Economics Bulletin

Volume 41, Issue 3

How much can we identify from repeated games?

Jose Miguel Abito
University of Pennsylvania, Business Economics and Public Policy

Cuicui Chen
State University of New York (SUNY) at Albany

Abstract

We propose a strategy to identify structural parameters in infinitely repeated games without relying on equilibrium selection assumptions. We exploit the extreme points of the equilibrium payoff set to construct bounds on the frequencies of stage game actions, which then impose restrictions on the parameters of interest. To illustrate the identification strategy, we use an infinitely repeated Prisoners Dilemma to get bounds on a utility parameter and a common discount factor.

We would like to thank the Editor and an anonymous referee for helpful comments. We also thank Mark Chicu, Matt Masten, and Tiago Pires for early conversations on this topic.

Citation: Jose Miguel Abito and Cuicui Chen, (2021) "How much can we identify from repeated games?", *Economics Bulletin*, Vol. 41 No. 3 pp. 1212-1222.

Contact: Jose Miguel Abito - abito@upenn.edu, Cuicui Chen - cchen27@albany.edu.

Submitted: December 06, 2020. **Published:** July 18, 2021.

1 Introduction

Game theoretical models and insights have been critical in capturing strategic interactions among economic agents in empirical settings. By combining assumptions on agents' rationality with an equilibrium concept, empirical researchers can link observed behavior with underlying model primitives such as agents' preference, technology, cost and constraints.

Equilibrium multiplicity typically precludes a unique mapping from primitives to observed behavior, making standard inference problematic without additional, often strong, assumptions. Thus, an important part of the literature on the econometrics of games has focused on the multiple equilibria problem (De Paula, 2013). While there has been substantial progress in analyzing static games, progress with dynamic games is more limited. One important class of dynamic games where identification results are virtually unexplored are infinitely repeated games.

Infinitely repeated games provide an important framework to model long-run relationships that lead to incentives and outcomes not otherwise captured by one-shot interactions. For example, the theory offers insights into how agents can cooperate in a non-cooperative environment without formal or explicit contracts. Despite its richness, the theory is often criticized for the lack of sharp predictions due to the equilibrium multiplicity (Dal Bó and Fréchette, 2011). The multiplicity problem can be so perverse in some circumstances (e.g., when the Folk Theorem holds) that it *seems* almost any kind of behavior can be rationalized in the data.

In this note, we derive an identified set for the structural parameters of an infinitely repeated game without assuming equilibrium selection or additional restrictions beyond subgame perfect Nash equilibrium. Our approach is general enough to accommodate games beyond two players, discrete or continuous actions, and different types of data available to the researcher. Data can come from a cross-section of repeated games, time series from a single game, or a combination of both. Finally, we also take a conservative approach to what kind of information the researcher has: only the frequencies of action profiles are observed and not individual histories.

We construct our identified set as follows. Actions chosen by players today, or at a given *history*, depend on what players expect to happen in the future. The one-stage deviation principle allows us to rationalize a given player's chosen action (conditional on the rival's) as an inequality involving the sum of the stage-game and *equilibrium* continuation payoffs. Bounds on the frequency of a given action profile can thus be constructed by first taking the extreme points of the equilibrium payoff set, and then finding necessary and sufficient conditions such that the given action profile is a Nash equilibrium in a particular normal form game. Finally, we can get the set of parameter values consistent with the bounds on the frequency of action profiles, and this is the identified set.

To the best of our knowledge, Lee and Stewart (2016) is the only paper that formally tackles identification in repeated games without stringent equilibrium selection assumptions. Their paper shows that payoffs in a repeated game can be point identified up to an affine transformation if one observes a player's full best response correspondence or at least the best response for a particular strategy. A player's best response correspondence (and strategies in repeated games in general) is a complicated object and is unlikely observable, since it involves specifying actions at each possible history, both on and *off* the equilibrium path. Our approach only supposes that the econometrician can estimate the frequency of action profiles that *occur in equilibrium* and remains agnostic about which actions are chosen off-equilibrium.

