Ambiguity in Multi-Stage Games: Uncertainty about the Opponent's Strategic Behaviour*

Jürgen Eichberger
Alfred Weber Institut,
Universität Heidelberg, Germany.

Simon Grant
Research School of Economics,
Australian National University, Australia.

David Kelsey
Department of Industrial Economics
Nottingham University Business School.

6th July 2026

Abstract

We propose Consistent-Planning Equilibrium under Ambiguity (CP-EUA), a solution concept for two-player multi-stage games with almost perfect information. Players evaluate continuation payoffs using a neo-additive representation centred on a behavioural-strategy theory of the opponent. The additive component is identified with the opponent's behavioural strategy, while the ambiguity component captures the player's lack of confidence in that theory. The lack-of-confidence parameter is revised by Generalized Bayesian Updating (GBU). Players take account of possible changes in their future ambiguity-sensitive evaluations by using consistent planning. We show that, in the centipede game, ambiguity about the opponent's strategic behaviour together with sufficient optimism can sustain cooperation for many periods. Similarly, in a non-cooperative bargaining game, ambiguity and optimism may generate delay before agreement is reached.

Keywords: optimism, neo-additive beliefs, extensive-form games, dynamic consistency, consistent planning, centipede game, bargaining.

JEL classification: D81

Corresponding Author Prof. em. Dr. Jürgen Eichberger, Alfred Weber Institute, Department of Economics, Heidelberg University, Bergheimer Str. 58, D-69115 HEIDELBERG, GERMANY
e-mail juergen.eichberger@awi.uni-heidelberg.de

*For comments and suggestions we would like to thank Pierpaolo Battigalli, Lorenz Hartmann, Philippe Jehiel, David Levine, George Mailath, Larry Samuelson, Marciano Siniscalchi, Ron Stauber, Rabee Tourky, Qizhi Wang, as well as participants in presentations at Adelaide, ANU, Bristol, Exeter, Nottingham, FUR (Warwick), DTEA (Paris), SAET (Faro), RUD (Heidelberg), Dynamic Models in Economics Workshop on Game Theory (NUS, Singapore), and finally the referees and editor of this journal.

1 Introduction

Ambiguity—uncertainty without known probabilities—has been studied extensively in single-person decision problems. In many economic environments, however, the relevant uncertainty is strategic. A player may be unsure how an opponent will behave, and this uncertainty may not be well summarized by a single fully trusted probability assessment. This issue is especially important in dynamic games, where actions observed along the way may change how a player evaluates the opponent's likely future behaviour.

The resulting dynamic problem is not the familiar one of changing tastes, non-exponential discounting, or ordinary Bayesian learning under expected utility. It arises because updated preferences may themselves be ambiguity-sensitive. Once a history is observed, a player may revise both her theory about the opponent's behaviour and her confidence in that theory. If she has a non-neutral attitude toward the ambiguity she perceives around the theory, then the updated evaluation of continuation actions need not agree with the evaluation implicit in her earlier plan. An action that was optimal as part of an ex ante strategy may therefore fail to be optimal once the relevant history is reached.

This paper proposes a solution concept for finite two-player multi-stage games with almost perfect information.[1] The concept, Consistent Planning Equilibrium under Ambiguity (CP-EUA), extends Eichberger and Kelsey's (2014) Equilibrium under Ambiguity for static games to dynamic strategic environments. Beliefs are updated using Generalized Bayesian Updating (GBU), following Eichberger et al. (2007), and dynamic choice is disciplined by consistent planning in the sense of Strotz (1955) and Siniscalchi (2011).[2]

The ambiguity in the model concerns the opponent's behaviour. A player has a theory about how the opponent will behave, but need not fully trust that theory. The model separates this lack of confidence from the player's attitude toward the ambiguity surrounding it. Thus the opponent's behaviour is the strategic object about which the player is uncertain; the player's perceived ambiguity reflects her confidence in the theory she holds about that behaviour; and her ambiguity attitude determines how pessimistically or optimistically she evaluates the consequences of that perceived ambiguity. In equilibrium, each player's theory about the opponent is required to coincide with the opponent's equilibrium behaviour.

A player's own behavioural strategy is treated differently. It is her plan for how she will act at histories at which she may be called upon to move. When she evaluates a current action, her own continuation plan is held fixed. A randomization over her current actions is evaluated as an ordinary randomization over the continuation payoffs associated with the actions realized. Once

[1] Osborne and Rubinstein (1994) (p102) refer to this class as extensive games with perfect information and simultaneous moves.

[2] Bose and Daripa (2009) employ a similar notion of consistent planning for their equilibrium with players whose preferences have a multiple-prior representation. Bose and Renou (2014) and Karni and Safra (1989) use versions of consistent planning in games.

an action has been realized, the player still faces ambiguity about the opponent's behaviour; but she does not face ambiguity about the probabilities generated by her own randomizing device. This separation is a central feature of the approach. It keeps strategic ambiguity focused on the behaviour of the opponent, while preserving the standard interpretation of a behavioural strategy as the player's own contingent plan.

The updating rule is chosen to keep this distinction clear. After a history is observed, the player updates her theory about the opponent and her confidence in that theory, while her underlying attitude toward ambiguity remains fixed. A surprising opponent-history lowers confidence in the theory and increases the weight placed on the ambiguity surrounding it. If the opponent's observed behaviour has zero likelihood under the theory, and the player initially perceived some ambiguity, the updated evaluation becomes one of complete ambiguity over the opponent's continuation behaviour. This feature is useful in dynamic games, where histories that were not anticipated by a player's theory must still be evaluated.

Consistent planning then provides the dynamic discipline. The player does not mechanically commit to a strategy that is optimal from the ex ante perspective. Instead, she anticipates how her future selves will evaluate later choices after updating. At each history, she compares the actions currently available, holding fixed her own continuation plan and the opponent's behavioural strategy. A behavioural strategy is a consistent plan if every action assigned positive probability is optimal under the corresponding updated continuation evaluation. CP-EUA requires both players' behavioural strategies to be consistent plans given the opponent's behavioural strategy.

When there is no ambiguity, CP-EUA reduces to ordinary expected-payoff optimality after every history. By the one-step-deviation principle for finite dynamic games, it then coincides with subgame perfect equilibrium (SPE). With positive ambiguity, the same sequential logic is applied using updated ambiguity-sensitive continuation payoffs. The concept therefore preserves the backward-induction discipline of dynamic games while allowing histories to affect perceived ambiguity about the opponent's future behaviour.

We study two applications. The first is the centipede game, where SPE predicts immediate termination but experimental subjects often continue for several rounds. We show that, when players perceive ambiguity about the opponent's behaviour and are sufficiently optimistic, CP-EUA can sustain cooperation for many periods. The second application is an alternating-offer bargaining game. There, ambiguity and optimism can generate inefficient delay before agreement is reached.

The analysis highlights a channel that is absent from standard backward induction. Under subgame perfection, actions by the opponent that are not sequentially optimal are ruled out. Under strategic ambiguity, a player need not place full confidence in a single prediction of the opponent's continuation behaviour. Updating then changes the player's confidence in that prediction, and consistent planning determines how the player acts when her later evaluations differ from the evaluation implicit in the original plan. This provides a way to analyse dynamic games in which players are sophisticated about their own future updating but remain ambiguity-sensitive about the opponent's behaviour.

Organization of the Paper Section 2 defines the class of multi-stage games and behavioural strategies. Section 3 introduces strategic ambiguity, updating, and conditional continuation payoffs. Section 4 defines CP-EUA. Section 5 establishes basic properties. Sections 6 and 7 analyse the centipede and bargaining applications. Section 8 discusses related literature, and Section 9 concludes.

2 Multi-stage Games

2.1 Behavioural Strategies and Histories

We consider finite two-player multi-stage games with public histories. There are two players, $i = 1, 2$, and finitely many stages $t = 1, \dots, T$. At each stage, both players observe the history of previous actions and then choose actions simultaneously.

Let H^t denote the set of histories of length t, that is, histories containing t previously chosen action profiles. The empty history is denoted by h^0, so $H^0 = \{h^0\}$. A history at the beginning of stage t has length $t-1$ and is written

$$h = \langle a^1, \dots, a^{t-1} \rangle \in H^{t-1},$$

where $a^\tau = \langle a_1^\tau, a_2^\tau \rangle$ is the action profile chosen at stage τ. For each earlier stage $\tau = 1, \dots, t-1$, let

$$h^{\tau-1} := \langle a^1, \dots, a^{\tau-1} \rangle$$

denote the history immediately before the action profile a^τ is chosen.

Set

$$H := \bigcup_{t=0}^{T-1} H^t, \qquad Z := H^T, \qquad \overline{H} := H \cup Z.$$

Thus H is the set of nonterminal histories, Z is the set of terminal histories, and $\overline{H}$ is the set of all histories. A generic terminal history is typically denoted by z.

At each nonterminal history $h \in H$, player i has a finite nonempty set of feasible actions A_i^h. Write $A^h := A_1^h \times A_2^h$ for the set of feasible action profiles at h. If $a \in A^h$, then $\langle h, a \rangle$ denotes the history obtained by appending the action profile a to h.

A terminal history $z \in Z$ is identified with the outcome generated by that play of the game.

Player i's material payoff is a function $u_i : Z \to \mathbb{R}$. The game form with material payoffs is

$$\Gamma = \left\langle \{1,2\}, \overline{H}, (u_1, u_2) \right\rangle . \tag{2.1}$$

A *pure strategy* for player i selects a feasible action after every nonterminal history. Formally, it is a function

$$s_i : H \to \bigcup_{h \in H} A_i^h$$

such that $s_i(h) \in A_i^h$ for every $h \in H$. Let S_i denote the finite set of pure strategies of player i. Each pure-strategy profile $s = \langle s_1, s_2 \rangle \in S_1 \times S_2$ generates a unique terminal history in Z. When no confusion can arise, we identify s with this terminal history.

A *behavioural strategy* for player i is a collection of probability distributions

$$\beta_i = (\beta_i(h))_{h \in H},$$

where $\beta_i(h) \in \Delta(A_i^h)$ for each $h \in H$. Let B_i denote the set of behavioural strategies of player i. A behavioural-strategy profile is written

$$\beta = \langle \beta_1, \beta_2 \rangle \in B_1 \times B_2.$$

A degenerate behavioural strategy assigns probability one to a single action at each history. We identify such behavioural strategies with pure strategies. Thus pure strategies are the deterministic special cases of the behavioural strategies in B_i.

For any behavioural-strategy profile $\beta = (\beta_1, \beta_2)$ and any history $h = \langle a^1, \ldots, a^t \rangle \in \overline{H}$, define the probability of reaching h under β by

$$P_\beta(h) := \prod_{\tau=1}^{t} \beta_1(h^{\tau-1})[a_1^\tau]\, \beta_2(h^{\tau-1})[a_2^\tau], \tag{2.2}$$

with $h^{\tau-1}$ as defined above. For the empty history h^0, $P_\beta(h^0) = 1$. Thus $P_\beta(h)$ is the probability that the play generated by β reaches h.

2.2 Continuation Games and Histories

For a non-terminal history $h \in H$, a continuation history after h is a sequence of action profiles h' such that $\langle h, h' \rangle \in \overline{H}$. Let H^h and Z^h denote the sets of non-terminal and terminal continuation histories after h, respectively. For a terminal continuation $z' \in Z^h$, define

$$u_i^h(z') := u_i(\langle h, z' \rangle).$$

Thus u_i^h is player i's payoff function in the continuation of the game after h.

For any behavioural-strategy profile $\beta = \langle \beta_i, \beta_{-i} \rangle$, let $P^h_\beta \in \Delta(Z^h)$ denote the probability distribution over terminal continuation histories induced by β after h. If $z' = \langle a^1, \ldots, a^m \rangle \in Z^h$, then

$$P^h_\beta(z') = \prod_{\tau=1}^{m} \beta_i(\langle h, a^1, \ldots, a^{\tau-1} \rangle)[a^\tau_i]\, \beta_{-i}(\langle h, a^1, \ldots, a^{\tau-1} \rangle)[a^\tau_{-i}], \tag{2.3}$$

with the convention that $\langle h, a^1, \ldots, a^0 \rangle = h$.

Player i's ordinary expected continuation payoff after h can therefore be expressed,[3]

$$U^h_i(\beta_i, \beta_{-i}) := \sum_{z' \in Z^h} u^h_i(z')\, P^h_\beta(z'). \tag{2.4}$$

Only the continuations of β_i and β_{-i} after h are payoff-relevant in this expression.

3 Strategic Ambiguity and Conditional Evaluation

The game form specifies what players can do and how terminal histories are evaluated. It remains to describe how a player evaluates uncertainty about the opponent's behaviour. Player i's "theory" about her opponent is represented by a behavioural strategy of the opponent. In equilibrium this theory coincides with the opponent's equilibrium behavioural strategy.

To model the ambiguity player i perceives around this theory, we use the neo-additive evaluation of Chateauneuf et al. (2007). The parameter $\delta_i \in [0, 1]$ measures the player's lack of confidence in her theory, while $\alpha_i \in [0, 1]$ measures how pessimistically she views the ambiguity surrounding it. The case $\alpha_i = 1$ corresponds to maximal pessimism, while $\alpha_i = 0$ corresponds to maximal optimism.

For any payoff function $f : B_{-i} \to \mathbb{R}$ that is affine in each of the opponent's local randomizations, the neo-additive evaluation around the behavioural theory β_{-i} is

$$(1 - \delta_i) f(\beta_{-i}) + \delta_i \left[\alpha_i \min_{\gamma_{-i} \in B_{-i}} f(\gamma_{-i}) + (1 - \alpha_i) \max_{\gamma_{-i} \in B_{-i}} f(\gamma_{-i}) \right]. \tag{3.1}$$

Here the first term evaluates f at the player's theory β_{-i}, while the second term captures the ambiguity surrounding that theory.

Since B_{-i} is a product of finite simplexes and f is affine in each of the opponent's local randomizations, the minimum and maximum are attained at extreme points of B_{-i}. These extreme points correspond to degenerate behavioural strategies, equivalent to pure strategies.

The evaluation in (3.1) is the ex ante evaluation. In a dynamic game, the same form is applied after histories, with confidence updated in light of the opponent's observed behaviour. For a history

$$h = \langle a^1, \ldots, a^t \rangle \in \overline{H}, \qquad a^\tau = \langle a^\tau_i, a^\tau_{-i} \rangle,$$

[3] By ordinary we mean before taking into account the effects of strategic ambiguity.

define the likelihood, according to player i's theory β_{-i}, of the opponent's component of the observed history by

$$L_i(h;\beta_{-i}) := \prod_{\tau=1}^{t} \beta_{-i}(h^{\tau-1})[a_{-i}^{\tau}], \tag{3.2}$$

with $L_i(h^0;\beta_{-i}) = 1$. The likelihood $L_i(h;\beta_{-i})$ differs from the full reach probability $P_\beta(h)$. The latter uses both players' behavioural strategies; the former uses only the opponent's component of the history. Thus a history may have zero full reach probability because of player i's own earlier behaviour even though the opponent's observed behaviour has positive likelihood.

We update ambiguity by Generalized Bayesian Updating. For neo-additive evaluations, GBU preserves the neo-additive form, leaves the ambiguity-attitude parameter α_i unchanged, and updates δ_i according to the likelihood of the observed event.[4] In the present behavioural-strategy formulation, the updated ambiguity parameter after h is

$$\delta_i^h(\beta_{-i}) = \frac{\delta_i}{\delta_i + (1-\delta_i)L_i(h;\beta_{-i})}, \quad \text{if } L_i(h;\beta_{-i}) > 0. \tag{3.3}$$

If $L_i(h;\beta_{-i}) = 0$ and $\delta_i > 0$, set

$$\delta_i^h(\beta_{-i}) = 1.$$

In that case the theory has assigned zero likelihood to the opponent's component of the observed history.

