}$, $Q(F) = \int_F q(\omega)\, dP(\omega)$. Consider the random variables $\tau(\cdot, a'), a' \in A_{-a}$, on the probability space $(\Omega, \mathcal{F}, Q)$. Because the density q of Q with respect to P is measurable and the random variables $\tau(\cdot, a'), a' \in A_{-a}$, have the Fubini property on $(\Omega \times A, \mathcal{F} \boxtimes \mathcal{A}, P \boxtimes \alpha)$, one easily verifies that they also have the Fubini property on $(\Omega \times A, \mathcal{F} \boxtimes \mathcal{A}, Q \boxtimes \alpha)$.

If $b_a(t_a)$ satisfies anonymity in beliefs, i.e., if, under this measure, the types $t_{a'}$ of agents $a' \in A_{-a}$ are essentially pairwise exchangeable, one also verifies easily that the random variables $\tau(\cdot, a'), a' \in A_{-a}$, on the probability space $(\Omega, \mathcal{F}, Q)$ are essentially pairwise exchangeable. By Proposition 3 of Qiao et al. (2016), it follows that these are essentially pairwise conditionally independent given the sub-σ-algebra $\mathcal{D}$ of $\mathcal{F}$ that is generated by the mapping

$$\omega \to D(\{\tau(\omega, a')\}_{a' \in A_{-a}}), \tag{A.3}$$

and, moreover, for α-almost every $a' \in A_{-a}$, the mapping (A.3) is a regular conditional distribution for $\tau(\cdot, a)$ given $\mathcal{D}$. The second statement of Proposition 2.7 follows upon translating this statement back into a statement about the random variables $t_{a'}, a' \in A_{-a}$, on the probability space $(R_a, \mathcal{R}_a, b_a(t_a))$. ■

Proof of Proposition 3.1. By the definition of a regular conditional distribution, one obtains that, for α-almost all $\hat{a} \in A_{-a}$, $D(\{t_{a'}\}_{a' \in A_{-a}}) \circ \sigma(\cdot, \hat{a})^{-1}$ is a regular conditional distribution of $\sigma(\cdot, \hat{a})$ given $\mathcal{D}$. By Theorem 1 of Qiao et al. (2016), it follows that

$$D(\{\sigma(t_{a'}, a')\}_{a' \in A_{-a}}) = \int_{\hat{a} \in A_{-a}} D(\{t_{a'}\}_{a' \in A_{-a}}) \circ \sigma(\cdot, \hat{a})^{-1} d\alpha(\hat{a})$$

for $b_a(t_a)$-almost all $\{t_{a'}\}_{a' \in A_{-a}} \in R_{\tau^a}$.[17] ■

Proof of Proposition 3.1, Statements (a) - (c). The proof proceeds along similar lines as the proof of Proposition 5.3 of Sun (2006). By Lemma A.5 in Sun (2006), there exists a measurable function f from $\mathcal{M}(T) \times [0,1]$ into T such that for any $\delta \in \mathcal{M}(T)$,

$$\ell \circ f(\delta, \cdot)^{-1} = \delta \tag{A.4}$$

where ℓ is the uniform distribution on $[0,1]$. Given this function f and the random variable $\tilde{\delta}$, define the mapping $\tau : \Omega \times A \to T$ such that, for any $\omega \in \Omega$ and $a' \in A$,

$$\tau(\omega, a') = f(\tilde{\delta}(\omega), h(\omega, a')), \tag{A.5}$$

where h is the function given by the richness of the Fubini extension $(\Omega \times A, \mathcal{F} \boxtimes \mathcal{A}, P \boxtimes \alpha)$.

I claim that τ is measurable with respect to $\mathcal{F} \boxtimes \mathcal{A}$. In fact, τ is the composition of the measurable function $f : \mathcal{M}(T) \times [0,1] \to T$ with the function $H : \Omega \times A \to \mathcal{M}(T) \times [0,1]$ that is given by setting

$$H(\omega, a') = (\tilde{\delta}(\omega), h(\omega, a'))$$

for any $\omega \in \Omega$ and $a' \in A$. Because the map $\omega \to \tilde{\delta}(\omega)$ is measurable with respect to $\mathcal{F}$ and the map $(\omega, a') \to h(\omega, a')$ is measurable with respect to $\mathcal{F} \boxtimes \mathcal{A}$, the map $(\omega, a') \to H(\omega, a')$ is measurable with respect to $\mathcal{F} \boxtimes \mathcal{A}$, and so is the map $(\omega, a') \to \tau(\omega, a') = f(H(\omega, a'))$.

Because T is a complete separable metric space, $\mathcal{M}(T)$ is also a complete separable metric space, and the σ-algebra $\mathcal{D}$ is countably generated. Because the random variables $h(\cdot, a')$, $a' \in A$, are essentially pairwise independent, Proposition 3 in Hammond and Sun (2006) implies that they are also essentially pairwise conditionally independent given $\mathcal{D}$. As in Remark 1 of Hammond and Sun (2008), it follows that the random pairs $(\tilde{\delta}(\cdot), h(\cdot, a'))$, $a' \in A$, are also essentially pairwise conditionally independent given $\mathcal{D}$, and so are the random variables $\tau(\cdot, a') = f(\tilde{\delta}(\cdot), h(\cdot, a'))$, $a' \in A$.

Moreover, because, for α-almost every $a' \in A$, the random variable $h(\cdot, a')$ has the uniform distribution ℓ, (A.4) and (A.5) imply that, for α-almost every $a' \in A$, conditional on the event $\tilde{\delta}(\cdot) = \delta$, the probability distribution of $\tau(\cdot, a')$ is almost surely equal to δ. For α-almost every $a' \in A$, therefore the function $\tilde{\delta}(\cdot)$ is a regular conditional distribution for $\tau(\cdot, a')$

[17] I thank a referee for suggesting this very elegant proof, which is much simpler than what I had before.

given the σ-algebra $\mathcal{D}$ that is generated by $\tilde{\delta}$. Statement (b) has thus been proved.

Statement (c) follows by Corollary 2 of Qiao et al. (2016) and the fact that, conditionally on $\mathcal{D}$, the random variables $\tau(\cdot, a'), a' \in A$, are essentially pairwise independent. Statement (a) follows by Proposition 2.5. ■

Proof of Proposition 3.1, Statements (d) and (e).

