A Memory Design Folk Theorem

Sami Petersen*

5 July 2026
Preliminary draft

Abstract

The logic behind the folk theorem goes from delivering an 'anything goes' result to an 'anything is implementable' result when applied to information design in games without perfect recall. I show that when agents (artificial ones, say) can be made to play a game multiple times without remembering their own past play, a designer can construct a *finite*-horizon repetition of the game with a public information structure that sustains any folk theorem payoff profile.

*Department of Economics, UC Berkeley.

Contact: sami.petersen@berkeley.edu
First version: 19 March 2026

My thanks to Federico Echenique for helpful discussion.

1 Introduction

There is a finite normal form base game. A designer is able to build extensive forms from this base game, including ones which limit the players' memory. Which payoffs can she implement? This question is particularly relevant to designing incentives around AI agents, whose memory can flexibly be restricted: simply make two copies and run them separately; neither will have any recollection of what the other copy did.

I study this question in the context of multi-stage games. The designer picks a number of stages in which the base game will be played, and makes a public announcement at each stage about the history of play thus far. I find that when players (can be made to) forget their own past actions, the designer's ability to incentivise the players is formidable. With only a finite number of stages, it is possible to implement any payoff whatsoever that can be sustained in some equilibrium of the infinitely repeated perfect-recall version of the base game. This holds both if the players receive a flow payoff at each stage, and if the payoffs are fixed by the actions taken at one 'real' stage.

The main result establishes that given a sufficiently long but finite horizon, any feasible and strictly individually rational profile of payoffs is implementable, in multiselves sequential equilibrium, provided the base game has full dimension (Theorem 1). In other words, a direct analogue of the Fudenberg and Maskin (1986) folk theorem holds under 'memory design' even when there is complete information and the time horizon is both finite and commonly known.[1]

Literature (terse). Piccione and Rubinstein (1997) launched the modern literature on imperfect recall in economic theory, focusing on (single-agent) decision problems. Lambert, Marple and Shoham (2019) generalised a collection of familiar solution concepts to (multi-player) games without perfect recall. Koh and Sanguanmoo (2025) build a general framework for designing memory in games, holding fixed the underlying extensive form.[2] This paper studies the joint design of memory and extensive forms.

[1] This also holds in the two-player case without full dimensionality (Proposition 1) and in the general multi-player case for Nash-reversion payoffs (Proposition 2).

[2] Relatedly, Chen, Ghersengorin and Petersen (2024) apply a form of constrained memory- and extensive-form design with a single agent to address the 'deceptive alignment' problem in AI. Levy and Szentes (2025) develop a revelation principle to study single-agent implementation in multi-stage decision problems with imperfect recall.

The important work of Salcedo (2013, 2017) and Doval and Ely (2020) considers the joint design problem, under perfect recall and for a particular class of extensive forms that partly excludes the present setting.[3] The environment studied in this paper is instead closely related to Forges (1986) and Myerson (1986), who studied implementation in multi-stage games with mediators, though with the extensive form held fixed. Since my designer is additionally 'omniscient' and I do not assume that players can commit, more closely related are the recent multi-stage extensions of Bayes correlated equilibrium which impose sequential rationality (Makris & Renou, 2023).

Finite-horizon folk theorems have been established before under particular conditions on the base game such as multiplicity of Nash payoffs (cf. Benoit & Krishna, 1985; Smith, 1995), or on what the players commonly know about each other (e.g. Kreps et al., 1982) or about the time horizon (e.g. Neyman, 1999; Hauser & Salmi, 2026). The results in this paper require no assumptions on the base game beyond those assumed in the standard infinite-horizon folk theorems with complete information, and they allow all players to commonly know the horizon. What the players do not know in this paper is their *current* location in the game (in particular, the current stage number).[4]

There are papers which obtain folk theorems in other non-standard settings. Tennenholtz (2004) formalises and generalises an observation from Rubinstein (1998) to prove a folk theorem for 'program games', a setting where players (computers) get to see each other's source code before playing a game. Kalai et al. (2010) show that folk theorem payoffs can be obtained when players can make conditional commitments. More recently, Kovarik, Oesterheld and Conitzer (2024) show that a folk theorem also appears when players observe a simulation of the game they are about to play, which itself includes a simulation of the next game down, and so on. This paper can be interpreted as pointing out another environment where folk-type dynamics emerge. But rather than emerging as vast indeterminacy of equilibrium in a fixed setting (an 'anything goes' prediction), they instead emerge as equilibria that can be deliberately selected by an information designer ('anything is implementable').

[3] An important exception: when payoffs in the multi-stage game depend only on play in the final stage.

[4] The kinds of incentives that can be generated by self-locating uncertainty were illustrated in a delightful example from Nishihara (1997), where cooperation in the one-shot prisoner's dilemma is achieved, despite recall being perfect, by jointly designing information and the extensive form.

2 Model

Environment. Fix a finite normal-form game $G = (I, (A_i, u_i)_{i \in I})$ which we shall call the *base game*. From it we may construct a multi-stage game

$$G_T(\tilde{u}) := (I, A, H, P, \mathcal{S}, (\tilde{u}_i)_{i \in I})$$

where $T < \infty$ is the number of stages, $A := \prod_{i \in I} A_i$ are the pure action profiles realisable at each stage, $Z := A^T$ are the terminal histories in H, the player correspondence $P : H \rightrightarrows I$ satisfies $P(h) = I$ at every $h \in H \setminus Z$, the information sets $\mathcal{S} = \{S_1, ..., S_{|I|}\}$ exhibit full amnesia ($S_i = H \setminus Z$ for every $i \in I$), and $\tilde{u}_i : Z \to \mathbb{R}$ is the payoff of each player $i \in I$. Call $G_T(\tilde{u})$ the *T-stage extension* of G with *extended payoffs* $\tilde{u} = (\tilde{u}_i)_{i \in I}$.

The interpretation is that the players are called to play the base game T times but without remembering any past play. In particular, the players do not recall their own past actions, nor do they know the calendar time (the current stage number $t \in \{1, ..., T\}$). I will consider two forms of payoff extensions. Given a base game G with payoffs $u = (u_i)_{i \in I}$, the *average-payoff extension* is given by $\tilde{u}(a_1, ..., a_T) = \frac{1}{T} \sum_{t=1}^{T} u(a_t)$ and the *final-stage payoff extension* is given by $\tilde{u}(a_1, ..., a_T) = u(a_T)$ for any sequence $(a_t)_{t=1}^{T}$ in A with $T \in \mathbb{N}$. Other specifications are conceivable but I will focus on these. The average-payoff extension is familiar from finitely-repeated games, and the final-stage extension has a natural interpretation: the players are 'simulated' multiple times before playing the 'real' game (e.g. Chen, Ghersengorin & Petersen, 2024). I will say that a payoff extension is *admissible* iff it coincides with one of these.

Information. An *information structure* for a multi-stage game $G_T(\tilde{u})$ is a pair (M, π) where M is a finite set and $\pi : \bigcup_{t=1}^{T}(M^{|I|} \times A)^{t-1} \to \Delta(M^{|I|})$ is a probability kernel. The interpretation is that a profile $m = (m_i)_{i \in I}$ of player-specific messages is drawn at each stage $t \in \{1, ..., T\}$ from a distribution $\pi(\cdot|a^{t-1}, m^{t-1})$ that depends on previous play and messages. We will interpret messages as being publicly observed whenever it is the case that messages always coincide across players under π.

A *repeated memory mechanism* for $(G, \tilde{u})$ is a triple $(G_T(\tilde{u}), M, \pi)$, where $G_T(\tilde{u})$ is the T-stage extension of G with extended payoffs $\tilde{u}$, and (M, π) is an associated public information structure. These formally induce another multi-stage game where, at each stage, players receive messages, act once simultaneously, and then forget both the messages and the actions before entering the next stage. Messages can therefore be identified with information sets. As a result, each player $i \in I$ is limited to (behavioural) strategies $\sigma_i : M \to \Delta(A_i)$ mapping individual messages to mixed actions. I will often use the term 'repeated memory mechanism' to refer to this induced game.

Implementation. The solution concept will be *strong memory correlated equilibrium* (sMCE) from Koh and Sanguanmoo (2025) which, in the current setting, coincides with *multiselves sequential equilibrium* from Lambert, Marple and Shoham (2019). Both coincide with sequential equilibrium when recall is perfect. I will say that a repeated memory mechanism *implements* a payoff profile $v \in \mathbb{R}^{|I|}$ iff there is a sMCE of the game induced by the mechanism whose associated expected payoffs are v.

Note that to interpret the implemented payoffs, we need to specify how the extended payoff functions $\tilde{u} = (\tilde{u}_i)_{i \in I}$ in the multi-stage extension relate to the ones $u = (u_i)_{i \in I}$ in the base game. But provided the extension is admissible, I will say that a payoff profile $v \in \mathbb{R}^{|I|}$ is *feasible* tout court iff $v \in \text{cvh}(\{u(a) : a \in A\})$, and that it is *strictly individually rational* iff $v_i > \underline{v}_i := \min_{\alpha_{-i} \in \prod_{j \neq i} \Delta(A_j)} \max_{a_i \in A_i} u_i(a_i, \alpha_{-i})$ for all $i \in I$. The base game is said to have *full dimension* iff the set of feasible and strictly individually rational payoffs has a non-empty interior as a subset of $\mathbb{R}^{|I|}$.

3 Results

Here is the main result.

Theorem 1. Fix any finite full-dimensional base game G. For any profile $v \in \mathbb{R}^{|I|}$ of feasible and strictly individually rational payoffs, and any admissible payoff extension $\tilde{u}$, there is a finite-horizon repeated memory mechanism for $(G, \tilde{u})$ which implements v.

Proof. Appendix A.1.[5] □

Remark 1. Memory design delivers an exact analogue of Benoit and Krishna's (1985) finite folk theorem but under the minimal conditions of Fudenberg and Maskin (1986). It is exact rather than approximate since the designer acts as a public randomisation device. The result holds when the designer only makes public recommendations, and it can be further strengthened to cover the 'correlated minmax' payoffs as well.

The proof of Theorem 1 is more cumbersome than Fudenberg and Maskin's because of the finite time horizon, but the incentives are completely analogous. The key observation is that (self-locating) uncertainty about the calendar time can be made to play the role of ('objective') uncertainty about how many times the game will be repeated. Because players cannot tell which stage $1, 2, ..., T$ they are in at any given moment, they are never certain that they currently find themselves at the final stage. Player i's self-locating belief $1 - \mu_i(T) > 0$ that they are at a non-terminal stage thereby acts like the discount factor $\delta > 0$ in the infinitely repeated case with perfect recall: both can be interpreted as the probability that the stage game will be played at least one more time. That is what prevents players from inducting backwards.