Our paper is related to the literature on equilibrium multiplicity in static games (Tamer, 2003). Conditional on equilibrium continuation payoffs, the game we analyze can be interpreted as a static game with multiple equilibria. Rosen (2006) derives an identified set for the marginal cost parameter in an oligopolistic quantity-setting model. In constructing the identified set, he summarizes multiple equilibria through a one-dimensional, "equilibrium selection" variable whose lowest value represents perfect competition and highest value full collusion. He does not consider dynamic incentives in the model and thus the researcher has to assume a range of values that the equilibrium selection variable can have. Our approach generalizes this idea of using a "pseudo" conduct parameter to dynamic (repeated) games by relying on the extreme points of the equilibrium payoff set to generate bounds.

In dynamic games, methods that rely on Markov perfect Nash equilibrium, e.g., Bajari, Benkard and Levin (2007), typically assume that the same equilibrium is played across markets. We depart from the Markovian assumption and instead assume subgame perfect Nash equilibria and allow different markets to have different equilibria. Using a repeated game lab experiment, Salz and Vespa (2020) estimate the size of counterfactual prediction errors arising from the Markovian assumption. We focus on the (partial) identification of the underlying structural parameters and thus complement their study.

We proceed as follows. In the next section, we use an infinitely repeated, 2×2 Prisoners' Dilemma to illustrate our identification strategy. We then formalize the identification strategy and characterize the identified set in a general setting in Section 3. Section 4 concludes.

2 Example: Prisoners' Dilemma

We illustrate the key idea behind our identification strategy using an infinitely repeated, 2×2 Prisoners' Dilemma. In each period, players choose whether to cooperate (C) or defect (D). Each player then receives stage game payoffs given in Figure 1, where α is a utility parameter to be estimated. For the stage game to be a Prisoners' Dilemma, $\alpha \in (0.5, 2)$.

Players maximize the sum of their present discounted payoffs assuming a common discount factor $\delta \in (0,1)$. The goal of the econometrician is to estimate (α, δ) using data on the frequencies of various action profiles. Finally, we assume that observed actions (or outcomes) are a result of subgame perfect Nash equilibrium (SPNE) behavior.

		P2: C	P2: D
P1	C	α, α	$-1, 2$
	D	$2, -1$	$0, 0$

Figure 1: Prisoners' Dilemma stage game payoffs with $\alpha \in (0.5, 2)$

We are interested in learning the kinds of restrictions SPNE imposes on the distribution of observed action profiles. Given these restrictions, what can we learn about the parameters α and δ? In particular, can we use these restrictions to reduce the identified set to a strict subset of $(0.5, 2) \times (0, 1)$?

By definition, a strategy profile is a SPNE if it prescribes strategies that constitute a Nash equilibrium at every history. Thus, if we observe the action profile (C, C) with positive probability, then there must be some history h and a set of *equilibrium* continuation payoffs $v_{(a_1,a_2)|h}$, such that (C, C) is a Nash equilibrium of the normal form game with the payoff matrix given by Figure 2, where $\Delta_{C|h} \equiv v_{(C,C)|h} - v_{(D,C)|h}$ and $\Delta_{D|h} \equiv v_{(D,D)|h} - v_{(D,C)|h}$.[1]

		P2: C	P2: D
P1	C	$(1-\delta)(\alpha-2) + \delta\Delta_{C|h}, (1-\delta)(\alpha-2) + \delta\Delta_{C|h}$	$(1-\delta)(-1) + \delta\Delta_{D|h}, 0$
	D	$0, (1-\delta)(-1) + \delta\Delta_{D|h}$	$0, 0$

Figure 2: Normal form game using the one-stage deviation principle: Prisoners' Dilemma

A necessary condition for (C, C) to be a Nash equilibrium in the normal form game depicted in Figure 2 is that, given equilibrium continuation strategies, C is a player's best response to the other player choosing C. If we restrict to pure strategies, this necessary

[1] To economize on notation, we assume $v_{(D,C)|h} = v_{(C,D)|h}$.

