Uncertain Reaction lag and Cooperation in Continuous Prisoners' Dilemma

In-Uck Park*
University of Bristol, U.K. and IAS, Princeton, U.S.A.

July 3, 2012

(Very Preliminary–do not circulate)

1 Introduction

Friedman-Oprea (2012) report experimental results of a continuous-time prisoners' dilemma game. They find that the cooperation is much higher in their experiment (above 90% of the duration of the game which is 60 seconds) than in the previous experiments in discrete time settings. They also report that, when both subjects start with Defect (D) at the beginning, typically one player switches to Cooperate (C) in the first few seconds with the other following suit within one second, and then one player switches back to D a few seconds before the end with the other quickly following suit. This pattern seems natural, yet has not been theoretically justified as an equilibrium (as opposed to -equilibrium).

As Friedman-Oprea point out, the asynchronous nature of moves may contribute to the increased cooperation in continuous-time environments where the reaction can be quick. That is, players may be more willing to continue with C, betting on the opponent doing the same, when they can respond quickly upon finding out that their bet was wrong, relative to the situations in which wrong bets cannot be mended for a fixed duration of time as in discrete time settings.

The basic insight may be depicted by the varying efficacy of the "trigger strategy until the end" that has been examined in previous studies: it is an -best response if the time lag to respond to "cheating" is very small (Radner, 1986; Bergin-MacLeod, 1993) and it is a best response if the time lag is nil (Simon-Stinchcombe, 1989).

A key difficulty in supporting cooperation in finite-horizon settings of the prisoners' dilemma game is to prevent unraveling from the end of the game. The aforementioned experimental data exhibit prolonged cooperation in spite of eventual defections by both players. The data also exhibit a small yet clear time lag in responding to the opponent's moves. If the length of the time lag is public information as in the papers cited above, the

*This project started while I was a Richard B. Fisher member at the Institute for Advanced Study, Princeton, and I thank their hospitality. I also thank Heski Bar-Isaac, Dan Friedman, Bentley MacLeod, Stephen Morris, Hyun Song Shin, and Aleksey Zinger for various helpful comments and suggestions. All remaining errors are mine.

temptation to pre-empt the opponent unravels cooperation incentives all the back to the beginning. Thus, the prolonged cooperation followed by eventual defections remain as a theoretical puzzle.

In this paper I resolve this puzzle by introducing ***heterogeneous reaction lags as private information.*** The key intuition is that the quicker a player is in responding, the longer she is willing to continue with cooperation. Then, each player perceives the opponent's defection time as a random variable, and at any point in time toward the end there always is a chance that the opponent, being a sufficiently quick type, will cooperate a little longer (unless one already defected). This, in turn, justifies each player's willingness to cooperate a little longer if her reaction lag is short enough.

I characterize an equilibrium that exhibits such features, and also calculate the actual equilibrium for parameter values corresponding to the experiments of Friedman-Oprea. The calculated equilibrium matches the experimental data closely. I also discuss some modelling issues in continuous-time and the approach taken to deal with them.

As the upper bound of the reaction lag vanishes, the limit of the aforesaid equilibrium prescribes cooperation throughout the whole duration of the game, which is the solution predicted by Simon-Stinchcombe (1989) and Bergin-MacLeod (1993) for the continuous-time game of prisoners' dilemma. Therefore, the current analysis supports their predictions as the limiting equilibrium as the reaction lag vanishes, which has not been established in the literature as far as I am aware.

In my model, cooperation is sustained until toward the end by way of endogenously creating and capitalizing on the prospect at any point in time that the opponent may continue with cooperation a little longer. In a famous article, Kreps, et al. (1982) also support cooperation in a discrete-time, finitely repeated game of prisoners' dilemma based on a similar idea: if each player is perceived to be committed to a "Tit-for-Tat" strategy with a small probability, a normal player will masquerade as such a type to induce cooperation from the opponent who may entertain the idea that she may facing such a commitment type. My approach differs crucially in that it does not introduce uncertainty on the opponent's preference, and consequently, does not rely on the idea of reputation building (i.e., some type trying to imitate another type) to be perceived as committed to a particular behavior pattern conducive to cooperation. Rather, in my model each type would defect at a uniquely optimal time that maximizes the same utility function within its own physical constraint. As a result, the "end play" of the equilibrium (i.e., after the cooperation breaks down) is very simple in my model, unlike in the model of Kreps, et al. (1982) in which the authors find it "very complex" that they do not characterize it in their article.

2 Illustration

A. Game

Two players, $i = 1; 2$, play the following game over the time interval $I = (0; 1]$:

1\ 2	C	D
C	(1,1)	(0,h)

D	$(h, 0)$	(ℓ, ℓ)

where $0 < \ell < 1 < h$.[1] The players start with (D, D) at $t = 0$ and at any time $t \in I = (0, 1]$, they may switch to the other action and back as they wish (subject to constraints below), which is observed by the other player instantaneously. The players receive flow payouts determined by the action pairs prevailing at each point $t \in [0, 1]$ without discounting.

The players face two related constraints concerning reaction time:

[A] If player j switches her action at $t \in (0, 1)$, then player $i \neq j$ cannot switch her action during $(t, t + \theta_i)$.

[B] If player i switches her action at $t \in (0, 1)$, then she cannot switch her action again during $(t, t + \;_i)$.

Part [A] means that $t + \theta_i$ is the first instance of time after t, at which player i may respond to her opponent's move at t (provided that $t + \theta_i \leq 1$). This reflects the time needed for player i to recognize the changed situation, decide how to react, and put the decision into action. The value of θ_i, which we call the "reaction lag," is private information (type) of player i, $i = 1, 2$, which is an independent draw from a common distribution represented by an atomless cdf $F(\theta)$ on $(0, \bar{\theta})$ where $\bar{\theta} > 0$ is small, with an associated density function $f(\theta) = F'(\theta)$ which is assumed to be continuous: i.e.,

(1) $$\theta_i \sim F \text{ on } (0, \bar{\theta}) \quad \text{where} \quad \lim_{\theta \to 0} F(\theta) = 0 \text{ and } f(\theta) = F'(\theta) \text{ is continuous.}$$

Part [B] postulates an analogous minimum time lag between any two consecutive moves by the same player, which reflects the physical constraint of the player. It seems plausible that $\;_i \leq \theta_i$. For expositional ease (to avoid two-dimensional types), we assume $\;_i = \;\theta_i$ for a small $\; \in (0, 1]$.

The game is conceptually quite simple, yet players may make multiple moves in a continuous-time framework. A rigorous formulation of the players' strategies in such environments requires laying down a nontrivial analytical structure, as is done in the next section. In this section, an informal description seems to suffice: A strategy of an "agent" (a player of a given type) specifies, contingent on every possible history, a probability that she will have switched her action by every future point in time, conditional on no player has switched her action by then after the referenced history.[2] An agent's strategy is required to respect [A], [B], and an inter-temporal consistency condition stipulating that an agent's plan (of switching action) contingent on a possible future history does not change as the history evolves toward that particular future history and reaches it. A profile of agent strategies of the two players constitutes a perfect Bayesian equilibrium if the usual mutual best-response property is satisfied and the belief profile is consistent with Bayes rule.

B. Description of equilibrium

We present a symmetric equilibrium for the case that F is uniform, i.e., $F(\theta) = \theta/\bar{\theta}$ and $f(\theta) = 1/\bar{\theta}$. It is described below as a sequence of several phases.

---

[1] We could assume $h < 2$ so that (C, C) is efficient, but not needed for our results.

[2] That is, it consists of an "open-loop" strategy for every history until the next move takes place.

1) ***War-of-attrition (WoA) phase:*** Starting from $(D; D)$ at the beginning of the game, both players switch to C with a common "flow rate" $q > 0$ provided that the time has not passed T and the other has not switched to C already, where $T < 1 - {}^{-}(h - \grave{} + \frac{1+h}{1\ \ \ell})$ is a fixed point in time. Thus, for each player the cumulative probability that she will have switched to C at or before $t \in (0; T]$ provided that the other has not switched to C by then, is $1 - e^{\ \ qt}$. We refer to this initial part of the game until one player switches to C at some point in time before T, as the "war-of-attrition (WoA) phase."