Having induced self-locating uncertainty, the designer then only needs to tune two things. First, raising the horizon T mimics an increase in patience δ. That makes punishments more lengthy and likely, which widens the set of implementable payoffs. Second, by sending messages in each stage which recommend actions to the players, the designer can replicate the carrot-and-stick strategy used to prove Theorem 2 of Fudenberg and Maskin (1986), without revealing the current stage number to the players.

This reasoning naturally extends to the standard settings without full dimensionality.

[5] Indeed the proof establishes that such a mechanism exists for any longer horizon too. See Appendix A.1 for explicit bounds.

Proposition 1 (Two players)**.** Fix any finite two-player stage game G. For any profile $v \in \mathbb{R}^{|I|}$ of feasible and strictly individually rational payoffs, and any admissible payoff extension $\tilde{u}$, there exists a finite-horizon repeated memory mechanism for $(G, \tilde{u})$ which implements v.

Proof. Appendix A.2. $\square$

Proposition 2 (Nash reversion)**.** Fix any finite stage game G. For any feasible payoff profile $v \in \mathbb{R}^{|I|}$ satisfying $v_i > u_i(\bar{\sigma})\ \forall i \in I$ for some Nash equilibrium $\bar{\sigma}$ of G, and for any admissible payoff extension $\tilde{u}$, there exists a finite-horizon repeated memory mechanism for $(G, \tilde{u})$ which implements v.

These results are, respectively, analogues of Theorem 1 and Theorem C in Fudenberg and Maskin (1986), the latter being attributed to Friedman (1971, 1977). Fortunately the proof of Proposition 2 is no longer than the proof of its analogue, the regular Nash reversion folk theorem. I therefore include it here.

Proof of Proposition 2. Fix a finite stage game $G = (I, (A_i, u_i)_{i \in I})$ and some Nash equilibrium $\bar{\sigma}$ of G. Let $v \in \mathbb{R}^{|I|}$ denote some feasible payoff profile satisfying $v_i > u_i(\bar{\sigma})$ for every player $i \in I$. Since v is feasible, it can be achieved by some (correlated) mixed action profile $\alpha \in \Delta(A)$, where $A := \prod_{i \in I} A_i$. Fix this profile.

Let us write the extended payoffs as $\tilde{u}(z) = \sum_{t=1}^{T} w_t u(a_t)$ for each $z = (a_1, ..., a_T) \in A^T$, where $w = (w_1, ..., w_T) \in \Delta(\{1, ..., T\})$ are normalised weights. This nests the two admissible payoff extensions.

Consider the following information structure. At the empty history, draw a pure action profile $a \sim \alpha$, and publicly recommend 'a'. At any history where play has been obedient throughout, independently draw $\tilde{a} \sim \alpha$ and recommend $\tilde{a}$. At any history where a player has disobeyed in the past, recommend the profile $\bar{\sigma} \in \prod_{i \in I} \Delta(A_i)$ of possibly mixed actions (that is, at every stage going forward).

If players are obedient, their payoffs will be $\sum_{t=1}^{T} w_t \mathbb{E}_{a_t \sim \alpha}[u(a_t)] = u(\alpha) \equiv v$ as required.

Each player has $|\operatorname{supp} \alpha| + 1$ sets of incentive constraints, one for each information set generated by the information structure. It suffices for sMCE to verify each of these separately, rather than invoking the one-shot deviation principle.

It is immediate that the off-path recommendation $\bar{\sigma}$ is obeyed, regardless of beliefs: it prescribes instantaneous mutual best replies and, regardless of the actions chosen, the same will be recommended in the future. It remains to verify the on-path incentives.

Supposing that players $I \setminus \{i\}$ always obey, player i's continuation payoff from obeying after receiving recommendation $a \in \operatorname{supp} \alpha$ is

$$\sum_{t=1}^{T} \mu_i(t \mid a) \left(w_t u_i(a) + \sum_{s=t+1}^{T} w_s u_i(\alpha) \right)$$

while i's continuation payoff from a 'phantom' deviation[6] that plays $a_i' \in A_i \setminus \{a_i\}$ is

$$\sum_{t=1}^{T} \mu_i(t \mid a) \left(w_t u_i(a_i', a_{-i}) + \sum_{s=t+1}^{T} w_s u_i(\bar{\sigma}) \right).$$

Every stage is eventually reached, and on-path recommendations are i.i.d., so $\mu_i(\cdot \mid a) = \text{Unif}\{1, ..., T\}$ by belief consistency. The a-obedience constraint for player i is then

$$\begin{aligned}
\sum_{t=1}^{T} \left(w_t u_i(a) + \sum_{s=t+1}^{T} w_s u_i(\alpha) \right) &\geq \sum_{t=1}^{T} \left(w_t u_i(a_i', a_{-i}) + \sum_{s=t+1}^{T} w_s u_i(\bar{\sigma}) \right) \\
\left[u_i(a) - u_i(a_i', a_{-i}) \right] \sum_{t=1}^{T} w_t &\geq \left[u_i(\bar{\sigma}) - u_i(\alpha) \right] \sum_{t=1}^{T} \sum_{s=t+1}^{T} w_s \\
\frac{u_i(a) - u_i(a_i', a_{-i})}{u_i(\bar{\sigma}) - v_i} &\leq \sum_{t=1}^{T} \sum_{s=t+1}^{T} w_s
\end{aligned}$$

since $(w_t)_{t=1}^{T} \in \Delta^{T-1}$ and $u_i(\alpha) > u_i(\bar{\sigma})$. By expressing the right side as

$$\sum_{t=1}^{T} \sum_{s=t+1}^{T} w_s = \sum_{t=1}^{T} \sum_{s=1}^{T} \mathbb{1}\{s > t\} w_s = \sum_{s=1}^{T} w_s \sum_{t=1}^{T} \mathbb{1}\{s > t\} = \sum_{s=1}^{T} w_s (s-1)$$

[6] A phantom deviation changes behaviour at whichever history is actual, while believing that play is otherwise on path (meaning obedient, in this case), even elsewhere in the current information set. See Lambert, Marple and Shoham (2019) for details and Koh and Sanguanmoo (2025) for discussion.

we obtain the required obedience constraints:

$$\sum_{t=1}^{T} t w_t \geq 1 + \frac{u_i(a_i', a_{-i}) - u_i(a)}{v_i - u_i(\bar{\sigma})} \qquad \forall a \in \operatorname{supp} \alpha,\ a_i' \in A_i,\ i \in I.$$

The game G is finite, so this is a finite collection of finite lower bounds.

The weights would all equal $w_t = 1/T$ under the average-payoff extension and $w_t = \mathbb{1}\{t = T\}$ under the final-stage payoff extension. The left side of the inequalities would therefore simplify to either $\frac{1}{T}\sum_{t=1}^{T} t = \frac{T+1}{2}$ or to $\sum_{t=1}^{T} t\mathbb{1}\{t = T\} = T$.

In either case, there is a finite choice of $T \in \mathbb{N}$ which ensures obedience in sMCE. □

References

Benoit, Jean-Pierre & Vijay Krishna. (1985). "Finitely Repeated Games". *Econometrica*, *53*(4), 905–922.

Carmona, Guilherme. (2008). "On the Full Dimensionality Assumption for the Discounted Folk Theorem". *Economics Letters*, *99*(2), 357–359.

Chen, Eric Olav, Alexis Ghersengorin, & Sami Petersen. (2024). "Imperfect Recall and AI Delegation" [GPI Working Paper Series].

Doval, Laura & Jeffrey C. Ely. (2020). "Sequential Information Design". *Econometrica*, *88*(6), 2575–2608.

Forges, Françoise. (1986). "An Approach to Communication Equilibria". *Econometrica*, *54*(6), 1375–1385.

Friedman, James W. (1971). "A Non-Cooperative Equilibrium for Supergames". *The Review of Economic Studies*, *38*(1), 1–12.

Friedman, James W. (1977). "Oligopoly and the Theory of Games" (Vol. 8). North-Holland Publishing Company.

Fudenberg, Drew & Eric Maskin. (1986). "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information". *Econometrica*, *54*(3), 533–554.

Hauser, Daniel & Julia Salmi. (2026). "Private Information About the Game Horizon" [Work in progress].

Kalai, Adam Tauman, Ehud Kalai, Ehud Lehrer, & Dov Samet. (2010). "A Commitment Folk Theorem". *Games and Economic Behavior*, *69*(1), 127–137.

Koh, Andrew & Sivakorn Sanguanmoo. (2025). "Memory Correlated Equilibrium" [Working paper].

Kovarik, Vojtech, Caspar Oesterheld, & Vincent Conitzer. (2024). "Recursive Joint Simulation in Games".

Kreps, David M., Paul R. Milgrom, John Roberts, & Robert Wilson. (1982). "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma". *Journal of Economic Theory*, *27*(2), 245–252.

Lambert, Nicolas S., Adrian Marple, & Yoav Shoham. (2019). "On Equilibria in Games with Imperfect Recall". *Games and Economic Behavior*, *113*, 164–185.

Levy, Matthew R. & Balázs Szentes. (2025, December). "Information Design for AI Proxies under Imperfect Recall" [Working paper].

Makris, Miltiadis & Ludovic Renou. (2023). "Information Design in Multi-Stage Games". *Theoretical Economics*, *18*(4), 1475–1509.

Myerson, Roger B. (1986). "Multistage Games with Communication". *Econometrica*, *54*(2), 323–358.

Neyman, Abraham. (1999). "Cooperation in Repeated Games When the Number of Stages Is Not Commonly Known". *Econometrica*, *67*(1), 45–64.

Nishihara, Ko. (1997). "A Resolution of N-Person Prisoners' Dilemma". *Economic Theory*, *10*(3), 531–540.

Piccione, Michele & Ariel Rubinstein. (1997). "On the Interpretation of Decision Problems with Imperfect Recall". *Games and Economic Behavior*, *20*(1), 3–24.

Rubinstein, Ariel. (1998). "Modeling Bounded Rationality". The MIT Press.

Salcedo, Bruno. (2013). "Implementation Without Commitment in Moral Hazard Environments".

