AN EFFICIENT DYNAMIC MECHANISM

By Susan Athey and Ilya Segal[1]

This paper constructs an efficient, budget-balanced, Bayesian incentive-compatible mechanism for a general dynamic environment with quasilinear payoffs in which agents observe private information and decisions are made over countably many periods. First, under the assumption of "private values" (other agents' private information does not directly affect an agent's payoffs), we construct an efficient, ex post incentive-compatible mechanism, which is not budget-balanced. Second, under the assumption of "independent types" (the distribution of each agent's private information is not directly affected by other agents' private information), we show how the budget can be balanced without compromising agents' incentives. Finally, we show that the mechanism can be made self-enforcing when agents are sufficiently patient and the induced stochastic process over types is an ergodic finite Markov chain.

KEYWORDS: Dynamic mechanism design, dynamic incentive compatibility, perfect Bayesian equilibrium, budget balance, Markov games with private information, folk theorems with private.

1. INTRODUCTION

THE RENOWNED VICKREY–GROVES–CLARKE (VCG) mechanism established the existence of an incentive-compatible, efficient mechanism for a general class of static mechanism design problems. In these problems, a public decision must be taken that affects the payoffs of agents, and agents have private information about the costs and benefits of the alternative decisions. The VCG mechanism provides incentives for truthful reporting of private information under the assumption that values are private (other agents' private information does not directly affect an agent's payoff) and that preferences are quasilinear so that incentives can be provided using monetary transfers. Subsequently, a pair of classic papers, Arrow (1979) and d'Aspremont and Gerard-Varet (1979) (AGV), constructed an efficient, incentive-compatible mechanism in which the transfers were also budget-balanced, using the solution concept of Bayesian-Nash equilibrium, under the additional assumption that private information is independent across agents. Together, these results have served as benchmarks and formed the building blocks for a large literature on static mechanism design as well as on folk theorems in repeated games with private information.

Of course, many real-world allocation problems are dynamic in nature. Private information may evolve over time, as when a firm's production costs change, and it may be influenced by allocation decisions, as when firms who produce more learn by doing. In this paper, we extend both the VCG and the

AGV mechanisms to general dynamic settings, under the natural extensions of the private values and independence assumptions considered by these authors. In particular, we consider a general infinite-horizon dynamic model in which agents observe private signals over time and decisions are made over time, with the distribution of private signals affected by both past signals and past decisions. Agents' payoffs are quasilinear, but otherwise can depend on the signals and decisions in an arbitrary way. Under the assumption of private values, we construct an efficient, incentive compatible mechanism that generalizes the VCG mechanism and under the additional assumption that private information is independent across agents conditional on the past public actions, we construct a mechanism where the transfers are budget- balanced. We also show that when agents are sufficiently patient and the efficient decision policy induces an ergodic finite Markov chain over types, the mechanism can be made self-enforcing (so that all the decisions and payments are chosen by the agents themselves without an external enforcer). In particular, the mechanism will satisfy a strong form of participation constraints.

We begin by ignoring budget balance, and observing that under private values, it is possible to induce truthtelling using the dynamic "team mechanism," where transfers give each agent the sum of the other agents' utilities in each period. Such transfers make each agent the residual claimant for total surplus and provide him with the incentive to be truthful as long as the mechanism prescribes an efficient decision rule. The problem with the team mechanism is that it is not budget-balanced. As we illustrate using a simple example in Section 3, a naive attempt to balance the budget using the idea of the AGV mechanism runs into the following difficulty. In a static setting, the AGV mechanism gives every agent an incentive to report truthfully given his beliefs about opponents' types, by giving him a transfer equal to the "expected externality" his report imposes on the other agents. Thus, an agent's current beliefs about opponents' types play an important role in determining his transfer. However, in a dynamic setting, these beliefs evolve over time as a function of opponent reports and the decisions those reports induce. If the transfers are constructed using the agents' prior beliefs at the beginning of the game, the transfers will no longer induce truthful reporting after agents have gleaned some information about each other's types. If, instead, the transfers are constructed using beliefs that are updated using earlier reports, this will undermine the incentives for truthful reporting at the earlier stages.

Despite these difficulties, we show that dynamic efficiency can be implemented with balanced budget in the case of independent types and private values. We construct a mechanism that achieves this, a mechanism that we call the *balanced team mechanism*. This mechanism sustains an equilibrium in truthful strategies by giving each agent in each period an incentive payment equal to the *change* in the expected present value (EPV) of the other agents' utilities that is induced by his current report. We show that on the one hand, these incentive payments cause each agent to internalize the expected externality imposed on the other agents by his reports. On the other hand, we show that the expected

incentive payment to an agent is zero when he reports truthfully no matter what reporting strategies the other agents use. The latter property allows us to balance the budget by letting the incentive payment to a given agent be paid by the other agents without affecting those agents' reporting incentives.

We also show that the balanced team mechanism can be made self-enforcing in the infinite-horizon setting in which the agents are sufficiently patient and the induced stochastic process is an ergodic finite Markov chain. Intuitively, the ergodicity condition ensures that agents' private information in a given period is not "too" persistent and so has relatively little effect on the continuation payoffs of patient agents (thus ruling out, for example, the polar case of perfectly persistent types, where the logic of Myerson and Satterthwaite (1983) implies that budget-balanced, incentive-compatible transfers cannot, in general, be made self-enforcing). Under these conditions, the payments in the balanced team mechanism stay bounded even as the continuation payoffs grow with the agents' patience. Hence we can ask agents to implement all the decisions and payments without relying on an external enforcer, instead punishing any detected deviation with a breakdown in cooperation.[2]

2. RELATED LITERATURE

Much of the existing literature on dynamic mechanism design has avoided dealing with the problem of contingent deviations by focusing on one of the following simple cases: (i) a single agent with private information (e.g., Courty and Li (2000), Battaglini (2005)), (ii) a continuum of agents with independent and identically distributed (i.i.d.) private information whose aggregate is predictable (e.g., Atkeson and Lucas (1992)), or (iii) information that is independent across periods, and preferences and technology that are time-separable (e.g., Fudenberg, Levine, and Maskin (1994), Wang (1995), Athey and Bagwell (2001), Athey and Miller (2007), and Miller (2012), but see Athey and Bagwell (2008) for an exception).[3] In each of these cases, an individual agent learns nothing in the course of the mechanism about the others' types that is relevant for the future, hence there is no need to consider contingent deviations. In more general settings, however, even if the mechanism hides the agents' reports from each other, an agent would typically be able to infer something about the other agents' types from the prescribed decisions and try to exploit this information in contingent deviations. Thus, a dynamic mechanism has to

---

[2]This conclusion obtains when the discrete-time stochastic process is held fixed while the parties grow patient. It need not apply to situations where we hold fixed a continuous-time stochastic process and increase the frequency of interactions, thus increasing the persistence of the discrete snapshots of the stochastic process at the same time as we increase the discount factor (see Skrzypacz and Toikka (2012)).

[3]Part of the literature on dynamic contracting considers the case of imperfect commitment (e.g., Bester and Strausz (2001), Battaglini (2007), and Krishna and Morgan (2008)). We sidestep this issue by allowing the agents to commit to the mechanism in advance.

satisfy more incentive constraints than a corresponding static mechanism, in which no information leaks out.

Among the literature that has considered more general dynamic games, most studies have analyzed incentive compatibility in more special settings than that considered in this paper. For example, Pavan, Segal, and Toikka (2014) used a first-order approach to characterize incentive-compatible and profit-maximizing mechanisms in situations where each agent observes a one-dimensional real-valued signal in every period (see also references therein). In contrast, this paper makes no assumptions about the nature of agents' private information, but instead of a general characterization of incentive-compatible mechanisms, it focuses on achieving two goals: efficiency and budget balance.

Subsequent to our working paper (Athey and Segal (2007a)), Bergemann and Valimaki (2010) proposed an alternative efficient dynamic mechanism, which is not budget-balanced and which requires the property of independent types, in contrast to our Proposition 1.[4] In fact, Pavan, Segal, and Toikka (2014) established that in continuous-type settings in which the first-order approach is valid and in which the agents' types follow independent and exogenous stochastic processes, all the mechanisms that implement the same decision rule must give every type sequence of an agent the same expected present value of payments, up to a constant, where the expectation is taken over the other agents' type sequences. Hence, under these assumptions, all the efficient mechanisms can only differ in either the "constants" chosen (e.g., Bergemann and Valimaki's (2010) mechanism chooses the constants to satisfy agents' ex post participation constraints and "efficient exit" conditions) or in how the expected payments are "spread" into the ex post payments (e.g., our balanced team mechanism spreads the payments so as to satisfy ex post budget balance), or both.