For any $\delta \in \mathcal{M}(T)$, define a mapping $\tau_\delta : \Omega \times A \to T$ such that, for any $\omega \in \Omega$ and $a' \in A$,

$$\tau_\delta(\omega, a') = f(\delta, h(\omega, a')), \tag{A.6}$$

where, as before, $f : \mathcal{M}(T) \times [0,1] \to T$ is the function given by Lemma A.5 in Sun (2006), satisfying

$$\ell \circ f(\delta, \cdot)^{-1} = \delta, \tag{A.7}$$

with ℓ equal to Lebesgue measure on $[0,1]$, and h is the function given by the richness of the Fubini extension $(\Omega \times A, \mathcal{F} \boxtimes \mathcal{A}, P \boxtimes \alpha)$. For any $a \in A$, set

$$\boldsymbol{\tau}_\delta^a := \{\tau_\delta(\omega, a')\}_{a' \in A_{-a}} \tag{A.8}$$

and

$$\hat{b}_a(\delta) := P \circ (\boldsymbol{\tau}_\delta^a)^{-1}. \tag{A.9}$$

Then, given the random variable $\tilde{\delta}$, for α-almost every $a \in A$, $\hat{b}_a(\tilde{\delta}(\cdot))$ is a regular conditional distribution for $\{\tau(\cdot, a')\}_{a' \in A_{-a}} = \{f(\tilde{\delta}(\cdot), h(\cdot, a'))\}_{a' \in A_{-a}}$ given $\tilde{\delta}$.

Because T and $\mathcal{M}(T)$ are complete separable metric spaces, there also exist functions $\beta_a : T \to \mathcal{M}(T), a \in A$, such that $\beta_a(\tau(\cdot, a))$ is a regular conditional distribution for $\tilde{\delta}$ given $\tau(\cdot, a)$. By Statement (a) of the proposition, the functions $\beta_a, a \in A$, are essentially identical, i.e., there exists a function $\beta^* : T \to \mathcal{M}(T)$ such that $\beta_a = \beta^*$ for α-almost all $a \in A$. Equation (3.2) follows by the law of iterated expectations. Statement (e) has thus been proved.

Statement (d) follows from Statements (e) and (a) and Proposition 2.4. ■

Proof of Proposition 3.2. The "only if" part of the proposition follows from the argument sketched in the text. To prove the "if" part of the proposition, let Ψ, Φ, Π be such that, for the given $\beta(\cdot)$, equations (3.3) and (3.4) hold for all $B_1 \in \mathcal{B}(T)$ and all $B_2 \in \mathcal{B}(\mathcal{M}(T))$.

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space, and let $(\Omega \times A, \mathcal{F} \boxtimes \mathcal{A}, P \boxtimes \alpha)$ be a rich Fubini extension of the product space $(\Omega \times A, \mathcal{F} \otimes \mathcal{A}, P \otimes \alpha)$.

$\tilde{\delta} : \Omega \to \mathcal{M}(T)$ be such that $\Phi = P \circ \tilde{\delta}^{-1}$, so that the distribution of $\tilde{\delta}$ is Φ. Let $\tau : \Omega \times A \to T$ be the mapping that is given by Proposition 3.1. Let β^* be the associated macro belief function and let Ψ^*, Φ^*, Π^* be the associated measures that are given by the "if" part of the proposition (for the macro belief function β^*).

By construction, $\Phi^* = \Phi$. By the "if" part of the proposition (for the macro belief function β^*)), it follows that

$$\Pi^*(B_1 \times T) = \Psi^*(B_1) \tag{A.10}$$

and

$$\Pi^*(B_1 \times B_2) = \int_{B_2} \delta(B_1)\, d\Phi(\delta) \tag{A.11}$$

for all $B_1 \in \mathcal{B}(T)$ and all $B_2 \in \mathcal{B}(\mathcal{M}(T))$. From (A.11) and the fact that Φ and Π satisfy (3.4), one infers that $\Pi^* = \Pi$. With $\Pi^* = \Pi$, (A.10) and the fact that Ψ and Π satisfy (3.4) for β, imply $\Psi^* = \Psi$. Thus, Ψ and Π satisfy (3.4) for both β and β^*. It follows that $\beta^*(\tau(\cdot, a)) = \beta(\tau(\cdot, a))$, P-almost surely and therefore, that β is a macro belief function for the family $\tau(\cdot, a), a \in A$, of random types. ■

Proof of Proposition 3.3. I will show that the condition of Proposition 3.2 is equivalent to Π, Ψ, Φ satisfying Statements (a), (b), and (c).

(a) Statement (a) asserts that, Φ-almost every measure $\delta \in \mathcal{M}(T)$ is absolutely continuous with respect to Ψ, i.e. that, for any $B_1 \in \mathcal{B}(T)$, $\Psi(B_1) = 0$ implies $\delta(B_1) = 0$. Because (3.3), with $B_2 = \mathcal{M}(T)$, yields $\Pi(B_1 \times \mathcal{M}(T)) = \Psi(B_1)$, (3.4) implies

$$\Psi(B_1) = \Pi(B_1 \times \mathcal{M}(T)) = \int_{\mathcal{M}(T)} \delta(B_1)\, d\Phi(\delta) \tag{A.12}$$

for all $B_1 \in \mathcal{B}(T)$. For any $B_1 \in \mathcal{B}(T)$, therefore, the assertion that $\Psi(B_1) = 0$ implies $\delta(B_1) = 0$ is true for Φ-almost all δ. It remains to be shown that the null set of distributions δ for which the implication is not true can be chosen independently of B_1.

For this purpose, I use Dynkin's $\pi{-}\lambda$ Theorem, as in the proof of Remark 2.1. Let $\mathcal{P} = \{B^k\}_{k=1}^{\infty}$ be a countable family of subsets of T that is closed under finite intersections and suppose that $\{B^k\}_{k=1}^{\infty}$ generates $\mathcal{B}(T)$. For any k, let Δ_k be the set of $\delta \in \mathcal{M}(T)$ for which $\Psi(B^k) = 0$ implies $\delta(B^k) = 0$. By (A.12), $\Phi(\Delta^k) = 1$. Because the family $\mathcal{P} = \{B^k\}_{k=1}^{\infty}$ is countble, it follows that $\Phi(\cap_{k=1}^{\infty} \Delta^k) = 1$.