Salcedo, Bruno. (2017, January). "Interdependent Choices" [Working paper].

Smith, Lones. (1995). "Necessary and Sufficient Conditions for the Perfect Finite Horizon Folk Theorem". *Econometrica*, *63*(2), 425–430.

Tennenholtz, Moshe. (2004). "Program Equilibrium". *Games and Economic Behavior*, *49*(2), 363–373.

A Appendix

A.1 Proof of Theorem 1

The proof is structured as follows. I first define some notation and an information structure that will be fixed throughout. Then I show that the off-path beliefs I will use are consistent with sMCE under this information structure when players are obedient (Lemma 1). Finally, I verify obedience. The last step has two parts; the final-stage payoff extension and the average flow payoff extension are handled separately.

Setup. Write V for the set of feasible and strictly individually rational payoff profiles of the finite stage game $G = (I, (A_i, u_i)_{i \in I})$. Fix some $v \in V$. Since v is feasible, it is attained by some (correlated) mixed action profile $\alpha \in \Delta(A)$. Fix this profile.

Write $\underline{v}_i := u_i(\beta^i)$ for the minmax payoff of player $i \in I$, meaning their payoff under a profile $\beta^i = (\beta^i_j)_{j=1}^{|I|}$ of mixed actions which satisfies

$$\beta^i_{-i} \in \underset{\alpha_{-i} \in \prod_{j \neq i} \Delta(A_j)}{\arg\min} \max_{a_i \in A_i} u_i(a_i, \alpha_{-i})$$
$$\beta^i_i \in \underset{a_i \in A_i}{\arg\max}\, u_i(a_i, \beta^i_{-i}).$$

By Theorem 1 of Carmona (2008), full dimensionality guarantees the existence of payoff vectors $w^1, ..., w^{|I|} \in V$ satisfying $w^i_i \in (\underline{v}_i, v_i)$ and $w^j_i > w^i_i$ for all $i, j \in I$ with $j \neq i$. Let $\alpha^j \in \Delta(A)$ denote a (correlated) mixed action profile satisfying $w^j = \mathbb{E}_{a \sim \alpha^j}[u(a)]$.

Consider an information structure which partitions the histories of G_T into three phases: good [I], stick [II], and carrot [III]. We begin by specifying what the designer announces at a given stage, phase by phase, before describing how each phase is reached.

Announcements:

[I] In the good phase, the designer independently draws $a \sim \alpha$ and announces $(a; \mathrm{I})$. This is interpreted as a recommendation to play a at the current stage of the good phase.

[II] In the stick phase, the designer independently draws $b^i \sim \beta^i$ from a profile of mixed actions which minmaxes the player $i \in I$ who most recently deviated from the designer's recommendation. The designer announces $(b^i; \mathrm{II}_i)$, interpreted as a recommendation to punish i by playing b^i at the current stage of the stick phase.

[III] In the carrot phase, the designer independently draws $a^i \sim \alpha^i$ and announces $(a^i; \mathrm{III}_i)$, where $i \in I$ is the player who deviated most recently.

Transitions:

[I] The game enters the good phase only at the null history, and remains there iff players remain obedient (until the game ends).[7]

[II] The game (re-)starts the stick phase iff it reaches a history ending in disobedience. Provided the players are subsequently obedient, this phase continues for $\ell \in \mathbb{N}$ stages or until the game ends, whichever comes first.

[III] The game starts the carrot phase only after ℓ stages of obedient play in the stick phase, and remains there provided players are obedient (until the game ends).

That completes our description of the information structure.

If players are obedient, their payoffs will be

$$v = \mathbb{E}_{a_t \sim \text{i.i.d.} \alpha}[\tilde{u}(a_1, \ldots, a_T)] = \begin{cases} \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}_{a \sim \alpha}[u(a)] & \text{under average-flow extended payoffs} \\ \mathbb{E}_{a_T \sim \alpha}[u(a_T)] & \text{under final-stage extended payoffs} \end{cases}$$

as required.

It remains to verify obedience in each of these two cases. Each player will have $|\operatorname{supp} \alpha| + \sum_{j \in I} \left(|\operatorname{supp} \beta^j| + |\operatorname{supp} \alpha^j|\right)$ sets of incentive constraints, one for each information set generated by the information structure. It suffices for MCE to verify each of these separately (rather than invoking the one-shot deviation principle). To confirm that the MCE is an sMCE, we need to verify that the specified beliefs during the stick and carrot phases are admissible. Hence the following lemma.

Lemma 1 (Off-path beliefs)**.** The following beliefs for each player $i \in I$ are consistent

[7] Given a history, let's say that players were *disobedient* at a particular stage iff exactly one player $i \in I$ deviated from the designer's recommendation at that stage. Players were *obedient* otherwise.

with sMCE under strategy profile $\sigma = (\sigma_i)_{i\in I}$ and information structure (M, π):

$$\mu_i(s,t|b^j;\mathrm{II}_j) = \frac{1}{|S|}, \qquad S := \{(s,t)\in\mathbb{N}^2 : 1\le s<t\le T,\ t-s\le \ell\};$$
$$\mu_i(s,t|a^j;\mathrm{III}_j) = \frac{1}{|C|}, \qquad C := \{(s,t)\in\mathbb{N}^2 : 1\le s<s+\ell<t\le T\}.$$

for any hypothesis (s,t) about the current stage $t\in\{1,...,T\}$ and most recent deviation $s\in\{1,...,T-1\}$, and any messages $(b^j;\mathrm{II}_j),(a^j;\mathrm{III}_j)\in M$ where $b^j\in\operatorname{supp}(\beta^j)$ and $a^j\in\operatorname{supp}(\alpha^j)$ with $j\in I$ a player who can deviate ($|A_j|>1$).

Proof. Write $\sigma=(\sigma_i)_{i\in I}$, where $\sigma_i : M\to\Delta(A_i)$, for the conjectured obedient (behavioural) strategy profile. Let $\hat\alpha_i := \mathrm{Unif}\,A_i$ and $\varepsilon_k := \frac{1}{k+2}$ and $\sigma_i^k := (1-\varepsilon_k)\sigma_i+\varepsilon_k\hat\alpha_i$, for $i\in I$ and $k\in\mathbb{N}$. Write μ_i^k for the resulting (self-locating) beliefs. Notice that $\lim_{k\uparrow\infty}\sigma^k=\sigma$.

The probability that j deviates at some particular stage $t\in\{1,...,T\}$ given σ^k is

$$\varepsilon_k\left(1-\frac{1}{|A_j|}\right) \;=:\; x_k^j.$$

The chance that only j deviates at a given stage is

$$x_k^j\prod_{i\neq j}(1-x_k^i) \;=:\; y_k^j$$

while the chance of exactly one deviator at a given stage is

$$\sum_{j\in I} y_k^j \;=:\; z_k.$$

The stick-phase posterior on $(s,t)\in S$ given just $(b^j;\mathrm{II}_j)$ (with $|A_j|>1$) is then

$$\begin{aligned}\mu_i^k(s,t|b^j;\mathrm{II}_j) &= \frac{\Pr_k\left[\{(s,t)\}\cap\{b^j;\mathrm{II}_j\}\right]}{\sum_{(s',t')\in S}\Pr_k\left[\{(s',t')\}\cap\{b^j;\mathrm{II}_j\}\right]} = \frac{y_k^j(1-z_k)^{t-s-1}\beta^j(b^j)}{\sum_{(s',t')\in S}y_k^j(1-z_k)^{t'-s'-1}\beta^j(b^j)}\\ &= \frac{(1-z_k)^{t-s-1}}{\sum_{(s',t')\in S}(1-z_k)^{t'-s'-1}} \longrightarrow \frac{1}{\sum_{(s',t')\in S}1} = \frac{1}{|S|}.\end{aligned}$$

Repeating the argument shows that the carrot-phase posterior on $(s,t) \in C$ is

$$\mu_i^k(s,t|a^j;\mathrm{III}_j) \longrightarrow \frac{1}{|C|}$$

as needed. $\square$

Proof of Theorem 1 (Part 1/2: final-stage payoff extension).

Let us now verify obedience under the final-stage payoff extension.

[I]

We begin with the good phase. Supposing that players $I \setminus \{i\}$ always obey, player i's payoff from obeying $a \in \operatorname{supp} \alpha$ is

$$\mu_i(T|a;\mathrm{I})u_i(a) + [1 - \mu_i(T|a;\mathrm{I})]\mathbb{E}_{a' \sim \alpha}[u_i(a')].$$

Deviating to a_i' instead yields

$$\mu_i(T|a;\mathrm{I})u_i(a_i', a_{-i}) + \sum_{t=T-\ell}^{T-1} \mu_i(t|a;\mathrm{I})\underline{v}_i + \sum_{t=1}^{T-\ell-1} \mu_i(t|a;\mathrm{I})w_i^i$$

since $\ell < T$. Because (i) every stage is always reached on path and (ii) on-path recommendations are i.i.d., players must have uniform beliefs $\mu_i(\cdot|a) = \frac{1}{T}$, for every $a \in \operatorname{supp} \alpha$. The a-obedience constraint for player i is therefore

$$u_i(a) + (T-1)v_i \geq u_i(a_i', a_{-i}) + \ell\underline{v}_i + (T-\ell-1)w_i^i \qquad \forall a \in \operatorname{supp} \alpha,\ a_i' \in A_i,\ i \in I.$$

Re-arranging, it suffices for $(T,\ell) \in \mathbb{N}^2$ to satisfy

$$T \geq 1 + \frac{u_i(a_i', a_{-i}) - u_i(a)}{v_i - w_i^i} - \ell\frac{w_i^i - \underline{v}_i}{v_i - w_i^i} \qquad \forall a \in \operatorname{supp} \alpha,\ a_i' \in A_i,\ i \in I.$$

Such a pair exists since the game G is finite and $v_i > w_i^i$.

[II]

We now verify obedience in the stick phase. Let $s \in \{1, ..., T-1\}$ denote the time of the most recent deviation, and consider the player $j \in I$ who deviated at s. (Recall that whenever there are multiple deviators, the players are considered to be obedient.) Player $i \in I$ has two sets of incentive constraints when j is being punished: one where $i = j$ and one where $i \neq j$. We check these in turn.