condition can be written as

$$(1-\delta)(\alpha-2)+\delta\Delta_{C|h}\geq 0. \tag{1}$$

Similarly, a necessary condition for observing (D,D) is that, given equilibrium continuation strategies, D is a player's best response to D, which gives

$$(1-\delta)(-1)+\delta\Delta_{D|h}\leq 0.$$

Finally, a necessary condition for observing (C,D) or (D,C) is that, given equilibrium continuation strategies, C is a best response to D, and vice-versa:

$$(1-\delta)(-1)+\delta\Delta_{D|h}\geq 0.$$

and

$$(1-\delta)(\alpha-2)+\delta\Delta_{C|h}\leq 0$$

To see how these necessary conditions can be used to create restrictions on α, suppose we have $\Pr(C,C)>0$ in the data. If we actually observe the equilibrium continuation payoffs and the discount factor, then from inequality (1), we have

$$\alpha\geq 2-\frac{\delta}{1-\delta}\Delta_{C|h}.$$

However, unless we assume additional equilibrium restrictions such as which types of strategies are selected in equilibrium, we actually do not observe $\Delta_{C|h}$ and hence cannot use the above inequality as is.[2] What we do know is that these equilibrium continuation payoffs belong to some equilibrium payoff set $\mathcal{V}$ which, in turn, is a subset of the set of feasible and individually rational stage game payoffs, $\mathcal{F}^*$. In fact, Folk Theorems show that for a sufficiently large δ, $\mathcal{V}=\mathcal{F}^*$ (Fudenberg and Maskin, 1986).

We now show how to derive restrictions implied by the frequencies of action profiles to construct an identified set for α and δ. Let $\underline{v}(\alpha)$ and $\overline{v}(\alpha)$ be the lower and upper bounds of $\mathcal{F}^*$, which can be functions of the unknown parameter α. In our example, $\overline{v}(\alpha)=\frac{3\alpha}{1+\alpha}$ and $\underline{v}(\alpha)=0$. The shaded region in Figure 3 illustrates $\mathcal{F}^*$ for our Prisoners' Dilemma. Thus, for any history h, $\Delta_{C|h}\in\left[\underline{\Delta}(\alpha),\overline{\Delta}(\alpha)\right]$ where $\underline{\Delta}(\alpha)=\underline{v}(\alpha)-\overline{v}(\alpha)$ and $\overline{\Delta}(\alpha)=\overline{v}(\alpha)-\underline{v}(\alpha)$. The proposition below gives the restrictions for each action profile.

[2] Notice that despite knowing the equilibrium continuation payoffs $(\Delta_{C|h},\Delta_{D|h})$, the model is still *incomplete* in that we cannot write the likelihood for (C,C) similar to static entry models (Tamer, 2003). In this sense, we have a static equilibrium selection problem on top of the dynamic equilibrium selection problem, i.e., knowing the equilibrium continuation payoffs. Our identification strategy deals with both static and dynamic equilibrium selection problems.

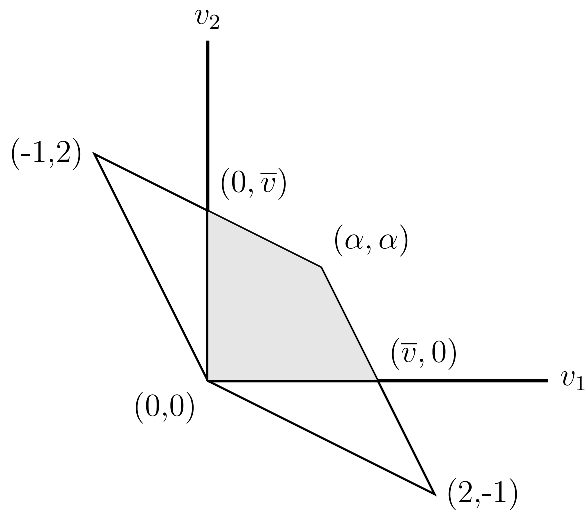

Figure 3: Shaded region: Set of feasible and individually rational payoffs $\mathcal{F}^*$ with $\alpha \in (0.5, 2)$

