Supplementary Information for: "Revision Games"

Yuichiro Kamada[†] Michihiro Kandori[‡]

November 21, 2019

S1 Additional Discussion for Example 2 (Cournot Duopoly)

First, we offer a story to motivate this example. Two fishing boats depart from a harbor, and they must return when the fish market, located at the harbor, opens at 6:00am. They catch fish at a fishing ground that is far from the fish market. They have to depart the fishing ground at 5:00am to reach the market at 6:00am. Hence, 5:00am is the end of the revision game, and assume that the cost for the fishermen is associated with the transportation cost of the final catch (the fish at hand at 5:00am) to the market (consider the case where catching fish itself is easy and the cost of catching fish is negligible).

The fishermen wish to collude (i.e., to restrict their catch) so as to increase the price at the fish market. They start with a small amount of catch (the collusive quantity). They operate side by side, closely monitoring each other's behavior. A revision opportunity of their quantities corresponds to the arrival of a fish school, which follows a Poisson process.[1] When the Poisson arrival rate is $\lambda = 0.1$ and the time unit is a minute, a fish school comes every ten minutes on average. According to the optimal trigger strategy equilibrium, the fishermen do not touch fish schools until 4:49am. In the last eleven minutes, however, whenever a fish school visits them, they catch additional fish. If anyone deviates from this equilibrium plan, they catch a large amount (so that each fisherman's total amount becomes the Nash quantity) when the next fish school arrives. In this way, the fishermen can attain an expected payoff that is 97% of fully collusive profit in expectation as we will show below.

[†]Haas School of Business, University of California Berkeley, 2220 Piedmont Avenue, Berkeley, CA 94720-1900, USA, e-mail: `y.cam.24@gmail.com`

[‡]Faculty of Economics, University of Tokyo, e-mail: `kandori@e.u-tokyo.ac.jp`

[1]Pun not intended.

This story does not exactly match the revision-game model in that (i) the fishermen may be able to discard fish at any time in reality and (ii) catching additional fish may be infeasible in reality if the deadline is too close. We can formally show that introducing the possibility of discarding fish does not alter our equilibrium. The main reason is that the fishermen would like to catch more, if possible, under any equilibrium action profile of the optimal trigger strategy before time 0. The formal proof is given at the end of this section. For the latter point (ii), the issue can be addressed by the argument we present in Section 7.1.

Next, we calculate the expected payoff from the optimal trigger strategy equilibrium. It is computed as:

$$\begin{aligned}
&\int_0^{t(x^*)} (a - 2bx(t) - c)\, x(t) \lambda e^{-\lambda t} dt + e^{-\lambda t(x^*)} (a - 2bx^* - c)\, x^* \\
&= \int_0^{t(x^*)} \left(a - 2b\frac{a-c}{3b}\left(5 - 4e^{\frac{\lambda}{18}t}\right) - c \right) \frac{a-c}{3b}\left(5 - 4e^{\frac{\lambda}{18}t}\right) \lambda e^{-\lambda t} dt + e^{-\lambda t(x^*)} \left(a - 2b\frac{a-c}{4b} - c \right) \frac{a-c}{4b} \\
&= \frac{(a-c)^2}{9b} \int_0^{t(x^*)} \left(-7 + 8e^{\frac{\lambda}{18}t}\right)\left(5 - 4e^{\frac{\lambda}{18}t}\right) \lambda e^{-\lambda t} dt + e^{-\lambda t(x^*)} \frac{(a-c)^2}{8b} \\
&= \frac{(a-c)^2}{9b} \left[35e^{-\lambda t} - 72e^{-\frac{17}{18}\lambda t} + 36e^{-\frac{8}{9}\lambda t} \right]_0^{t(x^*)} + e^{-\lambda t(x^*)} \frac{(a-c)^2}{8b} \\
&= \frac{(a-c)^2}{9b} \left(35\left(\frac{17}{16}\right)^{-18} - 72\left(\frac{17}{16}\right)^{-17} + 36\left(\frac{17}{16}\right)^{-16} + 1 \right) + \left(\frac{17}{16}\right)^{-18} \frac{(a-c)^2}{8b} \\
&= \frac{(a-c)^2}{8b} \left(\left(\frac{17}{16}\right)^{-18} \left(35 \cdot \frac{8}{9} - 64 \cdot \frac{17}{16} + 32\left(\frac{17}{16}\right)^2 + 1 \right) + \frac{8}{9} \right) \\
&= \frac{(a-c)^2}{8b} \left(\left(\frac{17}{16}\right)^{-18} \left(-\frac{323}{9} + 32\left(\frac{17}{16}\right)^2 \right) + \frac{8}{9} \right).
\end{aligned}$$