3. A TWO-PERIOD EXAMPLE

Consider a seller (agent 1) and a buyer (agent 2) who engage in a two-period relationship. In each period $t = 1, 2$, they can trade a contractible quantity $x_t \in [0, 1]$. Before the first period, the seller privately observes a random type $\tilde{\theta}_1$ in $[1, 2]$, whose realization θ_1 determines his cost function $\frac{1}{2}\theta_1(x_t)^2$ in each period $t = 1, 2$. The buyer's value per unit of the good in period 1 is equal to 1, and in period 2 it is given by a random type $\tilde{\theta}_2$ in $[0, 1]$ whose realization she privately observes between the periods.

An efficient (surplus-maximizing) mechanism must have trading decisions x_1 and x_2 determined by the decision rules $\chi_1(\theta_1) = 1/\theta_1$ and $\chi_2(\theta_1, \theta_2) =$

---

[4]Earlier versions of non-budget-balanced efficient mechanisms for more special dynamic settings were proposed by Friedman and Parkes (2003), Schwarz and Sonin (2003), Bapna and Weber (2005), and Cremer, Spiegel, and Zheng (2009).

θ_2/θ_1, respectively. Note, in particular, that the first-period trade will reveal the realization of the seller's realized type θ_1 to the buyer.

The problem of designing an efficient mechanism comes down to designing transfers to each agent as a function of their reports. In this simple setting, each agent makes only one report. Let us first consider the AGV mechanism for this problem, for the case where the buyer does not learn the seller's type before making his announcement. To give the buyer an incentive for truthful reporting, she is charged an "incentive payment" equal to the expected externality he imposes on the seller, that is, the seller's expected cost

$$\gamma_2(\theta_2) = -\mathbb{E}_{\tilde{\theta}_1}\left[\frac{1}{2}\tilde{\theta}_1\big(\chi_1(\tilde{\theta}_1)\big)^2 + \frac{1}{2}\tilde{\theta}_1\big(\chi_2(\tilde{\theta}_1, \theta_2)\big)^2\right]$$

$$= -\frac{1}{2}\mathbb{E}_{\tilde{\theta}_1}[1/\tilde{\theta}_1] \cdot \big(1 + (\theta_2)^2\big).$$

Similarly, the seller's incentives are provided by paying him an "incentive payment" equal to the expectation of the buyer's utility

$$\gamma_1(\theta_1) = \mathbb{E}_{\tilde{\theta}_2}\big[\chi_1(\theta_1) + \tilde{\theta}_2 \cdot \chi_2(\theta_1, \tilde{\theta}_2)\big] = \big(1 + \mathbb{E}_{\tilde{\theta}_2}[(\tilde{\theta}_2)^2]\big)/\theta_1.$$

Now, since each party's incentive payment does not depend on the other's report, we can balance the budget simply by charging each party's incentive payment to the other party, that is, letting the total transfer to each agent i be $\psi_i(\theta_i, \theta_{-i}) = \gamma_i(\theta_i) - \gamma_{-i}(\theta_{-i})$.

Now we turn to the case of interest, where the buyer makes his announcement after the seller's type θ_1 is revealed. If we use the AGV transfers described above, the buyer anticipates that the second-period trade will be determined by $\chi_2(\theta_1, \cdot)$. However, the buyer must pay (through $\gamma_2(\theta_2)$) the expectation (over $\tilde{\theta}_1$) of the cost of the seller, rather than the seller's actual cost. Then if the seller's type θ_1 is known to be high, the buyer has the incentive to induce inefficiently high trade by "overreporting" his value. Similarly, the buyer does not internalize the benefit of an unexpectedly low cost, and in that case he "underreports" his value to induce less-than-efficient trade.

To fix this problem, we could instead give the buyer an incentive transfer based on the actual externality he imposes given the seller's report: $\tilde{\gamma}_2(\theta_1, \theta_2) = -\frac{1}{2\theta_1} \cdot (1 + (\theta_2)^2)$. This will give the buyer the incentive to report θ_2 truthfully, no matter what θ_1 the seller reports. However, this transfer depends on the seller's report θ_1. Thus, if we attempted to balance the budget by having it be paid by the seller, making his total transfer $\tilde{\psi}_1(\theta_1, \theta_2) = \gamma_1(\theta_1) - \tilde{\gamma}_2(\theta_1, \theta_2)$, the seller would want to reduce $\tilde{\gamma}_2(\theta_1, \theta_2)$ by overreporting his cost θ_1 in period 1, thus exaggerating the externality imposed on him by the buyer.

The problem of contingent deviations illustrated here arises not only when types are persistent as in the above example, but in any dynamic setting in

which the agents' preferences and/or technology are not separable across periods.

In this paper, we propose a different way to construct transfers that resolves the problem. Similarly to the AGV mechanism, our construction proceeds in two steps: (i) construct incentive transfers $\gamma_1(\theta)$, $\gamma_2(\theta)$ to make each agent report truthfully if he expects the other to do so and (ii) charge each agent's incentive transfer to the other agent, making the total transfer to agent i equal $\psi_i(\theta) = \gamma_i(\theta) - \gamma_{-i}(\theta)$. However, in contrast to AGV transfers, the incentive transfer $\gamma_i(\theta)$ to agent i will now depend not just on agent i's announcements θ_i, but also on those of the other agents. How do we then ensure that step (ii) does not destroy incentives? For this purpose, we ensure that even though agent $-i$ can affect the other's incentive payment $\gamma_i(\theta_i, \theta_{-i})$, he cannot manipulate the *expectation* of that payment given that agent i reports truthfully. We achieve this by letting $\gamma_i(\theta_i, \theta_{-i})$ be the *change* in the expectation of agent $-i$'s utility, conditional on all the previous announcements, that is brought about by the report of agent i. (In the general model in which an agent reports in many periods, these incentive transfers would be calculated in each period for the latest report.) No matter what reporting strategy agent $-i$ adopts, if he believes agent i to report truthfully, his expectation of the change in his expected utility due to agent i's future announcements is zero by the law of iterated expectations. Hence agent $-i$ can be charged $\gamma_i(\theta_i, \theta_{-i})$ without affecting his incentives.

In our example, our construction entails giving the buyer an incentive transfer of

$$\gamma_2(\theta_1, \theta_2) = -\frac{1}{2\theta_1} \cdot \left((\theta_2)^2 - \mathbb{E}\left[(\tilde{\theta}_2)^2 \right] \right),$$

which, on the one hand, gives him correct incentives by letting him internalize the seller's expected cost and, on the other hand, ensures that the expectation of this transfer cannot be manipulated by seller: $\mathbb{E}_{\tilde{\theta}_2}[\gamma_2(\theta_1, \tilde{\theta}_2)] = 0$ for any θ_1. Therefore, we can now charge this incentive transfer to the seller— that is, let $\psi_1(\theta_1, \theta_2) = \gamma_1(\theta_1) - \gamma_2(\theta_1, \theta_2)$—without undermining the seller's incentives for truthful reporting. Also, letting then $\psi_2(\theta_1, \theta_2) = -\psi_1(\theta_1, \theta_2) = \gamma_2(\theta_1, \theta_2) - \gamma_1(\theta_1)$ balances the budget and provides incentives for the buyer to report truthfully.

In the rest of this paper, we generalize the idea of using incentive payments that give an agent the change in the EPV of opponent utilities induced by his report to design an efficient mechanism for a general dynamic model. The argument that the EPV of one agent's incentive payments cannot be manipulated by other agents remains the same in the general model, but it becomes more subtle to show that the EPV of an agent's own incentive payments provides him with the correct incentives.

4. THE SETUP

We consider a model with I agents and a countable number of periods, indexed by $t \in \mathbb{N} = \{0, 1, \ldots\}$. In each period t, each agent $i \in \{1, \ldots, I\}$ privately observes his realized private state (type) $\theta_t^i \in \Theta^i$ and then all agents observe the realized verifiable public state $\theta_t^0 \in \Theta^0$. The state space is thus $\Theta = \prod_{i=0}^I \Theta^i$. After the state $\theta_t \in \Theta$ is realized, a public decision $x_t \in X$ is made. (All the above sets will be treated as measurable spaces and the product sets will be endowed with the product measures.) Also, each agent i is given a transfer $y_t^i \in \mathbb{R}$.

The payoff of each agent i is given as a function of the sequences of states $(\theta_t)_{t=0}^\infty$, decisions $(x_t)_{t=0}^\infty$, and monetary transfers $(y_t)_{t=0}^\infty$, as follows,

$$\sum_{t=0}^\infty \delta^t [u^i(x_t, \theta_t) + y_t^i],$$

where $\delta \in (0, 1)$ is a discount factor, and the functions $u^i : X \times \Theta \to \mathbb{R}$ are assumed to be measurable and bounded.