Let $\mathcal{Q}$ be the family of subsets of T such that, for any $B \in \mathcal{Q}$ and any $\delta \in \cap_{k=1}^{\infty} \Delta_k$, $\Psi(B) = 0$ implies $\delta(B) = 0$. One easily verifies that $\mathcal{Q}$ is a λ-system. By the argument just given, every set in the π-system $\mathcal{P} = \{B^k\}_{k=1}^{\infty}$ is also contained in $\mathcal{Q}$. Because $\mathcal{P}$ generates $\mathcal{B}(T)$, Dynkin's $\pi - \lambda$ Theorem implies that every set in $\mathcal{P}$ is also contained in Q. Thus, for Φ-almost every $\delta \in \mathcal{M}(T)$, $\Psi(B) = 0$ implies $\delta(B_1) = 0$, for all $B_1 \in \mathcal{B}(T)$. For such δ, the existence of the density function g_{Ψ} satisfying follows by the Radon-Nikodym theorem.

(b) The proof of Statement (b) is similar. The statement asserts that, for Ψ-almost every $t \in T$, the measure $\beta(t)$ is absolutely continuous with respect to Φ, i.e. that, for every $B_2 \in \mathcal{B}(\mathcal{M}(T))$, $\Phi(B_2) = 0$ implies $\beta(B_2|t) = 0$. Because (3.4), with $B_1 = T$, yields $\Pi(T \times B_2) = \Phi$, (3.3) implies

$$\Phi(B_2) = \Pi(T \times B_2) = \int_T \beta(B_2|t)\, d\Psi(t) \tag{A.13}$$

for all $B_2 \in \mathcal{B}(\mathcal{M}(T))$. For any $B_2 \in \mathcal{B}(\mathcal{M}(T))$, therefore, the assertion that $\Phi(B_2) = 0$ implies $\beta(B_2|t) = 0$ is true for Ψ-almost all t.

Moreover, by the same argument as in the proof of Statement (a), using Dynkin's $\pi - \lambda$ Theorem, the null set of t for which the implication is not true can be chosen independently of B_2. For Ψ-almost all $t \in T$, therefore, $\Phi(B_2) = 0$ implies $\beta(B_2|t) = 0$. For any such t, the existence of the density function f_{Φ} follows by the Radon-Nikodym theorem.

(c) By Statements (a) and (b), (3.3) and (3.4) can be written as

$$\Pi(B_1 \times B_2) = \int_{B_2} \int_{B_1} g_{\Psi}(t|\delta)\, d\Psi(t) d\Phi(\delta) \tag{A.14}$$

and

$$\Pi(B_1 \times B_2) = \int_{B_1} \int_{B_2} f_{\Phi}(\delta|t)\, d\Phi(\delta) d\Psi(t). \tag{A.15}$$

Statement (c) follows by Fubini's theorem and the Radon-Nikodym theorem. ∎

Proof of Proposition 3.4. To prove the first claim of the proposition, let β be as stated and suppose that β admits a common prior. Let Ψ, Φ, Π be the measures given by Proposition 3.3 and let f_{Φ}, g_{Ψ}, π be the associated density functions, as specified in Proposition 3.3. I first claim that, for any $t_0 \in T$, Φ is absolutely continuous with respect to $\beta(t_0)$. To prove this claim, I note that, because the measures $\beta(t), t \in T$, are absolutely continuous with respect to $\beta(t_0)$, there exist density functions $f_0(\cdot|t), t \in T$, for $\beta(t)$ with respect to $\beta(t_0)$, such that, for any $t \in T$ and any $B_2 \in \mathcal{B}(\mathcal{M}(T))$,

$$\beta(B_2|t) = \int_{B_2} f_0(\delta|t)\; d\beta(\delta|t_0).$$

By (A.13), it follows that

$$\begin{aligned} \Phi(B_2) &= \int_T \int_{B_2} f_0(\delta|t)\; d\beta(\delta|t_0)\; d\Psi(t) \\ &= \int_{B_2} \int_T f_0(\delta|t)\; d\Psi(t)\; d\beta(\delta|t_0) \end{aligned}$$

for any $B_2 \in \mathcal{B}(\mathcal{M}(T))$. Absolute continuity of Φ with respect to $\beta(t_0)$ follows immediately. A density function φ_0 for Φ with respect to $\beta(t_0)$ is given by setting

$$\varphi_0(\delta) = \int_T f_0(\delta|t)\; d\Psi(t), \tag{A.16}$$

for Φ-almost any δ. Since $\beta(t_0)$ is also absolutely continuous with respect to Φ, the density $\varphi_0(\delta)$ is strictly positive for Φ-almost all δ and is in fact the inverse of the density $f_\Phi(\delta|t_0)$ of $\beta(t_0)$ with respect to Ψ.

I next show that, for $\Psi \otimes \Phi$-almost every pair (t, δ), the value $f_\Phi(\delta|t)$ of the density function $f_\Phi(\cdot|t)$ is strictly positive. For all $t \in T$, the density functions $f_\Phi(\cdot|t)$ and $f_\Phi(\cdot|t_0)$ are related by the equations

$$f_0(\delta|t) = f_\Phi(\delta|t) \cdot \varphi_0(\delta) \;\text{ and }\; 1 = f_\Phi(\delta|t_0) \cdot \varphi_0(\delta) \tag{A.17}$$

holding for Φ-almost all δ. By the mutual absolute continuity of $\beta(t)$ and $\beta(t_0)$, the value $f_0(\delta|t)$ of the density $f_0(\cdot|t)$ is strictly positive, for Φ-almost all δ. Hence $f_\Phi(\delta|t) > 0$ for all t, for Φ-almost all δ in the set

$$D_0 := \{\delta \in \mathcal{M}(T) | f_\Phi(\delta|t_0) > 0\}.$$

The definition of D_0 implies that $\int_{\mathcal{M}(T)\setminus D_0} f_\Phi(\delta|t_0) d\Phi(\delta) = 0$ and therefore $\beta(\mathcal{M}(T)\setminus D_0|t_0) = 0$. By the absolute continuity of Φ with respect to $\beta(t_0)$, it follows that $\Phi(\mathcal{M}(T)\setminus D_0) = 0$ and therefore $\Phi(D_0) = 1$. Thus, $f_\Phi(\delta|t_0) > 0$ for Φ-almost all δ and therefore $f_\Phi(\delta|t) > 0$ for $\Psi \otimes \Phi$-almost all (t, δ).

