Sleeping Newcomb

John W. E. Cremin[*†]

Aix-Marseille University, CNRS, AMSE

First version: 31^{st} of March 2026
Latest Version: July 5, 2026

Abstract

I study games with self-locating uncertainty in which an agent at a single information set is uncertain of his position even *within* a given information set and play of the game. In such games, there is an analogy to be drawn with Newcomb's problem: in both settings, locally rational (thirder) reasoning and globally optimal (planning) reasoning can prescribe different strategies. I call this a *Newcomb tension*, and present a representation theorem: a Bayesian with commitment power and an uncommitted agent holding incorrect 'one-boxer' beliefs are behaviourally equivalent, and I characterise exactly when such beliefs exist. In the single-agent case, randomisation always resolves the tension but in multi-agent games, in which planning and interim social weights diverge under some conditions, a multi-agent Newcomb tension can survive this randomisation resolution when there is 'payoff interdependence' or an asymmetric awakening structure (number of decision nodes) across agents. I consider the implications of this for the duplicating Sleeping Beauty problem, and a duplicating variant of the absent-minded driver.

Keywords— Absent-Minded Driver, Sleeping Beauty Problem, Newcomb's Problem, Self-Locating/Indexical Uncertainty, Imperfect Recall, Dynamic Consistency

JEL Codes— C72, D71, D81, D83

1 Introduction

Bored, one Sunday morning, of debating the virtues of a minimalist state, Nozick has decided to make a foray into Bayesian epistemology. He throws aside the draft of his latest letter to Rawls, heads to the blackboard, and draws the table in Figure 1. This is a representation of the *duplicating Sleeping Beauty problem.* In it, some scientists toss a fair coin (θ) on Sunday ($d = 0$), before putting Sleeping Beauty to sleep with a memory-erasing drug, and waking her on Monday ($d = 1$). They then, having chatted to her for a bit, put her to sleep again with the drug, and *if the coin landed tails*, proceed to wake the clone on Tuesday. The controversy of the canonical variant

[*]Email: `john-walter-edward.cremin@univ-amu.fr`. The project leading to this publication has received funding from the French government under the "France 2030" investment plan managed by the French National Research Agency (reference: ANR-17-EURE-0020) and from Excellence Initiative of Aix-Marseille University – A*MIDEX.
[†]Thanks to participants of the 2026 France-Spain Theory Meeting for helpful comments and suggestions.

(in which Beauty is simply awakened twice when $\theta = T$ rather than cloned[1]) is over whether Beauty should believe the coin has landed heads with probability $\frac{1}{2}$ or $\frac{1}{3}$ upon awakening, but here Nozick's interest is instead piqued by a different question. On the conventional debate, the *thirders* seem to have probably got the better end of the stick, but despite this, and despite the equivalence between the information structures across canonical and duplicating variants, he is not sure that Beauty should *act* on the basis of those thirder beliefs here.

Agent & day	$\theta = H(=0)$	$\theta = T(=1)$
$n = 1, d = 1$	•	•
$n = 2, d = 2$		•

Figure 1: Dot diagram for Duplicating Sleeping Beauty. Agent $n = 1$ (the original) wakes in both states; agent $n = 2$ (the clone) wakes only in $\theta = T$. Both agents share a single information set and cannot distinguish their awakenings. Upon waking up, the agent chooses a 'bet' $x \in [0, 1]$, and receives utility $-(x - \theta)^2$. They are expected utility maximisers.

Being an avid reader, he is also familiar with Harsanyi (1955) and his *Social Aggregation Theorem*, and decides that when it comes to deciding the optimal action, this is really more a social choice problem than a purely epistemological puzzle. Imagining Beauty behind a *veil of ignorance*, to borrow the Rawlsian terminology, he decides that one should balance the interests of the original and the clone. To do this, again following Harsanyi (1955), he assumes co-cardinal utilities, adopts a utilitarian social welfare function, and weights the original Beauty and the clone according to their probability-shares of the information set.

However, here he notices a problem. *Which* probabilities should she use? If she tries to decide the optimal solution on Sunday, she will want to assign weight $\frac{1}{4}$ to the clone's interests, since with 50% chance the clone will not wake up at all, and with 50% chance will be one of two agents waking up and meriting equal weight. Upon waking up however, she will disagree with her earlier weighting, as she will believe that each of the three dots should be given weight $\frac{1}{3}$ in line with her updated thirder beliefs: the clone should have social weight $\frac{1}{3}$. The changing social welfare weights imply that the optimal action is different when selected before the game begins than when selected during it. With this observation, he notices an eerie resemblance to his earlier Newcomb's problem.[2]

In Newcomb's problem, which is a paradox in decision theory, adherents of one of the famous responses, known as *two-boxing*, find themselves in the peculiar situation that though they consider it optimal to take one course of action (that is, to *two-box*), they would also choose to commit themselves to the alternative course of action (to *one-box*) before the problem if they could. I call this divergence a *Newcomb tension*, due to this analogy. In both cases (Newcomb and Duplicating Sleeping Beauty), the agent at the point of decision faces a tension between *interim rationality* (act on your current beliefs) and *planning rationality* (act as you would have committed to act before arriving at the information set). This Newcomb tension is not confined to the duplicating Sleeping Beauty and Newcomb problems, but can occur generally in the presence of *self-locating uncertainty*, which is one's uncertainty over one's own location within an information set when it contains multiple decision nodes on a given play of the game. The absent-minded driver (AMD) of Piccione

[1] C.f. Appendix A.1 for more discussion of this problem if it is unfamiliar to you.

[2] C.f. Appendix A.2 for readers unfamiliar with this problem. Though the elements mentioned in the following paragraph are all that are necessary to follow the argument, more familiarity with it will help.

and Rubinstein (1997) (c.f. Appendix A.3 for a discussion of this) exhibits a similar structure when randomisation is not permitted: the driver at the information set, reasoning with correct beliefs about the intersection at which they find themselves, does not consider the planning optimal action to be interim optimal. Beyond this, as was famously noted by Piccione and Rubinstein (1997) themselves in the special issue posing and discussing the problem, no interim optimal action can even be said to exist in that case.

The purpose of this paper is to study systematically under what conditions a Newcomb tension arises in games with self-locating uncertainty, which I study more generally in a companion paper: *Sleeping Beauty's Dismal Day Out*. When does the 'awakening structure' (set of decision nodes on each path of play within each information set; e.g. Figure 1) of a game create a divergence between planning-optimal and interim-optimal actions? I restrict attention to games with a single information set, allowing multiple agents to occupy 'dots' (decision nodes) at said set, and distinguishing intra-personal self-locating uncertainty (as in the absent-minded driver) from the inter-personal variety (as in Duplicating Sleeping Beauty). After defining the Newcomb tension, I prove a representation theorem (Theorem 2.2): a committed thirder is behaviourally equivalent to an uncommitted agent holding "one-boxer" beliefs, so commitment power can be modelled as a belief rather than as an explicit planning stage. Such beliefs exist exactly when no mixture of actions beats the planning optimum at every awakening it reaches: always, for a single agent choosing amongst behavioural strategies, though the guarantee can fail when multiple agents' payoffs are interdependent.

The paper's main results concern when the Newcomb tension can and cannot be resolved. In the single-agent case, behavioural strategies (randomisation) always achieve this: the planning-optimal behavioural strategy is also interim-optimal (Theorem 3.1). In the multi-agent case, I approach the question of what action an agent should choose as a social choice problem, aggregating payoffs with a utilitarian social welfare function. The central result (Theorem 4.1) is that the Newcomb tension requires either payoff interdependence or an asymmetric dot structure across agents, and generically exists whenever one or the other is present. In the duplicating absent-minded driver (a variant of the AMD in which two separate individuals act at the different intersections), Proposition 4.3 characterises when and by how much the planning and interim optima diverge. In the duplicating Sleeping Beauty problem, my approach implies that while thirder beliefs remain the correct credences, Beauty should not bet on them.

My interest in games with self-locating uncertainty is driven primarily by a desire to study models of social and individual learning involving *temporal ignorance*, where agents arrive in a game at a given time, but only observe some signal of this arrival time rather than the time itself. My earlier article *Bot Got Your Tongue?* features this, though in that instance it is not the main focus of the paper. Though the Sleeping Beauty and absent-minded driver problems, and especially their duplicating variants, may seem arcane problems of no practical application, their purpose is to aid the study of self-locating uncertainty generally, so that it can be applied to this applied-theoretical purpose.

2 Games with Self-Locating Uncertainty

The Newcomb tension described above arises in a class of extensive-form games in which an agent at an information set cannot determine which node they occupy, and in which this ignorance is greater than in standard games of imperfect information. I now formalise this class. The setting

is deliberately simple: I consider games with a *single* information set at which a single action must be chosen, but I allow *multiple agents* to occupy nodes at that information set, and allow individual terminal histories ('plays') of the game to feature multiple nodes within it. This is the simplest environment in which self-locating uncertainty and the resulting Newcomb tension can be studied in isolation, free from complications introduced by sequential play across multiple information sets. I also assume throughout that agents have *interpersonally comparable* (co-cardinal) von Neumann-Morgenstern utility. This assumption is without loss for single-agent games, but is essential for the multi-agent welfare analysis in Section 4.

The formulation follows the *primitive extensive form* of the companion paper *Sleeping Beauty's Dismal Day Out*, itself a modification of Kreps and Wilson (1982) that drops the requirements of perfect recall, and exclusive occupancy (one agent per information set).

2.1 The Primitive Extensive Form

The physical order of play is represented by a finite set of nodes T together with a binary relation $\prec$ of *precedence* on T. Following Kreps and Wilson (1982), I require $(T, \prec)$ to form an *arborescence*: $\prec$ is a partial order, and the predecessors of each $t \in T$ are totally ordered by $\prec$. I partition T into terminal nodes Z (outcomes, having no successors), decision nodes $X = T \setminus Z$, and initial nodes Θ (states, having no predecessors). For $t \notin \Theta$, I write $p_1(t)$ for the immediate predecessor of t. For $x \in X$ I write $S(x)$ for the set of immediate successors. Choices are represented by a finite set of *actions* A and a function $\alpha : T \setminus \Theta \to A$, where $\alpha(t)$ denotes the action taken to reach t from $p_1(t)$, with α one-to-one on $S(x)$ for each $x \in X$.

Definition 2.1 (Single-Information-Set Game with Self-Locating Uncertainty)**.** A *single-information-set SLU game* is a collection $\Gamma = \{T, \prec; A, \alpha; N, \iota; h; u, \rho\}$ where $(T, \prec)$ and (A, α) are as above, N is a finite set of *agents*, $\iota : h \to N$ is an *agent assignment* specifying which agent occupies each node in the information set, $h \subseteq X$ is the unique information set, $u : Z \times N \to \mathbb{R}$ is a payoff function assigning a co-cardinal von Neumann-Morgenstern utility to each terminal node for each agent, and ρ is a strictly positive probability measure on Θ (the prior). The information set satisfies:

(A1) *(Action consistency)* $\quad x, x' \in h \;\Rightarrow\; \alpha(S(x)) = \alpha(S(x'))$.

The common feasible action set is $A(h) = \alpha(S(x))$ for any $x \in h$. Perfect recall is not imposed: an agent may have multiple nodes in h that lie on the same play of the game. When $|N| = 1$, the game is a *single-agent* SLU game and the agent assignment is trivial.

The decision nodes $X \setminus h$ not belonging to the information set are for *nature*: their successors are determined by ρ, not by any agent's choice. In games without a nature move, $h = X$. A *play* of the game is a maximal chain $\theta = t_0 \prec t_1 \prec \cdots \prec t_k \in Z$ with $\theta \in \Theta$. Two nodes $x, x' \in T$ are *co-reachable* if there exists a play that passes through them both.

Definition 2.2 (Self-Locating Uncertainty)**.** The information set h exhibits *self-locating uncertainty* if it contains distinct co-reachable nodes that the occupying agent(s) cannot distinguish. Two forms arise:

(a) *Intra-personal self-locating uncertainty:* there exist co-reachable nodes $x, x' \in h$ with $\iota(x) = \iota(x')$. A single play of the game passes through h more than once, and the *same* agent occupies both nodes. This is the structure of the absent-minded driver, and in the language of Osborne and Rubinstein (1994), this is *absent-mindedness*, and not *forgetfulness*.

(b) *Inter-personal self-locating uncertainty:* there exist co-reachable nodes $x, x' \in h$ with $\iota(x) \neq \iota(x')$. A single play of the game passes through h more than once, and *different* agents occupy the two nodes. Each agent cannot distinguish their own node(s) from the other agents' nodes. This is the structure of Duplicating Sleeping Beauty: in the tails world, the original and the clone are visited sequentially on the same path of play, yet share an information set.

A game may exhibit both forms simultaneously.

The qualifier "co-reachable" is essential for self-locating uncertainty. In any extensive form game with imperfect information, a player may have multiple nodes in the same information set that lie on different branches of the tree. Such nodes are not co-reachable: no single play visits both. This is standard imperfect information and does not constitute self-locating uncertainty. Intra-personal self-locating uncertainty arises only when a single play can pass through the agent's information set more than once, as in the absent-minded driver of Piccione and Rubinstein (1997). The perfect recall condition (KW2) of Kreps and Wilson (1982) rules out this possibility.

