The dynamic cost of ex post incentive compatibility in repeated games of private information

David A. Miller*
UCSD

This version: September 27, 2005
First version: November 11, 2003

Abstract

In a repeated game with private information, a perfect public equilibrium (PPE) can break down if communication is not necessarily simultaneous or if players can "spy" on each others' information. An *ex post perfect public equilibrium* (EPPPE) is a PPE that is ex post incentive compatible in each stage game; unlike PPE, EPPPE is robust under to any communication protocol, and to spying. However, robustness comes at a cost to the players: in many games, efficient payoffs in the corresponding static mechanism design problem cannot be supported as average payoffs in an EPPPE, even when players are patient. In two-player repeated allocation games, an optimal EPPPE never employs a (static) efficient outcome function in any stage game. Instead, the players always prefer to give up some static efficiency by sometimes allocating to the player with the lower valuation. Under independent valuations, optimal equilibria are often stationary, but when valuations are globally interdependent, optimal equilibria are never stationary. Applied to the problem of collusion with hidden costs, these results yield new insights into the phenomenon of price wars in collusive equilibria.

JEL Classifications: C72 (Noncooperative games), C73 (Stochastic & dynamic games), D82 (Asymmetric & private information), L13 (Oligopoly & other imperfect markets)

*Address: Department of Economics, 9500 Gilman Dr., La Jolla, CA 92093–0508. E-mail: d9miller@ucsd.edu. Home page: http://dss.ucsd.edu/~d9miller. This paper is based on the second chapter of my Ph.D. dissertation at Stanford University. I thank Susan Athey, Ilya Segal, and Steven Tadelis, my dissertation committee, as well as Jon Levin, for their valuable guidance. I also thank Timothy Bresnahan, Harrison Cheng, Peter Hammond, Narayana Kocherlakota, Moritz Meyer-ter-Vehn, Paul Milgrom, Amalia Miller, Mark Wright, and seminar participants at Stanford, UT Austin, UCSD, Northwestern, Kellogg School of Management, Michigan, Brown, Yale School of Management, Wisconsin, UCLA, USC, UC Riverside, and the 2005 Econometric Society World Congress for helpful comments. I am grateful to the Stanford Institute for Economic Policy Research for financial support.

1 Introduction

Many an economic application involves a group of agents who wish to coordinate their activities for mutual advantage, but to do so they must often share private information. An external "mechanism design" institution can help by enforcing their desired outcome once they have revealed their information. However, in many important applications such institutions may be weak or absent. For example, several firms that wish to collude may experience private cost shocks that could disrupt their ability to divide the market efficiently, and since their collusion is illegal they cannot appeal to the government to organize them. Or the members of a community may wish to rent some resource to be shared among them as a public good, but experience private taste shocks that affect the optimal quantity for them to rent; local leaders who also experience taste shocks cannot serve as disinterested referees. In cases like these, the parties need to self-regulate the exchange of information and self-enforce informal agreements, which is possible if their interaction continues over time.

The usual approach to modelling these situations is to assume that the players have common knowledge of a common prior, and reveal their information simultaneously.[1] The passage of time, however, poses a distinct problem for simultaneity: when reporting their information to the group, the players each face a strong incentive to try to "speak last." By waiting until others have revealed their information, a player learns more about whether she can gain an advantage by manipulating her own report. The players will have similar problems if they lack a common prior, or if they can "spy" on one another. Through unobserved spying, for instance, a player can observe a noisy signal of other players' information, and by doing so can effectively force other players to unknowingly speak earlier. When players can spy on each other in this manner, there is little prospect for simultaneous communication. In this paper, I show that when communication is unstructured or the players do not share a common prior, the outcomes the players can support using informal enforcement amongst themselves are constrained compared to what they could achieve with simultaneous communication.

More specifically, I study *perfect public equilibrium* (PPE) under the constraint of *ex post incentive compatibility* (EPIC), yielding a refinement of PPE that I call *ex post perfect public equilibrium* (EPPPE). EPPPE models self-enforcing agreements that do not rely on any particular communication protocol or probability distribution. However, this robustness comes at a cost to the players. My first main result is that, for a wide range of economically interesting games, efficiency under EPPPE is generally an impossible goal: total welfare is bounded away from efficiency, and the bound applies uniformly even as players grow more

[1] Miller (2005) shows that under these assumptions, patient players in an infinitely repeated game can use informal enforcement amongst themselves to replicate the same outcomes that an external mechanism designer could provide in a static setting with mandatory participation.

patient. This result also implies that an optimal EPPPE can be characterized by the solution to a static mechanism design problem. My second set of results provides techniques for solving this problem within the class of games with separable utility, and in particular shows that in this class an optimal EPPPE is characterized by the solution to a linear program. My third set of results characterizes optimal EPPPEs for an economically important subclass of games. Most importantly, I show that in repeated allocation games (in which players must divide a rival good), in an EPPPE it is never optimal for the players to implement an efficient outcome function. Instead, a superior EPPPE can be constructed by giving up some allocative efficiency.

Recent work indicates that it can be fruitful to apply static mechanism design techniques to the analysis of infinitely repeated games with private information. The recursive structure of PPE, as demonstrated by Abreu, Pearce, and Stacchetti (1990), allows each player's equilibrium utility in any stage of the game to be decomposed into a dynamic program in which he maximizes his current payoffs plus his promised future utility, or "continuation reward." This dynamic program can be interpreted as a static mechanism design problem in which the continuation rewards promised by the equilibrium are absorbed into the static monetary transfers of a direct revelation mechanism. I call this "the mechanism design approach" to repeated games, and in Miller (2005) I proved that it is valid when the players are sufficiently patient. The approach is outlined in Section 3.

Fudenberg, Levine, and Maskin (1994) initiated the study of PPEs in repeated games with private information and communication.[2] Fudenberg, Levine, and Maskin prove a PPE folk theorem for games with independent private valuations and finite type spaces. Athey, Bagwell, and Sanchirico (2004) and Athey and Bagwell (2001) construct optimal collusive PPEs when costs are hidden. Aoyagi (2003a,b), and Martin and Vergote (2004) construct collusive PPEs in repeated auctions, when valuations may not be independent or type spaces may be intervals of real numbers. Levin (2003) and Rayo (2003) construct optimal equilibria for principal-agent and team relationships. Miller (2005) extends the folk theorem to PPEs in general games with private information and communication, including games with arbitrary type spaces. All these papers use PPE as their equilibrium concept, and therefore IIC as their incentive compatibility concept.[3]

However, IIC is a fragile incentive compatibility concept for static games with private information, so PPE is a correspondingly fragile equilibrium concept for repeated games with private information. In particular, IIC is not robust to several sensible modifications to the

[2] In addition, Compte (1998) and Kandori and Matsushima (1998) study communication in games with private monitoring, while Skrzypacz and Hopenhayn (2004), Blume and Heidhues (2002), and Hörner and Jamison (2004) focus on collusion with private information but no communication.

[3] Levin (2003) endows only one player with private information, so IIC and EPIC are identical. Rayo (2003) uses EPIC for the special case of independent types and separable production, but the contracts he constructs do not admit EPIC when extended to cases of interdependent production.

game. For instance, an IIC mechanism in a static game with rigidly enforced simultaneous messages will generally not be incentive compatible in an otherwise identical game with non-simultaneous messages. Similarly, an IIC mechanism will generally not be incentive compatible in an otherwise identical game in which players can spy on their opponents' private information. Finally, IIC mechanisms are sensitive to changes in the joint probability distribution over signals. Game modifications such as these allow players to know or discover information about each other in flexible and possibly complex ways, so that, for any one player, different types of private information may lead to the same interim expectations. Since an IIC mechanism induces players to report their signals truthfully by exploiting differences in interim expected utilities, IIC can break down when the relationship between information and interim expectations is not one to one.

EPIC concerns the players' incentives at the ex post phase, after the entire vector of private information has become common knowledge.[4] Thus it is robust to differences in players' beliefs about each other's information. EPIC is the appropriate concept of incentive compatibility when communication takes place face to face, in which case the ordering of announcements may not be fixed in advance, or when players have the opportunity to spy on their opponents' private information and by doing so subvert simultaneity. It is also the appropriate concept of incentive compatibility when the players lack a common prior, or lack common knowledge of a common prior. Bergemann and Morris (2003) show that, in typical settings, EPIC is the appropriate incentive compatibility concept for static mechanisms when players may have arbitrary higher order beliefs about each others' payoff relevant signals. In essence, the modeling choice between EPIC and IIC is a question of institutions: Is there some institutional mechanism to which the players can submit their private information, and trust that it will not be revealed until all players have submitted? Is there an institution that provides enough public information for the players to form common knowledge of a common prior? Is there a sufficiently inexpensive means by which players can prevent their private information from being surreptitiously observed? In many applications we might expect that such institutions are not available. In particular, if the relevant private information is embodied in an organizational characteristic that is in principle observable, such as production cost, then spying is likely to be a thorny problem.[5]

EPPPE's relationship with EPIC is analogous to PPE's relationship with IIC. The first main result of this paper, which is developed in Section 4, arises from a comparison of

[4]In an independent private valuations setting, EPIC is equivalent to dominant strategies incentive compatibility, but EPIC is weaker than dominant strategies when payoff functions are interdependent. For further discussion of this distinction, see Chung and Ely (2003a).

[5]For example: In 2003, a manager at Boeing was caught with confidential information from Lockheed Martin, regarding the financial and technical plans for an Air Force rocket contract for which both firms had bid. Perhaps based in part on the stolen plans, Boeing had won the larger share of the contract. See Kelley (2003).

these relationships. The usual static mechanism design approach assumes that there exists an outside authority to enforce monetary transfers, absorb any ex post imbalances in the transfers, and even provide a subsidy to convince the players to participate in the mechanism. But in the repeated game, continuation play offers future utility for the players to look forward to. Enforcement and participation can thus be endogenized by threatening players with possibility of losing this future utility should they deviate, while the budget can be brought into balance by substituting changes in future utility for unbalanced monetary transfers. Consider the set of expected utility profiles in a static mechanism design problem that can be implemented by some IIC or EPIC mechanism. Miller (2005) uses the fact that under IIC the budget can often be balanced ex post if participation constraints are ignored to prove that PPE can often support the highest levels of IIC-implementable utility. In contrast, although under EPIC the availability of future utility endogenizes enforcement and participation, shifting future utility in an attempt to ameliorate budget imbalances merely postpones the problem rather than solving it. Hence the Pareto frontier of EPPPE utilities is strictly worse than the Pareto frontier in the static mechanism design problem under EPIC (with the budget balanced in expectation).

The intuition for why EPIC is costly in repeated games is that if the allocation in any stage of an EPPPE is efficient, then the players as a group must generally receive different amounts of aggregate promised future utility after different ex post realizations of their signals.[6] Since aggregate promised future utility cannot exceed the optimum, after some realizations of signals it must be less than the optimum, and hence less than efficient. Although the quantity of aggregate promised future utility that must be foregone in order to support efficiency in any given period scales down in average terms as players become more patient, the players' greater patience places greater importance on the cost of supporting efficiency in future periods. These two effects exactly balance.

I show that "ex post unbalancedness" is necessary and sufficient for EPPPE to be inefficient: an efficient allocation in the static mechanism design problem under EPIC must not satisfy ex post budget balance. This condition is satisfied in many economically important models, and some applications are offered in later sections. A related example is provided by Gärtner and Schmutzler (2005), who consider mergers of asymmetrically informed firms; they show that efficiency and budget balance are generally irreconcilable with EPIC, even when the mechanism can employ not only monetary transfers, but also allocations of shares in the merged firm. Ex post unbalancedness requires that all players have payoff-relevant private information, and that the players cannot obtain insurance. For instance, if the players had access to actuarially fair insurance against budget imbalances, they could use

[6] In the terminology of Fudenberg, Levine, and Maskin (1994), efficiency under EPIC is generally not "enforceable" with respect to any hyperplane.

an efficient mechanism in every stage and impose the ex post imbalances on the insurer. Similarly, if one player did not possess private information, she could serve as an insurer inside the game. Both PPE and EPPPE are game theoretic equilibrium concepts that apply to fully modeled game situations (e.g., including any insurer as a strategic player), so they imply that ex post budget balance must be satisfied whenever an insurer is absent. However, the theorem still holds when there is free disposal of utility, since destroying utility at the end of the current period is a perfect substitute, in terms of providing incentives, for reducing aggregate promised future utility. I use the term "money burning" to refer to both means of providing incentives.[7]

Since EPIC efficiency is typically not attainable in EPPPE, the question of what is *optimal* (Pareto-best among EPPPEs) naturally arises. The second half of this paper, beginning with Section 5, examines this question in detail. The dynamic programming problem of designing an optimal EPPPE is equivalent to a static mechanism design problem in which the objective to be maximized is the welfare from the outcome function minus the expected quantity of money that must be burnt. That this problem has a solution under quite general conditions can be derived from a theorem of Balder (1996), as shown in Appendix A. To develop sharper characterizations, I restrict attention to games in the separable utility environment, as defined by Chung and Ely (2003a). I argue that for most interesting games in this class, an optimal EPPPE implements inefficient outcomes. Furthermore, I show that in this class the problem of designing an optimal EPPPE can be solved by linear programming.

Section 6 looks more closely at allocation games in the separable environment. Allocation games are those in which some object, benefit, or privilege is in limited supply and must be given to one player or divided among several players. Examples include repeated auctions, repeated trade, and collusion with a fixed market size. I prove that it is not optimal to employ an efficient outcome function in the first period. Instead, it is preferable to employ an outcome function that sometimes allocates inefficiently. Loosely speaking, it is always possible to improve on an efficient outcome function by partial "pooling"; i.e., sometimes asking the players to announce only that their information is within a particular range. Furthermore, in many (but not all) private valuations settings, it is optimal to pool over large regions so as to completely eliminate the need for money burning, yielding a stationary equilibrium. This indicates that in many cases the players are willing, as a group, to tolerate

[7] It is important to note that unless there is free disposal, EPPPE is not necessarily robust to non-common priors or non-common knowledge of common priors. This is because in the absence of free disposal EPPPE relies on changes in future utility on the equilibrium path in order to provide incentives. When players do not have common knowledge of a common prior, then even if the equilibrium provides appropriate incentives to every player, a player may not believe that other players face appropriate incentives. Under free disposal, all incentive payments can be executed at the end of the current period, so that heterogeneous beliefs about future utilities do not affect beliefs about incentives.

a substantial amount of inefficient allocation before they are willing to burn money. However, in the case of interdependent valuations, an optimal allocation always burns some money.

Section 7 applies these results to the problem of collusion with hidden costs to provide a new explanation for price wars. Under independent costs, the inefficiency inherent in an EPPPE is optimally manifested in the form of inefficient production, but when costs are interdependent an optimal EPPPE can also incorporate inefficiently low prices. Price wars have not previously been explained in the literature without resort either to impatience or to assumptions about imperfect observability that imply symmetric equilibria.

2 The model

The dynamic game consists of infinite repetitions of a stage game. All players share a common discount factor $\delta \in (0,1)$. The stage game has payoff-relevant private information, a forum for public announcements, and publicly observed actions, and players can exchange "money," which is a tradable good in zero net supply that enters their utility functions quasilinearly. For simplicity the only possible messages are the private signals. Signals are drawn independently over time and from the same distribution, although there may be interdependence among the signals observed by different players within any particular stage.[8]

Formally:

- $\mathcal{N}$ is a finite set of players, $i = 1, \ldots, N$;
- Θ_i is a set of possible private signals for player i, with $\Theta \equiv \Theta_1 \times \cdots \times \Theta_N$;
- ϕ is a common prior probability measure on Θ;
- $\mathcal{X}_i$ is a space of actions available to player i,with $\mathcal{X} \equiv \mathcal{X}_1 \times \cdots \times \mathcal{X}_N$;
- $\pi_i : \Theta \times \mathcal{X} \to \mathbb{R}_+$ is a measurable, uniformly bounded *payoff function* for player i, with $\pi \equiv (\pi_1, \ldots, \pi_N)$;
- $\mathbb{R}^N$ is the space of possible monetary transfers, where a positive component for player i indicates that player i receives a positive quantity of money.

The timing of the stage game is as follows: First, $\theta \in \Theta$ is realized according to ϕ, and each player i privately observes θ_i. Then each player sends a public announcement $\hat{\theta}_i \in \Theta_i$. After all announcements have been observed, each player i chooses an action from $\mathcal{X}_i$, which

[8] In Miller (2005) I allow a more general message space, and provide a more formal treatment of voluntary monetary transfers. There I also prove that the revelation principle is valid when the players are sufficiently patient, even though they reveal their information directly to each other rather than to a mechanism designer.

is publicly observed.[9] Finally, players can make zero-sum monetary transfers. Note that this setup allows interdependence among the payoff functions of different players.

A player's utility in the stage game is equal to her payoff plus any monetary transfer she receives. In the repeated game, each player seeks to maximize her discounted sum of stage game utilities. A strategy in this setting is, loosely speaking, a mapping from histories and private signals to announcements; from histories, private signals, and announcements to actions; and from histories, private signals, announcements, and actions to monetary transfers. PPE is the usual equilibrium concept for this type of game. A PPE is a sequential equilibrium in which players condition their strategies on only the public history and their private signal in the current period.[10] The formal notation for strategies and equilibria is fairly cumbersome and is laid out in Miller (2005).

This paper focuses on optimal equilibria, which achieve high payoffs along the equilibrium path by threatening to switch to a trigger punishment in the continuation game after any deviation from the equilibrium path is observed. Since the conditions imposed so far will not assure it, assume that there exists a perfect Bayesian equilibrium in the stage game, so that playing the stage game equilibrium in every period constitutes a sequential equilibrium that is suitable for use as a trigger punishment.[11]

Assumption 1. *There exists a perfect Bayesian equilibrium in the stage game.*

This equilibrium can be thought of as the default outcome when the players fail to communicate. By backward induction in a one-shot stage game, there can be no monetary transfers in a stage game equilibrium, so a player's utility in this equilibrium is merely her payoff from the outcome.

To make the model interesting, it is important to assume that players do not observe the entire vector of true signals ex post, since otherwise they could observe and punish any

[9]Later sections consider games in which the group must choose an action regarding—for example—who is to receive an object or whether a public good is to be provided. Such cases can be mapped onto this setup by setting the payoff functions to yield an extreme penalty if any player deviates from the group action. When the players are sufficiently patient, the inverse of this mapping can be seen as a reduced form interpretation of a grim trigger punishment that supports unanimous group actions in equilibrium.

[10]It is worth noting here that the set of attainable PPE payoffs does not generally equal the set of attainable sequential equilibrium payoffs. Although Abreu, Pearce, and Stacchetti (1990) showed that any pure strategy sequential equilibrium is payoff-equivalent to a PPE, the same property does not always hold for mixed strategy sequential equilibria. Kandori and Obara (2003) demonstrated that, in certain types of games, sequential equilibria in private strategies can outperform perfect public equilibria. Kandori and Obara construct private equilibria in which one player's private experiment can provide a better signal for her to monitor the other players. In games with private information and communication, however, the signal for each player's action is the announcement that he makes, so other players cannot influence its probability distribution by manipulating their own announcements. Whether there exist other types of private strategy equilibria that can improve on PPE in games with private information remains an open question.

[11]Since this paper focuses on the limit of Pareto-best equilibria as $\delta \to 1$, a simple grim trigger punishment is sufficient. More sophisticated punishments could enlarge the set of attainable payoffs, but would not improve the limit of the Pareto frontier as $\delta \to 1$.

deviation in which a player misrepresents her signal. For the same reason, it is also necessary to assume that players cannot use their ex post realized payoffs to infer their opponents' true signals. Since such inferences depend on the details of the players' conditional beliefs about each others' signals, this assumption is consistent with the type of robustness discussed in the introduction.[12] Finally, for simplicity I assume that randomizations over actions and transfers are public, so they can be observed and punished.

3 The mechanism design approach

The mechanism design approach described in Miller (2005) focuses on a special class of PPEs that can be represented as recursive mechanisms in which the players directly reveal their information to each other. A recursive mechanism tracks promised utilities along the equilibrium path as a state variable. The problem of designing a recursive mechanism can be interpreted as a static mechanism design problem in which the present value of the promised future utilities are folded into the monetary transfers of the current period. The static literature usually assumes the presence of an outside authority who operates the mechanism and enforces its prescriptions. However, since no such outside authority exists in the context of a repeated game, it is necessary to prove the validity of the mechanism design approach. Theorem 1 from Miller (2005), restated at the end of Section 3.1, shows that if the players are sufficiently patient then this approach yields the same Pareto frontier of utilities that are attainable in PPE. Sufficient patience is necessary because the fact that players reveal their information to each other rather than to a mechanism designer is restrictive for low discount factors.

3.1 Recursive mechanisms

PPEs are naturally unwieldy objects to work with, because in each stage they take the entire history of messages as an argument. Fortunately, Abreu, Pearce, and Stacchetti (1990) showed that PPEs can be constructed recursively without loss of generality with respect to attainable utilities. In addition, Miller (2005) showed that it is without loss of generality with respect to the Pareto frontier to work in the smaller and more convenient class of recursive mechanisms. A recursive mechanism employs promised utility as a payoff-irrelevant state variable, and in each stage maps the promised utility and the vector of players' announcements into prescribed actions, prescribed monetary transfers, and promised

[12] Mezzetti (2004) shows how to construct efficient EPIC mechanisms in a static setting with interdependent payoffs by adding a second communication phase in which the players announce their realized payoffs. However, Mezzetti's mechanisms cannot simultaneously satisfy EPIC and ex post budget balance, and so they cannot be used to construct efficient EPPPEs even when there is a second communication phase at the end of each stage.

continuation rewards for the succeeding stage. Thus each stage requires an outcome function x, a transfer function t, and a continuation reward function w; collectively, these three functions applied in a particular stage are called a "stage mechanism":

Definition 1. A *stage mechanism* is a measurable triplet $\langle x, t, w \rangle : \Theta \to \mathcal{X} \times \mathbb{R}^N \times \mathbb{R}^N$ such that $\sum_i t_i(\theta) = 0$ for all $\theta \in \Theta$.