2) ***Transition phase to cooperation:*** If one player, say i, switched to C in the WoA phase, say at $t \leq T$, then player j follows suit ¯ later, i.e., switches to C at $t + {}^{-}$ (and i adheres to C until then).[3] In case neither player switched to C by T, both players switch to C simultaneously at $T^{0} = T + \frac{\bar{\theta}}{1\ \ \ell} < 1 - \left(\frac{h}{1\ \ \ell} + h - \grave{}\right)^{-}$.[4]

3) ***Cooperation phase:*** After the two players both switched to C as explained in 2) above, they continue with $(C; C)$ until the time reaches $\widehat{T} = 1 - \frac{h}{1\ \ \ell}{}^{-}$. We refer to this part of the game as the "cooperation phase."

4) ***Defection phase:*** Once the time reaches $\widehat{T}$ through the cooperation phase, each player switches to D at an appropriate time point during $(\widehat{T}; 1)$ depending on her type, conditional on the other not having switched to D already: Specifically, a player of type defects when the remaining time is $r(\ ) = \frac{h}{1\ \ \ell}$, i.e., at $t = 1 - r(\ )$. Once one player, say i, defects at t, player $j \neq i$ follows suit as soon as possible, i.e., at $t + {}_{j}$, after which both players keep to D until the end of the game. (As shown below, both will switch to D before the game ends.) We refer to this part of the game as the "defection phase."

5) ***O -equilibrium:*** If history at any t departs from the equilibrium-paths described above, then both players switch to D as soon as they can, unless they were already playing D, and keep to D until the end of the game. A caveat is in order: in the off-equilibrium contingency that one player, say i, switched to C at t in the WoA phase, yet the other player did not follow suit at $t + {}^{-}$, the earliest time after this deviation that player i can switch back to $+$

verify that, off-the-equilibrium path, it is optimal for each player to keep or switch to D as soon as possible, given that the other will do the same (apart form the aforementioned exception case which we will deal with later). Below we verify optimality of the strategy on the equilibrium-path, backwards from the defection phase.

Defection phase: In the defection phase, once one player defects, again from the fact that D is a strictly dominant strategy of the static game, it is optimal for the other player to follow suit as soon as possible, and then for both to keep to D until the end, as specified above.

Next, consider any point $t \in (\widehat{T}; 1)$ in the defection phase with no defection having taken place from $(C; C)$ by either player until then. According to the defection strategy described above, either player may defect at any point between $\widehat{T}$ and 1 depending on her type, with a higher type defecting earlier. Letting $r = 1 - t$ denote the time remaining until the end of the game, therefore, the posterior belief of either player's type at $t = 1 - r$ is uniform on the interval $(0; \#(r))$ where $\#(r) = \frac{1\ \ell}{h} r$ is the type that is supposed to defect at $t = 1 - r$, because a player of a type higher than $\#(r)$ should have defected by $t = 1 - r$. Consequently, each player anticipates that the other player will defect with a uniform distribution during the remaining period of the game, $(t; 1)$.

Given this, let's calculate the expected payoff of a player, say i, from waiting an infinitesimal amount of time, Δ, from $t = 1 - r$ before defecting. In the contingency that she gets pre-emptied between t and $t + \Delta$ (i.e., player $-i$ defects in between), player i's payoff is higher the lower her type, $\ _i$, is (because lower $\ _i$ means that she can follow suit more quickly); but her payoff is the same regardless of her type if she does not get pre-emptied. Therefore, if the marginal net change in player i's expected payoff from waiting is 0 when her type is $\#(r)$, any lower type should strictly prefer to wait, as prescribed by the proposed equilibrium strategy. The change in payoff for a $\#(r)$-type agent is $\Delta(1 - `)$ if she does not get pre-emptied (because the wait prolongs the duration of $(C; C)$ by Δ, which reduces the duration of $(D; D)$ by the same amount of time at the end of the game); and it will be $-\#(r)h$ if she gets pre-emptied (because she, instead of her opponent, will suffer from the "sucker" payoff 0, as opposed to h, for a duration of the reaction lag, $\#(r)$). Since the latter happens with probability $\Delta{=}r$, the net change in her payoff is first-order approximated[5] as $\Delta(1 - `) - \Delta\#(r)h{=}r$, which is 0 because $\#(r) = \frac{1\ \ell}{h} r$. Hence, any type lower than $\#(r)$ strictly prefers to wait at t as explained above. Additionally, it is a straightforward calculation to show that it is optimal for a player of type $\#(r)$ to defect at $t = 1 - r$ (outlined in the previous footnote). This verifies the optimality of the defection phase.

Cooperation phase: Consider any point in time $t < \widehat{T}$ in the cooperation phase after

[5] Precisely, it is $P(r, \ , \theta_i) = \frac{\Delta}{r}\left[\frac{\Delta}{2} + \left(r - \frac{\Delta}{2} - \theta_i\right)\ell\right] + \frac{r-\Delta}{r}\left[\ \ + \frac{(r-\Delta)(1-\ell)}{2h}h + (r - \ \ - \frac{(r-\Delta)(1-\ell)}{2h})\ell\right]$ where θ_i is player i's type. Its rst and second derivatives wrt are calculated to be, respectively, [see cntnsPD-1.nb] $\frac{(r-\Delta)(1-\ell)-\theta_i h}{hr}\ell$ and $\frac{(\ell-1)\ell}{hr} < 0$. Note that the rst derivative decreases in θ_i, and is 0 when $\ = 0$ and $\theta_i = r(1 - \ell)/h = \vartheta(r)$, which implies that the continuation payo increases as increases from 0 for all $\theta_i < \vartheta(r)$, so such types should not switch to D at $t = 1 - r$. For player i of type $\vartheta(r)$, since $P(r, \ , \vartheta(r))$ is strictly concave in , waiting until is worse than defecting at $t = 1 - r$ so long as $r - \ /2 - \vartheta(r) > 0$. When $r - \ /2 - \vartheta(r) \leq 0$, the rst term of $P(r, \ , \vartheta(r))$ become $(\ /r)(\ /2)$ and the derivative of $P(r, \ , \vartheta(r))$ wrt is calculated to be $\ell(1 - h - \ell)(r - \ )/(hr) < 0$.

both players have switched to C. It is straightforward to verify that it is suboptimal for either player to switch to D before $\widehat{T}$ because she would be better off by doing so at $\widehat{T}$ instead, which would prolong the duration of $(C; C)$ by $\widehat{T} - t'$ and curtail the duration of $(D; D)$ by the same length later.

Transition phase to cooperation: Moving backward one step, consider player j after player i has switched to C at $t < T$ in the WoA phase: It is clearly optimal for j to switch to C at $t+\beta$ because otherwise i will switch back to D as soon as possible after $t+\beta$ (again, this point in time is not pinned down–see the next section) and keep to it until the game ends, which is worse. (Indeed, not switching at $t+\beta$ is dominated by switching at $t+\beta$ and then defecting at $\widehat{T}$.) Given this, it is optimal for player i to keep to C until $t+\beta$ because her switching back to D before $t+\beta$ is worse by the same reasoning.

Consider the other case that neither player switched to C until T. Then, by switching to C at T', the players get into cooperation phase until $\widehat{T}$, followed by the defection phase. The expected payoff from this is easily calculated to be larger than that for a player who behaves differently, because then the off-equilibrium strategy prescribes that both players would switch to D as soon as possible (unless they were already playing D) and keep to D until the end of the game.

WoA phase: The game starts with both players playing D at $t = 0$. Consider either player, say i, at $t < T$ in the contingency that neither player has switched to C by then. Relative to the option of switching to C right now (i.e., at t), let's compare that of waiting for an infinitesimal amount of time Δ and switching to C at $t + \Delta \leq T$ (provided that the other player has not switched in between). If the other player, j, switches in between, which happens with probability $1 - e^{-q\Delta}$, player i would enjoy a payoff of h until she follows suit herself β later, whereas she would get a payoff of 0 for the same duration if she switched to C at t. If player j did not switch in between, which happens with probability $e^{-q\Delta}$, the wait simply delay the transition to the cooperation phase: this prolongs the initial state of $(D; D)$ by Δ and curtails the duration of $(C; C)$ by the same amount of time in the cooperation phase, costing $1 - \ell$ both players for a duration of Δ. Therefore, the expected net gain from waiting is first-order approximated[6] by $(1 - e^{-q\Delta})\beta h - e^{-q\Delta}\Delta(1 - \ell)$ and its derivative wrt Δ is $q\beta h - (1 - \ell)$ when evaluated at $\Delta = 0$. By setting $q = \frac{1-\ell}{\theta h}$, therefore, we ensure that both players are indifferent between switching to C at any point $t < T$ and waiting until a later time $t' \leq T$, establishing the optimality of the strategy in the WoA phase.