The formula captures the role of surprise. The updated ambiguity parameter is larger when the observed opponent-history has lower likelihood. At the same time, conditioning on a history restricts attention to continuations after that history. Thus updating has two effects: it changes confidence in the player's theory and it determines the continuation problem to be evaluated. Hence an increase in the ambiguity weight should not be read as an expansion of the continuation problem; rather, updating both restricts the continuation problem and changes the weight placed on ambiguity within that problem.

We can now evaluate the actions available to player i at a history. The relevant comparison is local: what happens if player i chooses a particular action now, while retaining her continuation plan thereafter?

For each action $a_i \in A_i^h$, let $\langle \beta_i^{-h}, a_i \rangle$ denote the behavioural strategy that agrees with β_i at every history other than h, and assigns probability one to a_i at h. Formally,

$$\langle \beta_i^{-h}, a_i \rangle(h')[a'] := \beta_i(h')[a'], \qquad h' \neq h,$$

[4] See Eichberger et al. (2007) and Proposition 1 in Eichberger et al. (2010). Alternative rules such as Dempster–Shafer updating and optimistic updating are also used for ambiguous beliefs, but for neo-additive evaluations they force the updated ambiguity-attitude parameter to the extreme values $\alpha_i = 1$ and $\alpha_i = 0$, respectively. GBU is therefore better suited to the present analysis, where α_i represents a stable attitude toward ambiguity while δ_i changes with the likelihood of the observed history.

and

$$\langle \beta_i^{-h}, a_i \rangle (h)[a'] := \begin{cases} 1, & a' = a_i, \\ 0, & a' \neq a_i. \end{cases}$$

Thus $U_i^h(\langle \beta_i^{-h}, a_i \rangle, \gamma_{-i})$ is the ordinary expected continuation payoff when player i plays a_i at h, thereafter follows β_i, and the opponent follows γ_{-i}.

Given the opponent's behavioural strategy β_{-i}, the conditional neo-additive continuation payoff associated with choosing a_i at h and following β_i thereafter is

$$\begin{aligned} V_i^h(a_i; \beta_i, \beta_{-i}) = {} & \Big(1 - \delta_i^h(\beta_{-i})\Big) \, U_i^h\left(\langle \beta_i^{-h}, a_i \rangle, \beta_{-i}\right) \\ & + \delta_i^h(\beta_{-i}) \Bigg[\alpha_i \min_{s_{-i} \in S_{-i}} U_i^h\left(\langle \beta_i^{-h}, a_i \rangle, s_{-i}\right) \\ & \qquad + (1 - \alpha_i) \max_{s_{-i} \in S_{-i}} U_i^h\left(\langle \beta_i^{-h}, a_i \rangle, s_{-i}\right) \Bigg]. \end{aligned} \tag{3.4}$$

where each pure strategy $s_{-i} \in S_{-i}$ is identified with the corresponding degenerate behavioural strategy.[5] If $L_i(h; \beta_{-i}) = 0$ and $\delta_i > 0$, then $\delta_i^h(\beta_{-i}) = 1$, and (3.4) reduces to the Hurwicz evaluation

$$\alpha_i \min_{s_{-i} \in S_{-i}} U_i^h(\langle \beta_i^{-h}, a_i \rangle, s_{-i}) + (1 - \alpha_i) \max_{s_{-i} \in S_{-i}} U_i^h(\langle \beta_i^{-h}, a_i \rangle, s_{-i}).$$

Thus complete ignorance is triggered by a zero-likelihood opponent-history, not merely by a zero full reach probability.

The payoff $V_i^h(a_i; \beta_i, \beta_{-i})$ is therefore the action-contingent continuation value relevant at history h. It evaluates the realized current action a_i, holding fixed player i's continuation plan β_i and the opponent's behavioural strategy β_{-i}. The next section uses these local continuation values to define consistent planning.

4 Consistent-Planning Equilibrium under Ambiguity

4.1 Consistent Plans

At a history h, the relevant choice is local. Player i compares the actions currently available, holding fixed her continuation plan β_i and the opponent's behavioural strategy β_{-i}. Thus the relevant objects are the action-contingent continuation values $V_i^h(a_i; \beta_i, \beta_{-i})$ defined in (3.4).

If player i uses a behavioural randomization $\xi_i \in \Delta(A_i^h)$ at history h, this randomization is resolved before the selected action is implemented. Once an action a_i has been selected, player i still faces ambiguity about the opponent's behaviour, but not about the probabilities generated

[5] Equivalently, the extrema in the ambiguity term could be written over B_{-i} as in (3.1) above; by the extreme-point argument above. However, it follows that the values are unchanged when the extrema are taken over the pure strategies S_{-i}, identified with degenerate behavioural strategies.

by her own randomizing device. The value of ξ_i is therefore

$$\sum_{a_i \in A_i^h} \xi_i[a_i]\, V_i^h(a_i; \beta_i, \beta_{-i}).$$

This ordering of own randomization and ambiguity is in line with the representation studied by Monet and Vergopoulos (2024): own randomization is evaluated outside the ambiguity operator, while the ambiguity concerns the opponent's behaviour.

Linearity in ξ_i implies that an optimal behavioural randomization can assign positive probability only to actions that maximize $V_i^h(\cdot; \beta_i, \beta_{-i})$. This is not an assumption that an ex ante mixture over complete strategies is behaviourally equivalent to a pure strategy under ambiguity. Consistent planning evaluates choices at the histories at which they are made. A local randomization cannot serve as a commitment device that induces the player to implement an action she would not choose once that action is selected. Future randomizations may enter the continuation plan β_i, and hence affect $V_i^h(a_i; \beta_i, \beta_{-i})$, but they are credible only if the actions in their support are themselves locally optimal at the histories where those randomizations are used. This gives the consistency requirement.

Definition 4.1 *Given $\beta_{-i} \in B_{-i}$, a behavioural strategy $\beta_i \in B_i$ is a* consistent plan *for player i if, for every non-terminal history $h \in H$, every action $a_i \in A_i^h$, and every alternative action $a_i' \in A_i^h$,*

$$\beta_i(h)[a_i] > 0 \quad \Longrightarrow \quad V_i^h(a_i; \beta_i, \beta_{-i}) \geqslant V_i^h(a_i'; \beta_i, \beta_{-i}). \tag{4.1}$$

4.2 CP-EUA

A consistent plan is a one-player optimality requirement, taking the opponent's behavioural strategy as given. Equilibrium imposes this requirement on both players simultaneously.[6]

Definition 4.2 *A* Consistent Planning Equilibrium under Ambiguity *(CP-EUA) is a profile of behavioural strategies $\langle \beta_1, \beta_2 \rangle$ such that, for each player $i = 1, 2$, β_i is a consistent plan for player i given the opponent's behavioural strategy β_{-i}.*

If $\delta_1 = \delta_2 = 0$, then

$$V_i^h(a_i; \beta_i, \beta_{-i}) = U_i^h(\langle \beta_i^{-h}, a_i \rangle, \beta_{-i}),$$

so CP-EUA requires ordinary expected-payoff optimality after every history. By the one-step deviation principle for finite dynamic games, CP-EUA then coincides with SPE. For positive ambiguity, the same one-step-deviation requirement is imposed using the updated neo-additive continuation payoff.

[6] Because both the behavioural-strategy theory of the opponent and the player's own plan are specified directly at histories, the definition does not appeal to Kuhn-equivalence between mixed and behavioural strategies, nor to any analogous equivalence under ambiguous conjectures.

4.3 Existence

The existence argument uses the agent-normal-form logic implicit in consistent planning. Each history at which player i may move is treated as a decision point for an agent choosing a distribution over the actions available at that history. The payoff of the agent corresponding to player i at history h is the linear extension of V_i^h over actions at h:

$$\Psi_{i,h}(\xi_i;\beta_i,\beta_{-i}) := \sum_{a_i \in A_i^h} \xi_i[a_i] V_i^h(a_i;\beta_i,\beta_{-i}). \tag{4.2}$$

Applying Nash's theorem to this finite agent-normal-form game yields existence.

Theorem 4.1 *Let Γ be a finite two-player multi-stage game. Then Γ has at least one CP-EUA for any given parameters $\alpha_1, \alpha_2, \delta_1, \delta_2$, where $0 \leqslant \alpha_i \leqslant 1$ and $0 < \delta_i \leqslant 1$, for $i = 1, 2$.*[7]

5 Properties of CP-EUA

The definition of CP-EUA modifies dynamic optimality by replacing expected continuation payoffs with updated ambiguity-sensitive continuation payoffs. The next results record what this change preserves and where it departs from the standard benchmark.

5.1 Upper Hemi-Continuity

The first property is stability of the equilibrium correspondence with respect to the ambiguity parameters.

Definition 5.1 *Consider a set $A \subseteq \mathbb{R}^n$ and a compact set $B \subseteq \mathbb{R}^m$. A correspondence $\phi : A \to \mathcal{P}(B)$ is* upper hemi-continuous *at $x^0 \in A$ if, for every sequence $x^r \to x^0$ and every sequence $y^r \in \phi(x^r)$ with $y^r \to y^0$, we have $y^0 \in \phi(x^0)$.*

The following result establishes that the equilibrium correspondence is upper hemi-continuous in δ and α. This implies that an equilibrium will not suddenly disappear as δ or α changes, however new equilibria may appear.

Proposition 5.1 *Let $\Gamma = \langle \{1,2\}, \overline{H}, (u_1,u_2) \rangle$ be a two-player multi-stage game and let $\mathcal{E}$ denote the set of CP-EUAs of Γ. The equilibrium correspondence is upper hemi-continuous in $\langle \alpha_1, \alpha_2, \delta_1, \delta_2 \rangle$ on $[0,1]^2 \times (0,1]^2$.*

5.2 Vanishing Ambiguity

The behaviour of CP-EUA near full confidence is delicate. For every positive degree of ambiguity, continuation evaluations are defined after all opponent histories. If the opponent-history

[7] If $\delta_i = 0$ for each player, existence follows from the usual existence result for subgame perfect equilibrium in finite games.

has zero likelihood, then $L_i(h; \beta_{-i}) = 0$, and GBU sets $\delta_i^h(\beta_{-i}) = 1$. Thus the continuation payoff is determined by the Hurwicz component of (3.4). For pure pessimism, $\alpha_i = 1$, this corresponds to complete pessimism over possible opponent continuation behaviours; for pure optimism, $\alpha_i = 0$, it corresponds to complete optimism.

This creates a discontinuity at $\delta_i = 0$. With full confidence, if an unforeseen opponent-history occurs, Bayesian updating is not defined. With any positive grain of ambiguity, the update is well-defined and assigns full weight to the ambiguity component. Hence a sequence of CP-EUAs need not converge to a SPE as ambiguity vanishes.

5.3 Dominance

In this section we show that under a mild assumption on payoffs, CP-EUA implies that weakly dominated strategies will not be played. The intuition is as follows. If a player perceives her opponent's action to be ambiguous she cannot rule out the possibility that a weakly dominated strategy may lead to a lower payoff than the strategy which dominates it. Since the extrema in the ambiguity term are attained at degenerate behavioural strategies, dominance can be stated in terms of pure continuation strategies.

Fix a history h. Let S_i^h denote the set of player i's pure continuation strategies after h, identified with the corresponding degenerate behavioural continuation strategies. Given a behavioural strategy β_i, say that $s_i^h \in S_i^h$ is *played after h* under β_i if every action prescribed by s_i^h is assigned positive probability by β_i at the corresponding continuation history.

For $s_i^h \in S_i^h$, define

$$m(s_i^h) := \min_{s_{-i}^h \in S_{-i}^h} u_i^h(s_i^h, s_{-i}^h), \qquad M(s_i^h) := \max_{s_{-i}^h \in S_{-i}^h} u_i^h(s_i^h, s_{-i}^h).$$

Two continuation strategies $\hat{s}_i^h, \tilde{s}_i^h \in S_i^h$ have *different extremes* if

$$m(\hat{s}_i^h) = m(\tilde{s}_i^h) \quad \Longrightarrow \quad M(\hat{s}_i^h) \neq M(\tilde{s}_i^h).$$

A two-player multi-stage game has *different extremes* if all pairs of continuation strategies of each player have different extremes after every history.

Proposition 5.2 *Let Γ be a two-player multi-stage game with different extremes. Suppose $1 \geqslant \delta_i > 0$ and $0 < \alpha_i < 1$ for each player i. In any CP-EUA, no weakly dominated pure continuation strategy is played after any history. In particular, no weakly dominated pure strategy is played in CP-EUA.*

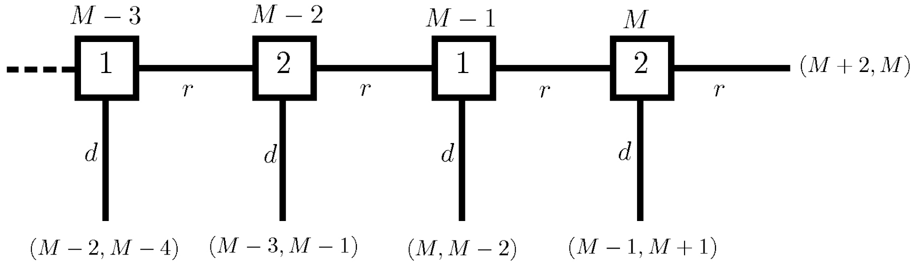

Figure 1: Centipede Game: Last 4 stages

6 The Centipede Game

The centipede game is a natural stress test for any dynamic equilibrium concept. Backward induction predicts immediate termination, while laboratory play often displays substantial continuation. We use the game to show how optimism under strategic ambiguity can sustain cooperation without changing material payoffs. The centipede game was introduced by Rosenthal (1981) and studied in laboratory experiments by McKelvey and Palfrey (1992). A survey of subsequent experimental research can be found in Krokow et al. (2016).[8]

6.1 The Game

The centipede game may be described as follows. There are two people, player 1 (she) and player 2 (he). Between them is a table which contains $2M$ one-pound coins and a single two-pound coin. They move alternately. At each move there are two actions available. The player whose move it is may either pick up the two-pound coin in which case the game ends; or she may pick up two one-pound coins keep one and give the other to her opponent; in which case the game continues. In the final round there is a single two-pound coin and two one-pound coins remaining. Player 2, who has the move, may either pick up the two-pound coin, in which case the game ends and nobody gets the one-pound coins; or he may pick up the two one-pound coins keep one and give the other to his opponent, in which case the opponent also gets the two pound coin and the game ends. We label an action that involves picking up two one-pound coins by r (right) and an action of picking up the two-pound coin by d (down). Figure 1 shows the final four decision nodes. A standard backward induction argument establishes that there is a unique iterated dominance equilibrium. At any node the player, whose move it is, picks up the 2-pound coin and ends the game.[9]

[8] There is a substantial literature (too large to be reviewed here) that, in contrast to our analysis based on ambiguity, appeals to altruistic or intrinsically cooperative preferences in order to explain the experimental results contradicting the predictions of backward induction. Kaneko and Ishikawa (2023), e.g. show that their theory of conscious inertial behaviour (CIB) may also result in substantial cooperation in the centipede game. However instead of introducing ambiguity about the other player's strategy, CIB involves relaxing backward induction and perfect comparability of payoffs.

[9] There are other Nash equilibria. However these only differ from the iterated dominance equilibrium off the equilibrium path.

