On Defining Ex Ante Payoffs in Games with Diffuse Prior*

Attila Ambrus† and Aaron Kolb‡

May 2, 2017

Abstract

We formalize the notion of a diffuse prior and show that, for a general class of games, strategies which admit well-defined expected payoffs under the diffuse prior satisfy a stationarity property. We define the diffuse prior through a limit construction, using sequences of well-defined priors that become increasingly dispersed and uniform. Admissible strategies are those for which ex ante payoffs along these sequences converge to a limit that does not depend on the particular sequence of priors. We show that a strategy is admissible if and only if it is almost stationary in a precise sense. A secondary contribution of the paper is a generalization of the concept of distributional strategies (Milgrom and Weber, 1985) to multistage games.

1 Introduction

The diffuse, or uninformative, prior is often interpreted informally as a uniform distribution on the real line. This prior has two advantages for use in economics: first, as it represents complete ignorance, it is appropriate for modeling situations in which agents have no advance knowledge of the environment; and second, it makes updating beliefs through Bayes' rule computationally simpler. This tractability comes from ex ante symmetry across all states, and not having to worry about corners of the state space. However, although the diffuse prior is commonly used, it is not formally defined: any uniform distribution must have constant density and must integrate to one, but any positive constant density, integrated over the real line, yields infinity, and zero density integrates to zero. This lack of a formal representation means that ex ante expected payoffs are not defined when driven by a random variable drawn from a diffuse prior distribution. The existing literature has circumvented this issue by leaving expected payoffs undefined and instead focusing on payoffs conditional on signal realizations (for example, Friedman (1991), Klemperer (1999), Morris and Shin (2002, 2003), Myatt and Wallace (2014)).

In this paper, we develop a method for formally defining expected payoffs under a diffuse prior, and thereby bringing them into the realm of traditional game theory, where expected payoffs are assumed to be well-defined for all strategy profiles of a game. We claim that the diffuse prior can be rigorously constructed as a limit of well-defined distributions, and that expected payoffs under a diffuse prior can be defined in certain cases. We define a class of games with a stationary information

*The authors thank Volodymyr Baranovskyi, Joel Sobel and Tom Wiseman for helpful comments and encouragement. All errors are our own.

†Duke University, *attila.ambrus@duke.edu*

‡Indiana University Kelley School of Business, *kolba@indiana.edu*

and payoff structures, where signals and payoffs are translation invariant in a precise sense. We then show that the strategies which admit expected payoffs in this manner all have a distinguishing characteristic: they are nearly stationary, in that there is an associated stationary strategy and the difference between these must vanish when evaluated over larger subsets of the real line.

We capture the main features of this uninformative prior by using a sequence of (proper) measures that *diffuse* in a formal sense (Definition 5). We say that a class of strategies is *admissible* (Definition 6) when ex ante expected payoffs, taken along any diffusing sequence of proper priors, have a well-defined limit that does not depend on the particular sequence.[1] Stationary strategies are unsurprisingly admissible, as given any stationary strategy profile, expected payoffs conditional on al signal realizations are the same. Our main result (Theorem 1) states roughly that in any class of admissible strategies that includes stationary strategies, every strategy is *nearly stationary* in a particular sense (Definition 4). Furthermore, every such strategy is payoff-equivalent to some stationary strategy.

A summary of our result is as follows. In stationary environments, where the diffuse prior is most appropriate, we show that admissible strategies are nearly stationary, and we argue that it is without loss of generality to restrict attention to stationary strategies.

The paper is structured as follows. Section 2 gives an overview of the main ideas and outlines the steps of the main proof, and Section 3 gives the formal analysis. Section 4 shows how our results apply to a companion paper on delegation. The appendix contains proofs not provided in the body of the paper.

2 Overview

To capture the diffuse prior as a limit object, we define sequences of proper measures to be diffusing if, roughly, the measures become increasingly uniform and spread out over the real line. Our definition allows for a large class of diffusing sequences, including sequences of uniform distributions on $[-n, n]$ or sequences of normal distributions with variance n.

We will define payoffs under a diffuse prior in cases where the limit of payoffs taken along any diffusing sequence exists and is independent of the sequence. Before defining the class of games we consider, we demonstrate how some concrete functions from $\mathbb{R}$ to $\mathbb{R}$ stand up to this criterion. Clearly, a constant function, when integrated with respect to any probability measure, integrates to that constant, and so all diffusing sequences result in the same limit, and thus a constant function is admissible. In addition, the function $x \mapsto \mathbb{1}_{[0,1]}(x)$, which takes the value 1 if $x \in [0, 1]$ and 0 otherwise, is also admissible; it is not difficult to show that along any diffusing sequence, the expected value approaches 0. On the other hand, a function like $x \mapsto \mathbb{1}_{[0,\infty)}$ is *not* admissible. One could obtain a limit of 0 by defining a diffusing sequence $(\mathbb{P}^1_n)_{n\in\mathbb{N}}$ with the densities $\frac{\mathbb{1}_{[-n^2,n]}}{n^2+n}$ and a limit of 1 by a different diffusing sequence $(\mathbb{P}^2_n)_{n\in\mathbb{N}}$ with densities $\frac{\mathbb{1}_{[-n,n^2]}}{n^2+n}$. The key property of

[1] In a paper largely unrelated to our work, Dale and Morgan (2015) consider specific sequences of proper priors diffusing in a similar sense as in our definition, in the context of a specific game from Morris and Shin (2002). They do not investigate the possibility of defining ex ante expected payoffs in the game, instead they are interested in comparing equilibrium predictions of the model with proper versus improper priors. Equilibrium (and in general, strategic) analysis is not part of the current paper.

admissible functions here is that they are constant or "nearly" constant in some formal sense. As we are interested in games of asymmetric information, and not real-valued functions per se, the exercise is more subtle than the above examples suggest. Nonetheless, the above intuition plays a key role in the analysis that follows.

We analyze multi-stage games with asymmetric information but observable actions, with n players. At the beginning of the game, the state of the world $\theta \in \mathbb{R}$ is drawn according to a diffuse prior, and players receive private, conditionally i.i.d. signals s_i about θ. Time is discrete and of finite or infinite horizon. Stages are indexed by $t = 1, 2, \ldots, \tau$ where $\tau \in \mathbb{N} \cup \{+\infty\}$. Each period, players simultaneously choose real-valued actions and observe realized actions before moving to the next period. We assume that the game has a stationary structure in terms of signals and payoffs: (i) for each player there is some distribution F_i with full support such that for all θ, $s_i - \theta$ is drawn from this distribution and (ii) payoffs are invariant to a translation of all actions (across all stages) and θ by a constant.

Given the stationary structure of the game, it is useful to define strategies as joint distributions over signals and actions at different histories for a single fixed state $\theta = 0$ (as we will see, under appropriate conditions this determines strategies for all other states, too). Since we start from strategy sets that are not restricted to be stationary, we need to provide a more careful formal definition of strategies. Since the state space is uncountable, it is not practical to define strategies as products of signal-dependent distributions over actions. The key tension is that desirable topologies should be both rich enough so that payoff functions are continuous in strategies, but also coarse enough so that the strategy space is compact. We follow the approach of Milgrom and Weber (1985) in using *distributional strategies*, which are measures μ over the product space of signals and actions with bounded support with respect to the latter. A distributional strategy induces a conditional distribution on actions for any given signal, and by integrating over signals, it induces a conditional distribution on actions for any given θ.

After defining strategies "anchored" at $\theta = 0$, it is useful to condition these distributions on any given θ, and thus for every strategy μ and state θ, we define "recentered" strategies μ^θ, which, roughly speaking, interpret the history of the game (which includes a player's own signal and all observed past actions of all players) as if it had been shifted by up by the constant θ. A stationary strategy, as we define it, is one which is translation invariant, in the sense that if the history is shifted by a constant, then actions are shifted by the same constant; it follows that for stationary strategies, all of its recentered strategies are that same strategy. We assume that the space of all distributional strategies, denoted $\mathcal{M}$, is compact,[2] and give an easy to check sufficient condition for this to hold. We say that a strategy μ has a *limit strategy* $\mu^* \in \mathcal{M}$ if $\lim_{\theta \to \pm\infty} \mu^\theta = \mu^*$ (Definition 4). We show that whenever this property holds, the strategy μ^* must be stationary, and hence we call μ *nearly stationary*.

