An Efficient Dynamic Mechanism*

Susan Athey[†] Ilya Segal[‡]

First Draft: November, 2004 This Draft: February, 2007

Abstract

This paper constructs an efficient, budget-balanced, Bayesian incentive-compatible mechanism for a general dynamic environment with private information. As an intermediate result, we construct an efficient, ex post incentive-compatible mechanism, which is not budget balanced. We also provide conditions under which participation constraints can be satisfied in each period, so that the mechanism can be made self-enforcing if the horizon is infinite and players are sufficiently patient.

In our dynamic environment, agents observe a sequence of private signals over a number of periods (either finite or countable). In each period, the agents report their private signals, and make public (contractible) and private decisions based on the reports. The probability distribution over future signals may depend on both past signals and past decisions. The construction of an efficient mechanism hinges on the assumption of "private values" (each agent's payoff is determined by his own observations). Balancing the budget relies on the assumption of "independent types" (the distribution of each agent's private signals does not depend on the other agents' private information, except through public decisions).

*We are grateful to a co-editor, three anonymous referees, Eric Maskin, Andy Postlewaite, and participants at the fall 2004 Stanford-Berkeley theory fest, the 2007 winter meetings of the AEA, and IDEI Toulouse for helpful comments, and to the National Science Foundation for generous financial support through grants SES-0351500 (Athey) and SES-0318447 (Segal).

[†]Professor of Economics, Harvard University, Littauer Center, 1875 Cambridge St., Cambridge, MA 02138. http://kuznets.fas.harvard.edu/~athey/

[‡]Roy and Betty Anderson Professor in the Humanities and Sciences, Stanford University, Department of Economics, 579 Serra Mall, Stanford, CA 94305-6072. http://www.stanford.edu/~isegal/

1. Introduction

Multi-agent mechanism design has been used to model many important economic settings. However, most of the existing literature on multi-agent mechanism design studies a one-time decision. In reality, a sequence of decisions often needs to be made, but a static mechanism cannot be employed if the parties receive information over time that should affect the decisions, and the agents' preferences and/or technology are not time-separable. For example, parties in a long-term relationship may need to make a sequence of trading and investment decisions in a changing environment. A procurement authority may wish to conduct a sequence of auctions, where bidders have serially correlated values or capacity constraints or learning-by-doing. Ongoing utilization decisions may have to be made for a renewable resource that confers private benefits to individuals. An electricity production schedule may need to be allocated among generating plants with privately known costs that include start-up and shut-down costs. Computational capacity or bandwidth may need to be allocated in a network whose users have private values for various time blocks. When different decisions could be made either simultaneously or sequentially, sequential timing is typically more efficient since later decisions can be made contingent on the information revealed over time. Even in settings in which the resource allocation decision is a one-time decision, the agents' costs of computation (which may be viewed as information acquisition) or communication may be reduced by making the computation and communication decisions sequentially, which helps explain the prevalence of dynamic allocation mechanisms such as iterative auctions.

This paper considers the problem of sustaining incentives, balancing the budget, and satisfying participation constraints in an efficient dynamic mechanism. The additional problem arising in a dynamic mechanism is that it faces more stringent incentive constraints. In a static mechanism, it suffices to prevent deviations in which an agent pretends that he has a different type. In a dynamic mechanism, however, an agent can make his reporting strategy at a point in time contingent on the information he has gleaned about the other agents' types from past interaction, and so each possible contingency has associated incentive constraints.

We construct incentive-compatible dynamic mechanisms for a general infinite-horizon dynamic model in which each agent observes a sequence of private signals over time, and public and private decisions are made over time. The distribution of signals at any point in time may depend on the previously observed signals (e.g., allowing serial correlation) and/or previously made decisions (e.g., allowing investments that stochastically affect payoffs or affect information about payoffs). We assume that agents' payoffs depend on the signals and decisions and are quasilinear in monetary transfers. Also, we make use of two assumptions, that extend "standard" assumptions in static mechanism design to the general dynamic setting: (1) "Private Values," which means that each agent's payoff does not depend on the other agents' private information, and (2) "Independent Values," which means that the distribution of each agent's private signals is independent of the other agents' private information, conditional on past public decisions.

The main result of this paper is that under these two assumptions, the additional incentive constraints do not rule out the implementation of efficient decision plans using budget-balanced transfers. To show this, we construct budget-balanced mechanisms that induce "truthfulness and obedience": they provide agents with the incentives to report private signals truthfully and to obey the mechanism's prescriptions for private decisions at all times and for all possible histories. Budget balance is desirable because there may not be an external source of funds, and burning money is inefficient for the agents. Thus, our result extends to the dynamic setting the static results of Arrow (1979) and d'Aspremont and Gerard-Varet (1979) (AGV), as well as Rogerson's (1992) results incorporating private actions. Under more restrictive conditions, we show that participation constraints can be satisfied as well.

We begin by ignoring both participation constraints and budget balance. We observe that with private values, it is possible to induce truthfulness and obedience throughout the dynamic mechanism using a "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 and obedient as long as the mechanism prescribes an efficient decision rule. This mechanism could be viewed as an extension of the famous Vickrey-Groves-Clarke mechanism to a fully dynamic setting.[1]

The problem with the Team Mechanism is that it is not budget-balanced. As we illustrate using a simple example in Section 2, 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 supports truthtelling in a Bayesian-Nash equilibrium: an agent has an incentive to report truthfully given his beliefs about opponents' types, because the expected value of his transfer is equal to the "expected externality" his report imposes on the other agents through its effect on decisions. Thus, an agent's current beliefs about opponent 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 transfer to one agent is conditioned on information that is revealed by a second agent earlier in the game, the second agent's own incentives to report truthfully will be undermined at that earlier stage.

Despite these difficulties, we show that dynamic efficiency can be implemented with balanced budget in the case of independent private values. We demonstrate a mechanism that achieves

[1] For some special dynamic settings, similar non-budget-balanced efficient mechanisms are proposed by Friedman and Parkes (2003), and subsequently to the present paper, by Bapna and Weber (2005) and Bergemann and Valimaki (2006).

this, a mechanism that we call "the Balanced Mechanism." We construct transfers that sustain an equilibrium in truthful and obedient strategies by giving each agent in each period an incentive payment equal to the *change* in the present value of the expected utilities of the other agents that is induced by his current report. We show that on the one hand, these incentive payments give incentives for truthful reporting and obedience, by letting each agent internalize the expected externality imposed on the other agents by his reports. On the other hand, these incentive payments have the property that the expected incentive payment to opposing agents is zero when he is truthful and obedient, no matter what the others report. Hence, the expected incentive payment to one agent cannot be manipulated by the other agents. This allows us to balance the budget by letting the incentive payment of a given agent be paid by the other agents without affecting those agents' reporting incentives.

A well-known shortcoming of the AGV mechanism is that it need not satisfy the agents' participation constraints. In fact, in some some well-known static cases, there exists no efficient balanced-budget mechanism satisfying participation constraints (e.g., Myerson and Satther-waite (1983), Mailath and Postlewaite (1990)). The same will obviously remain true for dynamic settings with a finite horizon or low patience. For this reason, we focus on an infinite-horizon model, and provide conditions under which participation constraints in each period can be satisfied when agents are sufficiently patient. Intuitively, the conditions ensure 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. Under these conditions, the payments in the Balanced Mechanism are bounded even as the continuation payoffs grow with the agents' patience, hence the mechanism can be made self-enforcing, with agents implementing all decisions and payments without an external enforcer, and punishing any detected deviation with a breakdown in cooperation.

Most 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 i.i.d. private information whose aggregate is predictable (e.g., Atkeson and Lucas (1993)), or (iii) information is independent across periods and preferences and technology are time-separable (Wang (1995), Athey and Bagwell (2001), Athey, Bagwell, and Sanchirico (2004), Levin (2003), Rayo (2003), Miller (2004), Athey and Miller (2004), but see Athey and Bagwell (2004) for an exception).[2] 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

[2]Part of the literature on dynamic contracting considers the case of imperfect commitment (e.g. Bester and Strausz (2001), Battaglini (2003), and Krishna and Morgan (2004)). We sidestep this issue by allowing the agents to commit to the mechanism in advance.

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 satisfy more incentive constraints than a corresponding static one, in which no information leaks out. Satisfying these additional incentive constraints constitutes the main contribution of this paper.

2. Some motivating examples

We consider a series of examples with alternative assumptions.

Example 1. *(Two-Period Trading Game) A seller (agent 1) and a buyer (agent 2) 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 observes a signal* ${}_1 \in [0; 1]$, *which determines his cost function* $\frac{1}{2 \, {}_1} (x_t)^2$ *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 signal* ${}_2 \in [0; 2]$ *that she observes between the periods.*

An efficient (surplus-maximizing) mechanism must have trading decisions x_1 and x_2 determined by the allocation rules ${}_1({}_1) = {}_1$, ${}_2({}_1; {}_2) = {}_1 {}_2$; respectively. Note in particular that the first-period trade will reveal the seller's type ${}_1$ 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 revelation, she is charged an "incentive payment" equal to the expected externality he imposes on the seller, i.e., the seller's expected cost:

$$ {}_2({}_2) = -\mathbb{E}_{\tilde{}_1}\left[\frac{1}{2\tilde{}_1}\left({}_1\left(\tilde{}_1\right)\right)^2 + \frac{1}{2\tilde{}_1}\left({}_2\left(\tilde{}_1; {}_2\right)\right)^2\right] = -\frac{1}{2}\mathbb{E}_{\tilde{}_1}\left[\tilde{}_1\right] \cdot \left(1 + ({}_2)^2\right) : $$

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

$$ {}_1({}_1) = \mathbb{E}_{\tilde{}_2}\left[{}_1({}_1) + \tilde{}_2 \cdot {}_2\left({}_1; \tilde{}_2\right)\right] = {}_1\left(1 + \mathbb{E}_{\tilde{}_2}[(\tilde{}_2)^2]\right) : $$

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, i.e., letting the total transfer to each agent i be ${}_i({}_i; {}_{-i}) = {}_i({}_i) - {}_{-i}({}_{-i})$.