Given this result, Statement (c) in Proposition 3.3 implies that $g_\Psi(t|\delta) > 0$ for $\Psi \otimes \Phi$-almost all (t, δ). By elementary set theory, it follows that, for Φ-almost all δ, $g_\Psi(t|\delta) > 0$ for Ψ-almost all t. For

$$D := \{\delta \in D_0 | \Psi(\{t \in T | g_\Psi(t|\delta) > 0\}) = 1\},$$

one thus has $\Phi(D) = \Phi(D_0) = 1$. For any $\delta \in D$, Statement (a) in Proposition 3.3 implies that, for any $B_1 \in \mathcal{B}(T)$, $\delta(B_1) = 0$ implies $\Psi(B_1) = 0$, so Ψ is absolutely continuous with respect to δ. Thus Ψ and any one of the measures in D are mutually absolutely continuous. Hence the measures in D are also mutually absolutely continuous. This completes the proof that β satisfies Statement (i) in the proposition.

Turning to Statement (ii), I note that Statement (c) in Proposition 3.3 implies

$$\begin{aligned} & f_\Phi(\delta_1|t_2) \cdot g_\Psi(t_2|\delta_2) \cdot f_\Phi(\delta_2|t_1) \cdot g_\Psi(t_1|\delta_1) \\ = \; & g_\Psi(t_2|\delta_1) \cdot f_\Phi(\delta_2|t_2) \cdot g_\Psi(t_1|\delta_2) \cdot f_\Phi(\delta_1|t_1) > 0, \end{aligned} \tag{A.18}$$

for Ψ-almost all t_1, t_2 in T and Φ-almost all δ_1, δ_2 in $\mathcal{M}(T)$. Using (A.17) with t_0 replaced by t_1, I find that, for any $t \in T$, the formula

$$f_1(\delta|t) = f_\Phi(\delta|t) \cdot \varphi_1(\delta) \tag{A.19}$$

defines a density function for $\beta(t)$ with respect to $\beta(t_1)$. By the same argument, based on the mutual absolute continuity of the measures Ψ and $\delta \in D$, for any $\delta \in D$, the formula

$$g_1(t|\delta) = g_\Psi(t|\delta) \cdot \psi_1(t) \tag{A.20}$$

defines a density function for δ with respect to δ_1, where ψ_1 is the density of Ψ with respect to δ_1. Upon using (A.18) and (A.19) to substitute for the densities f_Φ and g_Ψ in (A.18), one finds that the terms $\varphi_1(\delta_1), \varphi_1(\delta_2), \psi_1(t_1), \psi_1(t_2)$ cancel out, and one is left with the equation

$$f_1(\delta_1|t_2) \cdot g_1(t_2|\delta_2) \cdot f_1(\delta_2|t_1) \cdot g_1(t_1|\delta_1) = g_1(t_2|\delta_1) \cdot f_1(\delta_2|t_2) \cdot g_1(t_1|\delta_2) \cdot f_1(\delta_1|t_1).$$

Equation (3.10) follows because the density functions $f_1(\cdot|t_1)$ and $g_1(\cdot|\delta_1)$ for $\beta(t_1)$ and δ_1 with respect to themselves have the constant value one, leaving the equation

$$f_1(\delta_1|t_2) \cdot g_1(t_2|\delta_2) = f_1(\delta_2|t_2) \cdot g_1(t_1|\delta_2),$$

which must hold for Ψ-almost all t_1, t_2 and Φ-almost all δ_1, δ_2, or, equivalently, in view of the mutual-absolute-continuity properties of the different families of measures, for $\beta(t_0)$-almost all δ_1, δ_2 in $\mathcal{M}(T)$ and δ_1-almost all t_1, t_2 in T. This completes the proof that β satisfies Statement (ii) in the proposition.

To prove the second claim in the proposition, let β be as stated in the proposition and suppose that Statements (i) and (ii) hold. By Statement (ii), there exist $t_1 \in T$ and $\delta_1 \in D$ such that (3.10) holds for $\beta(t_1)$-almost all $\delta_2 \in \mathcal{M}(T)$ and δ_1-almost all $t_2 \in T$. Thus, one can define

$$\begin{aligned} \pi_1(t,\delta) &:= \lambda_1 \cdot \frac{f_1(\delta|t)}{f_1(\delta_1|t)} \text{ if } f_1(\delta_1|t) > 0, \\ \pi_1(t,\delta) &:= 0 \text{ if } f_1(\delta_1|t) = 0, \end{aligned} \tag{A.21}$$

with

$$\lambda_1 := \left[\int_T \int_{\mathcal{M}(T)} \frac{f_1(\delta|t)}{f_1(\delta_1|t)} \, d\beta(\delta|t_1) d\delta_1(t) \right]^{-1},$$

and

$$\Pi(B_1 \times B_2) = \int_{B_1} \int_{B_2} \pi_1(t,\delta) \, d\beta(\delta|t_1) d\delta_1(t) \tag{A.22}$$

for any $B_1 \in \mathcal{B}(T)$ and $B_2 \in \mathcal{B}(\mathcal{M}(T))$. From (A.21) and (A.22), one computes

$$\begin{aligned} \Psi(B_1) &= \Pi(B_1 \times \mathcal{M}(T)) = \int_{B_1} \int_{\mathcal{M}(T)} \lambda_1 \cdot \frac{f_1(\delta|t)}{f_1(\delta_1|t)} \, d\beta(\delta|t_1) d\delta_1(t) \\ &= \int_{B_1} \int_{\mathcal{M}(T)} \lambda_1 \cdot \frac{1}{f_1(\delta_1|t)} \, d\beta(\delta|t) d\delta_1(t) = \int_{B_1} \lambda_1 \cdot \frac{1}{f_1(\delta_1|t)} \, d\delta_1(t) \end{aligned} \tag{A.23}$$