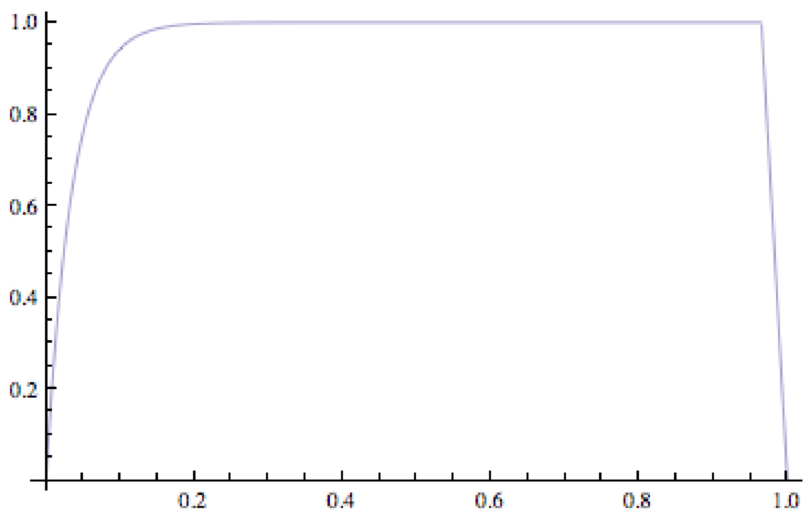
It remains to show that a player, i, is indifferent between switching to C at T and not switching at any $t \leq T$. Since the two options are identical if the other player, j, switches to C before T, consider the case that player j does not switch until T. (The contingency that player j switches exactly at T is a measure 0 event, so is left out in calculation.) By switching at T, player i receives a payoff of $\widehat{T} - T - \beta$ (calculated from a payout of 0 from T to $T + \beta$, then 1 until $\widehat{T}$) before the defection stage starts. By not switching at T, she

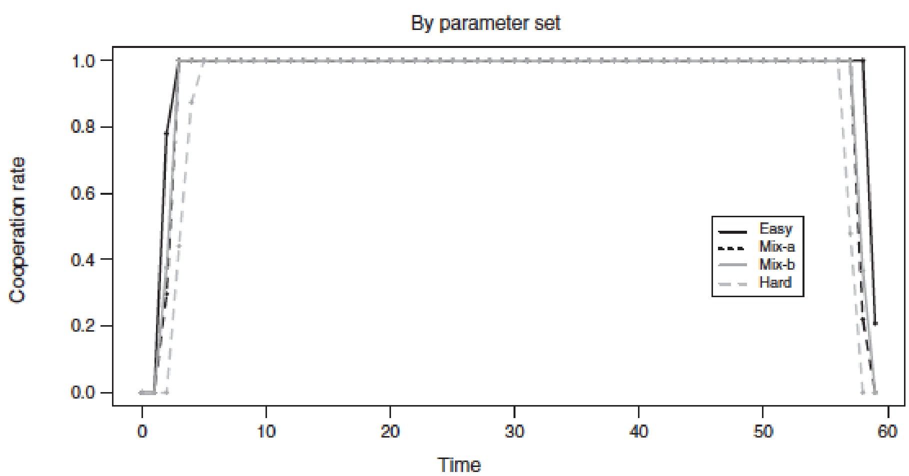
[6] A full expression is $\int_0^\Delta q e^{-q\tau}(\tau\ell + \theta h + (\widehat{T} - t - \tau - \theta) + V_{\widehat{T}})d\tau + e^{-q\Delta}\left(\Delta\ell + (\widehat{T} - t - \Delta - \theta) + V_{\widehat{T}}\right) = \frac{e^{-q\Delta}}{q}\left(1 - \ell - \theta qh - e^{q\Delta}(1 - \ell - \theta q(h-1) - q(\widehat{T} - t))\right)$ where $V_{\widehat{T}}$ denotes the continuation payoff from $\widehat{T}$. Its derivate with respect to Δ is $-e^{-q\Delta}(1 - \ell - \theta qh)$ which is identically 0 when $q = \frac{1-\ell}{\theta h}$.

receives the same payoff of $\widehat{T} - T -$ ¯ (calculated from a payout of ` from T to $T^{0} = T + \frac{\bar{\theta}}{1 \ \ell}$, then 1 until $\widehat{T}$) before the defection stage starts, establishing the desired indifference.

D. Discussion of equilibrium

This equilibrium portrays a picture very close to the experimental results reported in Friedman and Oprea (2012, AER) who estimate the median reaction lag at 0.6 second. We simulated the cooperation rate of our equilibrium at every $t \in (0; 1)$ for ¯ $= .015$ min or 0.9 second, which is illustrated in the first graph below in parallel with the corresponding graph of the experimental results. The parameter values used for our equilibrium, $h = 1.4$ and ` $= .4$, correspond to the "Easy" case of the experiment.[7] The theoretical prediction is very close to the experimental results for the "Mix-a" case as well ($h = 1.8;$ ` $= .4$), although it under-predicts cooperation in the "Mix-b" and "Hard" cases.





<u>Random $_i$:</u> We note that the equilibrium strategy in the defection phase is straightforwardly extended to the cases that the reaction time of a $_i$-type player is a random variable

[7] The graph from experiment is the median cooperation rate with initial states randomly assigned to all four possible action pairs, which biases the initial cooperation forward a little. But, our graph is for the rst switch to C, which also biases forward a little.

with a mean $_i$, as long as its support is narrow enough to ensure that a reaction of a $_i$-type player, upon observing a defection by the opponent at $t < 1 - r(\ _i)$, is realized before the end of the game. This is because the analysis only relies on the expected payoffs which remain the same in these variants of the model. Consequently, the equilibrium described in the previous section extends straightforwardly to this case as well, with the flow rate of switching in the WoA phase modified accordingly (slower because the time lag between the first and second switches to C will be longer to ensure that the slowest type can make it for sure).

Flickering/blitzing: It is conceivable that once in the cooperation phase, a player with a very short reaction lag (i.e., very low $_i$ and thus, low $_i$) would switch back and forth as quickly as she can, so as to benefit from her opponent not being able to respond in case her reaction lag is longer than $_i$. This is a byproduct of extending the standard practice of modelling a reaction lag in continuous-time settings, to allow for a seemingly realistic feature that players may differ in the length of their reaction lag. Consequently, the possibility that a player may block her opponent from responding by "flickering" appears to be an artificial, rather than realistic, concern. One can resolve this issue, for instance, by introducing "exhaustion" to the effect that if a player switches repeatedly within an interval less than $\bar{\ }$, then she bears a cost for each switch that increases sufficiently rapidly with the number of repetition; or conversely, a player can reduce her reaction time below $_i$ at a reasonable cost. Both costs do not seem unrealistic. However, we take a different approach described below.[8]

Punishing the opponent when she does not follow suit to cooperation: Consider the off-equilibrium contingency that one player, say i, switched to C at t in the WoA phase, yet the other player did not follow suit at $t + \bar{\ }$. Then, player i wishes to switch back to D at the earliest time after t, but such a time does not exist because the real numbers are not well-ordered and consequently, player i's equilibrium strategy in this continuation game is not determined in this manner. One can go around this technical issue by postulating a more involved continuation equilibrium. Specifically, player i switches back to D at $t + \bar{\ } +$ for a small $\ > 0$, then both players switch to C at some later time before $\widehat{T}$, say T^{00}, followed by the cooperation and defection phases of the equilibrium-path. If player i switches back to D either before or after $t +$, then the continuation would be for both players to keep to D until the end. Provided that T^{00} is sufficiently late, player $j \neq i$ prefers following the equilibrium-path to invoking this continuation game by failing to follow suit at $t + \bar{\ }$. However, we take an alternative approach as explained below.