An *action profile* at h is a function $\mathbf{a} : h \to A(h)$ assigning an action to each node in h; when the agent plays a behavioural strategy, the realised actions at different nodes may differ (though not if they play a pure or mixed strategy). For each state $\theta \in \Theta$ and action profile $\mathbf{a}$, let $D(\theta, \mathbf{a})$ denote the set of nodes in h that are visited on a play beginning at θ under $\mathbf{a}$ (i.e. the *dots* or *awakenings* in state θ under profile $\mathbf{a}$). Under a pure strategy in which the same action a is played at every node in h, I write $D(\theta, a)$ for this set. The game payoff in state θ under action profile $\mathbf{a}$ is $U(\mathbf{a}, \theta)$; under the pure strategy a, this is $U(a, \theta)$. $D(\theta, a)$, much like $U(\mathbf{a}, \theta)$, depends only on on-path actions.[3] Accordingly, I identify action profiles that agree at every visited dot, and sums over $\mathbf{a}$ are taken over the resulting equivalence classes; with this convention, $\Pr_\sigma(\mathbf{a}) = \prod_{d \in D(\theta, \mathbf{a})} \sigma(\mathbf{a}(d))$ defines a probability distribution over profiles in each state.

2.2 The Planning Stage

Before the game proper begins, I introduce a *planning stage* information set h_0, which precedes both the resolution of uncertainty and the main information set h. The planning stage precedes nature's choice of the state of the world: the temporal order is

$$h_0 \longrightarrow \text{Nature draws } \theta \in \Theta \text{ according to } \rho \longrightarrow h.$$

At h_0, the available actions are either $A(h_0) = \{\text{Commit}, \text{Not Commit}\}$ or $A(h_0) = \varnothing$ (the planning stage is inoperative). If $A(h_0) = \{\text{Commit}, \text{Not Commit}\}$, then choosing *Commit* binds the agent irrevocably to the 'optimal'(in the view of the planning agent) strategy at h before the state is realised, while *Not Commit* leaves the agent free to optimise at h given their beliefs upon arrival. If $A(h_0) = \varnothing$, the planning stage plays no role and the game begins directly with nature's move.

[3] If in the absent-minded driver the action-profile chooses 'Exit' at the first intersection, it does not make a difference what actions are then specified, there is only one 'dot'

The planning stage serves two purposes. Firstly, it makes the *value of commitment* concrete: the Newcomb tension is precisely the situation in which the agent at h_0 would strictly prefer to Commit, because the planning-optimal action differs from the interim-optimal action at h. Secondly, it allows a direct representation of Newcomb's Game, which is a modification of *Newcomb's Problem* (I provide a review of this in Appendix A.2), altered so that the state of the world is a straightforward function of $A(h_0)$.

Example (Newcomb's Game). Let $I = \{1\}$, with $A(h_0) = \{\text{Commit, Not Commit}\}$ and $A(h) = \{\text{One-box, Two-box}\}$. The state of the world $\theta \in \Theta = \{\text{Full}, \text{Empty}\}$ represents whether the opaque box contains \$1,000,000 or \$0, and is chosen by a predictor *as a function of the action at* h_0: if the agent commits to one-boxing, $\theta = \text{Full}$; if not, $\theta = \text{Empty}$. At h_0, the agent can guarantee a high expected payoff by committing to one-box. At h, two-boxing is a dominant strategy for the interim-agent. This is the Newcomb tension: $a^{\text{plan}} = \text{One-box} \neq a^{\text{int}} = \text{Two-box}$.

This is the only example in the paper where the state of the world depends on the action at the planning stage. In all other settings, and for all the formal results and definitions that follow, the state θ is drawn independently of h_0. The Newcomb tension in self-locating uncertainty games arises not because the agent's plan causally influences the state (as the predictor does in Newcomb's game), but because the awakening structure at h distorts the agent's beliefs about a state that is genuinely independent.

Note that this commitment/flexibility trade-off is reminiscent of but distinct from that of the principle behind mutually assured destruction. In that and similar settings, the advantage of commitment is that it allows deterrence and keeps certain decision nodes off-path. Conditional on actually arriving at such a node, planning and interim agents do not have different preferences. The Newcomb tension is about when the planning and interim agents disagree on the optimal action at *on-path* nodes.

Social welfare in multi-agent games. When $|N| > 1$, the planning agent must aggregate the payoffs of different agents. I assume that agents behind a veil of ignorance adopt a *utilitarian social welfare function* (SWF). For each agent $n \in N$, state $\theta \in \Theta$, and action profile $\mathbf{a} \in A(h)^{|h|}$, let $D(\theta, n, \mathbf{a}) \subseteq h$ denote the set of dots that agent n occupies in state θ when the actions realised along the path of play are given by $\mathbf{a}$, and let $|D(\theta, \mathbf{a})| = \sum_{n'} |D(\theta, n', \mathbf{a})|$ be the total number of dots. I assume throughout that $|D(\theta, \mathbf{a})|$ is deterministic for each state–action-profile pair $(\theta, \mathbf{a})$. In the Sleeping Beauty problem the number of awakenings is determined by the state alone; in the absent-minded driver it depends on the action profile; in both cases it is a fixed (non-random) function of $(\theta, \mathbf{a})$.

The planner treats each state-action-profile pair $(\theta, \mathbf{a})$ as a compound state with probability $\rho(\theta) \Pr_\sigma(\mathbf{a})$,[4] and within each such compound state distributes probability uniformly across its $|D(\theta, \mathbf{a})|$ dots. This formalises my treatment of duplicating Sleeping Beauty in the Introduction. Agent n's social weight is then the total probability falling on n's dots:

$$\lambda_n(\sigma) \;=\; \sum_{\theta \in \Theta} \sum_{\mathbf{a}} \rho(\theta) \Pr_\sigma(\mathbf{a}) \, \frac{|D(\theta, n, \mathbf{a})|}{|D(\theta, \mathbf{a})|}.$$

[4]Here $\Pr_\sigma(\mathbf{a}) = \prod_{d \in D(\theta, \mathbf{a})} \sigma(\mathbf{a}(d))$ is the probability that σ realises the on-path action profile $\mathbf{a}$.

These weights sum to one across all agents and reduce to $\lambda_n = 1/|N|$ when all agents share the same dots in every $(\theta, \mathbf{a})$. The planning payoff evaluates social welfare within each compound state $(\theta, \mathbf{a})$ using that compound state's social weights, then averages across compound states:

$$V^{\text{plan}}(\sigma) \;=\; \sum_{\theta\in\Theta}\sum_{\mathbf{a}} \rho(\theta)\, \Pr_{\sigma}(\mathbf{a}) \sum_{n\in N} \frac{|D(\theta, n, \mathbf{a})|}{|D(\theta, \mathbf{a})|}\, U_n(\mathbf{a}, \theta).$$

When $|N| = 1$ this reduces to the single-agent planning objective. The key feature of multi-agent games is that the interim agent at h uses *thirder* probabilities to compute the analogous weights, which generically differ from the planner's weights; this is the source of the Newcomb tension studied in Section 4. Note that the planner's social weights themselves vary with σ whenever the awakening structure is endogenous, as in the AMD.

2.3 The Newcomb Tension

To define the Newcomb tension formally, we must define the beliefs and optimal strategies of agents at planning and interim stages.

Definition 2.3 (Planning-Optimal Strategy)**.** A *behavioural strategy* $\sigma \in \Delta(A(h))$ specifies the probability $\sigma(a)$ with which the agent independently plays action a at each visit to h. The *planning-optimal strategy* maximises the ex ante expected game payoff:

$$\sigma^* \;=\; \arg\max_{\sigma\in\Delta(A(h))} V^{\text{plan}}(\sigma), \qquad V^{\text{plan}}(\sigma) \;=\; \sum_{\theta\in\Theta}\sum_{\mathbf{a}} \rho(\theta)\, \Pr_{\sigma}(\mathbf{a})\, U(\mathbf{a}, \theta).$$

When I restrict attention to pure strategies, I write $a^{\text{plan}} = \arg\max_{a\in A(h)} V^{\text{plan}}(a)$ with $V^{\text{plan}}(a) = \sum_\theta \rho(\theta)\, U(a, \theta)$.

I refer to the Bayesian beliefs an agent would form ex-interim as 'thirder' beliefs, in reference to the Sleeping Beauty problem. One can still find epistemologists who argue the halfer position, but the literature has largely settled on the thirder stance, and indeed the absent-minded driver literature in game theory implicitly assumes this position too. The belief assignment itself is not new: it is the *consistent* assessment of Piccione and Rubinstein (1997), formalised by Grove and Halpern (1997) as the distribution placing weight on each node in proportion to its reach probability, equivalently the long-run frequency with which the agent occupies it (if we imagine the game being played many times). The *halfer* assignment likewise coincides with the *Z-consistent*, or outcome-uninformative, belief of Piccione and Rubinstein (1997) and Grove and Halpern (1997). Ross (2010) later states generalised halfer and thirder principles that show how to follow halfer and thirder reasoning in more general situations of state-dependent awakening. What the present setting demands, however, is to combine two ingredients that these treatments keep apart: Grove and Halpern (1997) work within a single fixed game with no exogenous states of the world, while Ross (2010) allows the probability of awakening to depend on the state but not on the agent's own action profile. In general the two cannot be separated, since the number and identity of an agent's dots depend jointly on the state θ and the realised action profile $\mathbf{a}$. The definition below therefore forms beliefs over which dot d is occupied *conditional on a full state–action-profile tuple* $(\theta, \mathbf{a})$, with agents reasoning *in a given equilibrium* σ by assuming that all agents (and temporal parts of agents) act according to the equilibrium strategy-profile.

Definition 2.4 (Thirder Beliefs)**.** If strategy σ is being played, a *thirder* assigns to each pair $(\theta, \mathbf{a})$ a probability proportional to the prior times the probability of the action profile times the number of awakenings:

$$\pi(\theta, \mathbf{a}) \;=\; \frac{\rho(\theta)\ \Pr_\sigma(\mathbf{a})\,|D(\theta, \mathbf{a})|}{\displaystyle\sum_{\theta', \mathbf{a}'} \rho(\theta')\ \Pr_\sigma(\mathbf{a}')\,|D(\theta', \mathbf{a}')|}.$$

The marginal over states is $\pi(\theta) = \sum_{\mathbf{a}} \pi(\theta, \mathbf{a})$. When $|D(\theta, \mathbf{a})|$ does not depend on $\mathbf{a}$ (as in Sleeping Beauty), this reduces to $\pi(\theta) = \rho(\theta)|D(\theta)|/C$ with $C = \sum_{\theta'} \rho(\theta')|D(\theta')|$. The agent forms beliefs over dots visited conditional on a given state-action-profile pair $(\theta, \mathbf{a})$ by distributing the probability uniformly over all these dots.

Definition 2.5 (Interim-Optimal Strategy)**.** The *interim-optimal strategy* for a thirder is the behavioural strategy $\sigma^{\text{int}} \in \Delta(A(h))$ that maximises the expected continuation payoff at h under thirder beliefs. The agent at a node in h chooses their action a' at the current node to maximise:

$$V^{\text{int}}(a') \;=\; \sum_{\theta,\, \mathbf{a}} \pi(\theta, \mathbf{a})\ \frac{1}{|D(\theta, \mathbf{a})|} \sum_{d \in D(\theta, \mathbf{a})} v_d(a',\, \mathbf{a}_{-d},\, \theta),$$

where $v_d(a', \mathbf{a}_{-d}, \theta)$ is the continuation payoff at dot d when the agent plays a' at d, the actions at all other dots are as in $\mathbf{a}_{-d}$, and the state is θ. Since the thirder beliefs depend on the strategy σ through $\Pr_\sigma(\mathbf{a})$, the interim-optimal strategy is a fixed point: σ^{int} is interim-optimal if, when all temporal parts play σ^{int}, each finds σ^{int} optimal given the resulting beliefs.

Having defined these objects, I can now formally define the concept of a Newcomb tension:

Definition 2.6 (Newcomb Tension)**.** A single-information-set SLU game Γ exhibits a *Newcomb tension* if $\sigma^* \neq \sigma^{\text{int}}$: the planning-optimal and interim-optimal strategies diverge. The agent at the information set, reasoning correctly given their epistemically justified beliefs, chooses a strategy that is suboptimal from the planning perspective. The *value of commitment* is $V^{\text{plan}}(\sigma^*) - V^{\text{plan}}(\sigma^{\text{int}}) > 0$.

Note that for a single agent with dot-independent continuation payoffs (Section 2.4), the Newcomb tension can only arise when $|D(\theta, \mathbf{a})|$ varies across states or action profiles: if the number of awakenings is constant across all $(\theta, \mathbf{a})$, the thirder marginal over states is proportional to ρ, and the interim and planning objectives share a maximiser. When continuation payoffs are dot-dependent, the position of the dot is itself a source of divergence, as in the absent-minded driver; and with multiple agents, a constant awakening count still leaves open the interdependence channel of Theorem 4.1, since the occupant of a dot maximises their own payoff while the planner aggregates across agents.