In a recursive mechanism, some deviations are observable—for example, deviant actions or refusal to participate in monetary transfers. Such "off-menu" deviations are discouraged by terminating the use of stage mechanisms and switching to the punishment equilibrium. There is also the potential for "on-menu" deviations, in which a player misrepresents her private signal in a way that the other players cannot observe. On-menu deviations are discouraged by imposing incentive compatibility constraints, defined below.[13]

In addition to a collection of stage mechanisms, a recursive mechanism must also specify a set $\mathcal{V}$ of promised (average) utility state variables over which it is defined, as well as an initial promised utility v^0. Formally:

Definition 2. A *recursive mechanism* is a triplet $\big\langle \mathcal{V}, \big\{\langle x(\cdot; v), t(\cdot; v), w(\cdot; v)\rangle : v \in \mathcal{V}\big\}, v^0 \big\rangle$, abbreviated as $\big\langle \mathcal{V}, \big\{\langle x, t, w\rangle(\cdot; v)\big\}, v^0\big\rangle$, such that:

(i) $\mathcal{V} \subset \mathbb{R}^N$,

(ii) $\big\{\langle x, t, w\rangle(\cdot; v)\big\}$ is a collection of stage mechanisms indexed by $v \in \mathcal{V}$,

(iii) $v^0 \in \mathcal{V}$.

Such a mechanism is called "recursive" because in each period the stage mechanism is selected based on the promised utility v carried over from the previous period, and the stage mechanism generates a continuation reward $w(\hat{\theta}; v)$ that becomes the promised utility for the next period (where $\hat{\theta}$ is the vector of announced signals). This notation implies a deterministic mechanism. At several points the analysis raises the possibility that mechanisms may be randomized, but the extra notation is introduced only where necessary. Furthermore, in contexts that do not invite confusion, the $(\cdot; v)$ notation is dropped for clarity.

Given a signal profile $\theta \in \Theta$ and a vector of announcements $\hat{\theta} \in \Theta$ in a stage mechanism $\langle x, t, w\rangle$, each player's total utility is her stage game utility plus the present value of her continuation reward:

$$u_i(\theta, \hat{\theta}; \delta, x, t, w) \equiv \pi_i(\theta, x(\hat{\theta})) + t_i(\hat{\theta}) + \frac{\delta}{1-\delta} w_i(\hat{\theta}). \tag{1}$$

[13] Here, "on-menu" and "off-menu" have the same meaning as the terminology "on-schedule" and "off-schedule" developed by Athey, Bagwell, and Sanchirico (2004) and Athey and Bagwell (2001).

In order for a recursive mechanism to be equivalent to a PPE, it must be feasible, individually rational, and incentive compatible (see Miller (2005)). Here, individual rationality—the constraint that players do not wish to make off-menu deviations—is assured by the threat of the trigger punishment, so long as (i) the average payoffs in equilibrium exceed those of the punishment and (ii) the players are sufficiently patient.[14] As $\delta \to 1$, sufficient patience is assured, so for simplicity I require individual rationality only at the limit. Thus the individual rationality constraint can be folded into the feasibility constraint. Let p_i be player i's payoff in the trigger punishment. Define

$$\mathcal{P} \equiv \left\{ v \in \mathbb{R}^N : v_i > p_i \ \forall\, i \in \mathcal{N} \text{ and } \sum_i v_i \leq \max_{x:\Theta\to\mathcal{X}} \sum_i \mathbb{E}_\phi\big[\pi_i(\theta, x(\theta))\big] \right\} \tag{2}$$

as the set of attainable and individually rational payoffs in the stage game.[15] Feasibility, then, is the constraint that a recursive mechanism must deliver the utility that it promises, subject to individual rationality:

Definition 3. Given a discount factor $\delta \in (0,1)$, a recursive mechanism $\big\langle \mathcal{V}, \big\{\langle x,t,w\rangle(\cdot;v)\big\}, v^0\big\rangle$ is *feasible* if:

(i) $\mathcal{V} \subset \mathcal{P}$,

(ii) $v = (1-\delta)\,\mathbb{E}_\phi\big[u\big(\theta,\theta;\delta,x(\cdot;v),t(\cdot;v),w(\cdot;v)\big)\big]$ for all $v \in \mathcal{V}$,

(iii) $w(\theta;v) \in \mathcal{V}$ for all $\theta \in \Theta$ and all $v \in \mathcal{V}$.

Under incentive compatibility, each player must prefer to make a truthful announcement rather than deviate to some other on-menu announcement. EPIC can be thought of as an incentive compatibility concept that is robust to any possible order in which announcements might be made, while IIC can be thought of as an incentive compatibility concept that requires all announcements to be made simultaneously.

Definition 4. Given a discount factor $\delta \in (0,1)$, a stage mechanism $\langle x,t,w\rangle$ is *ex post incentive compatible (EPIC)* if

$$u_i\big(\theta,\theta;\delta,x,t,w\big) \geq u_i\big(\theta,(\hat\theta_i,\theta_{-i});\delta,x,t,w\big) \tag{3}$$

[14] This is individual rationality in the game theoretic sense rather than the usual mechanism design sense of participation constraints. However, condition (i) makes clear that in this context the two notions are equivalent if the punishment is interpreted as an outside option.

[15] The notation $\mathbb{E}_\phi\big[f(\cdot)\big]$ indicates $\int_\Theta f\,d\phi$. I use this somewhat non-standard notation because there are places in this paper where the more familiar notation $\mathbb{E}_\theta\big[f(\theta)\big]$ would invite misinterpretation.

for all $\hat{\theta}_i \in \Theta_i$, for all $i \in \mathcal{N}$, and for ϕ-almost all $\theta \in \Theta$.[16] It is *interim incentive compatible (IIC)* if

$$\mathbb{E}_\phi\left[u_i\big(\theta,\theta;\delta,x,t,w\big)\big|\theta_i\right] \geq \mathbb{E}_\phi\left[u_i\big(\theta,(\hat{\theta}_i,\theta_{-i});\delta,x,t,w\big)\big|\theta_i\right] \tag{4}$$

for all $\hat{\theta}_i \in \Theta_i$, for all $i \in \mathcal{N}$, and for ϕ_i-almost all $\theta_i \in \Theta_i$.[17] A recursive mechanism $\big\langle \mathcal{V}, \big\{\langle x,t,w\rangle(\cdot;v)\big\}, v^0\big\rangle$ is EPIC (resp. IIC) if $\langle x,t,w\rangle(\cdot;v)$ is EPIC (resp. IIC) for all $v \in \mathcal{V}$.

Note that EPIC implies IIC. EPIC also has two other important properties. First, given the probability measure ϕ, a stage mechanism is EPIC if and only if it is IIC for every probability measures that is mutually absolutely continuous with respect to ϕ.[18] Second, Bergemann and Morris (2003) show (in a model with a finite signal space Θ) that, for most games, a static mechanism is EPIC if and only if it is IIC in an otherwise equivalent game formulated with the universal type space—that is, the space in which players are allowed arbitrary higher order beliefs about their opponents' private information and beliefs.[19]

The following theorem restates Theorem 1 from Miller (2005), which shows that it is without loss of generality to consider instead the much simpler class of feasible and IIC recursive mechanisms. This result justifies the mechanism design approach to repeated games with private information. The main intuition is that when the players select their recursive mechanism (this selection is not modeled), they are collectively acting as their own mechanism designer. Their recursive mechanism specifies the equilibrium path they should follow; incentive compatibility discourages on-menu deviations while the trigger punishment discourages off-menu deviations. Hence a feasible and IIC recursive mechanism describes an equilibrium path and implies the existence of a PPE that supports that path. In the other direction, the equilibrium path of a PPE can be described by a history-dependent dynamic

[16] In the terminology of Balder (1996), this is "almost sure incentive compatibility." Although this is not the most complete notion of incentive compatibility, it enables Appendix A to apply Balder's existence result for an optimal equilibrium. Almost sure incentive compatibility is consistent with the construction of a PPE if, for any (zero-probability) event that some θ is realized on which incentive compatibility fails, there exists some sequential equilibrium in the continuation game. Since the ϕ-measure of the set of all such θ is zero, the expected value of the equilibrium prior to θ being realized is not affected.

[17] Where $\phi_i(\theta_i) = \int_{\Theta_{-i}} d\phi(\cdot,\theta_i)$.

[18] To see this, suppose f is EPIC given ϕ. Then there may exist a ϕ-null set E_0 on which incentive compatibility fails. However, f is clearly EPIC, and thus IIC, for any probability measure for which E_0 is a null set, including every probability measure with respect to which ϕ is absolutely continuous. Conversely, suppose f is IIC for any ϕ' that is absolutely continuous with respect to ϕ. Let E_0 be the maximal ϕ-null set on which incentive compatibility fails. Then, for any particular $\theta \in \Theta\backslash E_0$, f is IIC for any sequence $\{\phi'_k\}$ of probability measures, each absolutely continuous with respect to ϕ, that converges to the probability measure ϕ_θ that assigns probability 1 to $\{\theta\}$ (although the limit may not be absolutely continuous with respect to ϕ). This implies that f is IIC for ϕ_θ. Since this is true for every $\theta \in \Theta\backslash E_0$, it implies that f is EPIC given ϕ.

[19] Note that if the players do not share a common belief about ϕ, then in a recursive mechanism they may disagree over whether the promised future utilities are feasible. However, if in a repeated game they share the same belief about ϕ but do not have common knowledge of this belief, then EPIC may be an appropriate incentive compatibility concept. Chung and Ely (2003b) takes some first steps in this direction for the context of a static mechanism design problem with independent private valuations.

mechanism. Similar to the logic of Abreu, Pearce, and Stacchetti (1990), it is without loss of generality with respect to attainable payoffs to substitute a recursive mechanism for a dynamic mechanism.

Theorem 1. *Let $p \in \mathbb{R}^N$ be the expected utility profile of some stage game perfect Bayesian equilibrium. For any $\varepsilon > 0$, there exists $\underline{\delta} < 1$ such that, for all $\delta > \underline{\delta}$:*

(i) If a feasible and IIC recursive mechanism $\langle \mathcal{V}, \{\langle x, t, w\rangle (\cdot; v)\}, v^0\rangle$ satisfies $v_i > p_i + \varepsilon$ for all $i \in \mathcal{N}$ and all $v \in \mathcal{V}$, then there exists a PPE that yields the same announcements, actions, and net transfers along the equilibrium path;

(ii) If $\mathcal{V} \subset \mathbb{R}^N$ is the set of average utility profiles yielded at the beginning of any period along the equilibrium path of some PPE, and $v_i > p_i + \varepsilon$ for all $i \in \mathcal{N}$ and all $v \in \mathcal{V}$, then for any $v \in \mathcal{V}$ there exists a feasible and IIC recursive mechanism with initial promised utility v.

The proof is in Miller (2005). This theorem validates the mechanism design approach with the IIC concept of incentive compatibility. Notice that, since any EPIC recursive mechanism is also IIC, conclusion (i) of Theorem 1 implies that any feasible and EPIC recursive mechanism generates an equilibrium path that can be supported by a PPE. Accordingly, define:

Definition 5. An *ex post perfect public equilibrium (EPPPE)* is a PPE for which the equilibrium path can be generated by a feasible and EPIC recursive mechanism.

This is similar to defining an EPPPE as a PPE that satisfies EPIC (defined appropriately on the space of strategies) in every continuation game. Under this definition, it is appropriate shorthand to call any EPIC recursive mechanism an "EPPPE mechanism."

3.2 The equivalent static mechanism

The recursive problem can be transformed into an equivalent static problem as follows.

Definition 6. Given $\delta \in (0, 1)$ and a stage mechanism $\langle x, t, w\rangle$, the *equivalent static mechanism* is a pair $\langle x, y\rangle$ where

$$y(\theta) \equiv t(\theta) + \frac{\delta}{1-\delta}\big(w(\theta) - \mathbb{E}_\phi[w(\theta)]\big) \tag{5}$$

is the *equivalent static transfer function.*

In a static mechanism $\langle x, y\rangle$, players maximize $\pi_i(\theta, x(\hat{\theta})) + y_i(\hat{\theta})$; i.e., they maximize recursive utility in the stage mechanism to which the static mechanism is equivalent. A static mechanism satisfies EPIC (resp. IIC) if and only if it is equivalent to a stage mechanism that satisfies EPIC (resp. IIC). In a recursive mechanism, individual rationality requires $w \in \mathcal{P}$, but this imposes no constraints on the equivalent static transfer function y, because y expresses the deviation of $\frac{\delta}{1-\delta} w$ from its mean. When discussing static mechanisms this paper ignores individual rationality, which is not problematic when the goal is to analyze Pareto-best recursive mechanisms as $\delta \to 1$. When a larger set of attainable payoffs is of interest, the individual rationality requirement needs to be imposed; see Miller (2005) for an example of how to accomplish this.

Note that by construction an equivalent static mechanism is ex ante budget balanced: $\mathbb{E}_\phi\big[\sum_i y_i(\theta)\big] = 0$. Later sections construct recursive mechanisms from equivalent static mechanisms, and in such contexts it is crucial to impose expected budget balance as a constraint on the static problem.

3.3 Summary functions, efficiency, and optimality

As utility is transferrable and the focus is on Pareto-best mechanisms, much of the succeeding analysis will make use of aggregated summary functions. To simplify the notation, for most functions a capital letter indicates the aggregation across all agents of the corresponding lower case letter. In particular, for a recursive mechanism $\big\langle \mathcal{V}, \big\{\langle x, t, w\rangle(\cdot; v)\big\}, v^0\big\rangle$:

- $V \equiv \sum_i v_i$ is the *aggregate promised utility*,
- $V^0 \equiv \sum_i v_i^0$ is the *value* of the recursive mechanism,
- $W(\theta) \equiv \sum_i w_i(\theta)$ is the *aggregate continuation reward function*,
- $Y(\theta) \equiv \sum_i y_i(\theta)$ is the *aggregate equivalent transfer function.*

It is also helpful to name two more aggregate quantities, given a stage mechanism $\langle x, t, w\rangle$:

- $\sum_i \mathbb{E}_\phi[\pi_i(\theta, x(\theta))]$ is the *aggregate welfare*,
- $\frac{\delta}{1-\delta}\left(\max_\theta[W(\theta)] - \mathbb{E}_\phi[W(\theta)]\right)$ is the *surplus gap.*[20]

Aggregate welfare is, naturally, the expected aggregate payoff from a particular outcome function. The surplus gap is a little less intuitive, but the results in the next section demonstrate its critical importance: the value V^0 of an optimal EPPPE mechanism is equal to the aggregate welfare given $x(\cdot; v^0)$ minus the surplus gap given $w(\cdot; v^0)$. In a dynamic sense,

[20] Most of this paper uses notation as if maxima always exist. Most results can be applied to cases in which maxima do not exist by employing appropriate definitions and limit arguments. For specific conditions that assure the existence of an optimal EPPPE mechanism, see Appendix A.

the surplus gap is the difference in total future welfare between the best possible future path and the expected future path, given a continuation reward function w. If there is free disposal of utility, or "money burning," then the surplus gap can be interpreted as the amount of money that the players expect to burn. In a static sense, the surplus gap is the largest possible budget deficit given an equivalent static transfer function y, as stated by the following simple lemma:

Lemma 1. *Given a stage mechanism $\langle x, t, w \rangle$, if $\langle x, y \rangle$ is its equivalent static mechanism, then*

$$\frac{\delta}{1-\delta}\left(\max_{\theta}[W(\theta)] - \mathbb{E}_{\phi}[W(\theta)]\right) = \max_{\theta}[Y(\theta)]. \tag{6}$$

Proof. By construction of y from t and w. □

The notion of efficiency employed in this paper includes only those stage game allocations that can be implemented under the constraints of EPIC. Specifically, an outcome function x is *ex post implementable* if there exists an equivalent static transfer function y such that $\langle x, y \rangle$ is EPIC. "EPIC efficiency" is defined over the set of ex post implementable outcome functions, while optimality is defined over the set of EPPPE mechanisms:

Definition 7. An ex-post implementable outcome function is *EPIC-efficient* if no other ex post implementable outcome function attains higher aggregate welfare. An EPPPE mechanism is *optimal* if no other EPPPE mechanism attains higher value.

4 The dynamic cost of EPIC

4.1 Designing an optimal EPPPE mechanism

The main result of this section is Theorem 2, which characterizes the value of any optimal EPPPE mechanism. The following "ceiling" lemma is a key step in proving this result. It states an important fact about the promised continuation reward function $w(\cdot; v)$ in any optimal EPPPE mechanism: the highest possible aggregate continuation reward $W(\theta)$ that can be promised at the end of the first period must equal the value V^0 of the mechanism—i.e., initial promised utility at the start of the first period.

Lemma 2 (Ceiling lemma)**.** *If $\left\langle \mathcal{V}, \left\{ \langle x, t, w \rangle(\cdot; v) \right\}, v^0 \right\rangle$ is an optimal EPPPE mechanism, then $V^0 = \max_\theta[W(\theta; v^0)]$.*

The proof, in Appendix B, is somewhat heavy on notation, but its logic is relatively straightforward. If the highest aggregate continuation reward $\max_\theta[W(\theta; v^0)]$ were greater than the value V^0 of the mechanism, then it would be preferable to start the mechanism as if the players had already reached the continuation game in which $\max_\theta[W(\theta; v^0)]$ was the initial aggregate promised utility. If $\max_\theta[W(\theta; v^0)]$ were less than V^0, then it would be preferable to design an alternative mechanism that increases the entire schedule of aggregate continuation rewards $W(\cdot; v)$ by a small lump sum $\alpha > 0$. This would be feasible because, for any θ and small enough α, $W(\theta; v^0) + \alpha$ would not exceed V^0.[21] The bulk of the proof consists of constructing this alternative mechanism in detail.

The following theorem states a fundamental property of any optimal EPPPE: its value V^0 must equal the first period aggregate welfare minus the first period surplus gap. Recall that the surplus gap is defined as the maximum continuation reward minus the expected continuation reward, in aggregate present value terms: $\frac{\delta}{1-\delta}\left(\max_\theta[W(\theta; v^0)] - \mathbb{E}_\phi[W(\theta; v^0)]\right)$. This theorem indicates that the surplus gap is the quantity of aggregate utility that must be given up to satisfy EPIC in an optimal EPPPE.

Theorem 2. *If $\left\langle \mathcal{V}, \left\{\langle x, t, w\rangle(\cdot; v)\right\}, v^0\right\rangle$ is an optimal EPPPE mechanism, then*

$$V^0 = \sum_i \mathbb{E}_\phi[\pi_i(\theta, x(\theta; v^0))] - \frac{\delta}{1-\delta}\left(\max_\theta[W(\theta; v^0)] - \mathbb{E}_\phi[W(\theta; v^0)]\right). \tag{7}$$

Proof. First, feasibility requires that

$$V^0 = (1-\delta)\sum_i \mathbb{E}_\phi[\pi_i(\theta, x(\theta; v^0))] + \delta\mathbb{E}_\phi\left[W(\theta; v^0)\right]. \tag{8}$$

Second, the definition of the surplus gap implies that

$$\begin{aligned} V^0 = {} & (1-\delta)\sum_i \mathbb{E}_\phi[\pi_i(\theta, x(\theta; v^0))] \\ & + \delta\left(\max_\theta[W(\theta; v^0)] - \frac{1-\delta}{\delta}\frac{\delta}{1-\delta}\left(\max_\theta[W(\theta; v^0)] - \mathbb{E}_\phi[W(\theta; v^0)]\right)\right). \end{aligned} \tag{9}$$

Third, by Lemma 2, $V^0 = \max_\theta[W(\theta; v^0)]$. With some rearrangement, this yields Eq. 7. □

The theorem states a property of optimal EPPPE mechanisms, but by Theorem 1, if the discount factor is sufficiently high then the conclusion applies to optimal EPPPEs as well. It is important to note that the value V^0 of an optimal EPPPE mechanism does not change as the discount factor δ changes. Promised utility w enters both the EPIC constraints and the surplus gap only as $\frac{\delta}{1-\delta}w$; i.e., in present value terms rather than average terms. Although w enters the feasibility constraint in average terms, the feasibility constraint does not bind

[21] Recall that stage mechanisms may randomize, so the set of possible promised utilities is convex.

for sufficiently high δ. Thus, although the optimal continuation reward function must vary in average terms as δ varies, it must do so in a manner such that its present value remains fixed. Hence for an optimal EPPPE mechanism the first period outcome function, monetary transfer function, and surplus gap do not vary with δ, and thus Theorem 2 implies that the value of the mechanism does not vary with δ. The corollary below states this fact formally.

Corollary 1. *There exists* $\underline{\delta} < 1$ *such that, for all* $\delta > \underline{\delta}$ *and* $\delta' > \underline{\delta}$, *if* $\langle \mathcal{V}, \{\langle x, t, w\rangle(\cdot; v)\}, v^0\rangle$ *is an optimal EPPPE mechanism given* δ, *then there exists an EPPPE mechanism* $\langle \mathcal{V}, \{\langle x, t, w'\rangle(\cdot; v)\}, v^0\rangle$ *that is optimal given* δ', *with* $w'(\cdot; v^0) = \frac{\delta}{1-\delta}\frac{1-\delta'}{\delta'}\big(w(\cdot; v^0) - v^0\big) + v^0$.

The proof is omitted.[22] Since $\frac{\delta}{1-\delta}\left(\max_\theta[W(\theta)] - \mathbb{E}_\phi[W(\theta)]\right) = \max_\theta[Y(\theta)]$ and the feasibility constraint does not bind at the limit as $\delta \to 1$, Theorem 2 and Corollary 1 indicate that the main task in designing an optimal EPPPE mechanism when players are patient is to solve the static mechanism design problem of maximizing aggregate welfare *minus the surplus gap*. Formally, the task is to solve

$$\max_{\langle x,y\rangle}\left[\sum_i \mathbb{E}_\phi[\pi_i(\theta, x(\theta))] - \max_\theta[Y(\theta)]\right] \quad \text{s.t. (i) EPIC and (ii) } \mathbb{E}_\phi\big[Y(\theta)\big] = 0, \qquad (10)$$

where (ii) must hold by the definition of an equivalent static mechanism (Definition 6). Given a discount factor δ, the equivalent transfer function y can be decomposed into a monetary transfer function t and a continuation reward function w.