for any $B_1 \in \mathcal{B}(T)$. By (3.10), we also have

$$\begin{aligned} \pi_1(t,\delta) &= \lambda_1 \cdot \frac{g_1(t|\delta)}{g_1(t_1|\delta)} \text{ if } g_1(t_1|\delta) > 0, \\ \pi_1(t,\delta) &= 0 \text{ if } g_1(t_1|\delta) = 0, \end{aligned}$$

so (A.22) also yields

$$\begin{aligned} \Phi(B_2) &= \Pi(T \times B_2) = \int_T \int_{B_2} \lambda_1 \cdot \frac{g_1(t|\delta)}{g_1(t_1|\delta)} \, d\beta(\delta|t_1) d\delta_1(t) \\ &= \int_{B_2} \int_T \lambda_1 \cdot \frac{1}{g_1(t_1|\delta)} \, d\delta(t) d\beta(\delta|t_1) = \int_{B_2} \lambda_1 \cdot \frac{1}{g_1(t_1|\delta)} \, d\beta(\delta|t_1) \end{aligned} \tag{A.24}$$

From (A.23) and (A.24), one sees that Ψ and δ_1, as well as Φ and $\beta(t_1)$ are mutually absolutely continuous. Because δ_1 and any other $\delta \in D$ are mutually absolutely continuous, it follows that Ψ satisfies Statement (a) in Proposition 3.3. Because $\beta(t_1)$ and any other measure $\beta(t), t \in T$, are mutually absolutely continuous, it follows that Φ satisfies Statement (b) in Proposition 3.3. By inspection of (A.23) and (A.24), the densities of Ψ with respect to δ_1 and of Φ with respect to $\beta(t_1)$ are given as

$$\psi(t) = \lambda_1 \cdot \frac{1}{f_1(\delta_1|t)} \quad \text{and} \quad \varphi(\delta) = \lambda_1 \cdot \frac{1}{g_1(t_1|\delta)}. \tag{A.25}$$

For any t and δ the densities of $\beta(t)$ with respect to Φ and of δ with respect to Ψ can be computed from (A.25) and the densities of $\beta(t)$ with respect to $\beta(t_1)$ and of δ with respect to δ_1. This yields

$$f_\Phi(\delta|t) = \frac{f_1(\delta|t)}{\varphi(\delta)} = \frac{1}{\lambda_1} \cdot g_1(t_1|\delta) \cdot f_1(\delta|t)$$

and

$$g_\Psi(t|\delta) = \frac{g_1(t|\delta)}{\psi(t)} = \frac{1}{\lambda_1} \cdot f_1(\delta_1|t) \cdot g_1(t|\delta),$$

so (3.10) implies the validity of (3.9). By Proposition 3.3, it follows that β admits a common prior, with Π, Ψ, Φ as specified in the proposition.

To see that the common prior is unique, let Π^*, Ψ^*, Φ^* be any triple of distributions associated with a common prior for β. Using Proposition 3.3, let π^* be the density of Π^* with respect to $\Psi^* \otimes \Phi^*$. By the argument given above, Ψ^* and the measures $\delta \in D$ are mutually absolutely continuous, and so are Φ^* and the measures $\beta(t), t \in T$. Given a pair (t_0, δ_0), let ψ_0, φ_0 be the density functions for Ψ^* with respect to δ_0 and for Φ^* with respect to $\beta(t_0)$. Then Π^* has a density $\pi_0 = \pi^* \cdot \psi_0 \cdot \varphi_0$ with respect to $\delta_0 \otimes \beta(t_0)$. Using equation (3.14) in the text, one finds that

$$\begin{aligned}
\frac{\pi_0^*(t_2,\delta_2)}{\pi_0^*(t_1,\delta_1)} &= \frac{\pi^*(t_2,\delta_2) \cdot \psi_0(t_2) \cdot \varphi_0(\delta_2)}{\pi^*(t_1,\delta_1) \cdot \psi_0(t_1) \cdot \varphi_0(\delta_1)} \\
&= \frac{\pi^*(t_2,\delta_1) \cdot \psi_0(t_2)}{\pi^*(t_1,\delta_1) \cdot \psi_0(t_1)} \cdot \frac{\pi^*(t_2,\delta_2) \cdot \varphi_0(\delta_2)}{\pi^*(t_2,\delta_1) \cdot \varphi_0(\delta_1)} \\
&= \frac{g_\Psi(t_2|\delta_1) \cdot \psi_0(t_2)}{g_\Psi(t_1|\delta_1) \cdot \psi_0(t_1)} \cdot \frac{f_\Phi(\delta_2|t_2) \cdot \varphi_0(\delta_2)}{f_\Phi(\delta_1|t_2) \cdot \varphi_0(\delta_1)} \\
&= \frac{g_0(t_2|\delta_1)}{g_0(t_1|\delta_1)} \cdot \frac{f_0(\delta_2|t_2)}{f_0(\delta_1|t_2)}
\end{aligned}$$

for Ψ^*-almost all t_1, t_2 and Φ^*-almost all δ_1, δ_2. Up to modifications on sets of $\delta_0 \otimes \beta(t_0)$-measure zero, the ratio $\frac{\pi_0^*(t_2,\delta_2)}{\pi_0^*(t_1,\delta_1)}$ is thus uniquely determined by the density functions $f_0(\cdot|t), g_0(\cdot|\delta), t \in T, \delta \in D$. Because $\Pi^*(T \times M(T)) = 1$, it follows that, up to modifications on sets of $\delta_0 \otimes \beta(t_0)$-measure zero, the density π^* itself is uniquely determined by these density functions. Therefore Π^* is uniquely determined by these density functions. ■

B Appendix: Conditional Exchangeability

Let $(\Omega, \mathcal{F}, P)$ and $(I, \mathcal{I}, \lambda)$ be complete atomless probability spaces. Let $(\Omega \times I, \mathcal{F} \boxtimes \mathcal{I}, \mathcal{P} \boxtimes \lambda)$ be a Fubini extension of the product $(\Omega \times I, \mathcal{F} \otimes \mathcal{I}, \mathcal{P} \otimes \lambda)$ and let $f : \Omega \times I \to T$ be a process that is measurable with respect to $\mathcal{F} \boxtimes \mathcal{I}$ and that takes values in a complete separable metric space T. The random variables $f(\cdot, i)$, $i \in I$, are *essentially pairwise exchangeable* if there exists a probability measure π on T^2 such that, for λ-almost all $i_1 \in I$, one has

$$P(\{\omega \in \Omega | f(\omega, i_1) \in B_1\} \cap \{\omega \in \Omega | f(\omega, i_2) \in B_2\}) = \pi(B_1 \times B_2) = \pi(B_2 \times B_1)$$

for λ-almost all $i_2 \in I$ and all B_1, B_2 in $\mathcal{B}(T)$.