The *expectation paradox* of Grove and Halpern (1997), in which an agent's ex ante valuation of a game differs from the valuation he forms once at the information set, is an instance of the same phenomenon, cast as a change in *valuation* rather than of optimal choice. The two are tightly linked: it is precisely the revaluation at the information set that shifts the optimal action, so the valuation form and the choice form are two readings of a single tension, and their single-agent absent-minded driver is the relevant case. The Newcomb tension as defined here is more general in two respects: it accommodates multiple agents (Section 4), and it allows the state of the world itself to be determined by the planning-stage commitment, as in Newcomb's game (Appendix A.2).

A second contrast is that the Newcomb tension can be dissolved by enlarging the strategy space: in the single-agent case the planning optimum becomes interim-optimal once behavioural strategies are admitted (Theorem 3.1), with no change of beliefs. The expectation paradox, by contrast, is untouched by randomisation: at the behavioural optimum the interim and ex ante valuations still diverge under thirder beliefs, closing only if the agent adopts the outcome-uninformative belief instead (with each node valued, as in their framework, by the full payoff of the play passing through it, rather than by its continuation payoff). The two are resolved by different levers: strategy space versus belief. (In the multi-agent case even this lever is not decisive: the tension can survive randomisation, as the duplicating absent-minded driver shows.)

2.4 One-Boxer Beliefs: A Representation Theorem

I now formalise the relationship between commitment and beliefs. The central observation is that a Bayesian (thirder) agent who can commit to the planning-optimal action is behaviourally indistinguishable from an agent without commitment who holds different beliefs. I call these beliefs *one-boxer beliefs*, by analogy with one-boxing in Newcomb's problem.

To state the definition, I use the *continuation payoffs* already introduced in the interim-optimal strategy above. Recall that $v_d(a', \mathbf{a}_{-d}, \theta)$ denotes the payoff realised from dot d onwards when the agent plays a' at d, the actions at all other dots are $\mathbf{a}_{-d}$, and the state is θ. The continuation payoffs are *dot-independent* if $v_d(a', \mathbf{a}_{-d}, \theta)$ does not depend on d or on $\mathbf{a}_{-d}$ for any given a' and θ. In many games of interest the game payoff depends on the action and the state but not on which dot the agent occupies, so $v_d(a', \mathbf{a}_{-d}, \theta) = U(a', \theta)$ for all d and $\mathbf{a}_{-d}$, and the continuation payoffs are trivially dot-independent. This is the case in the duplicating Sleeping Beauty problem of the Introduction, where each agent is scored at their single awakening.[5] In other games, notably the absent-minded driver, the payoff depends on *where* in the game tree the action is taken: exiting at the first intersection yields payoff 0, while exiting at the second yields payoff 4.

Definition 2.7 (One-Boxer Beliefs)**.** A joint probability distribution P^{OB} over the triples $(\theta, \mathbf{a}, d)$ *reached under* σ^*, i.e. those with $\Pr_{\sigma^*}(\mathbf{a}) > 0$ and $d \in D(\theta, \mathbf{a})$, gives the *one-boxer beliefs for the planning-optimal strategy* σ^* if an interim agent holding P^{OB} and choosing their action finds σ^* optimal. Concretely, define the interim expected continuation payoff of

[5] In the canonical Sleeping Beauty problem, if the agent is scored at *each* awakening, continuation payoffs are not dot-independent: the continuation payoff at the Monday awakening in Tails includes the Tuesday score, and so depends on $\mathbf{a}_{-d}$. No difficulty arises for the results below (Theorem 3.1 covers the game regardless), but the universal-marginal property of Proposition 2.1(i) is specific to dot-independent payoffs: with per-awakening scoring the planner weights states by their awakening counts, and the prior marginal is no longer one-boxer.

playing a' at the current dot:

$$V^{\mathrm{OB}}(a') \;=\; \sum_{\theta,\,\mathbf{a},\,d} P^{\mathrm{OB}}(\theta, \mathbf{a}, d)\, v_d(a',\, \mathbf{a}_{-d},\, \theta).$$

Then P^{OB} is one-boxer if:

(a) for every a' in the support of σ^*, $V^{\mathrm{OB}}(a') \geq V^{\mathrm{OB}}(a'')$ for all $a'' \in A(h)$; and

(b) if σ^* is non-degenerate, then $V^{\mathrm{OB}}(a') = V^{\mathrm{OB}}(a'')$ for all a', a'' in the support of σ^* (indifference).

I write $P^{\mathrm{OB}}(\theta) = \sum_{\mathbf{a},d} P^{\mathrm{OB}}(\theta, \mathbf{a}, d)$ for the marginal over states.

This next definition simply notes that the thirder beliefs I have already defined lead the agent to select an interim-optimal strategy, and thus can also be called *two-boxer* beliefs to complete the Newcomb analogy that motivates the definition of *one-boxer* beliefs. In games with a Newcomb tension, interim EU-maximisation under two-boxer beliefs selects a strategy that is not planning-optimal.

Definition 2.8 (Two-Boxer (Bayesian) Beliefs)**.** The *two-boxer* or *Bayesian* beliefs at h are the thirder beliefs $\pi(\theta, \mathbf{a})$ defined in Definition 2.4.

Proposition 2.1 (Non-uniqueness of One-Boxer Beliefs)**.** One-boxer beliefs are not unique; in general, many conditionals over dots will do.

(i) **(Universal marginal.)** Suppose continuation payoffs are dot-independent: $v_d(a', \mathbf{a}_{-d}, \theta) = v(a', \theta)$ for all d, $\mathbf{a}_{-d}$, and θ. Then the prior marginal $P^{\mathrm{OB}}(\theta) = \rho(\theta)$, combined with *any* conditional over action profiles and dots, constitutes one-boxer beliefs for every dot-independent payoff specification v.

(ii) **(Halfer conditionals need not suffice.)** When continuation payoffs are dot-dependent, the conditional beliefs $P^{\mathrm{OB}}(d \mid \theta, \mathbf{a})$ matter, and the uniform ("halfer") conditional $P^{\mathrm{OB}}(d \mid \theta, \mathbf{a}) = 1/|D(\theta, \mathbf{a})|$ need not be one-boxer. In the absent-minded driver (restricting to pure strategies for illustration), the pure strategy Continue yields two dots $D = \{X, Y\}$. Halfer beliefs assign $P^{\mathrm{OB}}(X) = P^{\mathrm{OB}}(Y) = \frac{1}{2}$, but one-boxer beliefs require $P^{\mathrm{OB}}(X) \geq \frac{3}{4}$.

Proof. See Appendix B.1. □

Theorem 2.2 (Representation Theorem with Existence)**.** Let σ^* be a planning-optimal strategy (pure or behavioural). Say that a mixed action $\beta = \sum_a \eta(a)\, \delta_a$, with η a probability distribution over $A(h)$, *dot-wise dominates* σ^* if

$$v_d(\beta,\, \mathbf{a}_{-d},\, \theta) \;>\; v_d(\sigma^*,\, \mathbf{a}_{-d},\, \theta)$$

at every triple $(\theta, \mathbf{a}, d)$ reached under σ^*, where v_d denotes the continuation payoff of the occupant $\iota(d)$.

(i) **(Existence iff.)** The set of one-boxer beliefs for σ^* is non-empty if and only if no mixed action dot-wise dominates σ^*.

(ii) **(Sufficient conditions.)** When behavioural strategies are feasible, so that σ^* is planning-optimal over the whole simplex $\Delta(A(h))$ rather than over pure strategies alone, no dominating mixture exists, and hence one-boxer beliefs exist, whenever any of the following holds: *(a)* the game is single-agent; *(b)* the game is payoff-independent (Section 4) and the dot structure is exogenous (does not depend on the action profile); *(c)* the game is payoff-independent and the dot counts are symmetric across agents; *(d)* all agents have identical payoffs. In cases (a), (c), and (d) the thirder beliefs π are themselves one-boxer.

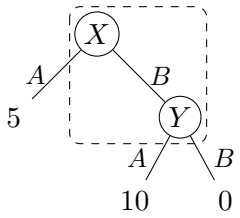
(iii) **(Representation.)** Whenever one-boxer beliefs exist, a Bayesian agent who holds thirder beliefs π at the information set but has commitment power is behaviourally equivalent to an interim expected-utility maximiser who holds one-boxer beliefs P^{OB} and has no commitment power.

Proof. See Appendix B.2. □

There are three distinct parts to this theorem: (i) a characterisation of when one-boxer beliefs exist: the only obstruction is a mixture that beats σ^* at every dot it reaches, a condition that can be checked directly; (ii) sufficient conditions under which planning optimality over behavioural strategies rules that obstruction out, so that existence is automatic; and (iii) the representation result, which observes that an agent with one-boxer beliefs acts *as-if* they were a Bayesian with commitment power. Part (iii) is the entire point of introducing these beliefs; parts (i) and (ii) delimit when the device is available. The sufficient conditions do not cover all multi-agent games: with payoff interdependence, a planning optimum can itself be dot-wise dominated, and then no beliefs, one-boxer or otherwise, rationalise it (Proposition 4.4).

Considering the representation theorem, we can observe that the result is true by definition. The committed thirder plays σ^* because they are bound to the plan; the uncommitted one-boxer plays σ^* because their beliefs P^{OB} are defined so as to make it interim-optimal. When the planning-optimal strategy is a non-degenerate mixture, the one-boxer is indifferent between the actions in its support: the belief revision from π to P^{OB} equalises the expected continuation payoffs that the thirder's beliefs would have ranked differently. Given this result, if we want to model agents with commitment power in a game, we do not need to formalise a planning stage and the ability to commit. Instead, we can use this *as-if* representation and just assume they hold one-boxer beliefs.

This representation result should not be conflated with Proposition 2.3 of Grove and Halpern (1997); it is reminiscent of this but distinct from it. Their result concerns *valuation*: under their 'outcome-uninformative/Z-consistent' beliefs, the interim expected value of a fixed strategy equals its ex ante value. Mine concerns *choice*: under a one-boxer belief, the interim agent's best response reproduces the planning-optimal strategy. The two conditions constrain different objects (the value of the incumbent strategy, versus the values of its deviations), and neither implies the other. The absent-minded driver makes the gap concrete. At the behavioural optimum, the optimal mixture is already interim-optimal under thirder beliefs (Theorem 3.1), so there is no choice tension; yet under those same thirder beliefs the interim valuation is $\frac{8}{5}$ against an ex ante value of $\frac{4}{3}$, so the Grove-Halpern valuation gap survives where my tension vanishes. Conversely, the outcome-uninformative (here, halfer) belief that closes their valuation gap leaves the pure



(a) Pure optimum truncates (Y off-path): relaxing the support rescues existence.

	a_2=0	a_2=1	a_2=2
a_1=0	10	20	5
a_1=1	20	0	30
a_1=2	5	30	0

(b) Pure optimum reaches both dots: the relaxed condition fails too.

Figure 2: Two pure-strategy failures of one-boxer existence. **(a)** A game on the information set $h = \{X, Y\}$ with payoffs at the leaves; the pure optimum is A. The dot Y at which A is locally best is off-path under pure A, so enlarging the belief support to reachable-in-principle dots restores existence. **(b)** A single-state game with both dots always reached; entries are $U(a_1, a_2)$ for the actions a_1, a_2 played at the two dots. The pure optimum $a^{\text{plan}} = 0$ is dot-wise dominated by action 1, so no belief rationalises it.

planning-optimum interim-suboptimal when the agent optimises over continuation payoffs, and so is not one-boxer. The two senses in which one might reconcile the interim agent with the plan are therefore genuinely distinct, and a single belief cannot in general satisfy both.

When existence fails. For a single agent, part (ii) makes existence automatic once behavioural strategies are allowed, so the criterion of part (i) can bite only in the pure-strategy world. And it does: in the absent-minded game of Figure 2(a), the pure planning-optimum is A (payoff 5 against 0), under which the play exits at X, so only X is reached; but the hybrid "B at X, A at Y" scores 10, so $\beta = \delta_B$ dot-wise dominates A at the only reached dot, and no one-boxer beliefs exist.[6] This failure is an artefact of the admissible support of one-boxer beliefs. Definition 2.7 lets P^{OB} weight only dots reached under σ^*, yet A is locally optimal precisely at Y (payoff 10), the dot that pure A truncates off-path. Relaxing the definition to allow weight on any dot reachable *in principle* would therefore rescue existence in Figure 2(a): a point mass at Y rationalises A.

The relaxation does not, however, buy existence in general. It merely replaces the criterion of part (i) with the same condition quantified over *all* dots of h, which can itself fail. Figure 2(b) is a single-state, single-agent game with two dots *both always reached* and three actions. Its pure planning-optimum is action 0 (the diagonal payoff 10 beats 0), yet a single-dot deviation to action 1 scores 20 at *either* dot; action 1 thus dot-wise dominates action 0 everywhere, and *no* belief, however supported, rationalises it. When the pure optimum reaches every dot, as in (b) where both are always visited, "reached" and "reachable" coincide and the two conditions are one. The support relaxation bites only when the optimum truncates a dot off-path, as in (a); in both regimes the universal resolution is randomisation (Theorem 3.1).