The initial state $\theta_0 \in \Theta$ is assumed to be publicly known. The distribution of subsequent states is governed by transition probability measure $\mu : X \times \Theta \to \Delta(\Theta)$ (where the function μ is measurable). Specifically, for any period $t \geq 0$, given the current state θ_t and the current decision x_t, the next-period state is a random variable $\tilde{\theta}_{t+1}$ distributed according to the probability measure $\mu(x_t, \theta_t) \in \Delta(\Theta)$ (so that for any measurable set $A \subseteq \Theta$, $\mu(x_t, \theta_t)(A)$ is the probability that $\tilde{\theta}_{t+1} \in A$).

We note that our Markov formulation is without loss of generality, since any dynamic model can be described using Markov notation by defining large enough private states to memorize complete histories of private signals, and a public state large enough to memorize calendar time and public decisions (see Appendix A for details).

We consider mechanisms in which, following a publicly observed initial state $\theta_0 \in \Theta_0$, a decision $x_0 \in X$ is made, and then in each period $t \geq 1$, every agent i makes a public report of his private state $\theta_t^i \in \Theta^i$, then the public state $\theta_t^0 \in \Theta^0$ is observed, and based on the reports and the public state, a decision $x_t \in X$ is implemented and a transfer $y_t^i \in \mathbb{R}$ is made to each agent i.[5] The *truthtelling strategy* of agent i always reports his current state θ_t^i in every period $t \geq 1$ truthfully, regardless of the observed past (in particular, regardless of whether he has lied in the past). We will consider perfect Bayesian equilibria (PBE) in truthtelling strategies, with beliefs that assign probability 1 to the other agents' latest reports being truthful.[6]

---

[5]See Appendix A for a discussion of the way we have specified timing in the Markov formulation.

[6]Our focus on sustaining truthful strategies in a direct revelation mechanism is innocuous since our goal is to propose particular mechanisms rather than to characterize the set of all possible

5. THE TEAM MECHANISM

A *decision policy* is a measurable function $\chi : \Theta \to X$, where $\chi(\theta)$ represents the decision made when the realized state in this period is θ. (Note that the focus on Markov decision policies is again without loss of generality, since the state space can be expanded to remember everything that a policy conditions upon.) Starting from an initial state $\theta_0 \in \Theta$, a decision policy χ together with the transition probability measures μ uniquely determine a probability measure over the sequence of states $(\theta_t)_{t=0}^{\infty} \in \Theta^{\mathbb{N}}$.[7]

We say that a measurable policy function $\chi^* : \Theta \to X$ is an *efficient decision policy* if it maximizes the total expected surplus for any starting state. Formally, we denote the total surplus in a period as a function of the period's decision x and state θ by $s(x, \theta) = \sum_{i=1}^{I} u^i(x, \theta)$. Then we can characterize an efficient decision policy χ^* and the associated measurable *social value function* $S : \Theta \to \mathbb{R}$ recursively using the principle of dynamic programming:[8,9]

$$(1) \qquad S(\theta) = s\big(\chi^*(\theta), \theta\big) + \delta \mathbb{E}^{\mu(\chi^*(\theta), \theta)}\big[S(\tilde{\theta})\big]$$
$$= \sup_{x \in X}\big\{s(x, \theta) + \delta \mathbb{E}^{\mu(x, \theta)}\big[S(\tilde{\theta})\big]\big\} \quad \text{for all } \theta \in \Theta.$$

The team mechanism consists of an efficient decision policy χ^* together with the transfer functions that pay each agent the sum of the other agents' utilities according to the announcements:

$$(2) \qquad \psi^{*i}(\theta) = \sum_{j \neq i} u^j\big(\chi^*(\theta), \theta\big).$$

DEFINITION 1: We have *private values* if the utility $u^i(x, \theta)$ of each agent i depends only on x, θ^0, and θ^i.

---

mechanisms. However, if we were interested in such a characterization, we could make use of a revelation principle proposed by Myerson (1986) for general dynamic games. In this revelation principle, agents' reports are kept private, so as to minimize the scope for contingent deviations. In contrast, the mechanisms proposed in this paper are incentive compatible even with all the reports being public and, therefore, would remain incentive compatible if agents were only shown imperfect signals of the reporting history.

[7]The existence and uniqueness of this measure follow from the Tulcea product theorem (Pollard (2002, Chapter 4, Theorem 49)).

[8]Due to our assumption of bounded flow payoffs and discounting, this characterization follows from Theorems 9.2 and 9.4 and Exercise 9.2 in Stokey and Lucas (1989). For sufficient conditions for (1) to have a measurable solution, see Stokey and Lucas (1989, Chapter 9). When the sets Θ and X are finite, (1) always has a (trivially measurable) solution (see Puterman (1994)).

[9]The notation $\mathbb{E}^{\nu}[\tilde{\theta}]$ denotes the expected value of a random variable $\tilde{\theta}$ distributed according to a probability measure $\nu \in \Delta(\Theta)$.

This property simply means that each agent can calculate his payoff as a function of his own information.[10]

PROPOSITION 1: *Suppose that we have private values. Then in the team mechanism, truthtelling strategies form a perfect Bayesian equilibrium. Furthermore, truthtelling strategies form a subgame-perfect equilibrium of the modified game in which the private types are publicly observed.*

PROOF: Since the result for publicly observed types implies the result for private types, we establish the former. By the one-stage deviation principle,[11] it suffices to check that a one-stage deviation of an agent i to reporting $\hat{\theta}^i$ in any given period t when the true state is θ is unprofitable. By private values, the agent's period-t team transfer from the deviation only depends on $\hat{\theta}^i$ through its impact on other agents' payoffs through the public decision,

$$\psi^{*i}(\hat{\theta}^i, \theta^{-i}) = \sum_{j \neq i} u^j(\chi^*(\hat{\theta}^i, \theta^{-i}), (\hat{\theta}^i, \theta^{-i})) = \sum_{j \neq i} u^j(\chi^*(\hat{\theta}^i, \theta^{-i}), \theta),$$

and so his period-t payoff from the deviation is $u^i(\chi^*(\hat{\theta}^i, \theta^{-i}), \theta^i) + \psi^{*i}(\hat{\theta}^i, \theta^{-i}) = s(\chi^*(\hat{\theta}^i, \theta^{-i}), \theta)$. Furthermore, reversion to truthtelling in period $t+1$ will yield flow payoffs of $u^i(\chi^*(\theta_\tau), \theta_\tau^i) + \psi^{*i}(\theta_\tau) = s(\chi^*(\theta_\tau), \theta_\tau)$ in the team mechanism in all periods $\tau \geq t+1$, whose EPV evaluated in period $t+1$ is $S(\theta_{t+1})$. Hence, the agent's period-t EPV of his deviation payoff can be expressed as $s(\chi^*(\hat{\theta}^i, \theta^{-i}), \theta) + \delta \mathbb{E}^{\mu(\chi^*(\hat{\theta}^i, \theta^{-i}), \theta)}[S(\tilde{\theta})]$. By (1), this expression is maximized by reporting $\hat{\theta}^i = \theta^i$, which induces the optimal decision $\chi^*(\theta)$.   *Q.E.D.*

----

[10]It is well known that when the property of private values is violated, efficiency may not be implementable even in a static setting (see, e.g., Jehiel and Moldovanu (2001)). Note, however, that the assumption of private values allows one agent's type to affect other agents' future payoffs through the stochastic process, as long as each agent observes his realized payoffs. For example, this assumption is satisfied in the setting of Mezzetti (2004), whose proposed efficient mechanism is a special case of our team mechanism.

[11]The one-stage deviation principle obtains by applying the principle of optimality to an agent's discrete-time dynamic stochastic program in which his state includes everything he observes, and his action (report) along with his current state determines the probability distribution of his next state (holding fixed the other agents' strategies). The principle of optimality implies that to verify that a given strategy is optimal, it is sufficient to verify that, in all states, the agent is deterred from deviating for one period and then following the given strategy. The principle of optimality for discrete-time dynamic stochastic programming can be found in Stokey and Lucas (1989, Theorem 9.2 and Exercise 9.2). While the setting covered by Stokey and Lucas (1989) is more general than the complete-information setting considered in Fudenberg and Tirole (1991, Theorem 4.2), the idea of the proof for the one-stage deviation principle is the same in both cases: it proceeds by first showing, using backward induction, that if one-stage deviations are deterred, finite deviations are deterred as well, and then extending this to infinite deviations using discounting and the assumption of bounded flow payoffs.

The intuition for the proof is that the team transfers make each agent into a claimant for the total expected surplus in each period, which is maximized when all agents adhere to truthful strategies. Hence, no agent has an incentive to deviate. This is a straightforward extension of the Vickrey–Groves–Clarke mechanism to our dynamic model.