Now we turn to the case of interest, where the buyer makes his announcement after the seller's type ${}_1$ is revealed. If we use the AGV transfers described above, the buyer anticipates that the second-period trade will be determined by ${}_2({}_1; \cdot)$. However, the buyer must pay (through ${}_2({}_2)$) the expectation (over $\tilde{}_1$) of the cost of the seller. Then, if the seller's cost ${}_1$ is known to be higher than average, the buyer "over-reports" his value to induce inefficiently high trade, since the buyer will not have to pay the actual cost of this trade. Similarly, the buyer does not internalize the benefit of an unexpectedly low cost, and in that case he "under-reports" 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: $\tilde{\ }_2(\ _1;\ _2) = -\frac{1}{2}\ _1 \cdot \left(1 + (\ _2)^2\right)$. This will give the buyer the incentive to report $\ _2$ truthfully, no matter what $\ _1$ the seller reports. However, this transfer depends on the seller's report $\ _1$. Thus, if we attempted to balance the budget by having it be paid by the seller, making his total transfer $\tilde{\ }_1(\ _1;\ _2) = \ _1(\ _1) - \tilde{\ }_2(\ _1;\ _2)$, the seller would want to reduce $\tilde{\ }_2(\ _1;\ _2)$ by overreporting his cost $\ _1$ 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: (1) Construct incentive transfers $\ _1(\ )$, $\ _2(\ )$ to make each agent report truthfully if he expects the other to do so, and (2) charge each agent's incentive transfer to the other agent, making the total transfer to agent i equal $\ _i(\ ) = \ _i(\ ) - \ _{-i}(\ )$. However, in contrast to AGV transfers, the incentive transfer $\ _i(\ )$ to agent i will now depend not just on agent i's announcements $\ _i$ but on those of the other agents. How then do we then ensure that step 2 does not destroy incentives? For this purpose, we ensure that even though agent $-i$ can affect the other's incentive payment $\ _i(\ _i;\ _{-i})$, he cannot manipulate the *expectation* of that payment given that agent $-i$ reports truthfully. We achieve this by letting $\ _i(\ _i;\ _{-i})$ be the *change* in the expectation of agent $-i$'s final 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 announcement strategy agent $-i$ adopts, if agent i announces truthfully, the expectation of agent $-i$'s utility follows a martingale with respect to agent i's announcements — that is, the changes have zero expectation. Hence agent $-i$ can be charged $\ _i(\ _i;\ _{-i})$ without affecting his incentives.

In Example 1, our construction entails giving the buyer an incentive transfer of

$$\ _2(\ _1;\ _2) = -\frac{1}{2}\ _1 \cdot \left((\ _2)^2 - \mathbb{E}\left[(\tilde{\ }_2)^2\right]\right);$$

which on the one hand gives him correct incentives by letting him internalize the seller's expected cost (since $\mathbb{E}_{\tilde{\ }_1}\left[\ _2\left(\tilde{\ }_1;\ _2\right)\right] = -\frac{1}{2}\mathbb{E}_{\tilde{\ }_1}\left[\tilde{\ }_1\right] \cdot (\ _2)^2 + \text{const}$), and on the other hand ensures that the expectation of this transfer cannot be manipulated by seller: $\mathbb{E}_{\tilde{\ }_2}\left[\ _2\left(\ _1;\tilde{\ }_2\right)\right] = 0$ for any $\ _1$. Therefore, we can now charge this incentive transfer to the seller — i.e., let $\ _1(\ _1;\ _2) = \ _1(\ _1) - \ _2(\ _1;\ _2)$ —without undermining the seller's incentives for truthful reporting. Also, letting then $\ _2(\ _1;\ _2) = -\ _1(\ _1;\ _2) = \ _2(\ _1;\ _2) - \ _1(\ _1)$ balances the budget and provides incentives for the buyer to report truthfully.

In the rest of this paper, we show that the idea of charging an agent for the change in opponent utilities induced by his report can be applied to yield an efficient mechanism in a general dynamic setting. The idea that agent i's incentive payments cannot be manipulated by others remains the same in the more general model, but it becomes more subtle to show that the anticipation by agent i of his own future incentive payments provides the correct incentives.

The general setting we study includes, for example, the following variations of Example 1.

Example 2. *(Serially Correlated Valuations) Suppose that the buyer privately observes his first-period valuation* ($_{2;1}$) *before the first period, and his second-period valuation* ($_{2;2}$) *before the second period; and that these two valuations are correlated.*[3]

Example 3. *(Decisions affect Valuations) In addition to "exogenous" correlation between the buyer's first and second period valuations, the buyer's consumption level in the first period* (x_1) *affects the probability distribution over his second-period valuation. This might arise due to the buyer's "habit formation."*

Example 4. *(Decisions Affect Information) The buyer never observes his valuation* $_2$, *but after the first period, she observes a signal* s *that is correlated with* $_2$ *and whose probability distribution (e.g., informativeness) depends on her first-period consumption* x_1. *This could describe the buyer's learning about the product by consuming it.*

Example 5. *(Hidden Actions affect Valuations) The buyer makes a private investment* a *in the first period that affects the probability distribution over her second-period valuation. This investment is not observable by the other agents and is not verifiable.*

Example 6. *(Hidden Actions affect Information) The buyer never observes his valuation* $_2$, *but after the first period, she observes a signal* s *that is correlated with* $_2$ *and whose probability distribution (e.g., informativeness) depends on a private investment* a *she makes in the first period. For example, the investment could model the buyer's gathering of information about or computation of his preferences.*

3. The Setup

3.1. Technology and Payoffs

We consider a model with I agents and a countable number of periods, indexed by $t = 1; 2; :::$. (A finite-horizon model with T periods is a special case.) In each period t, each agent $i \in \{1; :::; I\}$

[3] A model of dynamic procurement auctions with serially correlated values has been considered by Athey and Bagwell (2004).

observes a private signal $_{i;t} \in \Theta_{i;t}$, where $\Theta_{i;t}$ is a measurable space.[4] We use $\Theta_i^t = \prod^t{}_{=1}\Theta_{i;}$ to denote the space of possible signal histories of agent i at time t, and let $\Theta_t = \prod_{i=0}^{I} \Theta_{i;t}$ and $\Theta^t = \prod_{i=0}^{I} \Theta_i^t$. We interpret $\Theta_i = \Theta_i^\infty$ as the agent's type space, even though the agent only observes his type over time, and $\Theta = \prod_{i=0}^{I} \Theta_i$ as the state space. All these sets will be treated as measurable spaces with the product sigma-algebra.

In each period $t = 1; :::$, after this period's signals are observed, each agent $i \in \{1; :::; I\}$ makes a private decision $x_{i;t} \in X_{i;t}$, and there is a public (contractible) decision $x_{0;t} \in X_{0;t}$, where $X_{0;t}$ and $X_{i;t}$ for each i are measurable spaces. We let $X_t = \prod{}_{i=0}^{I} X_{i;t}$, $X^t = \prod^t{}_{=1} X$, and $X = X^\infty$.

The distribution of signals at any moment may depend both on the history of observed information $^{t-1} \in \Theta^{t-1}$ and the history of decisions $x^{t-1} \in X^{t-1}$. Thus, the evolution of uncertainty at time t is described by history-contingent probability measures over the period-t signal space Θ_t; denoted $_t|\, x^{t-1};\ ^{t-1}$. For any measurable set $\mathcal{F} \subset \Theta_t$; $_t(\mathcal{F})|\cdot;\cdot$ is a measurable function of the history on $X^{t-1} \times \Theta^{t-1}$.[5;6] (Note that $_1$ is a constant since there is no history in period 1.)

A *decision plan* is a measurable function $: \Theta \to X$, where each $_t(\ )$ represents the decision made at time t. We say that decision plan is *observationally measurable* if it is measurable and for each t, $_t(\ )$ depends only on ^{t} – the information observed by time t. We will sometimes write $_t(\ ^t)$ rather than $_t(\ )$ to emphasize observational measurability. The sequence of decisions up to time t given type vector is denoted $^t(\ ) = (\ _1(\ ); :::;\ _t(\ ))$.

Any decision plan uniquely determines a probability measure over Θ; which we denote $[\ ]$. The existence and uniqueness of this measure follow from the Tulcea product theorem.[7]

---

[4]We could also allow for publicly observed (contractible) signals. Formally, they could be incorporated, e.g., by introducing an additional "agent 0" who observes signals $\theta_{0,t}$ in each period t and always reports them truthfully.

[5]We can view these history-contingent distributions as describing Nature's behavioral strategy in the dynamic game in which Nature chooses types and the mechanism chooses decisions. Alternatively, Nature's mixed strategy could be described in a different way: e.g., in Example 4, we could be given a distribution of the buyer's value θ_2 and a distribution of the second-period signal s conditional on θ_2 for any given first-period decision x_1. From this description, we could deduce an equivalent behavioral strategy, which gives a distribution of the current signal conditional on the history (e.g., of $s|x_1$ and of $\theta|x_1, s$). However, we prefer to start directly with behavioral strategies: The problem with alternative approaches is that conditional distributions are not well-defined on zero-probability events (e.g., when signal spaces are continuous, almost all points would have zero probabilities), while it will be important for our analysis to have well-defined forward expectations at all times and events.

[6]One may want to allow the set X_t of possible decisions at time t and the set Θ_t of possible observations at time t to depend on the history $\left(\theta^{t-1}, x^{t-1}\right)$. (The latter dependence, for example, may occur in situations in which decisions determine which signal is observed, as in Example 4 above). However, these situations could be incorporated in our model simply by letting X_t describe the set of all potential decisions that could be made at time t at all histories, and similarly letting Θ_t be the set of all signals that could be potentially observed at time t at all histories. The actual history then determines which observations from Θ_t and which decisions from X_t are relevant for the agents' payoffs.

[7]Formally, the distributions $\nu_t|\chi_{t-1}\left(\theta^{t-1}\right), \theta^{t-1}$ are probability kernels that define a consistent family

Our results on the existence of a budget-balanced mechanism rely on the following assumption

Definition 1. *We have independent types if for each* t, x^{t-1};*and* $^{t-1}$, *the probability measure* $_{t}| x^{t-1};\ ^{t-1}$ *over* Θ_t *can be written in the form*

$$_{t}| x^{t-1};\ ^{t-1} = \prod_{i=0}^{I}\ _{i;t}| x_0^{t-1}; x_i^{t-1};\ _i^{t-1};$$

for some probability measures $_{i;t}$ *over* $\Theta_{i;t}$ *that depend only on the public decision history* x_0^{t-1} *and on agent* i*'s private history* $(\ _i^{t-1}; x_i^{t-1})$.

This definition means that an agent's private information does not have any stochastic effects on the other agents' private signals, conditional on the public decisions. Note that this assumption still allows one agent's private signals or decisions to affect another agent's future private signals through the implemented public decisions. Thus, the probability measure $[\ ]$ over Θ induced by a given decision rule need not be independent across agents.

Now we describe the agents' payoffs. The payoff of each agent i is given as a function of the stream x of decisions, state , and stream of monetary transfers $y_i = (y_{i;t})_{t=1}^{\infty}$; where $y_{i;t} \in \mathbb{R}$:

$$\sum_{t=1}^{\infty}\ ^{t}\left(u_{i;t}(x^t;\ ^t) + y_{i;t}\right);$$

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

Since each function $u_{i;t}$ in each period t can depend on the entire history of signals and decisions, writing the payoff as a present discounted value is not too restrictive. For example, in a setting with a finite horizon T, any utility functional $U_i(x;\ )$ can be represented as

$$U_i(x;\ ) = \sum_{t=1}^{\infty}\ ^{t}u_{i;t}(x^t;\ ^t); \tag{3.1}$$

simply by letting $u_{i;t} \equiv 0$ for $t \neq T$ and $u_{i;t}\left(x^T;\ ^T\right) \equiv\ ^{-T}U_i(x;\ )$. As for the general infinite-horizon case, in the Appendix we show that a functional $U_i : X \times \Theta \to \mathbb{R}$ can be written in the form (3.1) for some uniformly bounded sequence $(u_{i;t})_{t=1}^{\infty}$ if and only if U_i has the following property, which ensures that decisions and observations in the distant future have a vanishing effect on the payoffs, and which will be important for our analysis:[8]

$\{\mu^T\}_{T=1}^{\infty}$ of probability measures on Θ^T for all finite horizons T. The Tulcea product theorem (see, e.g., Pollard (2002), Chapter 4, Theorem 49) establishes that there exists a unique stochastic process on Θ whose marginals for any finite horizon T coincide with μ^T.