Reaction lag as the time to absorb updated situations: Although it serves the purpose, the off-equilibrium strategy above is devised to circumvent a purely technical issue inherent in continuous-time analysis. As a matter of fact, when the anticipated act of following suit

[8]Another possibility is to introduce \ rm moves" which each player may commit to a certain action with certainty at a predetermined action time at least θ_i prior to that time, regardless of what happens in between. (This does not include probabilistic moves such as in the WoA phase, so does not include closed-loop strategies.) However, this seems to create another issue in the defection stage because players would want to commit to defect at a certain time to reduce reaction time when pre-emptied. It is not clear that our equilibrium can be modi ed to acccommodate this.

does not take place, it seems plausible that the player who was anticipating it would find herself in an unexpected situation and, therefore, she would need a little time to absorb the updated situation before deciding how react. This line of reasoning may be extended a little further: if a player anticipates a move from her opponent with a positive probability at a certain future point in time and intends to react differently depending on whether the anticipated move indeed takes place or not, then upon reaching that future point in time, she will need a little time to absorb which of the two possible contingencies has actually realized before reacting, regardless of whether the anticipated move indeed took place at time t or not. Interpreting reaction lag in this way as a small lapse of time needed to absorb "updates" of the relevant situation, one is led to apply the reaction lag defined earlier in [A] when the opponent fails to take an anticipated move, as well as when the opponent takes a move (expectedly or unexpectedly).

This is the approach adopted in the next section to develop a formal model and formulate players' strategies. Note, however, that whether a player anticipates a move by her opponent at a certain time or not is determined by the other player's strategy. To accommodate this interdependence, first we formulate "new plans" as building blocks of an agent's strategy, where a new plan (i) starts either when the opponent makes a move or when the previous plan expires, and (ii) prescribes probabilities that the agent will have switched to the other action by every future points in time until its expiration time (which can be the end of the game). Then, we require that an equilibrium strategy profile be "internally consistent" in the sense that whenever a plan has its expiration before the end of the game, it is the first instance of time at which a move is anticipated by the opponent with a positive probability.

Additionally, we make the modelling choice that requires every new plan to prevail at least for a duration equal to the player's reaction lag, so as to ensure that she can react to the updated situation uninterrupted by the opponent's moves. This modelling choice is for the sake of resolving the aforementioned issue of flickering, and is not needed if either cost discussed there (on quick/frequent switches) is introduced. We make this choice so that our formulation of strategies resolves both of the technical issues explained above, namely, flickering and the nonexistence of the optimal time to switch back to D when the opponent fails to follow suit.

3 A Formal Model

The game starts with $(D; D)$ at $t = 0$. For $t \in (0; 1)$, a history $h^t = (h_1^t; h_2^t)$ consists of two sequences of switches/moves, one for each player: $h_i^t = ({}_1; {}_2; \cdots ; {}_n)$ is a list of sequential times at which player i switched her action, where $0 < {}_1 < \cdots < {}_n \leq t$. Hence, the odd-number-subscripted times are when the player switched from D to C and the even-number-subscripted times are when the player switched from C to D. Let $h_i^0 = \emptyset; \ i = 1; 2$, for completeness. A history $h^{t'}$ is an *extension* of h^t if $t < t'$ and the two histories are identical up to and including time t; and $h^{t'}$ is a *simple extension* of h^t if, in addition, neither player switched after t according to $h^{t'}$. Given a history, a "jump" refers to an instance of either one player switching or both players switching at the same time. Histories need to satisfy

certain properties to be consistent with admissible strategies to be formalized below.

A ***plan at*** h^t of player i of type θ_i, "agent θ_i" for short, is a right-continuous non-decreasing function $p_i(\cdot|h^t;\theta_i) : (t;T_i(h^t;\theta_i)] \to [0;1]$ where $p_i(t^+|h^t;\theta_i) := \lim_{s\#t} p_i(s|h^t;\theta_i) = 0$ and $T_i(h^t;\theta_i) \in (t;1]$, called the "expiration-time", is the point in time until which the agent θ_i considers this plan to be valid. Abusing notation, we use $p_i(\cdot|h^t;\theta_i)$ to mean a plan, i.e., as a shorthand for $(p_i(\cdot|h^t;\theta_i);T_i(h^t;\theta_i))$. The value $p_i(s|h^t;\theta_i)$ denotes the probability that, conditional on the history h^t has been reached, the agent θ_i will have switched by $s \in (t;T_i(h^t;\theta_i)]$ without previous switches by either player after h^t.

An ***agent-strategy*** of player i of type θ_i, denoted by ${}_i(\theta_i)$, is a collection of plans $(p_i(\cdot|h^t;\theta_i);T_i(h^t;\theta_i))$, one for every possible history h^t where $t \in [0;1)$, together with a subset ${}^N_i(\theta_i) \subset {}_i(\theta_i)$ of "new plans", with the following properties:

[C1] The initial plan, $p_i(\cdot|h^0;\theta_i)$, is in this subset, i.e., $p_i(\cdot|h^0;\theta_i) \in {}^N_i(\theta_i)$.

[C2] If $p_i(\cdot|h^t;\theta_i) \in {}^N_i(\theta_i)$ and $T_i(h^t;\theta_i) < 1$, then $p_i(\cdot|h^{T_i(h^t,\theta_i)};\theta_i) \in {}^N_i(\theta_i)$ when $h^{T_i(h^t,\theta_i)}$ is an extension of h^t with no jumps during $(t;T_i(h^t;\theta_i))$, although there may be a jump at $T_i(h^t;\theta_i)$. In this case, we say $p_i(\cdot|h^{T_i(h^t,\theta_i)};\theta_i)$ is a "new plan due to expiration."

[C3] If $p_i(\cdot|h^t;\theta_i) \in {}^N_i(\theta_i)$ and either (i) it is a new plan due to expiration, or (ii) h^t contains a move by player $-i$ during $(t-\theta_i;t]$,[9] then

$$\min\{1;t+\theta_i\} \le T_i(h^t;\theta_i) \text{ and } p_i(s|h^t;\theta_i) = 0;\ \forall s \in (t;t+\theta_i) \cap (0;1). \tag{2}$$

Furthermore, $p_i(\cdot|h^{t'};\theta_i) \notin {}^N_i(\theta_i)$ if $h^{t'}$ is an extension of h^t and $t' < t+\theta_i$. In this case, we say $p_i(\cdot|h^t;\theta_i)$ is a "new plan due to situation update."

[C4] If $p_i(\cdot|h^t;\theta_i) \in {}^N_i(\theta_i)$ and neither (i) nor (ii) above holds where $t > 0$, then h^t contains a move by player i and $p_i(s|h^t;\theta_i) = 0\ \forall s \in (t;t+\theta_i) \cap (t;T_i(h^t;\theta_i))$. In this case, we say $p_i(\cdot|h^t;\theta_i)$ is a "new plan solely due to own move."

[C5] If $p_i(\cdot|h^t;\theta_i) \in {}^N_i(\theta_i)$ and $h^{t'}$ is an extension of h^t with a first jump at $t' < T_i(h^t;\theta_i)$ since t, then $p_i(\cdot|h^{t'};\theta_i) \in {}^N_i(\theta_i)$ unless $p_i(\cdot|h^t;\theta_i)$ is a new plan due to situation update and $t' < t+\theta_i$; in the latter case, the jump at t' is by player $-i$ and $p_i(\cdot|h^{t+\theta_i};\theta_i) \in {}^N_i(\theta_i)$ where $h^{t+\theta_i}$ is any extension of $h^{t'}$.

The condition [C3] stipulates physical constraints of player i in responding to an updated situation namely, that i) it takes at least θ_i for her to absorb the updated situation and react to it, and that ii) each new plan lasts at least for a duration of θ_i if prompted by an updated situation, so that she has a chance to respond to that update of situation. The latter case of [C5] stipulates that if the other player takes another move before player i has a chance to respond to the updated situation, then yet another new plan starts at $t+\theta_i$ for player i (this respects the principle that any updated situation warrants a separate reaction lag for info processing and consequently, a new plan). These place some constraints on the

[9] This means that all \situation updates" that take place during $(t, t+\theta_i]$, i.e., during the reaction lag prompted by the previous situation update(s), are processed together during the same reaction lag $(t+\theta_i, t+2\theta_i)$ and responded collectively at $t+2\theta_i$.

histories to define $_i(\ _i)$ on (these constraints are rather complex to fully describe), e.g., agent i's move cannot be taken within $_i$ of a move by player $-i$, unless it was preceded by a sequence of new plans that lasted only for $_i$ each due to player $-i$'s move within the reaction lag of each new plan. These are common knowledge.