The fact that truthtelling is a subgame-perfect equilibrium under publicly observed types means that the mechanism is robust to agents' observation of each other's current types, i.e., truthtelling is a *within-period ex post equilibrium*. Note that requiring the mechanism to be robust to observation of *future* types would be too strong for the dynamic setting, even with a single agent: the agent might want to report differently and induce different decisions if he could foresee his future types.[12,13]

6. BALANCING

A major problem with the team mechanism is that its transfers are not budget-balanced. However, we now show that they can be balanced. Rather than focusing on the team mechanism, however, we assume that we are given a general direct Markov mechanism with a measurable decision policy $\chi : \Theta \to X$ and a bounded measurable transfer policy $\psi : \Theta \to \mathbb{R}^I$.[14] Supposing that truthtelling is a PBE of this mechanism, we construct a new mechanism that implements the same decision policy χ in a PBE, but whose transfers are budget-balanced. Our construction requires that the transition probabilities have the following property:

---

[12]One exception is given by settings in which the transition probabilities $\mu(x, \theta)$ are independent of decisions x and so an efficient decision policy χ^* simply maximizes the current surplus $s(x, \theta)$. In such settings, truthtelling is a subgame-perfect equilibrium even if the agents observe all the information before the game starts. In more general settings, Pavan, Segal, and Toikka (2014) found that it is sometimes possible to construct mechanisms that are robust to an agent's observation of *other agents'* future types, but not of his own future types.

[13]There does exist a "detail-free" version of the team mechanism if the mechanism can be supplemented by "cheap talk" among agents. In this version, agents report to the mechanism their realized nonmonetary payoffs (instead of states) and the mechanism gives each agent a transfer equal to the sum of the other agents' reported payoffs. (These transfers make the agents into a "team" in the terminology of Marschak and Radner (1972).) In addition, agents make non-verifiable public announcements of their states and implement decisions upon observing these announcements. The resulting game will have an equilibrium that implements an efficient decision policy. However, this mechanism heavily relies on common knowledge and coordination among agents, and requires (at least one of) them to be able to calculate an efficient decision policy.

[14]Recall that Markovness is merely for notational convenience, since we do not impose any restrictions on the state space. As for the boundedness restriction, we use it to ensure that transfers have an expected present value in the sense of double Lebesgue integration, permitting us to interchange expectation and infinite-horizon summation in calculating this expected present value.

DEFINITION 2: We have *independent types* if the transition probability measure can be written in the form $\mu(x, \theta) = \mu^0(x, \theta^0) \cdot \prod_{i=1}^{I} \mu^i(x, \theta^0, \theta^i)$ for each $x \in X$, $\theta \in \Theta$, where $\mu^0 : X \times \Theta^0 \to \Delta(\Theta^0)$ and $\mu^i : X \times \Theta^0 \times \Theta^i \to \Delta(\Theta^i)$ for each agent i.

This definition means that, conditional on decisions and public states, an agent's private information does not have any effect on the distribution of the current and future types of other agents or of the public states. (Note that it still allows one agent's reports to affect the future types of other agents as well as the future public state through the implemented decisions.)

To construct the balanced transfers, first let $\Psi^i(\theta)$ denote the EPV of agent i's transfers in the original mechanism in state θ, which is characterized by the recursive equation

$$(3) \qquad \Psi^i(\theta) = \psi^i(\theta) + \delta \mathbb{E}^{\mu(\chi(\theta), \theta)}\big[\Psi^i(\tilde{\theta})\big].$$

Now we define the following "balanced transfers" $\bar{\psi} : \Theta \times \Theta \to \mathbb{R}^I$, where $\bar{\psi}(\theta, \bar{\theta})$ describes the transfers paid when the current state is θ and the previous state was $\bar{\theta}$:[15]

$$(4) \qquad \bar{\psi}^i(\theta, \bar{\theta}) = \gamma^i(\theta^i, \bar{\theta}) - \frac{1}{I-1} \sum_{j \neq i} \gamma^j(\theta^j, \bar{\theta}), \quad \text{where}$$

$$(5) \qquad \gamma^i(\theta^i, \bar{\theta}) = \mathbb{E}^{\mu(\chi(\bar{\theta}), \bar{\theta})}\big[\Psi^i(\theta^i, \tilde{\theta}^{-i})\big] - \mathbb{E}^{\mu(\chi(\bar{\theta}), \bar{\theta})}\big[\Psi^i(\tilde{\theta})\big].$$

To understand the balanced transfers (4), note that $\gamma^j(\theta^j, \bar{\theta})$, interpreted as agent j's incentive term, gives the *change* in the EPV of agent j's transfers in the original mechanism that results from his current report given the previous state $\bar{\theta}$. As in the standard AGV mechanism, all the other agents pitch in the same amount $\gamma^j/(I-1)$ to pay agent j's incentive term, which ensures that the transfers are budget-balanced: $\sum_{i=1}^{I} \bar{\psi}^i(\theta, \bar{\theta}) = 0$ for all $\theta, \bar{\theta} \in \Theta$. (Also, note that if the original transfers ψ are bounded by K, then their EPVs Ψ are bounded by $K/(1-\delta)$, so the balanced transfers $\bar{\psi}^i$ are bounded by $4K/(1-\delta)$. This ensures that the balanced transfers have an EPV in the sense of double Lebesgue integration and that the one-stage deviation principle applies to the mechanism; see footnote 11.)

PROPOSITION 2: *If truthtelling is a PBE of the Markov mechanism (χ, ψ) and we have independent types, then truthtelling is a PBE of the balanced mechanism $(\chi, \bar{\psi})$, where transfers $\bar{\psi}$ are given by (4).*

---

[15]These transfers can be "Markovized" by expanding the state to be $(\theta, \bar{\theta})$, but we do not do it.

PROOF: For any agent i, $\theta^i \in \Theta$, and $\bar{\theta} \in \Theta$, let $\gamma^i_+(\theta^i, \bar{\theta}) = \mathbb{E}^{\mu(\chi(\bar{\theta}), \bar{\theta})}[\Psi^i(\theta^i, \tilde{\theta}^{-i})]$ and $\gamma^i_-(\bar{\theta}) = \mathbb{E}^{\mu(\chi(\bar{\theta}), \bar{\theta})}[\Psi^i(\tilde{\theta})]$, so that (5) is written as $\gamma^i(\theta^i, \bar{\theta}) = \gamma^i_+(\theta^i, \bar{\theta}) - \gamma^i_-(\bar{\theta})$. Observe that for any $\hat{\theta} \in \Theta$ such that $\hat{\theta}^i = \bar{\theta}^i$ and $\hat{\theta}^0 = \bar{\theta}^0$, using independent types, we have

$$
\begin{aligned}
\mathbb{E}^{\mu(\chi(\hat{\theta}), \bar{\theta})}\big[\gamma^i(\tilde{\theta}^i, \hat{\theta})\big] &= \mathbb{E}^{\mu^i(\chi(\hat{\theta}), \hat{\theta}^0, \hat{\theta}^i)}_{\tilde{\theta}^i}\big[\gamma^i_+(\tilde{\theta}^i, \hat{\theta})\big] - \gamma^i_-(\hat{\theta}) \\
&= \mathbb{E}^{\mu^i(\chi(\hat{\theta}), \hat{\theta}^0, \hat{\theta}^i)}_{\tilde{\theta}^i} \mathbb{E}^{\mu^{-i}(\chi(\hat{\theta}), \hat{\theta}^{-i})}_{\tilde{\theta}^{-i}}\big[\Psi^i(\tilde{\theta}^i, \tilde{\theta}^{-i})\big] \\
&\quad - \mathbb{E}^{\mu(\chi(\hat{\theta}), \hat{\theta})}\big[\Psi^i(\tilde{\theta})\big] \\
&= 0.
\end{aligned}
$$

Hence, using the law of iterated expectations, the time-t expectation of $\gamma^i(\tilde{\theta}^i_\tau, \tilde{\theta}_{\tau-1})$ in every period $\tau \geq t$ is zero if agent i reports truthfully in periods $\tau - 1$ and τ, regardless of the other agents' reports, or (crucially) of agent i's reports in the other periods.