[II$_i$]

When player i is being punished in the stick phase, the designer will recommend some $b^i \in \operatorname{supp} \beta^i$. Player i's obedience constraint is

$$\sum_{s=1}^{T-1} \sum_{t=s+1}^{\min\{s+\ell,T\}} \mu_i(s,t \mid b^i; \mathrm{II}_i) \Big[\mathbb{1}\{t=T\} u_i(b^i) + \mathbb{1}\{t < T \leq s+\ell\} \underline{v}_i + \mathbb{1}\{s+\ell < T\} w_i^i\Big]$$
$$\geq \sum_{s=1}^{T-1} \sum_{t=s+1}^{\min\{s+\ell,T\}} \mu_i(s,t \mid b^i; \mathrm{II}_i) \Big[\mathbb{1}\{t=T\} u_i(a_i', b_{-i}^i) + \mathbb{1}\{t < T \leq t+\ell\} \underline{v}_i + \mathbb{1}\{t+\ell < T\} w_i^i\Big].$$

By Lemma 1, we may take beliefs to be uniform. They then cancel:

$$[u_i(b^i) - u_i(a_i', b_{-i}^i)] \sum_{s=1}^{T-1} \sum_{t=s+1}^{\min\{s+\ell,T\}} \mathbb{1}\{t=T\}$$
$$\geq \sum_{s=1}^{T-1} \sum_{t=s+1}^{\min\{s+\ell,T\}} \Big[(\mathbb{1}\{t < T \leq t+\ell\} - \mathbb{1}\{t < T \leq s+\ell\}) \underline{v}_i + (\mathbb{1}\{t+\ell < T\} - \mathbb{1}\{s+\ell < T\}) w_i^i\Big].$$

Simplifying,

$$[u_i(b^i) - u_i(a_i', b_{-i}^i)] \sum_{s=1}^{T-1} \mathbb{1}\{T \leq s+\ell\}$$
$$\geq \sum_{s=1}^{T-1} \sum_{t=s+1}^{\min\{s+\ell,T\}} \Big[\mathbb{1}\{t < T,\ s+\ell < T \leq t+\ell\} \underline{v}_i - \mathbb{1}\{s+\ell < T \leq t+\ell\} w_i^i\Big].$$

Choosing a pair (T, ℓ) which satisfies $T \geq \ell + 1$ lets us simplify the left side:

$$[u_i(b^i) - u_i(a_i', b_{-i}^i)]\ell \geq \sum_{s=1}^{T-1} \sum_{t=s+1}^{\min\{s+\ell,T\}} \mathbb{1}\{s+\ell < T \leq t+\ell\}\Big(\mathbb{1}\{t < T\}\underline{v}_i - w_i^i\Big)$$
$$= (\underline{v}_i - w_i^i)\sum_{s=1}^{T-1}\sum_{t=s+1}^{s+\ell} \mathbb{1}\{s+\ell < T \leq t+\ell\}.$$

Requiring further that $T \geq 2\ell + 1$ lets us simplify the right:

$$\ell\frac{u_i(b^i) - u_i(a_i', b_{-i}^i)}{\underline{v}_i - w_i^i} \leq \sum_{s=1}^{T-1}\sum_{t=s+1}^{s+\ell} \mathbb{1}\{s+\ell < T \leq t+\ell\}$$
$$= \sum_{s=1}^{T-1} \mathbb{1}\{s+\ell < T\} \max\{0,\ s - [T-(2\ell+1)]\}$$
$$= \sum_{s=1}^{T-\ell-1} \max\{0,\ s - [T-(2\ell+1)]\}$$
$$= \sum_{s=1+[T-(2\ell+1)]}^{T-\ell-1} (s - [T-(2\ell+1)]) = \sum_{k=1}^{\ell} k = \frac{\ell(\ell+1)}{2}.$$

Therefore, the b^i-obedience (i's stick phase) constraint is satisfied whenever $(T, \ell) \in \mathbb{N}^2$ satisfy

$$T \geq 2\ell + 1$$
$$\ell \geq \max\left\{1,\ 2\frac{u_i(a_i', b_{-i}^i) - u_i(b^i)}{w_i^i - \underline{v}_i} - 1\right\} \qquad \forall b^i \in \operatorname{supp}\beta^i,\ a_i' \in A_i,\ i \in I.$$

[II$_j$]

Next, when player $j \neq i$ is being punished in the stick phase, the designer will recommend some $b^j \in \operatorname{supp} \beta^j$. Using Lemma 1, player i's obedience constraint is

$$\sum_{s=1}^{T-1} \sum_{t=s+1}^{\min\{s+\ell,T\}} \Big[\mathbb{1}\{t=T\}u_i(b^j) + \mathbb{1}\{t<T\leq s+\ell\}u_i(\beta^j) + \mathbb{1}\{t<T>s+\ell\}w_i^j\Big]$$
$$\geq \sum_{s=1}^{T-1} \sum_{t=s+1}^{\min\{s+\ell,T\}} \Big[\mathbb{1}\{t=T\}u_i(a_i', b_{-i}^j) + \mathbb{1}\{t<T\leq t+\ell\}\underline{v}_i + \mathbb{1}\{t<T>t+\ell\}w_i^i\Big].$$

By imposing $T \geq 2\ell+1$, this rearranges (as with II$_i$) to

$$\begin{aligned}[u_i(b^j) - u_i(a_i', b_{-i}^j)]\ell &\geq \sum_{s=1}^{T-1} \sum_{t=s+1}^{\min\{s+\ell,T\}} \Big[\mathbb{1}\{t<T\leq t+\ell\}\underline{v}_i - \mathbb{1}\{t<T\leq s+\ell\}u_i(\beta^j) \\ &\qquad + \mathbb{1}\{T>t+\ell\}w_i^i - \mathbb{1}\{T>s+\ell\}w_i^j\Big] \\ &= \sum_{s=1}^{T-1} \sum_{t=s+1}^{\min\{s+\ell,T\}} \Big[\mathbb{1}\{t<T\leq s+\ell\}[\underline{v}_i - u_i(\beta^j)] + \mathbb{1}\{T>t+\ell\}(w_i^i - w_i^j) \\ &\qquad + \mathbb{1}\{s+\ell<T\leq t+\ell\}(\underline{v}_i - w_i^j)\Big].\end{aligned}$$

Since $\underline{v}_i < w_i^j$ for all $i, j \in I$, it suffices that

$$[u_i(b^j) - u_i(a_i', b_{-i}^j)]\ell \geq \sum_{s=1}^{T-1} \sum_{t=s+1}^{\min\{s+\ell,T\}} \Big[\mathbb{1}\{t<T\leq s+\ell\}[\underline{v}_i - u_i(\beta^j)] + \mathbb{1}\{T>t+\ell\}(w_i^i - w_i^j)\Big].$$

Notice that

$$\sum_{s=1}^{T-1} \sum_{t=s+1}^{\min\{s+\ell,T\}} \mathbb{1}\{t<T\leq s+\ell\} = \sum_{s=T-\ell}^{T-1} (T-s-1) = \sum_{k=1}^{\ell-1} k = \frac{\ell(\ell-1)}{2};$$
$$\sum_{s=1}^{T-1} \sum_{t=s+1}^{\min\{s+\ell,T\}} \mathbb{1}\{T>t+\ell\} = \sum_{s=1}^{T-1} \sum_{t=1}^{\ell} \mathbb{1}\{T>t+s+\ell\} = \ell(T-2\ell-1) + \frac{\ell(\ell-1)}{2}.$$

Plugging these in, the inequality becomes

$$[u_i(b^j) - u_i(a_i', b_{-i}^j)]\ell \geq \frac{\ell(\ell-1)}{2}[\underline{v}_i - u_i(\beta^j)] + \left(\ell(T-2\ell-1) + \frac{\ell(\ell-1)}{2}\right)(w_i^i - w_i^j).$$

Simplifying,

$$\begin{aligned}
2[u_i(b^j) - u_i(a_i', b_{-i}^j)] &\geq (\ell-1)[\underline{v}_i - u_i(\beta^j)] + (2T - 3\ell - 3)(w_i^i - w_i^j) \\
2\frac{u_i(b^j) - u_i(a_i', b_{-i}^j)}{w_i^i - w_i^j} - (\ell-1)\frac{\underline{v}_i - u_i(\beta^j)}{w_i^i - w_i^j} &\leq (2T - 3\ell - 3) \\
\frac{u_i(a_i', b_{-i}^j) - u_i(b^j)}{w_i^j - w_i^i} + \frac{3}{2}(\ell+1) - \frac{(\ell-1)}{2}\frac{u_i(\beta^j) - \underline{v}_i}{w_i^j - w_i^i} &\leq T
\end{aligned}$$

Therefore, the b^j-obedience constraints with $j \neq i$ are satisfied whenever

$$T \geq \max\left\{2\ell+1,\ \frac{u_i(a_i', b_{-i}^j) - u_i(b^j) - \frac{\ell-1}{2}[u_i(\beta^j) - \underline{v}_i]}{w_i^j - w_i^i} + \frac{3}{2}(\ell+1)\right\} \qquad \substack{\forall i \in I,\ j \in I\setminus\{i\},\\ a_i' \in A_i,\ b^j \in \operatorname{supp}\beta^j.}$$

[III]

Finally, we verify incentives in the carrot phase. After player $j \in I$ has been punished for ℓ stages in the stick phase, the designer will recommend some $a^j \in \operatorname{supp}\alpha^j$. Applying Lemma 1, the obedience constraint of player $i \in I$ is

$$\begin{aligned}
&\sum_{s=1}^{T-\ell-1} \sum_{t=s+\ell+1}^{T} \Big[\mathbb{1}\{t = T\} u_i(a^j) + (1 - \mathbb{1}\{t = T\}) w_i^j\Big] \\
\geq &\sum_{s=1}^{T-\ell-1} \sum_{t=s+\ell+1}^{T} \Big[\mathbb{1}\{t = T\} u_i(a_i', a_{-i}^j) + \mathbb{1}\{t < T \leq t + \ell\}\underline{v}_i + \mathbb{1}\{t < T > t + \ell\} w_i^i\Big].
\end{aligned}$$

Simplifying,

$$\begin{aligned}
&[u_i(a^j) - u_i(a_i', a_{-i}^j)] \sum_{s=1}^{T-\ell-1} \sum_{t=s+\ell+1}^{T} \mathbb{1}\{t = T\} \\
\geq &\sum_{s=1}^{T-\ell-1} \sum_{t=s+\ell+1}^{T} \Big[\mathbb{1}\{t < T \leq t + \ell\}(\underline{v}_i - w_i^j) + \mathbb{1}\{T > t + \ell\}(w_i^i - w_i^j)\Big].
\end{aligned}$$