Note that when monetary transfers must sum to zero for every realization of θ, to construct an optimal EPPPE it is not sufficient to merely solve Eq. 10 for the first stage mechanism. It is also necessary to design stage mechanisms for the later periods that support the continuation rewards promised by $w(\cdot; v^0)$. As in the proof of Lemma 2, this process is somewhat simplified by allowing randomized mechanisms. It is possible, however, to keep the analysis focused on the equivalent static mechanism $\langle x, y\rangle$ by expanding the game to allow free disposal of utility, or money burning; i.e., to allow $\sum_i t_i(\theta) \leq 0$. Then the first period's stage mechanism can be used in every period, with the proviso that at the end of each period the players must collectively burn the amount $\frac{\delta}{1-\delta}\big(\max_\theta\big[W(\theta)\big] - W(\theta; v^0)\big)$ in order to satisfy EPIC. Thus the problem of designing an optimal EPPPE mechanism is solved implicitly by solving the problem of designing an optimal equivalent static mechanism. Appendix A demonstrates that there exists a solution to this problem under quite general conditions.

[22] The only subtlety is that it is necessary to show that the promised utilities indicated by $\{w'(\theta; v) : \theta \in \Theta\}$ can be supported by EPPPE mechanisms in their respective continuation games. This is accomplished by reasoning similar to that in Lemma 2.

4.2 Ex post unbalancedness

By Theorem 2, if the surplus gap is strictly positive for an EPIC-efficient outcome function, then EPIC efficiency cannot be attained by any EPPPE. Recall that "EPIC-efficient" means efficient subject to EPIC, so it is possible to implement an EPIC-efficient outcome in any particular period. However, the positive surplus gap means that aggregate continuation rewards must on average fall below the EPIC-efficient level of aggregate utility. The corollary below indicates that the surplus gap for an efficient outcome function is strictly positive under a simple condition called "ex post unbalancedness." An ex post unbalanced game is one in which no equivalent static mechanism can approximate both EPIC efficiency and ex post budget balance while satisfying EPIC.

Definition 8. Let Q^* be the EPIC-efficient level of aggregate welfare. A stage game is *ex post unbalanced* if there exist some $\varepsilon > 0$ and $\eta > 0$ such that if an equivalent static mechanism $\langle x, y \rangle$ satisfies EPIC and $\sum_i \mathbb{E}_\phi[\pi_i(\theta, x(\theta))] > Q^* - \varepsilon$, then there exists a subset $E \subset \Theta$ such that $\phi(E) > \eta$ and $\theta \in E$ implies $Y(\theta) > \eta$.

Section 5 will show that interesting games are usually ex post unbalanced. But, intuitively, for efficient mechanisms the two constraints of EPIC and ex post budget balance usually conflict because EPIC requires ex post imbalances in order to implement efficiency. Given a well-behaved game, this is likely to be true for approximate efficiency as well.

Corollary 2. *Let Q^* be the EPIC-efficient level of aggregate welfare, and assume that the set of EPPPEs is non-empty. Then the following conditions are equivalent:*

(i) The stage game is ex post unbalanced.

(ii) For all $\delta \in (0, 1)$, the set of aggregate average payoffs attainable in an EPPPE of the repeated game is uniformly bounded above by some $\overline{V} < Q^*$.

This corollary shows that EPIC efficiency is not attainable in an EPPPE mechanism whenever the stage game is ex post unbalanced. Complications in the proof, in Appendix B, arise from the possibility that an optimal EPPPE mechanism may not exist. If a optimum exists, then ex post unbalancedness reduces to the condition that no EPIC-efficient outcome rule can be implemented with ex post budget balance. In this case the proof is simple: either the optimum implements an EPIC-efficient outcome rule but burns money, or it implements an EPIC-inefficient outcome rule. Either way the players do not realize an EPIC-efficient level of welfare.

This contrasts with the conclusion in Miller (2005) that efficiency is generally attainable in a PPE, given three or more players and a sufficiently high discount factor. Since EPPPE is a more robust equilibrium concept than PPE, this identifies a trade-off between robustness and efficiency. The critical difference between IIC and EPIC is that IIC constrains expected monetary transfers at the interim stage rather than actual monetary transfers at the ex post stage. Thus under IIC and PPE there may be leeway to adjust the ex post monetary transfers to achieve ex post budget balance. In a sense, ex post unbalancedness can potentially be circumvented by institutions that support IIC as the appropriate concept of incentive compatibility.

5 The separable utility environment

The analysis thus far has dealt with games only in the abstract. To apply these results, this section introduces a class of games that contains the models most commonly used to study problems with private information. This class is called the *separable utility environment*, after Chung and Ely (2003a), and it is characterized by one-dimensional signals and multiplicatively separable utility.[23]

Definition 9. A game in the *separable utility environment* satisfies

(i) $\mathcal{X}$ is a compact and convex metric space;

and, for each $i \in N$:

(ii) $\Theta_i = \left[\underline{\theta}, \overline{\theta}\right] \subset \mathbb{R}$;

(iii) $\pi_i(\theta, x(\theta)) = \pi_i^x(x(\theta))\, \pi_i^\theta(\theta)$;

(iv) $\pi_i^x : \mathcal{X} \to [0,1]$ is continuous;

(v) $\pi_i^\theta : \Theta \to [-1,1]$ is differentiable in θ_i and strictly increasing in θ_i for ϕ_{-i}-almost every θ_{-i};

(vi) If x is efficient, then $\pi_i^x(x(\theta))$ is non-decreasing in θ_i for all $\theta_{-i} \in \Theta_{-i}$.

Condition (i) ensures the existence of an optimal EPPPE, by Proposition 6 in Appendix A, and includes the case in which $\mathcal{X}$ is the space of probability measures on some compact, but not necessarily convex, metric space. Conditions (ii–v) place the game in

[23] The separable utility environment also uses collective actions rather than individual actions. As mentioned in Section 2, a game with collective actions can be modeled by specifying that the payoff function provides extremely low utility if there is any disagreement among the players. That is, if some player disagrees with the collective action prescribed by a mechanism, then that player (and possibly the other players) will be penalized severely. Alternatively, these penalties could be enacted via trigger punishments in equilibrium rather than by assumption on the payoff functions. The approach here is to define the action space $\mathcal{X}$ over collective actions, and simply assume that players cannot unilaterally deviate from a collective action.

the separable utility environment as defined by Chung and Ely (2003a). Condition (vi) is added for convenience, so that unconstrained efficiency is also EPIC efficient: by Chung and Ely's Theorem 2, (vi) is necessary and sufficient for an outcome function to be ex post implementable in the separable utility environment.

Much of the analysis in the remainder of the paper concerns the equivalent static mechanism. For brevity I refer to an equivalent static mechanism simply as a "mechanism" wherever it will not be confused with a stage mechanism or a recursive mechanism, and refer to an equivalent static transfer function as a "transfer function."

Given an ex post implementable outcome function x, in an EPIC mechanism each player's transfer function must have a particular shape in order to discourage on-menu deviations. Specifically, Chung and Ely's Proposition 5 implies that the envelope theorem (for instance, Milgrom and Segal (2002) Corollary 1) fully determines each player's equilibrium utility for each θ_{-i}, up to an additive constant $y_i(\underline{\theta}, \theta_{-i})$: for all $\theta \in \Theta$,

$$\begin{aligned}\pi_i(\theta, x(\theta)) + y_i(\theta) &= \pi_i\big((\underline{\theta}, \theta_{-i}), x(\underline{\theta}, \theta_{-i})\big) + y_i(\underline{\theta}, \theta_{-i}) \\ &\quad + \int_{\underline{\theta}}^{\theta_i} \frac{\partial \pi_i^{\theta}(s_i, \theta_{-i})}{\partial s_i} \pi_i^x(x(s_i, \theta_{-i}))ds_i. \end{aligned} \tag{11}$$

Each player i's transfer $y_i(\theta)$ can be decomposed as

$$y_i(\theta) = y_i^0(\theta_{-i}) - r_i(\theta; x). \tag{12}$$

where $y_i^0(\theta_{-i}) \equiv y_i(\underline{\theta}, \theta_{-i})$ is her *fixed transfer*, the portion of her transfer that does not vary with θ_i; and $r_i(\theta; x)$ is her *EPIC payment*, the portion of her transfer that is pinned down by EPIC and the choice of outcome function x. Then rearranging Eq. 11 gives

$$r_i(\theta; x) = \pi_i(\theta, x(\theta)) - \pi_i\big((\underline{\theta}, \theta_{-i}), x(\underline{\theta}, \theta_{-i})\big) - \int_{\underline{\theta}}^{\theta_i} \frac{\partial \pi_i^{\theta}(s_i, \theta_{-i})}{\partial s_i} \pi_i^x(x(s_i, \theta_{-i}))ds_i. \tag{13}$$

As usual, aggregate EPIC payments and aggregate transfers are notated with capital letters: $R(\theta; x) \equiv \sum_i r_i(\theta; x)$ and $Y(\theta) = \sum_i y_i^0(\theta) - R(\theta; x)$.

5.1 The generality of ex post unbalancedness

In the separable utility environment, it is possible to address the generality of ex post unbalancedness, under which EPPPEs are unable to support EPIC efficiency. The first step in this discussion is to clarify how to identify whether a game is ex post unbalanced. Then I present examples of games that are ex post unbalanced, and which are clearly not knife-edge cases, as well as some examples of games that are not ex post unbalanced. I argue intuitively that certain characteristics that make games "interesting" are also likely to make them ex post unbalanced.

Recall that ex post unbalancedness is the condition that no mechanism can approximate both (EPIC) efficiency and ex post budget balance. In the separable utility environment, the aggregate EPIC payment function $R(\cdot;x)$ is fixed by the choice of outcome function x, so the question of whether ex post budget balance is possible is answered by determining whether there exist fixed transfer functions $\{y_i^0\}_{i\in\mathcal{N}}$ such that $\sum_i y_i^0(\theta) - R(\theta;x) = 0$ for all θ. Since $\sum_i y_i^0(\theta)$ is by construction additively separable in θ_i and θ_{-i} for all θ and all i—or "$(N-1)$–additively separable"—we conclude that a game is ex post unbalanced if and only if, for x^* efficient, $R(\theta;x^*)$ is not $(N-1)$–additively separable.[24]

Definition 10. Given a probability measure ϕ, a function $a : \Theta \to \mathbb{R}$ is $(N-1)$*–additively separable* if there exist functions $\alpha_i : \Theta_{-i} \to \mathbb{R}$, $i = 1,\ldots,N$, such that $a(\theta) = \sum_i \alpha_i(\theta_{-i})$ for ϕ-almost all θ.

Theorem 3. *In the separable utility environment, a game is ex post unbalanced if and only if, for all efficient outcome functions x^*, the aggregate EPIC payment function $R(\cdot;x^*)$ is not $(N-1)$–additively separable.*

This theorem implies that in order to determine whether a game is ex post unbalanced, it is sufficient to examine the EPIC payment functions for the efficient outcome functions and identify whether any is additively separable. The proof, in Appendix B, relies crucially on the fact that an optimal mechanism always exists in the separable utility environment, which is a consequence of Proposition 6 in Appendix A. The basic idea is simple: if $R(\cdot;x^*)$ is $(N-1)$–additively separable then it is possible to construct a mechanism with an ex post balanced budget, and otherwise it is not possible. The subtlety in the proof arises from accounting for potential problems on sets of ϕ-measure zero.[25]

To see that interesting games in the separable utility environment are generally ex post unbalanced, notice that $R(\cdot;x)$ depends on π^θ, which is not restricted to be $(N-1)$–additively

[24] It is important to distinguish the concept of $(N-1)$–additive separability, which refers to the functional forms of $R(\cdot;x)$ and $\sum_i y_i^0$, from the concept of (multiplicatively) separable utility, which refers to the functional form of $\pi_i(\theta, x(\theta))$ in the separable utility environment.

[25] In the private valuations case, and under the assumption that there exists an efficient outcome function that is differentiable in θ, Laffont and Maskin (1980, Theorem 4.1) go one step further to give conditions on the outcome function that guarantee exact ex post balance. In the setting considered here, a similar result could be obtained by imposing the differentiability assumption, applying integration by parts (in reverse) to Eq. 13, and noting the consequences of the aggregate EPIC payments being $(N-1)$–additively separable. Note, however, that differentiability rules out many of the most interesting cases.

separable.[26] Even in games with independent private valuations (i.e., π_i^θ depends only on θ_i, so that π^θ is by construction $(N-1)$–additively separable), $\pi^x(x(\theta))$ is generally not $(N-1)$–additively separable when efficiency involves non-trivial comparisons across players, as shown by Hurwicz and Walker (1990). For example, in a game of allocating an object to one of the players, an efficient outcome function x^* gives the object to the player with the highest value of $\pi_i^\theta(\theta)$—an assignment for which $\pi^x(x^*(\theta))$ is generally not additively separable when $\{\pi_i^\theta\}$ have overlapping ranges.[27] For a second example, in a game of providing a public good, an efficient outcome function x^* provides the good when the sum $\sum_i \pi_i^\theta(\theta)$ is positive; this yields a $\pi^x(x^*(\theta))$ that is generally not $(N-1)$–additively separable when $\{\pi_i^\theta\}$ have non-trivial ranges.[28] In a more abstract sense, games that are "smooth" are ex post unbalanced, as shown in the following proposition (proven in Appendix B).

Proposition 1. *In the separable utility environment, let $\mathcal{D} \equiv \left\{\pi^x(\chi) : \chi \in \mathcal{X}\right\}$. A game is ex post unbalanced if both*

(i) $\mathcal{D} \subset \mathbb{R}^N$ is a smooth, compact, convex set with non-empty interior;

(ii) For any subset $\overline{\Theta} \subset \Theta$ with $\phi(\overline{\Theta}) = 1$, $\left\{\pi^\theta(\theta) : \theta \in \overline{\Theta}\right\}$ spans $\mathbb{R}^N$.

On the other hand, several types of games are immediately indentifiable as "ex post balanced" (that is, not ex post unbalanced). Intuitively, the properties that make a game ex post balanced are properties that trivialize either the problem of balancing the budget or the problem of providing incentives, so in this sense ex post balanced games are "less interesting" than ex post unbalanced games. For an example, as suggested in the introduction, if one player does not have private information, then she can insure the other players against budget imbalances. This is commonly the case in auctions when the auctioneer has a fixed valuation of zero. A second example is when the efficient outcome does not vary with θ; then no EPIC payments need be made. This is the case in pure common value auctions, in which it does not matter which player wins the object because all players value it equally.[29] Third, if payoffs are identical across players—that is, what is best for any one is best for all—then the game is isomorphic to a one-player decision problem, and truthful revelation

[26] However, ex post unbalancedness is a non-generic property in the larger space of games with multi-dimensional signals. Jehiel, Meyer-ter-Vehn, Moldovanu, and Zame (2005) show that in this space EPIC allocation rules are generically trivial, and are thus implementable with ex post budget balance.

[27] Consider an allocation game (see Section 6.3 for definition) with one indivisible object and $\pi_i^\theta(\theta) = \theta_i$: For an efficient allocation rule, find a nonzero Nth cross-difference at $\theta_1 \in \{1-\varepsilon, 1+\varepsilon\}$ and $\theta_j \in \{1, 1+j\varepsilon\}$ for $j = 2, \ldots, N$, where $\varepsilon > 0$. At these points, whether player 1 is pivotal at θ_1 depends on θ_2, whether θ_2 makes a difference as to whether player 1 is pivotal depends on θ_3, and so on.

[28] Consider an indivisible public good with valuations $\pi_i^\theta(\theta) = \theta_i$: for an efficient provision rule, find a nonzero Ith cross-difference at $\theta_1 \in \{-\varepsilon, \varepsilon\}$ and $\theta_j \in \{0, j\varepsilon\}$ for $j = 2, \ldots, N$, where $\varepsilon > 0$.

[29] This is also the case in allocation games when the players' valuation ranges do not overlap, so that the same player always has the highest valuation.

may be taken for granted. These cases are formalized in the following proposition. (The proof is obvious and omitted.)

Proposition 2. *A game in the separable utility environment is ex post balanced if any of the following is satisfied:*

(i) There exists some player $i \in \mathcal{N}$ such that Θ_i is is a singleton;

(ii) There exists an efficient outcome function x^* such that $x^*(\theta) = x^*(\theta')$ for all $\theta \in \Theta$ and all $\theta' \in \Theta$;

(iii) For all $i \in \mathcal{N}$ and all $j \in \mathcal{N}$, for all $\chi \in \mathcal{X}$, and for all $\theta \in \Theta$, $\pi_i(\chi, \theta) = \pi_j(\chi, \theta)$.

For some types of games it is possible to construct knife-edge cases in which the players' incentive payments always balance under efficiency. Chung and Ely (2003a) give an example of this sort: a simple trading game between two players in which $x_i(\theta)$ is the probability that player i gets the object, and $\pi_i(\theta, x(\theta)) = (\theta_i + a\theta_{-i})\, x_i(\theta)$. The EPIC payments for an efficient outcome function sum to exactly zero for all θ if and only if $a = -1$. Laffont and Maskin (1980, Corollary 4.2) construct a larger class of games with three or more players and quadratic preferences in which the players' incentive payments exactly cancel out, but more general preferences lead to ex post unbalancedness.

5.2 Computation

In Section 3, the problem of designing an optimal EPPPE was successively simplified, first from a problem in the space of repeated game strategies to a dynamic programming problem in the space of recursive mechanisms, and thence to a static mechanism design problem. I now show that in the separable utility environment this static mechanism design problem can be solved by linear programming.

First, consider the problem when the outcome function x is fixed: the problem is then to find a *conditionally optimal* fixed transfer function $\sum_i y_i^{0*}(\cdot; x)$—i.e., one such that $\left\langle x, \sum_i y_i^{0*}(\cdot; x) - r(\cdot; x)\right\rangle$ is optimal among EPIC mechanisms that implement x. Formally, this is the following nonlinear program, which minimizes the surplus gap:

Definition 11 (Program NLP(x))**.**

$$\min_{\left\{\hat{y}_i^0 : \Theta_{-i} \to \mathbb{R}\right\}_{i=1}^N} \left[\max_\theta \left[\sum_i \hat{y}_i^0(\theta_{-i}) - R(\theta; x)\right]\right] \quad \text{s.t. } \mathbb{E}_\phi\left[\sum_i \hat{y}_i^0(\theta_{-i}) - R(\theta; x)\right] = 0. \qquad (14)$$

The definition and proposition below show that Program NLP(x) can be solved by a linear programming algorithm.

Definition 12 (Algorithm LP(x)). Let $\{\tilde{y}_i^{0*}(\cdot;x)\}_{i=1}^N$ solve

$$\max_{\left\{\tilde{y}_i^0:\Theta_{-i}\to\mathbb{R}\right\}_{i=1}^N} \left[\mathbb{E}_\phi\Big[\sum_i \tilde{y}_i^0(\theta_{-i})\Big]\right] \quad \text{s.t.} \ \sum_i \tilde{y}_i^0(\theta_{-i}) \le R(\theta;x) \ \forall\,\theta. \tag{15}$$

Then let

$$y_i^{0*}(\cdot;x) = \tilde{y}_i^{0*}(\cdot;x) - \frac{1}{N}\,\mathbb{E}_\phi\Big[\sum_i \tilde{y}_i^{0*}(\theta_{-i}) - R(\theta;x)\Big] \tag{16}$$

for each i. The output is $\{y_i^{0*}(\cdot;x)\}_{i=1}^N$.

Proposition 3. *Any output $y_i^{0*}(\cdot;x)$ from Algorithm LP(x) also solves Program NLP(x).*

The proof is in Appendix B. The intuition is that the regions of Θ on which the solution to Program NLP(x) attains a maximum of $\hat{Y}^0(\theta) - R(\theta;x)$ are the same regions of Θ on which the constraint is binding in the linear programming step of Algorithm LP(x). The lump sum adjustment in Eq. 16 must be made after solving the linear program in Eq. 15 because the size of the lump sum is determined by the solution.

For the same reasons, under one additional convexity assumption the problem of finding an (unconditionally) optimal mechanism can also be solved by an algorithm with a single linear program, as follows.

Assumption 2. *The set $\mathcal{C} \equiv \left\{\pi^x(\chi)\right\}_{\chi\in\mathcal{X}}$ is convex.*

Definition 13 (Algorithm LP*). Let x^{**} and $\{\tilde{y}_i^{0**}\}_{i=1}^N$ solve

$$\begin{aligned} &\max_{\pi^x(x(\cdot)):\Theta\to\mathcal{C},\ \{\tilde{y}_i^0:\Theta_{-i}\to\mathbb{R}\}_{i=1}^N} \mathbb{E}_\phi\left[\sum_i \pi_i^x(x(\theta))\pi_i^\theta(\theta) + \sum_i \tilde{y}_i^0(\theta_{-i}) - R(\theta;x)\right] \\ &\quad\text{s.t.}\ \begin{cases} \tilde{Y}^0(\theta) \le R(\theta;x)\ \forall\,\theta \\ \pi_i^x(x(\theta)) \ge \pi_i^x(x(\theta_i',\theta_{-i}))\ \forall\,\theta_i' \le \theta_i,\ \forall\,\theta,\ \forall\,i. \end{cases} \end{aligned} \tag{17}$$

Then let

$$y_i^{0**} \equiv \tilde{y}_i^{0**} - \frac{1}{N}\mathbb{E}_\phi\Big[\sum_i \tilde{y}_i^{0**}(\theta_{-i}) - R(\theta;x^{**})\Big] \tag{18}$$

for each i. The output is $\left\langle x^{**}, \{y_i^{0**}\}_{i=1}^N\right\rangle$.

Some notation: Let $y^*(\cdot;x) \equiv \sum_i y_i^{0*}(\cdot;x) - r(\cdot;x)$ be the conditionally optimal transfer function implied by the output of Algorithm LP(x); then $\langle x, y^*(\cdot;x)\rangle$ is a conditionally optimal mechanism given x, and for brevity I say that $\langle x, y^*(\cdot;x)\rangle$ "solves" both Algorithm LP(x) and Program NLP(x). Since $y_i^{0**} = y_i^{0*}(\cdot;x^{**})$ for all i, I also say that the optimal mechanism $\langle x^{**}, y^*(\cdot;x^{**})\rangle$ "solves" Algorithm LP*. The value of a conditionally optimal mechanism given x is $V^{0*}(x) \equiv \sum_i \mathbb{E}_\phi[\pi_i(\theta, x(\theta))] - \max_\theta[Y^*(\theta; x)]$, so the value of an optimal mechanism is $V^{0*}(x^{**}) = \sum_i \mathbb{E}_\phi[\pi_i(\theta, x^{**}(\theta))] - \max_\theta[Y^*(\theta; x^{**})]$.