The collusive payoff is:

$$\left(a - 2b\frac{a-c}{4b} - c \right) \frac{a-c}{4b} = \frac{(a-c)^2}{8b}.$$

Thus, the ratio of the former to the latter is:

$$\left(\frac{17}{16}\right)^{-18} \left(-\frac{323}{9} + 32\left(\frac{17}{16}\right)^2 \right) + \frac{8}{9} \simeq 0.96817\ldots$$

Model that allows discarding fish:

Assume that players can discard fish at any moment of time in $\{-t| -t = 0, -\Delta, -2\Delta, ...$ and $t < T\}$ for small $\Delta > 0$.[2] Consider the strategy profile of this game where (1) actions at the revision opportunities are the same as under the optimal trigger strategy of our original model, (2) players do not discard fish after any history before time 0, and (3) they take mutual best replies when they may discard fish at time 0.[3]

Let us elaborate on (3). When the profile of fish at hand is $(\overline{q}_1, \overline{q}_2)$ at time 0 (and a revision opportunity does not arise at time 0), players effectively play the Cournot game where their quantities are restricted to satisfy $q_i \leq \overline{q}_i$, $i = 1, 2$. Examination of the reaction curves in the restricted strategy spaces reveals that the unique Nash equilibrium of this fish discarding game at time 0 is (i) the Cournot Nash profile (q_1^N, q_2^N) if $q_i^N \leq \overline{q}_i$, $i = 1, 2$, (ii) $(\overline{q}_1, \overline{q}_2)$ if $R_i(q_{-i}) \geq \overline{q}_i$, $i = 1, 2$, where R_i is the reaction function of player i (they do not discard fish because they would like to catch more, if possible), and (iii) $(\overline{q}_i, R_{-i}(\overline{q}_i))$ if $q_i^N \geq \overline{q}_i$ and $R_{-i}(\overline{q}_i) \leq \overline{q}_{-i}$, $i = 1, 2$.

Given the above observation, one can check that if a player unilaterally discards fish at any moment before time 0 (after any history), her payoff at the fish discarding game at time 0 never increases. Hence, players have an incentive to follow the fish discarding rule specified by our strategies. Lastly, we show that players have an incentive to follow the revision rule specified by our strategies. When a revision opportunity arrives at $-t$, the equilibrium action $q(t)$ is no greater than the Cournot Nash quantity. Consider any unilateral deviation of player i at time t. If the deviation action is less than myopic best reply $R_i(q(t))$, it remains unchanged at the fish discarding game at time 0 (see case (ii) specified above). Alternatively, when player i deviates to a larger amount than the myopic best reply, the deviating action is later corrected towards the myopic best reply (case (iii) specified above). In either case, the player has no incentive to deviate because the trigger strategy profile constitutes an equilibrium in our original model (because the deviation to the myopic best reply is unprofitable under the optimal trigger strategy equilibrium profile in our model).

If a revision opportunity arrives (before time 0) when players are supposed to play the Cournot Nash equilibrium quantity, any unilateral deviation also remains to be

[2] This is to avoid technical issues associated with defining strategies in continuous time.

[3] In particular, actions in the revision opportunities depend only on what has happened on the previous revision opportunities (and the initial actions), as under the optimal trigger strategy.

unprofitable. This is because after any unilateral deviation, the opponent does not change his Nash action in the fish discarding game at time 0. Hence, the strategy profile we constructed under the possibility of discarding fish is an equilibrium, and by construction it achieves the same outcome as the optimal trigger strategy equilibrium of our original model.