6.2 Notation

Given the special structure of the centipede game, we can simplify our notation. The tree can be identified with the non-terminal nodes $H = \{1, ..., M\}$. The set of non-terminal nodes can be partitioned into the two player sets $H_1 = \{1, 3, ..., M - 1\}$ and $H_2 = \{2, 4, ..., M\}$. It will be a maintained hypothesis that M is an even number, $M \geqslant 4$.

Strategies and Payoffs. A behavioural strategy for player i is a mapping $\beta_i : H_i \to \Delta\{r, d\}$. A (pure) strategy s_i is a behavioural strategy, which assigns a deterministic action to each history i.e. $s_i : H_i \to \{r, d\}$. Given a strategy combination (s_1, s_2) set $m(s_1, s_2) := 0$ if d is never played, otherwise set $m(s_1, s_2) := m' \in H$ where m' is the first node where action d is played. The payoff of strategy combination (s_1, s_2) is: $u_1(s_1, s_2) = M + 2$, if $m(s_1, s_2) = 0$; $u_1(s_1, s_2) = m(s_1, s_2)+1$ if $m(s_1, s_2)$ is odd and $u_1(s_1, s_2) = m(s_1, s_2)-1$ otherwise. Likewise $u_2(s_1, s_2) = M$, if $m(s_1, s_2) = 0$; $u_2(s_1, s_2) = m(s_1, s_2) + 1$, if $m(s_1, s_2)$ is even; $u_2(s_1, s_2) = m(s_1, s_2) - 1$ otherwise.

6.3 General Results

We now characterize CP-EUA in the centipede game. The parameter restriction is deliberately simple: both players have the same degree of ambiguity and the same attitude toward it. Hence the results can be stated in terms of (α, δ). That is, throughout this section we take Γ to be an M stage centipede game, where both players are neo-expected payoff maximizers with $\delta_1 = \delta_2 = \delta \in [0, 1]$ and $\alpha_1 = \alpha_2 = \alpha \in [0, 1]$.

There are the following possibilities, cooperation continues until the final node, there is no cooperation at any node or there is a mixed equilibrium. The mixed equilibria also involve a substantial amount of cooperation. The first proposition shows that if there is sufficient ambiguity and players are sufficiently optimistic the equilibrium involves playing r (right) until the final node. At the final node Player 2 chooses d (down) since it is a dominant strategy.

Proposition 6.1 *Let Γ be an M-stage centipede game. For $\delta\,(1 - \alpha) \geqslant \frac{1}{3}$, there exists a CP-EUA $\langle \hat{\beta}_1, \hat{\beta}_2 \rangle$ such that $\hat{\beta}_i(m)[r] = 1$, for $1 \leqslant m \leqslant M - 1$, and $\hat{\beta}_2(M)[d] = 1$. This equilibrium will be unique provided the inequality is strict.*

This confirms our intuition. Ambiguity-loving preferences can lead to cooperation in the centipede game. Observe that $\delta\,(1 - \alpha)$ is the decision-weight on the best outcome in the neo-additive evaluation (equation 3.1). Cooperation does not require highly ambiguity loving preferences. A necessary condition for cooperation is that ambiguity aversion is not too high i.e. $\alpha \leqslant \frac{2}{3}$. Such ambiguity-attitudes are not implausible.[10]

[10] Kilka and Weber (2001) experimentally estimate that $\alpha = \frac{1}{2}$.

Recall that players do not cooperate in SPE. We would expect that ambiguity aversion makes cooperation less likely. Ambiguity aversion increases the attractiveness of playing down, which offers a low but ambiguity-free payoff. The next result finds that, provided players are sufficiently ambiguity-averse, non-cooperation at every node is an equilibrium.

Proposition 6.2 *Let Γ be an M-stage centipede game. For $\alpha \geqslant \frac{2}{3}$, there exists a CP-EUA $\langle \hat{\beta}_1, \hat{\beta}_2 \rangle$ such that $\hat{\beta}_i(m)[d] = 1$, for $1 \leqslant m \leqslant M$. This equilibrium will be unique provided $\alpha > \frac{2}{3}$.*

It is perhaps worth emphasising that pessimism must be large in order to induce players to exit at every node. If $\frac{1}{2} < \alpha < \frac{2}{3}$ the players overweight bad outcomes more than they overweight good outcomes. However non-cooperation at every node is not an equilibrium in this case even though players are fairly pessimistic about their opponent's behaviour.

For parameter values not covered by the two previous results the centipede game has no equilibrium in pure behavioural strategies, that is, at least one player will choose a strictly mixed behavioural strategy at one information set $h \in H$. Consequently we need to look for mixed equilibria.

Proposition 6.3 *Let Γ be an M-stage centipede game. Then for $\alpha < \frac{2}{3}$ and $\delta\,(1-\alpha) < \frac{1}{3}, \Gamma$ does not have a CP-EUA in pure behavioural strategies.*

Proposition 6.3 tells us that when $\alpha < \frac{2}{3}$ and $0 < \delta\,(1-\alpha) < \frac{1}{3}$ there is a mixed equilibrium. The general case is complex with a large number of types of equilibria depending on the parameter values. To illustrate the possibilities in the next section we shall study the 4-stage centipede game in detail.

6.4 The 4-stage Centipede Game

The remaining parameter region is where the centipede application is most informative: pure behavioural-strategy equilibria disappear, but the model still delivers sharp mixed-equilibrium predictions.

In the four-stage centipede game, there are four information sets $H = \{1, 2, 3, 4\}$, where players choose their behavioural strategy. At each information set the player who has to move chooses a probability distribution over r (right), β_i^t, and d (down), $(1-\beta_i^t)$.

Hence the behavioural strategy of Player 1 at stages $t = 1$ and $t = 3$ is denoted by $\langle \beta_1^1, \beta_1^3 \rangle$ and for Player 2 at stages $t = 2$ and $t = 4$ by $\langle \beta_2^2, \beta_2^4 \rangle$. These behavioural strategies yield neo-additive payoffs $\langle V_1^1(\beta_1^1, \beta_1^3, \beta_2^2, \beta_2^4), V_1^3(\beta_1^1, \beta_1^3, \beta_2^2, \beta_2^4) \rangle$ and $\langle V_2^2(\beta_1^1, \beta_1^3, \beta_2^2, \beta_2^4), V_2^4(\beta_1^1, \beta_1^3, \beta_2^2, \beta_2^4) \rangle$ for Player 1 and Player 2 at their respective information sets. For notational simplicity, we suppress the dependency of payoffs on the ambiguity parameters α and δ. The absence of a pure behavioural-strategy equilibrium in the remaining region does not mean that the model loses

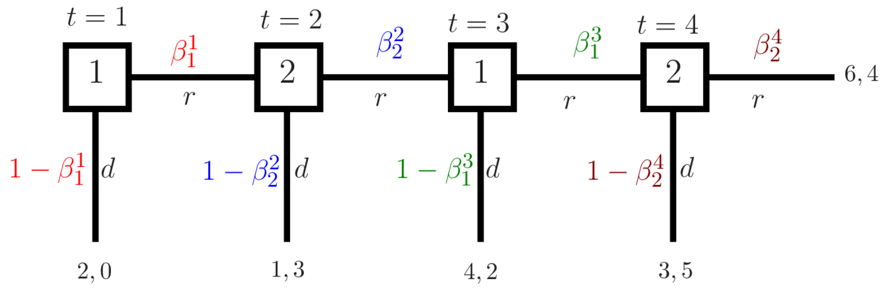

Figure 2: Four-stage centipede game

predictive content. In the 4-stage game, the mixed equilibria can be described explicitly for all parameter values $(\alpha, \delta) \in [0,1]^2$.

Proposition 6.4 *Let* Γ *be a 4-stage centipede game. Assume that* $0 \leqslant \delta(1-\alpha) < \frac{1}{3}$ *and* $\alpha < \frac{2}{3}$. *Then there are three possible types of CP-EUA* $\left\langle\left(\beta_1^1, \beta_1^3\right), \left(\beta_2^2, \beta_2^4\right)\right\rangle$ *as follows.*

1. *Case B:* $\frac{2}{3} \geqslant \alpha \geqslant \frac{9\delta-1}{11\delta}$. *Player 1 randomizes at node 1 playing* r *with probability* $\beta_1^1 = \frac{(2-3\alpha)\delta}{1-\delta}$, *and plays* d *at node* 3. *Player* 2 *plays* r *with probability* $\beta_2^2 = \frac{1-3\delta(1-\alpha)}{3(1-\delta)}$ *at node 2 and* d *at node 4.*

2. *Case C:* $\frac{9\delta-1}{11\delta} \geqslant \alpha \geqslant \frac{9\delta-1}{12\delta}$. *Player 1 randomizes at node 1 playing* r *with probability* $\beta_1^1 = \frac{\delta^2\alpha(3\alpha-2)}{3-5(6-7\alpha)-(35\alpha-27)\delta^2}$ *and at node 3 playing* r *with probability* $\beta_1^3 = \frac{3(4\alpha-3)\delta+1}{\alpha\delta}$. *Player* 2 *randomizes at node 2 playing* r *with probability* $\beta_2^2 = \frac{(2-3\alpha)\delta}{1-\delta}$ *and plays* d *at node 4.*

3. *Case D:* $\frac{9\delta-1}{12\delta} \geqslant \alpha \geqslant \frac{3\delta-1}{3\delta}$. *Player* 1 *plays* r *at node 1 and randomizes at node 3 playing* r *with probability* $\beta_1^3 = \frac{1-3\delta(1-\alpha)}{3(1-\delta)}$. *Player* 2 *randomizes at node 2 playing* r *with probability* $\beta_2^2 = \frac{(2-3\alpha)\delta}{1-\delta}$ *and plays* d *at node* 4.

The regions in α-δ space in which these types of equilibria occur are illustrated in Figure 3. Case A is the equilibrium of Proposition 6.2 where players play d (down) at every node and case E is the equilibrium of Proposition 6.1 where cooperation continues until the final node. The three cases B, C and D of Proposition 6.4 apply to successively lower degrees of ambiguity aversion.

Type B equilibrium For the profile of behaviour strategies specified in Proposition 6.4 to constitute a CP-EUA in case B, we require that player 1's "theory" about the "randomization" of player 2 at node 2 should make player 1 at node 1 indifferent between selecting either d or r. That is, $2 = 1 \cdot (1-\delta)\left(\left(1-\beta_2^2\right) + 4\beta_2^2\right) + \delta\left(\alpha + 4(1-\alpha)\right)$, which solving for β_2^2 yields[11],

$$\beta_2^2 = \frac{1-3\delta(1-\alpha)}{3(1-\delta)}. \tag{6.1}$$

[11] This is essentially the usual reasoning employed to determine the equilibrium 'mix' with standard expected payoff maximizing players.

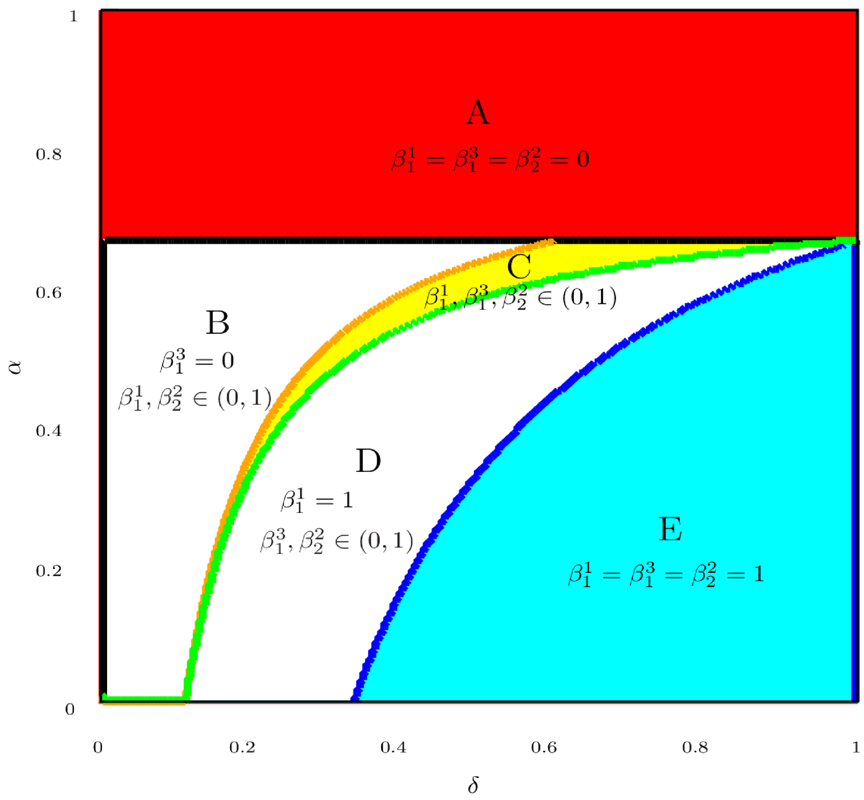

Figure 3: Equilibria of the 4-Stage Game

The situation for player 2 is different, however, since his perception of the "randomization" undertaken by player 1 over his actions at node 1 increases the ambiguity he experiences at node 2, $\delta_2 = \frac{\delta}{\delta + (1-\delta)\beta_1^1}$ according to GBU updating. This increase should generate enough ambiguity for player 2 so that he is indifferent between his two actions at node 2 given his "theory" that player 1 will choose d at node 3. More precisely, given the GBU of player 2's belief conditional on reaching node 2, player 2 should be indifferent between selecting either d or r; that is, $3 = 2(1 - \delta_2(1-\alpha)) + 5\delta_2(1-\alpha)$, where $\delta_2 = \frac{\delta}{\delta + (1-\delta)\beta_1^1}$. Solving for β_1^1 yields,

$$\beta_1^1 = \frac{(2 - 3\alpha)\delta}{1 - \delta}. \tag{6.2}$$

Substituting β_1^1 into the expression above we obtain $\delta_2 = \frac{1}{3(1-\alpha)}$ and $\delta_2(1-\alpha) = \frac{1}{3}$, as required.

The case B equilibria occur when $\frac{2}{3} \geqslant \alpha \geqslant \frac{9\delta - 1}{11\delta}$. These parameter values could be described as situations of low ambiguity and low pessimism. On the equilibrium path players are not optimistic enough, given the low degrees of ambiguity, in order to play r (right) at all nodes. However low pessimism makes them optimistic enough for playing r (right) once they are off the equilibrium path whenever it is not a dominated strategy. Player 2 randomizes at stage $t = 2$ enough to make player 1 indifferent about how to move at stage 1.

Type C equilibrium In case C players randomize at all nodes except the final node where action d is dominant for Player 2. Hence, these players must be indifferent between their actions r and d at these three nodes given their behavioural strategies where, in equilibrium, beliefs about the opponent's behavioural strategy coincide with the behavioural strategy actually chosen.

Type D equilibrium Case D is similar to Case B except that both players randomize one stage later. Here ambiguity δ is relatively high. Hence, induced by the low degree of pessimism, Player 1 will play right with probability one and the GBU of player 2's ambiguity will remain unchanged. Similar to Case B, Player 2 at node 2 and Player 1 at node 3 will now randomize. Given that down is less attractive to Player 1 at node 3, she is prepared to randomize at that node. This makes Player 2 willing to randomize at node 2.

Kilka and Weber (2001) experimentally estimate the parameters of the neo-additive model as $\delta = \alpha = \frac{1}{2}$. For α and δ close to these values the equilibrium would correspond to case D from Proposition 6.4. Player 1 will choose r with probability 1 at the first node and there will be randomization at the intermediate second and third nodes.