Choosing $T \geq 2\ell + 1$,

$$[u_i(a^j) - u_i(a_i', a_{-i}^j)](T - \ell - 1) \geq \frac{\ell(2T - 3\ell - 3)}{2}(\underline{v}_i - w_i^j) + \frac{(T - 2\ell - 2)(T - 2\ell - 1)}{2}(w_i^i - w_i^j).$$

Because $w_i^i \le w_i^j$ for all $i, j \in I$, it suffices that

$$[u_i(a^j) - u_i(a_i', a_{-i}^j)](T - \ell - 1) \ge \frac{\ell(2T - 3\ell - 3)}{2}(\underline{v}_i - w_i^j)$$
$$\frac{u_i(a_i', a_{-i}^j) - u_i(a^j)}{w_i^j - \underline{v}_i}(T - \ell - 1) \le \ell(T - \ell - 1) - \frac{1}{2}\ell(\ell + 1)$$
$$\left(\ell - \frac{u_i(a_i', a_{-i}^j) - u_i(a^j)}{w_i^j - \underline{v}_i}\right)(T - \ell - 1) \ge \frac{\ell(\ell + 1)}{2}$$

The $(a^j)_{j\in I}$ obedience constraints of player i are then satisfied whenever the following inequalities hold for all $i, j \in I$, $a_i' \in A_i,\ a^j \in \operatorname{supp} \alpha^j$:

$$\ell \ge 1 + \frac{u_i(a_i', a_{-i}^j) - u_i(a^j)}{w_i^j - \underline{v}_i}$$
$$T \ge \max\left\{ 2\ell + 1,\ 1 + \ell + \frac{\ell(\ell + 1)}{2}\left(\ell - \frac{u_i(a_i', a_{-i}^j) - u_i(a^j)}{w_i^j - \underline{v}_i}\right)^{-1} \right\}.$$

The bound on ℓ (and therefore on T) is finite since the game G is finite.

The constraints on $\ell \in \mathbb{N}$ in cases I, II, and III are all finite lower bounds. Fixing one compatible $\ell \in \mathbb{N}$, the constraints on $T \in \mathbb{N}$ are all finite lower bounds as well. There are finitely many of them. Therefore there exists a pair $(T, \ell) \in \mathbb{N}^2$ satisfying all obedience constraints simultaneously in the final-stage payoff setting. $\square$

Proof of Theorem 1 (Part 2/2: average flow payoff extension).

Let us now verify obedience under the average flow payoff extension.

[I]

We begin with the good phase. Supposing that players $I \setminus \{i\}$ always obey, player i's payoff from obeying $a \in \operatorname{supp} \alpha$ is

$$\frac{1}{T}\sum_{t=1}^{T} \mu_i(t|a; \mathrm{I})\Big(u_i(a) + (T - t)\mathbb{E}_{a\sim\alpha}[u_i(a)]\Big).$$

Because (i) every stage is always reached on path and (ii) on-path recommendations are i.i.d., players must have uniform beliefs $\mu_i(\cdot|a) = \frac{1}{T}$, for every $a \in \operatorname{supp}\alpha$. The a-obedience constraint for player i is therefore

$$\sum_{t=1}^{T}\big(u_i(a)+(T-t)v_i\big) \geq \sum_{t=1}^{T-\ell-1}\Big[u_i(a_i',a_{-i})+\ell\underline{v}_i+(T-\ell-t)w_i^i\Big]+\sum_{t=T-\ell}^{T}\Big[u_i(a_i',a_{-i})+(T-t)\underline{v}_i\Big].$$

Re-arranging:

$$\sum_{t=1}^{T-\ell-1}\Big[u_i(a)-u_i(a_i',a_{-i})+(T-t)v_i-\ell\underline{v}_i-(T-\ell-t)w_i^i\Big] \geq -\sum_{t=T-\ell}^{T}\Big[u_i(a)-u_i(a_i',a_{-i})+(T-t)(v_i-\underline{v}_i)\Big]$$

$$\sum_{t=1}^{T-\ell-1}\Big[u_i(a)-u_i(a_i',a_{-i})+\ell(v_i-\underline{v}_i)+(T-\ell-t)(v_i-w_i^i)\Big] \geq \sum_{t=T-\ell}^{T}\Big[u_i(a_i',a_{-i})-u_i(a)-(T-t)(v_i-\underline{v}_i)\Big].$$

Because $v_i > w_i^i$, a sufficient condition for this inequality to hold is that

$$\sum_{t=1}^{T-\ell-1}\Big[-[u_i(a_i',a_{-i})-u_i(a)]+\ell(v_i-\underline{v}_i)\Big] \geq \sum_{t=T-\ell}^{T}\Big[u_i(a_i',a_{-i})-u_i(a)-(T-t)(v_i-\underline{v}_i)\Big].$$

The stage game is finite, so define

$$M := \max_{\substack{i\in I\\ a_i'\in A_i\\ a\in\operatorname{supp}\alpha}} |u_i(a_i',a_{-i})-u_i(a)| \in [0,\infty) \qquad \text{and} \qquad d := \min_{i\in I}(v_i-\underline{v}_i) \in (0,\infty).$$

It then also suffices that

$$\sum_{t=1}^{T-\ell-1}[-M+\ell d] \geq \sum_{t=T-\ell}^{T}\Big[\underbrace{u_i(a_i',a_{-i})-u_i(a)}_{\leq M}+(T-t)\underbrace{(-1)(v_i-\underline{v}_i)}_{\leq 0}\Big]$$

and hence that

$$(T-\ell-1)(\ell d-M) \geq (T-[T-\ell]+1)M.$$

Therefore, the a-obedience (good phase) constraint is satisfied whenever

$$\ell > \frac{M}{d} \qquad \text{and} \qquad T \geq \frac{\ell(\ell+1)d}{\ell d - M}.$$

[II]

We now verify obedience in the stick phase. Let $s \in \{1, ..., T-1\}$ denote the time of the most recent deviation, and consider the player $j \in I$ who deviated at s. (Recall that whenever there are multiple deviators, the players are considered to be obedient.) Player $i \in I$ has two sets of incentive constraints when j is being punished: one where $i = j$ and one where $i \neq j$. We check these in turn.

[II$_i$]

When player i is being punished in the stick phase, the designer will recommend some $b^i \in \operatorname{supp} \beta^i$. Player i's obedience constraint is

$$\begin{aligned}
&\sum_{s=1}^{T-1} \sum_{t=s+1}^{\min\{s+\ell,T\}} \mu_i(s,t|b^i;\mathrm{II}_i)\Big[u_i(b^i) + (\min\{T, s+\ell\} - t)\underline{v}_i + \max\{0, T-(s+\ell)\}w_i^i\Big] \\
\geq &\sum_{s=1}^{T-1} \sum_{t=s+1}^{\min\{s+\ell,T\}} \mu_i(s,t|b^i;\mathrm{II}_i)\Big[u_i(a_i', b_{-i}^i) + \min\{\ell, T-t\}\underline{v}_i + \max\{0, T-(t+\ell)\}w_i^i\Big].
\end{aligned}$$

By Lemma 1, off-path beliefs are uniform over $S = \{(s,t) \in \mathbb{N}^2 : 1 \leq s < t \leq T,\ t - s \leq \ell\}$. Beliefs then cancel. Normalise $w_i^i = 0$, noting that because $w^i \in V$, this implies that $\underline{v}_i < 0$. Re-arranging:

$$\sum_{s=1}^{T-1} \sum_{t=s+1}^{\min\{s+\ell,T\}} \Big[u_i(b^i) - u_i(a_i', b_{-i}^i) + [(\min\{T, s+\ell\} - t) - \min\{\ell, T-t\}]\underline{v}_i\Big] \geq 0$$

$$\sum_{s=1}^{T-1} \sum_{t=s+1}^{\min\{s+\ell,T\}} \Big[-[u_i(a_i', b_{-i}^i) - u_i(b^i)] - \underline{v}_i(\min\{T-t, \ell\} - \min\{T-t, s+\ell-t\})\Big] \geq 0$$

Define

$$\tilde{M} := \max_{\substack{i\in I\\ a_i'\in A_i\\ b^i\in \operatorname{supp}\beta^i}} |u_i(a_i', b_{-i}^i) - u_i(b^i)| \in [0,\infty) \qquad \text{and} \qquad \tilde{d} := \min_{i\in I}(0 - \underline{v}_i) \in (0,\infty).$$

The inequality then holds whenever

$$\begin{aligned}
\sum_{s=1}^{T-1}\sum_{t=s+1}^{\min\{s+\ell,T\}} \Big[-\tilde{M} + \tilde{d}(\min\{T-t,\ell\} - \min\{T-t, s+\ell-t\})\Big] &\geq 0\\
\tilde{d}\sum_{s=1}^{T-1}\sum_{t=s+1}^{\min\{s+\ell,T\}} \min\{\max\{0, T-s-\ell\}, t-s\} &\geq \sum_{s=1}^{T-1}\sum_{t=s+1}^{\min\{s+\ell,T\}} \tilde{M}\\
\tilde{d}\sum_{s=1}^{T-\ell-1}\sum_{t=s+1}^{\min\{s+\ell,T\}} \min\{T-\ell-s, t-s\} &\geq \tilde{M}\sum_{s=1}^{T-1}\min\{\ell, T-s\}\\
\sum_{s=1}^{T-\ell-1}\sum_{k=1}^{\ell}\min\{s,k\} &\geq \frac{\tilde{M}}{\tilde{d}}\frac{1}{2}\ell(2T-\ell-1).
\end{aligned}$$

Choosing $T > 2\ell$, we can simplify the left side:

$$\begin{aligned}
\text{LHS} &= \sum_{j=1}^{T-\ell-1}\sum_{k=1}^{\ell}\min\{j,k\}\\
&= \sum_{j=1}^{\ell}\sum_{k=1}^{\ell}\min\{j,k\} + \sum_{j=\ell+1}^{T-\ell-1}\sum_{k=1}^{\ell}\min\{j,k\}\\
&= \sum_{j=1}^{\ell}\left(\sum_{k=1}^{j} k + \sum_{k=j+1}^{\ell} j\right) + \sum_{j=\ell+1}^{T-\ell-1}\sum_{k=1}^{\ell} k\\
&= \sum_{j=1}^{\ell}\left(\frac{j(j+1)}{2} + j(\ell-j)\right) + (T-\ell-1-\ell)\frac{\ell(\ell+1)}{2}\\
&= \frac{1}{2}\left((2\ell+1)\frac{\ell(\ell+1)}{2} - \frac{\ell(\ell+1)(2\ell+1)}{6} + (T-2\ell-1)\ell(\ell+1)\right)\\
&= \frac{1}{6}\ell(\ell+1)(3T-4\ell-2).
\end{aligned}$$