Our main result, Theorem 1, has two components. First, it says that the class of nearly stationary strategies is admissible, so near stationarity is sufficient for admissibility. Second, it says that in any admissible class of strategies which is at least as large as the set of stationary strategies, all

[2] We use the topology of weak convergence for measures and distributions, which is metrized by the Prokhorov distance, d_P, and we use the usual topology for $\mathbb{R}$; continuity and compactness are with respect to these topologies.

strategies are nearly stationary. Since these strategies are payoff-equivalent to stationary strategies, we argue that for a game with diffuse prior to have well-defined ex ante expected payoffs, essentially all strategies have to be stationary.

We begin the proof of Theorem 1 by establishing the existence of some distribution $\mu_i^* \in \mathcal{M}$ with the following property: for all $\eta > 0$, μ_i^θ is within η of μ_i^* for an infinite measure set of θ (Lemma 7). We then call μ_i^* an *attraction*. It is direct from the definition that this is a weaker condition than near stationarity; a necessary condition for μ_i to be near μ_i^* is that μ_i^* is an attraction.

Next, we argue that there can be at most one such attraction for any strategy. The proof of this claim is by contradiction and requires several steps. In particular, we show that there must exist some profile of stationary strategies of the rivals, μ_{-i}, against which these distributions yield distinct expected payoffs. We suppose that μ^* and $\hat{\mu}$ are two distinct attractions for player i's strategy μ_i. Given $\eta > 0$, we can construct a sequence of measures $(\mathbb{P}_n^1)_{n\in\mathbb{N}}$ (resp. $(\mathbb{P}_n^2)_{n\in\mathbb{N}}$) that places increasing mass on θ such that μ^θ is within η of μ^* (resp. $\hat{\mu}$). By continuity and translation, the limit of expected payoffs must be at most η from each of two distinct values. This violates admissibility, giving the desired contradiction.

Given the unique attraction μ^*, we show that μ^* is a limit strategy for μ. If μ has any limit strategy, that strategy must be an attraction, so μ^* is the only candidate. We show that if μ^* is not a limit strategy for μ, then there is a compact set of strategies that does not contain μ^* but contains μ^θ for infinitely many θ. This compact set itself contains an attraction, and this contradicts the uniqueness of the attraction μ^*.

To complete the proof, we argue that μ^* is stationary, so that μ_i is nearly stationary.

3 Model

Before analyzing the diffuse prior, we specify the class of games we consider, which consists of an information structure and a payoff structure.[3]

3.1 Setup

Let Γ denote the game. There are n players indexed by $i \in N := \{1, 2, \ldots, n\}$. The game has stages indexed by $t = 1, 2, \ldots$; the number of stages is $\tau \in \mathbb{N} \cup \{+\infty\}$. All players assign a diffuse prior to the state of the world (formalized below), $\theta \in \Theta := \mathbb{R}$. Players receive signals $s_i \in S_i := \mathbb{R}$ which are conditionally i.i.d. given θ with distributions $s_i - \theta \sim F_i$ for some full-support function F_i on the reals admitting a density; let $S := \times_{i\in\mathbb{N}} S_i$. We use ϕ_i to denote the measure induced by F_i, and $\phi = \phi_1 \times \phi_2 \times \cdots \times \phi_n$. In each stage t, players choose actions $a_i^t \in \mathbb{R}$. Players observe other players' past actions perfectly. The public history is a record of all players' past actions (but not signals) through period $t - 1$ and it is denoted h^t. A player's private history is the public history together with that player's initial signal s_i, and is denoted h_i^t or (s_i, h^t). We set $h_i^1 = s_i$. The set of actions available to player i in stage t after private history h_i^t is denoted $A_i^t(h_i^t)$.

[3] Since we are concerned with admissibility, equilibrium plays no role in our analysis, and we do not specify an equilibrium concept.

A pure strategy specifies for each private history h_i^t an action $a_i^t(h_i^t) \in A_i^t(h_i^t)$. Let $a_i := (a_i^t)_{t=1}^{\tau}$ denote the full sequence of realized actions by each player, and let A_i denote the set of all possible such sequences. Let $A := \times_i A_i$. Given a realization of θ and sequence of action profies taken in each stage, players receive payoffs $u_i(a_i, a_{-i}, \theta)$, where a_{-i} denotes the vector $(a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_n)$.[4] We assume the u_i are continuous in all arguments.

Informally, a mixed strategy is a randomization over pure strategies. However, in games of incomplete information with uncountable type spaces (such as this one), certain topological problems arise from interpreting strategies as mixed strategies (i.e., distributions over pure strategies) or behavioral strategies (i.e., products of history-contingent distributions over actions). As an alternative, Milgrom and Weber (1985) introduce the notion of *distributional strategies*, which we adapt in Definition 1.

Next, define a sequence $X := (X_k)_{k=1}^{n\tau}$ where each $X_{n(t-1)+j}$ is a copy of $\mathbb{R}$ for player $j \in N$ in stage $t \geq 1$. The set of public histories is contained in every X, and the set of private histories for a given player i is contained in $S_i \times X$.

Since $S \times X$ is complete and separable, the space of probability measures $\Delta(S \times X)$, under the topology of weak convergence, is metrized by the Prokhorov distance.[5]

We are now ready to define a (distributional) strategy. These strategies apply to the reduced strategic form;[6] we provide a definition for full strategic form strategies in the supplementary appendix.

Definition 1 (Strategies). *A strategy for player $i \in N$ is a probability measure μ_i on $S \times X$ such that:*

- *For all measurable $T \subset S$, $\mu_i(T \times X) = \phi(T)$.*
- *For all $j \in N$ and for all $t \geq 1$, given any private history h_j^t for player j, the support of $\mu_{i,X_{n(t-1)+j}}(\cdot|h_j^t)$ is a subset of $A_j^t(h_j^t)$, where $\mu_{i,X_{n(t-1)+j}}(\cdot|h_j^t)$ denotes the marginal distribution of μ_i in the $X_{n(t-1)+j}$ dimension (that is, the dimension of X corresponding to j's action in stage t) conditional on player j's private history.*
- *For all $j \neq i$ and for all $t \geq 1$, conditional on any private history h_i^t, the random variables induced by the marginal distributions in the $X_{n(t-1)+j}$ dimensions are mutually independent and uniformly distributed over $A_j^t(h_i^t)$.*

A few comments on Definition 1 are in order. The idea of a distributional strategy originates in Milgrom and Weber (1985) (hereafter "MW") to deal with topological problems that arise when private signals (or types) are drawn from a continuum. Although the standard way to define mixed strategies in games with finite or countable signal spaces is as products, over signals, of signal-dependent distributions, this approach runs into problems when signal spaces are uncountable. As an uncountable product, the strategy space is "too large" and there is no topology on this space

[4] Payoffs here are for the entire game. This allows, for example, payoffs calculated as a sum of discounted stage-level payoffs, but does not restrict payoffs to be separable across stages.

[5] See Billingsley (2009, Theorem 6.8).

[6] The reduced strategic form collapses pure strategies that yield equivalent payoffs for all strategy profiles of the other players into one pure strategy; see Fudenberg and Tirole (1991)

which simultaneously ensures that (i) the strategy space is compact and (ii) payoff functions are continuous in strategies. The solution to this problem posed by MW is to define strategies as joint distributions over signals and actions.