Although Algorithms LP(x) and LP* involve continua of variables and constraints in the separable utility environment, they suggest a simple method for computing approximations on a discretized signal space.[30] Section 6 develops general characterizations for the solutions of both Algorithms LP(x) and LP* as applied to the special case of allocation games. Section 6 also describes an approximate version of Algorithm LP* numerically for allocation games with a range of probability distributions and utility functions. Furthermore, the separable utility environment could easily be extended to games with finite signal and action spaces, on which these programs would have finite numbers of variables and constraints.

It is interesting to note that the maximization problems in Eq. 15 and Eq. 17 can be interpreted as static mechanism design problems in which there is a "no-subsidy" condition (i.e., $\tilde{Y}(\theta) = \tilde{Y}^0(\theta) - R(\theta; x) \leq 0$ for all θ). That is, when money burning is allowed the problem of designing an optimal EPPPE differs by a lump sum from the problem of designing a static EPIC mechanism when the mechanism designer is not allowed to provide any subsidy, and individual rationality is ignored.[31] In the static problem, money must be burnt to provide the correct incentives, and moving from the static problem to the repeated game does not alleviate this restriction.

6 Two-player allocation games

In an allocation game, there is some rival good that must be split among the players. Leading examples include auctions, in which some objects or benefits must be distributed; trade, in which players can exchange goods for money; and collusion, in which firms must divide a market. To show the process of designing an optimal (EPPPE) mechanism, I begin with a detailed example, which displays a number of interesting properties. Then I explore the generality of these properties in both analytically and numerically. These properties are, in decreasing order of generality:[32]

1. *An optimal mechanism exists.* (Proposition 6, Appendix A)

2. *It is impossible to implement an EPIC-efficient outcome function in every period unless money is burnt.* (Theorem 3)

[30] Simple computer programs for solving Algorithms LP(x) and LP* for two-player games are available on request from the author. They require *Mathematica*, version 5.0 or higher, from Wolfram Research, Inc.

[31] The lump sum difference accounts for the fact that the equivalent static transfer function subtracts out the mean of the promised future utility function.

[32] A similar set of properties applies to games with public goods as well. An addendum demonstrating this is available on my web site: http://dss.ucsd.edu/∼d9miller.

3. *Given an efficient outcome function x^*, the conditionally optimal fixed transfer function $y_i^{0*}(\cdot; x^*)$ for each player is monotonically increasing up to a cutoff point, and flat thereafter.* Proposition 4, below, shows that this is generally true in two-player allocation games, subject to some mild regularity conditions.

4. *The optimal mechanism involves partial "pooling."* That is, when both players' signals fall within a certain range, the outcome function is fixed and ignores their signals. Though I have not proven that this is true for all allocation games, it is the case in every example I have computed, across a wide range of allocation games with varying valuation functions (π^θ) and probability measures (ϕ).

5. *When valuations are interdependent, the optimal mechanism has a positive surplus gap.* That is, when both players' valuation functions depend on both θ_1 and θ_2, then an optimal mechanism requires either money burning or changes in aggregate continuation rewards.

6. *For many models with private valuations, the optimal mechanism has a zero surplus gap, and the optimal outcome function displays a "stair step" pattern.* A recursive mechanism with a zero surplus gap is stationary, since continuation rewards do not need to change in order to provide incentives. That is, the inefficiency implied by property 2 is achieved by choosing an inefficient outcome function in each period rather than by burning money. This indicates that any increase in allocative efficiency is very expensive in terms of the amount of money that must be burnt to support it. This property seems to hold in many, but not all, allocation games with private valuations.

The first main theoretical result of this section follows from a closer examination of property 4: Theorem 4 shows that, in two-player allocation games, an optimal mechanism never employs an efficient outcome function. Instead it is preferable for the players to give up some static efficiency in order to reduce the surplus gap. The proof operates by demonstrating that an efficient outcome function can always be improved by introducing a small amount of pooling in a region where both players' valuations are low.

The second main theoretical result shows that property 5 is general: Theorem 5, below, shows that under interdependent valuations it is always optimal to use a mechanism that requires money burning or changes in continuation rewards. This is particularly noteworthy in light of property 6—that under private valuations the optimal surplus gap is often zero.

A third result is conjectured: that property 6 holds for all allocation games with symmetric private valuations. Section 6.2 offers some suggestive evidence for this conjecture.

This section makes much use of the (equivalent static) transfer function, $y : \Theta \to \mathbb{R}^N$, so a review of its properties is in order. First, recall its definition:

$$y(\theta) \equiv t(\theta) + \frac{\delta}{1-\delta}\big(w(\theta) - \mathbb{E}_\phi[w(\theta)]\big). \tag{19}$$

By construction, the aggregate transfer function $Y(\theta) \equiv \sum_i y_i(\theta)$ satisfies ex ante budget balance: $\mathbb{E}_\phi[Y(\theta)] = 0$. In the separable utility environment, for each player i, y_i can be decomposed into the EPIC payment function $r_i(\cdot; x)$, which is pinned down by EPIC, and the fixed transfer function y_i^0, which does not vary with θ_i so as to avoid disrupting the incentives:

$$y_i(\theta) = y_i^0(\theta_{-i}) - r_i(\theta; x). \tag{20}$$

Given an outcome function x, the problem of designing a conditionally optimal mechanism (Algorithm LP(x) of Section 5.2) is to choose a fixed transfer function that minimizes the surplus gap, $\max_\theta[Y(\theta)] = \max_\theta[\sum_i y_i^0(\theta) - R(\theta; x)]$. A solution to this problem is a set of conditionally optimal fixed transfer functions $\{y_i^{0*}(\cdot; x)\}_{i\in\mathcal{N}}$, from which the conditionally optimal transfer function $y^*(\cdot; x) = \big(y_i^{0*}(\cdot; x) - r_i(\cdot; x)\big)_{i\in\mathcal{N}}$ is constructed.

6.1 Example: Linear private valuations, uniform distribution

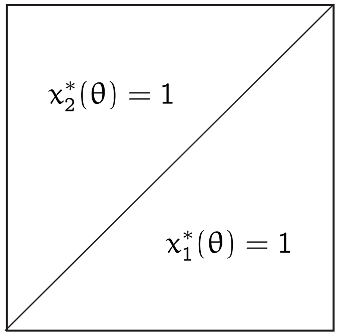
To illustrate the process of designing an EPPPE mechanism for an allocation game, consider an example in which two players must allocate one indivisible object each period. They have independent private valuations, each of which is distributed uniformly on the interval $[0, 1]$. An outcome function x, or "allocation rule," maps from the vector of private valuations to a probability distribution over who gets the object. Let $x_i(\theta)$ indicate the probability that player $i \in \{1, 2\}$ gets the object given a pair of valuations $\theta = (\theta_1, \theta_2)$. The players have separable utility of the form $\pi_i^\theta(\theta)\, \pi_i^x(x(\theta)) = \theta_i\, x_i(\theta)$.

Efficient allocation rule in the first period One approach to designing an EPPPE mechanism is to start with an efficient allocation rule x^* in the first period. This approach is exhibited in Figure 1. The allocation rule, which assigns the object to the player with the highest valuation (ignoring ties), is shown in Figure 1A. The resulting aggregate welfare is

$$\sum_i \mathbb{E}_\phi[\pi_i(\theta, x^*(\theta))] = \int_0^1 \int_0^1 \max_i[\theta_i]\, d\theta_1 d\theta_2 = \frac{2}{3}. \tag{21}$$

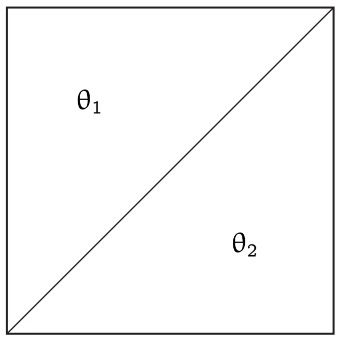
Figure 1B shows the aggregate EPIC payment function, $R(\cdot; x^*)$. Since each player's

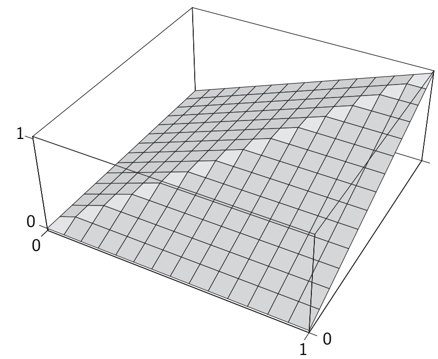
A. Efficient allocation $x^*(\theta)$



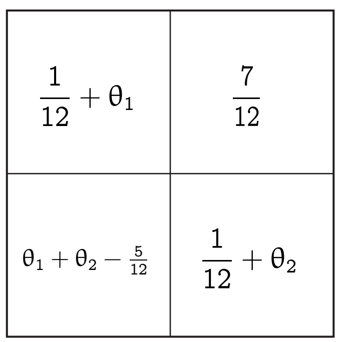
Each graph shows the signal space. The 2-dimensional graphs show player 1's signal on the horizontal axis and player 2's signal on the vertical axis. The 3-dimensional graphs show player 1's signal on the foreground axis and player 2's signal on the background axis.

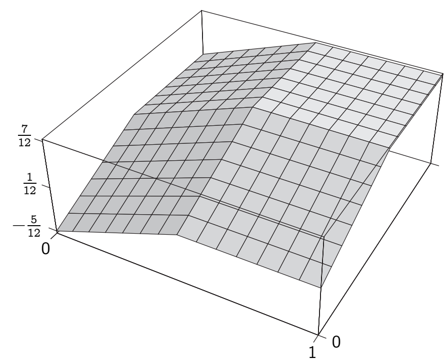
B. Aggregate EPIC payments $R(\theta; x^*)$



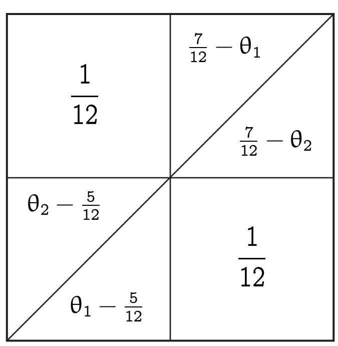


C. Aggregate fixed transfers $Y^{0*}(\theta; x^*)$





D. Aggregate transfers $Y^*(\theta; x^*)$ $= Y^{0*}(\theta; x^*)$ $- R(\theta; x^*)$



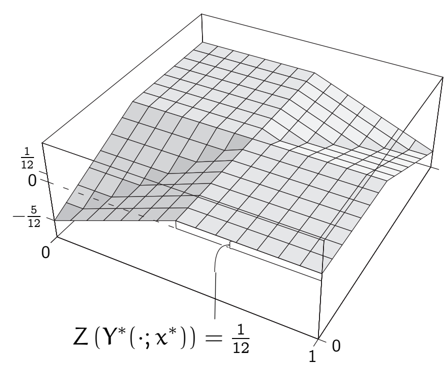


FIGURE 1: FIRST PERIOD EFFICIENCY IN THE ALLOCATION EXAMPLE

EPIC payment for this allocation rule follows the familiar second price auction,

$$R(\theta; x^*) = r_1(\theta; x^*) + r_2(\theta; x^*) = \min_i[\theta_i]. \tag{22}$$

Note that $R(\cdot; x^*)$ is not $(N-1)$–additively separable. Then, by Theorem 3, the game is ex post unbalanced, so the surplus gap for an efficient allocation rule must be positive.

Theorem 2 implies that the goal should be to maximize aggregate welfare minus the surplus gap. Since aggregate welfare and the EPIC payments are already fixed by the choice of x^*, the first task is to find fixed transfer functions that minimize the surplus gap. An aggregate solution $\sum_i y_i^{0*}(\cdot; x^*)$ is shown in Figure 1C. Notice how $\sum_i y_i^{0*}(\cdot; x^*)$ mimics the slope of $R(\cdot; x^*)$ in the quadrants off the diagonal; the $(N-1)$–additive separability of $\sum_i y_i^{0*}(\cdot; x^*)$ prevents it from also matching $R(\cdot; x^*)$ in the quadrants on the diagonal.

Figure 1D shows the aggregate transfer function, $Y^*(\cdot; x^*) = \sum_i y_i^{0*}(\cdot; x^*) - R(\cdot; x^*)$. By Lemma 1, the surplus gap is equal to the maximum of the aggregate transfer function, $\max_\theta[Y^*(\theta; x^*)] = \frac{1}{12}$. By Theorem 2, the value of the mechanism is

$$V^0 = \sum_i \mathbb{E}_\phi[\pi_i(\theta, x^*(\theta))] - \max_\theta[Y^*(\theta; x^*)] = \frac{7}{12}. \tag{23}$$

Given a discount factor δ, to transform $\langle x^*, y^*(\cdot; x^*)\rangle$ into the first stage of an EPPPE mechanism we must decompose the equivalent static transfer function as follows:

$$Y^*(\cdot; x^*) = \sum_i \Big(t_i(\cdot; v^0) + \frac{\delta}{1-\delta}\big(w_i(\cdot; v^0) - \mathbb{E}_\phi[w_i(\theta; v^0)]\big)\Big) \tag{24}$$

There are many ways to achieve this decomposition, but all must satisfy two constraints: (i) $\sum_i t_i(\theta; v^0) = 0$ for all θ and (ii) $(1-\delta)\,\frac{2}{3} + \delta \sum_i \mathbb{E}_\phi[w_i(\theta; v^0)] = \frac{7}{12}$. If money burning is allowed, then the continuation reward $w_i(\theta; v^0)$ can be delivered by having each player burn the appropriate amount of money, and then continuing the game with the initial promised utility v^0. If money burning is not allowed, then it is necessary to design additional stage mechanisms that can deliver any of the promised continuation rewards.

Optimal allocation rule Construction of the optimal mechanism is shown in Figure 2. For this example, the solution is to "pool" the allocation on regions where the players have similar valuations, as plotted in Figure 2A. "Pooling" refers to the idea that, conditional on

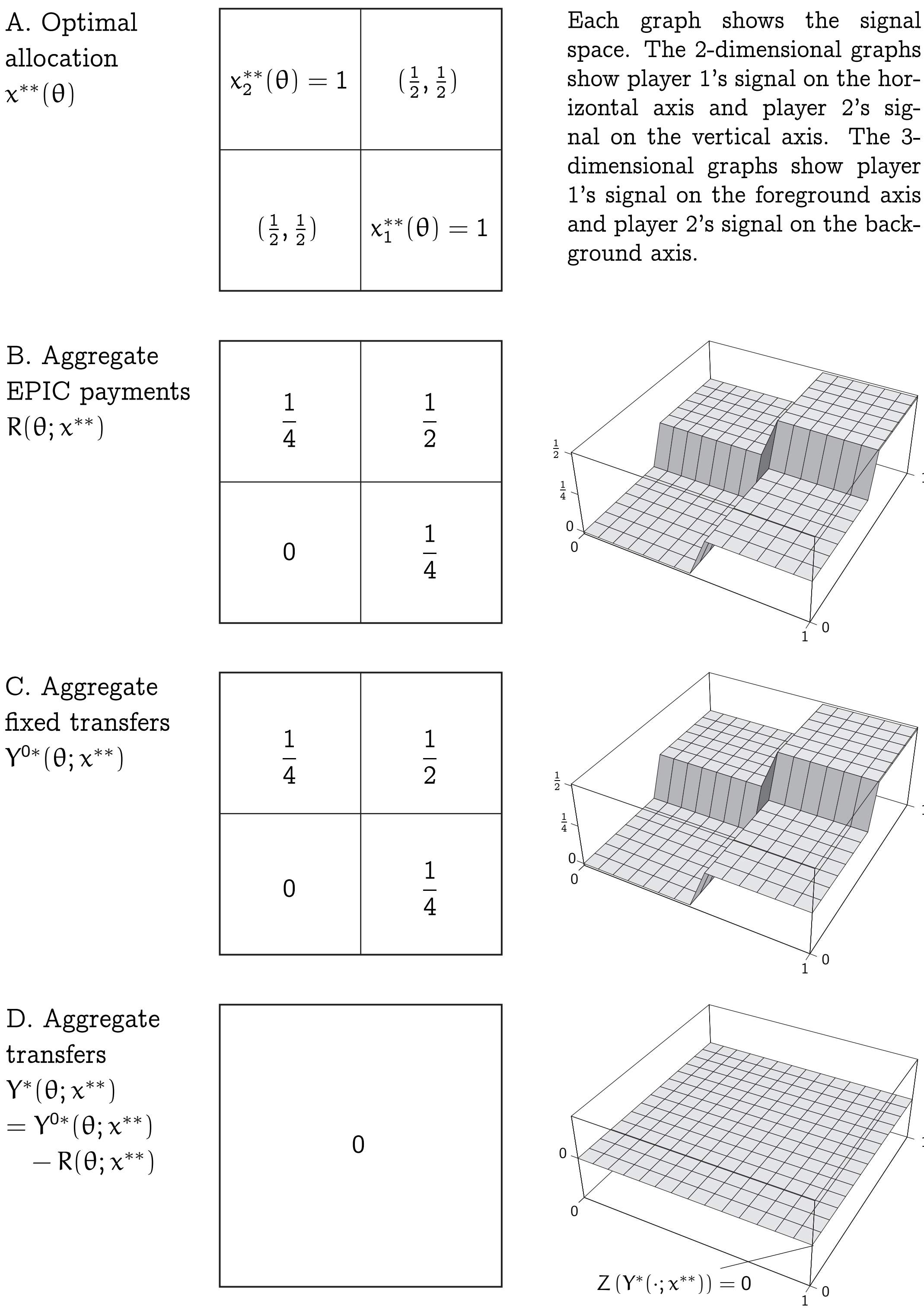


FIGURE 2: OPTIMALITY IN THE ALLOCATION EXAMPLE

θ falling within a particular rectangular region, the allocation ignores the players' signals:

$$x^{**}(\theta) = \begin{cases} (\frac{1}{2}, \frac{1}{2}) & \text{if } \theta_1 < \frac{1}{2} \text{ and } \theta_2 < \frac{1}{2}, \\ (\frac{1}{2}, \frac{1}{2}) & \text{if } \theta_1 \geq \frac{1}{2} \text{ and } \theta_2 \geq \frac{1}{2}, \\ x^*(\theta) & \text{otherwise.} \end{cases} \tag{25}$$

This allocation rule forms a rectangular "stair step" pattern when plotted in three dimensions. The aggregate welfare provided by this allocation rule is $\sum_i \mathbb{E}_\phi[\pi_i(\theta, x^{**}(\theta))] = \frac{5}{8}$. As for the EPIC payments, when $\theta_i < \frac{1}{2}$, player i wins the object with probability that depends only on θ_{-i}, so her EPIC payment is zero. When $\theta_i \geq \frac{1}{2}$, she wins the object with probability that is $\frac{1}{2}$ greater than when $\theta_i < \frac{1}{2}$, so she must pay an EPIC payment of $\frac{1}{4}$; this discourages her from claiming to have a valuation greater than $\frac{1}{2}$ when her valuation is actually less than $\frac{1}{2}$. Thus the aggregate EPIC payments, as plotted in Figure 2B, are:

$$R(\theta; x^{**}) = \begin{cases} 0 & \text{if } \theta_1, \theta_2 < \frac{1}{2}, \\ \frac{1}{2} & \text{if } \theta_1, \theta_2 \geq \frac{1}{2}, \\ \frac{1}{4} & \text{otherwise.} \end{cases} \tag{26}$$

Observe that $R(\cdot; x^{**})$ is $(N-1)$–additively separable. Thus $\sum_i y_i^{0*}(\cdot; x^{**}) = R(\cdot; x^{**})$, as shown in Figure 2C, and the surplus gap is zero (Figure 2D). The value of the mechanism is $V^{0*}(x^{**}) = \frac{5}{8}$, which exceeds the value of $\frac{7}{12}$ achieved by the efficient allocation rule analyzed above.

This optimal solution is easily transformed into an EPPPE mechanism by setting the stage game monetary transfer function to

$$t_i(\theta; v^0) = y_i^{0*}(\theta_{-i}; x^{**}) - r_i(\theta; x^{**}). \tag{27}$$

We can be sure that $\sum_i t_i(\theta; v^0) = 0$ because the surplus gap is zero. Then the continuation reward function is simply $w(\theta; v^0) = v^0$ for all θ; i.e., the recursive mechanism is stationary.

6.2 A range of examples

In order to gain a broader picture of optimal EPPPE mechanisms in two-player allocation games, I computed Algorithm LP* for a variety of probability distributions and pairs of valuation functions using a 14×14 discrete signal space. One set of calculations considers combinations of eight probability distributions (including the uniform distribution and eight truncated independent bivariate normal distributions with the same variance but different means) with five pairs of valuation functions (including symmetric and asymmetric private valuations, symmetric partial common valuations, and one combination in which one player has private valuations while the other has partial common valuations). A second set of calculations considers forty pairs of symmetric private valuations, under the uniform distribution. A third set of calculations considers eight randomly generated distributions, with five pairs of symmetric private valuations.

The first set of computed optimal allocation rules is shown in Figure 3. The probability distribution is displayed in the left column, while the valuation functions are listed across the top. Each graph displays $\theta_1 \in [0,1]$ on the foreground axis and $\theta_2 \in [0,1]$ on the background axis. The vertical axis gives player 2's portion of the optimal allocation, $x_2^{**}(\theta)$; i.e., a value of $x_2^{**}(\theta) = 0$ on the vertical axis implies that $x_1^{**}(\theta) = 1$, so player 1 gets the entire object.[33]

The most general regularity in Figure 3 is that none of the optimal mechanisms implement efficient allocation rules. That this holds for all two-player allocation games is proven in Theorem 4, below, which shows that an efficient allocation rule can always be improved by introducing a small amount of pooling "at the bottom" (when both players have low valuations).