In addition, we require that an agent-strategy $_i(\ _i)$ with $^N_i(\ _i)$ satisfy "inter-temporal consistency" conditions:

[C6] If $p_i(\cdot|h^t;\ _i) \in\ _i(\ _i)$ and $h^{t'}$ is a simple extension of h^t such that $t' < T_i(h^t;\ _i)$, then $p_i(\cdot|h^{t'};\ _i) \in\ _i(\ _i)$ has the same expiration-time as $p_i(\cdot|h^t;\ _i)$, i.e., $T_i(h^{t'};\ _i) = T_i(h^t;\ _i)$, and $p_i(\cdot|h^{t'};\ _i)$ is obtained from $p_i(\cdot|h^t;\ _i)$ via Bayesian updating:

$$p_i(s|h^{t'};\ _i) = \frac{p_i(s|h^t;\ _i) - p_i(t'|h^t;\ _i)}{1 - p_i(t'|h^t;\ _i)} \quad \forall s \in [t'; T_i(h^t;\ _i)] \tag{3}$$

provided that the denominator is non-zero. In this case, we say that $p_i(\cdot|h^{t'};\ _i)$ is a "truncated(?) plan" of $p_i(\cdot|h^t;\ _i)$.

[C7] If $p_i(\cdot|h^t;\ _i) \in\ ^N_i(\ _i)$ is a new plan due to situation update and $h^{t'}$ is a non-simple extension of h^t such that $t' < t +\ _i$, then $p_i(s|h^{t'};\ _i) = p_i(s|h^t;\ _i)$ $\forall s \in (t'; t +\ _i]$ and $T_i(h^{t'};\ _i) = t +\ _i$. ($h^{t'}$ is illegitimate if any move after t was by player i.)

A ***strategy*** of player i, denoted by $_i$, is a collection of agent-strategies, one for each type $_i \in (0;\ ^-)$, i.e., $_i \in \Sigma := \prod_{\theta \in (0,\bar{\theta})} \Sigma(\ )$ where $\Sigma(\ )$ is the set of all agent-strategies of a player i of type that satisfy [C1]–[C7]. Although we do not restrict the strategies to "legitimate" histories (relative to the given strategy), only such histories will be relevant in our analysis.

To determine the outcome from a strategy profile, we now formalize the set of all possible histories. Given a type profile $(\ _1;\ _2) \in (0;\ ^-)$, possible first jumps are represented by points in $J_1 := I^3$ where $I = (0; 1]$. For expositional ease, denote the first, second and third copy of I in J_1 by $I_x; I_y$ and I_z, respectively. Then, the interpretation of points in J_1 is: a point $x \in I_x = (0; 1]$ is the contingency that only agent $_1$ switches (from D to C) at $t = x$ (without previous moves); a point $y \in I_y = (0; 1]$ is the contingency that only agent $_2$ switches (from D to C) at $t = y$ (without previous moves); and a point $z \in I_z = (0; 1]$ is the contingency that both agents $_1$ and $_2$ switch (from D to C) at $t = z$ (without previous moves). For $j_1 \in J_1$, let $|j_1| \in (0; 1]$ denote the time that the jump j_1 takes place, i.e., without reference to the player(s) who make it.

Inductively, for any admissible history $h^{|j_n|} = (j_1; j_2; \cdots; j_n)$, the set of possible next jumps can be represented by $J_{n+1}(h^{|j_n|}) := I_x(h^{|j_n|}) \times I_y(h^{|j_n|}) \times I_z(h^{|j_n|}) \subset \mathcal{I} := I^3$ where each $I_{(\ )}(h^{|j_n|})$ is an interval of the form $[\ ; 1]$ where $> |j_n|$. This process stops for some n bounded above because any two moves by the same player is at least $_i$ apart. Thus, allowing illegitimate histories, we may define the set of "outcomes" as $\mathcal{O}(\ _1;\ _2) = \left(\bigcup_{n=1}^{M(\theta_1,\theta_2)} \mathcal{I}^n\right) \cup \{\emptyset\}$ for a type profile $(\ _1;\ _2)$, where $M(\ _1;\ _2)$ is the maximum number of jumps that may take place in any history (including off-equilibrium) according to any $(\ _1(\ _1);\ _2(\ _2))$, and $\emptyset$ is the contingency that no player ever switched until the end of game. As $M(\ _1;\ _2) \to \infty$ as $\min\{\ _1;\ _2\} \to 0$, the set of outcomes across all type profiles

is $\mathcal{O} = \left(\bigcup_{n=1}^{1} \mathcal{I}^n\right) \cup \{\emptyset\}$. Note that a history and all simple extensions of it correspond to the same outcome in $\mathcal{O}$.

Consider a strategy profile $({}_1, {}_2)$. We restrict attention to profiles that satisfy the following measurability conditions. For each player $i = 1, 2$, the initial plans $p_i(h^0, {}_i)$, ${}_i \in (0, 1)$, and other new plans that start at simple extensions of h^0, define a non-decreasing right-continuous function $P_i(h^0)$, called player i's plan at h^0, that represents the cumulative probability that player i (of any type) will have switched without a previous switch by either player, i.e., $P_i(s|h^0)$ is the integral of $p_i(s|h^0, {}_i)$ or relevant new plans that succeed it, relative to the prior distribution F on ${}_i$. For any simple extension $h^t_{,}$ of h^0, a posterior belief, denoted by $F_i(h^t_{,})$, on ${}_i$ is determined by Bayes rule if t is in the support of $P_i(h^0)$, or by arbitrary assignment otherwise. For any history $h^{|j_1|} = (j_1)$, a posterior belief, denoted by $F_i(h^{|j_1|})$, is determined in the same manner. Similarly, for any admissible history $h^{|j_n|} = (j_1, j_2, \cdots, j_n)$, player i's plan at $h^{|j_n|}$, $P_i(h^{|j_n|})$, is determined analogously, and so is a posterior belief at any simple extension of it or at any extension of the form $h^{|j_{n+1}|} = (j_1, j_2, \cdots, j_n, j_{n+1})$.

We now evaluate the expected payoff the agent ${}_i$ from a strategy profile $({}_1, {}_2)$. In this discussion, measurability issues arise at places. We proceed by presuming that necessary measurability conditions are satisfied, because they are indeed sastified in the equilibrium I characterize.

Consider a strategy profile $({}_1, {}_2)$ and an agent ${}_i$. Then, $P_{\ i}(h^0)$ and player i's initial plan $p_i(h^0, {}_i)$ and other new plans that start at simple extensions of h^0, define for each player (more precisely, for player $-i$ and for agent ${}_i$) a cumulative distribution function for her to have made a first jump by any point in time $t \in (0, 1]$ during the game. These two cdf's determine a cdf for each of I_x, I_y and I_z in J_1 (it is a discrete cdf for I_z), and thereby one Stiltjes measure each on I_x, I_y and I_z in J_1 (Kolmogorov-Fomin p.362). For each jump in the support of this measure, a posterior belief is determined for each player as mentioned above.

Consider a time point $t \in (0, 1)$ with no previous jumps, denoted by $h^t_{,}$ as a history in H^t, at which a posterior belief $F_{\ i}(h^t_{,})$ is determined as explained above. Then, from the plans starting from the history $h^t_{,}$ and its simple extensions, Stiltjes measures can be obtained on $I_x \cap (t, 1], I_y \cap (t, 1]$ and $I_z \cap (t, 1]$ in J_1 analogously to above.