For the class of games we consider, it is necessary to extend the MW definition of distributional strategies. MW consider a one-shot simultaneous-move game, so the only history that players need to know is their own private signals. Our technical problem is more severe, since we have a multi-stage game, so the history includes past actions, and actions are also drawn from a continuum. Hence, the force of the distributional strategy is needed at every stage. For this reason, player i's distributional strategy μ_i includes dimensions for all the other players' actions. This results in some arbitrary detail which serves a technical purpose only. The strategically meaningful information from μ_i is in the distributions on the $X_{n(t-1)+i}$ dimensions conditional on each private history. To incorporate all possible private histories, the strategy must assign probabilities to a player's own signal as well as all players' actions. The marginal distributions for other players' actions conditional on the history are *not* strategically meaningful, so we normalize those to the uniform distribution for concreteness and to ensure full support. This ensures that player i's behavior is always well-defined as a conditional distribution. We wish to emphasize that μ_i does not contain any conjecture about other players' behavior, it just allows us to specify continuation behavior for all histories of the other players, on path or not.[7] Finally, we include all other players' signals in this distribution for the sake of symmetry – this way, all players' strategies have the same dimensions. The assumption that the marginal over each signal is ϕ is consistent with MW but is not essential, as long as these marginals have full support. We specify ϕ as the marginal for concreteness.

Under the above definition, a pure strategy is one such that for all s_i, $\mu(\cdot|s_i)$ places all mass on a single point (which can depend on s_i), where $\mu(\cdot|s_i)$ here denotes the regular conditional probability. That is, for each s_i, there exists $x \in X_i$ such that for all measurable $Y \subset X_i$, $\mu_i(Y|s_i) = 1$ if $x \in Y$ and 0 otherwise. We let $\mathcal{M}_i \subset \Delta(S_i \times X)$ denote the space of all strategies. Going forward, we occasionally drop subscripts for convenience.

Compactness and metrizability of the strategy space is useful throughout our analysis. So far we have placed no restrictions on the action spaces $A_i^t(h_i^t)$, and it is cleanest to impose these properties as an assumption. However, this assumption is commonly satisfied. Lemma 1 gives an easy-to-check sufficient condition for Assumption 1.

Assumption 1. *The strategy spaces $\mathcal{M}_i$ are compact metric spaces.*

We prove Lemma 1 in the appendix. An alternative sufficient condition, one which neither implies nor is implied by the one below, is that there exist compact sets Y_i such that for all s_i, t and h^t, $A_i^t(s_i, h^t) = Y_i$. However, our later assumption (Assumption 2) rules that alternative condition out.

Lemma 1. *The following condition is sufficient for Assumption 1: there exist M_i such that for all s_i, t and h^t, $A_i^t(s_i, h^t)$ is a compact subset of $[s_i - M_i, s_i + M_i]$.*

[7] However, Definition 1 does not allow us to specify continuation behavior after an off-path history due to a player's own deviation; for that, we need the non-reduced strategic form, which we discuss in the supplementary appendix.

3.2 Stationarity

As the diffuse prior implies symmetry accross states, we focus on games where this symmetry holds for all components of the game. We label these games stationary. Below we define stationarity of signals, payoffs and strategies.

We assume that the signal structure of the game is stationary so that conditional on any realization θ, the values s_i are drawn i.i.d. with $s_i - \theta \sim F_i$ for some distribution F_i admitting a density f_i and having full support on the real line. Given a constant $x \in \mathbb{R}$, let $\mathbf{x}^t$ denote the $t-$dimensional vector with x, so that $a_i + \mathbf{x}^\tau$ represents a sequence of actions shifted up by x at every stage 1 through τ. With slight abuse of notation, $a_{-i} + \mathbf{x}^\tau$ denotes the collection of $a_j + \mathbf{x}^\tau$ for each $j \neq i$.

Definition 2 (Stationary Payoffs)**.** *Payoffs are stationary if for all $\theta \in \Theta$, $x \in \mathbb{R}$, $a \in A$ and $i \in N$, $u_i(a_i + \mathbf{x}^\tau, a_{-i} + \mathbf{x}^\tau, \theta + x) = u_i(a_i, a_{-i}, \theta)$.*

Definition 2 says that payoffs are invariant to a translation of all strategies together with the true state by a constant.

At face value, a strategy as defined above gives a distribution over signals and markups conditional on $\theta = 0$, since the marginal distribution on each signal is F_i. In order to characterize admissible strategies in a more transparent way, it is useful to define a "recentering function" which gives joint distributions over actions given an arbitrary θ. For any strategy $\mu_i \in \mathcal{M}$ and $\theta \in \mathbb{R}$ we define the measure μ_i^θ by $\mu_i^\theta(Y) = \mu_i(Y - \theta)$ where $Y \subseteq S \times X$ and $Y - \theta$ denotes the set Y shifted in all dimensions by $-\theta$. Note that in the case of $\theta = 0$, we have $\mu^0 = \mu$. Assumption 2 below ensures that for all $\mu \in \mathcal{M}$ and $\theta \in \mathbb{R}$, μ^θ is itself a valid strategy. The interpretation of this strategy is that μ_i^θ behaves, relative to θ, after history $h_i^t + \theta$ exactly as μ_i behaves after history h_i^t.

Assumption 2. *The set of available strategies $\mathcal{M}$ is stable under translation: for all $\mu \in \mathcal{M}$ and $\theta \in \mathcal{M}$, $\mu^\theta \in \mathcal{M}$.*

The next lemma will be used later in our analysis (see Lemma 3).

Lemma 2. *For all strategies $\mu_i \in \mathcal{M}$, the map $\theta \mapsto \mu_i^\theta$ from $\mathbb{R}$ to $\mathcal{M}$ is uniformly continuous.*

Proof. We first specify a few preliminaries. We metrize the space $S \times X$ using the metric $d_{S\times X}$ defined as follows. Writing elements of $S \times X$ as (possibly infinite) sequences $z = (z_k)_{k=1}^{n+n\tau}$, define $d_{S\times X}(z, z') = \sum_{k=1}^{n+n\tau} \frac{|z_k - z'_k|}{2^k}$. This definition ensures that $d_{S\times X}(z, z') \leq \max_{k:1\leq k\leq n+n\tau} |z_k - z'_k|$. For a subset $Y \subseteq S \times X$, let $Y^\eta := \bigcup_{z\in Y} N_\eta(z)$, the union of all η-neighborhoods centered at points in Y. It follows from these definitions that for any $\eta \in \mathbb{R}$ (possibly negative), $Y + \eta \subseteq Y^{|\eta|}$, where as above $Y + \eta$ denotes the translation of the set Y by η in all dimensions. The image space $\mathcal{M}$ is metrized by the Prokhorov metric, $d_P(\mu, \hat{\mu}) := \inf\{\eta > 0 : \mu(Y) \leq \hat{\mu}(Y^\eta) + \eta \text{ and } \hat{\mu}(Y) \leq \mu(Y^\eta) + \eta \text{ for all } Y \subseteq S \times X\}$.

For the proof, we show that for all $\eta > 0$, if $|\theta - \theta'| < \eta$, then $d_P(\mu^\theta, \mu^{\theta'}) < \eta$. We have

$$\mu^\theta(Y) = \mu^{\theta'}(Y + (\theta' - \theta)) \leq \mu^{\theta'}(Y^{|\theta'-\theta|}) < \mu^{\theta'}(Y^\eta) + \eta,$$

and by a symmetric argument $\mu^{\theta'}(Y) < \mu^\theta(Y^\eta) + \eta$. By the definition of d_P, $d_P(\mu^\theta, \mu^{\theta'}) < \eta$. □

Next we define stationarity of strategies and a weakening of stationarity that we will later prove to be both necessary and sufficient for admissibility. Recall that $\mathbf{x}^t$ denote the $t-$dimensional vector with x in each component, so that $\mathbf{x}^{(1+n(t-1))}$ has the same dimension as each player's private history, h_i^t.

Definition 3 (Stationary Strategies)**.** *A strategy μ_i is stationary if for any stage t, the distribution over i's action in stage t, i's private signal and all past actions are together translation invariant: for all $x \in \mathbb{R}$, $\mu_{i,X_{n(t-1)+i}}(a_i^t + x | h_i^t + \mathbf{x}^t)$ is independent of x.*

Definition 4 (Limit Strategies and Near Stationarity)**.** *Given a strategy μ_i, we say that a strategy μ^* is a limit strategy for μ_i if $\lim_{\theta\to+\infty} \mu_i^\theta = \lim_{\theta\to-\infty} \mu_i^\theta = \mu^*$, with limits taken with respect to the Prokhorov metric. If μ^* is stationary, then we classify μ_i as* nearly stationary.