3 Single Agent Results

I now characterise when the Newcomb tension arises in single-information-set SLU games, when it can be resolved, and how the halfer-thirder debate relates to commitment; as per the section

[6] Randomisation restores existence: $\sigma^* = \frac{3}{4}A + \frac{1}{4}B$ is the planning optimum on $\Delta(A(h))$, and one-boxer beliefs exist for it by Theorem 2.2(ii), the game being for a single agent.

title, I begin with single agent games. Throughout, I write Θ for the set of states (initial nodes), $D(\theta, \mathbf{a})$ for the set of dots at h in state θ under action profile $\mathbf{a}$, and $U(\mathbf{a}, \theta)$ for the game payoff. When the number of dots does not depend on the action profile (as in Sleeping Beauty), I write simply $D(\theta)$.

3.1 Behavioural Strategies Resolve the Newcomb Tension

The Newcomb tension described in Section 2 arises when the agent is restricted to pure strategies. I now show that when the agent can *randomise*, adopting a behavioural strategy $\sigma \in \Delta(A(h))$, the tension vanishes entirely, regardless of the number of states.

Theorem 3.1 (Randomisation Resolution)**.** In any single-agent SLU game with a single information set h, the planning-optimal behavioural strategy σ^* is also interim-optimal: an agent waking at an unknown dot in h, taking all other temporal parts' strategy as σ^* and maximising their own continuation payoff, finds σ^* optimal. This holds for any number of states $|\Theta| \geq 1$.

Proof. See Appendix B.3. □

This result generalises the classical resolution of the absent-minded driver problem. In the standard AMD, which has a single state ($|\Theta| = 1$), the planning-optimal behavioural strategy $q^* = 2/3$ is also the interim fixed point. Theorem 3.1 shows that this coincidence does not depend on there being only one state of the world, or on the specific properties of the AMD: for *any* single-agent, single-information-set SLU game, per-dot continuation-payoff reasoning recovers the planning optimum, regardless of how many states the game has or whether payoff-irrelevant dots are present.

A notable feature of the AMD is that the Newcomb tension under pure strategies is not merely a divergence between planning and interim optima: the pure-strategy interim fixed point does not even *exist*, as I explain in Appendix A.3. It is thanks to the introduction of randomisation that the interim fixed point can exist, and as Theorem 3.1 shows, this fixed point necessarily coincides with the planning optimum. In this sense, randomisation is a stone that kills two birds: it creates the interim fixed point and resolves the Newcomb tension in a single step.

In Appendix C I work through a multi-state example, *Sleeping Beauty Behind the Wheel*, together with its duplicating extension: the coincidence of planning optimum and interim fixed point does not depend on the AMD's having no state of the world, though the expectation paradox of Grove and Halpern (1997) persists at the behavioural optimum.

4 Multiple Agent Results

I now turn to games in which multiple agents occupy dots at the same information set, assuming co-cardinal utility throughout and evaluating the planning stage with the utilitarian social welfare function of Section 2.2.

The Newcomb tension arises because the planner and the interim agent assign different social weights to the agents. Both condition on the prior ρ over states of the world and the induced distribution over action profiles $\mathbf{a}$, and both distribute probability uniformly over the dots visited

in each state-action-profile pair $(\theta, \mathbf{a})$. The difference is that the thirder (Definition 2.4) adjusts the probability of each $(\theta, \mathbf{a})$ proportionally to the number of dots $|D(\theta, \mathbf{a})|$, while the planner does not. This means the thirder assigns larger social weight to agents who appear in state-action-profile pairs with more awakenings, while the planner does not make this adjustment. Since the number of dots varies across state-action-profile pairs whenever the dot structure is asymmetric, the two sets of social weights generically disagree, and the planning-optimal and interim-optimal actions diverge. This social-weight wedge is the *asymmetry* channel; payoff interdependence produces the tension by a distinct route, since when one agent's payoff is sensitive to another's action the occupant's interim first-order condition ceases to coincide with the planner's even when the dot counts are symmetric. To separate the two channels formally, call a game *payoff-independent* (across agents) if each agent's payoff u_n depends on the action profile $\mathbf{a}$ only through the actions taken at agent n's own dots (equivalently, if a single-agent deviation at one agent's dot leaves every other agent's payoff unchanged); otherwise it is *payoff-interdependent*.

Theorem 4.1 (Multi-Agent Newcomb Tension)**.** Consider a multi-agent SLU game with co-cardinal utility and a single information set h.

(i) (Payoff independence + symmetric counts ⇒ no tension.) If the game is payoff-independent and the dot counts are symmetric across agents ($|D(\theta, n, \mathbf{a})| = |D(\theta, n', \mathbf{a})|$ for all $n, n' \in N$, $\theta \in \Theta$, and $\mathbf{a}$), then the equation

$$\frac{dV^{\text{plan}}}{d\sigma} = \frac{C(\sigma)}{|N|} \cdot \text{FOC}^{\text{int}}(\sigma)$$

holds, with FOC^{int} the interim first-order condition, and the per-dot continuation-payoff reasoning of Theorem 3.1 recovers the planning optimum. There is no Newcomb tension.

(ii) (Payoff interdependence generically ⇒ tension.) If the game is *payoff-interdependent* (some agent's payoff is sensitive to another agent's action), or the dot counts are asymmetric across agents, then the above equation generically fails to hold, and the planning-optimal and interim-optimal strategies diverge.

Proof. See Appendix B.4. □

Part (i) shows that mere payoff *disagreement* between agents does not by itself create a Newcomb tension: provided the game is payoff-independent (no agent's payoff is sensitive to another's action) and every agent visits the same number of dots in every state-action-profile pair, the proportional planning–interim equation holds even when $u_{n_1} \neq u_{n_2}$. Part (ii) shows that the tension generically arises as soon as either condition fails: when agents' payoffs are *interdependent*, or when the dot structure is asymmetric across agents. The duplicating absent-minded driver (Proposition 4.3), which is both payoff-interdependent and count-asymmetric, exhibits such a tension for all but a knife-edge set of its payoff parameters.

4.1 Duplicating Sleeping Beauty

Having established this general result on Newcomb tensions in multi-agent games, I now discuss some important special cases. Firstly, this allows me to comment on the duplicating Sleeping Beauty problem, as I began to do in the Introduction:

Proposition 4.2 (Newcomb Tension in Duplicating Sleeping Beauty)**.** In Duplicating Sleeping Beauty, let an awakened agent report a credence $x \in [0, 1]$ that the coin landed Tails, rewarded by a strictly proper scoring rule.[7] Then the planning-optimal report is the halfer credence, $x^{\text{plan}} = \frac{1}{2}$, while the interim-optimal report is the thirder credence, $x^{\text{int}} = \frac{2}{3}$: the game exhibits a Newcomb tension, and randomisation over reports does not resolve it.

Proof. See Appendix B.5. □

Kierland and Monton (2005) discuss the canonical and duplicating Sleeping Beauty Problems where Beauty is attempting to maximise a Brier score (quadratic loss), and argue that halfer beliefs are justified in the latter problem since they minimise the expected average error. Conversely, in the standard problem one should be a thirder, they believe, since the solution $\frac{1}{3}$ minimises total expected error. In their view, the two positions simply amount to pursuing different epistemic goals, where the one is more reasonable in the canonical problem, and the other in the duplicating variant. The stance I argue here is of course different, I take for granted that the thirder position is the epistemically correct answer in both versions, but approach the question of what bet one should choose as a social choice problem.[8] With this social choice approach, the planning stage in fact recovers their halfer recommendation: the compound-state planner is precisely an expected-average-error minimiser, and so reports $\frac{1}{2}$ in the duplicating problem (Proposition 4.2), while the awakened agent, correctly holding thirder credences, reports $\frac{2}{3}$. My disagreement with the halfers is thus confined to credence: their bet is vindicated at the planning stage, but not at the interim stage. Janda (2024) argues instead simply that we can justify *different* answers to the canonical and duplicating problems because they should be modelled as different games, one with absent-mindedness and the other forgetfulness. Epistemically, I disagree, but agree that agents trying to maximise expected utility should make different bets in each case, and of course am arguing that whereas the one problem exhibits a Newcomb tension, the other does not. I discuss Janda's paper at greater length in *Sleeping Beauty's Dismal Day Out*, where I discuss how one should model games with self-locating uncertainty in general.

4.2 Duplicating AMD

The absent-minded driver normally has no state of the world: the game tree is deterministic, with dots X and Y at a single information set. In the *duplicating* version, a clone (Player 2) is created when Player 1 continues at X: the red dot on the Continue branch marks the clone's creation, and Player 2 occupies Y. I adopt the convention that red dots may only precede action nodes (never terminal nodes), and that a terminal node carries a payoff for the clone only if the clone has been created somewhere on its branch. Under this convention, Exit at X occurs before the clone exists, so its payoff is a scalar for Player 1 alone (see Figure 3). Since Exit at X precedes

[7] A scoring rule is *strictly proper* if the expected score is uniquely maximised when the report equals the agent's true belief. E.g. quadratic loss.

[8] To my knowledge, this is novel in the literature on Sleeping Beauty.

the clone's creation, $D(n{=}1) = \{X\}$ and $D(n{=}2) = \{Y\}$: Player 1 occupies only X, while the clone occupies only Y. Let us slightly abuse notation, using σ to denote both the behavioural strategy and the Continue-probability it prescribes; the normalising constant is then:

$$C(\sigma) = (1-\sigma)\cdot 1 + \sigma \cdot 2 = 1 + \sigma.$$

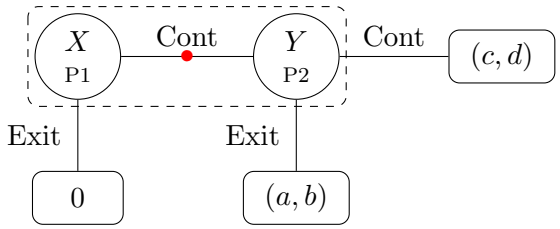


Figure 3: The duplicating AMD. A red dot on the Continue branch marks the creation of the clone (Player 2); since red dots may only precede action nodes, Exit at X occurs before the clone exists and carries a scalar payoff. Exit at Y gives (a, b) and Continue past Y gives (c, d).

General clone payoffs. Suppose exiting at Y gives (a, b) and continuing past Y gives (c, d) with $a, b, c, d > 0$, while exiting at X gives a scalar 0. Since the dot structure depends on the action profile, the full compound-state formula is required. The three action profiles and their contributions to V^{plan} are:

Profile	Prob	Dots	Weights (w_1, w_2)	Payoffs (U_1, U_2)
Exit at X	$1-\sigma$	1 (P1 only)	(1, 0)	(0, −)
Cont–Exit	$\sigma(1-\sigma)$	2	(1/2, 1/2)	(a, b)
Cont–Cont	σ^2	2	(1/2, 1/2)	(c, d)

The planning payoff is therefore

$$V^{\text{plan}}(\sigma) = \sigma(1-\sigma)\,\frac{a+b}{2} + \sigma^2\,\frac{c+d}{2} = \frac{(a{+}b)\,\sigma - (a{+}b{-}c{-}d)\,\sigma^2}{2},$$

which is quadratic in σ, with

$$\sigma^* = \frac{a+b}{2(a+b-c-d)}.$$

The thirder beliefs are $P(n{=}1, X) = \frac{1}{1+\sigma}$ and $P(n{=}2, Y) = \frac{\sigma}{1+\sigma}$, and the continuation-payoff derivatives are $g_X = a-(a-c)\sigma$ (Player 1's gain from continuing at X) and $g_Y = d-b$ (Player 2's gain from continuing at Y). The interim first-order condition gives

$$C(\sigma)\cdot \text{FOC}^{\text{int}}(\sigma) = a - (a-c+b-d)\sigma,$$

so $\hat{\sigma} = a/(a+b-c-d)$. The no-tension condition $\sigma^* = \hat{\sigma}$ requires $(a+b)/2 = a$, i.e. $b = a$: the tension vanishes whenever the two agents' payoffs coincide at the Exit terminal (Y), even if they differ at the Continue terminal. In particular, with fully identical payoffs (say $(a, b) = (4, 4)$ and $(c, d) = (1, 1)$) the per-compound-state social welfare is the common payoff regardless of the weights, the planner's problem reduces to a single-agent optimisation, and $\sigma^* = \hat{\sigma} = 2/3$, exactly the single-agent AMD numbers: duplication by itself does not create a Newcomb tension. For $b \neq a$:

Proposition 4.3 (Persistent Tension)**.** Suppose 'Exit at X' gives a scalar payoff of 0 for Player 1, 'Exit at Y' gives (a, b), and 'Continue past Y' gives (c, d) with $a, b, c, d > 0$, $a + b > 2(c + d)$, and $b > c + d$, so that both the planning optimum and the interim fixed point are interior. The Newcomb tension is present if and only if $b \neq a$, with

$$\sigma^* - \hat{\sigma} = \frac{b - a}{2(a + b - c - d)}.$$

The planner favours more continuation than the interim agent when $b > a$ (the clone benefits more from reaching Y than the original) and less when $b < a$.