I now define a concept of *essential conditional pairwise exchangeability.* Whereas essential pairwise exchangeability is defined in terms of the prior P, essential conditional pairwise exchangeability will refer to posteriors, i.e. conditional distributions, that are induced by some countably generated sub-σ-algebra $\mathcal{C}$ of $\mathcal{F}$.

For this purpose I define a new function $\varphi : \Omega \times I \times I \to T \times T$ by setting

$$\varphi(\omega, i_1, i_2) = (f(\omega, i_1), f(\omega, i_2)).$$

The measurability of f with respect to the Fubini extension $\mathcal{F} \boxtimes \mathcal{I}$ implies that φ is measurable with respect to the Fubini extension $\mathcal{F} \boxtimes \mathcal{I} \boxtimes \mathcal{I}$ of the product $(\mathcal{F} \boxtimes \mathcal{I}) \otimes \mathcal{I}$.

For any countably generated sub-σ-algebra $\mathcal{C}$ of $\mathcal{F}$, let $\mu_{\mathcal{C}}$ be a regular conditional distribution for φ given $\mathcal{C} \boxtimes (\mathcal{I}_1 \times \mathcal{I}_2)$ and, for any $(\omega, i_1, i_2) \in \Omega \times I \times I$, let $\mu_{\mathcal{C}, i_1, i_2}(\cdot, \omega)$ be the measure on $T \times T$ that is given by the value of $\mu_{\mathcal{C}}$ at (ω, i_1, i_2). I say that the function f exhibits *conditional essentially pairwise exchangeability* given $\mathcal{C}$ if, for P-almost all $\omega \in \Omega$, for λ-almost all $i_1 \in I$, under the measure $\mu_{\mathcal{C}, i_1, i_2}(\cdot, \omega)$, the random variables $f(\cdot, i_1)$ and $f(\cdot, i_2)$ are exchangeable, for λ-almost all $i_2 \in I$.

In the following, I discuss the relation between essential pairwise exchangeability and essential conditional pairwise exchangeability. I first note that, if the random variables $f(\cdot, i), i \in I$, are essentially pairwise exchangeable, then, by Proposition 3 of Qiao et al. (2016), they are essentially conditionally pairwise independent and identically distributed, and a conditional exact law of large numbers holds.

By the Fubini property, for P-almost every $\omega \in \Omega$, the cross-section distribution of $f(\omega, .)$ is well defined. Denote this cross-section distribution as $\delta(\omega)$, and let $\mathcal{D} \subset \mathcal{F}$ be the σ-algebra that is generated by the mapping

$$\omega \to \delta(\omega),$$

from Ω to $\mathcal{M}(T)$. Essential conditional pairwise independence and the conditional exact law of large numbers implies that, for P-almost all $\omega \in \Omega$, for λ-almost all $i_1 \in I$,

$$\mu_{\mathcal{D},i_1,i_2}(\cdot, \omega) = (\delta(\omega) \times \delta(\omega)) \tag{B.1}$$

for λ-almost all $i_2 \in I$.

Proposition B.1 *Assume that the random variables $f(\cdot, i), i \in I$, are essentially pairwise exchangeable. Let $\mathcal{C}$ be any countably generated sub-σ-algebra of $\mathcal{F}$ and let $\mathcal{A}(\mathcal{D},\mathcal{C}) \subset \mathcal{F}$ be the smallest σ-algebra that contains $\mathcal{C}$ as well as $\mathcal{D}$. If*

$$\mu_{\mathcal{A}(\mathcal{D},\mathcal{C})} = \mu_{\mathcal{D}}, \tag{B.2}$$

then the random variables $f(\cdot, i), i \in I$, are essentially conditionally pairwise exchangeable given $\mathcal{C}$.

Proof. The argument involves two steps. First, by combining (B.1) and (B.2), one obtains

$$\mu_{\mathcal{C}}(\omega, i_1, i_2) = \int_{\mathcal{M}(T)} \delta \times \delta \; db_{\mathcal{C}}(\delta|\omega) \tag{B.3}$$

for $P \boxtimes \lambda^2$-almost all (ω, i_1, i_2), where $b_{\mathcal{C}}(\cdot|\cdot)$ is a regular conditional distribution for $\delta(\cdot)$ given $\mathcal{C}$ and the integration is to be interpreted in such a way that

$$\mu_{\mathcal{C}}(B_1 \times B_2|\omega, i_1, i_2) = \int_{\mathcal{M}(T)} \delta(B_1) \cdot \delta(B_2) db_{\mathcal{C}}(\delta|\omega)$$

for $P \boxtimes \lambda^2$-almost all (ω, i_1, i_2) and and all B_1, B_2 in $\mathcal{B}(T)$.

Second, by Lemma 1 in Qiao et al. (2016), with I replaced by $I \times I$, equation (B.3) implies that, for λ^2-almost all $(i_1, i_2) \in I^2$, the mapping $\omega \to \mu_{\mathcal{C},i_1,i_2}(\cdot|\omega) = \mu_{\mathcal{C}}(\omega, i_1, i_2)$ is a regular conditional distribution for $\varphi(\cdot, i_1, i_2)$ given $\mathcal{C}$.[18] By (B.3), it follows that

$$\mu_{\mathcal{C},i_1,i_2}(B_1 \times B_2, \omega) = \int_{\mathcal{M}(T)} \delta(B_1) \cdot \delta(B_2) \; db_{\mathcal{C}}(\delta|\omega) \tag{B.4}$$

[18]In applying the lemma of Qiao et al. (2016), one must adapt the notation so that the product $I \times I$, with typical element (i_1, i_2), takes the place of their space I, with typical element i.

for $P \boxtimes \lambda^2$-almost all ω, i_1, i_2 and all B_1, B_2 in $\mathcal{B}(T)$. The proposition follows immediately. ■

Proposition B.1 includes the case where the sub-σ-algebra $\mathcal{C}$ is actually generated by one of the random variables $f(\cdot, i)$, $i \in I$.