6.4.1 Longer centipede games $M > 4$

Some features of *CP-EUA* will remain unchanged if there are more than four stages. In particular, if $\alpha > \frac{2}{3}$ players will choose d at all stages (Proposition 6.2) and for high optimism and high ambiguity players will play r until the final node (Proposition 6.1). Increasing the stages will however open possibilities for more $(\alpha, \delta)-$regions with mixed equilibria, that is, different drop-out regions and borderlines. The yellow region (case C) illustrates that some regions may be very small.

Consider the case of not too pessimistic players, e.g., $\alpha \sim 0.5$. For ex-ante high ambiguity levels, such as e.g. $\delta \sim 0.8$, equilibrium constellations of Type E will arise where players play right with probability 1 until the last stage when Player 2 will play down. For intermediate levels of ex-ante ambiguity such as $\delta \sim 0.5$ equilibrium constellations of Type D will occur where Player 1's equilibrium behavioural strategy β_1 will make her play right with probability 1 in the early stages of the game followed by randomizing in some intermediate stages and, in the later stages stopping to play right, since Player 2 will play down with probability 1 at the end of the game. If ex-ante ambiguity is low, e.g. $\delta \sim 0.1$, then players will play down with probability 1 at the later stages of the game. Only randomizing in the early stages will generate enough ambiguity for the degree of optimism inducing them to play right with positive probability. Notice that opponent actions that receive low likelihood under a player's behavioural-strategy theory increase the updated lack-of-confidence parameter through GBU. This is not ambiguity about the player's own randomizing device; it is ambiguity about the opponent's continuation behaviour after an unexpected opponent history. Hence, we conjecture that there will be sharing (playing r) with high probability in the early stages of the game followed by exit (playing d) in the final stages of the game.

7 Bargaining

Alternating-offer bargaining provides a second illustration of the model. In the standard complete-information bargaining game, agreement is reached immediately. In experiments, however, offers are sometimes rejected and agreement may be delayed. Strategic ambiguity offers a way to generate such delay without adding payoff uncertainty or irrational types.

The alternating-offer bargaining game, developed by Stahl (1972) and Rubinstein (1982), has become one of the most intensely studied models in economics, both theoretically and experimentally. In its shortest version, the ultimatum game, it provides a prominent example of a subgame-perfect equilibrium prediction that is at odds with experimental behaviour. The theoretical prediction is an initial offer of the smallest possible share of the surplus, often zero, followed by acceptance. Experimental results, however, show that initial offers often give the responder a substantial share of the surplus and are sometimes rejected.

In longer bargaining games, subgame-perfect equilibrium again predicts immediate agreement. The first-period offer depends on the discount factor and the length of the game, but it is accepted immediately. Experimental studies show, however, that players not only make larger offers than predicted by subgame perfection but also sometimes reject offers in the first round (Roth, 1995, p. 293). In a game of perfect information, standard backward induction leaves no room for costly delay.

A common explanation is incomplete information about the opponent's payoffs. Incomplete information can generate rejected offers, but it also requires the analyst to specify the type space that subjects are assumed to have in mind. The general objection raised by Forsythe et al. (1991) is therefore relevant:

> "In a series of recent papers, the Roth group has shown that even if an experiment is designed so that each bargainer knows his opponent's utility payoffs, the information structure is still incomplete. In fact, because we can never control the thoughts and beliefs of human subjects, it is impossible to run a complete information experiment. More generally, it is impossible to run an incomplete information experiment in which the experimenter knows the true information structure. Thus we must be willing to make conjectures about the beliefs which subjects might plausibly hold, and about how they may reasonably act in light of these beliefs. (p.243)"

The approach here keeps the material bargaining game one of complete information. The source of uncertainty is not the opponent's payoff type, but the opponent's future behaviour. A player may have a theory about the opponent's strategy while lacking full confidence in that theory.

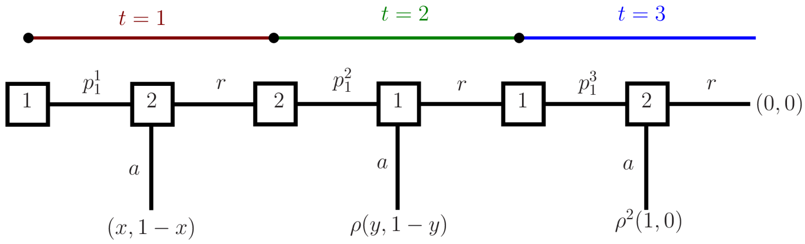

Figure 4: The Bargaining Game

If the player is sufficiently optimistic about this strategic ambiguity, an aggressive offer may be attractive because the opponent might accept it. When delay is not too costly, this gives the proposer an incentive to test whether the opponent will accept a favourable division.

Consider the three-period bargaining game in Figure 4. There is a unit surplus. The common discount factor is $\rho \in (0,1)$. In period $t=1$, player 1 proposes a division $\langle x, 1-x\rangle$; player 2 either accepts or rejects. If the offer is accepted, payoffs are $\langle x, 1-x\rangle$. If it is rejected, player 2 proposes a division $\langle y, 1-y\rangle$ in period $t=2$; player 1 either accepts or rejects. If this offer is accepted, payoffs are $\langle \rho y, \rho(1-y)\rangle$. If it is rejected, player 1 makes a final proposal $\langle z, 1-z\rangle$ in period $t=3$; player 2 either accepts or rejects. If the final offer is accepted, payoffs are $\langle \rho^2 z, \rho^2(1-z)\rangle$; if it is rejected, both players receive zero.

In line with the general model of this paper, we formulate the bargaining game on a finite grid of feasible offers. Let

$$\mathcal{G}_\varepsilon := \{0, \varepsilon, 2\varepsilon, \ldots, 1\}, \qquad \varepsilon > 0, \quad 1/\varepsilon \in \mathbb{N}. \tag{7.1}$$

be the set of possible shares for player 1. A proposal $q \in \mathcal{G}_\varepsilon$ denotes the division $\langle q, 1-q\rangle$. This finite-grid formulation keeps the bargaining application within the class of finite multi-stage games studied in this paper and has a natural interpretation in bargaining experiments, where offers are made in cents, tokens, or other discrete units. For simplicity of exposition, we assume common ambiguity parameters, $\alpha_1 = \alpha_2 = \alpha \in [0,1]$ and $\delta_1 = \delta_2 = \delta \in (0,1]$. For $r \in [0,1]$, write $\lceil r \rceil_\varepsilon := \min\{q \in \mathcal{G}_\varepsilon : q \geqslant r\}$ and $\lfloor r \rfloor_\varepsilon := \max\{q \in \mathcal{G}_\varepsilon : q \leqslant r\}$ for the nearest grid point above respectively below r.

After a period-2 offer by player 2, player 1 compares the payoff from accepting with the ambiguity-sensitive value of continuing to the final period. If player 2's period-2 proposal is the one prescribed by player 1's pure behavioural-strategy theory, then the observed opponent-history has likelihood one, player 1's updated ambiguity parameter remains δ, and the acceptance threshold is $\rho(1-\delta\alpha)$. If instead player 2's period-2 proposal has zero likelihood under player 1's theory, GBU sets the updated ambiguity parameter equal to one, and the corresponding threshold is $\rho(1-\alpha)$.

The second threshold is weakly lower than the first. With a continuum of offers, this difference would create an undercutting problem: player 2 could make a slightly lower unexpected offer that player 1 would accept after updating. On a finite offer grid, however, the two thresholds may lie in the same grid interval. In that case, the change in player 1's updated ambiguity parameter does not create a feasible lower accepted offer. Denote by $\bar{y} := \lceil \rho(1-\delta\alpha) \rceil_\varepsilon$ the smallest feasible offer that player 1 accepts in period 2, both when player 2's offer is expected and when it is a zero-likelihood surprise, i.e., that satisfies $\bar{y} - \varepsilon < \rho(1-\alpha) \leqslant \rho(1-\delta\alpha) \leqslant \bar{y}$.

The following no-undercutting condition says that there is no feasible grid offer that is rejected on the delayed-agreement path but accepted after a zero-likelihood period-2 proposal:

$$\mathcal{G}_\varepsilon \cap \big[\rho(1-\alpha), \rho(1-\delta\alpha)\big) = \varnothing. \tag{7.2}$$

We next derive sufficient conditions for delayed agreement in period 2. Consider player 2 in period 2 after rejecting a zero-likelihood initial offer by player 1. Denote by

$$x^A := \lfloor 1 - (1-\alpha)\rho \rfloor_\varepsilon \tag{7.3}$$

the largest initial demand by player 1 that player 2 accepts after such an offer.

The alternative initial demands that player 2 rejects after a zero-likelihood initial offer are collected in the set

$$\mathcal{X}^R := \{x \in \mathcal{G}_\varepsilon \setminus \{1\} : 1 - x < (1-\alpha)\rho\}. \tag{7.4}$$

If this set is nonempty, let

$$x^R := \max \mathcal{X}^R. \tag{7.5}$$

Thus x^R is the most favorable rejected alternative initial demand for player 1.

Finally, define

$$\overline{W}_1^A := (1-\delta)x^A + \delta(1-\alpha)\max\{x^A, \rho\}, \tag{7.6}$$

and, whenever $\mathcal{X}^R \neq \emptyset$,

$$\overline{W}_1^R := (1-\delta)\rho^2 + \delta(1-\alpha)\max\{x^R, \rho\}. \tag{7.7}$$

Then we obtain the following proposition.

Proposition 7.1 *Suppose that the no-undercutting condition* (7.2) *holds. Suppose further that*

$$(1-\delta)(1-\bar{y}) + \delta(1-\alpha)\max\{1-\bar{y}, \rho\} \geqslant \delta(1-\alpha), \tag{7.8}$$

and that

$$(1-\delta)\rho\bar{y} + \delta(1-\alpha) \geqslant \overline{W}_1^A. \tag{7.9}$$

If $\mathcal{X}^R \neq \emptyset$, suppose in addition that

$$(1-\delta)\rho\bar{y} + \delta(1-\alpha) \geqslant \overline{W}_1^R. \tag{7.10}$$

Then the bargaining game has a pure-strategy CP-EUA in which agreement is delayed until period $t = 2$. In this equilibrium, player 1 proposes $\langle 1, 0\rangle$ in period $t = 1$, player 2 rejects this offer, player 2 proposes $\langle \bar{y}, 1-\bar{y}\rangle$ in period $t = 2$, and player 1 accepts.

The conditions in Proposition 7.1 have a direct interpretation. Condition (7.2) ensures that player 2 cannot profit in period 2 by making a slightly lower unexpected offer that player 1 would accept only after the surprise update. Inequality (7.8) requires the prescribed period-2 offer to be optimal for player 2 after rejecting the initial demand. Inequality (7.9) requires player 1 to prefer the aggressive initial demand to the best alternative initial offer that player 2 would accept. If $\mathcal{X}^R \neq \emptyset$, inequality (7.10) further requires player 1 to prefer the aggressive initial demand to the best alternative initial offer that player 2 would reject. The equilibrium path therefore captures a simple economic mechanism: optimism under strategic ambiguity can make it worthwhile for player 1 to test whether player 2 will accept an extreme first-period demand; if player 2 rejects, agreement is reached one period later at the lowest grid offer acceptable to player 1.

8 Relation to the Literature

The closest literatures differ from the present paper along three dimensions: whether ambiguity is strategic, whether the game is dynamic, and how updating is handled after histories. We organize the comparison around these distinctions. We also relate the paper to recent work on rectangularity in multi-stage games with ambiguous incomplete information (Pahlke, 2022) and to self-confirming equilibrium under ambiguity (Battigalli et al., 2019).

The present paper extends our previous research on normal form games, Eichberger and Kelsey (2014), to a large class of extensive form games. Two earlier papers study a limited class of extensive form games, Eichberger and Kelsey (1999) and Eichberger and Kelsey (2004). These focus on signalling games in which each player only moves once. Consequently dynamic consistency is not a major problem. Signalling games may be seen as multi-stage games with only two stages and incomplete information. The present paper relaxes this restriction on the number of stages but has assumed complete information. The price of increasing the number of stages is that we are forced to consider dynamic consistency issues.

Rothe (2011) models each player's belief in an extensive form game as a neo-additive capacity with extreme pessimism (that is, $\alpha_i = 1$). However, he interprets the weight on the additive part of this capacity as the probability the player thinks her opponent is rational with the complementary weight being assigned to an irrational opponent. Furthermore, these (conditional) weights are specified exogenously for each decision node of the player. This shows how an equilibrium, of the kind studied in the present paper, can arise as the reduced form of a game when there is ambiguity over the types of the players. Building on Stauber (2011), Stauber (2017) provides a similar equilibrium notion for extensive games with ambiguity averse players where deviations from equilibrium play are assumed to identify irrational players.

Hanany et al. (2020) (henceforth HKM) also present a theory of ambiguity in multi-stage games. However they have made a number of different modelling choices. Firstly they consider games of incomplete information. In their model, there is ambiguity concerning the type of the opponent while their strategy is unambiguous. In contrast in our theory there is no type space and we focus on strategic ambiguity. However we believe that the difference between strategic ambiguity and ambiguity over types is not great. It would be straightforward to add a type space to our model, while HKM argue that strategic uncertainty can arise as a reduced form of a model with type uncertainty. Other differences are that HKM represent ambiguity by the smooth model, they use a different rule for updating beliefs and strengthen consistent planning to dynamic consistency. A cost of this is that they need to adopt a non-consequentialist decision rule. Thus current decisions may be affected by options which are no longer available. We conjecture that we could have obtained similar results using the smooth model. However, the GBU rule has the advantage that it defines beliefs both on and off the equilibrium path. In contrast, with the smooth rule, beliefs off the equilibrium path are to some extent arbitrary.

Jehiel (2005) proposes a solution concept which he refers to as *analogy-based equilibrium.* In this a player identifies similar situations and forms a single belief about their opponent's behaviour in all of them. These beliefs are required to be correct in equilibrium. For instance in the centipede game a player might consider their opponent's behaviour at all the non-terminal nodes to be analogous. Thus they may correctly believe that the opponent will play right with high probability at the average node, which increases their own incentive to play right. (The opponent perceives the situation similarly.) Jehiel predicts that either there is no cooperation or cooperation continues until the last decision node. This is not unlike our own predictions based on ambiguity.

What is common between his theory and ours is that there is an "averaging" over different decision nodes. In his theory this occurs through the perceived analogy classes, while in ours averaging occurs via the decision-weights in the neo-additive evaluation. We believe that an advantage of our approach is that the preferences we consider have been derived axiomatically and hence are linked to a wider literature on decision theory.

9 Conclusion

The paper develops a dynamic equilibrium concept for games in which players perceive ambiguity about their opponent's behaviour. The proposed equilibrium concept is defined over behavioural-strategy profiles. After each history, players update their confidence in their theories about the opponent's behaviour, and consistent planning disciplines the resulting dynamic choice problem. The analysis shows that strategic ambiguity can alter dynamic equilibrium behaviour in economically meaningful ways. Pure strategy equilibria may fail to exist even in finite games of complete and perfect information; in the centipede game, for example, some parameter values yield only mixed CP-EUA. Moreover, optimism under ambiguity can sustain cooperation in the centipede game and generate delay in non-cooperative bargaining, two outcomes that are difficult to reconcile with the standard subgame-perfect predictions in these games.

9.1 Irrational Types

A common way to reconcile backward-induction predictions with observed behaviour is to introduce a small probability of an "irrational" type. This approach has been used to explain cooperation in the repeated prisoner's dilemma and the chain-store paradox (Kreps et al., 1982), and also behaviour in the centipede game (McKelvey and Palfrey, 1992). By converting the original complete-information game into a game of incomplete information, one can make it optimal for a rational player to mimic behaviour associated with an irrational type in order to influence the opponent's future choices.