Proof. The planning payoff $V^{\text{plan}}(\sigma) = [(a+b)\sigma - (a+b-c-d)\sigma^2]/2$ is strictly concave with critical point $\sigma^* = (a + b)/[2(a + b - c - d)]$, which lies in $(0, 1)$ precisely when $a + b > 2(c + d)$. The interim first-order condition $C(\sigma) \cdot \text{FOC}^{\text{int}}(\sigma) = a - (a - c + b - d)\sigma$ is strictly decreasing in σ with root $\hat{\sigma} = a/(a + b - c - d)$, which lies in $(0, 1)$ precisely when $b > c + d$. Subtracting: $\sigma^* - \hat{\sigma} = (a + b - 2a)/[2(a + b - c - d)] = (b - a)/[2(a + b - c - d)]$. □

Outside the interior region, the optima clip to the boundary and the tension is read off the clipped values. When $b \leq c + d$ the interim agent continues for sure ($\hat{\sigma} = 1$), so a tension persists whenever $\sigma^* < 1$ (the configuration exploited in Proposition 4.4 below), while when both optima clip to the same corner the tension vanishes even for $b \neq a$.

The tension arises from the interaction between the dot asymmetry and the payoff asymmetry. Under the compound-state formula, 'Exit at X' contributes only Player 1's payoff (with full weight), while the Continue profiles split weights equally between the players. When $b > a$, the planner's equal weighting of the clone's higher payoff at Y makes continuation more attractive than the interim agent perceives; when $b < a$, the reverse holds. At $b = a$ the tension vanishes, even if $c \neq d$: the payoff asymmetry at the Continue terminal does not matter because the relevant wedge is driven by the Exit-at-Y payoffs, where the dot asymmetry bites. When additionally $c = d$, the game reduces to the identical-payoff case.

The duplicating absent-minded driver also shows that the one-boxer device of Section 2.4 has real limits in multi-agent games. Theorem 2.2(ii) guarantees the existence of one-boxer beliefs for behavioural planning optima in single-agent, payoff-independent, and identical-payoff games; payoff-interdependent games are excluded, and necessarily so:

Proposition 4.4 (No As-If Representation under Interdependence)**.** In the duplicating absent-minded driver of Figure 3, suppose $a, b, c, d > 0$ with $a + b > 2(c + d)$ and $d > b$. Then the planning-optimal behavioural strategy $\sigma^* = (a + b)/[2(a + b - c - d)]$ is interior, yet no one-boxer beliefs for σ^* exist: the pure action Continue dot-wise dominates σ^*, so no belief over awakenings, however chosen, makes σ^* interim-optimal. The value of commitment is strictly positive, and cannot be replicated by any belief revision.

Proof. See Appendix B.6. □

In the single-agent world there are two responses to the Newcomb tension: randomisation removes it outright (Theorem 3.1), while under pure strategies, where the tension itself persists, commitment can at least be emulated by one-boxer beliefs whenever the existence criterion of Theorem 2.2(i) holds. Under payoff interdependence both responses can fail at once: the tension survives randomisation, and there is no system of beliefs whose adoption reproduces the planning optimum. Commitment power is then irreplaceable: it must be modelled directly, not as a belief.

5 Literature Review

This paper touches on a few different areas of existing literature. First of all, it adopts the notation of Kreps and Wilson (1982). However, the companion article *Sleeping Beauty's Dismal Day Out* does more to develop Kreps and Wilson's article than this one, since in that article I consider how one should model games with self-locating uncertainty in general, and to what extent one can extend their sequential equilibrium existence result to them. In this paper, beyond adopting their notation, I do not consider similar questions or extend their results.

Beyond this, however, this article *does* contribute to the literature on the Sleeping Beauty problem, and specifically the section of that literature that discusses whether the credences of an awakened agent should be the same in the canonical and duplicating problems. I have already discussed this literature (specifically Kierland and Monton (2005); Janda (2024)) in Section 4.1, but note here that in my consideration of the novel game of the *duplicating absent-minded driver*, I extend the debate to general games with inter-personal self-locating uncertainty, and take a social choice approach to the problem. This literature has not discussed the formation of self-locating beliefs in games with interpersonal self-locating uncertainty with endogenous awakenings. To the best of my knowledge, each of these points is original to this paper.

Concerning the absent-minded driver, summarised in Appendix A.3, the main collection of articles within economic theory is that of the original special issue of *Games and Economic Behavior* on Imperfect Recall (Volume 20, Issue 1, July 1997). As suggested by Aumann et al. (1997), I take the *modified multi-selves approach* of Piccione and Rubinstein (1997) (one of several time-consistency notions catalogued by Halpern 1997) to be the straightforwardly correct approach in such games. An agent always considers changing the action they control in a given awakening assuming the candidate strategy will be followed in all other awakenings. My Theorem 3.1 builds on their analysis directly, and generalises the fact, which they of course observe themselves, that with behavioural strategies the 'paradox of the absent-minded driver' (what I call a Newcomb tension) disappears. I show that in single-information set games with only one agent, introducing multiple states does not defeat this resolution. In the multi-agent *duplicating absent-minded driver*, this then changes, and the paradox can defeat even behavioural strategies whenever there is payoff divergence. I have already discussed the relationship between this paper and Grove and Halpern (1997), along with related work such as Ross (2010), in Sections 2.3 and 2.4.

The planning-interim divergence at the heart of this paper is, at its most abstract, the founding problem of the dynamic-choice literature: Strotz (1955)'s planner, at a later date and with no new information, wishes to revise her own earlier plan, and his two responses, precommitment and consistent planning, are exactly those available to my agents, formalised respectively by the planning stage of Section 2.2 and the interim fixed point. The classical route to such dynamic inconsistency runs through non-expected-utility preferences (Hammond, 1988; Machina, 1989); the agents of this paper are expected-utility maximisers, and what generates the fork is instead the conditioning itself, on centred (self-locating) information that the ex ante stage cannot express. The closest decision-theoretic apparatus is that of Vergopoulos (2011), which distinguishes the state space Ω on which acts are defined from a finer space $\tilde{\Omega}$ on which events are defined: conditioning on $\tilde{\Omega}$ delivers dynamic consistency for free, the classical implication from dynamic consistency to the Sure-Thing Principle holds only when $\tilde{\Omega}$ is a relabelled copy of Ω, and when the two genuinely differ (his applications being ambiguity and causal structure, with his 'lazy student' a Newcomb problem in miniature) an agent can be dynamically consistent on the fine space while violating the STP on the coarse one. In the present paper, the planning stage evaluates uncentred acts over the

states θ while the awakened agent conditions on the centred information of which dot she occupies: the Newcomb tension is exactly an Ω-level violation coexisting with fine-space consistency, and Newcomb's game of Section 2.2 is a direct instance of his causal construction.

Beyond the economic literature, two articles in the philosophical literature bear directly on the absent-minded driver. Spohn (2025) argues, as per its title, that "*The Puzzle of the Absent-Minded Driver is Not About Absent-Mindedness, but about Indexical Belief*", and I agree: my own interest in these games is as infrastructure for learning under temporal ignorance, and his observation that his proportionality ('R-') principle, effectively Ross's Generalised Thirder Principle stated for all indexical propositions, applies equally to mere ignorance of one's temporal location points the same way. But the contrast between the canonical and duplicating drivers shows that different forms of indexical uncertainty are importantly different: whether agents behave *as-if* they form R-principle beliefs depends on whether they have commitment power.

Schwarz (2015) is the only other article of which I am aware that connects all three of the absent-minded driver, Sleeping Beauty, and Newcomb's problem. He crosses the halfer/thirder distinction with the causal/evidential one, finding that randomisation (the 'useless coins' of his title) helps the driver under some of the four combinations but not others; in my setting randomisation always resolves the single-agent tension (Theorem 3.1). I also sidestep the EDT/CDT axis that drives half his analysis, assuming causal decision theory throughout as is standard in game theory; and his treatment is single-agent, so the duplicating AMD, in which the tension survives behavioural strategies whenever payoffs diverge across agents (Proposition 4.3), lies outside the scope of his framework.

6 Conclusion

This paper has studied the Newcomb tension in games with self-locating uncertainty, restricting attention to games with a single information set. I have drawn an analogy between these games and Newcomb's problem: in both settings, an agent at the point of decision faces a conflict between locally rational reasoning and globally optimal planning for on-path decision nodes. In SLU games, the locally rational agent uses thirder beliefs (the epistemically correct self-locating credences in my view and the view of most epistemologists) while the planner evaluates strategies ex ante. The one-boxer belief representation theorem (Theorem 2.2) establishes when we can simply assume agents act with one-boxer beliefs in place of assuming commitment power: wherever such beliefs exist (and they always do for a single agent choosing behavioural strategies), a committed Bayesian who holds thirder beliefs and one who holds one-boxer beliefs are behaviourally equivalent. Under payoff interdependence this can fail (Proposition 4.4), and commitment must then be modelled directly. The one-boxer beliefs that restore planning optimality may coincide with halfer beliefs, but are instrumentally motivated; they are the beliefs an agent would need to hold in order for interim reasoning to reproduce the planning-optimal strategy without commitment power.

In the single-agent case, this divergence is always resolved by randomisation. Behavioural strategies kill two birds with one stone: they create the interim fixed point, which may not exist under pure strategies (as in the absent-minded driver), and they ensure that this fixed point coincides with the planning optimum (Theorem 3.1). This holds for any game corresponding to Definition 2.1 with a single agent, regardless of the number of states of the world. The multi-agent case is different, as the Newcomb tension in Duplicating Sleeping Beauty (Proposition 4.2) illustrates.

In multi-agent games, I have taken a social choice approach, aggregating the payoffs of different agents with a utilitarian social welfare function. This framework treats each state-action-profile pair $(\theta, \mathbf{a})$ as a compound state, distributes probability uniformly across the 'dots' or awakenings within each, and evaluates welfare accordingly. To my knowledge, this social choice approach is novel in the literature on Sleeping Beauty. With multiple agents, the Newcomb tension is generically present whenever agents' payoffs are interdependent or their awakening structure is asymmetric (Theorem 4.1). Awakening asymmetry drives a wedge between the planner's halfer-derived social weights and the interim agent's thirder-derived weights; payoff interdependence drives the tension through a separate channel, one agent's payoff depending on another's action.

A Background

A.1 Sleeping Beauty

> *Sleeping Beauty is in a windowless, clockless room on Sunday ($t = 0$) with a group of scientists, who explain to her that she will undergo the following experiment. The scientists shall toss a fair coin out of sight; if it lands heads, they shall put Beauty to sleep and awaken her on Monday ($t = 1$), before putting her to sleep again shortly afterwards for the rest of the experiment. If the coin lands tails, however, they shall also awaken her a second time on Tuesday ($t = 2$), before putting her to sleep for the rest of the experiment. Each time they put her to sleep they do so with a memory-wiping drug such that each awakening is indistinguishable to her. Upon every awakening they ask Beauty with what probability she believes that the coin landed heads. What should she say?* (Elga, 2000)

The two main schools of thought are *halfer* and *thirder*. Halfers, e.g. Lewis (2001), argue that upon awakening Sleeping Beauty has learnt no new information (she was certain to wake up whatever the state of the world) and so her credence in Heads should remain 1/2. Elga and his fellow thirders, however, argue that since each awakening is indistinguishable, she should assign equal probability to each of the three possible awakenings: {(Monday, Heads), (Monday, Tails), (Tuesday,Tails)}. She began with a uniform prior over states, and so assigns prior unconditional probability of $\frac{1}{2}$, $\frac{1}{2}$, and $\frac{1}{2}$ to each awakening occurring. Hence, since each awakening is equally probable, and there are three of them, she should simply normalise these probabilities to form a posterior: $\mathbb{P}(\theta = H) = 1/3$. Halfer and thirder positions are alternatively known as *uncentred* and *centred*, where this 'centre' terminology reflects that the halfer position produces beliefs that do not vary according to an agent's beliefs about their own self-location.

By referring to them as simply 'halfer' and 'thirder' positions, one risks giving the impression of two monolithic tribes which brook no internal dissent. Conversely, a dizzying array of varying arguments (often varying very subtly) have been offered in defence of each position, and the literature is enormous. Winkler (2017) offers a great review for those interested in investigating it. I myself am a thirder, in both this and the duplicating variant of the problem I discuss in the introduction (where in the event that the coin lands tails, we clone Sleeping Beauty and wake the clone up instead), and this is the stance on which the literature seems to have broadly settled. A final point worth noting is that the literature on the Sleeping Beauty problem was born out of that on the absent-minded driver discussed in Appendix A.3, to isolate the epistemological paradox it contains, since there are multiple paradoxical strands to the absent-minded driver paradox.