By the one-stage deviation principle (see footnote 11), to verify PBE it suffices to show that for any given t, a one-stage deviation of any agent i to reporting any $\hat{\theta}^i \in \Theta^i$ instead of his true realized type $\theta^i \in \Theta^i$ in period t is unprofitable following any reported state $\bar{\theta} = (\bar{\theta}^i, \bar{\theta}^{-i}) \in \Theta$ in period $t-1$. (The agent believes the reports $\bar{\theta}^j$ of all other agents $j \neq i$ to have been truthful; as for his own previous report $\bar{\theta}^i$, it does not matter whether it was truthful or not, since his true past types are payoff-irrelevant given the observed period-t type θ^i.) Consider the EPV of agent i's balanced transfers following this deviation. By the previous argument, the expectations of $\gamma^j(\tilde{\theta}^j_\tau, \tilde{\theta}_{\tau-1})$ for all agents $j \neq i$ in all periods $\tau \geq t$, as well as the expectations of $\gamma^i(\tilde{\theta}^i_\tau, \tilde{\theta}_{\tau-1})$ in periods $\tau \geq t+2$, are zero. Thus, using independent types, the EPV of agent i's balanced transfers following the deviation is

$$
\begin{aligned}
\gamma^i(\hat{\theta}^i, \bar{\theta}) &+ \delta \mathbb{E}^{\mu^{-i}(\chi(\bar{\theta}), \bar{\theta}^{-i})}_{\tilde{\vartheta}^{-i}} \mathbb{E}^{\mu^i(\chi(\hat{\theta}^i, \tilde{\vartheta}^{-i}), \tilde{\vartheta}^0, \theta^i)}_{\tilde{\theta}^i}\big[\gamma^i(\tilde{\theta}^i, (\hat{\theta}^i, \tilde{\vartheta}^{-i}))\big] \\
&= \gamma^i_+(\hat{\theta}^i, \bar{\theta}) - \gamma^i_-(\bar{\theta}) \\
&\quad + \delta \mathbb{E}^{\mu^{-i}(\chi(\bar{\theta}), \bar{\theta}^{-i})}_{\tilde{\vartheta}^{-i}} \mathbb{E}^{\mu^i(\chi(\hat{\theta}^i, \tilde{\vartheta}^{-i}), \tilde{\vartheta}^0, \theta^i)}_{\tilde{\theta}^i}\big[\gamma^i_+(\tilde{\theta}^i, (\hat{\theta}^i, \tilde{\vartheta}^{-i}))\big] \\
&\quad - \delta \mathbb{E}^{\mu^{-i}(\chi(\bar{\theta}), \bar{\theta}^{-i})}_{\tilde{\vartheta}^{-i}}\big[\gamma^i_-(\hat{\theta}^i, \tilde{\vartheta}^{-i})\big]
\end{aligned}
$$

$$
\tag{6} = \mathbb{E}^{\mu(\chi(\bar{\theta}), \bar{\theta})}_{\tilde{\vartheta}}\big[\Psi^i(\hat{\theta}^i, \tilde{\vartheta}^{-i})\big]
$$

$$
\tag{7} - \mathbb{E}^{\mu(\chi(\bar{\theta}), \bar{\theta})}_{\tilde{\vartheta}}\big[\Psi^i(\tilde{\vartheta})\big]
$$

$$
\tag{8} + \delta \mathbb{E}^{\mu(\chi(\bar{\theta}), \bar{\theta})}_{\tilde{\vartheta}} \mathbb{E}^{\mu(\chi(\hat{\theta}^i, \tilde{\vartheta}^{-i}), \theta^i, \tilde{\vartheta}^{-i})}_{\tilde{\theta}}\big[\Psi^i(\tilde{\theta})\big]
$$

$$
\tag{9} - \delta \mathbb{E}^{\mu(\chi(\bar{\theta}), \bar{\theta})}_{\tilde{\vartheta}} \mathbb{E}^{\mu(\chi(\hat{\theta}^i, \tilde{\vartheta}^{-i}), \hat{\theta}^i, \tilde{\vartheta}^{-i})}_{\tilde{\theta}}\big[\Psi^i(\tilde{\theta})\big].
$$

Expressing $\Psi^i(\hat{\theta}^i, \tilde{\vartheta}^{-i})$ with the recursive formula (3), terms (6) and (9) add up to $\mathbb{E}_{\tilde{\vartheta}}^{\mu(\chi(\bar{\theta}), \bar{\theta})}[\psi^i(\hat{\theta}^i, \tilde{\vartheta}^{-i})]$, which is the agent's expected period-t payment in the original mechanism following the deviation. Term (8) is the EPV of the agent's payments $\psi^i(\tilde{\theta}_\tau)$ in periods $\tau \geq t + 1$ in the original mechanism following the deviation. Finally, (7) does not depend on $\hat{\theta}^i$. Hence, the EPV of agent i's gain from the deviation in the balanced mechanism is the same as in the original mechanism, which is nonpositive by assumption. *Q.E.D.*

We define the *balanced team mechanism* to be a mechanism with an efficient decision plan, together with the balanced team transfers $\bar{\psi}^*$ that are constructed from the team transfers (2) according to (4).

COROLLARY 3: *With independent types and private values*, *the balanced team mechanism has a perfect Bayesian equilibrium in truthtelling strategies*.

7. A SELF-ENFORCING MECHANISM

In this section, we suppose that there is no external enforcer; instead, decisions and payments must be made by the agents themselves. Formally, we consider the *decentralized game*, which differs from the game described in Section 4 as follows: In each period, after each agent i reports his private type θ^i and the public state θ^0 is publicly observed, each agent i chooses a publicly observed action $x_t^i \in X^i$ and makes a publicly observed payment $z_t^{ij} \geq 0$. The agents' actions result in the decision $x_t = (x_t^1, \ldots, x_t^I) \in X = \prod_{i=1}^I X^i$ and the payments result in the total net payment of $y_t^i = \sum_{j \neq i}(z_t^{ji} - z_t^{ij})$ received by each agent i. (Note that $\sum_{i=1}^I y_t^i = 0$ by construction.) We provide sufficient conditions for the decentralized game to have an efficient equilibrium for discount factors close enough to 1. This can be viewed as a folk theorem-like result for Markov games with private states.[16]

To sustain an equilibrium in the decentralized game, agents must be deterred not only from misreporting their private information, but also from choosing incorrect actions and/or payments based on the preceding reports—so-called off-schedule deviations. For example, if agent i has a "nonparticipation action" $\hat{x}^i \in X^i$ that gives him a zero utility regardless of the other actions, then deterring his deviations to this action requires, in particular, that his continuation equilibrium expected payoff always be nonnegative, paralleling the traditional "participation constraints" in static mechanism design.[17] Note, however, that deterring off-schedule deviations is, in general, more difficult than the satisfaction of traditional participation constraints due to (a) the agent's ability to

---

[16]For a folk theorem for Markov games with *public* states, see Dutta (1995).

[17]In a follow-on paper (Athey and Segal (2007b)), we explore weaker sufficient conditions for participation constraints to hold in some special cases of our model.

misreport his type and learn the other agents' types before deciding whether to deviate off-schedule (as in Compte and Jehiel (2009)) and (b) his ability to receive payments from the other agents when he deviates off-schedule (while withholding any payments to them).

The usual way to deter off-schedule deviations is by showing that the short-term gain from any deviation is outweighed by the long-term loss of cooperation surplus when agents are sufficiently patient. When considering the balanced team mechanism, however, we must overcome the problem that the balanced team transfers in each period are expressed through the EPV of future payoffs, and so they could potentially grow with the discount factor, giving agents an incentive to withhold the prescribed payments. However, we show that if the stochastic process forms an ergodic finite Markov chain, the balanced team transfers are bounded *uniformly in the discount factor*. Intuitively, the assumption of ergodicity means that the starting state has a geometrically vanishing effect on the probability distributions of future states and, therefore, on the agents' expected future payoffs, which implies that the effect on the EPV of these payoffs is bounded uniformly in the discount factor. By the construction of balanced team transfers, they inherit this uniform bound, so sufficiently patient agents can be induced to make the prescribed transfers by offering them shares in the expected surplus from future cooperation.[18]

PROPOSITION 4: *Suppose that we have private values and independent types. Suppose also that Θ and X are finite sets, that χ^* is an efficient decision policy for any discount factor $\delta < 1$ close enough to 1, and that χ^* induces a Markov chain that has a unique ergodic set[19] and whose invariant distribution generates a positive expected total surplus. Finally, suppose that there exists an action profile $\hat{x} \in X$ such that $u^i(x^i, \hat{x}^{-i}, \theta) = 0$ for all i, $\theta \in \Theta$, and $x^i \in X^i$.[20] Then there exists $\delta^* < 1$ such that for all $\delta \in (\delta^*, 1)$, decision policy χ^* can be sustained in a PBE of the decentralized game.*

A decision policy that is efficient for δ close enough to 1 is known as a *Blackwell-optimal* policy, and it always exists and can be constructed using a

---

[18]This intuition suggests that ergodicity is crucial for preventing a given period's report from having a long-run impact. To see formally why the ergodicity assumption is indispensable for the proposition below, note that in the contrasting case in which states are perfectly persistent, a self-enforcing mechanism would have to satisfy the same incentive constraints and participation constraints as in the corresponding static model, and these constraints would preclude efficiency in many settings (see Myerson and Satterthwaite (1983), Segal and Whinston (2012)). The same would happen if agents were myopic (i.e., $\delta = 0$), which justifies our focus on patient agents.

[19]Recall that a set of states $S \subseteq \Theta$ is ergodic if, starting in S, the process remains in S with probability 1, and this is not true for any proper subset of S.