This explanation is powerful, but it depends on specifying which irrational types players regard as possible. In some games there may be a natural focal candidate, such as tit-for-tat behaviour in the repeated prisoner's dilemma. In many games, however, the relevant type is less obvious. Once enough freedom is allowed in specifying irrational types, a wide range of behaviour can be rationalized. The ambiguity approach developed here offers a different route. It does not require the analyst to add payoff types to the model. Instead, players retain the same material payoffs but may lack full confidence in their theories about their opponent's behaviour. The resulting equilibrium concept is based on axiomatic decision theory and can be applied to any game in the class we consider.

9.2 Directions for Future Research

The analysis has focused on finite two-player multi-stage games. A natural extension would be to allow incomplete information by adding type spaces for the players. Another extension would consider games with more than two players. This raises a nontrivial issue: with three or more players, one usually has to specify whether a player treats opponents' behaviour as independent. The appropriate analogue of independence for ambiguous beliefs is less clear and would need to be handled explicitly.

It should also be possible to extend the analysis beyond neo-additive preferences. The approach is most naturally suited to ambiguity models that maintain a separation between beliefs and tastes and admit a useful support notion. Multiple-prior and smooth models of ambiguity are natural candidates. For example, when beliefs are represented by a set of probabilities, a support notion can be defined using the intersection of the supports of the probabilities in that set, corresponding to the inner support of Ryan (2002).

Finitely repeated games provide another natural application. When the stage game has a unique Nash equilibrium, backward induction leaves no room for cooperation in the finitely repeated game. Yet in examples such as the repeated prisoner's dilemma, both intuition and experimental evidence suggest that some cooperation may occur. Strategic ambiguity about the opponent's future behaviour offers a possible way to study such behaviour without changing the players' material payoffs.

Appendix

A Multistage Games

A.1 Existence of Equilibrium

Proof of Theorem 4.1. For each non-terminal history $h \in H$ and each player $i = 1, 2$, introduce an agent (i, h) whose strategy set is the simplex $\Delta(A_i^h)$. A profile of these agents' mixed actions is identified with a behavioural strategy profile $\beta = \langle \beta_1, \beta_2 \rangle$ in the original game. The payoff of agent (i, h) is the linear extension of the history-h action-contingent values. That is, if agent (i, h) uses $\xi_i \in \Delta(A_i^h)$ while the other agents' choices induce the behavioural profile β, her payoff is

$$\Psi_{i,h}(\xi_i; \beta_i, \beta_{-i}) := \sum_{a_i \in A_i^h} \xi_i[a_i] V_i^h(a_i; \beta_i, \beta_{-i}). \tag{A.1}$$

The continuation plan β_i in (A.1) is the plan induced by the other agents of player i at histories different from h; once the realized action a_i is fixed, the probability assigned by ξ_i to other current actions at h is irrelevant for $V_i^h(a_i; \beta_i, \beta_{-i})$.

The strategy set of every agent is a non-empty compact convex simplex. The function $V_i^h(a_i; \beta_i, \beta_{-i})$ is continuous in the behavioural profile on the parameter domain $0 < \delta_i \leqslant 1$. The ordinary continuation payoff is a finite sum of products of local behavioural probabilities, the opponent-history likelihood $L_i(h;\beta_{-i})$ is a finite product of such probabilities, and

$$\delta_i^h(\beta_{-i}) = \frac{\delta_i}{\delta_i + (1-\delta_i)L_i(h;\beta_{-i})}$$

is continuous because the denominator is bounded below by $\delta_i > 0$. The minimum and maximum terms are taken over the finite set of pure opponent strategies and are therefore continuous. Hence each agent's payoff is continuous in the full profile and affine, hence quasi-concave, in her own mixed action.

By Nash's existence theorem, the finite agent game has a Nash equilibrium $\hat{\beta}$. At this equilibrium, if $\hat{\beta}_i(h)[a_i] > 0$, then a_i is in the support of an optimal mixed action for agent (i,h). Since the agent's payoff is linear in her own randomization, every action in the support of an optimal mixed action must maximize $V_i^h(\cdot;\hat{\beta}_i,\hat{\beta}_{-i})$ over A_i^h. Thus, for all $a_i' \in A_i^h$,

$$V_i^h(a_i;\hat{\beta}_i,\hat{\beta}_{-i}) \geqslant V_i^h(a_i';\hat{\beta}_i,\hat{\beta}_{-i}).$$

This is exactly the consistency condition in Definition 4.1. The same argument applies to every player and every non-terminal history. Hence $\hat{\beta}$ is a CP-EUA. ■

A.2 Upper hemi-continuity

Proof of Proposition 5.1. Fix a finite game Γ. Let $\theta^r = \langle \alpha_1^r, \alpha_2^r, \delta_1^r, \delta_2^r \rangle$ be a sequence in $[0,1]^2 \times (0,1]^2$ converging to $\bar{\theta} = \langle \bar{\alpha}_1, \bar{\alpha}_2, \bar{\delta}_1, \bar{\delta}_2 \rangle$ with $\bar{\delta}_i > 0$ for $i = 1, 2$. Let $\beta^r = \langle \beta_1^r, \beta_2^r \rangle$ be a CP-EUA at parameters θ^r, and suppose that $\beta^r \to \bar{\beta}$.

Consider any player i, non-terminal history h, and action $\bar{a}_i \in A_i^h$ with $\bar{\beta}_i(h)[\bar{a}_i] > 0$. For all sufficiently large r, $\beta_i^r(h)[\bar{a}_i] > 0$. Since β^r is a CP-EUA,

$$V_{i,r}^h(\bar{a}_i;\beta_i^r,\beta_{-i}^r) \geqslant V_{i,r}^h(a_i';\beta_i^r,\beta_{-i}^r) \qquad \text{for every } a_i' \in A_i^h,$$

where the subscript r records that the value is computed with parameters θ^r. On the domain with strictly positive δ_i, the value functions are continuous in the parameters and in the behavioural profile. Taking limits yields

$$V_i^h(\bar{a}_i;\bar{\beta}_i,\bar{\beta}_{-i}) \geqslant V_i^h(a_i';\bar{\beta}_i,\bar{\beta}_{-i}) \qquad \text{for every } a_i' \in A_i^h.$$

Thus every action used with positive probability by $\bar{\beta}_i$ is locally optimal at every history.

Therefore $\bar{\beta}$ is a CP-EUA at parameters $\bar{\theta}$, proving upper hemi-continuity. ∎

A.3 Dominance

Proof of Proposition 5.2. Let $\hat{\beta} = \langle \hat{\beta}_1, \hat{\beta}_2 \rangle$ be a CP-EUA. Fix a history h and suppose, contrary to the claim, that a weakly dominated pure continuation strategy $\hat{s}_i^h$ is played after h under $\hat{\beta}_i$. Let $\tilde{s}_i^h$ be a pure continuation strategy that weakly dominates $\hat{s}_i^h$ and strictly improves the payoff against at least one pure opponent continuation strategy.

For every opponent pure continuation strategy s_{-i}^h,

$$u_i^h(\tilde{s}_i^h, s_{-i}^h) \geqslant u_i^h(\hat{s}_i^h, s_{-i}^h),$$

with strict inequality for at least one s_{-i}^h. Consequently the expected component of the continuation value is weakly higher under $\tilde{s}_i^h$ than under $\hat{s}_i^h$. The minimum and maximum payoffs are also weakly higher. Since the game has different extremes, the strict improvement against some pure opponent strategy implies that at least one of the two extrema is strictly higher whenever the other is unchanged. Because $0 < \alpha_i < 1$ and $\delta_i^h(\hat{\beta}_{-i}) > 0$, the Hurwicz component of the neo-additive value is strictly higher for $\tilde{s}_i^h$ than for $\hat{s}_i^h$.

Thus replacing the continuation prescription $\hat{s}_i^h$ by $\tilde{s}_i^h$ yields a strictly larger updated continuation value at some history where $\hat{s}_i^h$ prescribes an action used with positive probability. This contradicts the local optimality condition in Definition 4.1. Hence no weakly dominated pure continuation strategy can be played after any history. Taking h to be the initial history gives the final claim for pure strategies. ∎

B Centipede Game

B.1 n-stage Centipede Game

Proof of Proposition 6.1. Let $\langle \hat{\beta}_1, \hat{\beta}_2 \rangle$ be the behavioural-strategy profile in which each player chooses r at every node $m < M$, and Player 2 chooses d at node M. We first verify that this profile is a CP-EUA when $\delta(1-\alpha) \geqslant \frac{1}{3}$.

At the final node M, Player 2 chooses d_M, which is a dominant action. This yields payoffs $\langle M-1, M+1 \rangle$.

Now consider Player 1's decision at node $M-1$. Under $\langle \hat{\beta}_1, \hat{\beta}_2 \rangle$, this node is reached with probability one, so the updated lack-of-confidence parameter is still δ. The continuation values are

$$V_1^{M-1}(d_{M-1};\hat{\beta}_1,\hat{\beta}_2) = M,$$

and

$$\begin{aligned} V_1^{M-1}(r_{M-1};\hat{\beta}_1,\hat{\beta}_2) &= (1-\delta)(M-1) + \delta\left[\alpha(M-1) + (1-\alpha)(M+2)\right] \\ &= M - 1 + 3\delta(1-\alpha). \end{aligned}$$

Thus r_{M-1} is a best response whenever $\delta(1-\alpha) \geqslant \frac{1}{3}$, and is the unique best response whenever the inequality is strict.

We next verify the remaining nodes by backward induction. Fix a node $\rho \leqslant M-3$ and suppose that r_κ is a best response at every later node κ with $\rho < \kappa < M$. We consider two cases.

First suppose that ρ is odd, so that Player 1 moves at node ρ. Under the proposed continuation profile, choosing d_ρ gives Player 1 the payoff $\rho+1$. Choosing r_ρ gives expected payoff $M-1$ under Player 1's behavioural-strategy theory of Player 2, worst payoff ρ, and best payoff $M+2$. Player 1's neo-expected pay-offs are, $V_1^\rho(d_\rho;\hat{\beta}_1,\hat{\beta}_2) = \rho+1$, and $V_1^\rho(r_\rho;\hat{\beta}_1,\hat{\beta}_2) = (1-\delta)(M-1) + \delta\left[\alpha\rho + (1-\alpha)(M+2)\right]$.
Therefore r_ρ is a best response if $(1-\delta\alpha)(M-\rho) \geqslant 2 - 3\delta + 2\delta\alpha$. Since $M-\rho \geqslant 3$, a sufficient condition is, $1 \geqslant 5\delta\alpha - 3\delta$. This follows from $\delta(1-\alpha) \geqslant \frac{1}{3}$, because then $\delta\alpha \leqslant \delta - \frac{1}{3}$, and hence

$$5\delta\alpha - 3\delta \leqslant 5\left(\delta - \frac{1}{3}\right) - 3\delta = 2\delta - \frac{5}{3} \leqslant \frac{1}{3} < 1.$$

Thus r_ρ is a best response. If $\delta(1-\alpha) > \frac{1}{3}$, the inequality is strict, so r_ρ is the unique best response.

Now suppose that ρ is even, so that Player 2 moves at node ρ. Under the proposed continuation profile, choosing d_ρ gives Player 2 the payoff $\rho+1$. Choosing r_ρ gives expected payoff $M+1$ under Player 2's behavioural-strategy theory of Player 1, worst payoff ρ, and best payoff $M+1$. Hence $V_2^\rho(d_\rho;\hat{\beta}_2,\hat{\beta}_1) = \rho+1$, whereas

$$V_2^\rho(r_\rho;\hat{\beta}_2,\hat{\beta}_1) = (1-\delta)(M+1) + \delta\left[\alpha\rho + (1-\alpha)(M+1)\right] = (1-\delta\alpha)(M+1) + \delta\alpha\rho.$$

Therefore r_ρ is a best response if $(1-\delta\alpha)(M-\rho) \geqslant \delta\alpha$. Since $\rho \leqslant M-2$, we have $M-\rho \geqslant 2$. Thus it is enough that $2(1-\delta\alpha) \geqslant \delta\alpha$, or equivalently $\delta\alpha \leqslant \frac{2}{3}$. This follows from $\delta(1-\alpha) \geqslant \frac{1}{3}$, since $\delta\alpha \leqslant \delta - \frac{1}{3} \leqslant \frac{2}{3}$. Thus r_ρ is a best response. If $\delta(1-\alpha) > \frac{1}{3}$, the inequality is strict, so r_ρ is the unique best response.

This completes the induction and proves existence of a CP-EUA in which cooperation continues until the final node.

It remains to establish uniqueness under the strict inequality. Let $\langle\beta_1, \beta_2\rangle$ be any CP-EUA and suppose $\delta(1-\alpha) > \frac{1}{3}$. At every history ρ, the updated lack-of-confidence parameter satisfies $\delta_i^\rho(\beta_{-i}) \geqslant \delta$, because the opponent-history likelihood is at most one. If the likelihood is zero, then $\delta_i^\rho(\beta_{-i}) = 1$ by definition.

The backward-induction argument above can therefore be repeated with $\delta_i^\rho(\beta_{-i})$ in place of δ. Since $\delta_i^\rho(\beta_{-i})(1-\alpha) \geqslant \delta(1-\alpha) > \frac{1}{3}$, the same strict inequalities imply that r_ρ is the unique best response at every node $\rho < M$. At node M, d_M is the unique best response for Player 2. Hence every CP-EUA must prescribe r at every node before M and d at M. Therefore the equilibrium described above is unique. ■

Proof of Proposition 6.2. Let $\langle\hat{\beta}_1, \hat{\beta}_2\rangle$ be the behavioural-strategy profile in which each player chooses d at every node at which she or he moves. We first verify that this profile is a CP-EUA when $\alpha \geqslant \frac{2}{3}$.

At the final node M, Player 2 chooses d_M, which is a dominant action. Now consider any earlier node $\rho < M$, and let i be the player who moves at ρ. Write

$$\eta_i^\rho := \delta_i^\rho(\hat{\beta}_{-i})$$

for player i's updated lack-of-confidence parameter at node ρ under the candidate profile. By definition, $\eta_i^\rho \in [0, 1]$.

If player i chooses d_ρ, the payoff is $\rho+1$, so $V_i^\rho(d_\rho; \hat{\beta}_i, \hat{\beta}_{-i}) = \rho+1$. If instead player i chooses r_ρ and then retains the continuation plan $\hat{\beta}_i$, the expected payoff under the behavioural-strategy theory $\hat{\beta}_{-i}$ is ρ, because the opponent is expected to choose d at the next node. The worst payoff over the opponent's pure continuation strategies is also ρ, while the best payoff is $\rho + 3$. Hence

$$\begin{aligned} V_i^\rho(r_\rho; \hat{\beta}_i, \hat{\beta}_{-i}) &= (1-\eta_i^\rho)\rho + \eta_i^\rho\,[\alpha\rho + (1-\alpha)(\rho+3)] \\ &= \rho + 3\eta_i^\rho(1-\alpha). \end{aligned}$$

Since $\alpha \geqslant \frac{2}{3}$ and $\eta_i^\rho \leqslant 1$,

$$\rho + 3\eta_i^\rho(1-\alpha) \leqslant \rho + 1.$$

Thus d_ρ is a best response at every node $\rho < M$. Together with the dominance of d_M at the final node, this proves that $\langle\hat{\beta}_1, \hat{\beta}_2\rangle$ is a CP-EUA.