It is immediate from Definition 4 that a strategy can have at most one limit strategy. However, a strategy need not have any limit strategy. For a counterexample, consider a game with a single player receiving some signal s where $s - \theta$ has distribution F conditional on θ, and suppose the action space given s is $\{s, s+1\}$. Then the following strategy, call it μ, has no limit strategy: assign action $a = s+1$ for all signals $s \geq 0$ and action $a = s$ for all signals $s < 0$. As $\theta \to +\infty$, $\mu \to \kappa_1$, where we use κ_x to denote the stationary strategy characterized by $a(s) = s + x$ with probability 1 for all $s \in \mathbb{R}$. As $\theta \to -\infty$, $\mu \to \kappa_0 \neq \kappa_1$, which fails Definition 4.

The next lemma provides a useful property equivalent to the one in Definition 4.

Lemma 3. *A strategy μ^* is a limit strategy for μ_i if and only if for all $\eta > 0$, $\lambda\{\theta : d_P(\mu_i^\theta, \mu^*) \geq \eta\} < \infty$, where d_P is the Prokhorov metric.*

Proof. For the "only if" direction, note that by the definition of a limit, $\lim_{\theta\to+\infty} \mu_i^\theta = \lim_{\theta\to-\infty} \mu_i^\theta = \mu^*$ implies that for all η, there exists $M > 0$ such that for all $\theta \leq -M$ and all $\theta \geq M$, $d_P(\mu_i^\theta, \mu^*) < \eta$. Hence $\{\theta : d_P(\mu_i^\theta, \mu^*) \geq \eta\} \subseteq [-M, M]$ and thus $\lambda\{\theta : d_P(\mu_i^\theta, \mu^*) \geq \eta\} < \infty$. For the "if" direction, we prove the contrapositive. Suppose that some arbitrary strategy μ^* is not a limit strategy for μ_i. Then given any $\eta > 0$, it is possible to construct a sequence $(\theta_j)_{j\in\mathbb{N}}$ such that (i) for all $j, k \in \mathbb{N}$ with $j \neq k$, $|\theta_j - \theta_k| > 1$ and (ii) for all $j \in \mathbb{N}$, $d_P(\mu^\theta, \mu^*) \geq 2\eta$. Since the map $\theta \mapsto \mu_i^\theta$ is uniformly continuous (Lemma 2), there exists $\delta(= \eta) > 0$ such that whenever $|\theta - \theta'| < \delta$, $d_P(\mu_i^\theta, \mu_i^{\theta'}) < \eta$. It follows that for all $k \in \mathbb{N}$ and all $\theta' \in (\theta_k - \delta, \theta_k + \delta)$, $d_P(\mu_i^{\theta'}, \mu^*) \geq d_P(\mu^\theta, \mu^*) - d_P(\mu_i^\theta, \mu_i^{\theta'}) > 2\eta - \eta = \eta$. Since the set $\lambda\left(\bigcup_{k\in\mathbb{N}} (\theta_k - \delta, \theta_k + \delta)\right) = \infty$, and thus $\lambda\{\theta : d_P(\mu_i^\theta, \mu^*) \geq \eta\} = \infty$, concluding the proof of the contrapositive. □

The following Lemma shows that the first part of Definition 4 implies the second.

Lemma 4. *If μ^* is a limit strategy of μ_i in the sense of Definition 4, then μ^* is stationary and thus μ_i is nearly stationary.*

Proof. If μ^* is not stationary, then there exists θ such that $\mu^{*,\theta} \neq \mu^*$. Now if μ_i is near μ^*, then μ_i^θ is near $\mu^{*,\theta}$. But also note that if μ_i is near μ^*, then $\mu_i^{\theta'}$ is also near μ^* for all $\theta' \in \mathbb{R}$, including $\theta' = \theta$, and thus μ_i^θ is near μ^*. Since a strategy can be near at most one strategy, we have $\mu^{*,\theta} = \mu^*$, a contradiction. We conclude that μ^* is stationary. □

Note that a profile of strategies induces ex interim (i.e., conditional on θ) expected payoffs for each player:

$$u_i(\mu_i, \mu_{-i}, \theta) := \int_{a \in X} u_i(a_i, a_{-i}, \theta)\, d(\times \mu_i(a_i; \theta))$$
$$= \int_{a \in X} u_i(a_i - \theta^\tau, a_{-i} - \theta^\tau, 0)\, d(\times \mu_i(a_i; \theta)),$$

where the integral is taken over all vectors a of actions across stages and players, and where the second line uses stationarity of payoffs.

In general, ex post payoffs can vary widely due to the noise in the signals. To ensure that the ex interim payoffs (given a realization of θ, but not the signal realization) are finite, we impose the following assumption, which is a joint condition on the payoff function and signal distribution.

Assumption 3 (Finite Interim Payoffs)**.** *Interim payoffs are finite: together, the signal distributions F_i and the payoff functions u_i are such that for all θ and for all strategy profiles $(\mu_1, \mu_2, \ldots, \mu_N)$, the interim expected payoff $u_i(\mu_i, \mu_{-i}, \theta)$ defined above is well-defined and finite.*

Assumption 3 is a relatively weak requirement. For example, in a single-stage game, any finite polynomial payoff function, such as quadratic-loss, together with a normally distributed signal satisfies this assumption. This stems from the fact that all moments of the normal distribution are finite.

Lemma 5. *For all profiles of strategies $(\mu_1, \mu_2, \ldots, \mu_N)$ and θ, $u_i(\mu_i, \mu_{-i}, \theta)$ is uniformly continuous in each μ_j and in θ.*

Proof. Since $u_i(a_i, a_{-i}, \theta)$ is continuous in a_j, it is continuous in each μ_j. Since the μ_j lie in a compact domain, we have uniform continuity in each μ_j. Next, since $u_i(\mu_i, \mu_{-i}, \theta) = u_i(\mu_i^{-\theta}, \mu_i^{-\theta}, 0)$ and by Lemma 2 each $\mu_j^{-\theta}$ term is uniformly continuous in θ, we have uniform continuity in θ. □

For ease of exposition, from now on we require that the game involves no redundant strategies – strategies which are indistinguishable in their payoff implications are labeled as the same strategy.

Assumption 4 (Irreducibility)**.** *The game is irreducible in that, for any two strategies $\mu_i \neq \mu_i'$ and state θ, there exists a profile μ_{-i} of rivals' stationary strategies such that $u_i(\mu_i, \mu_{-i}, \theta) \neq u_i(\mu_i', \mu_{-i}, \theta)$.*

3.3 The Diffuse Prior

The informal concept of diffuse prior has two key properties. The first property is diffuseness – that is, all real numbers are in the support of the prior. The second property is uniformity – all points are weighted equally. Hence, we define a sequence of proper measures to be *diffusing* if these properties hold in the limit – that is, sufficiently far into the sequence, the properties hold arbitrarily closely. The following definition formalizes this idea. We use λ to denote the Lebesgue measure.

Definition 5 (Diffusing Sequence)**.** *Consider a sequence* $(\mathbb{P}_n)_{n\in\mathbb{N}}$ *of Borel probability measures on* $\mathbb{R}$*. We say that this sequence is* diffusing *if for any set* X *of finite and strictly positive Lebesgue measure and any* $\eta > 0$*, there exists* $N \in \mathbb{N}$ *such that for all* $n \geq N$*,*

- $\mathbb{P}_n(X) > 0$*, and*
- *for all measurable* $Y \subseteq X$*,* $\left|\frac{\mathbb{P}_n(Y)}{\mathbb{P}_n(X)} - \frac{\lambda(Y)}{\lambda(X)}\right| < \eta$*.*

From the definition, we can establish the following property of diffusing sequences as a result.