Proposition B.2 *Assume that the random variables $f(\cdot, i), i \in I$, are essentially pairwise exchangeable. For any $a \in I$, let $\mathcal{C}(a)$ be the sub-σ-algebra of $\mathcal{F}$ that is generated by $f(\cdot, a)$. Then, for λ-almost every $a \in I$, the random variables $f(\cdot, i), i \in I$, are essentially conditionally pairwise exchangeable given $\mathcal{C}(a)$.*

Proof. By Proposition B.1, it suffices to show that equation (B.2),

$$\mu_{\mathcal{A}(\mathcal{D},\mathcal{C}(a))} = \mu_{\mathcal{D}}, \tag{B.5}$$

holds for λ-almost all $a \in I$. As mentioned above, by Proposition 3 of Qiao et al. (2016), conditionally on the σ-algebra $\mathcal{D}$ that is generated by the cross-section distribution mapping $\omega \to \delta(\omega)$, the random variables $f(\cdot, i), i \in I$, are essentially pairwise independent and identically distributed with the common conditional probability distribution $\delta(\omega)$. By Proposition 3 of Hammond and Sun (2006), it follows that, for λ-almost all $a \in I$, conditionally on $\mathcal{A}(\mathcal{D}, \mathcal{C}(a))$, the random variables $f(\cdot, i), i \in I$, are also essentially pairwise independent and essentially conditionally identically distributed. Moreover, for λ-almost all $a \in I$, conditionally on $\mathcal{D}$, the random variables $f(\cdot, a)$ and $f(\cdot, i), i \in I$, are independent, for λ-almost all $i \in I$. Therefore, equation (B.5) follows from Corollary 4 of Hammond and Sun (2006). ■

A converse of Proposition B.2 is also true.

Proposition B.3 *The random variables $f(\cdot, i), i \in I$, are essentially pairwise exchangeable if and only if, for λ-almost every $a \in I$, the random variables $f(\cdot, i), i \in I$, are essentially conditionally pairwise exchangeable given $\mathcal{C}(a)$, where $\mathcal{C}(a)$ is the sub-σ-algebra of $\mathcal{F}$ that is generated by $f(\cdot, a)$.*

Proof. Given Proposition B.2, it suffices to prove the "if" part of the proposition. Let $a \in I$ be such that the random variables $f(\cdot, i), i \in I$, are essentially conditionally pairwise exchangeable given $\mathcal{C}(a)$. Thus, for P-almost all $\omega \in \Omega$, for λ-almost all $i_1 \in I$, under the measure $\mu_{\mathcal{C}(a),i_1,i_2}(\cdot|\omega)$, the random variables $f(\cdot, i_1)$ and $f(\cdot, i_2)$ are exchangeable, for λ-almost all $i_2 \in I$. By Proposition 3 of Qiao et al. (2016), it follows that, for P-almost all $\omega \in \Omega$, for λ-almost all $i_1 \in I$, under the measure $\mu_{\mathcal{C}(a),i_1,i_2}(\cdot,\omega)$, conditionally on the σ-algebra $\mathcal{D}$ that is generated by the cross-section distribution mapping $\delta(\cdot)$, the random variables $f(\cdot,i_1)$ and $f(\cdot,i_2)$ are independent and identically distributed, for λ-almost all $i_2 \in I$, and that the common conditional distribution given $\mathcal{D}$ is equal to the cross-section distribution δ. Thus, for any B_1, B_2 in $\mathcal{B}(T)$, we have

$$\begin{aligned}
P(\{\omega &\in \Omega | f(\omega,i_1) \in B_1\} \cap \{\omega \in \Omega | f(\omega,i_2) \in B_2\}) \\
&= \int_\Omega \chi_{B_1}(f(\omega,i_1))\chi_{B_2}(f(\omega,i_2))dP(\omega) \\
&= \int_\Omega \int \chi_{B_1}(f(\cdot,i_1))\chi_{B_2}(f(\cdot,i_2))d\mu_{\mathcal{C}(a),i_1,i_2}(\cdot,\omega)dP(\omega) \\
&= \int_\Omega \int_{T\times T} \chi_{B_1}(t_1)\chi_{B_2}(t_2)d\delta(t_1)d\delta(t_2)\, db_{\mathcal{C}(a)}(\delta|\omega),
\end{aligned}$$

where χ_{B_1}, χ_{B_2} are the characteristic functions of the sets B_1 and B_2 and, as before, $b_{\mathcal{C}(a)}(\cdot|\cdot)$ is a regular conditional distribution for $\delta(\cdot)$ given $\mathcal{C}(a)$. Upon setting

$$\pi(B_1 \times B_2) = \int_\Omega \int_{T\times T} \chi_{B_1}(t_1)\chi_{B_2}(t_2)d\delta(t_1)d\delta(t_2)\, db_{\mathcal{C}(a)}(\delta|\omega),$$

one easily verifies that $\pi(B_1 \times B_2) = \pi(B_2 \times B_1)$, so the condition for exchangeability of $f(\cdot,i_1)$ and $f(\cdot,i_2)$ under the prior P is satisfied. ■

References

[1] Alesina, A., and G. Tabellini (1990), A Positive Theory of Fiscal Deficits and Government Debt, *Review of Economic Studies* 87, 403-414.

[2] Angeletos, G.-M., and I. Werning (2006), Crises and Prices: Information Aggregation, Multiplicity, and Volatility, *American Economic Review* 96 (5), 1720-1736.

[3] Aumann, R. (1998), Common Priors: A Reply to Gul, *Econometrica* 66(4), 929-938.

[4] Bertrand, M. and S. Mullainathan 2004), Are Emily And Greg More Employable Than Lakisha And Jamal? A Field Experiment On Labor Market Discrimination, *American Economic Review* 94(4), 991-1013.

[5] Bierbrauer, F.J. (2009), A Note on Optimal Income Taxation, Public Goods Provision and Robust Mechanism Design, *Journal of Public Economics* 93, 667-670.