It remains to prove uniqueness when $\alpha > \frac{2}{3}$. Let $\langle \beta_1, \beta_2 \rangle$ be any CP-EUA. At node M, d_M is the unique best response for Player 2. Now argue by backward induction. Suppose that at every node after ρ, the player who moves there uniquely chooses d. Let i be the player who moves at node $\rho < M$, and write

$$\eta_i^\rho := \delta_i^\rho(\beta_{-i}).$$

If player i chooses d_ρ, the payoff is $\rho+1$. If player i chooses r_ρ, then, by the induction hypothesis, the opponent's behavioural-strategy theory assigns probability one to d at the next node, and player i's own continuation plan assigns probability one to d at her or his next node. Thus the expected payoff and worst payoff from choosing r_ρ are ρ, while the best payoff is $\rho + 3$. Therefore

$$V_i^\rho(r_\rho; \beta_i, \beta_{-i}) = \rho + 3\eta_i^\rho(1 - \alpha).$$

Since $\alpha > \frac{2}{3}$ and $\eta_i^\rho \leqslant 1$, $\rho + 3\eta_i^\rho(1-\alpha) < \rho + 1 = V_i^\rho(d_\rho; \beta_i, \beta_{-i})$. Hence d_ρ is the unique best response at node ρ. The induction therefore implies that every CP-EUA prescribes d at every node. The equilibrium is unique. ■

Proof of Proposition 6.3. Suppose, to the contrary, that there exists a CP-EUA in pure behavioural strategies. Let $\langle \beta_1, \beta_2 \rangle$ be such an equilibrium. At node M, Player 2 chooses d_M, since d_M is a dominant action.

First suppose that no node before M is assigned action d. Then $\beta_i(m)[r] = 1$ for every $m < M$. In particular, at node $M-1$ Player 1's updated lack-of-confidence parameter is δ, since Player 2's component of the preceding history has likelihood one under β_2. The continuation values are

$$V_1^{M-1}(d_{M-1}; \beta_1, \beta_2) = M$$

and

$$\begin{aligned} V_1^{M-1}(r_{M-1}; \beta_1, \beta_2) &= (1-\delta)(M-1) + \delta\left[\alpha(M-1) + (1-\alpha)(M+2)\right] \\ &= M - 1 + 3\delta(1-\alpha). \end{aligned}$$

Since $\delta(1-\alpha) < \frac{1}{3}$, r_{M-1} is not a best response. This contradicts the supposition that $\beta_1(M-1)[r] = 1$.

Hence some node before M must be assigned action d. Let $\lambda < M$ be the last node before M at which action d is assigned. Thus every node $\lambda+1, \ldots, M-1$ is assigned action r.

First suppose that λ is even. Then Player 2 moves at λ, and $\lambda \leqslant M-2$. Since the profile is pure, $\delta_2^\lambda(\beta_1)$ is either δ or 1. If Player 2 chooses d_λ, his payoff is $\lambda+1$, so

$$V_2^\lambda(d_\lambda;\beta_2,\beta_1)=\lambda+1.$$

If he chooses r_λ, then, under Player 2's behavioural-strategy theory of Player 1, play reaches node M, where Player 2 chooses d_M. The expected payoff is therefore $M+1$. The worst payoff over Player 1's pure continuation strategies is λ, and the best payoff is $M+1$. Hence

$$\begin{aligned}V_2^\lambda(r_\lambda;\beta_2,\beta_1)&=(1-\delta_2^\lambda(\beta_1))(M+1)+\delta_2^\lambda(\beta_1)\left[\alpha\lambda+(1-\alpha)(M+1)\right]\\&=(M+1)-\delta_2^\lambda(\beta_1)\alpha(M+1-\lambda).\end{aligned}$$

Since $\delta_2^\lambda(\beta_1)\leqslant 1$,

$$V_2^\lambda(r_\lambda;\beta_2,\beta_1)\geqslant\lambda+(1-\alpha)(M+1-\lambda).$$

Because $\lambda\leqslant M-2$ and $\alpha<\frac{2}{3}$, we have $(1-\alpha)(M+1-\lambda)>1$. Thus

$$V_2^\lambda(r_\lambda;\beta_2,\beta_1)>\lambda+1=V_2^\lambda(d_\lambda;\beta_2,\beta_1),$$

contradicting $\beta_2(\lambda)[d]=1$.

Now suppose that λ is odd and $\lambda\leqslant M-3$. Then Player 1 moves at λ. Again, since the profile is pure, $\delta_1^\lambda(\beta_2)$ is either δ or 1. If Player 1 chooses d_λ, her payoff is $\lambda+1$, so

$$V_1^\lambda(d_\lambda;\beta_1,\beta_2)=\lambda+1.$$

If she chooses r_λ, then, under Player 1's behavioural-strategy theory of Player 2, play reaches node M, where Player 2 chooses d_M. The expected payoff is therefore $M-1$. The worst payoff over Player 2's pure continuation strategies is λ, and the best payoff is $M+2$. Hence

$$V_1^\lambda(r_\lambda;\beta_2,\beta_2)=(1-\delta_1^\lambda(\beta_2))(M-1)+\delta_1^\lambda(\beta_2)\left[\alpha\lambda+(1-\alpha)(M+2)\right].$$

For $\delta_1^\lambda(\beta_2)=1$, this gives

$$V_1^\lambda(r_\lambda;\beta_1,\beta_2)-(\lambda+1)=(1-\alpha)(M+2-\lambda)-1>0,$$

because $M+2-\lambda\geqslant 5$ and $\alpha<\frac{2}{3}$.

For $\delta_1^\lambda(\beta_2) = \delta$, write $n := M - \lambda$. Since $\lambda \leqslant M - 3$, we have $n \geqslant 3$. Then

$$\begin{aligned} V_1^\lambda(r_\lambda; \beta_1, \beta_2) - (\lambda + 1) &= n - 2 + \delta\,[3 - \alpha(n+2)] \\ &= n(1 - \delta\alpha) - 2 + 3\delta - 2\delta\alpha. \end{aligned}$$

The right-hand side is increasing in n, since $1 - \delta\alpha > 0$. Thus it is bounded below by its value at $n = 3$, namely

$$1 + 3\delta - 5\delta\alpha.$$

Since $\alpha < \frac{2}{3}$, we have $3 - 5\alpha > -\frac{1}{3}$, and hence

$$1 + 3\delta - 5\delta\alpha = 1 + \delta(3 - 5\alpha) > 1 - \frac{\delta}{3} \geqslant \frac{2}{3} > 0.$$

Thus r_λ is strictly better than d_λ, again contradicting $\beta_1(\lambda)[d] = 1$.

It remains to consider the case $\lambda = M - 1$. Thus Player 1 chooses d_{M-1}, and all nodes after any earlier node assigned d and before $M - 1$ are assigned r.

At node $M - 1$, if Player 1's updated lack-of-confidence parameter were equal to 1, then

$$V_1^{M-1}(r_{M-1}; \beta_1, \beta_2) = \alpha(M-1) + (1-\alpha)(M+2) = M + 2 - 3\alpha > M = V_1^{M-1}(d_{M-1}; \beta_1, \beta_2),$$

since $\alpha < \frac{2}{3}$. Therefore, for d_{M-1} to be a best response, Player 1's updated lack-of-confidence parameter at node $M - 1$ must be δ. This means that Player 2's component of the preceding history has likelihood one under β_2. Hence Player 2 assigns action r at every even node before M.

If Player 1 assigns action r at every odd node before $M - 1$, then at node $M - 2$ Player 2's updated lack-of-confidence parameter is δ. Since Player 1 assigns d at node $M - 1$, Player 2's continuation values at node $M - 2$ are

$$V_2^{M-2}(d_{M-2}; \beta_2, \beta_1) = M - 1$$

and

$$\begin{aligned} V_2^{M-2}(r_{M-2}; \beta_2, \beta_1) &= (1-\delta)(M-2) + \delta\,[\alpha(M-2) + (1-\alpha)(M+1)] \\ &= M - 2 + 3\delta(1-\alpha). \end{aligned}$$

Since $\delta(1-\alpha) < \frac{1}{3}$, r_{M-2} is not a best response. This contradicts the fact that Player 2 assigns action r at node $M - 2$.

It follows that Player 1 must assign action d at some odd node before $M-1$. Let $\mu < M-1$ be the last such odd node. Since Player 2 assigns action r at every even node before M, Player 1's updated lack-of-confidence parameter at node μ is δ. Moreover, after choosing r_μ and following her continuation plan, Player 1 reaches node $M-1$ and then chooses d_{M-1}. Thus the expected payoff from r_μ is M. The worst payoff over Player 2's pure continuation strategies is μ, and the best payoff is M. Hence

$$V_1^\mu(r_\mu;\beta_1,\beta_2) = (1-\delta)M + \delta\left[\alpha\mu + (1-\alpha)M\right] = M - \delta\alpha(M-\mu).$$

Since d_μ gives payoff $\mu+1$,

$$V_1^\mu(r_\mu;\beta_1,\beta_2) - V_1^\mu(d_\mu;\beta_1,\beta_2) = (M-\mu)(1-\delta\alpha) - 1.$$

Now $\mu \leqslant M-3$, so $M-\mu \geqslant 3$. Also $\delta\alpha < \frac{2}{3}$, because $\delta \leqslant 1$ and $\alpha < \frac{2}{3}$. Therefore

$$(M-\mu)(1-\delta\alpha) - 1 > 3\left(1-\frac{2}{3}\right) - 1 = 0.$$

Thus r_μ is strictly better than d_μ, contradicting $\beta_1(\mu)[d] = 1$.

All possible cases lead to a contradiction. Therefore, when $\alpha < \frac{2}{3}$ and $\delta(1-\alpha) < \frac{1}{3}$, the centipede game has no CP-EUA in pure behavioural strategies. ■

B.2 4-Stage Centipede Game

B.2.1 Best Responses

This section of the appendix studies mixed equilibria in the 4-stage centipede game. The notation is defined in the main text.

Proof of Proposition 6.4 Given behavioural strategies $\langle\beta_1^1,\beta_1^3\rangle$ and $\langle\beta_2^2,\beta_2^4\rangle$, the neo-additive payoffs $V_i^t(a_i;\beta_i,\beta_{-i})$ for Player 1 and Player 2 from choosing $a_i \in A_i^t = \{r,d\}$ at stages $t=1,3$ and $t=2,4$, respectively, can be obtained from Equation (3.4). If player i uses a behavioural randomization $\beta_i^t \in \Delta(A_i^t)$ at history t, one obtains the expected neo-additive payoff values $V_i^t(\beta_i,\beta_{-i}) = \sum_{a_i\in A_i^t}\beta_i^t[a_i]\,V_i^t(a_i;\beta_i,\beta_{-i})$.

The four expected neo-additive payoffs $V_i^t(\beta_i,\beta_{-i})$ are:

- Player 2 at stage 4:
$$V_2^4(\beta_1^1,\beta_1^3,\beta_2^2,\beta_2^4) = 4\cdot\beta_2^4 + (1-\beta_2^4)\cdot 5,$$

- Player 1 at stage 3:

  $$V_1^3(\beta_1^1,\beta_1^3,\beta_2^2,\beta_2^4) = 4(1-\beta_1^3)+\beta_1^3\cdot\left\{(1-\delta^3(\beta_2^2))\left[(1-\beta_2^4)3+\beta_2^4 6\right]+\delta^3(\beta_2^2)\left[\alpha 3+(1-\alpha)6\right]\right\}$$

  with updated degree of ambiguity $\delta^3(\beta_2^2) = \frac{\delta}{\delta+(1-\delta)\beta_2^2}$,

- Player 2 at stage 2:

  $$\begin{aligned} V_2^2(\beta_1^1,\beta_1^3,\beta_2^2,\beta_2^4) &= 3\cdot(1-\beta_2^2) \\ &\quad+\beta_2^2\cdot(1-\beta_2^4)\left\{(1-\delta^2(\beta_1^1))\left[(1-\beta_1^3)2+\beta_1^3 5\right]+\delta^2(\beta_1^1)\left[\alpha 2+(1-\alpha)\,5\right]\right\} \\ &\quad+\beta_2^2\cdot\beta_2^4\left\{(1-\delta^2(\beta_1^1))\left[(1-\beta_1^3)2+\beta_1^3 4\right]+\delta^2(\beta_1^1)\left[\alpha 2+(1-\alpha)\,4\right]\right\} \end{aligned}$$

  with updated degree of ambiguity $\delta^2(\beta_1^1) = \frac{\delta}{\delta+(1-\delta)\beta_1^1}$,

- Player 1 at stage 1:

  $$\begin{aligned} V_1^1(\beta_1^1,\beta_1^3,\beta_2^2,\beta_2^4) &= 2\cdot(1-\beta_1^1) \\ &\quad+\beta_1^1\cdot(1-\beta_1^3)\left\{(1-\delta)\left[(1-\beta_2^2)1+\beta_2^2 4\right]+\delta\left[\alpha 1+(1-\alpha)4\right]\right\} \\ &\quad+\beta_1^1\cdot\beta_1^3\left\{(1-\delta)\left[(1-\beta_2^2)1+\beta_2^2(1-\beta_2^4)3+\beta_2^2\beta_2^4 6\right]+\delta\left[\alpha 1+(1-\alpha)6\right]\right\}. \end{aligned}$$

From these neo-additive payoffs, one obtains the following best-reply correspondences at the four stages.

At Stage $t=4$, Player 2 faces no uncertainty (and, hence, no ambiguity) about the other player's behaviour. Hence, Player 2 has a dominant behavioural strategy,

$B_2^4(\beta_1^1,\beta_1^3,\beta_2^2,\beta_2^4) := \arg\max_{\beta_2^4\in[0,1]} V_2^4(\beta_1^1,\beta_1^3,\beta_2^2,\beta_2^4) = \{0\}$.

At $t=3$, Player 1 has the best reply correspondence,

$$B_1^3(\beta_1^1,\beta_1^3,\beta_2^2,\beta_2^4) := \arg\max_{\beta_1^3\in[0,1]} V_1^3(\beta_1^1,\beta_1^3,\beta_2^2,\beta_2^4) = \begin{cases} \{1\} & \text{for } V_1^3(\beta_1^1,1,\beta_2^2,\beta_2^4) > V_1^3(\beta_1^1,0,\beta_2^2,\beta_2^4)=4 \\ [0,1] & \text{for } V_1^3(\beta_1^1,1,\beta_2^2,\beta_2^4) = V_1^3(\beta_1^1,0,\beta_2^2,\beta_2^4)=4. \\ \{0\} & \text{for } V_1^3(\beta_1^1,1,\beta_2^2,\beta_2^4) < V_1^3(\beta_1^1,0,\beta_2^2,\beta_2^4)=4 \end{cases}$$

Straight-forward computation yields $V_1^3(\beta_1^1,0,\beta_2^2,\beta_2^4)=4$ and $V_1^3(\beta_1^1,1,\beta_2^2,\beta_2^4)=3\left[\frac{\delta(2-\alpha)+(1-\delta)\beta_2^2}{\delta+(1-\delta)\beta_2^2}\right]$. Hence, Player 1 will choose $\beta_1^3=1$ for $\left[\frac{\delta(2-3\alpha)}{1-\delta}\right] > \beta_2^2$ and $\beta_1^3=0$ for $\left[\frac{\delta(2-3\alpha)}{1-\delta}\right] < \beta_2^2$. For $\left[\frac{\delta(2-3\alpha)}{1-\delta}\right] = \beta_2^2$ any $\beta_1^3\in[0,1]$ will be a best reply.