Inductively, consider any admissible history $h^{|j_n|} = (j_1, j_2, \cdots, j_n) \in \mathcal{I}^n$, together with associated posteriors $F_i(h^{|j_n|})$ obtained above. Then, $P_{\ i}(h^{|j_n|})$ and agent ${}_i$'s plan at $h^{|j_n|}$, together with other plans that start at a simple extension of $h^{|j_n|}$, determine a cdf for each of I_x, I_y and I_z in the $(n+1)$-th copy of $\mathcal{I}$ (with its support contained in $[|j_n|, 1]$), and thereby one Stiltjes measure each on I_x, I_y and I_z the $(n+1)$-th copy of $\mathcal{I}$. For any $h^{|j_n|+\tau}$ which is a simple extension of $h^{|j_n|}$, Stiltjes measures can be obtained on $I_x \cap (|j_n| + \ , 1], I_y \cap (|j_n| + \ , 1]$ and $I_z \cap (|j_n| + \ , 1]$ in the $(n+1)$-th copy of $\mathcal{I}$ in an analogous manner.

In the previous 5 paragraphs we determined, for each admissible history $h^t \in H^t$, a posterior belief $F_{\ i}(h^t)$ on ${}_i$ and a Stiljes measure, denoted by (h^t) on the set of next jumps to be generated by player $-i$ and agent ${}_i$ according to $({}_1, {}_2)$. Note that this can be done in the same manner when the prior on ${}_i$ was restricted to $[\ , \bar{\ }]$ for any small > 0.

Suppose we did this with the prior on ${}_i$ restricted to $[\ , \bar{\ }]$. Then, there is an upper

bound of the number of jumps that may occur in any admissible history. Let M_ϵ denote the tight upper bound. For any admissible history $h^{|j_{M_\epsilon-1}|} = (j_1; j_2; \cdots; j_{M_\epsilon\ 1})$, a continuation value can be calculated for each agent $_i$ by integrating the flow payouts over the time period $(|j_{M_\epsilon\ 1}|; 1]$ relative the aforementioned Stiltjes measure (with the residual measure applied to the contingency that there is no further move until the end of the game), which is well-defined because the payout is a simple function over the period $(|j_{M_\epsilon\ 1}|; 1]$. Similarly, for any simple extension $h^{|j_{M_\epsilon-1}|+\tau}$ of $h^{|j_{M_\epsilon-1}|}$, a continuation value of agent $_i$ can be calculated analogously. In this manner, one can calculate a continuation value of agent $_i$ for all history with $M_\epsilon - 1$ jumps.

Then, consider a history h^t with $M_\epsilon - 2$ jumps. Recall that a Stiljes measure is determined on the set of next jumps from this history to be generated by player $-i$ and agent $_i$ according to $(\ _1;\ _2)$, as explained above. By integrating the value of agent $_i$ at these possible next jumps obtained above (with the residual measure applied to the contingency that there is no further move until the end of the game), one can calculate the value of agent $_i$ at h^t. In this manner, one can calculate a continuation value of agent $_i$ at every history with $M_\epsilon - 2$ jumps (or more jumps). Continuing recursively, one can calculate the value of agent $_i$ at all histories.

Recall that these values of agent $_i$ have been calculated with the prior on $_i$ restricted to $[\ ;\ \bar{}\ ]$. For each h^t, take the limit of agent $_i$'s value at h^t this obtained as $\to 0$. This limit must exist as the prior F on $_i$ is atomless. We define this limit as the value of agent $_i$ at h^t from the strategy profile $(\ _1;\ _2)$.

Insofar as continuation values can be calculated as above for all admissible strategy profiles, we can define perfect Bayesian equilibrium. However, this is not guaranteed because measurability needs to be checked at various stages for the integrations to be well-defined that determine continuation values, although failure of measurability would be very rare if at all.[10] Since it is not crucial for the purpose of this article to fully characterize the environments in which the measurability holds for all relevant cases, we proceed by focusing on strategy profiles for which the measurability holds whenever needed to define and establish an equilibrium. This is sufficient for our purpose.

Consider a strategy profile $(\ _1;\ _2)$, together with a profile of posterior beliefs defined for every (admissible) history as explained above. We introduce an "internal-consistency condition": for any plan $(p_i(\cdot|h^t;\ _i); T_i(h^t;\ _i))$ at history h^t of $_i$, the expiration-time $T_i(h^t;\ _i)$ is the first time point in $(t; 1]$ at which player $-i$ is expected to make a jump with a strictly positive probability as calculated from the plans in $_i$ starting from h^t and its simple extensions relative to the posterior $F_i(h^t)$, unless it is prior to the lapse of reaction lag from the latest start of a new plan due to situation update, in which case $T_i(h^t;\ _i)$ is the end of the relevant reaction lag; if player $-i$ is not expected to make a jump with a strictly positive probability during $(t; 1]$, then $T_i(h^t;\ _i) = 1$. A strategy profile $(\ _1;\ _2)$ is said to be "compatible" if every plan satisfies this condition. In this case we also say that $_i$ is compatible with $_i$.

A compatible strategy profile $(\ _1;\ _2)$, together with a belief profile, is a *perfect Bayesian*

[10] A bounded real function is integrable if measurable (Kolmogorov and Fomin, p.298), and existence of non-measurable subsets of real numbers rely on the Axiom of Choice (Wikipedia).

equilibrium (PBE) if for any history h^t, (i) the posterior belief on the players' types are obtained from (σ_1, σ_2) via Bayes rule whenever possible, and (ii) $\sigma_i(\theta_i)$ is optimal for player i of type θ_i in the continuation game among all agent-strategies that are compatible with σ_i.

4 Equilibrium

Consider the following symmetric strategy for both players where $q = (1-\ell)/(\beta h)$ and $\widehat{T} < 1 - \beta$ will be determined later:

a) Until some $T \in (0, \widehat{T} - 2\beta]$, player i switches to C with a flow rate $q > 0$ provided there has been no previous move. That is, the initial plan is $p_i(t|h^0; \theta_i) = 1 - e^{-qt}$ and $T_i(h^0; \theta_i) = T$ for all $\theta_i \in (0, \beta)$ and $i = 1, 2$; and if h^t is a simple extension of h^0 and $s < T$, then $T_i(h^s; \theta_i) = T$ and $p_i(t|h^s; \theta_i) = 1 - e^{-q(t-s)}$ for $t \in (s, T)$.

b) If player i switched to C at $s \leq T$, then player i's new plan is to keep to C until the expiration-time $s + \beta$; and player j's new plan is to keep D until the same expiration-time $s + \beta$ at which point to switch to C with certainty (provided that i did not move between s and $s + \beta$).

c1) If one player switched to C at $s \leq T$ and the other at $s + \beta$, then both players adopt identical new plans (that start at $s + \beta$) according to which they keep C (so long as the other does) until the expiration-time $\widehat{T}$.

c2) If no switch has been made at every $t \leq T$, both players adopt identical new plans (that start at T) according to which they keep D (so long as the other does) until the expiration-time T' at which point both switch to C with certainty, where $T' = T + \frac{\bar{\theta}}{1-\ell} < \widehat{T}$.

c2') If both switch to C at T' without any previous switches, then both players adopt identical new plans (that starts at T') according to which they keep to C (so long as the other does) until the expiration-time $\widehat{T}$.

d) If $\widehat{T}$ is reached through a history described in c1) or in c2'), then both players adopt identical, type-contingent new plans (that start at $\widehat{T}$) according to which either player of type $\theta \in (0, \beta)$ keeps to C until the expiration-time $1 - \phi(\theta) > \widehat{T}$ at which point she switches to D with certainty (provided that the other has not switched since $\widehat{T}$), where $\phi(\cdot)$ is a continuously increasing function to be determined below with $\phi(0) = 0$ and $\phi(\beta) = 1 - \widehat{T}$.

e) If player i switched to D at $s > \widehat{T}$ as per d), then player i's new plan (that starts at s) is to keep to D until the expiration-time 1, and player j's new plan (that starts at s) is to switch to D at $s + \tau_j$.

f) If player i switched to D at $s > \widehat{T}$ as per d) and player j switched to D at a later time $t < 1$, then both players adopt identical new plans (that start at t) according to which both player keep to D until the expiration-time 1.