S2 Discussions for Section 4.2

S2.1 A Microfoundation of the Voter Behavior

The winning probability (9) is derived from the following behavior of voters. There is a unit mass of voters $i \in [0,1]$, where fraction θ is pro-THAAD and the others are anti-THAAD. We assume that the pro-THAAD voters constitute a majority, i.e.,

$$\theta \in (1/2, 1]. \tag{38}$$

Pro-THAAD voter i votes for A if

$$x_A + \widetilde{\eta}_A + \widetilde{\varepsilon}_{iA} > x_B + \widetilde{\eta}_B + \widetilde{\varepsilon}_{iB},$$

where $\widetilde{\eta}_A$ and $\widetilde{\eta}_B$ are candidate-specific shocks that are common to all voters, and $\widetilde{\varepsilon}_{iA}$ and $\widetilde{\varepsilon}_{iB}$ are voter-specific idiosyncratic shocks (there is no abstention). Pro-THAAD voter i votes for B under the symmetric condition.[4] Similarly, anti-THAAD voter i votes for A if

$$1 - x_A + \widetilde{\eta}_A + \widetilde{\varepsilon}_{iA} > 1 - x_B + \widetilde{\eta}_B + \widetilde{\varepsilon}_{iB},$$

and anti-THAAD voter i votes for B under the symmetric condition. We assume uniform distributions: $\widetilde{\eta}_B - \widetilde{\eta}_A \sim U[-c, c]$ and $\widetilde{\varepsilon}_{iB} - \widetilde{\varepsilon}_{iB} \sim U[-h, h]$. The latter is interpreted as "i.i.d." shocks across voters and we adopt the convention that the law of large numbers holds in the continuum population. We also assume that the voter-specific shocks are more important than the common shocks in the following sense:

$$h \geq 1 + c. \tag{39}$$

[4] The behavior of the voter i such that $x_A + \widetilde{\eta}_A + \widetilde{\varepsilon}_{iA} = x_B + \widetilde{\eta}_B + \widetilde{\varepsilon}_{iB}$ does not affect the analysis because the measure of such voters is zero.. The same remark applies to the anti-THAAD voter i satisfying an analogous equality.

In the population of pro-THAAD voters, the fraction of voters voting for A given η_A and η_B is equal to

$$\begin{aligned}\Pr(x_A + \eta_A + \widetilde{\varepsilon}_{iA} > x_B + \eta_B + \widetilde{\varepsilon}_{iB}) &= \Pr(x_A - x_B + \eta_A - \eta_B > \widetilde{\varepsilon}_{iB} - \widetilde{\varepsilon}_{iA}) \\ &= \frac{1}{2} + \frac{x_A - x_B + \eta_A - \eta_B}{2h}.\end{aligned}$$

Here we used our assumption that $\widetilde{\varepsilon}_{iB} - \widetilde{\varepsilon}_{iB} \sim U[-h, h]$ and (39), which ensures $-h \leq x_A - x_B + \eta_A - \eta_B \leq h$. The symmetric expression applies to the anti-THAAD voters, and therefore the total mass of A-voters given common shocks η_A and η_B is

$$\theta\left(\frac{1}{2} + \frac{x_A - x_B + \eta_A - \eta_B}{2h}\right) + (1-\theta)\left(\frac{1}{2} + \frac{(1-x_A) - (1-x_B) + \eta_A - \eta_B}{2h}\right)$$

$$= \frac{1}{2} + \frac{(2\theta - 1)(x_A - x_B) + \eta_A - \eta_B}{2h}.$$

Candidate A's winning probability, denoted $P_A(x_A, x_B)$, is the probability that the above fraction is more than 1/2. This reduces to

$$P_A(x_A, x_B) = \Pr((2\theta - 1)(x_A - x_B) > \widetilde{\eta}_B - \widetilde{\eta}_A). \tag{40}$$

Recall that $\widetilde{\eta}_B - \widetilde{\eta}_A \sim U[-c, c]$ and assume that the support is large enough so that

$$2\theta - 1 \leq c. \tag{41}$$

Then, (40) is equal to

$$P_A(x_A, x_B) = \frac{1}{2} + \frac{(2\theta - 1)(x_A - x_B)}{2c}.$$

By defining $\delta := (2\theta - 1)/c$, we have obtained the proposed winning probability (9). Parameter restrictions (38) and (41) imply that $\delta \in (0, 1]$.