[8] This is a mild strengthening of the usual "continuity at infinity" assumption in infinite-horizon games (e.g., Fudenberg and Tirole (1991a), p. 110).

Definition 2. *A functional* $F : \prod_{t=1}^{\infty} Z_t \to \mathbb{R}$ *is Lipshitz Continuous at Infinity with discount factor* $\in (0;1)$ *(denoted LCI) if it is Lipshitz continuous in the metric* $(z;z') = ^{\inf\{t:z_t \neq z'_t\}}$, *that is, there exists* $C > 0$ *such that* $|F(z) - F(z')| \leq C \quad (z;z')$ *for all* $z;z' \in Z$:

Our main results will depend on the following property of payoffs:

Definition 3. *We have* private values *if each agent* i*'s utility* $U_i(x; \ )$ *depends only on the public decisions* x_0 *and the agent's private history* $(x_i; \ _i)$

This definition simply means that each agent observes his payoff, i.e., can calculate it as a function of his own information. It is less restrictive than appears at first because it allows an agent's *expected* utility conditional on time-t history $\left(\ ^t; x^t \right)$ to depend on the other agents' information, as long his *final* utility is fully determined by his observations (as, e.g., in Mezzetti (2004)).[9] We could also allow some cases in which agent i may not observe his final utility (as, e.g., in Example 4), by letting $U_i(x; \ )$ represent the agent's expectation of his utility conditional on the state , provided that this expectation does not depend on the other agents' private signals or decisions.

3.2. Mechanisms and Strategies

We consider mechanisms in which in each period, agents report their private information, and based on the reports, decisions are implemented, and transfers are made.[10] Formally, a *mechanism* is described with two observationally measurable functions: a decision plan $: \Theta \to X$ and a uniformly bounded transfer plan $: \Theta \to \left(\mathbb{R}^I\right)^{\infty}$, which prescribe the current decisions and transfers to the agents as a function of the reporting history. The public decisions $_0(\ )$ prescribed by the plan are implemented directly, while the prescribed private decisions $_i(\ )$ for agents $i \geq 1$ are nonbinding recommendations that the agents are free to disobey. The transfer plan is implemented directly. Its observational measurability means that the transfer $_{i;t}(\ )$ made to agent i in period t depends only on the current history ^{t}. The total discounted payments in the mechanism with a discount factor $\in (0;1)$ can be calculated as

$$\Psi_i(\ ) = \sum_{t=0}^{\infty} \ ^t \ _{i;t}(\ ^t); \tag{3.2}$$

[9]The property of Private Values is violated in some well-known settings, such as auctions for mineral rights, For such settings, however, it is known that efficiency is in general not implementable even in a static problem, as in Akerlof's (1970) classic example of trading with adverse selection. See also Jehiel and Moldovanu (2001).

[10]Thus, we focus on direct revelation mechanisms in which agents report truthfully in equilibrium. This focus is innocuous since our goal is to propose particular mechanisms rather than characterize everything that is implementable. However, if we were interested in such characterization, we could restrict attention to direct revelation mechanisms with truthful reporting by making use of a revelation principle proposed by Myerson (1986) for general dynamic games.

where uniform convergence of the series is ensured by the uniform boundedness of the transfer plan. Note that the agents' reporting incentives in a mechanism depend on the transfer plan only through the total discounted payments $\Psi(\ )$ and not on how they are spread over time. (However, if agents have opportunities to exit the mechanism during the course of the game, the satisfaction of their participation constraints will depend on the spread of payments - see Section 6 below.)

We say that the mechanism is *budget balanced* if $\sum_i \ _{i,t}(\ ) \equiv 0$ for all t.

In principle, the mechanism also determines what information agents observe about the history of reports. In general, the less information is revealed to an agent, the smaller is the set of contingent deviations available to him, and the easier it is to satisfy his incentive constraints (this idea underlies Myerson's (1986) "communication equilibrium"). In the extreme, if nothing is revealed to the agent, the problem is essentially static. In practice, the agent will observe at least some components of the decisions—e.g., the history of his own consumption—and so he will have access to at least some contingent deviations. We will propose mechanisms in which truthful-obedient play is incentive-compatible even if the agents observe the complete reporting history, and therefore will remain incentive-compatible no matter what the agents observe about the history (so long as agents observe enough information to implement their private decision plans).

The mechanism induces a multiperiod game, in which each period t consists of three stages:

Stage $t.0$: Each agent i privately observes his private signal $\ _{i,t}$.

Stage $t.1$: Each agent i makes a public report $\hat{\ }_{i,t}$.

Stage $t.2$: Each agent i makes a private decision $x_{i,t}$. The mechanism implements the public decision $x_{0,t} = \ _{0,t}\left(\hat{\ }^t\right)$ and the payments $y_{i,t} = \ _{i,t}\left(\hat{\ }^t\right)$ to each agent i.

Now we turn to describing the agents' strategies in the mechanism. By stage $t.1$ each agent i observes the reporting history $\hat{\ }^{t-1} \in \Theta^{t-1}$ and the history of own private signals $\ _i^t \in \Theta_i^t$ and private decisions $x_i^{t-1} \in X_i^{t-1}$. Thus, the agent's reporting strategy can be represented by a function $\ _i : \Theta \times \Theta_i \times X_i \to \Theta_i$, where the report of agent i in period t, $\ _{i,t}\left(\hat{\ };\ _i; x_i\right) = \left(\ _i\left(\hat{\ };\ _i; x_i\right)\right)_t$, is "observationally measurable" in that it depends only on the history of reports $\hat{\ }^{t-1}$ and the agent's private history $\ _i^t; x_i^{t-1}$. (For our results about the "Team Mechanism" below, we will allow deviations to more general strategies $\ _i$ in which $\ _{i,t}$ can depend on the reporting history $\hat{\ }_i^t$ that includes the concurrent reports of the other agents, which we call "measurable with respect to *current* observations.")

In stage $t.2$, each agent i observes in addition the reports from stage $t.1$. Thus, his private decision strategy can be represented by $\ _i : \Theta \times \Theta_i \times X_i \to X_i$, where in each period t, $\ _{i,t}\left(\hat{\ };\ _i; x_i\right) = \left(\ _i\left(\hat{\ };\ _i; x_i\right)\right)_t$, is "observationally measurable" in that it depends only on the reporting history $\hat{}^{t}$ and the agent's private history $_{i}^{t}; x_i^{t-1}$. The complete strategy of agent i is given by $(_i; _i)$.

A given observationally measurable strategy $(_i; _i)$ of agent i induces a *strategic plan* $\left(\bar{}_i; \bar{}_i\right)$, where $\bar{}_i(\ )$ denotes the agent's reports and $\bar{}_i(\ )$ his decisions from following the strategy given that his type is $_i$ and the opponents report $_{-i}$. This plan is constructed recursively, by letting for each $t = 1; ::,$ $\bar{}_{i;t}(\ ) = {}_{i;t}\left(\left(\bar{}_i^{t-1}\left(_i^{t-1}\right); _{-i}^{t-1}\right); _i^{t}; \bar{}_i^{t-1}\left(_i^{t-1}\right)\right)$, and $\bar{}_{i;t}(\ ) = {}_{i;t}\left(\left(\bar{}_i^{t}\left(_i^{t}\right); _{-i}^{t}\right); _i^{t}; \bar{}_i^{t-1}\left(_i^{t-1}\right)\right)$. Observational measurability means that $\bar{}_{i;t}(\ )$ depends only on $\left(_i^{t}; _{-i}^{t-1}\right)$ and $\bar{}_{i;t}(\ )$ depends only on ^{t}. Observe that agent i's expected payoff in the mechanism from using strategy $(_i; _i)$ depends only on the induced strategic plan $\left(\bar{}_i; \bar{}_i\right)$, for any strategies chosen by the other agents. Thus, all the strategies giving rise to the same strategic plan are strategically equivalent in the normal-form representation of the mechanism. (However, they may have different implications for extensive-form refinements, as discussed in Section 3.3 below).

Definition 4. *For a given decision plan , agent i's strategy $(_i; _i)$ is a truthful-obedient strategy if the corresponding strategic plan has $\bar{}_i(\ ) = {}_i$ and $\bar{}_i(\ ) = {}_i(\ )$ for all $\in \Theta$:*

This definition means that if an agent has been truthful and obedient so far he will continue to be truthful and obedient. It does not specify anything about reporting after a lie (i.e., $\hat{}_i^{t-1} \neq {}_i^{t-1}$) or a disobedience (i.e., $x_i^{t-1} \neq {}_i^{t-1}\left(\hat{}^{t-1}\right)$), and so there are many truthful-obedient strategies.

Our proofs will make use of the following simple observation: given a decision plan and transfer function ; when opposing agents use truthful-obedient strategies, any strategy $(_i; _i)$ for agent i induces the same decisions and transfers in every state as those that arise when agent i uses a truthful-obedient strategy but the decision plan and transfers are defined as follows:

$$\hat{}(\ ) \equiv \left(_{-i}(\bar{}_i(\ ); _{-i}); \bar{}_i(\ )\right) \tag{3.3}$$

$$\hat{}(\ ) \equiv \left(\bar{}_i(\ ); _{-i}\right); \text{ and } \hat{\Psi}(\ ) \equiv \sum_{t=0}^{\infty} {}^{t}\,\hat{}_{i;t}(\ ): \tag{3.4}$$

Note that if decision plan and payment plan are observationally measurable, then the decision plan $\hat{}$ and payment plan $\hat{}$ given by (3.3-3.4) are also observationally measurable.

When agent i uses reporting strategy $_i$ with the associated $\bar{}_i$, $\hat{}$; and $\hat{\Psi}$; while the other agents are truthful, agent i's payoff in the mechanism is

$$\mathbb{E}_{\tilde{}}^{[\hat{}]}\left[U_i\left(\hat{}\left(\tilde{}\right); \tilde{}\right) + \hat{\Psi}_i\left(\tilde{}\right)\right]: \tag{3.5}$$

We use $\mathbb{E}_{\tilde{}}$ to denote the expectation of random variable $\tilde{}$ whose probability distribution is .
Note that different reporting strategies by the agent induce different distributions over types, and this is captured by the fact that we take expectations using probability distribution $[\hat{}]$ where $\hat{}$ varies with the reporting strategy.

3.3. Solution Concept

In this paper, we focus on the solution concept of Bayesian Nash Equilibrium (BNE), so that each agent's ex ante choice of a history-contingent reporting strategy must be a best response to the history-contingent reporting strategies of opponents. Thus, the existence of a truthful-obedient Bayesian-Nash equilibrium means that expression (3.5) is maximized when agent i uses a truthful reporting strategy, and so $\hat{\ } =$ and $\hat{\Psi}_i = \Psi_i$.