A second regularity in Figure 3 is that the optimal mechanisms all pool at the bottom. This matches well with the intuition of Theorem 4, which is that pooling is a good strategy for reducing the surplus gap without giving up much efficiency. However, I also computed a similar range of examples for team decision games (the results are available in an addendum on my web site), and found examples (with interdependent valuations) in which the optimal mechanism does not pool at all. This suggests that pooling may not always be optimal, since team decision games share many similarities with allocation games. It is an object of ongoing research to find conditions under which pooling is, or is not, optimal.

A closer examination of Figure 3 reveals an even more striking, if somewhat less general, regularity. Excluding column 5 (in which both players have interdependent valuations), nearly all the examples exhibit a zero surplus gap. A zero surplus gap can be recognized by a stair step allocation rule—such as any example in column 1. The cases in which the surplus gap is not zero, such as example D.2 or any example in the fifth column, display allocation rules that are more rounded, and which sometimes split the object between the two players. In all cases, the surplus gap is quite small compared to the inefficiency of the allocation (example E.5 exhibits the highest ratio of money burning to allocative inefficiency at $\max_\theta[Y^{**}(\theta)]/\sum_i \mathbb{E}\left[\pi_i(\theta, x^*(\theta)) - \pi_i(\theta, x^{**}(\theta))\right] \approx 0.01$). That the surplus gap is so small at the optimum indicates that a small increase in allocative efficiency is very expensive, in terms of the amount of money that must be burnt to support it, even when the overall allocation is very far from efficient. This suggests that the logic of Theorem 4—that an

[33]It is inconvenient to graph x_1^{**} because the height of the surface in the foreground would obscure important features in the background.

efficient allocation can always be improved upon by pooling—applies also to many inefficient allocations. On the other hand, Theorem 5, below, shows that when the players have interdependent valuations—as in the fifth column—an optimal mechanisms always features a positive surplus gap.

These examples raise the possibility that it is common for a zero surplus gap to be optimal, particularly under symmetric private valuations. I am currently working to prove the conjecture that in allocation games with symmetric private valuations, a zero surplus gap is optimal, subject to some regularity conditions. For now, in lieu of proof, I offer the second and third sets of computations as suggestive evidence. The second set of computed optimal allocation rules is shown in Figure 4. The purpose of this set of computations is to investigate whether allocation games with symmetric private valuations optimally have zero surplus gaps under the uniform distribution. The third set of computed optimal allocation rules is shown in Figure 5.[34] The purpose of this set of computations is to investigate whether allocation games with symmetric private valuations optimally have zero surplus gaps regardless of the distribution. In both exercises, since each optimal allocation rule in the figure displays the stair-step pattern, no counterexample to the conjecture was generated.

6.3 General allocation games

Definition 14. A game in the separable utility environment is an *allocation game* if:

(i) $\mathcal{X} = \{\chi \in [0,1]^N : \sum_i \chi_i \leq 1\}$,

(ii) $\pi_i^x(\chi) = \chi_i$ for all χ and all i,

(iii) $\pi_i^\theta(\theta) \in [0,1]$ with $\pi_1^\theta(0,0) = \pi_2^\theta(0,0)$.

In this setup, π_i^θ is player i's valuation function. Part (iii) is merely a normalization. I impose the following assumption for convenience. Parts (i) and (iii) are mild regularity conditions; part (ii) requires monotonicity and that own signals be more important than others' signals, as is commonly considered in the literature.

Assumption 3. *There exists $B \in (1,\infty)$ such that, for all θ, for all $j \neq i$, and for all i,*

(i) $\pi_i^\theta(\theta)$ is twice continuously differentiable with $\frac{1}{B} < \frac{\partial \pi_i^\theta(\theta)}{\partial \theta_i} < B$;

(ii) $0 \leq \frac{\partial \pi_i^\theta(\theta)}{\partial \theta_j} < \frac{\partial \pi_i^\theta(\theta)}{\partial \theta_i}$ and $\frac{\partial \pi_i^\theta(\theta)}{\partial \theta_j} < \frac{\partial \pi_j^\theta(\theta)}{\partial \theta_j}$;

(iii) ϕ is a twice continuously differentiable probability density with $\frac{1}{B} < \phi(\theta) < B$.

[34] Due the computational load of the calculations, some of the examples use a 12×12 or 10×10 grid.

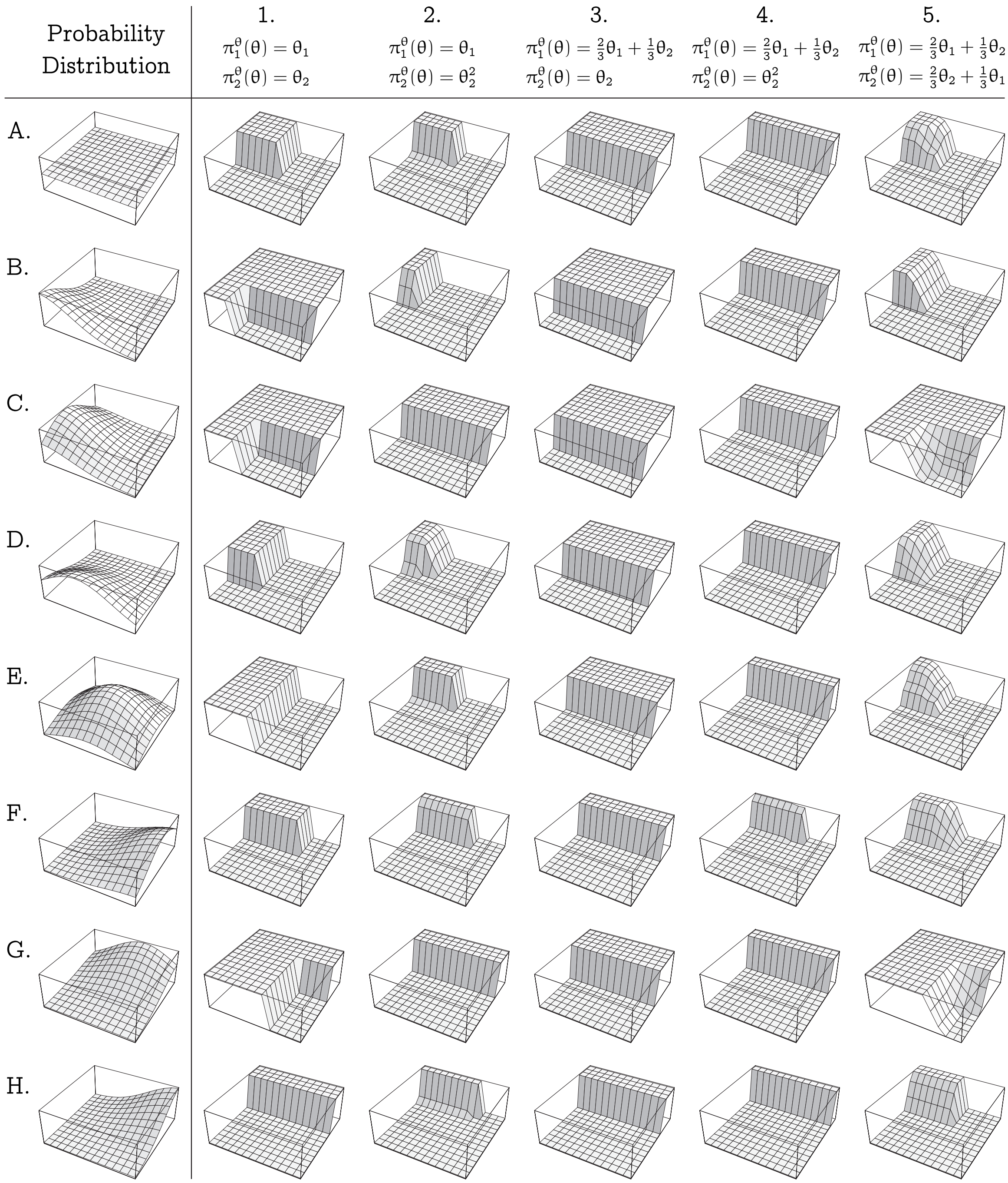


Figure 3: Computed optimal outcome functions for allocation games

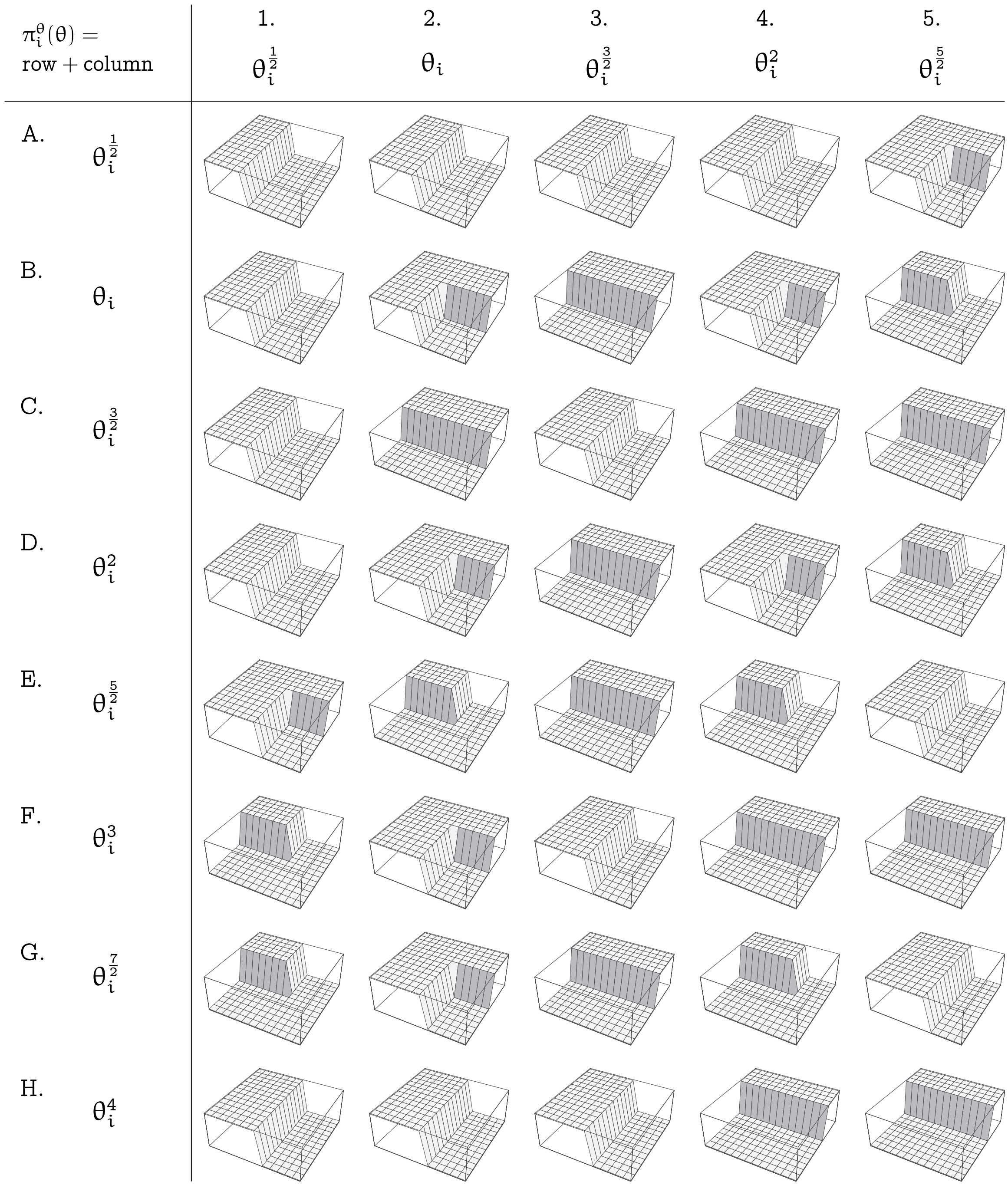


Figure 4: Computed optimal outcome functions for allocation games with symmetric private valuations and the uniform distribution

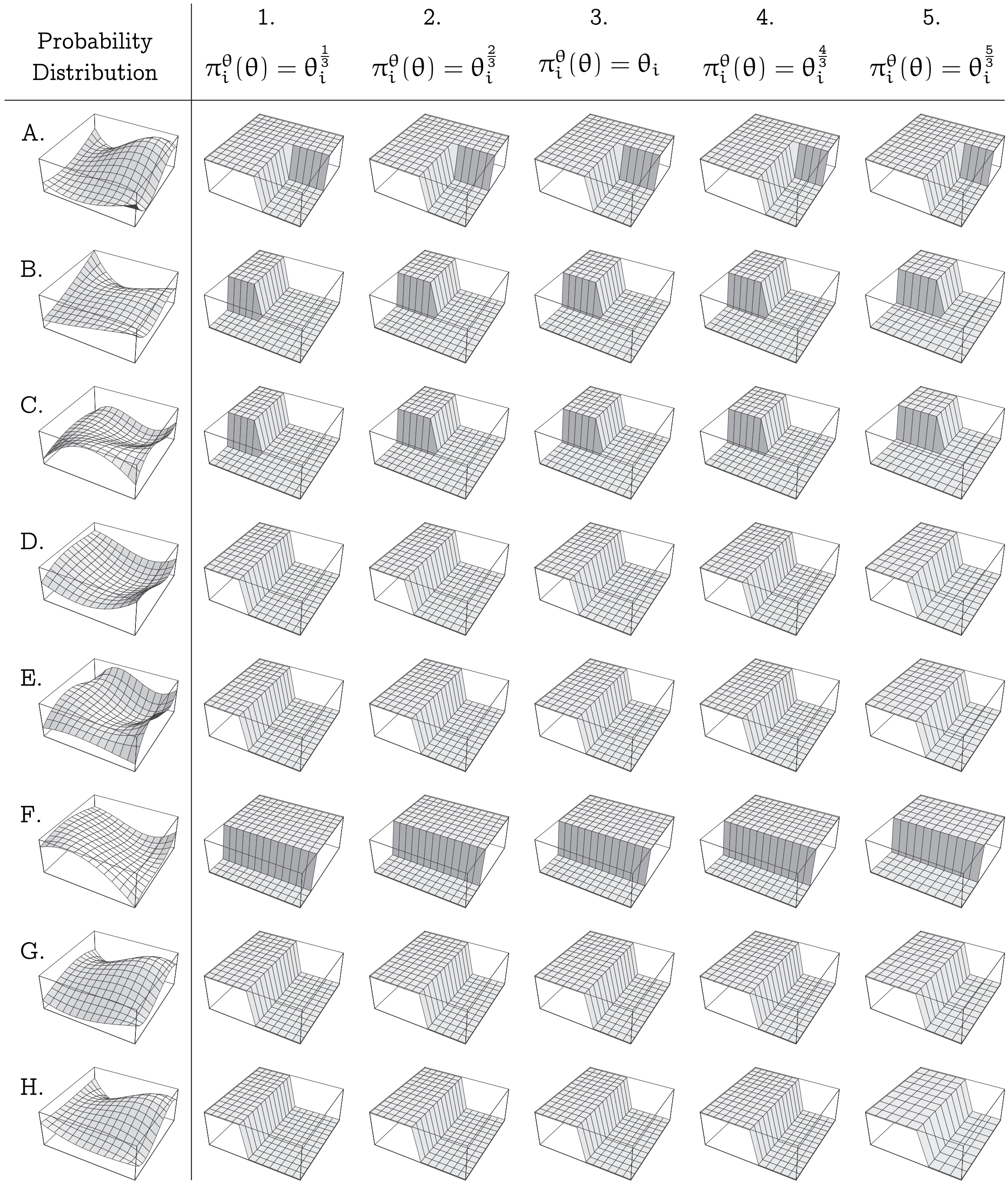


FIGURE 5: COMPUTED OPTIMAL OUTCOME FUNCTIONS FOR ALLOCATION GAMES WITH SYMMETRIC PRIVATE VALUATIONS AND RANDOMLY GENERATED DISTRIBUTIONS

6.3.1 Implications of efficiency

An allocation rule x^* is efficient if and only if $x_i^*(\theta) = 0$ for all $i \notin \arg\max_j[\pi_j^\theta(\theta)]$, for ϕ-almost all θ; i.e., if x^* allocates the object to a player with a highest valuation. The EPIC payments for an efficient allocation take the form of a Vickrey-Clarke-Groves (VCG) mechanism:[35] Let $\lambda_i^*(\theta_{-i})$ be the solution to $\pi_i^\theta(\lambda_i, \theta_{-i}) = \pi_{-i}^\theta(\lambda_i, \theta_{-i})$; that is, given θ_{-i}, $\lambda_i^*(\theta_{-i})$ is the signal for player i at which she is pivotal. Then the aggregate EPIC payments are

$$R(\theta; x^*) \equiv \sum_i r_i(\theta; x^*) = \sum_i \mathbb{I}[x_i^*(\theta) = 1]\, \pi_i^\theta(\lambda_i^*(\theta_{-i}), \theta_{-i}). \tag{28}$$

Given an allocation x, recall that $\{y_i^{0*}(\cdot; x)\}_{i\in\mathcal{N}}$ are the conditionally optimal fixed transfer functions that solve Algorithm LP(x), and $Y^*(\cdot; x) \equiv \sum_i y_i^{0*}(\cdot; x) - R(\cdot; x)$ is the conditionally optimal aggregate transfer function. The following proposition shows that in a two-player allocation game, subject to some mild regularity conditions, for any efficient allocation rule x^* each player i's conditionally optimal fixed transfer function $y_i^{0*}(\cdot; x^*)$ is increasing in θ_{-i} up to a cutoff point, and flat thereafter.

Proposition 4. *In a two-player allocation game under Assumption 3, for any efficient decision rule x^* there exists a conditionally optimal fixed transfer function $\sum_i y_i^{0*}(\cdot; x^*)$, a point b on the interior of Θ, and a pair of constants $a_1 \in \mathbb{R}$ and $a_2 \in \mathbb{R}$, such that*

(i) $\pi_1^\theta(b) = \pi_2^\theta(b)$;

(ii) $y_i^{0*}(\theta_{-i}; x^*) = a_i - \mathbb{I}\big[\theta_{-i} \le b_{-i}\big]\big(\pi_i(b) - \pi_i^\theta(t_i^*(\theta_{-i}), \theta_{-i})\big)$;

(iii) $\max_\theta[Y^*(\theta; x^*)] = a_1 + a_2$.

Furthermore, for all $\hat\theta \in [b_1, 1]\times[0, b_2]\cup[0, b_1]\times[b_2, 1]$, the aggregate transfer function attains the surplus gap: $Y^*(\hat\theta; x^*) = \max_\theta[Y^*(\theta; x^*)]$.

The proof, in Appendix B, operates by identifying the constraints that bind in Algorithm LP(x^*). In particular, given any $y_{-i}^0(\theta_i)$ function, for each θ_{-i} there are $|\Theta_i|$ constraints on $y_i^0(\theta_{-i})$, but only two are of any concern: the constraint that binds when θ_i is in the region in which player i gets the object and another that binds on the region in which player $-i$ gets the object. One of them binds for $\theta_{-i} \in [0, b_{-i}] \subset \Theta_{-i}$ while the other binds for $\theta_{-i} \in [b_{-i}, 1] \subset \Theta_{-i}$.

The main result of this section is that an efficient allocation is never optimal. An efficient allocation leaves a large surplus gap because its aggregate EPIC payment function is not $(N-1)$–additively separable. It is always possible to improve on an efficient allocation,

[35] Vickrey (1961); Clarke (1971); Groves (1973).

giving up some allocative efficiency in order to reduce the amount of money that must be burnt.

Theorem 4. *In a two-player allocation game under Assumption 3, an efficient allocation rule cannot be optimal.*

The proof improves on an efficient allocation rule by "pooling at the bottom": for a small rectangle close to the origin, the allocation rule always assigns the object to one of the players, without regard to θ. Proposition 4 implies that changes to R outside the pooling region can be offset by changes in $\{y_i^0\}_{i\in\mathcal{N}}$, so the proof proceeds by calculating the first-order changes in welfare, R, and $\{y_i^0\}_{i\in\mathcal{N}}$ on the pooling region. In particular, the benefit of pooling is that all EPIC payments are eliminated on the pooling region, and if the region is small enough this entire effect goes toward reducing the surplus gap. The cost is that some efficiency is given up. For a sufficiently small region of pooling, the benefit is greater than the cost.

6.3.2 Implications of a zero surplus gap

In the numerical examples discussed above, it is common for the surplus gap to be very small or even exactly zero in an optimal mechanism. A zero surplus gap implies that the aggregate EPIC payments $R(\cdot; x)$ must be additively separable, and we can use this fact to derive an important property: when players have interdependent valuations the surplus gap must be positive for an optimal mechanism. As discussed in Section 6.2, optimal mechanisms typically have very small surplus gaps, so this result is particularly important in that it provides a sufficient condition for a positive surplus gap to be optimal.

Theorem 5. *In a two-player allocation game under Assumption 3, suppose that both players have "globally interdependent valuations"; i.e.,* $\frac{\partial \pi_i^\theta}{\partial \theta_{-i}} > 0$ *for almost all* θ *and for all* i. *Then an optimal mechanism must have a strictly positive surplus gap.*

The proof, in Appendix B, begins with Lemma 3, which shows that under globally interdependent valuations the only way to achieve a zero surplus gap is to employ a constant outcome function. However, it is always preferable to increase efficiency by a small amount by allocating to player i when θ_i is very high and θ_{-i} is very low, even though this requires a positive surplus gap.

7 Application: Collusion with hidden costs

This section applies the abstract results generated thus far to the problem of collusion, yielding new insights into a problem that has been discussed in the literature: the "price wars" that are commonly observed under collusive arrangements.

Green and Porter (1984) note that standard full-information models of collusion normally yield optimal equilibria in which behavior along the equilibrium path is stationary and fully efficient, while in practice cartels typically exhibit intermittent periods in which they offer inefficiently low prices. Such episodes are called *price wars*, and can be explained in a model with hidden costs if EPPPE is used as the equilibrium concept.