Our sufficient condition thereby becomes

$$\frac{1}{6}\ell(\ell+1)(3T-4\ell-2) \geq \frac{\tilde{M}}{\tilde{d}}\frac{1}{2}\ell(2T-\ell-1)$$
$$T\left(\ell-2\frac{\tilde{M}}{\tilde{d}}+1\right) \geq \frac{\ell+1}{3}\left(4\ell-3\frac{\tilde{M}}{\tilde{d}}+2\right).$$

Therefore, the b^i-obedience (i's stick phase) constraint is satisfied whenever

$$\ell > 2\frac{\tilde{M}}{\tilde{d}} \qquad \text{and} \qquad T \geq \max\left\{2\ell+1,\ \frac{(4\ell+2)\tilde{d}-3\tilde{M}}{(\ell+1)\tilde{d}-2\tilde{M}}\frac{\ell+1}{3}\right\}.$$

[II$_j$]

Next, when player $j \neq i$ is being punished in the stick phase, the designer will recommend some $b^j \in \operatorname{supp}\beta^j$. Player i's obedience constraint is

$$\sum_{s=1}^{T-1}\sum_{t=s+1}^{\min\{s+\ell,T\}} \mu_i(s,t|b^j;\mathrm{II}_j)\Big[u_i(b^j)+(\min\{T,s+\ell\}-t)u_i(\beta^j)+\max\{0,T-(s+\ell)\}w_i^j\Big]$$
$$\geq \sum_{s=1}^{T-1}\sum_{t=s+1}^{\min\{s+\ell,T\}} \mu_i(s,t|b^j;\mathrm{II}_j)\Big[u_i(a_i',b_{-i}^j)+\min\{\ell,T-t\}\underline{v}_i+\max\{0,T-(t+\ell)\}w_i^i\Big].$$

Lemma 1 again lets us impose uniform off-path beliefs over $S := \{(s,t)\in\mathbb{N}^2 : 1 \leq s < t \leq T,\ t-s \leq \ell\}$, so beliefs cancel. Also normalise $w_i^i = 0$. Re-arranging:

$$\sum_{s=1}^{T-1}\sum_{t=s+1}^{\min\{s+\ell,T\}} \Big[u_i(b^j)-u_i(a_i',b_{-i}^j)+(\min\{T,s+\ell\}-t)u_i(\beta^j)-\min\{\ell,T-t\}\underline{v}_i$$
$$+\max\{0,T-(s+\ell)\}w_i^j\Big] \geq 0$$

$$\sum_{s=1}^{T-1}\sum_{t=s+1}^{\min\{s+\ell,T\}} \Big[-[u_i(a_i',b_{-i}^j)-u_i(b^j)]-\underline{v}_i\big(\min\{\ell,T-t\}-\min\{T-t,s+\ell-t\}\big)\Big]$$
$$\geq \sum_{s=1}^{T-1}\sum_{t=s+1}^{\min\{s+\ell,T\}} \Big[(\underline{v}_i-u_i(\beta^j))\min\{T-t,s+\ell-t\}-w_i^j\max\{0,T-(s+\ell)\}\Big]$$

Define

$$\begin{aligned}
\hat{M} &:= \max_{\substack{i\in I\\ j\in I\setminus\{i\}\\ a_i'\in A_i\\ b^j\in\operatorname{supp}\beta^j}} |u_i(a_i', b_{-i}^j) - u_i(b^j)| \in [0,\infty)\\
\hat{d} &:= \min_{i\in I}(0 - \underline{v}_i) \in (0,\infty)\\
\hat{B} &:= \max_{\substack{i\in I\\ j\in I\setminus\{i\}}} (\underline{v}_i - u_i(\beta^j))\\
\hat{W} &:= \min_{\substack{i\in I\\ j\in I\setminus\{i\}}} w_i^j \in (0,\infty).
\end{aligned}$$

Noting that $w_i^i = 0$ implies both $w_i^j > 0$ and $\underline{v}_i < 0$, the inequality then holds whenever

$$\begin{aligned}
&\sum_{s=1}^{T-1}\sum_{t=s+1}^{\min\{s+\ell,T\}} \Big[-\hat{M} + \hat{d}\Big(\min\{\ell, T-t\} - \min\{T-t, s+\ell-t\}\Big)\Big]\\
&\qquad \geq \hat{B}\underbrace{\sum_{s=1}^{T-1}\sum_{t=s+1}^{\min\{s+\ell,T\}} \min\{T-t, s+\ell-t\}}_{=\frac{1}{6}\ell(\ell-1)(3T-2\ell-2)} - \hat{W}\underbrace{\sum_{s=1}^{T-1}\sum_{t=s+1}^{\min\{s+\ell,T\}} \max\{0, T-(s+\ell)\}}_{=\frac{1}{2}\ell(T-\ell-1)(T-\ell)}.
\end{aligned}$$

$$\begin{aligned}
&\sum_{s=1}^{T-1}\sum_{t=s+1}^{\min\{s+\ell,T\}} \Big[\min\{\ell, T-t\} - \min\{T-t, s+\ell-t\}\Big]\\
&\geq \frac{1}{\hat{d}}\left(\frac{\hat{B}}{6}\ell(\ell-1)(3T-2\ell-2) - \frac{\hat{W}}{2}\ell(T-\ell-1)(T-\ell) + \sum_{s=1}^{T-1}\sum_{t=s+1}^{\min\{s+\ell,T\}}\hat{M}\right).
\end{aligned}$$

Following the derivation in [II$_i$],

$$\underbrace{\sum_{s=1}^{T-\ell-1}\sum_{k=1}^{\ell}\min\{s,k\}}_{=\frac{1}{6}\ell(\ell+1)(3T-4\ell-2)\text{ if }T>2\ell} \geq \frac{1}{\hat{d}}\left(\frac{\hat{B}}{6}\ell(\ell-1)(3T-2\ell-2) - \frac{\hat{W}}{2}\ell(T-\ell-1)(T-\ell) + \hat{M}\frac{\ell}{2}(2T-\ell-1)\right).$$

So, by choosing $T > 2\ell$ we obtain

$$\frac{1}{6}\ell(\ell+1)(3T-4\ell-2)-\frac{\hat{M}}{\hat{d}}\frac{1}{2}\ell(2T-\ell-1) \geq \frac{1}{\hat{d}}\left(\frac{\hat{B}}{6}\ell(\ell-1)(3T-2\ell-2)-\frac{\hat{W}}{2}\ell(T-\ell-1)(T-\ell)\right).$$

If $\hat{B} \leq 0$ then it's immediate from [II$_i$] that $\ell > 2\hat{M}/\hat{d}$ and a large enough choice of T will suffice. So, suppose that $\hat{B} > 0$. Select T and ℓ to satisfy the conditions in [II$_i$]. It's then sufficient to show that under these conditions, we can make

$$\frac{\hat{B}}{6}\ell(\ell-1)(3T-2\ell-2) \leq \frac{\hat{W}}{2}\ell(T-\ell-1)(T-\ell).$$

By choosing $T \geq 2\ell+2$, we have

$$\text{RHS} \geq \frac{\hat{W}}{2}\ell\frac{T}{2}\frac{T}{2} \quad \text{and} \quad \text{LHS} \leq \frac{\hat{B}}{2}\ell(\ell-1)T$$

so it suffices that

$$4\hat{B}\ell(\ell-1)T \leq \hat{W}\ell T^2$$
$$T \geq 4\frac{\hat{B}}{\hat{W}}(\ell-1).$$

Therefore, the b^j-obedience constraint is satisfied whenever

$$\ell > 2\frac{\hat{M}}{\hat{d}} \qquad \text{and} \qquad T \geq \max\left\{2\ell+2,\ \frac{(4\ell+2)\hat{d}-3\hat{M}}{(\ell+1)\hat{d}-2\hat{M}}\frac{\ell+1}{3},\ 4\frac{\hat{B}}{\hat{W}}(\ell-1)\right\}.$$

[III]

Finally, we verify incentives in the carrot phase.

After player $j \in I$ has been punished for ℓ stages in the stick phase, the designer will recommend some $a^j \in \operatorname{supp} \alpha^j$. The obedience constraint of player $i \in I$ is

$$
\begin{aligned}
&\sum_{s=1}^{T-\ell-1} \sum_{t=s+\ell+1}^{T} \mu_i(s,t|a^j;\mathrm{III}_j)\Big[u_i(a^j) + (T-t)w_i^j\Big] \\
\geq &\sum_{s=1}^{T-\ell-1} \sum_{t=s+\ell+1}^{T} \mu_i(s,t|a^j;\mathrm{III}_j)\Big[u_i(a_i', a_{-i}^j) + \min\{\ell, T-t\}\underline{v}_i + \max\{0, T-(t+\ell)\}w_i^i\Big].
\end{aligned}
$$

Once more, Lemma 1 lets us specify uniform beliefs over $C = \{(s,t) \in \mathbb{N}^2 : 1 \leq s < s+\ell < t \leq T\}$. Normalise $w_i^i = 0$ so that $\underline{v}_i < 0$, and $w_i^j \geq 0$ (strictly iff $i \neq j$). We then require

$$
\sum_{s=1}^{T-\ell-1} \sum_{t=s+\ell+1}^{T} \Big[-[u_i(a_i', a_{-i}^j) - u_i(a^j)] + (T-t)w_i^j - \min\{\ell, T-t\}\underline{v}_i\Big] \geq 0.
$$

Define

$$
\bar{M} := \max_{\substack{i,j \in I \\ a_i' \in A_i \\ a^j \in \operatorname{supp} \alpha^j}} |u_i(a_i', a_{-i}^j) - u_i(a^j)| \in [0, \infty) \qquad \text{and} \qquad \bar{d} := \min_{i \in I}(0 - \underline{v}_i) \in (0, \infty).
$$