At Stage 2, Player 2's neo-additive payoff leads to the best reply correspondence

$$B_2^2(\beta_1^1,\beta_1^3,\beta_2^2,\beta_2^4) := \arg\max_{\beta_2^2\in[0,1]} V_2^2(\beta_1^1,\beta_1^3,\beta_2^2,\beta_2^4) = \begin{cases} \{1\} & \text{for } V_2^2(\beta_1^1,\beta_1^3,1,\beta_2^4) > V_2^2(\beta_1^1,\beta_1^3,0,\beta_2^4)=3 \\ [0,1] & \text{for } V_2^2(\beta_1^1,\beta_1^3,1,s) = V_2^2(\beta_1^1,\beta_1^3,0,\beta_2^4)=3. \\ \{0\} & \text{for } V_2^2(\beta_1^1,\beta_1^3,1,\beta_2^4) < V_2^2(\beta_1^1,\beta_1^3,0,\beta_2^4)=3 \end{cases}$$

Depending on β_1^1 and β_1^3, the behavioural strategy of Player 1 at $t=1$, Player 2 may find it optimal either to choose $\beta_2^2 = 1$ or $\beta_2^2 = 0$:

$$V_2^2(\beta_1^1, \beta_1^3, 1, 0) = \left[\frac{\delta\,(5-3\alpha) + (1-\delta)(2+3\beta_1^3)\beta_1^1}{\delta + (1-\delta)\beta_1^1}\right] \gtreqless 3 = V_2^2(\beta_1^1, \beta_1^3, 0, 0).$$

Notice that optimal behaviour for Player 2 at Stage 2 depends both the opponent's behaviour in following periods β_1^3 and in past periods β_1^1 since Player 2's degree of ambiguity $\delta^2(\beta_1^1) = \frac{\delta}{\delta+(1-\delta)\beta_1^1}$ depends on β_1^1.

At the first stage of the game, $t = 1$, Player 1's neo-additive payoff yields the best reply correspondence:

$$B_1^1(\beta_1^1, \beta_1^3, \beta_2^2, \beta_2^4) := \arg\max_{\beta_1^1 \in [0,1]} V_1^1(\beta_1^1, \beta_1^3, \beta_2^2, \beta_2^4) = \begin{cases} \{1\} & \text{for } V_1^1(1, \beta_1^3, \beta_2^2, \beta_2^4) > V_1^1(0, \beta_1^3, \beta_2^2, \beta_2^4) = 2 \\ [0,1] & \text{for } V_1^1(1, \beta_1^3, \beta_2^2, \beta_2^4) = V_1^1(0, \beta_1^3, \beta_2^2, \beta_2^4) = 2. \\ \{0\} & \text{for } V_1^1(1, \beta_1^3, \beta_2^2, \beta_2^4) < V_1^1(0, \beta_1^3, \beta_2^2, \beta_2^4) = 2 \end{cases}$$

Note that $\beta_1^1 = 1$ or $\beta_1^1 = 0$ may be optimal depending on whether

$$V_1^1(1, \beta_1^3, \beta_2^2, 0) = (1-\delta)\left[3\beta_2^2 + \beta_2^2\beta_1^3\right] + \delta(1-\alpha)\left[2\beta_1^3 + 3\right] + 1 \gtreqless 2 = V_1^1(0, \beta_1^3, \beta_2^2, 0).$$

In particular, even for $\beta_1^3 = 0$, $\beta_2^2 \gtreqless \frac{1-3\delta(1-\alpha)}{3(1-\delta)}$, $\beta_1^1 > 0$ may be optimal.

B.2.2 CP-EUA of the 4-stage Centipede Game

A CP-EUA $(\beta_1^1, \beta_1^3, \beta_2^2, \beta_2^4) \in [0,1]^4$ is a fixed point of the best response correspondence. This can be expressed formally

$$(\hat\beta_1^1, \hat\beta_1^3, \hat\beta_2^2, \hat\beta_2^4) \in \underbrace{B_1^1(\hat\beta_1^1, \hat\beta_1^3, \hat\beta_2^2, \hat\beta_2^4) \times B_1^3(\hat\beta_1^1, \hat\beta_1^3, \hat\beta_2^2, \hat\beta_2^4)}_{=B_1(\hat\beta_1^1, \hat\beta_1^3, \hat\beta_2^2, \hat\beta_2^4)} \times \underbrace{B_2^2(\hat\beta_1^1, \hat\beta_1^3, \hat\beta_2^2, \hat\beta_2^4) \times B_2^4(\hat\beta_1^1, \hat\beta_1^3, \hat\beta_2^2, \hat\beta_2^4)}_{=B_2(\hat\beta_1^1, \hat\beta_1^3, \hat\beta_2^2, \hat\beta_2^4)}.$$

Conditional on the degree of ambiguity δ and ambiguity attitude α, the following types of CP-EUA are possible for the respective regions of parameter values (α, δ).

Type of CP-EUA/Region		$\hat\beta_1^1 \in$	$\hat\beta_1^3 \in$	$\hat\beta_2^2 \in$	$\hat\beta_2^4 \in$
A	$\alpha \geqslant a_0(\delta) = \frac{2}{3}$	$\{0\}$	$\{0\}$	$\{0\}$	$\{0\}$
B	$a_0(\delta) > \alpha \geqslant a_1(\delta)$	$[0,1]$	$\{0\}$	$[0,1]$	$\{0\}$
C	$a_1(\delta) \geqslant \alpha \geqslant a_2(\delta)$	$[0,1]$	$[0,1]$	$[0,1]$	$\{0\}$
D	$a_2(\delta) \geqslant \alpha \geqslant a_3(\delta)$	$\{1\}$	$[0,1]$	$[0,1]$	$\{0\}$
E	$a_3(\delta) \geqslant \alpha$	$\{1\}$	$\{1\}$	$\{1\}$	$\{0\}$

Table 1: *CP-EUA* Equilibria $\left(\beta_1^1, \beta_1^3, \beta_2^2, \beta_2^4\right)$

The following three equations show the difference of the neo-additive payoffs of the actions r and d (or, equivalently, of the degenerate behavioural strategies $\beta_i^t = 1$ and $\beta_i^t = 0$) at the three non-trivial stages of the game:[12]

$$
\begin{aligned}
E_1^1(\beta_1^1,\beta_1^3,\beta_2^2,\beta_2^4) &:= V_1^1(1,\beta_1^3,\beta_2^2,\beta_2^4) - V_1^1(0,\beta_1^3,\beta_2^2,\beta_2^4) = (1-\alpha)\delta\left(3+2\beta_1^3\right) + \left(3-\beta_1^3\right)(1-\delta)\beta_2^2 - 1,\\
E_2^2(\beta_1^1,\beta_1^3,\beta_2^2,\beta_2^4) &:= V_2^2(\beta_1^1,\beta_1^3,1,\beta_2^4) - V_2^2(\beta_1^1,\beta_1^3,0,\beta_2^4) = \frac{\big(3(1-\delta)\beta_1^3-(1-\delta)\big)\,\beta_1^1+\delta(2-3\alpha)}{\delta+(1-\delta)\,\beta_1^1},\\
E_1^3(\beta_1^1,\beta_1^3,\beta_2^2,\beta_2^4) &:= V_1^3(\beta_1^1,1,\beta_2^2,\beta_2^4) - V_1^3(\beta_1^1,0,\beta_2^2,\beta_2^4) = \frac{(2-3\alpha)\delta-(1-\delta)\,\beta_2^2}{\delta+(1-\delta)\,\beta_2^2}.
\end{aligned}
$$

In mixed equilibria players are indifferent between their two actions at one or more nodes. Depending on the parameter values (α,δ), the behavioural strategy combinations $(\beta_1^1,\beta_1^3,\beta_2^2,\beta_2^4)$ satisfying these equations as equalities and/or inequalities determine the different regions of CP-EUA equilibria.

Region B Equilibrium conditions for Region B:

$a_0(\delta)=\frac{2}{3}\geqslant\alpha\geqslant a_1(\delta)=\frac{9\delta-1}{11\delta}$: $\hat\beta_1^1\in[0,1]$,$\hat\beta_1^3=0$,$\hat\beta_2^2\in[0,1]$, $\hat\beta_2^4=0$:

$$
\begin{aligned}
&E_1^1(\hat\beta_1^1,0,\hat\beta_2^2,0)=3(1-\alpha)\delta+3(1-\delta)\hat\beta_2^2-1 && =0, && \hat\beta_1^1=\tfrac{(2-3\alpha)\delta}{1-\delta},\\
&E_2^2(\hat\beta_1^1,0,\hat\beta_2^2,0)=\tfrac{((1-\delta))\hat\beta_1^1+\delta(2-3\alpha)}{\delta+(1-\delta)\hat\beta_1^1} && =0, && \hat\beta_2^2=\tfrac{1-3(1-\alpha)\delta}{3(1-\delta)},\\
&E_1^3(\hat\beta_1^1,0,\hat\beta_2^2,0)=\tfrac{1-(9-12\alpha)\delta}{1-3\alpha\delta} && \leqslant 0, && \hat\beta_1^3=0,\\
&E_2^4(\hat\beta_1^1,0,\hat\beta_2^2,0)=-1 && <0, && \hat\beta_2^4=0.
\end{aligned}
$$

Region C Equilibrium conditions for Region C:

$a_1(\delta)=\frac{9\delta-1}{11\delta}\geqslant\alpha\geqslant a_2(\delta)=\frac{9\delta-1}{12\delta}$: $\hat\beta_1^1\in[0,1]$,$\hat\beta_1^3\in[0,1]$,$\hat\beta_2^2\in[0,1]$, $\hat\beta_2^4=0$:

$$
\begin{aligned}
E_1^1(\hat\beta_1^1,\hat\beta_1^3,\hat\beta_2^2,0) &= \left(2(1-\alpha)\delta-(1-\delta)\,\hat\beta_2^2\right)\hat\beta_1^3\\
&\quad +3\,(1-\delta)\,\hat\beta_2^2+3(1-\alpha)\delta-1 && =0, && \hat\beta_1^1=\tfrac{(3\alpha-2)\alpha\delta^2}{3-5(6-7\alpha)-(35\alpha-27)\delta^2},\\
E_2^2(\hat\beta_1^1,\hat\beta_1^3,\hat\beta_2^2,0) &= \tfrac{(2-3\alpha)\delta+(1-\delta)\left(3\hat\beta_1^3-1\right)\hat\beta_1^1}{\delta+(1-\delta)\hat\beta_2^2} && =0, && \hat\beta_2^2=\tfrac{(2-3\alpha)\delta}{1-\delta},\\
E_1^3(\hat\beta_1^1,\hat\beta_1^3,\hat\beta_2^2,0) &= \tfrac{(2-3\alpha)\delta-(1-\delta)\hat\beta_2^2}{\delta+(1-\delta)\hat\beta_2^2} && =0, && \hat\beta_1^3=\tfrac{3(4\alpha-3)\delta+1}{\alpha\delta},\\
E_2^4(\hat\beta_1^1,\hat\beta_1^3,\hat\beta_2^2,0) &= -1 && <0, && \hat\beta_2^4=0.
\end{aligned}
$$

Region D Equilibrium conditions for Region D:

$$
a_2(\delta)=\frac{9\delta-1}{12\delta}\geqslant\alpha\geqslant a_3(\delta)=\frac{3\delta-1}{3\delta}:\quad \hat\beta_1^1=1,\quad \hat\beta_1^3\in[0,1],\quad \hat\beta_2^2\in[0,1],\quad \hat\beta_2^4=0.
$$

$$
\begin{aligned}
&E_1^1(1,\hat\beta_1^3,\hat\beta_2^2,0)=\tfrac{3(\alpha^2+11\alpha-9)\delta^2+5(6-7\alpha)\delta-3}{3(1-\delta)} && \geqslant 0, && \hat\beta_1^1=1,\\
&E_2^2(1,\hat\beta_1^3,\hat\beta_2^2,0)=3(1-\delta)\hat\beta_1^3+3(1-\alpha)\delta-1 && =0, && \hat\beta_1^3=\tfrac{1-3(1-\alpha)\delta}{3(1-\delta)},\\
&E_1^3(1,\hat\beta_1^3,\hat\beta_2^2,0)=\tfrac{(2-3\alpha)\delta-(1-\delta)\hat\beta_2^2}{\delta+(1-\delta)\hat\beta_2^2} && =0, && \hat\beta_2^2=\tfrac{(2-3\alpha)\delta}{1-\delta},\\
&E_2^4(1,\hat\beta_1^3,\hat\beta_2^2,0)=-1 && <0, && \hat\beta_2^4=0.
\end{aligned}
$$

[12] Player 2's behaviour at node 4 is trivially determined $\beta_2^4=0$ since he has a dominant strategy.

These parameter regions of the CP-EUA are summarized in Table 1 and illustrated in Figure 3.

C Bargaining

Proof of Proposition 7.1. We verify the local optimality conditions in the definition of CP-EUA. The calculation proceeds from later bargaining histories to earlier ones only because later continuation values enter earlier comparisons. The proof does not treat a history as a zero-likelihood history for player i merely because it has zero reach probability under the candidate profile. The updated ambiguity parameter is determined by the likelihood of the opponent's observed behaviour.

Let

$$T_\delta := \rho(1-\delta\alpha), \qquad T_1 := \rho(1-\alpha).$$

Since $\delta \leqslant 1$, we have $T_1 \leqslant T_\delta$. By definition, $\bar{y} = \lceil T_\delta \rceil_\varepsilon$. The no-undercutting condition (7.2) implies that no grid point lies in $[T_1, T_\delta)$. Hence $\bar{y}$ is also the smallest grid point weakly above T_1. Thus

$$\bar{y} = \lceil T_\delta \rceil_\varepsilon = \lceil T_1 \rceil_\varepsilon . \tag{C.1}$$

We construct the following pure behavioural-strategy profile $\hat{\beta} = \langle \hat{\beta}_1, \hat{\beta}_2 \rangle$.

In period $t = 1$, player 1 proposes $\langle 1, 0 \rangle$. Player 2 rejects this offer. At any other initial offer $\langle x, 1-x \rangle$, with $x \in \mathcal{G}_\varepsilon \setminus \{1\}$, player 2 accepts if and only if

$$1 - x \geqslant (1-\alpha)\rho. \tag{C.2}$$

In period $t = 2$, after rejection of the initial offer $\langle 1, 0 \rangle$, player 2 proposes $\langle \bar{y}, 1-\bar{y} \rangle$. After rejection of any other initial offer, player 2 proposes $\langle 0, 1 \rangle$. At a period-2 response history h at which player 2 has proposed $\langle y, 1-y \rangle$, player 1 accepts if and only if

$$y \geqslant \lceil \rho\big(1 - \eta_1(h)\alpha\big) \rceil_\varepsilon , \qquad \eta_1(h) := \delta_1^h(\hat{\beta}_2). \tag{C.3}$$

Thus the acceptance threshold is $\bar{y}$ when $\eta_1(h) = \delta$, and also $\bar{y}$ when $\eta_1(h) = 1$, by (C.1).

In period $t = 3$, player 1 proposes $\langle 1, 0 \rangle$ at every final proposal history, and player 2 accepts every final proposal $\langle z, 1-z \rangle$.

Because $\hat{\beta}$ is pure, the opponent-history likelihood for player i is either one or zero at every history. If the opponent's observed behaviour agrees with $\hat{\beta}_{-i}$ up to that history, then $L_i(h; \hat{\beta}_{-i}) = 1$ and $\delta_i^h(\hat{\beta}_{-i}) = \delta$. If the opponent's observed behaviour deviates from $\hat{\beta}_{-i}$, then $L_i(h; \hat{\beta}_{-i}) = 0$ and $\delta_i^h(\hat{\beta}_{-i}) = 1$. A deviation by player i herself does not affect $L_i(h; \hat{\beta}_{-i})$.

We now verify optimality at each class of histories.

Period $t = 3$. Consider first a final response history at which player 1 has proposed $\langle z, 1-z \rangle$. If player 2 accepts, he obtains $\rho^2(1-z) \geqslant 0$; if he rejects, he obtains zero. Hence acceptance is optimal for player 2 after every final proposal.