A.2 Newcomb's Problem and Newcomb's Game

> *You walk into a room. Upon doing so, you are greeted by a friendly, super-intelligent robot, whose honesty you consider beyond dispute. He explains to you that you have a choice, and gestures towards two boxes on the desk. One of them is transparent, and you can see it contains \$1000. The other contains either a gazillion trillion dollars, or nothing at all. Your choice is between either taking only the mystery box (one-boxing), or both of them (two-boxing). The final crucial detail is that the mystery box contains the fortune if the robot predicted you would one-box, and nothing if he predicted you would two-box. He has an accuracy of near 100%. What do you do?* (Nozick, 1969)

Nozick (1969) introduced this problem and attributed it to Newcomb, famously noting that: "*To almost everyone, it is perfectly clear and obvious what should be done. The difficulty is that these people seem to divide almost evenly on the problem, with large numbers thinking that the opposing half is just being silly.*" One could say a similar thing of the debate over the Sleeping Beauty problem, though Nozick himself alas did not live to witness it.

On the one hand, one could argue that the correct solution is clearly to two-box. The contents of the box are already fixed when one enters the room, so it is a strictly dominant strategy to take both boxes whatever the mystery box contains: only a muppet of the highest order would do otherwise. This position follows from *causal decision theory*, which is implicitly assumed in game theory, for example. On the other, *evidential decision theory* asks: if you were to learn of your own decision after the fact, what would make you happier? To learn you had one-boxed or two-boxed? Since the robot is so accurate, whatever you did he likely predicted it. Hence, if you one-boxed you are likely rich and should be thrilled, but if you two-boxed you have covered this month's rent. Whilst this is not nothing, it is a real shame given the alternative. If your reaction, like mine was, is that one should two-box, you can nonetheless observe that the 'foolish' one-boxers will fare much better. If before walking into the room you learnt of the scheme ahead of time, and could somehow commit yourself to one-boxing, it would clearly be a good idea to do so.

This is the phenomenon I call a *Newcomb tension*: an agent who would commit to one action, if given the chance, finds that their locally rational reasoning at the point of decision leads them to a different action. The value of commitment is positive, and the interim-rational agent cannot capture it.

Before exploiting Newcomb's problem as a motivating example, however, we can observe that it is actually a poorly defined decision problem. Given this, whatever one's initial reaction as to the 'correct' solution, neither is truly correct: the assumption that the robot manages near-perfect projection is incompatible with the agent having free will (an essential assumption in decision theory: asking what one *should* do presupposes the capacity to actually choose amongst distinct options). To see this, suppose the agent considers playing a uniform mixed strategy (one-box with probability $1/2$) by flipping a fair coin in the room. If the predictor truly has no causal influence over the outcome of this coin flip and the agent has genuine free will, then the predictor can be right with probability at most $1/2$, not $1 - \varepsilon$. The premise that the predictor has accuracy near 100% uniformly over the agent's available strategies therefore implicitly assumes that the agent cannot, in fact, freely randomise: either the predictor has access to the realisation of the coin flip (in which case it is causally downstream of the prediction, and we are no longer in Newcomb's problem), or the agent's deliberation is itself determined in a way that the predictor can read off (in which case the language of free choice is misleading).

Newcomb's game. I therefore work with the following variant as a motivating example, which I call *Newcomb's game*. There is one agent, one binary action $A(h) = \{\text{One-box, Two-box}\}$ at the interim information set, and a planning stage $A(h_0) = \{\text{Commit, Not Commit}\}$ as in Section 2.2. If the agent chose Commit at h_0, then the predictor sets the contents of the opaque box *equal to whatever action the agent committed to*: $\theta = \text{Full}$ if the agent committed to One-box, $\theta = \text{Empty}$ if the agent committed to Two-box. If the agent chose Not Commit at h_0, then the predictor sets $\theta = \text{Empty}$ (i.e. predicts Two-box). The agent then plays at h, and the realised payoff is $U(\text{One-box}, \text{Full}) = M$, $U(\text{Two-box}, \text{Full}) = M + 1$, $U(\text{One-box}, \text{Empty}) = 0$, $U(\text{Two-box}, \text{Empty}) = 1$, with M large.

This is admittedly not as fun as Nozick's original story, since it is the very fact that the original problem is poorly defined that allows each camp to consider their stance reasonable, but it has the merit of being a well-defined game featuring a Newcomb tension. The predictor's behaviour is given by a deterministic function of the planning action, not by a mysterious near-perfect oracle that the agent's mixed strategies would defeat. Commit-to-One-box yields the planning payoff M; Commit-to-Two-box and Not Commit both yield 1; so for $M > 1$ the planning optimum is Commit-to-One-box, with payoff M. At the interim stage, however, Two-box dominates One-box point-wise, so the interim-optimal action is Two-box, with payoff 1 under Not Commit. The Newcomb tension is therefore present in Newcomb's game: $a^{\text{plan}} \neq a^{\text{int}}$, with strictly positive value of commitment $M - 1$.

A.3 The Absent-Minded Driver

The absent-minded driver (AMD), introduced by Piccione and Rubinstein (1997), is a single-player game with imperfect recall (specifically, *absent-mindedness* and not *forgetfulness*) in which the player cannot distinguish two decision nodes that lie on the same path of play. The game tree has the structure of Figure 3, but with a single agent occupying both intersections.

A driver, having foolishly decided to down several pints of lager before entering his vehicle, starts at intersection X and must choose between Exit and Continue. If he exits at X, the game ends with payoff 0. If he continues, he arrives at intersection Y, where he faces the same choice. Crucially, however, he cannot distinguish Y from X: in his drunken stupor they seem identical, and he both can never remember passing an intersection, and knows this fact upon arriving at any intersection (he is a logically omniscient drunk...). Formally, the intersections X and Y form a single information set h. If the driver exits at Y, the payoff is 4; if he continues past Y, the payoff is 1.

Since the driver cannot distinguish X from Y, any (pure) strategy must prescribe the same action at both intersections. Exit yields a game payoff of $U(\text{Exit}) = 0$ (the driver exits immediately at X), while Continue yields $U(\text{Continue}) = 1$ (the driver continues at both intersections). The planning-optimal pure action is therefore Continue.

However, the continuation payoffs are 'dot-dependent' (in the main body of the text I refer to different 'awakenings' or nodes within one branch of a play of the game as 'dots'): $v_X(\text{Exit}) = 0$, $v_Y(\text{Exit}) = 4$, $v_X(\text{Continue}) = 1$, $v_Y(\text{Continue}) = 1$. Under halfer beliefs $P(X) = P(Y) = 1/2$, the interim expected payoff from Exit is $\frac{1}{2} \cdot 0 + \frac{1}{2} \cdot 4 = 2$, exceeding the expected payoff from Continue of 1. The fact that each pure strategy implies a belief such that the opposite strategy is optimal is the paradox of the absent-minded driver. Piccione and Rubinstein (1997) discuss different approaches to reasoning in this model but note that under the *modified multi-selves approach*, adopting a behavioural strategy of continuing with probability 2/3 resolves the problem.

In addition to providing an interesting early example of what I call a *Newcomb tension*, this game also serves as an example of why sequential equilibrium may not exist in games of imperfect recall (Kreps and Wilson, 1982). In standard games, an equilibrium is defined in mixed strategies, which randomise which pure strategy is played at each information set, but do not independently randomise each time one arrives at that information set in the same play of the game. As per the famous Kuhn's Theorem (Kuhn, 1953), in games of perfect recall this distinction is unimportant.

B Omitted Proofs

B.1 Proof of Proposition 2.1

Proof. **Part (i).** When continuation payoffs are dot-independent, $v_d(a', \mathbf{a}_{-d}, \theta) = v(a', \theta)$ for all d and $\mathbf{a}_{-d}$, so the interim objective reduces to

$$V^{\mathrm{OB}}(a') \;=\; \sum_{\theta,\,\mathbf{a},\,d} P^{\mathrm{OB}}(\theta, \mathbf{a}, d)\, v(a', \theta) \;=\; \sum_{\theta} P^{\mathrm{OB}}(\theta)\, v(a', \theta),$$

since the beliefs over action profiles and dots sum out. Taking $P^{\mathrm{OB}}(\theta) = \rho(\theta)$ makes the interim and planning objectives identical. For a non-degenerate σ^*, dot-independence also makes the planning payoff linear in the behavioural strategy, $V^{\mathrm{plan}}(\sigma) = \sum_\theta \rho(\theta)\, \mathbb{E}_{a\sim\sigma}[v(a, \theta)]$, so every action in the support of a planning-optimal mixture maximises $\sum_\theta \rho(\theta)\, v(a, \theta) = V^{\mathrm{OB}}(a)$, and conditions (a) and (b) of Definition 2.7 hold as before.

Part (ii). In the absent-minded driver, there is a single state of the world. Under the pure strategy Continue, the action profile is $\mathbf{a} = (C, C)$ and $D(\theta, \mathbf{a}) = \{X, Y\}$. The continuation payoffs are $v_X(\text{Exit}, \mathbf{a}_{-X}, \theta) = 0$, $v_Y(\text{Exit}, \mathbf{a}_{-Y}, \theta) = 4$, $v_X(\text{Continue}, \mathbf{a}_{-X}, \theta) = 1$, $v_Y(\text{Continue}, \mathbf{a}_{-Y}, \theta) = 1$. The planning-optimal pure action is Continue ($U(\text{Continue}) = 1 > 0 = U(\text{Exit})$). Under halfer beliefs $P(X) = P(Y) = \frac{1}{2}$:

$$V^{\mathrm{OB}}(\text{Exit}) = \tfrac{1}{2}\cdot 0 + \tfrac{1}{2}\cdot 4 = 2 > 1 = \tfrac{1}{2}\cdot 1 + \tfrac{1}{2}\cdot 1 = V^{\mathrm{OB}}(\text{Continue}),$$

so the halfer chooses Exit, which is not planning-optimal. For one-boxer beliefs, I require $P(X)\cdot 0 + (1 - P(X))\cdot 4 \le P(X)\cdot 1 + (1 - P(X))\cdot 1$, yielding $P(X) \ge \frac{3}{4}$. □

B.2 Proof of Theorem 2.2

Proof. **Part (i): existence iff.** A belief P^{OB} on the triples reached under σ^* is one-boxer if and only if

$$V^{\mathrm{OB}}(\sigma^*) - V^{\mathrm{OB}}(a) \;=\; \sum_{\theta,\mathbf{a},d} P^{\mathrm{OB}}(\theta, \mathbf{a}, d)\,\big[v_d(\sigma^*, \mathbf{a}_{-d}, \theta) - v_d(a, \mathbf{a}_{-d}, \theta)\big] \;\ge\; 0 \qquad \forall\, a \in A(h), \tag{B.1}$$

where $v_d(\cdot, \mathbf{a}_{-d}, \theta)$ is the continuation payoff to the occupant $\iota(d)$ at dot d: since V^{OB} is linear in the mixing weights, (B.1) for all a gives condition (a) of Definition 2.7, and when σ^* is non-degenerate, condition (b) is automatic, because $V^{\mathrm{OB}}(\sigma^*) = \sum_{a\in\mathrm{supp}(\sigma^*)} \sigma^*(a)\, V^{\mathrm{OB}}(a)$ is a convex combination of values each at most $V^{\mathrm{OB}}(\sigma^*)$, forcing equality across the support (a convex combination that attains the common upper bound of its terms must place all its weight on terms attaining it). Existence is therefore a finite linear feasibility problem. By the theorem of the alternative (Motzkin's

transposition theorem[9]), (B.1) is feasible if and only if there is no probability distribution η over $A(h)$ whose mixed action $\beta = \sum_a \eta(a)\,\delta_a$ satisfies $v_d(\beta, \mathbf{a}_{-d}, \theta) > v_d(\sigma^*, \mathbf{a}_{-d}, \theta)$ at every reached triple; that is, if and only if no mixed action dot-wise dominates σ^*. When σ^* is a pure action a^{plan}, the reached triples are exactly $\{(\theta, d) : d \in D(\theta, a^{\text{plan}}),\ \rho(\theta) > 0\}$ with the action profile equal to a^{plan} everywhere, so the profile index may be suppressed. The dual certificate η is a multiplier in the linear programme, not a strategy the agent plays; it may be a single pure action or a non-trivial mixture, and its availability as an *analytical* object does not presuppose that the agent can randomise.

Part (ii): sufficient conditions. These sufficient conditions are inspired by Theorems 3.1 and 4.1, so the reader may find this part of the proof easier to follow if they return to it after reading those results and their proofs; it is in any case understandable as it stands. Let σ^* be planning-optimal over $\Delta(A(h))$ and suppose, for contradiction, that some β dot-wise dominates σ^*. In each case I show that the perturbation $\sigma_\varepsilon = (1-\varepsilon)\sigma^* + \varepsilon\,\beta$ strictly increases the planning payoff at $\varepsilon = 0$, contradicting optimality. Note first that pointwise dominance at every reached triple implies dominance on average over the other dots' actions: $v_d(\beta, \sigma^*_{-d}, \theta) > v_d(\sigma^*, \sigma^*_{-d}, \theta)$ at every dot reached with positive probability under σ^*.