S2.2 Catering to the middle

If $\gamma = 0$ and $\delta = \frac{1}{2}$, the model can also be interpreted in the following manner: Suppose that the policy space is $[-1, 1]$. Candidate 1's bliss point is -1 and candidate 2's is 1. Candidate 1 is faithful to her party's identity ("left wing"), so she chooses a policy from $[-1, 0]$ while never wanting to choose a "right wing" policy in $(0, 1]$ possibly because of reputational concern. Symmetrically, candidate 2 chooses a policy from $[0, 1]$.[5] Each candidate's winning probability is determined by the standard Hotelling rule where the median voter is uniformly distributed over $[-1, 1]$, and a candidate's utility from the implemented policy is given by $\max\{1-z, 0\}$ where z is the distance from the bliss point to the implemented policy. Then, each candidate's payoff from any given policy profile $(y_A, y_B) \in [-1, 1]^2$ is given by (10) with substitutions that $x_A = 1 + y_A$ and $x_B = 1 - y_B$. Hence, our model nests such a model as a special case.

We note that, in this alternative model, the utility function for the implemented policy is convex. Such policy preferences are especially relevant for issues that provoke strongly opinionated reactions (e.g. same-sex marriage, abortion, gun control, and so forth). This is because, for these policy issues, it is natural to assume that one's utility arising from policy preferences sharply decreases as the implemented policy moves away from her bliss point.[6] The functional form $\max\{1 - z, 0\}$ is the simplest way to capture this possibility. This convexity allows for there to be a potential room for cooperation in the revision game as long as the Nash profile is not $(-1, 1)$.[7]

With this reformulation of the model, Proposition 3 shows that if the office motivation is not too large or too small, each candidate starts from announcing their most preferred policies (policy -1 for candidate 1 and policy 1 for candidate 2). They stick to such announcements until a certain time before the election day, and then begin catering to the middle towards the end of the campaign period.

[5] Technically, if each candidate can choose a policy from $[-1, 1]$, then there does not exist a pure Nash equilibrium in the component game because a best response does not necessarily exist. We conjecture that, if we allowed for candidates to choose their policies from $[-1, 1]$, there would exist a nontrivial equilibrium in which at each opportunity candidates mix across multiple policies. Kamada and Sugaya (2019) analyze a mixed strategy plan in the context of their model, and their simulation result shows quite complicated dynamics of mixing probabilities. For this reason, here we do not delve into such an analysis.

[6] See Osborne (1995) for a criticism on the use of concave utility functions for preferences over electoral policies. Kamada and Kojima (2014) discuss implications of convex voter utility functions.

[7] Note that, on the other hand, there would be no room for cooperation if the utility is concave as traditionally assumed in the political science literature.

S3 Appendix for Section 7

S3.1 Details for Section 7.1

S3.1.1 Soft Deadline

The precise statements of the claims about introducing a soft deadline in Section 7.1 are as follows:

Proposition 8. *The following hold for any $T \geq 0$ and $\gamma \in (0, 1]$ in the soft deadline game.*

1. *Under (13), for any $\epsilon > 0$, cooperation can be sustained by a trigger strategy equilibrium.*

2. *Under (14), there exists $\bar{\epsilon} < 1$ such that for all $\epsilon < \bar{\epsilon}$, there is a unique trigger strategy equilibrium, and it has the plan $x(t) = a^N$ for each t.*

Proof. For any trigger strategy equilibrium plan $x(\cdot)$, define action a^x by an arbitrary action in A that satisfies the following equation:

$$\pi(a^x) = \sup_{t \geq 0} \pi(x(t)). \tag{42}$$

This is well defined because the action space is compact and π is continuous. Plan $x(\cdot)$ satisfies the following incentive constraint at any time $t \geq 0$,

$$\pi(x(t))P + \sum_{s=1}^{\infty} \pi(x(t+s))R_s \geq (\pi(x(t)) + d(x(t))\, P + \pi^N(1-P), \tag{43}$$

where P is the probability of no revision in the future and R_s is the probability that the last revision opportunity arrives s periods later. By definition, $\sum_{s=1}^{\infty} R_s = 1 - P$, and therefore (42) implies that the left-hand side of (43) is less than or equal to $\pi(a^x)$. By considering the right-hand side of (43) for a convergent subsequence $t = t_1, t_2, ...$ where $\lim_{n\to\infty} \pi(x(t_n)) = \pi(a^x)$ and $\lim_{n\to\infty} d(x(t_n)) = d(a^x)$ (both convergence can simultaneously hold because d is continuous as well), we obtain