[6] Bierbrauer, F.J. (2014), Optimal Tax and Expenditure Policy with Aggregate Uncertainty, *American Economic Journal: Microeconomics* 6, 205-257.

[7] Bierbrauer, F.J., and M.F. Hellwig (2015), Public-Good Provision in Large Economies, Preprint 2015-12, Max Planck Institute for Research on Collective Goods, Bonn, http://www.coll.mpg.de/pdf_dat/2015_12online.pdf.

[8] Billingsley, P. (1995), *Probability and Measure*, 3rd Edition, Wiley, New York.

[9] de Finetti, B. (1931), Funzione caratteristica di un fenomeno allatorio, *Atti della R. Accademia Nazionale dei Lincii Ser. 6, Memorie, Classe di Scienze, Fisiche, Matematiche e Naturali, Vol. 4*, 251-299.

[10] de Finetti, B. (1970/74), *Teoria della Probabilità*, Einaudi, Turin 1970, translation published as: *Theory of Probability*, Wiley, London 1974.

[11] Doob, J.L. (1937), Stochastic Process Depending on a Continuous Parameter, *Transactions of the American Mathematical Society* 42, 107-140.

[12] Doob, J.L. (1953), *Stochastic Processes,* Wiley, New York.

[13] Feinberg, Y. (2000), Characterizing common priors in the form of posteriors, *Journal of Economic Theory* 91, 127-179.

[14] Feldman, M., and C. Gilles (1985), An expository note on individual risk without aggregate uncertainty, *Journal of Economic Theory* 35, 26-32.

[15] Goldstein, I., and A. Pauzner (2005), Demand Deposit Contracts and the Probability of a Bank Run, *Journal of Finance* 60 (3), 1293-1327.

[16] Gul, F. (1998), A Comment on Aumann's Bayesian View, *Econometrica* 66(4), 923-928.

[17] Hammond, P.J., and Y. Sun (2003), Monte Carlo Simulation of Macroeconomic Risk with a Continuum of Agents: the Symmetric Case, *Economic Theory* 21, 743-766

[18] Hammond, P.J., and Y. Sun (2006), The Essential Equivalence of Pairwise and Mutual Conditional Independence, *Probability Theory and Related Fields* 135, 415-427

[19] Hammond, P.J., and Y. Sun (2008), Monte Carlo Simulation of Macroeconomic Risk with a Continuum of Agents: the General Case, *Economic Theory* 36, 303-325.

[20] Harsanyi, J.C. (1967/68), Games with Incomplete Information Played by Bayesian Players I, II, III, *Management Science* 14, 159 - 182; 320 - 334; 486-502.

[21] Hellman, Z., and D. Samet (2012), How Common are Common Priors?, *Games and Economic Behavior* 74, 517-525.

[22] Hellwig, C. (2002), Public Information, Private Information, and the Multiplicity of Equilibria in Coordination Games, *Journal of Economic Theory* 107(2), 191-222.

[23] Hellwig, M.F. (2011), Incomplete-Information Models of Large Economies with Anonymity: Existence and Uniqueness of Common Priors, No. 2011/08, Max Planck Institute for Research on Collective Goods, Bonn. http://www.coll.mpg.de/pdf_dat/2011_08online.pdf.

[24] Hellwig, M.F. (2013), From Posteriors to Priors via Cycles: An Addendum, *Economics Letters* 118, 455-458.

[25] Hellwig, M.F. (2019), Incomplete-Information Games in Large Populations with Anonymity, Preprint No. 2019/14, Max Planck Institute for Research on Collective Goods, Bonn, Bonn. http://www.coll.mpg.de/pdf_dat/2019_14online.pdf.

[26] Judd, K.L. (1985), The law of large numbers with a continuum of IID random variables, *Journal of Economic Theory* 35, 19-25.

[27] Khan, M.A., and Y. Sun (1999), Non-cooperative games on hyperfinite Loeb spaces, *Journal of Mathematical Economics* 31, 455–492.

[28] Kyle, A.S. (1985), Continuous Auctions and Insider Trading, *Econometrica* 53(6), 1315-1336.

[29] Kyle, A.S. (1989), Informed Speculation with Imperfect Competition, *Review of Economic Studies* 56 (3), Pages 317-355.

[30] Lindbeck, A., and J. Weibull (1987), Balanced-Budget Redistribution as the Outcome of Political Competition, *Public Choice* 52(3), 273-297.

[31] Mertens, J.-F., and S. Zamir (1985), Formulation of Bayesian analysis for games with incomplete information, *International Journal of Game Theory* 14, 1-29.

[32] Milgrom, P.R., and R.J.Weber (1985), Distributional Strategies for Games with Incomplete Information, *Mathematics of Operations Research* 10(4), 619-632.

[33] Mirrlees, J.A. (1971), An Exploration in the Theory of Optimum Income Taxation, *Review of Economic Studies* 38(2), 175-208.

[34] Morris, S., and H.S. Shin (1998), Unique Equilibrium in a Model of Self-Fulfilling Currency Attacks, *American Economic Review* 88(3), 587-597.

[35] Podczeck, K. (2010), On Existence of Rich Fubini Extensions, *Economic Theory* 45, 1-22.

[36] Qiao, L., Y. Sun, and Z. Zhang (2016), Conditional Exact Law of Large Numbers and Asymmetric Information Economies with Aggregate Uncertainty, *Economic Theory* 62(1-2), 43-64.

[37] Rochet, J.-C., and X. Vives (2004), Coordination Failures and the Lender of the Last Resort, *Journal of the European Economic Association* 2(6), 1116 - 1147.

[38] Rodrigues-Neto, J.A. (2009), From Posteriors to Priors via Cycles, *Journal of Economic Theory* 144, 876-883.

[39] Samet, D. (1998a), Iterative Expectations and Common Priors, *Games and Economic Behavior* 24, 131-141.

[40] Samet, D. (1998b), Common Priors and the Separation of Convex Sets, *Games and Economic Behavior* 24, 172-174.

[41] Sun, Y. (2006), The Exact Law of Large Numbers via Fubini Extension and Characterization of Insurable Risks, *Journal of Economic Theory* 126, 31-69.

[42] Sun, Y., and Y.Zhang (2009), Individual Risk and Lebesgue Extension without Aggregate Uncertainty, *Journal of Economic Theory* 144, 432-442.