It then suffices to make

$$
\begin{aligned}
\sum_{s=1}^{T-\ell-1} \sum_{t=s+\ell+1}^{T} \Big[-\bar{M} + \bar{d}\min\{\ell, T-t\}\Big] &\geq 0 \\
\underbrace{\sum_{s=1}^{T-\ell-1} \sum_{t=s+\ell+1}^{T} \min\{\ell, T-t\}}_{\geq \sum_{s=1}^{T-2\ell-1} \sum_{t=s+\ell+1}^{T-\ell} \ell} &\geq \underbrace{\frac{1}{\bar{d}} \sum_{s=1}^{T-\ell-1} \sum_{t=s+\ell+1}^{T} \bar{M}}_{= \frac{\bar{M}}{\bar{d}} \frac{1}{2}(T-\ell-1)(T-\ell)}.
\end{aligned}
$$

or simply

$$
\begin{aligned}
\sum_{s=1}^{T-2\ell-1} \sum_{t=s+\ell+1}^{T-\ell} \ell &\geq \frac{\bar{M}}{\bar{d}} \frac{1}{2}(T-\ell-1)(T-\ell) \\
\frac{\ell}{2}(T-2\ell-1)(T-2\ell) &\geq \frac{\bar{M}}{\bar{d}} \frac{1}{2}(T-\ell-1)(T-\ell).
\end{aligned}
$$

Whenever $T \geq 4\ell + 2$, we get $T - 2\ell - 1 \geq \frac{T}{2}$, so we only need

$$\frac{\ell}{2}\frac{T}{2}\frac{T}{2} \geq \frac{\bar{M}}{\bar{d}}\frac{1}{2}\underbrace{(T-\ell-1)(T-\ell)}_{\leq T^2}$$

$$\ell \geq 4\frac{\bar{M}}{\bar{d}}.$$

Notice that we have thereby covered both III_i and III_j.

Since G is finite, constraints on $\ell \in \mathbb{N}$ are all finite lower bounds. Fixing one compatible $\ell \in \mathbb{N}$, the constraints on $T \in \mathbb{N}$ are all finite lower bounds as well. There are finitely many of them, so there exists a pair $(T, \ell) \in \mathbb{N}^2$ satisfying all obedience constraints simultaneously in the average flow payoff setting. $\square$

A.2 Proof of Proposition 1

Fix a finite stage game $G = (\{1,2\}, (A_i, u_i)_{i\in\{1,2\}})$. Let $v \in \mathbb{R}^{|I|}$ denote some feasible and strictly individually rational payoff profile. Since v is feasible, it can be achieved by some (correlated) mixed action profile $\alpha \in \Delta(A)$, where $A := \prod_{i\in I} A_i$. Fix this profile.

Consider the information structure (M, π) described as follows. The message space is $M = \{(a;r), (a;p) : a \in A\}$. The possible paths of play in G_T are partitioned into 'reward' and 'punishment' phases. The null history is in a reward phase. Whenever the game is in a reward phase, the designer (independently) draws $a \sim \alpha$ and announces $(a;r)$. This is interpreted as a recommendation to play a during a reward phase.

If and only if a history ends in exactly one player deviating from the designer's latest recommendation does the game (re-)enter a punishment phase. Whenever the game is in a punishment phase, the designer (independently) draws $b \sim \beta = (\beta_1, \beta_2) \in \Delta(A_1) \times \Delta(A_2)$ where

$$\beta_i \in \operatorname*{arg\,min}_{\tilde{\alpha}_i \in \Delta(A_i)} \max_{\tilde{a}_{-i} \in A_{-i}} u_{-i}(\tilde{a}_{-i}, \tilde{\alpha}_i)$$

for both $i \in \{1,2\}$; and announces $(b;p)$. If the game has not ended after ℓ stages of obedient play during a punishment phase, the game returns to a reward phase.

We begin by specifying beliefs in the punishment phase assuming obedience holds.

Lemma 2 (Off-path beliefs)**.** The following beliefs for each player $i \in \{1, 2\}$ are consistent with sMCE under an obedient strategy profile for information structure (M, π):

$$\mu_i(s, t|b; p) = \frac{1}{|S|}, \qquad S := \{(s, t) \in \mathbb{N}^2 : 1 \leq s < t \leq T,\ t \leq \ell + s\}$$

for any hypothesis (s, t) about the current stage $t \in \{1, ..., T\}$ and most recent deviation $s \in \{1, ..., T - 1\}$, and any punishment-phase message $(b; p) \in M$ with $b \in \operatorname{supp} \beta$.

Proof. We construct a limit of fully mixed strategies. Write $\sigma = (\sigma_1, \sigma_2)$, where $\sigma_i : M \to \Delta(A_i)$, for the conjectured obedient (behavioural) strategy profile. Let $\hat{\alpha}_i :=$ Unif A_i and $\varepsilon_k := \frac{1}{k+2}$ and $\sigma_i^k := (1 - \varepsilon_k)\sigma_i + \varepsilon_k \hat{\alpha}_i$, for $i \in \{1, 2\}$ and $k \in \mathbb{N}$. Write μ_i^k for the resulting (self-locating) beliefs. Notice that $\lim_{k\uparrow\infty} \sigma^k = \sigma$.

Under (M, π), the message $(b; p)$ is announced at t if and only if $t \geq 2$ and exactly one player most recently disobeyed, at $s \in \{\max\{1, t - \ell\}, ..., t - 1\}$. Therefore $\mu_i^k(s, 1|b; p) = 0$ for all $k, s \in \mathbb{N}$. The probability that exactly one player disobeys at any particular $t \in \{1, ..., T\}$ given profile σ^k is

$$\sum_{i\in\{1,2\}} \varepsilon_k \left(1 - \frac{1}{|A_i|}\right)\left([1 - \varepsilon_k] + \varepsilon_k \frac{1}{|A_{-i}|}\right) \ =: \ x_k.$$

The chance that the most recent deviation was at $s \in \{\max\{1, t - \ell\}, ..., t - 1\}$ given that $(b; p)$ is announced at stage $t \in \{2, 3, ..., T\}$ is proportional to

$$x_k(1 - x_k)^{t-s-1}.$$

The posterior on $(s, t) \in S$ given just $(b; p)$ is therefore

$$\begin{aligned}
\mu_i^k(s, t|b; p) &= \frac{\mu_i^k(s, t)}{\sum_{(s',t')\in S} \mu_i^k(s', t')} \\
&= \frac{\mu_i^k(s|t)\mu_i^k(t)}{\sum_{(s',t')\in S} \mu_i^k(s'|t')\mu_i^k(t')} = \frac{x_k(1 - x_k)^{t-s-1}/T}{\sum_{(s',t')\in S} x_k(1 - x_k)^{t'-s'-1}/T} \overset{k\uparrow\infty}{\longrightarrow} \frac{1}{|S|}
\end{aligned}$$

as needed. $\square$

We will handle the two admissible payoff extensions separately.

Proof of Proposition 1 (Part 1/2: final stage payoff extension).

Payoffs will be $\mathbb{E}_{a_t \sim \alpha}[\tilde{u}(a_1, ..., a_T)] = \mathbb{E}_{a_T \sim \alpha}[u(a_T)] = v$, as claimed, if players are obedient under (M, π). Each player has $|\operatorname{supp}\alpha| + |\operatorname{supp}\beta|$ sets of incentive constraints, one for each information set generated by the information structure. It suffices for sMCE to verify each of these separately (rather than invoking the one-shot deviation principle).

Supposing that players $I \setminus \{i\}$ always obey, player i's payoff from obeying $a \in \operatorname{supp}\alpha$ is

$$\sum_{t=1}^{T} \mu_i(t|a)\Big(\mathbb{1}\{t = T\}u_i(a) + \mathbb{1}\{t < T\}\mathbb{E}_{a\sim\alpha}[u_i(a)]\Big).$$

Choose $T \geq \ell + 2$. Because every stage is always reached on path and on-path recommendations are i.i.d., players must have uniform beliefs $\mu_i(\cdot|a) = \frac{1}{T}$, for every $a \in \operatorname{supp}\alpha$. The a-obedience constraint for player i is therefore

$$\begin{aligned} u_i(a) + (T-1)v_i &\geq \sum_{t=1}^{T}\Big(\mathbb{1}\{t = T\}u_i(a_i', a_{-i}) + \mathbb{1}\{T - \ell \leq t < T\}u_i(\beta) + \mathbb{1}\{t < T - \ell\}v_i\Big) \\ &= u_i(a_i', a_{-i}) + \ell u_i(\beta) + (T - \ell - 1)v_i \end{aligned}$$

which re-arranges to a reward-phase obedience condition:

$$\begin{aligned} &T \geq \ell + 2 \\ &\ell \geq \frac{u_i(a_i', a_{-i}) - u_i(a)}{v_i - u_i(\beta)} \qquad \forall i \in \{1, 2\},\ a_i' \in A_i,\ a \in \operatorname{supp}\alpha. \end{aligned}$$

We now verify incentives in the punishment phase. Let $s \in \{1, ..., T-1\}$ denote the time of the most recent deviation. For $b \in \operatorname{supp}\beta$, the b-obedience constraint is

$$\begin{aligned} &\sum_{s=1}^{T-1}\sum_{t=s+1}^{\min\{s+\ell,T\}} \mu_i(s,t|b)\Big(\mathbb{1}\{t = T\}u_i(b) + \mathbb{1}\{t < T \leq s + \ell\}u_i(\beta) + \mathbb{1}\{t < T > s + \ell\}v_i\Big) \\ \geq &\sum_{s=1}^{T-1}\sum_{t=s+1}^{\min\{s+\ell,T\}} \mu_i(s,t|b)\Big(\mathbb{1}\{t = T\}u_i(a_i', b_{-i}) + \mathbb{1}\{t < T \leq t + \ell\}u_i(\beta) + \mathbb{1}\{t < T > t + \ell\}v_i\Big). \end{aligned}$$