$$\pi(a^x) \geq (\pi(a^x) + d(a^x))\, P + \pi^N(1-P), \tag{44}$$

which can be arranged as

$$\frac{d(a^x)}{\pi(a^x) - \pi^N} \leq \frac{1}{P} - 1 \tag{45}$$

if $\pi(a^x) > \pi^N$. This is a necessary condition for any $x(\cdot)$ to be sustained by a trigger strategy equilibrium. Conversely, if this condition is satisfied and $x(\cdot)$ is a constant plan $\forall t \geq 0\ x(t) = a$, then $x(\cdot)$ constitutes an equilibrium in the subgame after period 0.[8]

Now consider condition (13). For any $\epsilon > 0$, there is a positive probability of a revision in the future, and therefore P, the probability of no revision in the future, is strictly less than 1, so that $\frac{1}{P} - 1 > 0$. Note that condition (13) implies $\lim_{a \to a^N} \frac{d(a)}{\pi(a) - \pi^N} = 0$. Hence, under (13), there exists $\varepsilon > 0$ such that for any $x(\cdot)$ such that $a^x \neq a^N$ and a^x is in the ε-neighborhood of a^N, condition (45) holds. Note that such $x(\cdot)$ exists because there is $a \in A \setminus \{a^N\}$ in the ε-neighborhood of a^N such that $\pi(a) > \pi^N$ by assumption. Hence, we conclude that cooperation can be sustained from time 0 on (this also implies that (weakly more) cooperation can be sustained in all periods before).

Next, consider condition (14). Note that P, the probability of no revision in the future, is equal to 1 if $\epsilon = 0$ (i.e., the game ends immediately), and P is continuous in ϵ.[9] Therefore, there exists $\bar{\epsilon} > 0$ such that for all $\epsilon < \bar{\epsilon}$, the right-hand side of (45) is strictly less than the infimum in condition (14), which means that (45) is violated for any $x(\cdot)$ such that $\pi(a^x) > \pi^N$. Hence, in any trigger strategy equilibrium, players play a^N from period 0 on. By backward induction, a^N is also played in any period. $\square$

S3.1.2 ε-Willingness to Punish

Recall that we consider a model in which time is discrete $-t = -T, ..., -1, 0$, and a revision opportunity arrives with probability $\gamma = 0.1$ in each period. In the continuous time model with Poisson arrival rate λ, a revision opportunity arrives approximately with probability $\lambda\Delta$ in a small time interval Δ. To compare the discrete- and continuous-time models, we assume $\lambda = 1$ and the length of one period in the discrete time models is $\Delta = 0.1$ (so that $\gamma = \lambda\Delta = 0.1$).

In the continuous time models with selfish players ($\varepsilon = 0$), the optimal plan in

[8] This is because (44) is the incentive constraint for the constant plan.

[9] This can be verified by $P = \sum_{t=0}^{\infty}(1-\epsilon)(1-\gamma)^t\epsilon^t$.