Proposition 1. In an infinitely repeated Prisoner' Dilemma with the stage game payoffs given in Figure 1 and a common discount factor of $\delta \in (0,1)$, let $\overline{v}(\alpha) = \frac{3\alpha}{1+\alpha}$ and define:

$$\begin{aligned} f(\Delta, \alpha, \delta) &= (\alpha - 2) + \frac{\delta}{1-\delta}\Delta \\ g(\Delta, \alpha, \delta) &= -1 + \frac{\delta}{1-\delta}\Delta. \end{aligned}$$

Observed frequencies of action profiles impose the following restrictions on α and δ:

1. If $\Pr(C,C) > 0$, then $f(\overline{\Delta}(\alpha), \alpha, \delta) \geq 0$.

2. If $\Pr(C,D) > 0$ or $\Pr(D,C) > 0$, then $f(\underline{\Delta}(\alpha), \alpha, \delta) \leq 0$ and $g(\overline{\Delta}(\alpha), \alpha, \delta) \geq 0$.

3. If $\Pr(D,D) > 0$, then $g(\underline{\Delta}(\alpha), \alpha, \delta) \leq 0$.

Proof. Consider observing $\Pr(C,C) > 0$. Then there must be some history h and a set of equilibrium continuation payoffs $v_{\cdot,\cdot|h}$ such that (C,C) is a Nash equilibrium in this history. This implies that C is a best response to C and hence there exists a $\Delta_{C|h} \in [\underline{\Delta}, \overline{\Delta}]$ such that $f(\Delta_{C|h}, \alpha, \delta) \geq 0$. Since $f(\overline{\Delta}, \alpha, \delta) \geq f(\Delta_{C|h}, \alpha, \delta)$, then $\Pr(C,C) > 0$ implies $f(\overline{\Delta}, \alpha, \delta) \geq 0$.

A similar argument can be used to prove the other cases. $\square$

The shaded region in Figure 4 gives the combinations of α and δ that satisfy the restrictions imposed by $\Pr(C,C) = 1$, e.g., if Grim-Trigger strategies are played in equilibrium.

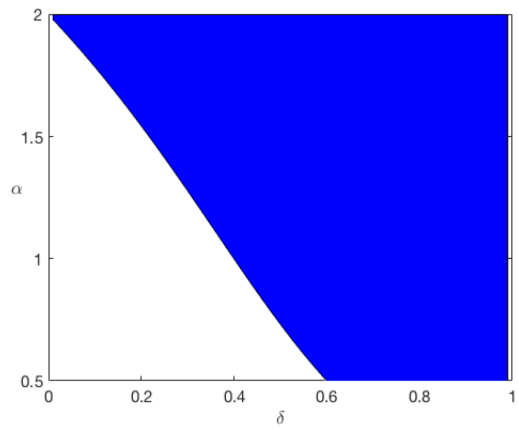

Figure 4: Shaded region: Identified set for α and δ given $\Pr(C, C) = 1$ in infinitely repeated Prisoners' Dilemma

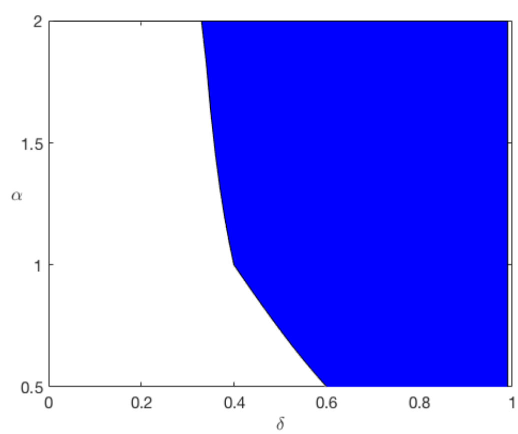

Figure 5: Shaded region: Identified set for α and δ given $\Pr(C, C) > 0$ and $\Pr(C, D) > 0$ in infinitely repeated Prisoners' Dilemma

If δ were known, say $\delta = 0.5$, the restrictions would imply that $\alpha \in [0.7321, 2]$. Otherwise, the restrictions would imply that α and δ must be in the shaded region. Conditional on δ, the restrictions are more informative (i.e., the identified set is a stricter subset of $(0.5, 2)$) the lower δ is. This is because for a low discount factor, both the contemporaneous and continuation payoffs from C (given the other player chooses C) must be sufficiently high for the action profile (C, C) to occur in the data.