(a) Single agent ($|N| = 1$). The forward reference to Theorem 3.1 here is harmless: its proof is self-contained and makes no use of the present theorem. By the first-order calculation of Step 5 of the proof of Theorem 3.1, extended linearly from pure directions $\delta_{a'}$ to the mixture β,

$$\frac{d}{d\varepsilon}\, V^{\text{plan}}(\sigma_\varepsilon)\bigg|_{\varepsilon=0} \;=\; \sum_{\theta} \sum_{d} \rho(\theta)\,\sigma_d(\theta, \sigma^*)\big[v_d(\beta, \sigma^*_{-d}, \theta) - v_d(\sigma^*, \sigma^*_{-d}, \theta)\big] \;>\; 0,$$

every summand with $\sigma_d(\theta, \sigma^*) > 0$ being strictly positive, and at least one such dot existing since $C(\sigma^*) > 0$. Moreover, Theorem 3.1 shows directly that σ^* is interim-optimal under the thirder beliefs π of Definition 2.4, so in this case π itself is one-boxer.

(b) Payoff independence with exogenous dot structure. Suppose each agent's payoff depends on the action profile only through the actions at that agent's own dots, and the dot structure does not depend on the action profile, so that

$$V^{\text{plan}}(\sigma) \;=\; \sum_{\theta} \rho(\theta) \sum_{n} w_n(\theta)\; \mathbb{E}_{\mathbf{a}\sim\sigma}\big[U_n(\mathbf{a}, \theta)\big]$$

with weights $w_n(\theta) = |D(\theta, n)|/|D(\theta)|$ fixed by the state. Under payoff independence a substitution at a dot not occupied by n leaves U_n unchanged, so the first-order expansion of $\mathbb{E}[U_n]$ collects only agent n's own dots:

$$\frac{d}{d\varepsilon}\, \mathbb{E}_{\mathbf{a}\sim\sigma_\varepsilon}\big[U_n(\mathbf{a}, \theta)\big]\bigg|_{\varepsilon=0} \;=\; \sum_{d \in D(\theta, n)} \big[v_d(\beta, \sigma^*_{-d}, \theta) - v_d(\sigma^*, \sigma^*_{-d}, \theta)\big],$$

strictly positive whenever $D(\theta, n) \neq \varnothing$ (every dot is reached, the structure being exogenous). Weighting by $\rho(\theta)\, w_n(\theta) \geq 0$ and summing over θ and n gives $dV^{\text{plan}}/d\varepsilon|_{\varepsilon=0} > 0$.

[9] Motzkin's transposition theorem extends the Farkas lemma to systems mixing strict and weak inequalities: exactly one of (I) $Ax > 0,\ Bx \geq 0$ and (II) $A^\top y + B^\top z = 0,\ y \geq 0,\ z \geq 0,\ y \neq 0$ is solvable. The mixture is what is needed here: the belief system (B.1) is weak, while the dominance certificate must be strict at every reached triple, since a weakly dominating mixture would not contradict (B.1). The Farkas lemma covers weak–weak systems, and Gordan's theorem all-strict ones; Motzkin's statement subsumes both.

(c) Payoff independence with symmetric dot counts. Suppose the game is payoff-independent and the dot counts are symmetric, $|D(\theta, n, \mathbf{a})| = |D(\theta, n', \mathbf{a})|$ for all $n, n' \in N$, θ, and $\mathbf{a}$. By part (i) of Theorem 4.1, $dV^{\text{plan}}/d\sigma = (C(\sigma)/|N|)\,\text{FOC}^{\text{int}}(\sigma)$, so the planning derivative is non-positive in every direction exactly when the occupant interim first-order condition is. The planning optimum σ^* is therefore interim-optimal under thirder beliefs, $P^{\text{OB}} = \pi$ is one-boxer, and a dominating β would make the occupant FOC, and hence the planning derivative, strictly positive in the direction β, the desired contradiction.

(d) Identical payoffs. With $u_n = u$ for all n, the within-compound-state weights sum to one and $V^{\text{plan}}(\sigma) = \sum_\theta \rho(\theta)\,\mathbb{E}_{\mathbf{a}\sim\sigma}[U(\mathbf{a}, \theta)]$ is the single-agent planning objective for the common payoff, while every occupant's continuation payoff is the common continuation payoff. The argument of case (a) applies verbatim, and π is again one-boxer.

Part (iii): representation. Given existence of a one-boxer belief P^{OB}, the equivalence is immediate from the definition (Definition 2.7): one-boxer beliefs are, by construction, the beliefs under which interim optimisation reproduces the planning-optimal strategy. The committed Bayesian plays σ^* because it is bound to the plan; the uncommitted agent holding P^{OB} and maximising V^{OB} finds σ^* optimal by conditions (a) and (b) of Definition 2.7. The two are therefore behaviourally indistinguishable. □

B.3 Proof of Theorem 3.1

Proof. The argument proceeds in five steps.

Step 1: All interim agents form the same belief. Since the game has a single information set h, every temporal part conditions on the same event (being at h) and hence forms the same interim belief over state-dot pairs. Under behavioural strategy σ, this belief is $P(\theta, d) = \rho(\theta)\,\sigma_d(\theta, \sigma)/C(\sigma)$, where $\sigma_d(\theta, \sigma)$ is the reach probability of dot d in state θ and $C(\sigma) = \sum_{\theta', d'} \rho(\theta')\,\sigma_{d'}(\theta', \sigma) > 0$.

Step 2: Existence of an interim best response. Fix a behavioural strategy σ used by all other temporal parts. The interim agent at a random dot chooses an action $a \in A(h)$ to maximise the expected continuation payoff $\sum_{\theta,d} P(\theta, d)\,v_d(a, \sigma_{-d}, \theta)$. Since $A(h)$ is finite, a maximiser exists. Moreover, since this is a single-agent game, all temporal parts share the same utility over terminal nodes, so every temporal part agrees on which action (or mixture over actions) is optimal given σ. This defines a best-response correspondence $\text{BR} : \Delta(A(h)) \rightrightarrows \Delta(A(h))$.

Step 3: Existence of an interim equilibrium. The set $\Delta(A(h))$ of behavioural strategies is compact and convex. The best-response correspondence BR is upper hemicontinuous (since the continuation payoffs are continuous in σ) and convex-valued (since the agent's expected payoff is linear in their own mixing probabilities). By Kakutani's fixed-point theorem, there exists a behavioural strategy $\hat{\sigma}$ with $\hat{\sigma} \in \text{BR}(\hat{\sigma})$: when all temporal parts play $\hat{\sigma}$, each finds $\hat{\sigma}$ optimal.

Step 4: Existence of a planning optimum. The planning objective,

$$V^{\text{plan}}(\sigma) \;=\; \sum_{\theta} \rho(\theta) \sum_{\mathbf{a}} \Pr_{\sigma}(\mathbf{a})\,U(\mathbf{a}, \theta),$$

is a continuous function on the compact set $\Delta(A(h))$, so a maximiser σ^* exists.

Step 5: The planning optimum is interim-optimal. Suppose for contradiction that σ^* is planning-optimal but not interim-optimal: there exists an action $a' \in A(h)$ such that the interim agent

strictly prefers deviating to a'. That is,

$$\sum_{\theta,d} P(\theta,d)\, v_d(a', \sigma^*_{-d}, \theta) \;>\; \sum_{\theta,d} P(\theta,d)\, v_d(\sigma^*, \sigma^*_{-d}, \theta).$$

Consider the perturbed behavioural strategy $\sigma_\varepsilon = (1-\varepsilon)\sigma^* + \varepsilon\,\delta_{a'}$ for small $\varepsilon > 0$. The first-order change in the planning payoff is:

$$\frac{d}{d\varepsilon} V^{\text{plan}}(\sigma_\varepsilon)\Big|_{\varepsilon=0} \;=\; \sum_{\theta}\sum_{d} \rho(\theta)\,\sigma_d(\theta,\sigma^*)\big[v_d(a', \sigma^*_{-d}, \theta) - v_d(\sigma^*, \sigma^*_{-d}, \theta)\big].$$

Since $P(\theta,d) = \rho(\theta)\,\sigma_d(\theta,\sigma^*)/C(\sigma^*)$ and the interim agent strictly prefers a', this equals

$$C(\sigma^*)\sum_{\theta,d} P(\theta,d)\big[v_d(a', \sigma^*_{-d}, \theta) - v_d(\sigma^*, \sigma^*_{-d}, \theta)\big] \;>\; 0.$$

But this contradicts σ^* being a planning optimum, since V^{plan} could be improved by shifting toward a'. □

B.4 Proof of Theorem 4.1

Proof. For part (i), suppose the game is *payoff-independent* and the dot counts are symmetric: $|D(\theta,n,\mathbf{a})| = |D(\theta,n',\mathbf{a})|$ for all $n, n', \theta, \mathbf{a}$. The argument has two steps: a symmetric-count step relating the planning derivative to the per-agent sum of interim advantages, and a payoff-independence step identifying that per-agent sum with the genuine occupant interim FOC.

Step 1 (symmetric counts). Each agent's dot share is $|D(\theta,n,\mathbf{a})|/|D(\theta,\mathbf{a})| = 1/|N|$ in every $(\theta,\mathbf{a})$. Fix $a' \in A(h)$ and perturb $\sigma_\varepsilon = (1-\varepsilon)\sigma + \varepsilon\,\delta_{a'}$. Following the Step 5 calculation of Theorem 3.1, the first-order change in the planning payoff is

$$\frac{d}{d\varepsilon} V^{\text{plan}}(\sigma_\varepsilon)\Big|_{\varepsilon=0} \;=\; \frac{C(\sigma)}{|N|}\sum_{n}\big(V^{\text{int}}_n(a') - V^{\text{int}}_{n,\text{inc}}(\sigma)\big),$$

where $V^{\text{int}}_n(a') = \sum_{\theta,\mathbf{a},\, d\in D(\theta,\mathbf{a})} P(\theta,\mathbf{a},d)\, u_n(\mathbf{a}_{d\to a'},\theta)$ is agent n's thirder-expected payoff from deviating to a' at the occupied dot, $V^{\text{int}}_{n,\text{inc}}(\sigma) = \sum_{a'}\sigma(a')\,V^{\text{int}}_n(a')$ is the incumbent value, $P(\theta,\mathbf{a},d) = \rho(\theta)\Pr_\sigma(\mathbf{a})/C(\sigma)$, and $\mathbf{a}_{d\to a'}$ denotes $\mathbf{a}$ with its action at dot d replaced by a'.

Step 2 (payoff independence). The actual interim agent wakes at a single dot d and earns only the *occupant's* payoff $u_{\iota(d)}$, so its first-order condition is the occupant sum

$$\text{FOC}^{\text{int}}(\sigma) \;=\; \sum_{\theta,\mathbf{a},\, d\in D(\theta,\mathbf{a})} P(\theta,\mathbf{a},d)\,\Big[\,u_{\iota(d)}(\mathbf{a}_{d\to a'},\theta) - \sum_{a''}\sigma(a'')\,u_{\iota(d)}(\mathbf{a}_{d\to a''},\theta)\,\Big].$$

The difference between the per-agent sum of Step 1 and this occupant sum is, at each $(\theta,\mathbf{a},d)$, the cross-term $\sum_{n\neq\iota(d)} u_n(\mathbf{a}_{d\to a'},\theta)$. Under payoff independence, u_n for $n \neq \iota(d)$ does not depend on the action at d (which is not agent n's own dot), so $u_n(\mathbf{a}_{d\to a'},\theta) = u_n(\mathbf{a},\theta)$ is *independent of* a'. Being constant in a', this cross-term contributes equally to $V^{\text{int}}_n(a')$ and to its σ-weighted incumbent (since $\sum_{a'}\sigma(a') = 1$), and so cancels in every advantage. Hence the per-agent sum equals the occupant FOC, and

$$\frac{dV^{\text{plan}}}{d\sigma} \;=\; \frac{C(\sigma)}{|N|}\cdot \text{FOC}^{\text{int}}(\sigma).$$

Since $C(\sigma)/|N| > 0$, the planning derivative is non-positive in every direction a' if and only if the occupant interim FOC is; the planning optimum is therefore interim-optimal, and there is no Newcomb tension. (Symmetric counts are used only in Step 1; payoff independence only in Step 2.)

For part (ii), suppose $|D(\theta_0, n_0, \mathbf{a}_0)| \neq |D(\theta_0, n_1, \mathbf{a}_0)|$ for some $(\theta_0, \mathbf{a}_0)$ and agents n_0, n_1. Then the social weights $|D(\theta_0, n, \mathbf{a}_0)|/|D(\theta_0, \mathbf{a}_0)|$ differ across agents in the compound state $(\theta_0, \mathbf{a}_0)$. In the planning derivative $dV^{\text{plan}}/d\sigma$, the compound-state formula weights each agent's payoff by their dot share *within each action profile*; agents with more dots in a given compound state receive proportionally larger weight. In the interim FOC, the thirder adjusts the probability of each $(\theta, \mathbf{a})$ proportionally to $|D(\theta, \mathbf{a})|$, which systematically overweights compound states in which more dots are reached. Because the dot shares differ across agents under $(\theta_0, \mathbf{a}_0)$, the planning and thirder weightings of the agents' payoffs diverge. Within the duplicating-AMD family this gap has the closed form $(b-a)/[2(a+b-c-d)]$ (Proposition 4.3), vanishing exactly on the measure-zero set $\{b = a\}$; the same reasoning applies family-by-family, which is the sense in which the failure is generic.