In the dynamic setting, one may be interested in dynamic equilibrium refinements. Here we focus on Perfect Bayesian Equilibrium (PBE) as defined in Fudenberg and Tirole (1991) and modified to fit our model (which has types learned over time rather than at the beginning of the game, as in their model). We argue that under some conditions, if a truthful-obedient BNE exists, we can also construct a truthful-obedient PBE.

To simplify the discussion, we focus on the special case where the following hold: independent types; decisions and signals are drawn from finite sets in each period; and all signals have positive probabilities for any history. Then, the requirements of PBE are: (i) Player i's belief about the types of other agents is the product of agent i's beliefs about the types of individual agents. These beliefs do not depend upon agent i's type. (ii) Players i and j always have the same beliefs about the type of any third agent k, and agent i's belief about agent j's type is consistent with the understanding that the behavior of any other agent cannot signal this information. (iii) Beliefs are determined by Bayes' rule whenever possible. This includes circumstances in which observed play in a previous period is incompatible with the beliefs and equilibrium strategies for that period. That is, after a zero-probability event, agents form beliefs, but then they continue to use Bayes' rule when possible from there. (iv) Given any history and beliefs, the agents' strategies from that point forward are a Bayesian-Nash equilibrium of the continuation game.

Note that when all signals have positive probabilities, opponents cannot report anything "unexpected," and so all zero-probability information sets of an agent arise following his own deviations. Since the agent cannot signal something to himself that he does not know, his beliefs about the others are unchanged following such a deviation. Then, since the definition of a truthful-obedient strategy does not restrict the agent's behavior following own deviations, we can modify his strategy to be a best-response to such beliefs at this information set, while still being truthful-obedient.[11] Doing this for all information sets and for all agents, we obtain a truthtful-obedient PBE.

[11]To see that a best response is guaranteed to exist, note that since U_i is LCI_δ, the agent's objective is continuous in his strategy $\sigma_i = (\alpha_i, \beta_i)$ in the metric $\rho_\delta(\sigma_i, \sigma_i') = \delta^{\inf\{t:\sigma_{i,t}\neq\sigma'_{i,t}\}}$, and his strategy space is compact in the metric when the action set is finite in each period. To establish compactness, take any sequence of strategies, find a set of first-period strategies that occurs infinitely often, then find a strategy with the same first-period strategy and that has a second-period strategy that is used infinitely often, etc., and this will be a converging subsequence in the metric.

4. The Team Mechanism

We want to implement an efficient decision plan *, that is, one that maximizes the total expected surplus

$$\mathbb{E}_{\tilde{}}^{[\ ]}\left[\sum_i U_i\left(\ (\tilde{\ });\tilde{\ }\right)\right]; \tag{4.1}$$

among all observationally measurable decision plans. The problem of calculating * is a stochastic dynamic programming problem, but for the purpose of this paper we simply assume that a solution exists.[12]

The Team Mechanism consists of the efficient decision plan * together with the transfer functions

$${}^{M}_{i;t}(\ ) = \sum_{j\neq i} u_{j;t}(\ {}^*(\ );\ );$$

which give each agent the sum of other agents' utilities. (Note that the per-period transfers are uniformly bounded since per-period utilities are uniformly bounded.) The resulting total discounted transfers are

$$\Psi^{M}_{i}(\ ) = \sum_{j\neq i} U_j(\ {}^*(\ );\ ):$$

We say that a strategy profile is a *within-period ex post equilibrium* if each agent's reporting strategy is a best response even among strategies that are measurable with respect to *current* observations (recall that in general, we restrict reporting strategies to be measurable only with respect to *past* opponents' reports). This is a refinement of Bayesian-Nash equilibrium, which does not consider deviations based on the opponents' current types. Note that the usual concept of ex post equilibrium considered in the static mechanism design literature allows deviations based on all the information eventually obtained. This would be too strong for the dynamic setting, where, even with one agent, this agent might want to report differently and make different decisions if he could foresee the future signals.[13] A notable exception is given by settings in which there exists an efficient decision rule that is "distribution-free," i.e., maximizes the total surplus for all states of the world, rather than just its expectation. An example is a case where $u_{i;t}\left(x^t;\ {}^t_i\right)$ and ${}_t|x^{t-1};\ {}^{t-1}$ do not depend on the past decisions x^{t-1}, and the efficient

[12] For example, it is guaranteed to exist when all the signal spaces and decision spaces are finite, by the same argument as in footnote 11 above.

[13] The refinement of ex post Nash equilibrium has been considered mainly in static mechanism design, where it requires that an agent's announcement must be a best response to the realized values of opponent announcements, when opponents follow the proposed equilibrium strategies. In a static setting, private-values context ex post equilibrium implementation is equivalent to implementation in dominant strategies, but this equivalence breaks down in a dynamic model where strategies might specify that a player responds to an opponent's announcement in pathological ways. See Miller (2004), who motivates and applies "within-period" ex post equilibrium to dynamic games.

policy $^*_t(\ ^t)$ maximizes period-by-period surplus $\sum_i u_{i;t}(\ ^*_t(\ ^t);\ ^t_i)$. In such settings, the Team Mechanism we propose sustains truthful reporting as a "true" ex post equilibrium.

Proposition 1. *Suppose that we have private values. Then for any mechanism with an efficient decision plan* * *and total discounted transfers* Ψ *that equal the total discounted transfers* $\Psi_i^M(\ )$ *in the Team Mechanism, any truthful-obedient strategy profile is a Bayesian-Nash equilibrium and a within-period ex post equilibrium.*

Proof. Consider a mechanism satisfying the conditions of the proposition. Suppose that agent i deviates to induce a strategic plan $(\bar{\ }_i; \bar{\ }_i)$ that is measurable with respect to current observations, while the other agents use truthful-obedient strategies. The decision and transfer plans following the deviation are given by $\hat{\ }$ and $\hat{\Psi}$ from (3.3) and (3.4), with $\hat{\ }$ being observationally measurable. Note that

$$\begin{aligned}
\hat{\Psi}_i(\ ) &= \sum_{j \neq i} U_j\left(\ ^*_{-i}(\bar{\ }_i(\ );\ _{-i});\ ^*_i(\bar{\ }_i(\ );\ _{-i});\ \bar{\ }_i(\ );\ _{-i}\right) \\
&= \sum_{j \neq i} U_j\left(\hat{\ }_{-i}(\ );\ ^*_i(\bar{\ }_i(\ );\ _{-i});\ \bar{\ }_i(\ );\ _{-i}\right) \\
&= \sum_{j \neq i} U_j(\hat{\ }(\ );\ );
\end{aligned}$$

where the first equality is by construction of the transfers, the second by definition of $\hat{\ }$, and the third by the assumption of private values, which means that U_j does not depend directly on $_i$ or x_i: Thus, agent i's expected payoff from the deviation (3.5) is given by

$$\mathbb{E}_{\tilde{\ }}^{[\hat{\ }]}\left[U_i\left(\hat{\ }\left(\tilde{\ }\right);\tilde{\ }\right) + \sum_{j \neq i} U_j\left(\hat{\ }\left(\tilde{\ }\right);\tilde{\ }\right)\right]:$$

Since $\hat{\ }(\cdot) = \ ^*(\cdot)$ when $(\bar{\ }_i; \bar{\ }_i)$ is a truthful-obedient strategy, and since by construction * maximizes expected total surplus among all observationally measurable decision plans, the deviation cannot raise agent i's expected payoff. ■

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

A simple way to implement the Team Mechanism is by asking the agents to report their realized payoffs (i.e., viewing these payoffs as signals) and giving each agent a transfer equal to the sum of the other agents' reports. One advantage of this implementation is that constructing the transfers does not require knowing any more details about the environment than what is needed to calculate the decision plan. Furthermore, if the designer wants to calculate the decision plan but does not know the agents' utility functions, he can ask each agent to report his own utility function (which could be modeled as a “rich” signal available in period 1), and by Proposition 1 the agents can be relied upon to report truthfully. Recall, however, that in general the efficient decision rule depends on the agents’ beliefs as well as the their utility functions. If agents’ beliefs are subjective and different agents may hold different beliefs, there is no “detail-free” efficient decision rule, let alone a detail-free mechanism.

5. 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, provided that the types are independent (as defined in Subsection 3.1).

Our construction of the balanced mechanism relies on the following notation: For any period t and history $^{t-1}$, a given decision plan induces a probability measure over Θ, which assigns probability 1 to $\tilde{}^{t-1} = {}^{t-1}$, and which we denote by ${}_t[\ ]|\ {}^{t-1}$: This measure exists and is unique by the Tulcea Product Theorem (see footnote 7). Also, for each agent i, and each ${}_{i;t}$, let

$$ {}^{i}_{t}[\ ]|\ {}^{t}_{i};\ {}^{t-1}_{-i} = \mathbb{E}_{\ddagger_t}^{\,t|\ \ t-1(\ t-1);\ t-1}\left[\ {}_t[\ ]|\ {}^{t}_{i};\left({}^{t-1}_{-i};\ \dagger_{-i;t}\right)\right]: $$

This is the measure probability measure over Θ when agent i’s private signals are given through period t and all the other signals are given through period $t-1$.[14]

Using this notation, for any mechanism that implements decision plan using transfers , we construct a new mechanism with the same decision plan and the following “Balanced Transfers.”

$$ {}^{B}_{i;t}(\ {}^{t}) \quad = \quad {}_{i;t}\left(\ {}^{t}_{i};\ {}^{t-1}_{-i}\right) - \frac{1}{I-1}\sum_{j\neq i}\ {}_{j;t}(\ {}^{t}_{j};\ {}^{t-1}_{-j});\ \text{where} \tag{5.1} $$

$$ {}_{j;t}(\ {}^{t}_{j};\ {}^{t-1}_{-j}) \quad = \quad {}^{-t}\left(\mathbb{E}_{\tilde{}}^{\,j}{}_{t}[\ ]|\ {}^{t}_{j};\ {}^{t-1}_{-j}\left[\Psi_j\left(\tilde{}\ \right)\right] - \mathbb{E}_{\tilde{}}^{\,t}[\ ]|\ {}^{t-1}\left[\Psi_j\left(\tilde{}\ \right)\right]\right) \tag{5.2} $$

(where conditioning on $\tilde{}^{t}_{i} = {}^{t}_{i}$ with $t = 0$ is interpreted as vacuous, i.e., no conditioning at all).