As a second example, the shaded region in Figure 5 gives the identified set for α and δ when $\Pr(C, C) > 0$ and $\Pr(C, D) > 0$. Except for low values of δ, the identified set coincides with the identified set under $\Pr(C, C) = 1$. In fact, there are no $\alpha < 2$ and $\delta < \frac{1}{3}$ that would satisfy the restrictions that $\Pr(C, D) > 0$ imposes. Intuitively, for an impatient player receiving a contemporaneous payoff of -1, even the maximum possible continuation payoff, which is less than $\bar{v}(2) = 2$, would yield a total payoff less than 0; the player would be better off choosing D.

We now proceed to generalize these ideas to formalize our identification strategy.

3 Identification

Consider an infinitely repeated game with perfect monitoring. There are N players. In each period (stage game), player i chooses an action $a_i \in A_i$, which can be discrete or continuous, and receives a payoff $\pi_i(a_i, a_{-i}; \alpha)$ where a_{-i} is the vector of actions for all players except i, and α is a vector of utility parameters. Players discount payoffs according to a common discount factor $\delta \in (0, 1)$.

Suppose we observe the action profile $\mathbf{a} = (a_1, a_2, ..., a_N)$ at some history. To rationalize observed actions, we rely on the one-stage deviation principle, which allows us to focus on single-stage deviations to fully characterize players' incentives in choosing specific actions at a given history. Since a subgame perfect Nash equilibrium (SPNE) of the game prescribes strategies that are Nash equilibrium at each possible history, there must be some history h such that for each player i,

$$(1 - \delta)\pi_i(a_i, a_{-i}; \alpha) + \delta v_{(a_i, a_{-i})|h} \geq (1 - \delta)\pi_i(a'_i, a_{-i}; \alpha) + \delta v_{(a'_i, a_{-i})|h} \tag{2}$$

for all $a'_i \in A_i$, and where $v_{(a_i, a_{-i})|h}$ is player i's *equilibrium* continuation payoffs following history h and stage game actions (a_i, a_{-i}).

For each player, the inequalities given by (2) constitute necessary conditions for observing the action profile $\mathbf{a}$ with positive probability. We can rewrite the inequality as

$$(1 - \delta) \left[\pi_i(a_i, a_{-i}; \alpha) - \pi_i(a'_i, a_{-i}; \alpha)\right] + \delta \Delta_{i|h} \geq 0 \tag{3}$$

where $\Delta_{i|h} \equiv v_{(a_i,a_{-i})|h} - v_{(a_i',a_{-i})|h}$ is the vector of player i's unilateral deviation payoff losses. We can then define the set of (α, δ) that satisfies these necessary conditions by $\Psi(\Delta_h)$ where $\Delta_h = (\Delta_{1|h}, \Delta_{2|h}, ..., \Delta_{N|h})$. Note that $\Psi(\Delta_h) \subseteq \Psi(\Delta_h')$ where $\Delta_{i|h}' \geq \Delta_{i|h}$ for all i, because satisfying the conditions in $\Psi(\Delta_h)$ implies satisfying the conditions in $\Psi(\Delta_h')$ in inequalities like (2). Because the *equilibrium* payoff set is invariant with respect to history, we use $\overline{\Delta}_i$ to denote the upper bound of $\Delta_{i|h}$ from the equilibrium payoff set for i. Then, $\Psi(\Delta_h) \subseteq \Psi(\overline{\Delta})$ where $\overline{\Delta} = (\overline{\Delta}_1, \overline{\Delta}_2, ..., \overline{\Delta}_N)$.