We now describe off-equilibrium strategy:

g1) If player i switched to C at $s < T$ as per a) and $s + {}^{-}$ has been reached without j having followed suit, then player i's new plan (that starts at $s + {}^{-}$) is to switch back to D at her expiration-time $s + {}^{-} + {}_{i}$ with certainty, and player j's new plan (that starts at $s + {}^{-}$) is to keep to D until her expiration-time 1.

g2) If player i switched to C at $s < T$ as per a) and $s + {}^{-}$ has been reached without j having followed suit, and player i switched back to D at $s + {}^{-} + {}_{i}$, then both players adopt identical new plans (that start at $s + {}^{-} + {}_{i}$) according to which both keep to D until the expiration-time 1.

g3) If history at any t is inconsistent with any of the contingencies described above, then both players' continuation plans dictate that they switch to D as soon as they can, unless they were already playing D, and then keep to D until the end.

Note that a)-g3) fully describe a strategy profile. We now verify that it constitutes a PBE. Given that D is a strictly dominant strategy of the stage game, it is straightforward that g1)–g3) describe, respectively, a PBE of the continuation game off-the-equilibrium path.

Focus on the equilibrium path. By the same token, once one player switches to D at $t > \widehat{T}$, it is a continuation PBE for the other to switch to D as soon as possible, and for both to keep to D until the end as e) and f) prescribe. Next we verify optimality of the remaining part of the strategy profile in reverse time order.

Optimality of d): For notational ease, let $\#(r) = {}^{1}(r)$ denote the type of either player that would switch back to D at $t = 1 - r$, i.e., when the remaining time is r, unless the other player switched back first. We proceed by postulating that $\#(r) < r \Leftrightarrow (\ ) >$ for all $r \in (0; 1 - \widehat{T})$, which we will confirm later for an open set of environments. Consider $t = 1 - r \in (\widehat{T}; 1)$, conditional on the game has been played according to c1) and $(C; C)$ was played up to t (excluding t). By d), the posterior belief of either player's type at $t = 1 - r$ is represented by the cdf $F|_r(\ ) = F(\ )/F(\#(r))$ on $(0; \#(r))$. Therefore, calculated at $t = 1 - r$, the expected continuation payoff of player i's strategy of switching to D at $t + \Delta \in (1 - r; 1)$ unless j switches first, is

(4)

$$\int_{\theta_j=0}^{\vartheta(r\ \Delta)} \Big[\Delta + {}_{j}h + (r - \Delta - {}_{j})`\Big] dF|_r({}_{j}) + \int_{\vartheta(r\ \Delta)}^{\vartheta(r)} \Big[r - ({}_{j}) + \max\{({}_{j}) - {}_{i}; 0\} \cdot `\Big] dF|_r({}_{j}):$$

The first term depicts that if player j's type is below $\#(r - \Delta)$ so that she would not switch to D until $t + \Delta$, then player i receives a payout of 1 until she switches to D at $t + \Delta$, then player j follows suit ${}_{j}$ later (during which time player i receives h), and then i receives ` from $(D; D)$ for the remaining period. Note that j follows suit before the game ends

because we postulated that $\ (\ ) >\ $. The second term captures that if player j's type is above $\#(r-\Delta)$ so that player j switches to D at $1-\ (\ _j) < t+\Delta$, then player i receives 1 from $(C;C)$ during $(1-r;1-\ (\ _j))$, then 0 for a duration of $\ _i$ until i also switches to D in response to j's switch, and then receives ` from $(D;D)$ for the remaining period.

From $F|_r(\ _j) = F(\ _j)\text{=}F(\#(r))$, the first term of (4) is

$$\frac{1}{F(\#(r))}\int_{\theta_j=0}^{\vartheta(r\ \ \Delta)}\Big[\Delta(1-`)+\ _j(h-`)+r`\Big]f(\ _j)\,d\ _j$$

and thus, its first derivative wrt Δ is

$$\frac{1}{F(\#(r))}\Big[(1-`)\Big(F(\#(r-\Delta))-\Delta f(\#(r-\Delta))\#^0(r-\Delta)\Big)-\Big(\#(r-\Delta)(h-`)+r`\Big)f(\#(r-\Delta))\#^0(r-\Delta)\Big]:$$

Similarly, for Δ small so that $\ (\ _j)-\ _i > 0$, the first derivative of the second term of (4) wrt Δ is

$$\frac{1}{F(\#(r))}\Big[\Delta+\big(r-\Delta-\ _i\big)`\Big]f(\#(r-\Delta))\#^0(r-\Delta):$$

Summing the two up and rearranging, the first derivative of (4) wrt Δ is

(5) $$\frac{1}{F(\#(r))}\Big[(1-`)F(\#(r-\Delta))-\big(\#(r-\Delta)(h-`)+\ _i`\big)f(\#(r-\Delta))\#^0(r-\Delta)\Big]:$$

Evaluating this at $\Delta = 0$, we get

(6) $$\frac{1}{F(\#(r))}\Big[(1-`)F(\#(r))-\big(\#(r)(h-`)+\ _i`\big)f(\#(r))\#^0(r)\Big]:$$

Note that this strictly decreases in $\ _i$. Therefore, if (6) assumes 0 at $\ _i = \#(r)$, i.e.,

(7) $$(1-`)F(\#(r)) = \#(r)hf(\#(r))\#^0(r);$$

then either player of type below $\#(r)$ would strictly prefer to wait at $t = 1-r$. Consequently, the function $\#(\cdot)$ that solves the differential equation (7) is a candidate for equilibrium $\#$. The additional condition for it to satisfy is that (5) decreases in Δ when $\ _i = \#(r)$. For example, this is satisfied under uniform prior, i.e., when $F(\ ) = \ \text{=}\bar{\ }$: the solution to (7) is $\#(r) = r(1-`)\text{=}h$ and therefore, the expression in the bracket of (5) when $\ _i = \#(r)$ is calculated to be $-\Delta(1-`)^2`\text{=}(h^2\,\bar{\ })$ as desired.

Actually, we also need to take care of large Δ such that $r-\Delta-\#(r) < 0$ so that $\max\{\ (\ _j)-\ _i;0\} = 0$ for some $\ _j$ in (4): since the argument above implies i) that a player of type $\#(r)$ finds it suboptimal to wait until $t+\Delta^0$ such that $r-\Delta^0-\#(r) = 0$, and ii) at $t+\Delta^0 = 1-(r-\Delta^0)$ a player of type $\#(r-\Delta^0)$ would strictly prefer not to wait further and, therefore, neither does a player of type $\#(r)$ (because the payoff from switching to D is the same for both types while that from waiting is lower for higher types who is slow to respond when the other player switches to D), it follows that a player of type $\#(r)$ finds it suboptimal to wait past $t+\Delta^0$, either.

In conclusion, if $\#(\cdot)$ solves the differential equation (7) and (5) decreases in Δ when $_i = \#(r)$, then d) is optimal when $\;\; = \#^{\;1}$. We have seen that when F is uniform then $(\;) = \frac{h}{1\;\ell} \;>\;$ as postulated. Thus, this postulation is verified at minimum for an open set of F including the uniform distribution.

Optimality of c1): If either player switches to D at some point $t \in (s + {}^{-}; \widehat{T})$, then compared with the strategy of switching at $\widehat{T}$, it reduces the duration of $(C; C)$ by $\widehat{T} - t$ and increases the duration of $(D; D)$ by the same length. Therefore, it is optimal for either player to keep C at least until $\widehat{T}$ so long the other player does the same.

Optimality of b): It is optimal for j to switch to C at $s + {}^{-}$ because otherwise $(D; D)$ will prevail during $(s + \;_i; 1)$ as per g1), which is worse. Given this, it is optimal for i to keep C until $s + {}^{-}$ because if she switches back to D at $t < s + {}^{-}$ then $(D; D)$ will prevail during $(t + \;_j; 1)$ as per g3), which is worse.