[20]Interpreting action $\hat{x}^i \in X^i$ as "quitting" by agent i, this assumption fixes an agent's utility (and normalizes it to zero) when all the other agents quit (but does not restrict his utility when others do not quit). This assumption ensures that, for any beliefs, the stage game has a Bayesian-Nash equilibrium in which all agents quit and make no payments, allowing us to use this equilibrium to punish off-schedule deviations.

simple algorithm (Puterman (1994, Theorem 10.1.4 and Section 10.3)). Similarly, checking the ergodicity of a Markov chain is a well known problem. For example, a Markov chain has a unique ergodic set if and only if there exists a state $\theta \in \Theta$ such that the chain has a positive probability of eventually visiting θ starting from any state $\theta_0 \in \Theta$ (Stokey and Lucas (1989, Theorem 11.2)).[21]

8. CONCLUSION

In conclusion, we mention several extensions of our results. One extension permits agents to take private actions, such as investments or other choices that influence their payoffs directly or through their effects on agents' future types.[22] In our working paper (Athey and Segal (2007a)), we showed that all of the present paper's results extend to this setting, provided that the private values assumption incorporates that an agent's private actions have no direct effect on other agents' payoffs, while the independent types assumption incorporates that these private actions have no direct effect on the distribution of other agents' types. This extends Rogerson's (1992) efficient and budget-balanced mechanism in a two-period setting with private actions.

Our construction also yields an efficient mechanism for settings with dynamic populations, where agents may enter and exit. While our team mechanism has the potentially problematic feature of making transfers to agents after they have exited, under the assumption of independent types, the incentive terms (5) in the balanced team mechanism implement efficiency without making transfers to agents who have exited (by construction, the term is zero for agents who do not have reports).[23]

While we have focused on direct revelation mechanisms for notational simplicity, our Proposition 2 easily extends to mechanisms with other message spaces (for example, with message spaces reduced due to communication costs, as in Fadel and Segal (2009)).

While we have made heavy use of quasilinear utilities, in the case of general payoffs it is possible to approximate the balanced team transfers by transferring the agents' future continuation utilities when they are sufficiently patient (similarly to Fudenberg, Levine, and Maskin's (1994) approximation of static

---

[21]A substantial part of the literature on Markov decision processes (MDPs) has focused on so-called *unichain* MDPs, in which *every* decision policy induces a Markov chain with a unique ergodic set (Puterman (1994, Sections 8.3–8.8)). Even though checking whether a given MDP is unichain is, in general, NP-hard (Tsitsiklis (2007)), it is easy in some important special cases (described in Kallenberg (2002) and Feinberg and Yang (2008)).

[22]In particular, private actions could involve information acquisition about own types, as in Cremer, Spiegel, and Zheng (2009) or as in the model of computational efforts of Larson and Sandholm (2001).

[23]As pointed out by Gershkov and Moldovanu (2009), when agents' types are correlated, it is generally impossible to implement efficiency without making transfers to agents who have exited.

budget-balanced mechanisms in repeated games with serially i.i.d. private information). This approach is used by Horner, Takahashi, and Vielle (2013) to prove a folk theorem for dynamic games with private information.[24]

Propositions 2 and 4 may be viewed as considering the polar opposite cases of "complete contracting" and "no contracting," respectively. Lewis and Kuribko (2010) considered an intermediate case of "incomplete contracting" in which agents can only contract on the current decision and on simple "property rights" that determine their participation constraints in the next period. They constructed a procedure for renegotiating property rights every period that sustains efficiency by mimicking the balanced team mechanism. This result (which requires neither ergodicity nor patience) extends the observation of Cramton, Gibbons, and Klemperer (1987) to a dynamic setting.

Finally, our Propositions 2 and 4 have focused on the setting of independent types. In static mechanism design, this case is viewed as the "hardest" case; with correlated types, efficiency with budget balance is achievable under generic conditions, as shown by d'Aspremont, Cremer, and Gerard-Varet (2004, 2003). We expect these results to extend to dynamic settings, but this would require different techniques than those we have used, so we leave this extension for future research.[25]

APPENDIX A: MARKOVIZATION

Consider a general model in which, in every period t, each agent i observes a private signal $s_t^i \in S_t^i$, a public verifiable signal $s_t^0 \in S_t^0$ is observed, and a decision $x_t \in X$ is made. This model can be "Markovized" (that is, written in Markov notation without changing the underlying game) by defining each agent i's state space as $\Theta^i = \prod_{t=0}^{\infty} S_t^i$ to memorize past signals, and using the public state to memorize calendar time, public signals, and past decisions, so that $\Theta^0 = \mathbb{N} \times \prod_{t=0}^{\infty} S_t^0 \times X^{\mathbb{N}}$. The transition probability measure $\mu(x, (t, \mathbf{s}^0, \mathbf{x}), (\mathbf{s}^i)_{i=1}^I)$ puts probability 1 on the next state $((\tilde{t}, \tilde{\mathbf{s}}^0, \tilde{\mathbf{x}}), (\tilde{\mathbf{s}}^i)_{i=1}^I)$ satisfying $\tilde{t} = t + 1$, $\tilde{x}_t = x$, $\tilde{x}_\tau = x_\tau$ for all $\tau \neq t$, and $\tilde{s}_\tau^i = s_\tau^i$ for all i and all $\tau \neq t + 1$.[26,27]

---

[24]See also Escobar and Toikka (2013), who established such a folk theorem using a different proving method.

[25]For a two-period setting with a period of private actions followed by a period of private signals, budget-balanced implementation of approximate efficiency that exploits correlation was examined by Obara and Rahman (2006).

[26]With a finite horizon T, we can give agent i a total utility of $U^i((x_\tau)_{\tau=1}^T, (s_\tau)_{\tau=1}^T)$ by giving it in the final period T. In an infinite-horizon model, the utility functions $U^i((x_\tau)_{\tau=1}^\infty, (s_\tau)_{\tau=1}^\infty)$ need to be Lipschitz continuous at $t = \infty$ so as to be representable as the present values of bounded utility flows; see Athey and Segal (2007a).

[27]Note, in particular, that the constructed public state s^0 evolves deterministically based on the previous state and the previous decision (which itself will be a deterministic function of the previous public state and agents' reports). Because of this, it makes no difference whether the agents observe the public state before reporting their types or only after (but observe the previous

The advantage of using Markov notation to describe a dynamic model is that, as noted by Townsend (1982), it permits us to restrict attention to sustaining reporting strategies in which every agent reports his whole payoff-relevant state truthfully at all histories, including those in which he lied in the past. To verify that such truthtelling strategies form a PBE, it suffices to establish that one-stage lies followed by reversion to truthtelling are unprofitable at any history.[28] The disadvantage of this approach is that it may result in a very large state space, necessitating the need to check a very large number of incentive constraints. (For example, in the above example, agent i should be induced to report all signals $s_1^i, \ldots, s_t^i$ truthfully in period t, even if he previously misreported some of these signals.) Since our results (with the exception of Section 7) do not depend on the size or structure of the state space, the Markov notation proves convenient for our purposes.[29]

APPENDIX B: PROOF OF PROPOSITION 4

B.1. *Technical Preliminaries*

To bound the incentive payments (5) in a way that is independent of the discount factor δ, we establish the following technical result, which makes use of the fundamental concepts in the theory of Markov chains (see Stokey and Lucas (1989) and Behrends (2000)).

LEMMA 5: *Let Π be a Markov transition matrix on a finite set Θ with a unique ergodic set $S \subseteq \Theta$, let $\lambda \in \Delta(\Theta)$ be the invariant distribution of Π,[30] and let $(\pi_{\theta_0,\theta}^{(t)})_{\theta_0,\theta \in \Theta} = \Pi^t$ (the t-step transition matrix). Then there exists $C \geq 0$ such that for all $\delta \in (0,1)$ and all $\theta, \theta_0 \in \Theta$,*

$$\left| \sum_{t=0}^{\infty} \delta^t \left(\pi_{\theta_0,\theta}^{(t)} - \lambda_\theta \right) \right| \leq C.$$

PROOF: Since Θ is a finite set, it suffices to establish a bound for any given pair $\theta, \theta_0 \in \Theta$. If $\theta \in \Theta \backslash S$, then it is a "transient" state, in which case

---

public state before reporting). In the paper's main model, we make the latter assumption merely for notational simplicity.

[28]In contrast, Myerson's (1986) revelation principle for general settings, in which agents report only *new* information in every period, does not restrict their reporting following lies. To establish that such strategies form a Bayesian-Nash equilibrium would require checking multistage lying deviations, while to establish that they form a PBE would require specifying the agents' sequentially optimal misreporting following own lies.

[29]Our working paper Athey and Segal (2007a) obtains the results of this paper without Markovizing the model, instead asking each agent to report only the new information observed in a given period. The proofs in that paper are sufficiently more complicated notationally, despite being similar substantively.