We can also derive a set of sufficient conditions for observing $\mathbf{a}$ with positive probability. These sufficient conditions are the incentive compatibility constraints that imply that actions in $\mathbf{a}$ are dominant strategies (conditional on equilibrium continuation strategies). Define the set of (α, δ) that satisfies these sufficient conditions by $\Xi(\Delta_h)$ where $\Xi(\cdot)$ is increasing in the components of Δ_h. Then, $\Xi(\underline{\Delta}) \subseteq \Xi(\Delta_h)$ where $\underline{\Delta} = (\underline{\Delta}_1, \underline{\Delta}_2, ..., \underline{\Delta}_N)$ is the vector of lower bounds from the equilibrium payoff set for all players.

The goal of the econometrician is to estimate (α, δ). We assume that the econometrician has access to data on frequencies of stage game action profiles, i.e. $\Pr(\mathbf{a})$. The data can be generated by observations across markets, across time for a given player or multiple players, or a combination of both. In the Prisoners' Dilemma example in Section 2, equilibrium selection is the only source of variation driving equilibrium behavior. For example, the econometrician might observe $0 < \Pr(C, C) < 1$ which implies that in some markets, an equilibrium with less than full cooperation was selected. In constructing the identified set in the Prisoners' Dilemma example, we essentially exploited the upper and lower bounds of the support of the distribution representing equilibrium selection, i.e. a distribution over equilibrium payoffs. Because we do not have other sources of variation nor additional information about equilibrium selection, we can only use the fact that either $\Pr(\mathbf{a}) > 0$ or $\Pr(\mathbf{a}) = 0$ to generate restrictions on (α, δ); the exact value of $\Pr(\mathbf{a})$ was not informative about (α, δ). In actual empirical applications, there is likely additional variation that influences observed equilibrium behavior as data are pooled across markets, periods and players. Our identification strategy in the general setting exploits this additional variation.

We assume that additional variation is captured by vectors of observables x and unobservables ϵ, with the joint cumulative distribution function $F(x, \epsilon)$. The equilibrium payoff set and the necessary and sufficient conditions Ψ and Ξ now depend on (x, ϵ).[3] We now derive the restriction on $\Pr(\mathbf{a})$. Since observing $\mathbf{a}$ implies that the set of necessary conditions that define $\Psi(\overline{\Delta})$ holds, then

$$\Pr(\mathbf{a}) \leq \int \Psi(\overline{\Delta}_{(x,\epsilon)}; x, \epsilon) \mathrm{d}F(x, \epsilon).$$

[3] In practice, one needs to compute the upper and lower bounds of Δ_h for each realization of (x, ϵ).

where we have made explicit the dependence on (x, ϵ). Similarly, since conditions that define $\Psi(\underline{\Delta})$ are sufficient for observing $\mathbf{a}$, then

$$\int \Xi(\underline{\Delta}_{(x,\epsilon)}; x, \epsilon) \mathrm{d}F(x, \epsilon) \leq \Pr(\mathbf{a}).$$

Thus, we can use these "worse case bounds" (Manski, 1990) to obtain an identified set for (α, δ).

We end this section with a proposition that summarizes our identification results.

Proposition 2. Let

$$\Pr(\Psi_{\mathbf{a}}) \equiv \int \Psi(\overline{\Delta}_{(x,\epsilon)}; x, \epsilon) \mathrm{d}F(x, \epsilon)$$

and

$$\Pr(\Xi_{\mathbf{a}}) \equiv \int \Xi(\underline{\Delta}_{(x,\epsilon)}; x, \epsilon) \mathrm{d}F(x, \epsilon)$$

where the necessary and sufficient conditions Ψ and Ξ are tailored to each action profile $\mathbf{a}$. Suppose we have data on the frequencies of a set of action profiles $\mathcal{A}$. Then

$$\mathcal{H} = \bigcap_{\mathbf{a} \in \mathcal{A}} \{(\alpha, \delta) : \Pr(\mathbf{a}) \in [\Pr(\Xi_{\mathbf{a}}), \Pr(\Psi_{\mathbf{a}})]\}$$

is an identified set for (α, δ).