Lemma 6. *If* $(\mathbb{P}_n)$ *is diffusing, then for any finite-measure set* $E \subset \mathbb{R}$*,* $\lim_{n\to\infty} \mathbb{P}_n(E) = 0$*.*

Proof. Consider any finite-measure subset $E \subset \mathbb{R}$ and any diffusing sequence $(\mathbb{P}_n)$. Choose any arbitrarily small $\eta > 0$. Choose $M > 1$ sufficiently large that (i) $\frac{1}{M} < \eta$ and (ii) $\lambda(E \setminus [-M, M]) < \eta\lambda(E)$. Let $X_M := [-M, M]$ and $X_{M^2} := [-M^2, M^2]$. By applying Definition 5 twice, first with the set E playing the role of X in the definition and again with X_{M^2} playing the role of X, there exists N such that for all $n \geq N$,

$$\mathbb{P}_n(X_{M^2}) > 0$$
$$\mathbb{P}_n(E) > 0$$
$$\left|\frac{\mathbb{P}_n(E \setminus X_M)}{\mathbb{P}_n(E)} - \frac{\lambda(E \setminus X_M)}{\lambda(E)}\right| < \eta \tag{1}$$
$$\left|\frac{\mathbb{P}_n(X_M)}{\mathbb{P}_n(X_{M^2})} - \frac{\lambda(X_M)}{\lambda(X_{M^2})}\right| < \eta. \tag{2}$$

In (1), recall that by construction $\frac{\lambda(E\setminus X_M)}{\lambda(E)} < \eta$ and thus rearranging (1) yields

$$\mathbb{P}_n(E \setminus X_M) < 2\eta\mathbb{P}_n(E) < 2\eta. \tag{3}$$

In (2), we have $\frac{\lambda(X_M)}{\lambda(X_{M^2})} = \frac{1}{M} < \eta$ and thus

$$\mathbb{P}_n(X_M) < 2\eta\mathbb{P}_n(X_{M^2}) < 2\eta. \tag{4}$$

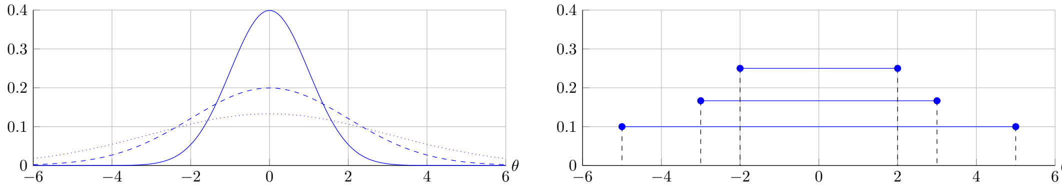
Adding (3) and (4) then yields

$$\mathbb{P}_n(E) \leq \mathbb{P}_n(E \setminus X_M) + \mathbb{P}_n(X_M) < 4\eta.$$

Since η is arbitrary, we have $\mathbb{P}_n(E) \to 0$. □

For illustrative purposes, we highlight two specific diffusing sequences (see Figure 1). As one would expect, flattening sequences of the uniform distribution or normal distribution are diffusing according to our definition.

Example 1. *Both* $(\mathbb{P}_n^1)$ *given by the density* $\frac{1}{2n}\mathbb{1}\{[-n, n]\}$ *and* $(\mathbb{P}_n^2)$ *given by* $N(0, n)$ *are diffusing.*



(a) Flattening Normal Distributions

(b) Flattening Uniform Distributions

Figure 1: Examples of Diffusing Sequences

Proof. If X is bounded and nonempty, then $X \subset [-n, n]$ for some n, and then $\mathbb{P}^1_n(X) = \lambda(X)/(2n) > 0$. Moreover, if $Y \subset X$, then $\mathbb{P}^1_n(Y) = \lambda(Y)/(2n)$ and thus $\left|\frac{\mathbb{P}^1_n(Y)}{\mathbb{P}^1_n(X)} - \frac{\lambda(Y)}{\lambda(X)}\right| = 0$. Now $\mathbb{P}^2_n(X) > 0$ for all $n \in \mathbb{N}$. Let h_n denote the density of $N(0, n)$. Note that for all $\eta \in (0, 1)$, for sufficiently large n, $\frac{\inf_{x \in X} h_n(x)}{\sup_{x \in X} h_n(x)} > 1 - \eta$. Hence $(1-\eta)^2 \frac{\lambda(Y)}{\lambda(X)} < \frac{\mathbb{P}^2_n(Y)}{\mathbb{P}^2_n(X)} < (1-\eta)^{-2} \frac{\lambda(Y)}{\lambda(X)}$, and by algebra the result follows. □

3.4 Admissibility

Recall the example from Section 2, where we asked which functions of θ are "admissible" in the sense that they yield consistent limits when integrated along certain kinds of sequences of distributions. We suggested that admissible functions are nearly constant in a particular way: there exists some constant such that for any $\epsilon > 0$, the set of θ on which function deviates from that constant by more than ϵ has finite measure. As we show in our main result, the spirit of this example extends to admissibility of strategies. Since players only observe noisy signals of θ, their strategies are not directly functions of θ but functions of their private signals. Nonetheless, having well-defined expected payoffs will place restrictions on players strategies, similar to how admissible functions above are nearly constant.

Definition 6 (Admissibility)**.** *A class $\mathcal{M}_0 \subset \mathcal{M}$ of strategies is said to be* admissible *if for any profile $(\mu_1, \mu_2, ..., \mu_N)$ of strategies $\mathcal{M}_0$, there exists $u^* \in \mathbb{R}$ such that for any sequence $(\mathbb{P}_n)$, $\lim_{n\to\infty} \int_{\theta \in \mathbb{R}} u_i(\mu_i, \mu_{-i}, \theta)\, d\mathbb{P}_n(\theta) = u^*$.*

Let $\mathcal{K}$ denote the class of stationary strategies. Note that for stationary strategies, the function $u(\mu_i, \mu_{-i}, \theta)$ defined above is constant in θ, and hence the sequence of integrals in Definition 6 trivially converges to the same limit for every diffusing sequence.

For the proof of the main result, we make use of a weaker property than that of Definition 4.

Lemma 7. *Let (Y, d) be a compact metric space, $X \subseteq \mathbb{R}$ with $\lambda X = \infty$, and $\pi : X \to Y$ a Lebesgue measurable function. There exists $y \in Y$ with the property that for all $\eta > 0$, $\lambda\{x \in X : d(\pi(x), y) < \eta\} = \infty$. We say that any such y is an* attraction*.*

We now present the main result.

Theorem 1. *Let Γ be an irreducible game with stationary payoffs.*

- *(Sufficiency) The class of nearly stationary strategies is admissible.*
- *(Necessity) If $\mathcal{M}_0$ is admissible and $\mathcal{K} \subset \mathcal{M}_0$, then every $\mu \in \mathcal{M}_0$ is nearly stationary and payoff equivalent to some stationary strategy.*

In the case of a single-stage game, Theorem 1 can be specialized. A stationary strategy in this case can be described by a single distribution over $[-M, M]$, interpreted as a (possibly random) markup x so that the action chosen is $a = s_i + x$, and any admissible strategy must be near, and payoff equivalent to some such stationary strategy.

4 Application: Delegation to One of Two Imperfectly Informed Experts

Note that the assumption that players' signals are drawn conditionally from the same distribution has played no role in any of our arguments. The baseline model can be generalized immediately to allow heterogeneous (but still conditionally independent) signal distributions. As a special case, this allows us to model an uninformed player as one who receives a trivial signal that conveys no information about the true state. We can thus extend the baseline model to cover sender-receiver games.

To illustrate, we consider a game with two stages: one stage in which n senders simultaneously choose actions after observing private signals, and a second stage in which the receiver observes the senders' actions and chooses an action. Depending on the application in mind, the receiver's action could be interpreted as a continuous variable (as in games of cheap talk) or as a sender's identity (as in games of delegation). In the case of cheap talk, the baseline model with heterogeneous signal distributions is already sufficient.

In the case of delegation, the receiver's available action choices may depend on the actions chosen by the senders. For example, in a companion paper by Ambrus et al. (2017), the senders are two experts who propose action choices, and the receiver must choose one of them. All players have quadratic loss functions. Senders have biases b_i and the selected sender gets a constant bonus $B \geq 0$. The game unfolds as follows:

1. Senders (experts) $i = 1, 2$ receive conditionally i.i.d. signals $s_i \sim N(\theta, \sigma^2)$.
2. Senders choose markups $k_i \in [-M, M]$ where M is some large constant.
3. The receiver (principal or decision maker) observes $a_i = s_i + k_i$ for $i \in \{1, 2\}$ and chooses which action will be implemented. Specifically, the receiver chooses an action $C(a_1, a_2)$ from $\{a_1, a_2\}$.[8]
4. Payoffs are realized: receiver gets $-(s_j + k_j - \theta)^2$, sender j gets $B - (s_j + k_j - \theta - b_j)^2$, sender $i \neq j$ gets $-(s_j + k_j - \theta - b_i)^2$.