Optimality of a): At any $t < T$, player i's continuation payoff of switching to C at $t + \Delta < T$, unless j switched first, is (see cntnsPD-1.nb)

$$(8) \quad \int_0^{\Delta} qe^{\;q\tau}\left(\;` + {}^{-}h + (\widehat{T} - t - \;\; - {}^{-}) + V_{\widehat{T}}\right)d\;\; + e^{\;q\Delta}\left(\Delta` + (\widehat{T} - t - \Delta - {}^{-}) + V_{\widehat{T}}\right)$$
$$= \frac{e^{\;q\Delta}}{q}\left(1 - ` - {}^{-}qh - e^{q\Delta}(1 - ` - {}^{-}q(h-1) - q(\widehat{T} - t))\right) + V_{\widehat{T}}:$$

The first term of (8) is for the contingency that player j switches at $t + \;\; < t + \Delta$, in which case i receives $`$ during $(t; t + \;\;)$, then h until ${}^{-}$ later when she also switches to D, and then 1 from $(C; C)$ until $\widehat{T}$, followed by the continuation payoff from $\widehat{T}$, denoted by $V_{\widehat{T}}$. Similarly, the second term is for the alternative contingency that player i switches first to D at $t + \Delta$. The derivative of (8) wrt Δ is

$$-e^{\;q\Delta}(1 - ` - {}^{-}qh) = 0$$

for all Δ because $q = (1 - `)=({}^{-}h)$. Therefore, either player is indifferent between switching to D at $t < T$ and waiting until $t + \Delta < T$, justifying the strategy described in a).

Optimality of c2): The expected payoff of either player from switching to C at $t = T$ without previous moves, is $\widehat{T} - T - {}^{-}$ from evaluating (8) at $\Delta = 0$ and $t = T$. That from not switching at T but switching at T^0 along with the other player is $(T^0 - T)` + (\widehat{T} - T^0) + V_{\widehat{T}}$. The two options are equivalent if $T^0 = T + \frac{\bar{\theta}}{1\;\ell} < \widehat{T}$ (ensure the inequality holds). Switching at some $t = (T; T^0)$, or not switching at T^0 is worse because then $(D; D)$ would prevail for at least $(t + {}^{-}; 1)$ and $(T^0 + {}^{-}; 1)$, respectively.

Optimality of c2'): The same argument applies as that given for optimality of c1) above.

This equilibrium can be extended to asymmetric players.

Suppose the two players are different: h_i and $`_i$ are the utility levels of player $i = 1; 2$ for relevant state. In addition, let F and f represent the distribution of player 1's type on

$(0; \bar{\ }_1)$ and let G and g represent the distribution of player 2's type on $(0; \bar{\ }_2)$. Let $\#_i(r)$ denote player i's type who defects at $t = 1 - r$, for $i = 1; 2$.

Then, the marginal net gain of player 2 from infinitesimal wait before defecting, i.e., formula (6) in the symmetric case, becomes

$$\frac{1}{F(\#_1(r))}\Big[(1 - `_2)F(\#_1(r)) - \big(\#_1(r)(h_2 - `_2) + {}_2`_2\big) f(\#_1(r))\#_1^0(r)\Big]: \tag{9}$$

Thus, the equilibrium condition that this value is 0 at ${}_2 = \#_2(r)$ becomes

$$(1 - `_2)F(\#_1(r)) = \big(\#_1(r)(h_2 - `_2) + \#_2(r)`_2\big) f(\#_1(r))\#_1^0(r): \tag{10}$$

The corresponding condition for player 1 is

$$(1 - `_1)G(\#_2(r)) = \big(\#_2(r)(h_1 - `_1) + \#_1(r)`_1\big) g(\#_2(r))\#_2^0(r): \tag{11}$$

We need to find the solution to the system of differential equations (10) and (11).[11] Setting $a_i = (1 - `_i)=`_i$ and $b_i = (h_i - `_i)=`_i$ for $i = 1; 2$, the system of differential equations becomes:

$$a_2F(\#_1(r)) = \Big(b_2\#_1(r) + \#_2(r)\Big) f(\#_1(r))\#_1^0(r); \quad 0 < a_2 < b_2 \tag{12}$$

$$a_1G(\#_2(r)) = \Big(b_1\#_2(r) + \#_1(r)\Big) g(\#_2(r))\#_2^0(r); \quad 0 < a_1 < b_1 \tag{13}$$

The solution to this system, if exists, can be illustrated in a *phase plane*, treating $\#_1$ in the x-axis and $\#_2$ in the y-axis. For notational ease, set $\#_1(r) = x(t)$ and $\#_2 = y(t)$. Then, (12) and (13) become

$$\dot{x} = \frac{a_2}{(b_2x + y)}\frac{F(x)}{f(x)} \quad \text{and} \quad \dot{y} = \frac{a_1}{(b_1y + x)}\frac{G(y)}{g(y)}: \tag{14}$$

We may set

$$F(x) = (x)x \quad \text{and} \quad G(y) = (y)y:$$

Then, the system (14) becomes

$$\dot{x} = \frac{a_2x}{(b_2x + y)}\frac{(x)}{f(x)} \quad \text{and} \quad \dot{y} = \frac{a_1y}{(b_1y + x)}\frac{(y)}{g(y)}: \tag{15}$$

For any $y \geq 0$, the first equation determines a solution $x(t|y)$ and for any $x \geq 0$, the second equation determines a solution $y(t|x)$. These determine the gradient vectors at every point $(x; y) > 0$ in the first quadrant of the phase plane. We assume that F and G satisfy

$$\lim_{x\,!\,0}\frac{(x)}{f(x)} > 0 \quad \text{and} \quad \lim_{y\,!\,0}\frac{(y)}{g(y)} > 0: \tag{16}$$

[11] These are *autonomous* DE's in the sense that the functions to solve for are in one variable only, i.e., time. I gather that these are relatively well-behaved.

This condition is widely satisfied so long as F and G do not vanish to an infinite order at 0,[12] including $F(x) = x^n$ for all $0 \leq n < \infty$.

Since $\dot{x} > 0$ and $\dot{y} > 0$ for all $(x; y) \gg 0$, starting from any initial point $(x_0; y_0) \gg 0$, the solution to (15) is unique and, as time goes back,

(i) will reach $(0; 0)$ within finite time, and

(ii) does not hit either axis until before reaching $(0; 0)$.

The conclusion (i) follows because $\dot{x}$ and $\dot{y}$ are bounded away from 0 by (16),[13] and (ii) follows because if it hit either axis, then the solution starting from that point would be unique and move along that axis in either direction of time, contradicting the supposition that the axis were reached from outside of the axis.

Let $\bar{x}(t)$ and $\bar{y}(t)$ be the solution to (15) from the starting point $(\bar{\ }_1; \bar{\ }_2)$. Let $t_0 < 0$ be such that $\bar{x}(t_0) = \bar{y}(t_0) = 0$. Then, $\#_1(r) = \bar{x}(t_0 + r)$ and $\#_2(r) = \bar{y}(t_0 + r)$ constitute a unique solution to the system (12) and (13).

References

Bergin, James, and W. Bentley MacLeod. 1993. "Continuous Time Repeated Games." *International Economic Review* **34**: 21–37.

Kreps, David M., Paul Milgrom, John Roberts, and Robert Wilson. 1982. "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma." *Journal of Economic Theory* **27**: 245–252.

Friedman, Daniel, and Ryan Oprea. 2012. "A Continuous Dilemma." *American Economic Review* **102**: 337–363.

Radner, Roy. 1986. "Can Bounded Rationality Resolve the Prisoners Dilemma?" In *Contributions to Mathematical Economics,* edited by Andreu Mas-Colell and Werner Hildenbrand, 387–399. Amsterdam: North-Holland.

Simon, Leo K. and Maxwell B. Stinchcombe. 1989. "Extensive Form Games in Continuous Time: Pure Strategies." *Econometrica* **57**: 1171-1214.

[12] An example that vanishes to an in nite order at 0 is $F(x) = e^{-1/x}$ if $x > 0$ and $F(0) = 0$. Then, $\xi(x)/f(x) = x \to 0$ as $x \to 0$.

[13] To be complete: If $y(t) \leq x(t)$ for some t and thus $y(t) = \gamma(t)x(t)$ with $\gamma(t) \leq 1$, then $ax/(bx + y) = a/(b + \gamma(t)) \geq a/(b + 1)$. So, this coe cient is bounded away from 0 and thus $\dot{x}$ is bounded away from 0. Thus, as long as $y \leq x$, $x(t)$ keeps going down as t decreases at a rate bounded away from 0. It must get to 0 (or become smaller than $y(t)$ in a nite time). Similarly, if $x(t) \leq y(t)$, $y(t)$ keeps going down as t decreases. One could say that perhaps it cross the line $y = x$ in nitely many times, thus never getting to $(0, 0)$, but by above it can spend only a nite amount of time on each side of the line $y = x$.