Consider a simple model of collusion with hidden costs. There is a passive consumer who demands exactly one unit at any price up to and including 1. There are two firms, $i \in \{1, 2\}$, each with a cost function $c_i : \Theta \to [0, 1]$. In each period, the shocks θ_1 and θ_2 that affect the firms' costs are realized, with each firm i observing only θ_i. The distribution of $\theta \in [0, 1]^2$, which is identical and independent across periods, is given by the probability density ϕ. In each period, each firm sets a price $p_i \in \mathbb{R}_+$, and the consumer buys from the firm with the lower price. If both firms post the same price, then the market is split equally. The profit function for each firm is

$$\pi_i(\theta, (p, \alpha)) = \begin{cases} p_i - c_i(\theta) & \text{if } p_i < p_{-i} \text{ and } p_i \leq 1 \\ \frac{1}{2}(p_i - c_i(\theta)) & \text{if } p_i = p_{-i} \leq 1, \\ 0 & \text{otherwise.} \end{cases} \tag{29}$$

Note that this game is essentially an allocation game. In fact, if each firm i always sets a price $p_i \in \{1, 2\}$, then by randomizing publicly the firms can achieve any division of the expected monopoly profits. Hence the conclusions in Section 6 apply; namely, under EPPPE the firms should not seek to implement an efficient allocation, and, depending on the utility functions and the probability distribution of costs, they may need to burn money.

Proposition 5. *In collusion with hidden costs, there exists $\underline{\delta} < 1$ such that for all $\delta > \underline{\delta}$, if an optimal EPPPE mechanism has a strictly positive surplus gap, then there exists an optimal EPPPE with price wars on the equilibrium path.*

The proof is omitted; the idea is that since the firms face a positive surplus gap, after some first period realizations of their production costs they must receive less than optimal utility in the continuation game. They can accomplish this by burning money if such a technology is available; alternatively they can coordinate to set inefficiently low prices in the second period. Ultimately, it does not matter what form of inefficiency they employ, so considerations from outside the model may play a role in their selection. For instance, it might be difficult to convince shareholders that money burning is a good idea, while price wars can be blamed (falsely) on unobservable, exogenous market conditions.

The proposition requires a positive surplus gap, because a zero surplus gap would imply that the firms could use the same mechanism in every period, and no price wars would be necessary. The results discussed in Section 6.3 indicate that the surplus gap is positive in settings with interdependent valuations. In the context of hidden costs, interdependent valuations can be interpreted as representing an industry with underlying common shocks as well as idiosyncratic shocks, in which the shocks are observed imperfectly and privately, and such that the uncertainty is not resolved until after the communication stage.

Proposition 5 implies that price wars can occur in an optimal EPPPE even though firms can make side payments, and that if price wars occur then they do not disappear as the firms grow more patient. These properties contrast with previous explanations for price wars, all of which rely on either impatience or an inability to statistically detect the identity of a deviator. Green and Porter (1984) generate price wars among patient firms in a hidden action model, under a restriction to symmetric equilibria and in a situation in which the identity of a deviator cannot be statistically detected. But Fudenberg, Levine, and Maskin (1994) show that this result does not hold for asymmetric equilibria.[36] Sannikov and Skrzypacz (2004) generate price wars that worsen as the discount factor approaches unity even in asymmetric equilibria, in a model in which information arrives continuously and the discount factor is interpreted as a measure of how often the firms can adjust their production. However, for any fixed frequency of adjustment, price wars disappear as firms become more patient. Furthermore, Abreu, Pearce, and Stacchetti (1986) show that a richer information structure, in which deviators can be statistically identified, allows PPEs to approach efficiency at the limit as $\delta \to 1$. Abreu, Pearce, and Stacchetti do indicate that there will be periods of extremely inefficient pricing, due to the bang-bang nature of optimal equilibria, but this inefficient pricing occurs with vanishing probability as the firms become more patient.

Rotemberg and Saloner (1986) and Bagwell and Staiger (1997) generate intermittent periods of suboptimal pricing among collusive firms in an environment with exogenous public demand shocks, but the effect is driven by the individual rationality constraints, so it disappears as the firms become more patient.

Athey, Bagwell, and Sanchirico (2004) consider a model with hidden costs and independent private values, using IIC rather than EPIC and restricting attention to situations in which the identity of a deviator cannot be statistically identified. They demonstrate that optimal collusive equilibria do not exhibit price wars, but rather display excessive price rigidity and inefficient production (i.e., pooling). Athey and Bagwell (2001) add a richer information structure the same model so that deviators can be statistically detected, and show that price wars do not occur and that efficiency can be attained given sufficient patience. Athey and Bagwell (2004) then add persistence of cost shocks from one period to the next, and show that price wars do not occur. When cost shocks are perfectly persistent,

[36] Although Cheng (2004) notes that for quantity-setting models the Fudenberg, Levine, and Maskin (1994) argument relies on an unrealistic upper bound on production.

production is optimally inefficient regardless of δ, but when cost shocks are only partially persistent the firms can support first-best collusion if they are sufficiently patient.

Cartels that actively communicate and exchange side payments (or employ other sorts of asymmetric punishments)—i.e., a "strong cartel" in the terminology of McAfee and McMillan (1992)—are not restricted to symmetric equilibria, and so the Green and Porter (1984) theory cannot explain why strong cartels that are worried about secret price cutting (or secret overproduction) might engage in price wars. And if the cartel firms set prices (or quantities) often enough then impatience will not play a role since the profit to be made by deviating over any short period of time will be small, so that the Abreu, Pearce, and Stacchetti (1986), Fudenberg, Levine, and Maskin (1994), Rotemberg and Saloner (1986), and Bagwell and Staiger (1997) theories do not apply. Sannikov and Skrzypacz (2004) do generate price wars that are robust to asymmetric equilibria and frequent adjustment, but they rely on the same sort of observability restrictions as Green and Porter (1984). The Athey, Bagwell, and Sanchirico (2004), Athey and Bagwell (2001), and Athey and Bagwell (2004) theories, on the other hand, apply to situations with observable production and pricing but private cost information, indicate that price wars should not occur under IIC. EPIC, however, is a natural restriction on equilibrium in situations with private costs whenever the cartel communicates face to face or the members can engage in industrial espionage. Thus Proposition 5 provides a rationale for price wars in a situation to which the previous explanations do not apply: full observability of prices and quantities, private cost information, and communication.

8 Conclusion

This paper contributes new results regarding cooperation when there is private information and continued interaction over time. The paper introduces ex post perfect public equilibrium (EPPPE), a repeated game equilibrium concept that allows players to communicate in any order, so they do not need to hire or develop an institution to enforce simultaneity and prevent spying. When modeling applications in which such institutions are unavailable, infeasible, or expensive, using EPPPE rather than PPE yields predictions that may be more realistic. On the other hand, the first main result of this paper—that efficiency typically cannot be attained by any EPPPE—indicates that such institutions may be valuable. A second main result indicates that an optimal EPPPE typically needs to employ an inefficient outcome function, so implementing static efficiency is not always the appropriate goal in repeated interactions. Applied to the problem of collusion with hidden costs, these ideas offer new insights into the phenomenon of price wars.

We could further expand the example of collusion to cover repeated procurement auctions in which the auctioneer participates in the game as a strategic agent. The auctioneer may have private information about her valuation for the item to be procured, while the firms that bid to supply the item have private information about their production costs. Because all the players have payoff-relevant private information, the results of Section 6 suggest that they may be able to benefit from allocating production inefficiently. A model of this sort may lead to new conclusions about the optimal design of procurement auctions.

Similarly, EPPPE may yield new insights into principal-agent problems when both the principal and the agent are endowed with private information. For example, suppose that a principal and an agent experience private taste shocks over how to implement any particular project. (These preferences should not relate to the project's profitability—for instance, the agent may wish to polish skills that can give her a better outside option, while the principal may wish to have more control so as to earn more credit for any success.) Each privately learns her own taste shock, and then they must jointly decide on a method of implementation. We might at first guess that efficiency is the outcome of their optimal PPE; however, if the institutional setting is not strong enough to support simultaneity, then EPPPE is likely to lead to a more realistic conclusion that the optimal equilibrium is not efficient.

Two important questions regarding the characteristics of an optimal EPPPE remain unanswered. First, there is the question of when an optimal EPPPE always involves pooling. Theorem 4 shows that in two-player allocation games it is preferable to employ pooling compared to employing an efficient outcome function, but this does not imply that optimal mechanisms always employ pooling. Numerical results indicate that optimal mechanisms often employ pooling, but it remains to be seen whether conditions can be found that guarantee or rule out pooling in an optimal mechanism.

The second question concerns when an optimal mechanism employs a zero surplus gap. Theorem 4 indicates that in two-player allocation games, when the players have interdependent valuations, the optimal surplus gap is strictly positive. However, for games with private valuations, numerical results indicate that the optimal EPPPE commonly, but not universally, has a zero surplus gap. These results are suggestive, but it would be useful to find conditions under which the optimal mechanism features a zero surplus gap, in order to clarify intuition about the trade-off between money burning and allocative efficiency, as well as to simplify the problem of calculating an optimal mechanism.

Appendix A Existence of an optimum

As explained in Section 4, allowing free disposal simplifies the problem of designing optimal EPPPE mechanisms to a static mechanism design problem. This appendix demonstrates the existence of a solution to this problem under minimally restrictive assumptions. The existence result relies on a few simple assumptions and a general existence theorem due to Balder (1996) that does not impose any structure on Θ. The first two assumptions cast the outcome space $\mathcal{X}$ as the space of lotteries on a compact space of events $\mathcal{A}$. This implies the existence of a public randomization device.[37]

Assumption 4. *$\mathcal{X}$ is the space of probability measures on some compact metric space $\mathcal{A}$.*

Assumption 5. *$\pi_i(\theta,\chi) = \int_{\mathcal{A}} \pi_i^a(\theta,a)\,d\chi(a)$, where $\pi_i^a(\theta,\cdot)$ is bounded and continuous on $\mathcal{A}$ for all θ and all i.*

Assumption 6. *There exists a mechanism that satisfies EPIC.*

This final assumption is satisfied whenever there exists a fixed outcome $\chi \in \mathcal{X}$ such that $E_\phi[\pi(\theta,\chi)] \in \mathcal{P}$, since any individually rational fixed outcome can be implemented under EPIC without employing monetary transfers.

Proposition 6. *Under Assumptions 4–6, there exists a mechanism $\langle \hat{x}, \hat{y}\rangle$ that satisfies (i) EPIC and (ii) $\sum_i \mathbb{E}_\phi[\pi_i(\theta,\hat{x}(\theta))] - \max_\theta[\hat{Y}(\theta)] = \max\big[\sum_i \mathbb{E}_\phi[\pi_i(\theta,x(\theta))] - \max_\theta[Y(\theta)]\big]$, where the maximum is taken over all mechanisms $\langle x, y\rangle$ that satisfy EPIC.*

Proof. The proof follows Balder (1996) Example 3.3. Define the topology $\mathscr{T} \equiv \{\emptyset, \mathcal{N}\} \times \mathscr{F}$ on $\mathcal{N} \times \Theta$, where $\mathscr{F}$ is the topology for ϕ. Let $\mathcal{K} \equiv \mathcal{X} \times [-B,B]^N$, where $B \in \mathbb{R}_+$ is large relative to the uniform bound on π. Let $\hat{u}(i,\theta,(\chi,\psi)) \equiv \pi_i(\theta,\chi) + \psi_i$, $\hat{U}(i,\theta,(\chi,\psi)) \equiv \sum_j \hat{u}(j,\theta,(\chi,\psi))$, and $\Omega(i,\theta) \equiv \big\{(\chi,\psi) \in \mathcal{K} : \sum_j \psi_j \le Z^*\big\}$, where $Z^* \in \mathbb{R}_+$.

First, I establish that Balder's Assumptions 2.1–2.7 are satisfied. Assumption 4 implies that $\mathcal{X}$ is a compact, convex metric space (by Aliprantis and Border (1999), Theorem 14.11), so $\mathcal{K}$ is convex, metrizable, and compact, satisfying Balder's Assumption 2.1. Assumption 5 implies that $\pi_i(\theta,\cdot)$ is a linear function on $\mathcal{X}$ for all θ and all i, so that Balder's Assumption 2.3 is satisfied; it also implies that $\Omega(i,\theta)$ is a closed and convex-valued correspondence, satisfying Balder's Assumption 2.2. Although $\hat{u}(\cdot)$ is not $\mathscr{T}$-measurable, $\hat{u}(i,\cdot)$ clearly is, and so $\hat{U}(i,\theta,(\chi,\psi))$ is $\mathscr{T}$-measurable—satisfying Balder's Assumption 2.4. Since π is uniformly bounded, $\hat{U}$ is uniformly bounded as well, and so Balder's Assumption 2.5 is satisfied. The linearity of $\pi_i(\theta,\cdot)$ also implies that $\hat{U}$ is concave and continuous, satisfying Balder's Assumption 2.6. Finally, Assumption 6 is identical to Balder's Assumption 2.7.

Now apply Balder's Theorem 2.1, which implies that for any $Z^* \in \mathbb{R}$ there exists a mechanism $\langle \hat{x}(\cdot;Z^*), \hat{y}(\cdot;Z^*)\rangle$ that solves

$$\begin{gathered}\max_{\langle x,y\rangle}\Big[\mathbb{E}_\phi\big[\hat{U}\big(i,\theta,(x(\theta),y(\theta))\big)\big]\Big] = \max_{\langle x,y\rangle}\Big[\sum_i \mathbb{E}_\phi[\pi_i(\theta,x(\theta))] + \mathbb{E}_\phi[Y(\theta)]\Big] \\ \text{s.t. EPIC and } Y(\theta) \le Z^* \ \forall\, \theta.\end{gathered} \tag{30}$$

Given a sequence $\{\langle x_k, y_k\rangle\}_{k=1}^\infty$ with

$$\lim_{k\to\infty}\Big(\sum_i \mathbb{E}_\phi[\pi_i(\theta,x_k(\theta))] - \sup_\theta[Y_k(\theta)]\Big) = \sup\Big[\sum_i \mathbb{E}_\phi[\pi_i(\theta,x(\theta))] - \sup_\theta[Y(\theta)]\Big] \tag{31}$$

[37] A similar existence theorem could be proven for non-randomized mechanisms by imposing convexity on the space of events and linearity on the $\pi_i(\theta,\cdot)$ functions.

where the supremum, of course, is over all mechanisms satisfying EPIC and $\mathbb{E}_\phi[Y_k(\theta)] = \mathbb{E}_\phi[Y_k(\theta)] = 0$ for all k, choose a subsequence $\{\langle x_\ell, y_\ell\rangle\}_{\ell=1}^\infty$ such that $\lim_{\ell\to\infty} \max_\theta[Y_\ell(\theta)]$ exists,[38] and set Z^* equal to this limit. Then $\langle \hat{x}, \hat{y}\rangle(\cdot; Z^*)$ is optimal:

$$\begin{aligned}\sum_i \mathbb{E}_\phi[\pi_i(\theta, \hat{x}(\theta; Z^*))] - \sup_\theta[\hat{Y}(\theta; Z^*)] &= \sum_i \mathbb{E}_\phi[\pi_i(\theta, \hat{x}(\theta; Z^*))] - Z^* \\ &= \lim_{\ell\to\infty}\Big(\sum_i \mathbb{E}_\phi[\pi_i(\theta, x_\ell(\theta))] - \sup_\theta[Y_\ell(\theta)]\Big) \qquad (32) \\ &= \sup\Big[\sum_i \mathbb{E}_\phi[\pi_i(\theta, x(\theta))] - \sup_\theta[Y(\theta)]\Big].\end{aligned}$$

The first equality holds because $\mathbb{E}_\phi[\hat{Y}(\theta; Z^*)]$ must be zero or else $\langle \hat{x}, \hat{y}\rangle(\cdot; Z^*)$ would be dominated in its maximization problem (Eq. 30) by a member of the sequence $\{\langle x_\ell, y_\ell\rangle\}_{\ell=1}^\infty$. The second equality, similarly, holds because $\sum_i \mathbb{E}_\phi[\pi_i(\theta, \hat{x}(\theta; Z^*))]$ must equal $\lim_{\ell\to\infty} \sum_i \mathbb{E}_\phi[\pi_i(\theta, x_\ell(\theta))]$ or else $\langle \hat{x}, \hat{y}\rangle(\cdot; Z^*)$ would be dominated by a member of the sequence. □

Appendix B Proofs

Proof of Lemma 2 (page 15). Let $W_{\max} \equiv \max_\theta[W(\theta; v^0)]$, and let $G = \big\langle \mathcal{V}, \{\langle x, t, w\rangle(\cdot; v)\}, v^0\big\rangle$ be an optimal EPPPE mechanism. First, if $W_{\max} > V^0$, then G promises aggregate continuation reward $W(\theta; v^0) > V^0$ in some second period subgames, contrary to the optimality of G.

So suppose that $W_{\max} < V^0$; then construct an alternative mechanism $\tilde{G}$ as follows. Let $\tilde{v}^0 = v^0 + \delta\alpha$ be the initial promised utility, where α is a vector with $\alpha_i = \frac{1}{N}\big(V^0 - W_{\max}\big)$ for $i \in \mathcal{N}$. To support $\tilde{v}^0$ as the initial promised utility, let

$$\langle \tilde{x}, \tilde{t}, \tilde{w}\rangle(\cdot; v^0) = \big\langle x(\cdot; v^0), t(\cdot; v^0), w(\cdot; v^0) + \alpha\big\rangle \qquad (33)$$

be the stage mechanism in the first period. Since α does not vary with the signal vector, this stage mechanism inherits EPIC from the original stage mechanism $\langle x, t, w\rangle(\cdot; v^0)$.

In the second period, fix for the moment a first-period realization $\theta \in \Theta$. Then $\tilde{w}(\theta; \tilde{v}^0) = w(\theta; v^0) + \alpha$ is the initial promised utility in the continuation game. Note that $W(\theta; v^0) < \tilde{W}(\theta; v^0) \leq V^0$, so there exists a zero-sum vector of lump sum monetary transfers $\xi(\theta) \in \mathbb{R}^N$ such that $\tilde{w}(\theta; \tilde{v}^0)$ is a convex combination of $w(\theta; v^0)$ and $v^0 + \xi(\theta)$; i.e., there exist $\xi(\theta) \in \mathbb{R}^N$ and $\beta(\theta) \in [0, 1]$ such that $\sum_i \xi_i = 0$ and

$$\beta(\theta) w(\theta; v^0) + \big(1 - \beta(\theta)\big)\big(v^0 + \xi(\theta)\big) = \tilde{w}(\theta; \tilde{v}^0). \qquad (34)$$

Thus $\tilde{w}(\theta; \tilde{v}^0)$ can be supported by a lottery between $w(\theta; v^0)$ and $v^0 + \xi(\theta)$, with probabilities β and $1 - \beta$ respectively. The first of these promised utility vectors, $w(\theta; v^0)$, can be supported as the initial promised utility by the EPPPE mechanism $\big\langle \mathcal{V}, \{\langle x, t, w\rangle(\cdot; v)\}, w(\theta; v^0)\big\rangle$. The second,

[38] Note that $\max_\theta[Y(\theta)] \geq 0$ for all Y. Since $\{\sum_i \mathbb{E}_\phi[\pi_i(\theta, x_k(\theta))] - \max_\theta[Y_k(\theta)]\}_{k=1}^\infty$ is without loss of generality a uniformly bounded increasing sequence, if $\max_\theta[Y_k(\theta)] \to +\infty$ then $\sum_i \mathbb{E}_\phi[\pi_i(\theta, x_k(\theta))] \to -\infty$, contradicting the assumption that π is uniformly bounded. Thus $\{\max_\theta[Y_k(\theta)]\}_{k=1}^\infty$ contains a convergent subsequence.

$v^0 + \xi(\theta)$, is contained in set of attainable and individually rational stage game payoffs, $\mathcal{P}$, because $w(\theta; v^0) \in \mathcal{P}$ by assumption, and $w_i(\theta; v^0) < w_i(\theta; v^0) + \alpha_i \le v_i^0 + \xi_i(\theta)$ for all $i \in \mathcal{N}$. Furthermore, it is easy to verify that $v^0 + \xi(\theta)$ can be supported as a the initial promised utility of an EPPPE mechanism in the continuation game by the stage mechanism $\big\langle x(\cdot; v^0), t(\cdot; v^0) + \frac{1}{1-\delta}\xi(\theta), w(\cdot; v^0)\big\rangle$ in the first period followed by stage mechanisms drawn from G in all subsequent periods. Since all this holds for any $\theta \in \Theta$, any $\tilde{w}(\theta; \tilde{v}^0)$ can be supported as the initial promised utility in a second period continuation game.