By Lemma 2, off-path beliefs may be set to be uniform over $S = \{(s,t) \in \mathbb{N}^2 : 1 \leq s < t \leq T,\ t \leq s+\ell\}$. Beliefs then cancel and we can simplify:

$$u_i(\beta)\sum_{s=1}^{T-1}\sum_{t=s+1}^{\min\{s+\ell,T\}} [\mathbb{1}\{t<T\leq s+\ell\} - \mathbb{1}\{t<T\leq t+\ell\}]$$
$$\geq [u_i(a_i', b_{-i}) - u_i(b)]\sum_{s=1}^{T-1}\sum_{t=s+1}^{\min\{s+\ell,T\}} \mathbb{1}\{t=T\} + v_i \sum_{s=1}^{T-1}\sum_{t=s+1}^{\min\{s+\ell,T\}} [\mathbb{1}\{t<T>t+\ell\} - \mathbb{1}\{t<T>s+\ell\}]$$

$$\sum_{s=1}^{T-1}\sum_{t=s+1}^{\min\{s+\ell,T\}} \mathbb{1}\{t<T\}\Big(-\mathbb{1}\{s+\ell<T\leq t+\ell\}u_i(\beta) + \mathbb{1}\{s+\ell<T\leq t+\ell\}v_i\Big)$$
$$\geq \sum_{s=1}^{T-1}\sum_{t=s+1}^{\min\{s+\ell,T\}} \mathbb{1}\{t=T\}[u_i(a_i', b_{-i}) - u_i(b)]$$

$$[v_i - u_i(\beta)]\sum_{s=1}^{T-1}\sum_{t=s+1}^{\min\{s+\ell,T\}} \mathbb{1}\{t<T\}\mathbb{1}\{s+\ell<T\leq t+\ell\} \geq [u_i(a_i', b_{-i}) - u_i(b)]\underbrace{\sum_{s=1}^{T-1}\sum_{t=s+1}^{\min\{s+\ell,T\}} \mathbb{1}\{t=T\}}_{=\ell \text{ if } T\geq \ell+1}.$$

By choosing $T \geq 2\ell+1$, we can simplify further:

$$\begin{aligned}
\frac{u_i(a_i', b_{-i}) - u_i(b)}{v_i - u_i(\beta)}\ell &\leq \sum_{s=1}^{T-1}\sum_{t=s+1}^{\min\{s+\ell,T\}} \mathbb{1}\{t<T\}\mathbb{1}\{s+\ell<T\leq t+\ell\} \\
&= \sum_{s=1}^{T-1}\sum_{t=s+1}^{s+\ell} \mathbb{1}\{s+\ell<T\leq t+\ell\} \\
&= \sum_{s=1}^{T-1} \mathbb{1}\{s+\ell<T\}\max\{0,\ s-[T-(2\ell+1)]\} \\
&= \sum_{s=1}^{T-\ell-1} \max\{0,\ s-[T-(2\ell+1)]\} \\
&= \sum_{s=1+[T-(2\ell+1)]}^{T-\ell-1} (s-[T-(2\ell+1)]) = \sum_{k=1}^{\ell} k = \frac{\ell(\ell+1)}{2}.
\end{aligned}$$

Therefore the b-obedience constraints are satisfied whenever

$$\begin{aligned} &T \geq 2\ell + 1 \\ &\ell \geq 2\frac{u_i(a_i', b_{-i}) - u_i(b)}{v_i - u_i(\beta)} - 1 \qquad \forall i \in \{1,2\},\ a_i' \in A_i,\ b \in \operatorname{supp}\beta. \end{aligned}$$

Because G is finite, there are finitely many constraints in the reward and punishment phases, each imposing finite lower bounds on T and ℓ. Thus there is a pair $(T, \ell) \in \mathbb{N}^2$ which enforces obedience. $\square$

Proof of Proposition 1 (Part 2/2: average flow payoff extension).

If players are obedient under (M, π), their payoffs will be $\frac{1}{T}\sum_{t=1}^{T} \mathbb{E}_{a\sim\alpha}[u(a)] = v$ as claimed. Each player has $|\operatorname{supp}\alpha| + |\operatorname{supp}\beta|$ sets of incentive constraints, one for each information set generated by the information structure. It suffices for sMCE to verify each of these separately (rather than invoking the one-shot deviation principle).

Let us first verify the on-path incentives. Supposing that players $I \setminus \{i\}$ always obey, player i's payoff from obeying $a \in \operatorname{supp}\alpha$ is

$$\frac{1}{T}\sum_{t=1}^{T} \mu_i(t|a)\Big(u_i(a) + (T-t)\mathbb{E}_{a\sim\alpha}[u_i(a)]\Big).$$

Because every stage is always reached on path and on-path recommendations are i.i.d., players must have uniform beliefs $\mu_i(\cdot|a) = \frac{1}{T}$, for every $a \in \operatorname{supp}\alpha$. The a-obedience constraint for player i is therefore

$$\sum_{t=1}^{T}\big(u_i(a)+(T-t)v_i\big) \geq \sum_{t=1}^{T-\ell-1}\Big[u_i(a_i', a_{-i})+\ell u_i(\beta)+(T-\ell-t)v_i\Big]+\sum_{t=T-\ell}^{T}\Big[u_i(a_i', a_{-i})+(T-t)u_i(\beta)\Big].$$

This re-arranges to

$$T \geq \frac{v_i - u_i(\beta)}{\ell[v_i - u_i(\beta)] - [u_i(a_i', a_{-i}) - u_i(a)]}\frac{\ell(\ell+1)}{2}$$

provided $T \in \mathbb{N}$ and $\ell < T/2$ are chosen sufficiently large.

We now verify incentives in the punishment phase. Let $s \in \{1, ..., T-1\}$ denote the time of the most recent deviation. For $b \in \operatorname{supp} \beta$, the b-obedience constraint is

$$\sum_{s=1}^{T-1} \sum_{t=s+1}^{\min\{s+\ell,T\}} \mu_i(s,t|b)\Big[u_i(b) + (\min\{T, s+\ell\} - t)u_i(\beta) + \max\{0, T-(s+\ell)\}v_i\Big]$$
$$\geq \sum_{s=1}^{T-1} \sum_{t=s+1}^{\min\{s+\ell,T\}} \mu_i(s,t|b)\Big[u_i(a_i', b_{-i}) + \min\{\ell, T-t\}u_i(\beta) + \max\{0, T-(t+\ell)\}v_i\Big].$$

By Lemma 2, off-path beliefs may be set to be uniform over $S = \{(s,t) \in \mathbb{N}^2 : 1 \leq s < t \leq T,\ t \leq s+\ell\}$. Beliefs then cancel. By also normalising $u_i(\beta) = 0$ (noting that this makes $v_i > 0$), we can re-arrange as follows.

$$\sum_{s=1}^{T-1} \sum_{t=s+1}^{\min\{s+\ell,T\}} v_i\Big[\max\{0, T-(s+\ell)\} - \max\{0, T-(t+\ell)\}\Big] \geq \sum_{s=1}^{T-1} \sum_{t=s+1}^{\min\{s+\ell,T\}} [u_i(a_i', b_{-i}) - u_i(b)]$$
$$v_i \sum_{s=1}^{T-\ell-1} \sum_{t=s+1}^{s+\ell} \Big[\max\{0, T-(s+\ell)\} - \max\{0, T-(t+\ell)\}\Big] \geq [u_i(a_i', b_{-i}) - u_i(b)]\frac{\ell(2T-\ell-1)}{2}.$$

Choose $T \geq 2\ell + 1$ so that the left side simplifies:

$$\begin{aligned}
\text{LHS} &= v_i \sum_{s=1}^{T-\ell-1} \sum_{t=s+1}^{s+\ell} \Big[T - s - \ell - \max\{0, T - (t+\ell)\}\Big] \\
&= v_i \sum_{j=1}^{T-\ell-1} \sum_{k=1}^{\ell} \Big[j - \max\{0, j-k\}\Big] = v_i \sum_{j=1}^{T-\ell-1} \sum_{k=1}^{\ell} \min\{j,k\} \\
&= v_i \sum_{j=1}^{\ell} \sum_{k=1}^{\ell} \min\{j,k\} + v_i \sum_{j=\ell+1}^{T-\ell-1} \sum_{k=1}^{\ell} \min\{j,k\} \\
&= v_i \sum_{j=1}^{\ell} \left(\sum_{k=1}^{j} k + \sum_{k=j+1}^{\ell} j \right) + v_i \sum_{j=\ell+1}^{T-\ell-1} \sum_{k=1}^{\ell} k \\
&= v_i \sum_{j=1}^{\ell} \left(\frac{j(j+1)}{2} + j(\ell - j) \right) + v_i(T - \ell - 1 - \ell)\frac{\ell(\ell+1)}{2} \\
&= \frac{v_i}{2} \left((2\ell+1)\frac{\ell(\ell+1)}{2} - \frac{\ell(\ell+1)(2\ell+1)}{6} + (T - 2\ell - 1)\ell(\ell+1) \right) \\
&= \frac{v_i}{6} \ell(\ell+1)(3T - 4\ell - 2).
\end{aligned}$$

Therefore the b-obedience inequality is

$$\begin{gathered}
\frac{v_i}{6}\ell(\ell+1)(3T - 4\ell - 2) \geq [u_i(a_i', b_{-i}) - u_i(b)]\frac{\ell(2T - \ell - 1)}{2} \\
T\Big((\ell+1)v_i - 2[u_i(a_i', b_{-i}) - u_i(b)]\Big) \geq (\ell+1)\Big(v_i(4\ell+2) - 3[u_i(a_i', b_{-i}) - u_i(b)]\Big)\frac{1}{3}.
\end{gathered}$$

For ℓ sufficiently large, we thus require

$$T \geq \frac{(4\ell+2)v_i - 3[u_i(a_i', b_{-i}) - u_i(b)]}{(\ell+1)v_i - 2[u_i(a_i', b_{-i}) - u_i(b)]} \frac{\ell+1}{3}.$$

Applying the $u_i(\beta) = 0$ normalisation to both of the inequalities just derived, the joint obedience constraint, for $\ell < T/2$ sufficiently large, is

$$T \geq (\ell+1) \max_{\substack{i \in \{1,2\} \\ a \in \operatorname{supp} \alpha \\ b \in \operatorname{supp} \beta \\ a_i' \in A_i}} \max\left\{ \frac{\ell v_i}{\ell v_i - [u_i(a_i', a_{-i}) - u_i(a)]}\frac{1}{2},\ \frac{(4\ell+2)v_i - 3[u_i(a_i', b_{-i}) - u_i(b)]}{(\ell+1)v_i - 2[u_i(a_i', b_{-i}) - u_i(b)]}\frac{1}{3} \right\}.$$

The constraints on $\ell \in \mathbb{N}$ are all finite lower bounds. Fixing one compatible $\ell \in \mathbb{N}$, the constraints on $T \in \mathbb{N}$ are all finite lower bounds as well. There are finitely many of them, so there exists a pair $(T, \ell) \in \mathbb{N}^2$ satisfying all obedience constraints simultaneously. $\square$