By construction, the Balanced Transfers are budget-balanced. Also, note that if the transfers in the original mechanism are uniformly bounded, then by Lemma 2 (in the Appendix) Ψ_j is LCI with some constant C, which implies that, for all t, $\left|\ {}_{j;t}(\ {}^{t}_{j};\ {}^{t-1}_{-j})\right| \leq 2C$, and therefore the balanced transfers are uniformly bounded by $2C$:

To understand the Balanced Transfers, note ${}_{j;t}$, interpreted as agent j’s “incentive payment” in period t, gives the change in the expectation of agent j’s present discounted transfers ${}^{-t}\Psi_j\left(\tilde{}\ \right)$ in the original mechanism that results from his report ${}_{j;t}$ in period t, given the

[14]We could interpret measures $\mu_t[\chi]|\,\theta^{t-1}$ and $\mu^i_t[\chi]|\,\theta^t_i, \theta^{t-1}_{-i}$ as conditional measures given θ^{t-1} and $\left(\theta^t_i, \theta^{t-1}_{-i}\right)$, respectively. However, conditional probability measures are not well-defined at any given zero-probability point (this is the source of the famous Borel-Kolmogorov paradox – see, e.g., Proschan and Presnell (1998)), while these concepts are uniquely defined.

signal history $^{t-1}$.[15] As in the standard AGV mechanism, all the other agents pitch in the same amount $\frac{1}{I-1}\ _{j;t}$ to pay agent j's incentive payment. Subtracting all these payments that a given agent i has to make from his own incentive payment yields the balanced transfers (5.1).

Proposition 2. *Consider any mechanism (;) in which there is a truthful-obedient BNE. If the agents' types are independent, then this is also a truthful-obedient BNE in the balanced mechanism* $(\ ; \ ^{B})$.

Remark 1. *Recall that all that is important for truthful reporting to be a BNE is the present value of the transfers, which is given by*

$$\Psi_i^B(\ ) = \sum_{t=1}^{\infty} {}^{t} \left[{}_{i;t}({}_i^t; {}_{-i}^{t-1}) - \frac{1}{I-1} \sum_{j \neq i} {}_{j;t} \left({}_j^t; {}_{-j}^{t-1} \right) \right] :$$

There are many ways other than (5.1) to spread $\Psi_i^B(\ )$ over time while ensuring budget balance in each period, and all these ways will preserve incentive compatibility.

Remark 2. *If decisions are made only up to period T, then the balanced team mechanism depends only on ^{T}, and agents need not report $_t$ for $t > T$ (they need not be observed).*

We define the *Balanced Team Mechanism* to be a mechanism with an efficient decision plan, together with balanced team transfers ^{MB} constructed from ^{M} using (5.1).

Corollary 1. *With independent types and private values, there is a truthtelling BNE of the Balanced Team Mechanism.*

Before proceeding with a formal proof of Proposition 2, we provide a sketch of the main ideas. We can write each agent j"s incentive payment in period t as

$$_{j;t}({}_j^t; {}_{-j}^{t-1}) = {}^{-t} \left[{}_{j;t}^{+}({}_{j;t}; {}^{t-1}) - {}_{j;t}^{-}({}^{t-1}) \right] ; \text{ where}$$

$$_{j;t}^{+}({}_{j;t}; {}^{t-1}) = \mathbb{E}_{\sim}^{{}^{j}_{t}[\ ]|\ {}^{t}_{j};\ {}^{t-1}_{-j}} \left[\Psi_j \left({}^{\sim} \right) \right] ; \quad {}_{j;t}^{-}({}^{t-1}) = \mathbb{E}_{\sim}^{{}_{t}[\ ]|\ {}^{t-1}} \left[\Psi_j \left({}^{\sim} \right) \right] : \tag{5.3}$$

The two terms are expectations of the same function $\Psi_j\left({}^{\sim} \right)$, and differ only in the history used to construct beliefs, where ${}_{j;t}^{-}({}^{t-1})$ uses only period $t-1$ information while ${}_{j;t}^{+}({}_{j;t}; {}^{t-1})$ uses, in addition, agent j's period t information.

[15] In the case where all types in Θ have positive probability, so that the conditional expectations are uniquely defined everywhere, and also using (3.2), agent j's incentive payment can be written as

$$\gamma_{j,t}(\theta_j^t, \theta_{-j}^{t-1}) = \sum_{\tau=t}^{\infty} \delta^{\tau-t} \left(\mathbb{E}_{\tilde{\theta}}^{\mu[x]} \left[\psi_{j,\tau}\left(\tilde{\theta}\right) | \left(\tilde{\theta}_j^t, \tilde{\theta}_{-j}^{t-1}\right) = \left(\theta_j^t, \theta_{-j}^{t-1}\right) \right] - \mathbb{E}_{\tilde{\theta}}^{\mu[x]} \left[\psi_{j,\tau}\left(\tilde{\theta}\right) | \tilde{\theta}^{t-1} = \theta^{t-1} \right] \right).$$

Suppose that agent i deviates to a different strategy, while the other agents are truthful and obedient. The proof shows that, for any such deviation, the present expected value of agent i's incentive payments $_{i,t}$ in the balanced mechanism equals, up to a constant, to that in the original mechanism (Claim 1). Also, the proof shows that the present expected value of the other agents' incentive payments $_{j,t}$ (for $j \neq i$) is zero regardless of agent i's deviation (Claim 2). Thus, if the agent had no profitable deviation in the original mechanism, he will have none in the balanced mechanism.

The proofs of both Claim 1 and Claim 2 hinge on the assumption of independent types, which implies that for any agent $j \neq i$, the mechanism designer forms correct beliefs over the type $\tilde{}_{j,t}$ of agent j in period t based on the public history in that period, $\hat{}^{t-1}$, even when agent i has deviated from truthful-obedient behavior. Thus, for any public history $\hat{}^{t-1}$ and private history $^{t}_{i}$, agent i's expectation of $^{+}_{j,t}(\tilde{}_{j,t}, \hat{}^{t-1})$ equals $\bar{}_{j,t}(\hat{}^{t-1})$. By the Law of Iterated Expectations, this implies that the ex ante expectation of all the other agents' incentive payments equals zero for any possible deviation of agent i, which establishes Claim 2.

To show Claim 1, we examine the present expected value of agent i's own incentive payments following his deviation (which need not be zero, since the mechanism designer may have incorrect beliefs over agent i's future signals following his deviation). For this purpose, we compare the terms $^{+}_{i,t}(\hat{}_{i,t}, \hat{}^{t-1})$ and $\bar{}_{i,t+1}(\hat{}^{t})$, which differ only in that $\bar{}_{i,t+1}$ incorporates period t reports from i's opponents, $\hat{}_{-i,t}$. Recall again that by independence of types, that the mechanism designer has correct beliefs over the opponents' reports $\tilde{}_{-i,t}$ in period t based on the public history $\hat{}^{t-1}$, despite agent i's deviation. Then, the expectation of $\bar{}_{i,t+1}(\hat{}_{i,t}, \tilde{}_{-i,t}, \hat{}^{t-1})$ in period t is equal to $^{+}_{i,t}(\hat{}_{i,t}, \hat{}^{t-1})$, irrespective of agent i's private history. By the Law of Iterated Expectations, the same two terms have the same ex ante expectations as well, regardless of agent i's deviation. Using this, and for simplicity considering a finite horizon T, the present expected incentive payment of agent i is equal to the expectation of

$$\sum_{t=1}^{T}\left(\bar{}_{i,t+1}\left(\tilde{\hat{}}^{t}\right) - \bar{}_{i,t}\left(\tilde{\hat{}}^{t-1}\right)\right) = \bar{}_{i,T+1}\left(\tilde{\hat{}}^{T}\right) - \bar{}_{i,1} = \Psi_i\left(\tilde{\hat{}}^{T}\right) - \mathbb{E}_{\tilde{}}\left[\Psi_i\left(\tilde{}^{T}\right)\right],$$

where $\tilde{\hat{}} = \left(\bar{}_i\left(\tilde{}\right), \tilde{}_{-i}\right)$ represents the agents' random reports following the deviation (with $\tilde{}$ distributed according to $[\hat{}]$, with $\hat{}$ defined in (3.3)). Thus, for any possible deviation of agent i, the present expected value of the agent's incentive payments in the balanced mechanism equals, up to a constant, the present expected payment in the original mechanism, which establishes Claim 1. The formal proof makes these arguments more precise and extends them to the general infinite-horizon case.

Proof of Proposition 2: First note that by the construction of the induced stochastic process $^{t}[\;]|^{t-1}$ and by independence, the probability distribution of $\tilde{}_{i,t}$ given history $^{t-1}$ can be written as $_{i,t}|\,^{t-1}_{0}(^{t-1}), \,^{t-1}_{i}(^{t-1}), \,^{t-1}_{i}$. Similarly, by independence of types the probability distribution of $\tilde{}_{-i;t}$ given history ${}^{t-1}$ can be written as ${}_{-i;t}| \ {}^{t-1}_{-i}({}^{t-1}); {}^{t-1}_{-i} \equiv \prod_{j \neq i} {}_{j;t}| \left({}^{t-1}_{0}({}^{t-1}); {}^{t-1}_{j}({}^{t-1}); {}^{t-1}_{j} \right)$. To clarify observational measurability, we sometimes write ${}^{-t}_{i} ({}_{i;t}; {}^{t-1})$ rather than ${}^{-t}_{i} (\ ):$

Suppose that agent i deviates to induce an observationally measurable strategic plan $\left(\bar{}_i; \bar{}_i \right)$, while the other agents use truthful-obedient strategies. The decision and transfer plans following the deviation are given by $\hat{}$ and $\hat{\Psi}$ from (3.3) and (3.4), respectively.

Claim 1: Agent i's expected discounted sum of incentive payments ${}_{i;t}$ in the Balanced Mechanism equals his present expected transfer in the original mechanism from the same deviation plus a constant:

$$\mathbb{E}_{\tilde{}}^{[\hat{}]} \left[\sum_{t=1}^{\infty} {}^{t} \ {}_{i;t} \left({}^{-t}_{i} \left(\tilde{} \right); \tilde{}^{t-1}_{-i} \right) \right] = \mathbb{E}_{\tilde{}}^{[\hat{}]} \left[\Psi_i \left(\bar{}_i \left(\tilde{} \right); \tilde{}_{-i} \right) \right] - \mathbb{E}_{\tilde{}}^{[\ ]} \left[\Psi_i \left(\tilde{} \right) \right]:$$

Proof of Claim 1: First, applying LIE, it follows that

$$\begin{aligned} {}^{+}_{i;t} \left({}^{-t}_{i} \left({}_{i;t}; {}^{t-1} \right); {}^{t-1}_{-i} \right) &= \mathbb{E}_{+_{-i,t}}^{-i,t| \ {}^{t-1}_{-i} \left({}^{-t-1}_{i} \left({}^{t-1}_{i}; {}^{t-2} \right) \right); \ {}^{t-1}_{-i}} \left[\bar{}_{i;t+1} \left({}^{-t}_{i} \left({}^{t-1}; {}_{i;t} \right); \left({}^{t-1}_{-i}; +_{-i;t} \right) \right) \right] \\ &= \mathbb{E}_{+_{-i,t}}^{-i,t| \hat{}^{t-1}_{-i} \left({}^{t-1} \right); \ {}^{t-1}_{-i}} \left[\bar{}_{i;t+1} \left({}^{-t}_{i} \left({}^{t-1}; {}_{i;t} \right); \left({}^{t-1}_{-i}; +_{-i;t} \right) \right) \right]; \qquad (5.4) \end{aligned}$$

where the second inequality follows by the definition of $\hat{}$:

Now, taking expectations over $\tilde{}_{i;t}$ with respect to probability distribution induced by ${}_{t}[\hat{}]| \ {}^{t-1}$, we can write

$$\begin{aligned} &\mathbb{E}_{\tilde{}_{i,t}}^{t[\hat{}]| \ {}^{t-1}} \left[{}^{+}_{i;t} \left({}^{-t}_{i} \left(\tilde{}_{i;t}; {}^{t-1} \right); {}^{t-1}_{-i} \right) \right] \\ &= \mathbb{E}_{\tilde{}_{i,t}}^{t[\hat{}]| \ {}^{t-1}} \mathbb{E}_{+_{-i,t}}^{-i,t| \hat{}^{t-1}_{-i} \left({}^{t-1} \right); \ {}^{t-1}_{-i}} \left[\bar{}_{i;t+1} \left({}^{-t}_{i} \left(\tilde{}_{i;t}; {}^{t-1} \right); \left({}^{t-1}_{-i}; +_{-i;t} \right) \right) \right] \\ &= \mathbb{E}_{\tilde{}_{t}}^{t[\hat{}]| \ {}^{t-1}} \left[\bar{}_{i;t+1} \left({}^{-t}_{i} \left(\tilde{}_{i;t}; {}^{t-1} \right); \left({}^{t-1}_{-i}; \tilde{}_{-i;t} \right) \right) \right]: \end{aligned}$$

Thus,

$$\begin{aligned} &\mathbb{E}_{\tilde{}_{i,t}}^{t[\hat{}]| \ {}^{t-1}} \left[{}^{t} \ {}_{i;t} \left({}^{-t}_{i} \left(\tilde{}_{i;t}; {}^{t-1} \right); {}^{t-1}_{-i} \right) \right] \\ = \ &\mathbb{E}_{\tilde{}_{t}}^{t[\hat{}]| \ {}^{t-1}} \left[\bar{}_{i;t+1} \left({}^{-t}_{i} \left(\tilde{}_{i;t}; {}^{t-1} \right); \left({}^{t-1}_{-i}; \tilde{}_{-i;t} \right) \right) \right] - \bar{}_{i;t} \left({}^{-t-1}_{i} \left({}_{i;t-1}; {}^{t-2} \right); {}^{t-1}_{-i} \right): \end{aligned}$$

Taking expectations over $\tilde{}^{t-1}$ with respect to the probability distribution $[\hat{}]$ and using LIE yields

$$\mathbb{E}_{\tilde{}}^{[\hat{}]} \left[{}^{t} \ {}_{i;t} \left({}^{-t}_{i} \left(\tilde{}_{i;t}; \tilde{}^{t-1} \right); \tilde{}^{t-1}_{-i} \right) \right] = \Gamma_{t+1} - \Gamma_t,$$

$$\text{where } \Gamma_t = \mathbb{E}_{\tilde{}}^{[\hat{}]} \left[\bar{}_{i;t} \left({}^{-t-1}_{i} \left(\tilde{}_{i;t-1}; \tilde{}^{t-2} \right); \tilde{}^{t-1}_{-i} \right) \right]:$$

Thus, we can write

$$\begin{aligned} \mathbb{E}_{\tilde{}}^{[\hat{}]} \left[\sum_{t=1}^{\infty} {}^{t} \ {}_{i;t} \left({}^{-t}_{i} \left(\tilde{}_{i;t}; \tilde{}^{t-1}_{-i} \right); \tilde{}^{t-1}_{-i} \right) \right] &= \lim_{T \to \infty} \sum_{t=1}^{T} \left(\Gamma_{t+1} - \Gamma_t \right) \\ = \lim_{T \to \infty} \Gamma_{T+1} - \Gamma_1 &= \mathbb{E}_{\tilde{}}^{[\hat{}]} \left[\Psi_i \left(\bar{}_i \left(\tilde{} \right); \tilde{}_{-i} \right) \right] - \mathbb{E}_{\tilde{}}^{[\ ]} \left[\Psi_i \left(\tilde{} \right) \right]; \end{aligned}$$

where the first equality obtains from the previous derivation using the Lebesgue Dominated Convergence Theorem and the LCI of Ψ_i, the second by canceling intermediate terms in the partial sum, and the third by the Dominated Convergence Theorem, the LCI of Ψ_i, and the fact that distribution $_{T+1}[\ ]|\left(^{-T}_{i}\left(_{iT};\ ^{T-1}_{-i}\right);\ ^{T}_{-i}\right)$ puts probability 1 on the event $\tilde{\tilde{}}^{T} = \left(^{-T}_{i}\left(_{iT};\ ^{T-1}_{-i}\right);\ ^{T}_{-i}\right)$.

Now consider a second claim:

Claim 2: The present expected value of agent j's incentive payments ($_{j;t}$*) is equal to zero when agent* j *follows a truthtelling strategy no matter what reporting strategies the other agents use:*

$$\mathbb{E}_{\sim}^{[\hat{}]}\left[\sum_{t=1}^{\infty}\ ^{t}\ _{j;t}\left(\tilde{}^{t}_{j};\ ^{-t-1}_{-j}\left(\tilde{}^{t-1}\right)\right)\right] = 0:$$

Proof of Claim 2: For any $j \neq i$; let $^{-}_{-j}(\ ) = \left(^{-}_{i}(\ );\ _{-i-j}\right)$ denote the reporting strategy of agents other than j. We write $^{-t}_{-j}(\ ^{t})$ for the reporting strategy up to period t. Using (5.2), we can write for any t, and any $^{t-1}$,

$$\begin{aligned}
&\mathbb{E}_{\sim_{j,t}}^{t[\hat{}]|\ ^{t-1}}\left[\ ^{+}_{j;t}\left(\left(^{t-1}_{j};\ \tilde{}_{j;t}\right);\ ^{-t-1}_{-j}\left(\ ^{t-1}\right)\right)\right] \qquad (5.5)\\
=\ &\mathbb{E}_{\sim_{j,t}}^{t[\hat{}]|\ ^{t-1}}\mathbb{E}_{\dagger_{-j,t}}^{-j,t|\ ^{t-1}_{-j}(\ ^{t-1}_{j};^{-t-1}_{-j}(\ ^{t-1}));^{-t-1}_{-j}(\ ^{t-1})}\mathbb{E}_{\tilde{\sim}}^{t+1[\ ]|\left(^{t-1}_{j};^{-t-1}_{-j}(\ ^{t-1});\tilde{}_{j,t};\dagger_{-j,t}\right)}\left[\Psi_j\left(\tilde{}\right)\right]\\
=\ &\mathbb{E}_{\dagger_{j,t}}^{j,t|\ ^{t-1}_{0}\left(^{t-1}_{j};^{-t-1}_{-j}(\ ^{t-1})\right);\ ^{t-1}_{j}\left(^{t-1}_{j};^{-t-1}_{-j}(\ ^{t-1})\right);\ ^{t-1}_{j}}\left[\begin{array}{c}\mathbb{E}_{\dagger_{-j,t}}^{-j,t|\ ^{t-1}_{-j}(\ ^{t-1}_{j};^{-t-1}_{-j}(\ ^{t-1}));^{-t-1}_{-j}(\ ^{t-1})}\\ \left[\mathbb{E}_{\tilde{\sim}}^{t+1[\ ]|\left(^{t-1}_{j};^{-t-1}_{-j}(\ ^{t-1});\dagger_{t}\right)}\left[\Psi_j\left(\tilde{}\right)\right]\right]\end{array}\right]\\
=\ &\mathbb{E}_{\sim}^{t[\ ]|(\ ^{t-1}_{j};^{-t-1}_{-j}(\ ^{t-1}))}\mathbb{E}_{\tilde{\sim}}^{t+1[\ ]|\left(^{t-1}_{j};^{-t-1}_{-j}(\ ^{t-1});\tilde{}_{t}\right)}\left[\Psi_j\left(\tilde{}\right)\right]\\
=\ &\mathbb{E}_{\sim}^{t[\ ]|\left(^{t-1}_{j};^{-t-1}_{-j}(\ ^{t-1})\right)}\left[\Psi_j\left(\tilde{}\right)\right] = \ ^{-}_{j;t}\left(^{t-1}_{j};\ ^{-t-1}_{-j}\left(\ ^{t-1}\right)\right):
\end{aligned}$$

The first equality involves substitutions and definitions, and the LIE. The second is by definition of $_{t}[\hat{}]|\ ^{t-1}$ and $\hat{}$ and independence of types (which implies that the distribution of $\tilde{}_{j;t}$ in $_{t}[\hat{}]|\ ^{t-1}$ depends only on the public decision history and on the private history of agent j). The third follows by definition of $_{t}[\ ]|\ ^{t-1}$, LIE and independent types. The fourth uses LIE and definitions, and the fifth uses definitions. By LIE and the fact that $_{j;t} = \ ^{-t}\left(^{+}_{j;t} - \ ^{-}_{j;t}\right)$, this implies that

$$\mathbb{E}_{\sim}^{[\hat{}]}\left[\ _{j;t}\left(\tilde{}^{t}_{j};\ ^{-t-1}_{-j}\left(\tilde{}^{t-1}\right)\right)\right] = \mathbb{E}_{\tilde{\sim}}^{[\hat{}]}\left[\mathbb{E}_{\sim}^{t[\hat{}]|\tilde{\tilde{}}^{t-1}}\left[\ _{j;t}\left(\tilde{}^{t}_{j};\ ^{-t-1}_{-j}\left(\tilde{}^{t-1}\right)\right)\right]\right] = 0:$$

Then, the Lebesgue Dominated Convergence Theorem and LCI of Ψ_j imply that

$$\mathbb{E}_{\sim}^{[\hat{}]}\left[\sum_{t=1}^{\infty}\ ^{t}\ _{j;t}\left(\tilde{}^{t}_{j};\ ^{-t-1}_{-j}\left(\tilde{}^{t-1}\right)\right)\right] = \sum_{t=0}^{\infty}\mathbb{E}_{\sim}^{[\hat{}]}\left[\ ^{t}\ _{j;t}\left(\tilde{}^{t}_{j};\ ^{-t-1}_{-j}\left(\tilde{}^{t-1}\right)\right)\right] = 0;$$

completing the proof of Claim 2.