Finally, let $\tilde{\mathcal{V}}$ contain, in addition to $\mathcal{V}$, all the other promised utilities implied by this construction. Thus this alternative mechanism $\tilde{G}$ is an EPPPE mechanism. Hence the fact that $\tilde{V}^0 = V^0 + \delta \sum_i \alpha_i$ indicates that the original EPPPE mechanism G could not have been optimal. □

Proof of Corollary 2 (page 18). The proof shows that (i) implies (ii), while failure of (i) implies failure of (ii). Suppose that (i) holds: the game is ex post unbalanced. Let $\big\langle \mathcal{V}, \big\{\langle x, t, w\rangle(\cdot; v)\big\}, v^0\big\rangle$ be any EPPPE mechanism, and let $\langle x, y\rangle$ be the equivalent static mechanism for $\langle x, t, w\rangle(\cdot; v^0)$ given δ. Then there exist $\varepsilon > 0$ and $\eta > 0$ such that if $\sum_i \mathbb{E}_\phi[\pi_i(\theta, x(\theta))] > Q^* - \varepsilon$ then $\max_\theta[Y(\theta)] > \eta$, in which case $V^0 < Q^* - \eta$ regardless of δ. If instead $\sum_i \mathbb{E}_\phi[\pi_i(\theta, x(\theta))] \le Q^* - \varepsilon$ then $V^0 \le Q^* - \varepsilon$ because $\max_\theta[Y(\theta)] \ge 0$ by construction. Let $\overline{V} = Q^* - \min\{\varepsilon, \eta\}$ to obtain (ii).[39]

Suppose that (i) fails, so the game is not ex post unbalanced. Then for any $\overline{V} < Q^*$ there exists some ex post implementable outcome function $\tilde{x}$ and an equivalent static transfer function $\tilde{y}$ such that $\tilde{V}^0 = \sum_i \mathbb{E}_\phi[\pi_i(\theta, \tilde{x}(\theta))] - \max_\theta[\tilde{Y}(\theta)] > \overline{V}$ and $\langle \tilde{x}, \tilde{y}\rangle$ satisfies EPIC. Suppose that (ii) holds, and choose $\overline{V} < Q^*$ sufficiently high that there exists an EPPPE mechanism $G = \big\langle \mathcal{V}, \big\{\langle x, t, w\rangle(\cdot; v)\big\}, v^0\big\rangle$ with $V^0 \le \overline{V}$. Given $\delta < 1$ sufficiently high, $\tilde{V}^0$ can be implemented as the value of a first stage mechanism $\langle \tilde{x}, \tilde{t}, \tilde{w}\rangle(\cdot; \tilde{v}^0)$ of a recursive mechanism $\tilde{G}$ by arranging the monetary transfer function $\tilde{t}(\cdot; v^0)$, subject to $\tilde{y} = \tilde{t}(\cdot; \tilde{v}^0) + \frac{\delta}{1-\delta}\tilde{w}(\cdot; \tilde{v}^0)$, so that $\big\{\tilde{w}(\cdot; \tilde{v}^0) : \theta \in \Theta\big\}$ falls on a line segment in the direction $(1, \ldots, 1)$. Since this line segment becomes arbitrarily short as $\delta \to 1$, there exists some $\delta < 1$ sufficiently high that it fits between $\tilde{v}^0$ and some zero-sum translation of v^0. Then the promised future surpluses indicated by $\tilde{w}(\cdot; \tilde{v}^0)$ can be realized by randomizing between the stage mechanisms $\langle x, t, w\rangle(\cdot; v^0)$ and $\langle \tilde{x}, \tilde{t}, \tilde{w}\rangle(\cdot; \tilde{v}^0)$ supplemented by appropriate zero sum vector of lump sum transfers (the construction is similar to the detailed argument in the proof of Lemma 2). The recursive mechanism $\tilde{G}$ also satisfies EPIC, by assumption on $\tilde{y}$. Thus $\tilde{G}$ is an EPPPE mechanism, so the fact that $\tilde{V}^0 > \overline{V}$, where $\overline{V} < Q^*$ was selected arbitrarily, is contrary to (ii). □

Proof of Theorem 3 (page 21). First, I introduce a more operational definition of $(N-1)$–additive separability. This definition is equivalent to the one in the text. Given a function $a : \Theta \to \mathbb{R}$, let

$$\Delta^1(a; \theta^1, \theta^2) \equiv a(\theta^1) - a(\theta_1^2, \theta_2^1, \ldots, \theta_N^1), \tag{35}$$

$$\Delta^k(a; \theta^1, \theta^2) \equiv \Delta^{k-1}(a; \theta^1, \theta^2) - \Delta^{k-1}\big(a; \theta^1, (\theta_1^1, \ldots, \theta_{k-1}^1, \theta_k^2, \theta_{k+1}^1, \ldots, \theta_N^1)\big). \tag{36}$$

[39]Note that if δ is so low that individual rationality constraints bind, then the value of an optimal EPPPE can only be lower.

Given a probability measure ϕ, a function $a' : \Theta \to \mathbb{R}$ is $(N-1)$–*additively separable* if there exists $a : \Theta \to \mathbb{R}$ such that (i) $\Delta^N(a; \theta^1, \theta^2) = 0$ for all $\theta^1 \in \Theta$ and all $\theta^2 \in \Theta$, and (ii) $a(\theta) = a'(\theta)$ for ϕ-almost all θ.

The proof proceeds by demonstrating that, in the separable utility environment, a game is not ex post unbalanced if and only if, for all efficient outcome functions x^*, the aggregate EPIC payment function $R(\cdot; x^*)$ is $(N-1)$–additively separable. This is equivalent to the claim in the theorem.

"If": Suppose there exists x^* efficient such that $R(\cdot; x^*)$ is $(N-1)$–additively separable; then by definition there exist functions $\tilde{r}_i : \Theta \to \mathbb{R}$, $i = 1, \ldots, N$, such that (i) $\Delta^N(\sum_i \tilde{r}_i; \theta^1, \theta^2) = 0$ for all $\theta^1, \theta^2 \in \Theta$ and (ii) $\sum_i \tilde{r}_i(\theta) = R(\theta; x^*)$ for ϕ-almost all θ. By (i), there exist fixed transfer functions $y_i^0 : \Theta_{-i} \to \mathbb{R}^N$, for $i = 1, \ldots, N$, such that $\sum_i y_i^0(\theta_{-i}) = \sum_i \tilde{r}_i(\theta)$ for ϕ-almost all θ. Let $y_i(\theta) = y_i^0(\theta) - r_i(\theta; x^*)$; then $\langle x^*, y \rangle$ is EPIC by construction. As claimed, the game is not ex post unbalanced because x^* is efficient and $Y(\theta) = 0$ for ϕ-almost all θ.

"Only if": Suppose that a game is not ex post unbalanced. Then there exists some sequence of (possibly randomized) EPIC mechanisms $\{\langle x_k, y_k \rangle\}_{k=1}^{\infty}$ such that, as $k \to \infty$, (i) x_k approaches EPIC efficiency and (ii) the set on which $Y_k(\theta) > 0$ has ϕ-measure zero for all k. This implies that an optimal mechanism $\langle x^*, y^* \rangle$ must achieve both EPIC efficiency and $Y^*(\theta) = 0$ for ϕ-almost all θ, since any other mechanism can be outperformed by some member of the sequence. Since by Proposition 6 (Appendix A) an optimal mechanism must exist, y^* can be decomposed into $Y^*(\theta) = \sum_i y_i^{0*}(\theta_{-i}) - R(\theta; x^*)$, which, since $Y^*(\theta) = 0$ for ϕ-almost all θ, implies—as claimed—that $R(\theta; x^*)$ is additively separable. □

Proof of Proposition 1 (page 22). Suppose that x^* is efficient; then by Definition 9 it is ex post implementable and $\pi^x(x^*(\theta)) \in \arg\max_{d \in \mathcal{D}}[\pi^\theta(\theta) \cdot d]$ is well defined because $\mathcal{D}$ is compact. Choose θ^1 and θ^2 such that $\{\pi^\theta(s)\}_{s \in \mathcal{S}}$, where $\mathcal{S} = \left\{\theta \in \Theta : \theta_i \in \{\theta_i^1, \theta_i^2\} \ \forall\ i\right\}$, spans $\mathbb{R}^N$ (which is possible because $\{\pi^\theta(\theta)\}_{\theta \in \Theta}$ spans $\mathbb{R}^N$). Then, by condition (i), $\pi^x(x^*(s^1)) \neq \pi^x(x^*(s^2))$ for all $s^1 \in \mathcal{S}$ and $s^2 \in \mathcal{S}$ with $s^1 \neq s^2$. Then $\triangle^N(R(\cdot; x^*); \theta^1, \theta^2) \neq 0$, and furthermore, since by condition (ii) in every subset $\overline{\Theta} \subset \Theta$ with $\phi(\overline{\Theta}) = 1$ there exist θ^1 and θ^2 that satisfy $\triangle^N(R(\cdot; x^*); \theta^1, \theta^2) \neq 0$, there does not exist $a' : \Theta \to \mathbb{R}$ that is both $(N-1)$–additively separable and satisfies $a'(\theta) = R(\theta; x^*)$ for ϕ-almost all $\theta \in \Theta$. Hence by Theorem 3 the game is ex post unbalanced. □

Proof of Proposition 3 (page 24). Let $\hat{y}^{0*}(\cdot; x)$ be a solution and $\zeta(x)$ be the value of Program NLP(x); let $\tilde{y}^{0*}(\cdot; x)$ be the solution and $-\tilde{\zeta}(x)$ be the value of Eq. 15 (i.e., Algorithm LP(x) before its lump sum adjustment). Then $\hat{y}^{0*}(\cdot; x) - \zeta(x)$ satisfies the constraints of Eq. 15 and yields a value of $-\zeta(x)$, so it must be that $-\zeta(x) \leq -\tilde{\zeta}(x)$. Similarly, $\tilde{y}^{0*}(\cdot; x) + \tilde{\zeta}(x)$ satisfies the constraints of Program NLP(x) and yields a value of $\tilde{\zeta}(x)$, so $\zeta(x) \leq \tilde{\zeta}(x)$. Hence $\zeta(x) = \tilde{\zeta}(x)$, so the solution to Algorithm LP(x), $y_i^{0*}(\cdot; x) = \tilde{y}^{0*}(\cdot; x) + \tilde{\zeta}(x)$, is also a solution to Program NLP(x). □

Proof of Proposition 4 (page 37). The proof proceeds in two steps. First, I establish that if $y_1^{0*}(\cdot; x^*)$ takes the form described in the proposition, then there exists $y_2^{0*}(\cdot; x^*)$ such that conditions in the proposition are satisfied. Second, I establish that for any $y_1^{0*}(\cdot; x^*)$ there exists $y_2^{0*}(\cdot; x^*)$ that takes the form described in the proposition.

Suppose that $y_1^{0*}(\cdot;x^*)$ satisfies condition (ii) in the proposition; i.e., there exists $b_2 \in \Theta_2$ and $\tilde{a}_1 \in \mathbb{R}$ such that

$$\tilde{y}_1^{0*}(\theta_2;x^*) = \tilde{a}_1 - \mathbb{I}\big[\theta_2 \leq b_2\big]\,\big(\pi_1^\theta(\lambda_1^*(b_2),b_2) - \pi_1^\theta(\lambda_1^*(\theta_2),\theta_2)\big), \tag{37}$$

where $\mathbb{I}$ is the indicator function, is one component of a solution to Eq. 15 given x^*. Since $\tilde{y}_1^{0*}(\cdot;x^*)$ is fixed, solving Eq. 15 given x^* amounts to maximizing $\tilde{y}_2^{0*}(\theta;x^*)$ pointwise subject to the constraint $\tilde{y}_2^{0*}(\theta_1;x^*) \leq R(\theta;x^*) - \tilde{y}_1^{0*}(\theta_2;x^*)$ for all θ. It is without loss of generality to assume that $\tilde{a}_1 = 0$, because otherwise we could simply subtract $\tilde{a}_1$ from $\tilde{y}_1^{0*}(\cdot;x^*)$ and add it to $\tilde{y}_2^{0*}(\cdot;x^*)$. Let $b_1 = \lambda_1^*(b_2)$. Then the constraint requires that $\tilde{y}_2^{0*}(\theta_1;x^*) \leq \pi_1^\theta(\lambda_1^*(b_2),b_2)$ on $[b_1,1]$, because otherwise it would be violated on $[b_1,1]\times[0,b_2]$:

$$\sum_i \tilde{y}_i^{0*}(\theta_{-i};x^*) = \pi_1^\theta(\lambda_1^*(\theta_2),\theta_2) - \pi_1^\theta(\lambda_1^*(b_2),b_2) + \tilde{y}_2^{0*}(\theta_1;x^*) > R(\theta;x^*) = \pi_1^\theta(\lambda_1^*(\theta_2),\theta_2). \tag{38}$$

Similarly, $\tilde{y}_2^{0*}(\theta_1;x^*) \leq \pi_2^\theta(\theta_1,\lambda_2^*(\theta_1))$ on $[0,b_1]$, because otherwise the constraint would be violated on $[0,b_1]\times[b_2,1]$:

$$\sum_i \tilde{y}_i^{0*}(\theta_{-i};x^*) = \tilde{y}_2^{0*}(\theta_1;x^*) > R(\theta;x^*) = \pi_2^\theta(\theta_1,\lambda_2^*(\theta_1)) \tag{39}$$

Since these are the only constraints, it follows immediately that setting

$$\tilde{y}_2^{0*}(\theta_1;x^*) = \pi_2^\theta(\theta_1,\lambda_2^*(\theta_1)) - \mathbb{I}\big[\theta_1 \leq b_1\big]\,\big(b_1, \pi_2^\theta(\lambda_2^*(b_1)) - \pi_2^\theta(\theta_1,\lambda_2^*(\theta_1))\big) \tag{40}$$

solves Eq. 15 given x^*. The conditions in the proposition follow after making the lump sum adjustment to satisfy ex ante budget balance.

Now consider any $\tilde{y}_1^{0*}(\cdot;x^*)$ and assume nothing about its form. Without loss of generality, let $\sup_{\theta_2}\big[\tilde{y}_1^{0*}(\theta_2;x^*)\big] = 0$. As above, since $\tilde{y}_1^{0*}(\cdot;x^*)$ is fixed the solution to Eq. 15 given x^* is to maximize $\tilde{y}_2^{0*}(\theta_1;x^*)$ pointwise subject to the constraints. For any particular θ_2, the constraint $\tilde{y}_2^{0*}(\theta_1;x^*) \leq R(\theta;x^*) - \tilde{y}_1^{0*}(\theta_2;x^*)$ requires that

$$\tilde{y}_2^{0*}(\theta_1;x^*) \leq \begin{cases} \pi_2^\theta(\theta_1,\lambda_2^*(\theta_1)) - \tilde{y}_1^{0*}(\theta_2;x^*) & \text{if } \theta_1 \in [0,\lambda_1^*(\theta_2)], \\ \pi_1^\theta(\lambda_1^*(\theta_2),\theta_2) - \tilde{y}_1^{0*}(\theta_2;x^*) & \text{otherwise.} \end{cases} \tag{41}$$

The "otherwise" part of the constraint is flat, while the part of the constraint for $\theta_1 \in [0,\lambda_1^*(\theta_2)]$ is increasing in θ_1 under the regularity conditions. Since $\pi_2^\theta(\theta_1,\lambda_2^*(\theta_1)) = \pi_1^\theta(\theta_1,\lambda_2^*(\theta_1))$, the two parts of the constraint coincide at $\theta_1 = \lambda_1^*(\theta_2)$ and at no other point. Thus the constraint implied by a particular θ_2 can be written equivalently as

$$\tilde{y}_2^{0*}(\theta_1;x^*) \leq \min\big\{\pi_2^\theta(\theta_1,\lambda_2^*(\theta_1)), \pi_1^\theta(\lambda_1^*(\theta_2),\theta_2)\big\} - \tilde{y}_1^{0*}(\theta_2;x^*). \tag{42}$$

Then the lower envelope of the constraints implied by all $\theta_2 \in \Theta_2$ can be written as follows:

$$\begin{aligned}\tilde{y}_2^{0*}(\theta_1; x^*) &\leq \min \left\{ \begin{array}{c} \pi_2^\theta(\theta_1, \lambda_2^*(\theta_1)) - \sup_{\theta_2} \big[-\tilde{y}_1^{0*}(\theta_2; x^*)\big], \\ \inf_{\theta_2} \big[\pi_1^\theta(\lambda_1^*(\theta_2), \theta_2) - \tilde{y}_1^{0*}(\theta_2; x^*)\big] \end{array} \right\} \\ &= \min\Big\{\pi_2^\theta(\theta_1, \lambda_2^*(\theta_1)), \inf_{\theta_2} \big[\pi_1^\theta(\lambda_1^*(\theta_2), \theta_2) - \tilde{y}_1^{0*}(\theta_2; x^*)\big]\Big\}.\end{aligned} \tag{43}$$

Let $\tilde{a}_1 = 0$ and let b_1 solve $\pi_2^\theta(b_1', \lambda_2^*(\theta_1)) = \inf_{\theta_2}\big[\pi_1^\theta(\lambda_1^*(\theta_2), \theta_2) - \tilde{y}_1^{0*}(\theta_2; x^*)\big]$. Then $\tilde{a}_1$, b_1, and $\tilde{y}_2^{0*}(\theta_1; x^*)$—after the lump sum adjustment to satisfy ex post budget balance—satisfy condition (ii), completing the proof. □

Remark 1. In the context of Proposition 4, suppose that an ex post implementable allocation rule x is similar to x^* except on a rectangle $[0, \varepsilon_1] \times [0, \varepsilon_2]$ with ε_1 and ε_2 sufficiently small. Then, using a substantially unchanged proof, we can demonstrate that $\tilde{Y}^*(\theta; x) = \tilde{Y}^*(\theta; x^*)$ for all $\theta \notin [0, \varepsilon_1] \times [0, \varepsilon_2]$.

Proof sketch. Use the proof of Proposition 4, but replace $\pi_i^\theta(\lambda_i^*(\theta_{-i}), \theta_{-i})$ with $r_i((1, \theta_{-i}); x)$. Since $r_i((1, \theta_{-i}); x) = r_i((1, \theta_{-i}); x^*)$, for $\theta_{-i} \in [\varepsilon_{-i}, 1]$, this will lead to the conclusion that $\tilde{y}_i^{0*}(\theta_{-i}; x^*) = \tilde{y}_i^{0*}(\theta_{-i}; x)$. For $\theta_{-i} \in [0, \varepsilon_{-i}]$, it will lead to the conclusion that $r_i((1, \theta_{-i}); x) - r_i((1, \theta_{-i}); x^*) = \tilde{y}_i^{0*}(\theta_{-i}; x^*) - \tilde{y}_i^{0*}(\theta_{-i}; x)$, so that $y_i^*(\theta; x^*) = y_i^*(\theta; x)$ for all $\theta \notin [0, \varepsilon_1] \times [0, \varepsilon_2]$. □

Proof of Theorem 4 (page 38). On a sufficiently small rectangle that borders the origin, first order approximations are valid under the regularity conditions. Under a first order approximation close to the origin, let β solve $\pi_1^\theta(\theta_1, \beta\theta_1) = \pi_2^\theta(\theta_1, \beta\theta_1)$. For $\varepsilon > 0$ small, for all $\theta \in E^\varepsilon \equiv [0, \varepsilon] \times [0, \beta\varepsilon]$ assume that

$$\pi_1^\theta(\theta) = c_{10} + c_{11}\theta_1 + c_{12}\theta_2, \tag{44}$$

$$\pi_2^\theta(\theta) = c_{20} + c_{21}\theta_1 + c_{22}\theta_2, \tag{45}$$

$$\phi(\theta) = c_{30} + c_{31}\theta_1 + c_{32}\theta_2. \tag{46}$$

By the regularity conditions, $c_{30} > 0$; $c_{10}, c_{20} \geq 0$; $0 \leq c_{12}, c_{21} < c_{11}, c_{22}$; and $\beta = \frac{c_{11} - c_{21}}{c_{22} - c_{12}} \in (0, \infty)$.

Let $x_1^\varepsilon(\theta) = \chi_1 \in [0, 1]$ if $\theta \in E^\varepsilon$, and $x_1^\varepsilon(\theta) = x_1^*(\theta)$ otherwise; note that any $\chi_1 \in [0, 1]$ will preserve monotonicity. By Remark 1, it can be shown that $\tilde{Y}^*(\theta; x^*) = \tilde{Y}^*(\theta; x^\varepsilon)$ on all of $\Theta \backslash E^\varepsilon$. Since, according to Algorithm LP*, $V^0 = \sum_i \mathbb{E}_\phi[\pi_i(\theta, x(\theta))] + \mathbb{E}_\phi[\sum_i \tilde{y}^0(\theta_{-i}) - R(\theta; x)]$, to demonstrate an improvement we need only to show that $\Delta V > 0$, where

$$\begin{aligned}\Delta V \equiv \int_{E^\varepsilon} \Big(\big(\pi_1^\theta(\theta) - \pi_2^\theta(\theta)\big)\big(\chi_1 - x_1^*(\theta)\big) + \sum_i \tilde{y}^{0*}(\theta_{-i}; x^\varepsilon) - R(\theta; x^\varepsilon) \\ + \sum_i \tilde{y}^{0*}(\theta_{-i}; x^*) - R(\theta; x^*) \Big)\, d\phi(\theta).\end{aligned} \tag{47}$$

Define the following, and note their values for $\theta \in E^\varepsilon$:

$$\Delta\tilde{y}_2^{0*}(\theta_1) \equiv \tilde{y}_2^{0*}(\theta_1; x^\varepsilon) - \tilde{y}_2^{0*}(\theta_1; x^*) = \pi_2(\theta_1, \beta\varepsilon)\,\chi_1 - \pi_2(\theta_1, \beta\theta_1), \tag{48}$$

$$\Delta\tilde{y}_1^{0*}(\theta_2) \equiv \tilde{y}_1^{0*}(\theta_2; x^\varepsilon) - \tilde{y}_1^{0*}(\theta_2; x^*) = \pi_1(\varepsilon, \theta_2)\,(1-\chi_1) - \pi_1(\tfrac{1}{\beta}\theta_2, \theta_2), \tag{49}$$

$$\Delta r_1(\theta) \equiv r_1(\theta; x^\varepsilon) - r_1(\theta; x^*) = -\pi_1(\tfrac{1}{\beta}\theta_2, \theta_2)\,\mathbb{I}\big[\theta_1 > \tfrac{1}{\beta}\theta_2\big], \tag{50}$$

$$\Delta r_2(\theta) \equiv r_2(\theta; x^\varepsilon) - r_2(\theta; x^*) = -\pi_2(\theta_1, \beta\theta_1)\,\mathbb{I}\big[\theta_2 > \beta\theta_1\big]. \tag{51}$$

Divide E^ε into two parts, $E_1^\varepsilon \equiv \{\theta \in E^\varepsilon : \theta_2 < \beta\theta_1\}$ and $E_2^\varepsilon \equiv \{\theta \in E^\varepsilon : \theta_2 \geq \beta\theta_1\}$; then

$$\begin{aligned}\Delta V = &\int_{E_1^\varepsilon} \Big(\big(\pi_1^\theta(\theta) - \pi_2^\theta(\theta)\big)\,(\chi_1 - 1) + \Delta\tilde{y}_2^{0*}(\theta_1) + \Delta\tilde{y}_1^{0*}(\theta_2) - \Delta r_1(\theta) \Big)\, d\phi(\theta) \\ &+ \int_{E_2^\varepsilon} \Big(\big(\pi_1^\theta(\theta) - \pi_2^\theta(\theta)\big)\chi_1 + \Delta\tilde{y}_2^{0*}(\theta_1) + \Delta\tilde{y}_1^{0*}(\theta_2) - \Delta r_2(\theta) \Big)\, d\phi(\theta).\end{aligned} \tag{52}$$