[30]The invariant distribution is uniquely defined; see Stokey and Lucas (1989, Theorem 11.2).

$\sum_{t=0}^{\infty} \delta^t \pi_{\theta_0,\theta}^{(t)} < \sum_{t=0}^{\infty} \pi_{\theta_0,\theta}^{(t)} < \infty$ and $\lambda_\theta = 0$, hence the desired bound obtains. If instead $\theta \in S$, let $G \subseteq S$ denote the cyclically moving subset of Θ that contains θ, let $t_0 = \min\{t \in \mathbb{N} : \pi_{\theta_0,\theta}^{(t)} > 0\}$, and let $n \geq 1$ denote the period of state θ, so that $\pi_{\theta_0,\theta}^{(t)} > 0$ only if $t = t_0 + rn$, $r \in \mathbb{N}$. Then the n-step transition process Π^n forms an irreducible aperiodic chain on G with the invariant distribution $(n\lambda_\theta)_{\theta \in G}$, and it converges at a geometric rate to its invariant distribution, that is, there exist $\rho \in (0,1)$ and $a \geq 0$ such that $|\pi_{\theta_0,\theta}^{(t_0+rn)} - n\lambda_\theta| \leq a\rho^r$ for all $\theta \in G$. (See Stokey and Lucas (1989, Theorem 11.4) and Behrends (2000, Chapter 7).) Therefore,

$$
\left| \sum_{t=0}^{\infty} \delta^t \left(\pi_{\theta_0,\theta}^{(t)} - \lambda_\theta \right) \right|
$$

$$
\leq \sum_{r=0}^{\infty} \delta^{t_0+rn} \left| \pi_{\theta_0,\theta}^{(t_0+rn)} - n\lambda_\theta \right| + \lambda_\theta \left| n \sum_{r=0}^{\infty} \delta^{t_0+rn} - \sum_{t=0}^{\infty} \delta^t \right|
$$

$$
\leq \sum_{r=0}^{\infty} a\rho^r + \sum_{t=0}^{t_0-1} \delta^t + \delta^{t_0} \sum_{r=0}^{\infty} \delta^{rn} \left(n - \sum_{t=0}^{n-1} \delta^t \right)
$$

$$
\leq \frac{a}{1-\rho} + t_0 + \frac{1}{1-\delta^n}(n - n\delta^{n-1})
$$

$$
\leq \frac{a}{1-\rho} + t_0 + n. \qquad\qquad \text{Q.E.D.}
$$

We apply the lemma to the transition matrix Π of the Markov chain induced by decision policy χ^*. Letting λ be the Markov chain's invariant distribution, letting, for every agent i,

$$
\bar{u}^i = \sum_{\theta \in \Theta} \lambda_\theta u^i \left(\chi^*(\theta), \theta \right)
$$

be the agent's expected flow utility in that distribution, and letting

$$
D^i(\theta_0; \delta) = \sum_{t=0}^{\infty} \delta^t \sum_{\theta \in \Theta} \pi_{\theta_0,\theta}^{(t)} \left(u^i \left(\chi^*(\theta), \theta \right) - \bar{u}^i \right),
$$

the lemma implies that there exists $C \geq 0$ such that for all i, $\theta_0 \in \Theta$, and $\delta \in (0,1)$,

$$
(10) \qquad \left| D^i(\theta_0; \delta) \right| \leq M|\Theta|C, \quad \text{where} \quad M = \max_{i,\theta \in \Theta, x \in X} \left| u^i(x,\theta) \right|.
$$

In particular, since the EPV of team transfers (2) can be expressed as

$$\Psi^{*i}(\theta) = \sum_{j \neq i} \left(D^j(\theta; \delta) + \frac{\overline{u}^j}{1 - \delta} \right),$$

(10) permits us to bound the incentive terms (5) for the team mechanism:

$$(11) \qquad \left| \gamma^{*i}(\theta^i, \bar{\theta}) \right| = \left| \mathbb{E}^{\mu(\bar{\theta}, \chi(\bar{\theta}))} \left[\sum_{j \neq i} \left(D^j(\theta^i, \tilde{\theta}^{-i}; \delta) - D^j(\tilde{\theta}; \delta) \right) \right] \right|$$

$$\leq 2(I - 1)M|\Theta|C.$$

B.2. *Equilibrium Strategies*

Define *on-schedule* histories as those in which all agents in all the previous periods have made decisions and payments consistent with the equilibrium strategies given the reports and the public states, and call all other histories *off-schedule* histories. Consider the following strategies:

• At any off-schedule history, agent i reports his true type, chooses action $\hat{x}^i$, and makes zero payments to all other agents.

• At any on-schedule history, agent i reports truthfully. Also, given the last reports $\hat{\theta}_t$ and the previous reports $\hat{\theta}_{t-1}$ (regardless of whether or not he has been reporting truthfully), he chooses action $x_t^i = \chi^{*i}(\hat{\theta}_t)$ and makes a payment of

$$z^{ij}(\hat{\theta}_t^j, \hat{\theta}_{t-1}) = \frac{1}{I - 1} \gamma^j(\hat{\theta}_t^j, \hat{\theta}_{t-1}) + \frac{\overline{u}^i + M}{I} + 2M|\Theta|C$$

to every agent $j \neq i$. Observe that by construction and (11), for all $\delta \in (0, 1)$ and all $(\hat{\theta}_t^j, \hat{\theta}_{t-1})$,

$$0 \leq z^{ij}(\hat{\theta}_t^j, \hat{\theta}_{t-1}) \leq 4M|\Theta|C + 2M/I.$$

Observe that if all agents follow these strategies, only on-schedule histories are realized and the game implements decision policy χ^*.

B.3. *Verification of PBE*

Now we show that when δ is close enough to 1, the strategies described above form a PBE, coupled with beliefs in which each agent puts probability 1 on the other agents to have reported truthfully in the previous period. First, note

that an agent has no profitable deviation at any off-schedule history, since his continuation payoff is then zero regardless of his beliefs or actions. Now we show that no agent i has a profitable one-stage deviation at any on-schedule history. First we observe that at all on-schedule information sets, no agent i has a profitable misreport. This is because the net prescribed payment to agent i is

$$(12) \qquad \sum_{j \neq i} \left[z^{ji}(\hat{\theta}_t^i, \hat{\theta}_{t-1}) - z^{ij}(\hat{\theta}_t^j, \hat{\theta}_{t-1}) \right] = \bar{\psi}^{*i}(\hat{\theta}_t, \hat{\theta}_{t-1}) - \bar{u}^i + \frac{1}{I} \sum_j \bar{u}^j$$

and so the agent faces a balanced team mechanism (up to a constant payment). Hence, by Proposition 2, if δ is close enough to 1 so that χ^* is an efficient decision policy, truthtelling is optimal at all on-schedule information sets.

Now we show that each agent has no profitable one-stage deviation at an on-schedule public history in period t in which, upon observing his true state θ_t^i and making an announcement $\hat{\theta}_t^i$ (which may or may not be equal to θ_t^i), he makes a public decision and/or payments that differ from those prescribed by his strategy given the announcements. Observe first that the maximal period-t gain from such a deviation is bounded above by

$$(I - 1) \sup_{i,j,\theta,\bar{\theta}} z^{ij}(\theta^j, \bar{\theta}) + 2 \max_{i,\theta,x} |u^i(x, \theta)|$$

$$\leq (I - 1)(4M|\Theta|C + 2M/I) + 2M$$

and that the deviation is followed by zero payoffs in all periods $\tau \geq t + 1$. Now we calculate the time-t EPV of the agent's future payoff if he does not deviate. Given that the total payments to the agent are given by (12) and that the period-t expectation of $\bar{\psi}^{*i}(\theta_\tau, \theta_{\tau-1})$ for $\tau \geq t + 2$ is zero (see the first paragraph in the proof of Proposition 2), the EPV of the agent's future payoff if he does not deviate can be written as

$$\sum_{\tau=t+1}^{\infty} \delta^{\tau-t} \sum_{\theta \in \Theta} \pi_{\theta_t,\theta}^{(\tau-t)} \left[u^i(\chi^*(\theta), \theta) - \bar{u}^i + \frac{1}{I} \sum_j \bar{u}^j \right]$$

$$+ \delta \sum_{\theta \in \Theta} \pi_{\theta_t,\theta} \bar{\psi}^{*i}(\theta, \theta_t)$$

$$= \delta \sum_{\theta \in \Theta} \pi_{\theta_t,\theta} \left[D^i(\theta; \delta) + \bar{\psi}^{*i}(\theta, \theta_t) \right] + \frac{\delta}{1 - \delta} \cdot \frac{1}{I} \sum_j \bar{u}^j.$$

By (10) and (11), the first term on the right-hand side is bounded uniformly in δ, while due to the assumption $\sum_j \bar{u}^j > 0$, the second term goes to infinity as

$\delta \to 1$. Hence, for δ close enough to 1 the expected future payoff sacrificed by the deviation exceeds the maximum possible present gain from it.