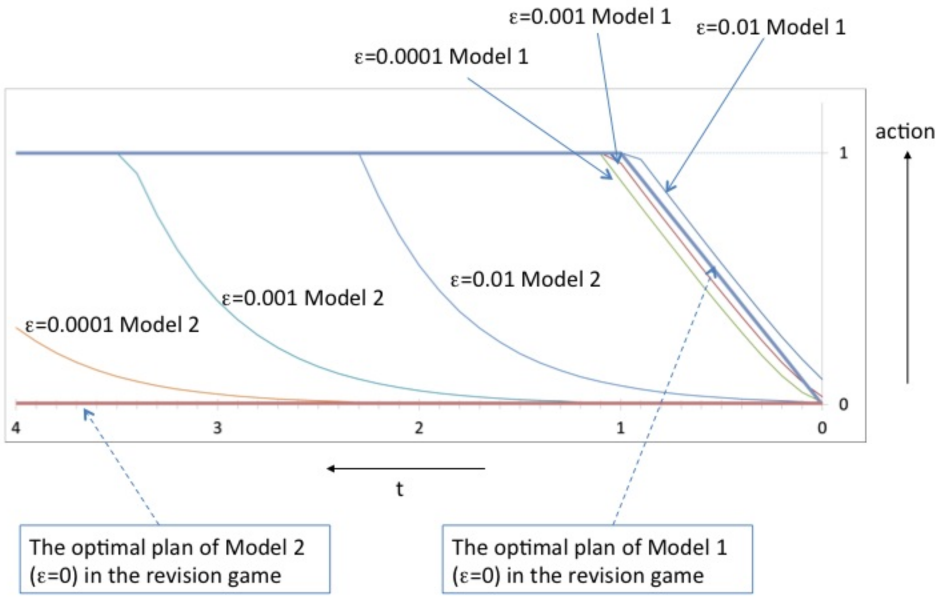

Figure 5: Finite Time Condition

Model 1 turns out to be $x(t) = t$ until it hits the optimal action 1.[10] In Model 2, the optimal plan exhibits no cooperation: $x(t) = 0$ for all t. Those predictions of the revision games are depicted by the thick curves in Figure 5.[11]

The discrete time models with ε-incentive to punish can be numerically solved backwards, and the solutions are depicted by thin curves in the figure (the discrete points are interpolated). As we can see, substantial cooperation is sustained even under very small incentive to punish (ε) in Model 1. In contrast, in Model 2, cooperation becomes very hard to sustain as the incentive to punish decreases.

Let us summarize the implications of our analysis above. In realistic situations, the assumption of the revision game that there is always a positive probability that another revision is possible may not be satisfied. In such a situation, however, if players have small incentive to punish a deviator, the predictions of the revision game survive: substantial cooperation is possible if and only if cooperation is sustained in the revision game.

[10]This is the solution to the differential equation $\frac{dx}{dt} = \lambda\frac{d(x)+\pi(x)-\pi^N}{d'(x)} = \frac{x^2+(2x-x^2)}{2x} = 1$.

[11]In the figure, t corresponds to the remaining time in the continuous-time models (revision games).

S3.2 Detailed Discussion for Section 7.2

Revision games have some similarities to stochastic games (e.g., Dutta (1995) and Hörner, Sugaya, Takahashi and Vieille (2011)) and to repeated games with a decreasing discount factor (Bernheim and Dasgupta, 1995). We compare those models in the framework of a general stochastic-game model, where players obtain flow payoffs in *each* period $n = 0, 1, 2, ...$ Although a payoff accrues only once (at the deadline) in the revision game, it can be interpreted as a special case of this general stochastic game. This observation facilitates a clear-cut comparison.

In the general stochastic-game model, the flow playoff to player i in period $n = 0, 1, ...$ is

$$u_i(a(n), s(n))$$

where $a(n)$ and $s(n) \in S$ are the action profile and state in period n, respectively. The probability distribution of $s(n+1)$ depends on $s(n)$ and $a(n)$. The game is terminated at the end of period n with probability $r(s(n), a(n))$. The flow payoff in period n is discounted by $D(s(n), n)$. Therefore, if the game is terminated at the end of period N, the realized payoff is

$$\sum_{n=0}^{N} u_i(a(n), s(n)) D(s(n), n).$$

Dynamic games that have some similarities to revision games are special cases of this general model.

- The first stochastic-game model proposed by Shapley (1953): $D(s(n), n) \equiv 1$ (discounting comes from random termination).
- Stochastic games where the folk theorems have been proved (Dutta (1995) and Hörner, Sugaya, Takahashi and Vieille (2011)): $D(s(n), n) = \delta^n$ and $r(s(n), a(n)) \equiv 0$ (no random termination).
- Repeated games with a decreasing discount factor (Bernheim and Dasgupta, 1995): S is a singleton (there is no state variable), $D(n) = \delta(0) \times \cdots \times \delta(n)$ ($\lim_{n \to \infty} \delta(n) = 0$) and $r(a(n)) \equiv 0$ (no random termination).