[8]Note that $\{a_1, a_2\}$ is not a set, but a multiset, because it can contain the same element multiple times (if $a_1 = a_2$).

A stationary strategy for the receiver must have the property that $C(a_1, a_2) = C(a_1 + x, a_2 + x)$ for all x. This includes mixtures of the following fundamental strategies: (i) always choose the higher offer, (ii) always choose the lower offer, (iii) always choose the offer from sender 1, and (iv) always choose the offer from sender 2. Ambrus et al. (2017) show that these are essentially the only possible best responses of the principal to a pair of stationary strategies of the experts.

Note that once we have shown irreducibility on the side of the senders, Theorem 1 can be applied to the game between senders, which implies that all admissible strategies are nearly stationary, and in particular, are nearly "constant markup strategies" – that is, each player's strategy is nearly a stationary strategy in which there is a single "markup" distribution G_i over $[-M, M]$ and actions are simply s_i plus the draw from G_i. It is easy to show that the receiver's best response to any such strategy profile is stationary and equivalent to one of the above forms, and that the set of stationary receiver strategies is also irreducible.

Proposition 1. *The game of Ambrus et al. (2017) satisfies the stationarity and irreducibility assumptions.*

Proof. Stationarity of signals and payoffs is immediate from the definition of the game, so we focus on irreducibility. We show irreducibility for sender 1, which by symmetry implies irreducibility for sender 2. Irreducibility for the principal is easy to show, since for two distinct principal strategies, one can marginally adjust one or both of the sender strategies and change the principal's payoffs differentially. Suppose the principal always chooses the lower offer: given offers a_1 and a_2, she chooses $C(a_1, a_2) = \min\{a_1, a_2\}$. Consider two strategies of expert 1, μ_1 and $\hat{\mu}_1$. WLOG, we consider only $\theta = 0$ and $b_1 = 0$. We show that there exists a constant markup strategy κ_m for player 2 such that $u_1(\mu_1, \kappa_m, 0) \neq u_1(\hat{\mu}_1, \kappa_m, 0)$. By the definition of payoffs and signals in the model, $u_1(\mu_1, \kappa_m, C, 0) = -\int_{\mathbb{R}^2} (\min\{a_1, s_2 + m\})^2 (Q \otimes F)(d(a_1, s_2))$, where Q is the CDF over player 1's action induced by the strategy μ_1, defined by $Q(x) = \mu_1(\{a_1 : a_1 \leq x\})$, and where F is the CDF of $N(0, \sigma^2)$. Define $\hat{Q}(x)$ for $\hat{\mu}_1$ likewise. Next, we show that we can differentiate w.r.t. m under the integral. Write $\mathbf{a} := (a_1, s_2), g(\mathbf{a}, m) := (\min\{a_1, s_2 + m\})^2$, and $\nu = Q \otimes F$. Note that for all $\mathbf{a} \in \mathbb{R}^2$, $g(\mathbf{a}, m)$ is absolutely continuous in m on bounded intervals. We have

$$a(m) := \int_{\mathbb{R}^2} g(\mathbf{a}, m) \nu(d\mathbf{a}) \tag{5}$$

$$= \int_{\mathbb{R}^2} \left[g(\mathbf{a}, m_0) + \int_{m_0}^{m} g_m(\mathbf{a}, z) dz \right] \nu(d\mathbf{a}) \tag{6}$$

$$= \int_{\mathbb{R}^2} g(\mathbf{a}, m_0) \nu(d\mathbf{a}) + \int_{m_0}^{m} \int_{\mathbb{R}^2} g_m(\mathbf{a}, z) \nu(d\mathbf{a}) dz \tag{7}$$

$$\implies a'(m) := \int_{\mathbb{R}^2} g_m(\mathbf{a}, m) \nu(d\mathbf{a}) \qquad \text{a.e. } m, \tag{8}$$

where $m_0 < m$ can be chosen arbitrarily. To obtain (6) we have used the fact that absolutely continuous functions are the integral of their derivatives.[9] Fubini's Theorem is used to obtain (7). Differentiability-a.e. of the integral yields (8).[10] By the same arguments, we obtain $a''(m) =$

[9] See Royden and Fitzpatrick (1988, Corollary 5.15).
[10] See Royden and Fitzpatrick (1988, Theorem 5.10).

$\int g_{mm}(\mathbf{a}, m)\nu(d\mathbf{a}) = 2\int_{\{\mathbf{a}:a_1>s_2+m\}} \nu(d\mathbf{a})$ a.e. m. Applying this to $\nu = Q \otimes F$ and $\hat{\nu} = \hat{Q} \otimes F$, it follows that if $u_1(\mu_1, \kappa_m, C, 0) = u_1(\hat{\mu}_1, \kappa_m, C, 0)$ for all $m \in \mathbb{R}$, then $\int_{\{\mathbf{a}:a_1>s_2+m\}} (Q \otimes F)(d(a_1, s_2)) = \int_{\{\mathbf{a}:a_1>s_2+m\}} (\hat{Q} \otimes F)(d(a_1, s_2))$ a.e. m. Rearranging, we have

$$\begin{aligned}\int_{S_2} (1 - Q(s_2 + m))a(s_2)ds_2 &= \int_{S_2} (1 - \hat{Q}(s_2 + m))a(s_2)ds_2 \\ \implies 0 &= \int_{S_2} (Q(m - s_2) - \hat{Q}(m - s_2))a(s_2)ds_2, \qquad (9)\end{aligned}$$

by the evenness of the normal distribution. Letting $K := Q - \hat{Q}$, (9) is a convolution equation:

$$[K * f](m) = 0 \qquad \text{a.e. } m.$$

Let $\mathcal{F}$ denote the normalized Fourier transform, $\mathcal{F}(g)(z) := \int_{-\infty}^{\infty} e^{-2\pi izx} g(x)dx$. Both K and f are Lebesgue integrable functions, and thus by the convolution theorem,[11] $\mathcal{F}(K * f) = \mathcal{F}(K) \cdot \mathcal{F}(f)$. But since $K * f \equiv 0$, $\mathcal{F}(K * f) \equiv 0$. Since $\mathcal{F}(f)(z) = e^{-2(\pi\sigma z)^2} > 0$ for all z, we must have $\mathcal{F}(K) \equiv 0$. Applying the inverse Fourier transform, we have $K = 0$ almost everywhere. This contradicts the assumption that $Q \neq \hat{Q}$, so it must be that $u_1(\mu_1, \kappa_m, C, 0) \neq u_1(\hat{\mu}_1, \kappa_m, C, 0)$ for some m. □

A Proofs

Proof of Lemma 1. Note that $\mathcal{M}_i$ is a metric space, as it is a subspace of the metric space $\mathcal{P}$ consisting of measures over the complete and separable metric space $S_i \times \mathbb{R}^{n\tau}$. We prove that $\mathcal{M}_i$ is compact by showing that it is relatively compact and that it is closed. Below we show that $\mathcal{M}_i$ is relatively compact; to verify that $\mathcal{M}_i$ is also closed requires several tedious steps of showing that convergent sequences of strategies converge to limits which satisfy the key properties of Definition 1.