Now consider a final proposal history for player 1, and let $\eta \in \{\delta, 1\}$ denote her updated ambiguity parameter at that history. Given player 2's continuation strategy of accepting every final proposal, a proposal $z \in \mathcal{G}_\varepsilon$ gives player 1 the ordinary continuation payoff $\rho^2 z$. Over player 2's pure continuation strategies, the worst payoff from proposing z is zero and the best payoff is $\rho^2 z$. Hence the ambiguity-sensitive value of proposing z is

$$(1-\eta)\rho^2 z + \eta(1-\alpha)\rho^2 z = \rho^2(1-\eta\alpha)z. \tag{C.4}$$

This is maximized at $z = 1$. Thus player 1's prescribed final proposal is optimal.

Period $t = 2$: player 1's response. Consider a period-2 response history h at which player 2 has proposed $\langle y, 1-y\rangle$. Let

$$\eta_1(h) := \delta_1^h(\hat{\beta}_2).$$

Acceptance gives player 1 the payoff ρy. If player 1 rejects, the game proceeds to period $t = 3$, where her own continuation strategy prescribes the final proposal $\langle 1, 0\rangle$. Given player 2's continuation strategy, this gives player 1 the ordinary continuation payoff ρ^2. Over player 2's pure continuation strategies, the worst payoff from rejection is zero and the best payoff is ρ^2. Hence the value of rejecting is

$$(1-\eta_1(h))\rho^2 + \eta_1(h)(1-\alpha)\rho^2 = \rho^2\big(1-\eta_1(h)\alpha\big). \tag{C.5}$$

Thus acceptance is optimal if and only if $\rho y \geqslant \rho^2\big(1-\eta_1(h)\alpha\big)$, or equivalently

$$y \geqslant \rho\big(1-\eta_1(h)\alpha\big). \tag{C.6}$$

Since offers must lie in $\mathcal{G}_\varepsilon$, this is exactly the response rule in (C.3). Therefore player 1's period-2 response strategy is optimal at every period-2 response history.

Period $t = 2$: player 2's proposal after rejection of $\langle 1, 0\rangle$. Consider the period-2 proposal history reached after player 2 rejects the initial offer $\langle 1, 0\rangle$. Player 1's observed behaviour is her prescribed initial proposal, so player 2's opponent-history likelihood is one and his updated ambiguity parameter is δ.

If player 2 proposes $y \in \mathcal{G}_\varepsilon$, player 1 accepts if and only if $y \geqslant \bar{y}$. This follows from (C.1): on the prescribed offer, $\eta_1(h) = \delta$, while after an unprescribed period-2 offer, $\eta_1(h) = 1$; in both cases the smallest accepted grid offer is $\bar{y}$.

If $y \geqslant \bar{y}$, player 2's ordinary continuation payoff is $\rho(1-y)$. If $y < \bar{y}$, player 1 rejects and player 2's ordinary continuation payoff under $\hat{\beta}_1$ is zero. Over player 1's pure continuation strategies, the worst payoff for player 2 is zero. The best payoff from proposing y is

$$\max\{\rho(1-y), \rho^2\}.$$

Therefore the value to player 2 of proposing y at this history is

$$V_2(y) = \begin{cases} (1-\delta)\rho(1-y) + \delta(1-\alpha)\max\{\rho(1-y), \rho^2\}, & \text{if } y \geqslant \bar{y}, \\ \delta(1-\alpha)\max\{\rho(1-y), \rho^2\}, & \text{if } y < \bar{y}. \end{cases} \tag{C.7}$$

For $y \geqslant \bar{y}$, the expression

$$(1-\delta)\rho(1-y) + \delta(1-\alpha)\max\{\rho(1-y), \rho^2\}$$

is weakly decreasing in y. Hence among accepted offers, player 2's value is maximized at $y = \bar{y}$.

For $y < \bar{y}$, the value is

$$\delta(1-\alpha)\max\{\rho(1-y), \rho^2\}.$$

Since $0 < \rho < 1$ and $1 - y \leqslant 1$, this is bounded above by $\delta(1-\alpha)\rho$, with equality at $y = 0$. The value of the prescribed offer $\bar{y}$ is

$$(1-\delta)\rho(1-\bar{y}) + \delta(1-\alpha)\max\{\rho(1-\bar{y}), \rho^2\}.$$

Dividing by $\rho > 0$, condition (7.8) is exactly the requirement that this value be at least $\delta(1-\alpha)\rho$. Therefore player 2 cannot profitably deviate to any rejected period-2 offer. Since $\bar{y}$ also maximizes the value among accepted offers, player 2's prescribed period-2 proposal is optimal after rejection of $\langle 1, 0\rangle$.

Period $t = 2$: player 2's proposal after rejection of any other initial offer. Now consider a period-2 proposal history reached after rejection of an initial offer $\langle x, 1-x\rangle$ with $x \in \mathcal{G}_\varepsilon \setminus \{1\}$. Player 1's observed initial offer then has zero likelihood under player 2's theory, so player 2's updated ambiguity parameter is one.

At such a history the additive component has zero weight. If player 2 proposes $y \in \mathcal{G}_\varepsilon$, the worst payoff over player 1's pure continuation strategies is zero and the best payoff is $\max\{\rho(1-y), \rho^2\}$. Thus the value of proposing y is

$$(1-\alpha)\max\{\rho(1-y), \rho^2\} = (1-\alpha)\rho\max\{1-y, \rho\}. \tag{C.8}$$

This is maximized at $y = 0$. Hence the prescribed offer $\langle 0, 1\rangle$ is optimal after every such history.

Period $t = 1$: player 2's response. Consider player 2's response to an initial offer $\langle x, 1-x\rangle$.

If $x = 1$, then player 1's observed behaviour has likelihood one under player 2's theory, and player 2's updated ambiguity parameter is δ. Acceptance gives player 2 the payoff zero. Rejection leads, under the prescribed continuation profile, to the period-2 offer $\langle \bar{y}, 1-\bar{y}\rangle$, which player 1 accepts. The resulting ordinary continuation payoff to player 2 is $\rho(1-\bar{y})$. The worst payoff over player 1's pure continuation strategies is zero, and the best is $\max\{\rho(1-\bar{y}), \rho^2\}$. Hence the value of rejecting is

$$(1-\delta)\rho(1-\bar{y}) + \delta(1-\alpha)\max\{\rho(1-\bar{y}), \rho^2\} \geqslant 0. \tag{C.9}$$

Thus rejection of $\langle 1, 0\rangle$ is optimal.

If $x \neq 1$, then player 1's observed initial offer has zero likelihood under player 2's theory, and player 2's updated ambiguity parameter is one. Acceptance gives player 2 the certain payoff $1 - x$. If player 2 rejects, his prescribed continuation offer is $\langle 0, 1\rangle$. Over player 1's pure continuation strategies, the worst payoff from rejection is zero and the best payoff is ρ, obtained if player 1 accepts the period-2 offer $\langle 0, 1\rangle$. Hence the value of rejection is

$$(1-\alpha)\rho. \tag{C.10}$$

It is therefore optimal for player 2 to accept if and only if $1 - x \geqslant (1-\alpha)\rho$, which is exactly the rule in (C.2). Thus player 2's period-1 response strategy is optimal at every initial offer.

Period $t = 1$: player 1's initial proposal. At the initial history, player 1's ambiguity parameter is δ.

If player 1 makes the prescribed initial proposal $\langle 1, 0\rangle$, player 2 rejects, then proposes $\langle \bar{y}, 1-\bar{y}\rangle$ in period $t = 2$, and player 1 accepts. The ordinary payoff to player 1 is therefore $\rho\bar{y}$. Over player 2's pure strategies, the worst payoff from this initial proposal is zero and the best payoff is one, obtained if player 2 accepts the initial offer. Hence the value of the prescribed initial proposal is

$$(1-\delta)\rho\bar{y} + \delta(1-\alpha). \tag{C.11}$$

Now consider an alternative initial proposal $\langle x, 1-x\rangle$ with $x \in \mathcal{G}_\varepsilon \setminus \{1\}$. Under player 2's prescribed response rule, this offer is accepted if $1 - x \geqslant (1-\alpha)\rho$. For such an accepted offer, player 1's ordinary payoff is x. Over player 2's pure strategies, the worst payoff is zero and the best payoff is $\max\{x, \rho\}$: player 2 can either accept the initial offer, giving player 1 the payoff x, or reject and make a period-2 offer giving player 1 at most ρ. Hence the value of such an accepted deviation is

$$(1-\delta)x + \delta(1-\alpha)\max\{x, \rho\}. \tag{C.12}$$

Among accepted alternative demands, this expression is weakly increasing in x. The largest such demand is x^A, as defined in (7.3). Hence every accepted alternative initial proposal has value at most $\overline{W}_1^A$. By condition (7.9), the prescribed initial proposal has value at least $\overline{W}_1^A$.

It remains only to consider alternative initial proposals that player 2 rejects. These are precisely the offers with $1 - x < (1-\alpha)\rho$, or equivalently $x \in \mathcal{X}^R$, where $\mathcal{X}^R$ is defined in (7.4). If $\mathcal{X}^R = \emptyset$, there are no rejected alternative initial proposals, and there is nothing further to verify.

Suppose, then, that $\mathcal{X}^R \neq \emptyset$. If player 1 makes a rejected alternative initial proposal, player 2 rejects and then makes the period-2 offer $\langle 0, 1\rangle$. Under the prescribed continuation strategy, player 1's ordinary continuation payoff is at most ρ^2. Over player 2's pure strategies, the worst payoff is zero and the best payoff is at most $\max\{x, \rho\}$: player 2 can accept the initial offer, giving player 1 the payoff x, or reject and later make an acceptable period-2 offer giving player 1 at most ρ. Hence the value of such a rejected deviation is bounded above by

$$(1-\delta)\rho^2 + \delta(1-\alpha)\max\{x, \rho\}. \tag{C.13}$$

Since $x \leqslant x^R$ for every $x \in \mathcal{X}^R$, every rejected alternative initial proposal has value at most $\overline{W}_1^R$. By condition (7.10), the prescribed initial proposal has value at least $\overline{W}_1^R$.

Combining the accepted-deviation and, when relevant, rejected-deviation comparisons, player 1 cannot profitably deviate from the prescribed initial proposal $\langle 1, 0\rangle$.

We have shown that at every bargaining history, the action prescribed by $\hat{\beta}_i$ maximizes player i's updated ambiguity-sensitive continuation payoff, given the opponent's behavioural strategy. Thus the one-step optimality condition in Definition 4.1 is satisfied for both players at every history. Therefore $\hat{\beta} = \langle \hat{\beta}_1, \hat{\beta}_2\rangle$ is a CP-EUA. Along its path, player 1 proposes $\langle 1, 0\rangle$ in period $t = 1$, player 2 rejects, player 2 proposes $\langle \bar{y}, 1-\bar{y}\rangle$ in period $t = 2$, and player 1 accepts. Agreement is therefore delayed until period $t = 2$. ■

References

Battigalli, P., Catonini, E., Lanzani, G., and Marinacci, M. (2019). Ambiguity attitudes and self-confirming equilibrium in sequential games. *Games and Economic Behavior*, 115:1–29.

Bose, S. and Daripa, A. (2009). A dynamic mechanism and surplus extraction under ambiguity. *Journal of Economic Theory*, 144:2084–2114.

Bose, S. and Renou, L. (2014). Mechanism design with ambiguous communication devices. *Econometrica*, 82(5):1853–1872.

Chateauneuf, A., Eichberger, J., and Grant, S. (2007). Choice under uncertainty with the best and worst in mind: Neo-addditive capacities. *Journal of Economic Theory*, 137(1):538 –567.

Eichberger, J., Grant, S., and Kelsey, D. (2007). Updating choquet beliefs. *Journal of Mathematical Economics*, 43:888–899.

Eichberger, J., Grant, S., and Kelsey, D. (2010). Comparing three ways to update Choquet beliefs. *Economics Letters*, 107:91–94.

Eichberger, J. and Kelsey, D. (1999). Education signalling and uncertainty. In Nau, Grønn, and Machina, editors, *Beliefs Interactions and Preferences*. Kluwer.

Eichberger, J. and Kelsey, D. (2004). Sequential two-player games with ambiguity. *International Economic Review*, 45:1229–1261.

Eichberger, J. and Kelsey, D. (2014). Optimism and pessimism in games. *International Economic Review*, 55(2):483–505.

Forsythe, R., Kennan, J., and Sopher, B. (1991). Dividing a shrinking pie: An experimental study of strikes in bargaining games with complete information. *Research in Experimental Economics*, 4:223–268.

Hanany, E., Klibanoff, P., and Mukerji, S. (2020). Incomplete information games with ambiguity averse players. *American Economic Journal: Microeconomics*, 12:135–187.

Jehiel, P. (2005). Analogy-based expectation equilibrium. *Journal of Economic Theory*, 123:81–104.

Kaneko, M. and Ishikawa, R. (2023). A resolution of the centipede paradox. *Discussion Paper Series No. 1388 in University of Tsukuba,*.

Karni, E. and Safra, Z. (1989). Ascending bid auctions with behaviorally consistent bidders. *Annals of Operations Research*, 19:435–446.

Kilka, M. and Weber, M. (2001). What determines the shape of the probability weighting function under uncertainty? *Management Science*, 47:1712–1726.

Kreps, D., Milgrom, P., Roberts, J., and Wilson, R. (1982). Rational cooperation in the finitely repeated prisoner's dilemma. *Journal of Economic Theory*, 27:253–279.

Krokow, E. M., Colman, A. M., and Pulford, B. (2016). Cooperation in repeated interactions: A systematic review of centipede game experiments, 1992-2016. *European Review of Social Psychology*, 27:231–282.

McKelvey, R. D. and Palfrey, T. R. (1992). An experimental study of the centipede game. *Econometrica*, 60:803–836.

Monet, B. and Vergopoulos, V. (2024). Ambiguity, randomization and the timing of resolution of uncertainty. *Economic Theory*, 78:1021–1045.

Osborne, M. J. and Rubinstein, A. (1994). *A Course on Game Theory.* MIT Press, Cambridge, Massachusetts.

Pahlke, M. (2022). Dynamic consistency in incomplete information games with multiple priors. *Games and Economic Behavior*, 133:85–108.

Rosenthal, R. (1981). Games of perfect information, predatory pricing, and the chain store. *Journal of Economic Theory*, 25(1):92–100.

Roth, A. E. (1995). Bargaining experiments. In Kagel, J. H. and Roth, A. E., editors, *Handbook of Experimental Economics*, chapter 4, pages 253–348. Princeton University Press.

Rothe, J. (2011). Uncertainty aversion and equilibrium in extensive games. In Petrosyan, L. A. and Zenevich, N. A., editors, *Contributions to game theory and management*, volume IV of *The Fourth International Conference Game Theory and Management.*, pages 389–406, St Petersburg, Russia. St Petersburg University Press.

Rubinstein, A. (1982). Perfect equilibrium in a bargaining model. *Econometrica*, 50(1):97–109.

Ryan, M. J. (2002). What do uncertainty-averse decision-makers believe? *Economic Theory*, 20:47–65.

Siniscalchi, M. (2011). Dynamic choice under ambiguity. *Theoretical Economics*, 6:379–421.

Stahl, I. (1972). *Bargaining Theory.* Stockholm School of Economics. Stockholm.

Stauber, R. (2011). Knightian games and robustness to ambiguity. *Journal of Economic Theory*, 146(1):248–274.

Stauber, R. (2017). Irrationality and ambiguity in extensive games. *Games and Economic Behavior*, 102:409–432.

Strotz, R. (1955). Myopia and inconsistency in dynamic utility maximization. *Review of Economic Studies*, 23:165–180.