So, adding up, we have

$$\mathbb{E}_{\tilde{}}^{[\hat{}]}\left[\Psi_i^B\left(\bar{}_i\left(\tilde{}\right);\tilde{}_{-i}\right)\right] = \mathbb{E}_{\tilde{}}^{[\hat{}]}\left[\Psi_i\left(\bar{}_i\left(\tilde{}\right);\tilde{}_{-i}\right)\right] - \mathbb{E}_{\tilde{}}^{[\ ]}\left[\Psi_i\left(\tilde{}\right)\right]: \qquad (5.6)$$

Since the second term does not depend on agent i's strategy ($_i$; $_i$), the expected transfer to agent i for any strategy $_i$ chosen by him is equal, up to a constant, to the expected transfer under the original mechanism, assuming that other agents use truthful-obedient strategies. Thus, if the original mechanism has a truthtelling BNE, it remains a BNE in the balanced mechanism. ■

6. Efficient Equlibria in Markov Games with Private States

In this section, we apply the results to show a Folk-Theorem like result for Markov games with private states — namely, that efficiency can be sustained in equilibrium when the discount factor is close enough to 1. In the games, there is no external enforcer, and so the decisions and payments must be self-enforcing. For simplicity we restrict attention to games with transfers so that our results can be applied directly.[16] The requirement of decisions and transfers to be self-enforcing means that we need to consider deviations in which they not only misreport, but also choose off-equilibrium decisions and transfers. Such deviations can trigger a permanent "punishment phase" of the game. Making sure that such deviations are unprofitable is similar to ensuring participation constraints in the mechanism where agents can choose to exit in each period. In a static context or in a model with very impatient agents, it is well known that whether or not participation constraints can be satisfied depends on the structure of the game. Here, we focus instead on the case where agents are patient and the horizon is infinite, and identify conditions under which participation constraints could be satisfied in this model. In a follow-on paper (Athey and Segal (2007)), we explore a special case of our model in greater depth, providing weaker sufficient conditions for participation constraints to hold.

To see the problem with guaranteeing participation constraints even with patient agents, note that the balanced transfers in each period are a function of discounted sums of future payoffs, and so they could potentially grow with the discount factor at a rate proportional to future surplus. Thus, in general the future surplus may not be enough to guarantee that all agents will wish to make the necessary transfers. In this section, we provide conditions under which the balanced transfers remain bounded as patience increases, and so the growing surplus from future cooperation can be used to make these transfers self-enforcing.

Formally, we consider an infinite-horizon game where each period t consists of three stages:

Stage t:0 : Each agent i simultaneously observes a private signal $_{i;t} \in \Theta_{i;t}$,

Stage t:1 : Each agent i sends a report $\hat{}_{i;t} \in \Theta_{i;t}$.

[16]In games without transfers, one could in principle extend the approach of Fudenberg, Levine, and Maskin (1994), using small changes in continuation play mimic the role of transfers.

Stage $t{:}2$: Each agent i chooses a publicly observable action $x_{0;i;t} \in X_{0;i;t}$; a privately observed action $x_{i;t} \in A_{i;t}$, and a publicly observable payment $z_{i;j;t} \geq 0$ to each agent j. (Note that the payments are not wasted by construction).

Letting $x_{0;t} = (x_{0;i;t}; :::; x_{0;I;t})$, the collective decision at time t is $x_t \equiv (x_{0;t}; x_{1;t}; :::; x_{I;t}) \in \prod_{i=0}^{I} X_{i;t} \equiv X_t$, and the total payment received by agent i in period t is $y_{i;t} = \sum_{j=1}^{I} (z_{j;i;t} - z_{i;j;t})$. The payoff of agent i in this game takes the same form as before,

$$\sum_{t=1}^{\infty} {}^{t} \left(u_{i;t}(x^t; \ {}^t) + y_{i;t}\right) :$$

We want to construct an efficient equilibrium for this game, i.e., one that implements an observationally measurable decision plan maximizing the expected joint surplus (4.1). We construct an equilibrium in which agents use truthful-obedient reporting strategies. Let denote the transfer plan sustained in the equilibrium.

One class of deviations is that an agent uses a different reporting strategy and/or private decision policy, but uses the same public decision and payment strategies. These deviations are unprofitable as long as the reporting strategies and private decision strategies form a BNE of the mechanism with enforcement. However, with self-enforcement, we also need to deter "off-schedule deviations," which involve, in addition to possible misreporting and deviation in private decisions, choosing public decisions or payments different from those required by the equilibrium strategies given past reports. Off-schedule deviations are publicly observable as deviations, and as soon as they are observed they can be punished by the agents switching to a "punishment" decision plan.

To deter off-schedule deviations, we need to make sure that the benefits of changing a decision or withholding a payment in one period are outweighed by the future gain from cooperation. As mentioned above, the balanced transfers could potentially grow large as the discount factor increases, since they depend on the expected discounted value of future utility, and so in general the payments are not guaranteed to be smaller than the future gains from cooperation. However, the balanced transfers are defined as functions of the *change* in future expected discounted utility for a particular agent induced by the arrival of a single agent's signal in a particular period. We now specialize the model in a way that guarantees that a particular period's signal and decisions have vanishing effect on the continuation expected utility. This will guarantee that the transfers stay within a fixed bound even as the discount factor grows, while the present value of cooperation grows without a bound, hence off-schedule deviations are deterred.

Assumptions

(i) $\Theta_t = \overline{\Theta}$ and $X_t = \overline{X}$ for all t, and all the sets are finite.

(ii) Per-period utilities $u_{i,t}(x^t, \theta^t)$ depend only on period-t variables and do not depend on t directly, and so can be written as $u_i(x_t, \theta_t)$.

(iii) The contingent probability measures $\mu_t|x^{t-1}, \theta^{t-1}$ depend only on period $t-1$ variables and do not depend on t directly, and so can be written as $\mu|x_{t-1}, \theta_{t-1}$ (where for $t = 1$ we take some fixed x_0, θ_0).

(iv) Types are independent (Definition 1) and values are private (Definition 3).

(v) There exists $\hat{x}_t \in X_t$ such that $u_i(x_{i,t}, \hat{x}_{-i,t}, \theta_t) = 0$ for all i, all $x_{i,t} \in \bar{X}$, and all $\theta_t \in \bar{\Theta}$.

Assumptions (i)-(iii) imply that we have a Markov Decision Process (MDP), as studied, e.g., in Puterman (1994).[17] In such problems, it is natural to focus on Markov policies. Furthermore, it turns out that there exists a single Markov policy that is optimal for δ close enough to 1:

Definition 5. *Decision policy χ is a Markov policy if each $\chi_t(\theta^t)$ depends only on the last state θ_t and does not depend directly on t, and so can be written as $\chi(\theta_t)$. Decision policy χ^* is a Blackwell policy if it is a Markov policy and there exists $\bar{\delta} < 1$ such that for any $\delta \in (\bar{\delta}, 1)$, χ^* achieves a weakly greater present expected total surplus starting at any state than any other observationally measurable decision policy χ.*

Puterman (1994) shows that a Blackwell policy exists (Theorem 10.1.4) and describes an algorithm that constructs it with a finite number of steps (Corollary 10.3.7, p.518).

Assumption (v) can be understood by interpreting decision $\hat{x}_{i,t} \in X_{i,t}$ as "non-participation" by agent i and saying that an agent's utility is fixed when no other agent participates. For simplicity, the fixed utility is normalized to zero. This ensures that all agents choosing non-participation and making zero payments is a BNE of the stage game regardless of agents' types and beliefs about others' types. This allows us to use such BNE as a punishment for the observed off-schedule deviations.[18]

Observe that in the Markov game, the team transfers can be written as $\psi_{i,t}(\theta_t) = \sum_{j \neq i} u_j(\chi^*(\theta_t), \theta_t)$, and the balanced team transfers can be written as functions of the current and most recent

[17] Note that dependence on $k > 1$ periods of history could be incorporated by expanding the sets Θ_i and X to remember k periods of past types and decisions.

[18] More generally, in particular games without a natural nonparticipation option, there may be punishment equilibria that can be used in place of nonparticipation; however, if the equilibria depend on the initial state of the game, one must guarantee the existence of an equilibrium for each possible state of the game. Equilibria where decisions do not depend on the state (such as the pooling equilibria constructed by Athey and Bagwell (2004) in their study of dynamic collusion with persistent private information) are especially convenient to analyze in this context.

states as follows:

$$\begin{aligned} {}_i^B(\ {}_{t-1};\ {}_t) &= {}_i^B(\ {}_{i;t};\ {}_{t-1}) - \frac{1}{I-1}\sum_{j\neq i} {}_j^B(\ {}_{j;t};\ {}_{t-1});\ \text{where for each } j; \\ {}_j^B(\ {}_{j;t};\ {}_{t-1}) &= \sum_{k\neq j}\sum_{=t}^{\infty}\ {}^{-t}\left\{ \begin{array}{l} \mathbb{E}_{\tilde{}}^{\ {}_t^j[\ ^*]|\ {}_{j,t};\ {}_{t-1}}\left[u_k\left(\ ^*\left(\tilde{}\ \right);\ \ \right)\right] \\ -\mathbb{E}_{\tilde{}}^{\ {}_t[\ ^*]|\ {}_{t-1}}\left[u_k\left(\ ^*\left(\tilde{}\ \right);\tilde{}\ \right)\right] \end{array} \right\}: \end{aligned} \tag{6.1}$$

Thus, in the Markov environment, period t transfers have the attractive feature that they depend only on the announcements in periods t and $t-1$; since these announcements fully determine beliefs.

Any Markov decision policy induces a Markov chain on states with transition probabilities described by $\ |\ ^*(\ {}_{t-1});\ {}_{t-1}$. We assume that the process induced by the Blackwell policy has limited history dependence, in the sense of having a single ergodic set:

Definition 6. *Given a Markov process, a set of states $S \subseteq \bar{\Theta}$ is called ergodic if, starting in S, the system remains there with probability one, and this is not true for any proper subset of S.*

Proposition 3. *Consider a dynamic game satisfying assumptions (i)-(v) above. Let $\ ^*$ be a Blackwell decision policy for the game, and suppose that the Markov process induced by $\ ^*$ has a single ergodic set (plus a possibly empty transient set), and that its invariant distribution yields a positive expected total surplus. Then there exists $\ ^* < 1$ such that, for all $\ \in (\ ^*; 1)$; the game has an efficient PBE that sustains decision plan $\ ^*$.*

Proof. We focus on those close enough to 1 for which $\ ^*$ is a Blackwell policy. Define "on-schedule" public histories as those in which all agents in all period have made public decisions and payments consistent with the equilibrium strategies given the public histories, and call any other public histories "off-schedule" histories. We construct an equilibrium in which strategies satisfy the following for each period t and each agent i :

- For an off-schedule history, in Stage t:1, the agent reports his true type. In Stage t:2, he chooses the nonparticipation action $\hat{x}_{i;t}$, and makes zero payments to all other agents.
- For an on-schedule history, in Stage t:1, the agent reports his true signal $\ {}_{i;t}$. In Stage t:2, given the current reports $\hat{}\ {}_t$, he chooses his public action $x_{0;i;t} = \ {}_{0;i}^*\left(\hat{}\ {}_t\right)$; he chooses his private action $x_{i;t} = \ {}_i^*\left(\hat{}\ {}_t\right)$ if in Stage t:1 he reported truthfully $\hat{}\ {}_{i;t} = \ {}_{i;t}$, otherwise he chooses a private action $x_{i;t}$ that maximizes the present expected value of his utility given the belief that in the future all agents are truthful and obedient. Also, the agent makes payments determined from the reports in periods $t-1$ and t as follows:

$$z_{i;j;t}\left(\hat{}\ {}_{i;t};\hat{}\ {}_{t-1}\right) = \frac{1}{I-1}\left[\ {}_j^B\left(\hat{}\ {}_{j;t};\hat{}\ {}_{t-1}\right)\right] + A_i \text{ to all agents } j \neq i;$$

where the constants $A_i \geq \frac{1}{I-1} \max_{i;(\ _{t-1};\ _{i,t})} \left| \ _i^B (\ _{i;t};\ _{t-1}) \right|$ are specified below.