Note first that X is compact: by standard results,[12] X is a countable product of compact metric spaces Y_i and is thus itself a compact metric space under the product metric defined by, for example,

$$d(x_i, \hat{x}_i) := \sum_{k=1}^{n\tau} \frac{\min\{1, |x_{i,k} - \hat{x}_{i,k}|\}}{2^k}.$$

Given any $\eta > 0$, choose compact interval $[-K, K]$ such that $T := [-K, K]^n$ satisfies $\phi(T) > 1 - \eta$. Define $\bar{a}_i := M_i + K$ and $\underline{a} := -M_i - K$. Let $Y := \times_{j=1}^{n} \times_{t=1}^{\tau} [\underline{a}_i, \bar{a}_i]$. Then $T \times Y$ is compact and for all $\mu_i \in \mathcal{M}_i$, $\mu_i(T \times Y) = \phi(T) > 1 - \eta$; thus $\mathcal{M}_i$ is relatively compact. □

Proof of Lemma 7. Suppose on the contrary that for each $y \in Y$, there exists $\eta_y > 0$ such that $\lambda\{x \in \mathbb{R} : d(\pi(x), y) < \eta_y\} < \infty$. The collection $\{N_{\eta_y}(y) : y \in Y\}$ is an open covering of Y, and by compactness, it has a finite subcovering denoted $\{N_{\eta_i}(y_i)\}_{i=1}^{n}$ for some $n \in \mathbb{N}$. It follows that $\mathbb{R} \subseteq \cup_{i=1}^{n} \pi^{-1}(N_{\eta_i}(y_i))$ and thus $\lambda X \leq \sum_{i=1}^{n} \lambda\pi^{-1}(N_{\eta_i}(y_i)) < \infty$, a contradiction. □

[11] See Reed and Simon (1980, Theorem IX.3b).
[12] For example, see (Ok, 2007, Theorem C.4).

Proof of Theorem 1. Sufficiency: Let $\mathcal{M}_0$ be the class of nearly stationary strategies, and consider any profile $(\mu_1, \mu_2, \ldots, \mu_N)$ of strategies in $\mathcal{M}_0$. By definition, for each $\mu_i \in \mathcal{M}_0$, there exists a strategy μ_i^* such that μ_i^* is a limit strategy for μ_i; by Lemma 4, this μ_i^* is stationary and by Lemma 3, for all $\eta > 0$,

$$\lambda\{\theta : d_P(\mu_i^\theta, \mu_i^*) \geq \eta\} < \infty. \tag{10}$$

Next, define $\Theta_{\geq\eta} = \{\theta \in \mathbb{R} : \max_{i\in N} d_P(\mu_i^\theta, \mu_i^*) \geq \eta\}$, which has finite measure by (10) and the fact that there are finitely many players. Let $\mathbb{P}_n$ be any diffusing sequence. For all n, by definition of the recentering function we have

$$\int_{\theta\in\mathbb{R}} u_i(\mu_i, \mu_{-i}, \theta)\, d\mathbb{P}_n(\theta) = \int_{\theta\in\mathbb{R}} u_i(\mu_i^{-\theta}, \mu_{-i}^{-\theta}, 0)\, d\mathbb{P}_n(\theta). \tag{11}$$

We can write the RHS of (11) as

$$\int_{\theta\in\Theta_{\geq\eta}} u_i(\mu_i^{-\theta}, \mu_{-i}^{-\theta}, 0)\, d\mathbb{P}_n(\theta) + \int_{\theta\in\mathbb{R}\setminus\Theta_{\geq\eta}} u_i(\mu_i^{-\theta}, \mu_{-i}^{-\theta}, 0)\, d\mathbb{P}_n(\theta) \tag{12}$$

In the first term of (12), by Lemma 5, the integrand $u_i(\mu_i^{-\theta}, \mu_{-i}^{-\theta}, 0)$ is bounded in absolute value by some quantity, call it M. For the second term of (12), note that by uniform continuity of ex interim payoffs, for any $\epsilon > 0$, there exists $\eta > 0$ such that for any profile of strategies $(\tilde{\mu}_1, \tilde{\mu}_2, \ldots, \tilde{\mu}_N)$, if $d_P(\tilde{\mu}_i, \mu_i^*) < \eta$ for all $i \in N$, then

$$|u_i(\tilde{\mu}_i, \tilde{\mu}_{-i}, 0) - u_i(\mu_i^*, \mu_{-i}^*, 0)| < \epsilon.$$

By definition, for $\theta \in \mathbb{R}\setminus\Theta_{\geq\eta}$ we have $\max_{i\in N} d_P(\mu_i^\theta, \mu_i^*) < \eta$ and thus $|u_i(\mu_i^{-\theta}, \mu_-^{-\theta}, 0) - u_i(\mu_i^*, \mu_{-i}^*, 0)| < \epsilon$. Putting these together,

$$\begin{aligned}
\left|\int_{\theta\in\mathbb{R}} u_i(\mu_i, \mu_{-i}, \theta)\, d\mathbb{P}_n(\theta) - u^*\right| \leq\ & \int_{\theta\in\Theta_{\geq\eta}} |u_i(\mu_i^{-\theta}, \mu_{-i}^{-\theta}, 0) - u^*|\, d\mathbb{P}_n(\theta) \\
&+ \int_{\theta\in\mathbb{R}\setminus\Theta_{\geq\eta}} |u_i(\mu_i^{-\theta}, \mu_{-i}^{-\theta}, 0) - u^*|\, d\mathbb{P}_n(\theta) \\
\leq\ & \mathbb{P}_n(\Theta_{\geq\eta})(M + |u^*|) + \epsilon \cdot \mathbb{P}_n(\mathbb{R} \setminus \Theta_{\geq\eta}) \\
\leq\ & \mathbb{P}_n(\Theta_{\geq\eta})(M + |u^*|) + \epsilon
\end{aligned} \tag{13}$$

Since $\Theta_{\geq\eta}$ has finite Lebesgue measure, by Lemma 6 there exists N such that for $n \geq N$, $\mathbb{P}_n(\Theta_{\geq\eta}) < \frac{\epsilon}{M+|u^*|}$, and thus the RHS of (13) is less than 2ϵ. Since ϵ is arbitrary, we have shown that $\lim_{n\to\infty} \int_{\theta\in\mathbb{R}} u_i(\mu_i, \mu_{-i}, \theta)\, d\mathbb{P}_n(\theta) = u^*$, so by definition, the class of nearly stationary strategies is admissible.

Necessity: First, we show that μ is nearly stationary, and then we show payoff equivalence. Suppose otherwise that μ is not nearly stationary. Now the map $\theta \mapsto \mu^\theta$ is a measurable function from $\mathbb{R}$ to $\mathcal{M}$, which is compact by Assumption 1. By Lemma 7, there exists an attraction $\mu^* \in \mathcal{M}$ for μ.

Next, given the existence of an attraction μ^*, we establish uniqueness. Suppose there also exists an attraction $\hat{\mu} \in \mathcal{M}$ for μ with $\hat{\mu} \neq \mu^*$. We show that there exists a profile of stationary strategies of the rivals such that if player i plays μ, he does not have a well-defined expected payoff in the limit. By irreducibility of payoffs (Assumption 4), there exists a profile μ_{-i} of stationary rival strategies such that $u(\mu^*, \mu_{-i}, 0) \neq u(\hat{\mu}, \mu_{-i}, 0)$. We show that this contradicts admissibility.

Note that by Lemma 5, for any $\epsilon > 0$, there exists $\eta > 0$ such that if $d_P(\mu^\theta, \mu^*) < \eta$, then

$$\epsilon > |u_i(\mu^\theta, \mu_{-i}, 0) - u_i(\mu^*, \mu_{-i}, 0)| \tag{14}$$

$$= |u_i(\mu, \mu_{-i}, -\theta) - u_i(\mu^*, \mu_{-i}, -\theta)|, \tag{15}$$

where the equality uses the definition of the recentering function and the fact that μ^* and the strategies in μ_{-i} are stationary. An analogous statement holds for $\hat{\mu}$, so let us redefine η so that for $\tilde{\mu} \in \{\mu^*, \hat{\mu}\}$, $d_P(\mu^\theta, \tilde{\mu}) < \eta$ implies $|u_i(\mu^\theta, \mu_{-i}, \theta) - u_i(\tilde{\mu}, \mu_{-i}, \theta)| < \epsilon$.