4 Concluding remarks

In this note, we propose a strategy to identify the structural parameters of an infinitely repeated game without relying on equilibrium selection assumptions beyond subgame perfect Nash equilibrium. The strategy exploits the extreme points of the equilibrium payoff set to derive bounds on the frequency of actions. These bounds in turn provide restrictions on the parameters of interest.

The identified set that we have constructed is not sharp. If one can compute the exact equilibrium payoff set, then one can construct a sharp identified set by examining all the restrictions for each possible equilibrium payoff. The repeated games literature characterizes the equilibrium payoff set at given discount factors and provides algorithms to compute this set (e.g., Abreu et al. (1990); Judd et al. (2003); Abreu and Sannikov (2014); Abreu et al. (2016)). Although computing all the elements of this set is likely computationally intensive or even infeasible in most empirical settings, our approach only relies on the extreme points of this set. To what extent restrictions from our identification strategy yield an informative identified set will depend on the actual empirical application and the available data. One important application of our identification strategy is the flexible estimation of firm conduct to detect collusion.

Finally, empirical settings involve both observed and unobserved state variables. Including observed exogenous state variables is theoretically straightforward (and in fact useful for identification as discussed in Section 3) but can be computationally demanding since this involves computing the (extreme points) of the equilibrium payoff set for each value of the state variable. Adding unobserved exogenous shocks is also feasible (but may also be computationally demanding) as long as one is willing to assume that these shocks come from a known distribution. Given the ability to accommodate states, it will be interesting to investigate how to adapt our identification strategy to general dynamic games.

References

Abreu, D., B. Brooks and Y. Sannikov (2016), "A Pencil Sharpening Algorithm for Two Player Stochastic Games with Perfect Monitoring," Stanford GSB Working Paper 3428.

Abreu, D., D. Pearce, and E. Stacchetti (1990), "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica*, 58:5, 1041-1063.

Abreu, D. and Y. Sannikov (2014), "An Algorithm for Two-Player Repeated Games with Perfect Monitoring," *Theoretical Economics*, 9:2, 313-338.

Bajari, P., C. L. Benkard and J. Levin (2007), "Estimating Dynamic Models of Imperfect Competition," *Econometrica*, 75:5, 1331-1370.

Dal Bó, P. and G. R. Fréchette (2011), "The Evolution of Cooperation in Infinitely Repeated Games: Experimental Evidence," *American Economic Review*, 101:1, 411-429.

De Paula, Á. (2013), "Econometric Analysis of Games with Multiple Equilibria", *Annual Review of Economics*, 5, 107-131.

Fudenberg, D. and E. Maskin (1986), "The Folk Thorem in Repeated Games with Discounting or with Incomplete Information," *Econometrica*, 54:3, 533-554.

Judd, K. L., Yeltekin, S. and Conklin, J. (2003), "Computing Supergame Equilibria," *Econometrica*, 71:4, 1239-1254.

Lee, B. S. and C. Stewart (2016), "Identification of payoffs in repeated games," *Games and Economic Behavior*, 99, 82-88.

Manski, C. F. (1990), "Nonparametric Bounds on Treatment Effects," *American Economic Review Papers & Proceedings*, 80:2, 319-323.

Rosen, A. M. (2006), "Identification and Estimation of Firms' Marginal Cost Functions with Incomplete Knowledge of Strategic Behavior," CeMMAP working paper CWP03/07.

Salz, T. and E. Vespa (2020), "Estimating Dynamic Games of Oligopolistic Competition: An Experimental Investigation," *RAND Journal of Economics*, 51:2, 447-469.

Tamer, E. (2003), "Incomplete Simultaneous Discrete Response Model with Multiple Equilibria," *Review of Economic Studies*, 70, 147-165.