We will show that the revision game is also strategically equivalent to a special case of this model, which we call "Game F."[12] To transform the revision game into Game F, first interpret the beginning of the revision game as period $n = 0$, and interpret the n^{th} arrival of a Poisson revision opportunity as period $n > 0$. Define $s(n)$, the state in period n, as the remaining time in the revision game (state space is $S = [0, T]$), and define the termination probability at a given remaining time (=state) by the probability of no future revision opportunity at such a remaining time in the revision game.[13] Each player i obtains in each period n discounted flow payoff

$$\pi_i(a(n))e^{-\lambda s(n)}.$$

Let us now show that the revision game and Game F share the same expected payoff. We present such an analysis for the case in which the current action depends only on the current state (the remaining time).[14] In general, if payoff $f(t)$ accrues each time a Poisson arrival happens, the expected sum of realized payoffs is

$$\int_0^T f(t)\lambda dt.$$

Heuristically, this is true because the probability of one Poisson arrival in a small time interval Δ is approximately equal to $\lambda\Delta$. If we apply this formula for $f(t) = \pi_i(a(t))e^{-\lambda t}$ (Game F's payoff) where $a(t)$ denotes the action taken at state t, we can see that the expected payoff in Game F is

$$\pi_i(a(T))e^{-\lambda T} + \int_0^T \pi_i(a(t))e^{-\lambda t}\lambda dt, \tag{46}$$

which is exactly the same as that in the revision game.

The equivalence between the revision game and Game F enables us to directly compare the revision game with related dynamic games. First, the mechanism to sustain cooperation in all those models are the same: the deviation gain today is wiped out by the reduction of the future payoff. Secondly, discounting of the future payoff is obtained by random termination of the game. This is the sole mechanism

[12] The "F" stands for flow playoffs.

[13] This provides the following state transition and termination probability. The initial state is $s(0) = T$. Given $s(n)$, the probability of $s(n+1) \in [s(n), T]$ is zero, the density of $s(n+1) \in [0, s(n))$ is $e^{-\lambda(s(n+1)-s(n))}\lambda$, and the probability of termination of the game is $e^{-\lambda s(n)}$.

[14] A more general case is given in Appendix S3.2.1.

of discounting in Shapley (1953), and it also partially accounts for the discounting in Game F ($\simeq$ the revision game). Milnor and Shapley (1957) provided a stochastic game model of gamblers' ruin where a payoff accrues only when the game is terminated (i.e., one of the players is ruined). The timing of payoff realization of this game is similar to that of the revision game in that a payoff accrues only once.[15]

Thirdly, the revision game ($\simeq$ Game F) is not a special case of the stochastic games where the folk theorems have been proved, such as Dutta (1995) and Hörner, Sugaya, Takahashi and Vieille (2011). In the latter, it is crucial that the state transition is *irreducible*: any state s is reachable from any state s' in a finite number of periods with probability one. This is clearly not the case in Game F, where the state, which corresponds to the remaining time in the revision game, never increases. Hence, the revision game belongs to the class of stochastic games (non-irreducible ones) for which the set of equilibria has not been fully characterized.

Fourthly, there is a similarity and difference between the revision game and the repeated-game model with a decreasing discount factor (Bernheim and Dasgupta, 1995). When the current state is $s \in [0, T]$ in Game F, by the same argument to derive formula (46), the continuation playoff is

$$\int_0^s \pi_i(a(t))e^{-\lambda t}\lambda dt. \tag{47}$$

In contrast, the continuation payoff in the Bernheim-Dasgupta model in period m is

$$\sum_{n=m+1}^{\infty} \pi_i(a(n)) \prod_{k=m+1}^{n} \delta(k). \tag{48}$$

Those formulae clarify the similarity and difference between Game F ($\simeq$ the revision game) and the Bernheim-Dasgupta model.

- In both models, the magnitude of continuation payoff decreases as time passes by: both (47) and (48) tend to zero as $s \to 0$ (s corresponds to the remaining time in the revision game) and $m \to \infty$ (because $\lim_{k\to\infty} \delta(k) = 0$) if $|\pi(a)|$ is

[15]The discussion paper version of this paper (Kamada and Kandori, 2017), Kamada and Kandori (2019), and Sherstyuk, Tarui and Saijo (2013) observe that the discounted repeated game is strategically equivalent to the game where (i) players prepare actions of the stage game in each period $t = 0, 1, ..,$ (ii) the game is terminated at the end of each period with some probability $r > 0$, and (iii) players obtain the stage game payoff in the last period only. Kamada and Kandori (2017) call this game the stationary revision game.

uniformly bounded.

- In the continuation payoffs, smaller weights are attached to the future flow payoffs in the Bernheim-Dasgupta model, while the opposite is true in Game F ($\simeq$ the revision game).