Recall that by admissibility, there is some u^* such that $\lim_{n\to\infty} \int u_i(\mu_i, \mu_{-i}, \theta) d\mathbb{P}_n(\theta) = u^*$ for all diffusing sequences $(\mathbb{P}_n)$. For the contradiction, we construct two sequence of measures $(\mathbb{P}_n^1)$ and $(\mathbb{P}_n^2)$ along which the limits differ. Let $v^* = u_i(\mu^*, \mu_{-i}, 0)$ and $\hat{v} = u_i(\hat{\mu}, \mu_{-i}, 0)$, and recall that $v^* \neq \hat{v}$. Since μ^* is an attraction for μ and $\lambda\{\theta : d_P(\mu^\theta, \mu^*) < \eta\} = \infty$, for each $n \in \mathbb{N}$, there exists $C_n^1 \subset \{\theta : d_P(\mu^\theta, \mu^*) < \eta\} \setminus [-n, n]$ with $\lambda(C_n^1) = 2n^2$. Define $B_n^1 := [-n, n] \cup C_n^1$, and define $\mathbb{P}_n^1(\theta) := \mathbb{1}_{B_n^1}(\theta)/\lambda(B_n^1)$. By construction, $(\mathbb{P}_n^1)_{n\in\mathbb{N}}$ is a diffusing sequence of measures, and by the assumption that μ_i is admissible, we must have $\lim_{n\to\infty} \int u_i(\mu^\theta, \mu_{-i}, \theta) d\mathbb{P}_n^1 = u^*$. Pick any $C > |v^*|$ such that C is an upper bound, over all θ, on the magnitude of $u_i(\mu^\theta, \mu_{-i}, \theta)$. We have

$$\begin{aligned}
\left| \int_{\theta\in\mathbb{R}} u_i(\mu^\theta, \mu_{-i}, \theta) d\mathbb{P}_n^1(\theta) - v^* \right| \leq\ & \int_{\theta\in\mathbb{R}} \left| u_i(\mu^\theta, \mu_{-i}, \theta) - v^* \right| d\mathbb{P}_n^1(\theta) \\
=\ & \int_{\theta\in B_n^1} \left| u_i(\mu^\theta, \mu_{-i}, \theta) - v^* \right| d\mathbb{P}_n^1(\theta) \\
=\ & \int_{\theta\in C_n^1} \left| u_i(\mu^\theta, \mu_{-i}, \theta) - v^* \right| d\mathbb{P}_n^1(\theta) \\
& + \int_{\theta\in[-n,n]} \left| u_i(\mu^\theta, \mu_{-i}, \theta) - v^* \right| d\mathbb{P}_n^1(\theta) \\
\leq\ & \frac{\epsilon \cdot 2n^2 + 2C \cdot 2n}{2n^2 + 2n} \to \epsilon
\end{aligned}$$

and thus $u^* \in [v^* - \epsilon, v^* + \epsilon]$. Likewise, $\hat{\mu}$ is an attraction, so $\lambda\{\theta : d_P(\mu^\theta, \hat{\mu}) < \eta\} = \infty$, and we define $C_n^2 \subset \{\theta : d_P(\mu^\theta, \hat{\mu}) < \eta\} \setminus [-n, n]$ with $\lambda C_n^2 = 2n^2$, $B_n^2 = [-n, n] \cup C_n^2$, and $\mathbb{P}_n^2(\theta) := \mathbb{1}_{B_n^2}(\theta)/\lambda(B_n^2)$, such that $|\int u(\theta) d\mathbb{P}_n^2 - v'| \to \epsilon$, and thus $u^* \in [\hat{v} - \epsilon, \hat{v} + \epsilon]$. Since η is arbitrary, we choose $\epsilon < \frac{|v^* - \hat{v}|}{2}$ and obtain a contradiction. Hence we conclude that there is a unique attraction μ^*.

We now prove that μ^* is a limit strategy for μ in the sense of Definition 4. By definition, this claim is that for all $\eta > 0$, $\lambda(\Theta_{\geq\eta}) < \infty$, where $\Theta_{\geq\eta} := \{\theta : d_P(\mu^\theta, \mu^*) \geq \eta\}$. We derive a contradiction by showing that otherwise, the uniqueness result above would be violated. Suppose by way of contradiction that μ^* is not a limit strategy of μ, and hence by Lemma 3, for some $\eta > 0$, $\lambda(\Theta_{\geq\eta}) = \infty$ where $\Theta_{\geq\eta} := \{\theta : d_P(\mu^\theta, \mu^*) \geq \eta\}$. Let $\mathcal{Q} := \{\tilde{\mu} \in \mathcal{M} : d_P(\tilde{\mu}, \mu^*) \geq \eta\}$. By Lemma 7, there exists an attraction $\hat{\mu} \in \mathcal{Q}$ for μ_i. By construction, $\hat{\mu} \neq \mu^*$. This contradicts the uniqueness of the attraction μ^* as argued previously, so μ^* must be a limit strategy for μ, as desired.

Finally, μ^* is stationary by Lemma 4, and μ is nearly stationary. □

B Full (non-Reduced) Strategic Form Strategies

Here we extend Definition 1 to the full strategic form. To illustrate how Definition 1 fails, consider the following minimal example. There is a single player who takes an action in two stages. In stage 1, he chooses from $\{a_1, b_1\}$ and in stage 2, he chooses from $\{a_2, b_2\}$. The reduced strategic form identifies the following two strategies as equivalent: (i) choose a_1 w.p. 1, then choose a_2 w.p. 1 regardless of the action chosen in period 1, and (ii) choose a_1 w.p. 1, then choose a_2 in period 2 if and only if a_1 was chosen in period 1. The distributional strategy under Definition 1 that represents these strategies is a CDF over $\{a_1, b_1\} \times \{a_2, b_2\}$ that places mass 1 on the point (a_1, a_2). In order to describe behavior in period 2 after a deviation in period 1, it is necessary to include another CDF, for which a conditional probability distribution is well-defined for each period 1 action. Thus, we define a non-reduced strategy as a pair (F_1, F_2) of CDFs over $\{a_1, b_1\} \times \{a_2, b_2\}$ such that F_2 has uniform marginal distribution in the first dimension. Intuitively, this approach splits each player into a sequence of distinct players, one for each stage of the game.

To generalize to the n-player, τ-stage game, we define a strategy as a sequence of CDFs $(\mu_{i,t})_{t=1}^{\tau}$ with the following properties. First, $\mu_{i,1}$ is an exact copy of the reduced-form strategy as before. Next, $\mu_{i,2}$ is formed by assigning independent uniform distributions to the dimensions corresponding to $A_1^1, A_2^1, \ldots, A_n^1$ and keeping the conditional distributions consistent with those of $\mu_{i,1}$ (when the conditional distributions under $\mu_{i,1}$ exist). For general t, $\mu_{i,t}$ assigns independent uniform distributions to dimensions $A_j^{t'}$ for all $t' < t$ and all $j \in \{1, 2, \ldots, n\}$.

References

Ambrus, A., V. Baranovskyi, and A. Kolb (2017). A delegation-based theory of expertise. *Unpublished working paper*.

Billingsley, P. (2009). *Convergence of probability measures*, Volume 493. John Wiley & Sons.

Dale, D. J. and J. Morgan (2015). A comment on "The welfare effects of public information" by Stephen Morris and Hyun Song Shin. *mimeo, University of California, Berkeley*.

Friedman, D. (1991). A simple testable model of double auction markets. *Journal of Economic Behavior and Organization 15*, 47–70.

Fudenberg, D. and J. Tirole (1991). Game theory. 1991.

Klemperer, P. (1999). Auction theory: a guide to the literature. *Journal of Economic Surveys 13*, 227–286.

Milgrom, P. and R. Weber (1985). Distributional strategies for games with incomplete information. *Mathematics of Operations Research 10*, 619–632.

Morris, S. and H. Shin (2002). The social value of private information. *The American Economic Review 92*, 1521–1534.

Morris, S. and H. Shin (2003). Global games: Theory and applications. Econometrics Society Monographs.

Myatt, D. and C. Wallace (2014). Central bank communication design in a Lucas-Phelps economy. *Journal of Monetary Economics 63*, 64–79.

Ok, E. A. (2007). *Real analysis with economic applications*, Volume 10. Princeton University Press.

Reed, M. and B. Simon (1980). *Methods of modern mathematical physics: Functional analysis*, Volume 1. Gulf Professional Publishing.

Royden, H. and P. Fitzpatrick (1988). *Real analysis*, Volume 4. Prentice Hall New York.