Note that ΔV is linear in χ_1, so that a maximum is to be found either at $\chi_1 = 0$ or at $\chi_1 = 1$. Thus it suffices to show that $\overline{\Delta V} \equiv \frac{1}{2}\big(\Delta V|_{\chi_1=0} + \Delta V|_{\chi_1=1}\big) > 0$. Solving explicitly for $\overline{\Delta V}$ yields

$$\overline{\Delta V} = \frac{\beta}{12}(c_{11} + \beta c_{22})c_{30}\varepsilon^3 + A\varepsilon^4, \tag{53}$$

where A is a term that does not vary with ε. Since the first term is an order of ε larger than the second term, it suffices to note that the first term is strictly positive under the regularity conditions. □

Lemma 3. *In a two-player allocation game, suppose that the players' valuation functions satisfy globally interdependent valuations, as defined in Theorem 5. Then any EPIC mechanism with a zero surplus gap must employ an allocation rule that is constant (except possibly at isolated points).*

Proof. Given any points $\theta^1 \in \Theta$ and $\theta^2 \in \Theta$, let $\overline{\Delta}^2 \equiv \Delta^2(R(\cdot; x); \theta^1, \theta^2)$ (see Eq. 36). We can write $\overline{\Delta}^2$ as a function of $x_1(\theta^1)$, $x_1(\theta^2)$, $x_1(\theta_1^1, \theta_2^2)$, and $x_1(\theta_1^2, \theta_2^1)$ as follows:

$$\begin{aligned}\overline{\Delta}^2 = &\big(x_1(\theta^2)\pi_1^\theta(\theta^2) - x_1(\theta_1^2, \theta_2^1)\pi_1^\theta(\theta_1^2, \theta_2^1)\big) - \big(x_1(\theta_1^1, \theta_2^2)\pi_1^\theta(\theta_1^1, \theta_2^2) - x_1(\theta^1)\pi_1^\theta(\theta^1)\big) \\ &+ \Big(\big(1 - x_1(\theta^2)\big)\,\pi_2^\theta(\theta^2) - \big(1 - x_1(\theta_1^1, \theta_2^2)\big)\,\pi_2^\theta(\theta_1^1, \theta_2^2)\Big) \\ &- \Big(\big(1 - x_1(\theta_1^2, \theta_2^1)\big)\,\pi_2^\theta(\theta_1^2, \theta_2^1) - \big(1 - x_1(\theta^1)\big)\,\pi_2^\theta(\theta^1)\Big) \\ &- \int_{\theta_1^1}^{\theta_1^2} \left(\frac{\partial\pi_1^\theta(s_1, \theta_2^2)}{s_1} x_1(s_1, \theta_2^2) - \frac{\partial\pi_1^\theta(s_1, \theta_2^1)}{s_1} x_1(s_1, \theta_2^1) \right) ds_1 \\ &- \int_{\theta_2^1}^{\theta_2^2} \left(\frac{\partial\pi_2^\theta(\theta_1^2, s_2)}{s_2}\big(1 - x_1(\theta_1^2, s_2)\big) - \frac{\partial\pi_2^\theta(\theta_1^1, s_2)}{s_2}\big(1 - x_1(\theta_1^1, s_2)\big) \right) ds_2.\end{aligned} \tag{54}$$

Multiply out and combine terms to yield:

$$\begin{aligned}\overline{\Delta}^2 &= x_1(\theta^2)\big(\pi_1^\theta(\theta^2) - \pi_2^\theta(\theta^2)\big) - x_1(\theta_1^2,\theta_2^1)\big(\pi_1^\theta(\theta_1^2,\theta_2^1) - \pi_2^\theta(\theta_1^2,\theta_2^1)\big) \\ &\quad - x_1(\theta_1^1,\theta_2^2)\big(\pi_1^\theta(\theta_1^1,\theta_2^2) - \pi_2^\theta(\theta_1^1,\theta_2^2)\big) + x_1(\theta^1)\big(\pi_1^\theta(\theta^1) - \pi_2^\theta(\theta^1)\big) \\ &\quad - \int_{\theta_1^1}^{\theta_1^2} \left(\frac{\partial \pi_1^\theta(s_1,\theta_2^2)}{s_1} x_1(s_1,\theta_2^2) - \frac{\partial \pi_1^\theta(s_1,\theta_2^1)}{s_1} x_1(s_1,\theta_2^1) \right) ds_1 \\ &\quad + \int_{\theta_2^1}^{\theta_2^2} \left(\frac{\partial \pi_2^\theta(\theta_1^2,s_2)}{s_2} x_1(\theta_1^2,s_2) - \frac{\partial \pi_2^\theta(\theta_1^1,s_2)}{s_2} x_1(\theta_1^1,s_2) \right) ds_2. \end{aligned} \tag{55}$$

Given $\theta^1 \in [0,1) \times [0,1)$ and $\delta > 0$, let $\theta_1^2 = \theta_1^1 + 2\delta$ and $\theta_2^2 = \theta_2^1 + \delta$. For any $\theta^1 \in [0,1) \times [0,1)$ there exists $\delta > 0$ sufficiently small that $\theta^2 \in \Theta$. Let $E \equiv [\theta_1^1, \theta_1^2] \times [\theta_2^1, \theta_2^2]$. Under the regularity conditions, for $\delta > 0$ small we can take the first order approximation of π^θ on E as $\pi_i^\theta(\theta) = c_i + a_i\theta_i + b_i\theta_{-i}$. Let $\overline{\pi}_i \equiv \pi_i^\theta(\theta^1)$, and let $l \equiv x_1(\theta^1)$, $L \equiv x_1(\theta_1^1, \theta_2^2)$, $H \equiv x_1(\theta_1^2, \theta_2^1)$, and $h \equiv x_1(\theta^2)$. By monotonicity of $\pi_i^x(x(\theta))$, $L \leq l \leq H$ and $L \leq h \leq H$. Allow x to be discontinuous or non-differentiable on E along a line of discontinuity or non-differentiability with slope $k \in [1, \infty]$ locally. For $\delta > 0$ sufficiently small, this is without loss of generality.[40] We can construct E so that the line passes through the center of E, passing first between θ^1 and (θ_1^2, θ_2^1) and then between (θ_1^1, θ_2^2) and θ^2. So:

$$\begin{aligned}\overline{\Delta}^2 &\approx (h - H - L + l)\,(\overline{\pi}_1 - \overline{\pi}_2) + (h - H)\,2\delta\,(a_1 - b_2) + (h - L)\,\delta\,(b_1 - a_2) \\ &\quad - a_1 \int_{\theta_1^1}^{\theta_1^1 + 2\delta} \big(x_1(s_1,\theta_2^2) - x_1(s_1,\theta_2^1)\big)\, ds_1 + a_2 \int_{\theta_2^1}^{\theta_2^1+\delta} \big(x_1(\theta_1^2,s_2) - x_1(\theta_1^1,s_2)\big)\, ds_2. \end{aligned} \tag{56}$$

Let $\gamma_{lH} \geq 0$ be the size of the discontinuity between θ^1 and (θ_1^2, θ_2^1). Define $\rho_1^L \equiv \frac{\partial x_1(\theta)}{\partial \theta_1}\big|_{\theta^1} > 0$, $\rho_2^L \equiv \frac{\partial x_1(\theta)}{\partial \theta_2}\big|_{\theta^1} < 0$, $\rho_1^H \equiv \frac{\partial x_1(\theta)}{\partial \theta_1}\big|_{\theta^2} > 0$, and $\rho_2^H \equiv \frac{\partial x_1(\theta)}{\partial \theta_2}\big|_{\theta^2} < 0$. Define $\omega_{11}^L \equiv \frac{\partial^2 x_1(\theta)}{\partial(\theta_1)^2}\big|_{\theta^1}$, $\omega_{11}^H \equiv \frac{\partial^2 x_1(\theta)}{\partial(\theta_1)^2}\big|_{\theta^2}$, $\omega_{22}^L \equiv \frac{\partial^2 x_1(\theta)}{\partial(\theta_2)^2}\big|_{\theta^1}$, $\omega_{22}^H \equiv \frac{\partial^2 x_1(\theta)}{\partial(\theta_2)^2}\big|_{\theta^2}$, $\omega_{12}^L \equiv \frac{\partial^2 x_1(\theta)}{\partial\theta_1\partial\theta_2}\big|_{\theta^1}$, and $\omega_{12}^H \equiv \frac{\partial^2 x_1(\theta)}{\partial\theta_1\partial\theta_2}\big|_{\theta^2}$. Then:

$$H \approx l + \int_0^{\delta - \frac{\delta}{2k}} \big(\rho_1^L + \omega_{11}^L s_1\big)\, ds_1 + \gamma_{lH} + \int_0^{\delta + \frac{\delta}{2k}} \big(\rho_1^H - \omega_{12}^H \delta + \omega_{11}^H s_1\big)\, ds_1, \tag{57}$$

$$L \approx l + \int_0^{\delta} \big(\rho_2^L + \omega_{22}^L s_2\big)\, ds_2, \tag{58}$$

$$h \approx H + \int_0^{\delta} \big(\rho_2^H - \omega_{22}^H s_2\big)\, ds_2; \tag{59}$$

[40] If the line has slope less than 1, switch the players. If there is discontinuity or non-differentiability on any connected region including more than one point, it must form a line that is locally differentiable because of the monotonicity of x as required for EPIC.

and

$$\int_{\theta_1^1}^{\theta_1^1+2\delta} \big(x_1(s_1,\theta_2^2) - x_1(s_1,\theta_2^1)\big) ds_1 \approx \int_0^{\delta+\frac{\delta}{2k}} \Big(L - H + \big(\rho_1^L + \rho_1^H + (\omega_{12}^L - \omega_{12}^H)\,\delta\big)s_1\Big)\, ds_1$$

$$+ \int_0^{\delta+\frac{\delta}{2k}} \frac{1}{2}(\omega_{11}^L + \omega_{11}^H) s_1^2\, ds_1 \tag{60}$$

$$+ \int_0^{\delta-\frac{\delta}{2k}} \Big(h - l - \big(\rho_1^H + \rho_1^L + \frac{1}{2}(\omega_{11}^H + \omega_{11}^L)\, s_1\big)s_1\Big)\, ds_1,$$

$$\int_{\theta_2^1}^{\theta_2^1+\delta} \big(x_1(\theta_1^2,s_2) - x_1(\theta_1^1,s_2)\big) ds_2 \approx \int_0^{\delta} \Big(H - l + \big(\rho_2^H - \rho_2^L - \frac{1}{2}(\omega_{22}^H + \omega_{22}^L)\, s_2\big)s_2\Big)\, ds_2. \tag{61}$$

Divide $\overline{\Delta}^2$ by δ and take the limit as $\delta \to 0$:

$$\lim_{\delta\to 0} \frac{\overline{\Delta}^2}{\delta} = \gamma_{lH}\Big(b_1 + \frac{1}{k}a_1\Big) + \big(\rho_2^H - \rho_2^L\big)\big(\overline{\pi}_1 - \overline{\pi}_2\big) \tag{62}$$

If there is no money burning, then $\lim_{\delta\to 0} \frac{\overline{\Delta}^2}{\delta} = 0$ except possibly at isolated points. Suppose that θ^1 is not such an isolated point; this tells us that either

1. $\gamma_{lH} = 0$ and either $\rho_2^H = \rho_2^L$ or $\overline{\pi}_1 = \overline{\pi}_2$;
2. $\gamma_{lH} > 0$, $b_1 = 0$, $k = \infty$, and either $\rho_2^L = \rho_2^H$ or $\overline{\pi}_1 = \overline{\pi}_2$;
3. $\gamma_{lH} > 0$, $b_1 \geq 0$, $k \leq \infty$, and $\gamma_{lH}\big(b_1 + \frac{1}{k}a_1\big) = (\rho_2^H - \rho_2^L)\,(\overline{\pi}_2 - \overline{\pi}_1)$.

In Case 1, we must have $\lim_{\delta\to 0}(\rho_2^H - \rho_2^L) = \lim_{\delta\to 0}(\omega_{22}^H - \omega_{22}^L) = 0$ everywhere in order to satisfy monotonicity. So divide $\overline{\Delta}^2$ by δ^2 and take the limit as $\delta \to 0$:

$$\lim_{\delta\to 0} \frac{\overline{\Delta}^2}{\delta^2} = -\frac{1}{2}\big(4b_2\rho_2^L - 2b_1(\rho_1^H + \rho_1^L)\big) \tag{63}$$

Since $\rho_2^L \leq 0$ while $\rho_1^H + \rho_1^L \geq 0$, this tells us that we must have $\rho_2^H = \rho_2^L = 0$ or $b_2 = 0$, as well as $\rho_1^H = \rho_1^L = 0$ or $b_1 = 0$. Hence under globally interdependent valuations, since x must be continuous almost everywhere, it must be constant almost everywhere. Case 2 does not concern situations with globally interdependent valuations. In Case 3, with $b_1 > 0$ for globally interdependent valuations, it must be that $(\rho_2^H - \rho_2^L)\,(\overline{\pi}_1 - \overline{\pi}_2) > 0$. If $\overline{\pi}_1 \neq \overline{\pi}_2$, then either $\rho_2^H \neq 0$ or $\rho_2^L \neq 0$, which is incompatible with the fact that Case 1 must hold on either side of the discontinuity. If $\overline{\pi}_1 = \overline{\pi}_2$, then the inequality cannot be satisfied. Hence Case 3 cannot apply to globally interdependent valuations, and so there can be no discontinuities except possibly at isolated points. □

Proof of Theorem 5 (page 38). By Lemma 3, under interdependent valuations it is not possible to implement a zero surplus gap except with a constant outcome function. Consider a constant outcome function with $\overline{x}_1(\theta) = \chi_1 \leq \frac{1}{2}$ for all θ. Since $r_i(\theta; \overline{x}) = 0$ for all θ and all i, let $\overline{y}_i^0(\theta) = 0$ for all θ and all i as well. Now, for $\varepsilon > 0$ small, consider an alternative outcome function x, such that $x_1(\theta) = 1$ for all $\theta \in E \equiv [1-\varepsilon, 1] \times [0, \varepsilon]$, and $x_1(\theta) = \chi_1$ otherwise. Choose $y_2^0(\theta_1) = (1-\chi_2)\,\pi_2(\theta_1, \varepsilon) + c$ for $\theta_1 \geq 1-\varepsilon$, and $y_2^0(\theta_1) = c$ otherwise; choose $y_1^0(\theta_2) = 0$ for all θ. Since $r_2(\theta) = (1-\chi_1)\,\pi_2(\theta_1, \varepsilon)$ for $\theta \in [1-\varepsilon, 1] \times (\varepsilon, 1]$ and $r_2(\theta) = 0$ elsewhere, while $r_1(\theta) = (1-\chi_1)\,\pi_1(1-\varepsilon, \theta_2)$ for $\theta \in E$ and $r_1(\theta) = 0$ elsewhere, the surplus gap is simply c, which must be chosen such that $\mathbb{E}_\phi[Y(\theta)] = 0$:

$$\max_\theta [Y(\theta)] = c = \int_E \big(1-\chi_1\big)\big(\pi_1(1-\varepsilon, \theta_2) - \pi_2(\theta_1, \varepsilon)\big)\, d\phi(\theta). \tag{64}$$

At the same time, the increase in welfare is

$$\sum_i \mathbb{E}_\phi[\pi_i(\theta, x(\theta))] - \sum_i \mathbb{E}_\phi[\pi_i(\theta, \overline{x}(\theta))] = \int_E \big(1-\chi_1\big)\big(\pi_1(\theta) - \pi_2(\theta)\big)\, d\phi(\theta). \tag{65}$$

Hence the improvement in V^0 is

$$\sum_i \mathbb{E}_\phi[\pi_i(\theta, x(\theta))] - \sum_i \mathbb{E}_\phi[\pi_i(\theta, \overline{x}(\theta))] - \max_\theta [Y(\theta)]$$
$$= (1-\chi_1) \int_E \Big(\pi_1(\theta) - \pi_1(1-\varepsilon, \theta_2) + \pi_2(\theta_1, \varepsilon) - \pi_2(\theta)\Big)\, d\phi(\theta), \tag{66}$$

which is strictly positive under the regularity conditions. □

References

Dilip Abreu, David Pearce, and Ennio Stacchetti. Optimal cartel equilibria with imperfect monitoring. *Journal of Economic Theory*, 39(1):251–69, June 1986.

Dilip Abreu, David Pearce, and Ennio Stacchetti. Toward a theory of discounted repeated games with imperfect private monitoring. *Econometrica*, 58(5):1041–1063, September 1990.

Charalambos D. Aliprantis and Kim C. Border. *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer, Berlin, 2nd edition, 1999.

Masaki Aoyagi. Bid rotation and collusion in repeated auctions. *Journal of Economic Theory*, 112(1):79–105, September 2003a.

Masaki Aoyagi. Efficient collusion in repeated auctions with communication. URL http://www.iser.osaka-u.ac.jp/~aoyagi/efficient.pdf. Working paper, December 2003b.

Susan Athey and Kyle Bagwell. Optimal collusion with private information. *RAND Journal of Economics*, 32(3):428–465, September 2001.

Susan Athey and Kyle Bagwell. Collusion with persistent cost shocks. URL http://www.stanford.edu/~athey/CollusionPersistent.pdf. Working Paper, August 2004.

Susan Athey, Kyle Bagwell, and Chris Sanchirico. Collusion and price rigidity. *Review of Economic Studies*, 71(247):317–350, April 2004.

Kyle Bagwell and Robert W. Staiger. Collusion over the business cycle. *RAND Journal of Economics*, 28(1):82–106, Spring 1997.

Erik J. Balder. On the existence of optimal contract mechanisms for incomplete information principal-agent models. *Journal of Economic Theory*, 68(1):133–148, 1996.

Dirk Bergemann and Stephen Morris. Robust mechanism design. URL http://cowles.econ.yale.edu/P/cd/d14a/d1421.pdf. Cowles Foundation Discussion Paper No. 1421, May 2003.

Andreas Blume and Paul Heidhues. Private monitoring in auctions. URL http://www.pitt.edu/~ablume/images/private12.pdf. Working Paper, December 2002.

Harrison Cheng. Inefficiency in repeated Cournot oligopoly games. Working Paper, 2004.

Kim-Sau Chung and Jeffrey C. Ely. Ex-post incentive compatible mechanism design. URL http://faculty-web.at.northwestern.edu/economics/chung/file/expost.pdf. Working Paper, June 2003a.

Kim-Sau Chung and Jeffrey C. Ely. Foundations of dominant strategy mechanisms. URL http://www.kellogg.northwestern.edu/research/math/dps/1372.pdf. Working Paper, June 2003b.

E. H. Clarke. Multipart pricing of public goods. *Public Choice*, 11:17–33, 1971.

Olivier Compte. Communication in repeated games with imperfect private monitoring. *Econometrica*, 66(3):597–626, May 1998.

Drew Fudenberg, David K. Levine, and Eric S. Maskin. The folk theorem with imperfect public information. *Econometrica*, 62(5):997–1039, September 1994.

Dennis Gärtner and Armin Schmutzler. Ex-post implementable merger mechanisms. Working paper, August 2005.

Edward J. Green and Robert H. Porter. Noncooperative collusion under imperfect price information. *Econometrica*, 52(1):87–100, January 1984.

Theodore Groves. Incentives in teams. *Econometrica*, 41(4):617–631, July 1973.

Johannes Hörner and Julian S. Jamison. Collusion with (almost) no information. URL http://www.kellogg.northwestern.edu/faculty/horner/collusion.pdf. Working Paper, January 2004.

Leonid Hurwicz and Mark Walker. On the generic nonoptimality of dominant-strategy allocation mechanisms: A general theorem that includes pure exchange economies. *Econometrica*, 58(3):683–704, May 1990.

Philippe Jehiel, Moritz Meyer-ter-Vehn, Benny Moldovanu, and William R. Zame. The limits of ex-post implementation. URL http://www.econ.ucla.edu/zame/ExPost.pdf. Working Paper, February 2005.

Michihiro Kandori and Hitoshi Matsushima. Private observation, communication and collusion. *Econometrica*, 66(3):627–652, 1998.

Michihiro Kandori and Ichiro Obara. Efficiency in repeated games revisited: The role of private strategies. URL http://www.econ.ucla.edu/workingpapers/wp826.pdf. UCLA Dept. of Economics Working Paper No. 826, January 2003.

Matt Kelley. Air Force shoots down Boeing, July 25, 2003. URL http://www.cbsnews.com/stories/2003/07/25/national/main565160.shtml.

Jean-Jacques Laffont and Eric S. Maskin. A differential approach to dominant strategy mechanisms. *Econometrica*, 48(48):1507–1520, September 1980.

Jonathan Levin. Relational incentive contracts. *American Economic Review*, 93(3):835–847, June 2003.

Alberto Martin and Wouter Vergote. Efficient bid rotation in repeated auctions. URL http://www.columbia.edu/~wrv13/CollusionAuctions.pdf. Working Paper, October 2004.

R. Preston McAfee and John McMillan. Bidding rings. *American Economic Review*, 82(3):579–599, June 1992.

Claudio Mezzetti. Mechanism design with interdependent valuations: efficiency. *Econometrica*, 72(5):1617–1626, September 2004.

Paul R. Milgrom and Ilya Segal. Envelope theorems for arbitrary choice sets. *Econometrica*, 70(2):583–601, March 2002.

David A. Miller. Attainable payoffs in repeated games with interdependent private information. URL http://dss.ucsd.edu/~d9miller/papers/d9miller_Folk.pdf. Working Paper, August 2005.

Luis Rayo. Relational team incentives and ownership. URL http://gsbwww.uchicago.edu/fac/luis.rayo/research/RayoTEAMS.pdf. Working Paper, May 2003.

Julio J. Rotemberg and Garth Saloner. A supergame-theoretic model of price wars during booms. *American Economic Review*, 76(3):390–407, June 1986.

Yuliy Sannikov and Andrzej Skrzypacz. Impossibility of collusion under imperfect monitoring with flexible production. Working Paper, October 2004.

Andrzej Skrzypacz and Hugo A. Hopenhayn. Tacit collusion in repeated auctions. *Journal of Economic Theory*, 114(1):153–169, January 2004.

William Vickrey. Counterspeculation, auctions, and competitive sealed tenders. *Journal of Finance*, 16(1):8–37, March 1961.