To see the latter point, note that, in Game F, the discounted flow payoff at state s (= the remaining time in the revision game) is

$$\pi_i(a(s))e^{-\lambda s},$$

whose coefficient of discounting $e^{-\lambda s}$ *increases* as time passes by (as s decreases). This fact translates into different efficiency properties of those games: the optimal trigger strategy equilibrium attains approximate efficiency in the Bernheim-Dasgupta model, while this is not true in the revision game.[16] In summary, the revision game has some similarities to the Bernheim-Dasgupta model, but the former is not a reformulation of the latter in a continuous time/state framework.

S3.2.1 The Non-Markov Case

We formally show that the revision game and Game F share the same strategy spaces and payoffs. At remaining time/state s, players in those games observe a sequence of past events and the current state

$$((T, a(T)), (s(1), a(1)), ..., (s(n), a(n)), s),$$

where $T > s(1) > \cdots > s(n) > s$ and n is an arbitrary positive integer. This records that players were called upon to move when the remaining time/state was $T, s(1), \ldots, s(n)$, and it also shows what action profiles were chosen in those occasions. Let H_s be the set of all those histories at s. A (pure) strategy of player i, denoted σ_i, both in the revision game and in Game F, is a mapping from histories to current actions $\sigma_i : \bigcup_{s=0}^{T} H_s \to A$, where H_T is a singleton set of the initial dummy history to determine the initial action.

A strategy profile σ and Poisson arrival rate λ determines the probability measure $P_s^{\sigma,\lambda}$ over the set of possible histories at s, H_s. Since a Poisson arrival at time $-s$ is

[16] The action distribution induced by the optimal trigger strategy equilibrium (Proposition 1) fails to attach probability close to 1 to the efficient action.

independent of the past history, $P_s^{\sigma,\lambda}$ is also equal to the probability measure of the past history conditional on a Poisson arrival at time $-s$. Hence, the expected payoff associated with the action profile chosen at s is expressed as

$$\int_{H_s} \pi_i(\sigma(h))dP_s^{\sigma,\lambda}(h) =: \pi_i(s).$$

Given this definition, by the same argument as in Appendix S3.2, the expected payoff to player i, both in the revision game and in Game F, is given by

$$\pi_i(T)e^{-\lambda T} + \int_0^T \pi_i(s)e^{-\lambda s}\lambda ds.$$

References

[1] Bernheim, B. D. and A. Dasgupta (1995): "Repeated Games with Asymptotically Finite Horizons," *Journal of Economic Theory*, 67(1): pp.129-152.

[2] Dutta, P. K. (1995): "A Folk Theorem for Stochastic Games," *Journal of Economic Theory*, 66: 1-32.

[3] Hörner, J, T. Sugaya, S. Takahashi, and N. Vieille (2011): "Recursive Methods in Discounted Stochastic Games: An Algorithm for $\delta \to 1$ and a Folk Theorem," *Econometrica*, 79(4): 1277-1318.

[4] Kamada Y. and M. Kandori (2017): "Revision Games, Part I: Theory," mimeo.

[5] Kamada Y. and M. Kandori (2019): "Cooperation in Revision Games and Some Applications," mimeo.

[6] Kamada, Y. and F. Kojima (2014): "Voter preferences, Polarization, and Electoral Policies," *American Economic Journal: Microeconomics*, 6(4): 203-236.

[7] Kamada, Y. and T. Sugaya (2019): "Valence Candidates and Ambiguous Platforms in Policy Announcement Games," mimeo.

[8] Milnor, J. and Shapley, L. S. (1957): "On Games of Survival," *Contributions to the Theory of Games*, 3, 15-46.

[9] Osborne, Martin J., “Spatial Models of Political Competition Under Plurality Rule: A Survey of Some Explanations of The Number of Candidates and The Positions They Take,” *Canadian Journal of Economics*, 1995, 2, 261-301.

[10] Shapley, L. S. (1953): “Stochastic Games,” *Proceedings of the National Academy of Sciences*, 39(10): 1095-1100.

[11] Sherstyuk, K., N. Tarui, and T. Saijo (2013): “Payment Schemes in Infinite-Horizon Experimental Games,” *Experimental Economics*, 16(1): 125-153.