For the payoff-interdependence channel, symmetric dot counts do not restore the equation. Consider a single-state game in which agent 1 occupies dot d_1 and agent 2 occupies dot d_2, both reached on every play, with $A(h) = \{C, D\}$ and prisoner's-dilemma payoffs: each agent receives 3 if both play C, 1 if both play D, and a defector receives 5 against a cooperator's 0. Dot counts are symmetric (one dot per agent in every profile), so the thirder and the planner agree on the probabilities of the compound states. Writing p for the probability assigned to C, the planning payoff is $V^{\text{plan}}(p) = 3p^2 + 5p(1-p) + (1-p)^2$, with $dV^{\text{plan}}/dp = 3 - 2p > 0$, so the planner plays C for sure. The occupant's advantage to D at their own dot, however, is $[5p + (1-p)] - 3p = 1 + p > 0$ for every p, so the unique interim fixed point is pure D. The planning and interim optima diverge, and the proportional planning–interim equation fails at every σ: the cross-terms cancelled in Step 2 of part (i) are exactly what is missing, since each agent's payoff moves with the *other* agent's action, and the occupant's first-order condition omits the externality that the planner internalises. □

B.5 Proof of Proposition 4.2

Proof. Let $S(x, \theta)$ denote the scoring rule, where $x \in [0, 1]$ is the reported credence in Tails; strict properness means that $\mathbb{E}_{\theta \sim q}[S(x, \theta)]$ is uniquely maximised at $x = q$. In state H (probability $1/2$) the original has one awakening and the clone none; in state T each has one. The awakening structure does not depend on the report, so the planner's distribution over (state, awakening) pairs is $\frac{1}{2}$ on (H, orig), $\frac{1}{4}$ on (T, orig), and $\frac{1}{4}$ on (T, clone), and the planning payoff is the expected occupant score under this distribution:

$$V^{\text{plan}}(x) = \tfrac{1}{2}\, S(x, H) + \tfrac{1}{4}\, S(x, T) + \tfrac{1}{4}\, S(x, T) = \tfrac{1}{2}\, S(x, H) + \tfrac{1}{2}\, S(x, T).$$

This is the expected score under the credence $(\frac{1}{2}, \frac{1}{2})$, so strict properness gives $x^{\text{plan}} = \frac{1}{2}$.

The interim agent assigns thirder probability $1/3$ to each of the three awakenings, so

$$V^{\text{int}}(x) = \tfrac{1}{3}\, S(x, H) + \tfrac{1}{3}\, S(x, T) + \tfrac{1}{3}\, S(x, T) = \tfrac{1}{3}\, S(x, H) + \tfrac{2}{3}\, S(x, T).$$

This is the expected score under the thirder credence $(\frac{1}{3}, \frac{2}{3})$, so strict properness gives $x^{\text{int}} = \frac{2}{3} \neq \frac{1}{2} = x^{\text{plan}}$.

Finally, randomisation does not close the gap. Each agent has at most one payoff-relevant awakening per state, so no visit would ideally act differently from any other, and mixing over

reports is weakly dominated by the optimal pure report under both objectives. The Newcomb tension therefore persists for any behavioural strategy. □

B.6 Proof of Proposition 4.4

Proof. Since $a + b > c + d$, the planning payoff $V^{\text{plan}}(\sigma) = [(a+b)\sigma - (a+b-c-d)\sigma^2]/2$ is strictly concave, and $a+b > 2(c+d)$ places its critical point σ^* in $(0, 1)$, so σ^* is the planning optimum and is non-degenerate. At X, the occupant (Player 1) obtains 0 from Exit and either a or c from Continue (according to the action at Y), both strictly positive; at Y, the occupant (Player 2) obtains b from Exit and $d > b$ from Continue. Continue therefore yields a strictly higher continuation payoff than Exit to the occupant of every reached dot, and hence strictly exceeds the mixture σ^*, which places positive weight on Exit. For any belief P^{OB} it follows that $V^{\text{OB}}(\text{Continue}) > V^{\text{OB}}(\text{Exit})$, so the indifference across the support of σ^* required by Definition 2.7(b) is unattainable; equivalently, $\beta = \delta_{\text{Continue}}$ is a dominating mixture in the sense of Theorem 2.2(i). (The interim fixed point is the corner $\hat{\sigma} = 1$: both occupants strictly prefer to continue.) □

C Multi-State Example: Sleeping Beauty Behind the Wheel

Consider a game with two equiprobable states, $\theta \in \{0, 1\}$ with $\rho(0) = \rho(1) = 1/2$. In state $\theta = 0$, the game is the standard AMD: the agent traverses intersections X_0 and Y_0, with payoffs 0 (Exit at X_0), 4 (Exit at Y_0), and 1 (Continue past Y_0). In state $\theta = 1$, an additional intersection Z_1 is inserted between X_1 and Y_1, with Exit at Z_1 leading to its own terminal node with payoff 3. Exit at Y_1 gives 5 (rather than 4), and the remaining payoffs are unchanged. See Figure 4.

Under behavioural strategy σ, which continues with probability q at any intersection, the expected payoffs for each state are $V_0(q) = 4q - 3q^2$ and $V_1(q) = 3q + 2q^2 - 4q^3$. Thus, computing overall expected utility and differentiating, we can find the optimal q:

$$q^* \;=\; \frac{-1+\sqrt{85}}{12} \;\approx\; 0.685.$$

To compute the interim fixed point, we need to compute thirder beliefs and thus the normalising constant, which is $C(q) = \frac{1}{2}(1+q)+\frac{1}{2}(1+q+q^2) = 1+q+\frac{1}{2}q^2$. The continuation-payoff derivatives at each dot are:

$$\begin{aligned} g_{X_0} &= 4 - 3q, & g_{Y_0} &= -3, \\ g_{X_1} &= 3 + 2q - 4q^2, & g_{Z_1} &= 2 - 4q, \\ g_{Y_1} &= -4. \end{aligned}$$

The interim first-order condition is

$$\text{FOC}^{\text{int}}(q) \;=\; \frac{7 - 2q - 12q^2}{2\,C(q)} \;=\; 0 \qquad \Longrightarrow \qquad \hat{q} \;=\; \frac{-1+\sqrt{85}}{12} \;\approx\; 0.685.$$

The planning optimum and the interim fixed point coincide, as guaranteed by Theorem 3.1. The additional intersection Z_1 in $\theta = 1$ inflates the number of awakenings in that state (three dots versus two in $\theta = 0$), causing the thirder to overweight $\theta = 1$. Yet this belief distortion does not create a Newcomb tension: the per-dot continuation-payoff reasoning of Theorem 3.1 absorbs

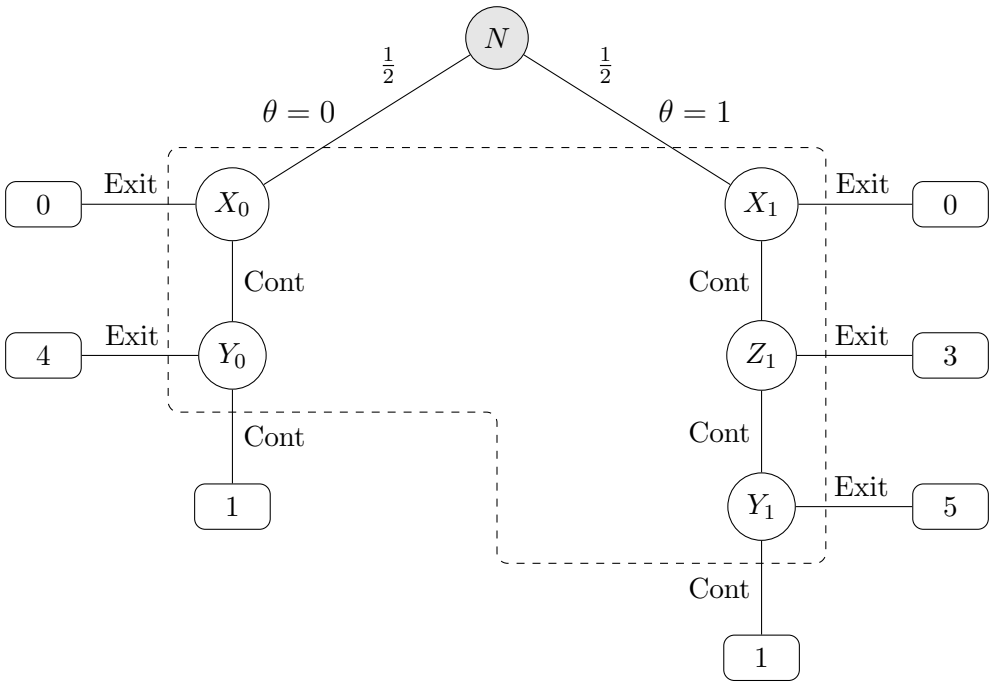


Figure 4: Sleeping Beauty Behind the Wheel. In $\theta = 0$, the standard AMD with intersections X_0 and Y_0. In $\theta = 1$, an additional intersection Z_1 is inserted, with Exit at Z_1 leading to its own terminal node with payoff 3. Exit at Y_1 pays 5 (versus 4 at Y_0). All five decision nodes lie in a single information set (dashed box).

it. The expectation paradox of Grove and Halpern (1997) discussed in Section 2.3 persists at q^*, where the thirder's interim valuation of the game, $\frac{127\sqrt{85}-120}{565} \approx 1.86$, still exceeds its ex ante value, $V^{\text{plan}}(q^*) = \frac{85\sqrt{85}-127}{432} \approx 1.52$.

Duplicating Sleeping Beauty Behind the Wheel. Finally, the two features of this section's examples combine without incident. Take the multi-state game of Figure 4 and create a clone (Player 2) at the first Continue branch in each state (at Y_0 in $\theta = 0$, at Z_1 in $\theta = 1$), with equal payoffs for the two players at every terminal reached after the cloning. Since $u_1 = u_2$ wherever both are defined, the planning payoff is exactly that of the non-duplicating game, and the occupant interim FOC sums the same dot–derivative products to the same polynomial, $C \cdot \text{FOC}^{\text{int}} = (7 - 2\sigma - 12\sigma^2)/2$, so the interim fixed point again coincides with the planning optimum, $\hat{\sigma} = \sigma^* \approx 0.685$. Despite the asymmetric dot counts, identical payoffs leave the game on the knife-edge at which the generic tension of Theorem 4.1(ii) vanishes, just as in the identical-payoff duplicating AMD; payoff divergence at any terminal restores it, by the same mechanism as Proposition 4.3.

References

Aumann, R. J., Hart, S., and Perry, M. (1997). The absent-minded driver. *Games and Economic Behavior*, 20(1):102–116.

Grove, A. J. and Halpern, J. Y. (1997). On the expected value of games with absentmindedness. *Games and Economic Behavior*, 20(1):51–65.

Halpern, J. Y. (1997). On ambiguities in the interpretation of game trees. *Games and Economic Behavior*, 20(1):66–96.

Hammond, P. J. (1988). Consequentialist foundations for expected utility. *Theory and Decision*, 25(1):25–78.

Harsanyi, J. C. (1955). Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. *Journal of political economy*, 63(4):309–321.

Janda, P. (2024). Doppelgänger changes the game. *Episteme*, 21(4):1182–1207.

Kierland, B. and Monton, B. (2005). Minimizing inaccuracy for self-locating beliefs. *Philosophy and Phenomenological Research*, 70(2):384–395.

Kreps, D. M. and Wilson, R. (1982). Sequential equilibria. *Econometrica: Journal of the Econometric Society*, pages 863–894.

Kuhn, H. W. (1953). Extensive games and the problem of information. In Kuhn, H. W. and Tucker, A. W., editors, *Contributions to the Theory of Games*, volume II of *Annals of Mathematics Studies*, pages 193–216. Princeton University Press, Princeton.

Lewis, D. (2001). Sleeping beauty: reply to elga. *Analysis*, 61(3):171–176.

Machina, M. J. (1989). Dynamic consistency and non-expected utility models of choice under uncertainty. *Journal of Economic Literature*, 27(4):1622–1668.

Nozick, R. (1969). Newcomb's problem and two principles of choice. In Rescher, N., editor, *Essays in Honor of Carl G. Hempel*, pages 114–146. D. Reidel, Dordrecht.

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

Piccione, M. and Rubinstein, A. (1997). On the interpretation of decision problems with imperfect recall. *Games and Economic Behavior*, 20(1):3–24.

Ross, J. (2010). Sleeping beauty, countable additivity, and rational dilemmas. *Philosophical Review*, 119(4):411–447.

Schwarz, W. (2015). Lost memories and useless coins: revisiting the absentminded driver. *Synthese*, 192(9):3011–3036.

Spohn, W. (2025). The puzzle of the absent-minded driver is not about absent-mindedness, but about indexical belief. *Homo Oeconomicus*, pages 1–13.

Strotz, R. H. (1955). Myopia and inconsistency in dynamic utility maximization. *The Review of Economic Studies*, 23(3):165–180.

Vergopoulos, V. (2011). Dynamic consistency for non-expected utility preferences. *Economic Theory*, 48(2-3):493–518.

Winkler, P. (2017). The sleeping beauty controversy. *The American Mathematical Monthly*, 124(7):579–587.