REFERENCES

ARROW, K. (1979): "The Property Rights Doctrine and Demand Revelation Under Incomplete Information," in *Economics and Human Welfare*, ed. by M. Boskin. New York: Academic Press. [2463]

ATHEY, S., AND K. BAGWELL (2001): "Optimal Collusion With Private Information," *RAND Journal of Economics*, 32, 428–465. [2465]

——— (2008): "Collusion With Persistent Cost Shocks," *Econometrica*, 76, 493–540. [2465]

ATHEY, S., AND D. MILLER (2007): "Efficiency in Repeated Trade With Hidden Valuations," *Theoretical Economics*, 2, 299–354. [2465]

ATHEY, S., AND I. SEGAL (2007a): "An Efficient Dynamic Mechanism," Working Paper, Stanford University. [2466,2477-2479]

——— (2007b): "Designing Efficient Mechanisms for Dynamic Bilateral Trading Games," *American Economic Review: Papers and Proceedings*, 97, 131–136. [2475]

ATKESON, A., AND R. LUCAS (1992): "On Efficient Distribution With Private Information," *Review of Economic Studies*, 59, 427–453. [2465]

BAPNA, A., AND T. WEBER (2005): "Efficient Dynamic Allocation With Uncertain Valuations," Working Paper, Stanford University. [2466]

BATTAGLINI, M. (2005): "Long-Term Contracting With Markovian Consumers," *American Economic Review*, 95, 637–658. [2465]

——— (2007): "Optimality and Renegotiation in Dynamic Contracting," *Games and Economic Behavior*, 60, 213–246. [2465]

BEHRENDS, E. (2000): *Introduction to Markov Chains, With Special Emphasis on Rapid Mixing*. Braunschweig: Vieweg. [2479,2480]

BERGEMANN, D., AND J. VALIMAKI (2010): "The Dynamic Pivot Mechanism," *Econometrica*, 78, 771–789. [2466]

BESTER, H., AND R. STRAUSZ (2001): "Contracting With Imperfect Commitment and the Revelation Principle: The Single Agent Case," *Econometrica*, 69, 1077–1098. [2465]

COMPTE, O., AND P. JEHIEL (2009): "Veto Constraint in Mechanism Design: Inefficiency With Correlated Types," *American Economic Journal: Microeconomics*, 1, 182–206. [2476]

COURTY, P., AND H. LI (2000): "Sequential Screening," *Review of Economic Studies*, 67, 697–717. [2465]

CRAMTON, P., R. GIBBONS, AND P. KLEMPERER (1987): "Dissolving a Partnership Efficiently," *Econometrica*, 55, 615–632. [2478]

CREMER, J., Y. SPIEGEL, AND C. ZHENG (2009): "Auctions With Costly Information Acquisition," *Economic Theory*, 38, 41–72. [2466,2477]

D'ASPREMONT, C., AND L. GERARD-VARET (1979): "Incentives and Incomplete Information," *Journal of Public Economics*, 11, 25–45. [2463]

D'ASPREMONT, C., J. CREMER, AND L. GERARD-VARET (2003): "Correlation, Independence, and Bayesian Incentives," *Social Choice and Welfare*, 21, 281–310. [2478]

——— (2004): "Balanced Bayesian Mechanisms," *Journal of Economic Theory*, 115, 385–396. [2478]

DUTTA, P. K. (1995): "A Folk Theorem for Stochastic Games," *Journal of Economic Theory*, 66, 1–32. [2475]

ESCOBAR, J., AND J. TOIKKA (2013): "Efficiency in Games With Markovian Private Information," *Econometrica*, 81, 1887–1934. [2478]

FADEL, R., AND I. SEGAL (2009): "The Communication Cost of Selfishness," *Journal of Economic Theory*, 144, 1895–1920. [2477]

FEINBERG, E. A., AND F. YANG (2008): "On Polynomial Cases of the Unichain Classification Problem for Markov Decision Processes," *Operations Research Letters*, 36, 527–530. [2477]

FRIEDMAN, E., AND D. C. PARKES (2003): "Pricing WiFi at Starbucks—Issues in Online Mechanism Design," in *Fourth ACM Conference on Electronic Commerce*. New York: Association for Computing Machinery, 240–241. [2466]

FUDENBERG, D., AND J. TIROLE (1991): *Game Theory*. Cambridge, MA: MIT Press. [2471]

FUDENBERG, D., D. LEVINE, AND E. MASKIN (1994): "The Folk Theorem With Imperfect Public Information," *Econometrica*, 62, 997–1039. [2465,2477]

GERSHKOV, A., AND B. MOLDOVANU (2009): "Learning About the Future and Dynamic Efficiency," *American Economic Review*, 99, 1576–1587. [2477]

HORNER, J., S. TAKAHASHI, AND N. VIELLE (2013): "Truthful Equilibria in Dynamic Bayesian Games," Working Paper, Yale University. [2478]

JEHIEL, P., AND B. MOLDOVANU (2001): "Efficient Design With Interdependent Valuations," *Econometrica*, 69, 1237–1259. [2471]

KALLENBERG, L. C. M. (2002): "Classification Problems in MDPs," in *Markov Processes and Controlled Markov Chains*, ed. by Z. How, J. A. Filar, and A. Chen. Boston: Kluwer, 151–165. [2477]

KRISHNA, V., AND J. MORGAN (2008): "Contracting for Information Under Imperfect Commitment," *RAND Journal of Economics*, 39, 905–925. [2465]

LARSON, K., AND T. SANDHOLM (2001): "Costly Valuation Computation in Auctions," in *Proceedings of the Theoretical Aspects of Reasoning About Knowledge (TARK)*. San Francisco, CA: Morgan Kaufmann, 169–182. [2477]

LEWIS, T., AND N. KURIBKO (2010): "Assigning Property Rights to Ensure Efficient Exchange and Investment," Working Paper, Duke University. [2478]

MARSCHAK, J., AND R. RADNER (1972): *Economic Theory of Teams*. New Haven: Yale University Press. [2472]

MEZZETTI, C. (2004): "Mechanism Design With Interdependent Valuations: Efficiency," *Econometrica*, 72, 1617–1626. [2471]

MILLER, D. (2012): "Robust Collusion With Private Information," *Review of Economic Studies*, 79, 778–811. [2465]

MYERSON, R. (1986): "Multistage Games With Communication," *Econometrica*, 54, 323–358. [2470,2479]

MYERSON, R. B., AND M. A. SATTERTHWAITE (1983): "Efficient Mechanisms for Bilateral Trading," *Journal of Economic Theory*, 29, 265–281. [2465,2476]

OBARA, I., AND D. RAHMAN (2006): "Approximate Implementability With ex post Budget Balance," Working Paper, UCLA Department of Economics. [2478]

PAVAN, A., I. SEGAL, AND J. TOIKKA (2014): "Dynamic Mechanism Design: A Myersonian Approach," *Econometrica* (forthcoming). [2466,2472]

POLLARD, D. (2002): "A User's Guide to Measure-Theoretic Probability," Cambridge: Cambridge University Press. [2470]

PUTERMAN, M. (1994): *Markov Decision Processes*. Hoboken, NJ: Wiley. [2470,2477]

ROGERSON, W. (1992): "Contractual Solutions to the Hold-Up Problem," *Review of Economic Studies*, 59, 777–793. [2477]

SCHWARZ, M., AND K. SONIN (2003): "Auctions and Actions," Working Paper, New Economic School. [2466]

SEGAL, I., AND M. WHINSTON (2012): "Property Rights," in *Handbook of Organizational Economics*, ed. by R. Gibbons and J. Roberts. Princeton: Princeton University Press. [2476]

SKRZYPACZ, A., AND J. TOIKKA (2012): "Patience and Persistence in Repeated Trade," Working Paper, Stanford University. [2465]

STOKEY, N., AND R. LUCAS (1989): *Recursive Methods in Economic Dynamics*. Cambridge: Harvard University Press. [2470,2471,2477,2479,2480]

TOWNSEND, R. (1982): "Optimal Multiperiod Contracts and the Gain From Enduring Relationships Under Private Information," *Journal of Political Economy*, 90, 1166–1186. [2479]

TSITSIKLIS, J. N. (2007): "NP-Hardness of Checking the Unichain Condition in Average Cost MDPs," *Operations Research Letters*, 35, 319–323. [2477]

WANG, C. (1995): "Dynamic Insurance With Private Information and Balanced Budgets," *Review of Economic Studies*, 62, 577–595. [2465]

Dept. of Economics, Graduate School of Business, Stanford University, Stanford, CA 94305-6072, U.S.A.; athey@stanford.edu
and
Dept. of Economics, Stanford University, 579 Serra Mall, Stanford, CA 94305-6072, U.S.A.