Observe that if all agents follow these strategies, we will only have on-schedule histories, and the agents will always be truthful and obedient, which implements the Blackwell decision plan *. Now we show that these strategies are part of a PBE, for some appropriately chosen constants A_i, when is close enough to 1. The accompanying PBE beliefs of each agent i in each period t about the other agents' current types if there has been no off-schedule deviation are described by $\Pi_{j \neq i}\ _j| \ ^*\left(\hat{}_{t-1}\right); \hat{}_{j;t-1}$ (i.e., the agent believes that the other agents were truthful and obedient in the previous period). The beliefs following an off-schedule deviation are described by $\Pi_{j \neq i}\ _j|x_{0;j;t-1}; \hat{x}_{j;t-1}; \hat{}_{j;t-1}$. It is easy to see that these beliefs satisfy PBE conditions (i)-(iii) of Subsection 3.3. It remains to show that the strategies described above form a best response to each other given these beliefs at any information set.

First, it is easy to see that the strategies form a best response to each other for any off-schedule public history, since at this point any possible strategy of agent i will give him a nonpositive expected continuation payoff regardless of his beliefs. Now consider the information sets with on-schedule public histories. First we argue that at such information sets, there is no profitable "on-schedule" deviation for each agent i, i.e., one in which he makes the public decisions and payments prescribed by the equilibrium strategies given the reports. For information sets in the beginning of a period t, this holds because the Blackwell policy by definition must be efficient starting in period t given any initial beliefs, and so together with the total prescribed payments

$$\sum_{j \neq i} \left[z_{j;i;t}\left(\hat{}_{j;t}; \hat{}_{t-1}\right) - z_{i;j;t}\left(\hat{}_{i;t}; \hat{}_{t-1}\right)\right] = \ _i^B\left(\hat{}_{t-1}; \hat{}_{t}\right) - (I-1)A_i + \sum_{j \neq i} A_j$$

the agent faces a Balanced Team Mechanism, in which truthfulness and obedience are optimal for him within the class of on-schedule strategies according to Proposition 2. In the middle of a period t after reports are made but before decisions are chosen, agent i has no profitable on-schedule deviation by the Principle of Dynamic Programming, since we already know that truthfulness and obedience will be optimal for him starting in period $t+1$ no matter what private action $x_{i;t}$ he chooses now, and his choice of $x_{i;t}$ is optimal given truthfulness and obedience in the future.

Now we need to show that, for some appropriately chosen constants A_i, each agent has no profitable "off-schedule" deviations, i.e., those in which he plans a public decision and/or payment that differ from the equilibrium ones prescribed given the reporting histories, at any on-schedule information set. We do this by showing that if at any future information set the agent plans a first off-schedule public action and/or payment, we can modify his strategy to be truthful and obedient starting from this information set without reducing his expected continuation payoff. By doing it for any information set at which a first off-schedule deviation is planned, we would obtain an on-schedule deviation that is at least as profitable as the original off-schedule deviation. Since we already know that there are no profitable on-schedule deviations, this will complete the proof.

We choose the constants A_i so that the expected surplus in the game is shared equally among the agents, and since the surplus is positive and very large from any starting state when the agents are very patient, this ensures that the agents will want not to "quit" the game by resorting to an off-schedule deviation. For this purpose, recall that since by assumption the Markov process induced by * has a single ergodic set, it induces a unique invariant distribution over $\tilde{}_t$ (Stokey and Lucas, 1989), which we denote by . Let $\bar{u}_i$ denote the expected per-period utility of agent i in this distribution:

$$\bar{u}_i \equiv \mathbb{E}_{\tilde{}_t}\left[u_i\left(^*\left(\tilde{}_t\right); \tilde{}_t\right)\right]:$$

Since the expectation of ^B_j is zero for all $j \neq i$, we let $A_i = \bar{u}_i = I + K$, where the constant K should be large enough to ensure that $z_{i;j;t}(\ ) \geq 0$ for all . We show that such K can be found independently of , for which purpose we demonstrate that the transfers $^B_j(\ _{j;t};\ _{t-1})$ are bounded uniformly in . We do it by making use of the following lemma, whose proof (in the Appendix) relies on the structure of Markov chains with a single ergodic subset - see, e.g., Behrends (2000, Chapter 7) and Stokey and Lucas (1989):

Lemma 1. *Take a finite Markov chain with a single ergodic set, let* $p^{(t)}_{;\ '} = \Pr\{\tilde{}_t = \ ' | \tilde{}_0 = \ \}$ *(the t-step transition probabilities), and let denote the invariant distribution. The expression* $\left|\sum_{t=0}^{\infty} {}^t \left(p^{(t)}_{;\ '} - \ _{'}\right)\right|$ *is bounded uniformly across all* $\in (0;1)$ *and all* $;\ ' \in \bar{\Theta}$.

The Lemma implies that there exists $C < \infty$ such that for all $_0 \in \bar{\Theta}$, $\in (0;1)$;

$$\left|\mathbb{E}_{\tilde{}}^{[\ ^*]|\ _0}\left[\sum_{t=0}^{\infty} {}^t\left\{u_j\left(^*\left(\tilde{}_t\right); \tilde{}_t\right) - \bar{u}_j\right\}\right]\right| \leq C: \tag{6.2}$$

In particular, (6.2) implies that the transfers are bounded:

$$\begin{aligned} \left| ^B_j(\ _{j;t};\ _{t-1})\right| &\leq \left|\mathbb{E}_{\tilde{}}^{^j_t[\ ^*]|\ _{j,t};\ _{t-1}}\left[\sum_{=t}^{\infty} {}^{-t}\left\{u_j\left(^*\left(\tilde{}\right); \tilde{}\right) - \bar{u}_j\right\}\right]\right| \\ &\quad + \left|\mathbb{E}_{\tilde{}}^{_t[\ ^*]|\ _{t-1}}\left[\sum_{=t}^{\infty} {}^{-t}\left\{u_j\left(^*\left(\tilde{}\right); \tilde{}\right) - \bar{u}_j\right\}\right]\right| \\ &\leq 2\sup_{_t}\left|\mathbb{E}_{\tilde{}}^{_t[\ ^*]|\ _t}\left[\sum_{=t}^{\infty} {}^{-t}\left\{u_j\left(^*\left(\tilde{}\right); \tilde{}\right) - \bar{u}_j\right\}\right]\right| \leq 2C; \end{aligned}$$

where the second inequality uses the Law of Iterated Expectations and the fact that $|\mathbb{E}g(X)| \leq \sup|g(X)|$. Therefore, letting $K = 2C = (I-1) + M = I$, where $M = \max_{i;x_t;\ _{i,t}} |u_i(x_t;\ _t)|$, ensures that $0 \leq z_{i;j;t}(\ _{i;t};\ _{t-1}) \leq 2K$ for any $\in (0;1)$ and any $(\ _{j;t};\ _{t-1})$.

Now, given these constants, compare the expected continuation payoff in period t from making a first off-schedule deviation in period t with that from reverting to truthfulness and obedience in period t. An off-schedule deviation will bring nonpositive payoffs starting in period $t+1$, while the expected continuation payoff from truthfulness and obedience in period $t+1$ conditional on any starting state $\ _t$ can be bounded below as follows:

$$
\begin{aligned}
& \sum_{=t+1}^{\infty} \ \mathbb{E}_{\sim}^{t+1[\ ^*]|\ t} \left[u_i\left(\ ^*\left(\tilde{\ }\right); \tilde{\ } \right) + \ _i^B\left(\tilde{\ }_{-1}; \tilde{\ } \right) + \frac{1}{I} \sum_j \bar{u}_j - \bar{u}_i \right] \\
= & \sum_{=t+1}^{\infty} \ \mathbb{E}_{\sim}^{t+1[\ ^*]|\ t} \left[u_i\left(\ ^*\left(\tilde{\ }\right); \tilde{\ } \right) + \frac{1}{I} \sum_j \bar{u}_j - \bar{u}_i \right] \\
\geq & \frac{\ ^{t+1}}{1-\ } \frac{1}{I} \sum_j \bar{u}_j - \left| \sum_{=t+1}^{\infty} \ \left\{ \mathbb{E}_{\sim}^{[\ ^*]|\ t} \left[u_i\left(\ ^*\left(\tilde{\ }\right); \tilde{\ } \right) \right] - \bar{u}_i \right\} \right| \\
\geq & \ ^{t+1} \left[\frac{1}{I(1-\ )} \sum_i \bar{u}_i - C \right]: \qquad (6.3)
\end{aligned}
$$

where the equality obtains since

$$
\mathbb{E}_{\sim}^{[\ ^*]|\ t} \left[\ _i^B\left(\tilde{\ }_{-1}; \tilde{\ } \right) \right] = 0
$$

for $\ \geq t+1$ by the Law of Iterated Expectations, and the last inequality is by (6.2). This should be weighed against the possible gain in period t from the off-schedule deviation rather than being truthful and obedient in this period, which is at most

$$
\ ^t \left(I \sup_{i;j;t;\ _{t-1};\ _{i,t}} z_{i;j;t}(\ _{i;t};\ _{t-1}) + 2 \max_{i;\ _t;x_t} |u_i(x_t;\ _t)| \right) = \ ^t (I \cdot 2K + 2M):
$$

Thus, we see that for $\ =(1-\ ) > I(C + 2IK + 2M) = \left(\sum_i \bar{u}_i \right)$, the off-schedule deviation is dominated by becoming truthful and obedient. Together with the previous argument, this shows that the proposed strategies and beliefs form a PBE for $\ $ close enough to 1. ■

We conclude this section with a couple of observations about how the results can be generalized. First, we restricted attention to efficient allocation rules. We could also consider other decision rules that could be implemented using transfers (not necessarily budget-balanced) that are uniformly bounded in $\ $ and that have a Markov structure. Then, our proof can be modified to establish that the balanced transfers satisfy satisfy participation constraints for sufficiently patient agents.

Second, the Markov structure is more restrictive than necessary. The important feature of the Markov structure is that today's signals and reports have a vanishing impact on the expected value of future payoffs and transfers. This feature could be preserved in models that are nonstationary and have dependence on longer and possibly infinite histories, provided that this dependence is bounded.

7. Conclusions

This paper generalizes to a dynamic setting the results of Arrow (1979) and d'Aspremont and Gerard-Varet (1979) that, given an allocation rule and incentive-compatible transfers, it is possible to construct budget-balanced and incentive-compatible transfers. The result holds for a very general class of dynamic games, with arbitrary intertemporal dependence of agents' signals, arbitrary effects of the public decision on the distribution of signals and their informativeness, and privately observed decisions (moral hazard problems) so long as the decisions do not have direct externalities on other agents. The main substantive restrictions are additive separability of payoffs in money, and independence of types across agents (conditional on past public decisions). Although the analysis of participation constraints depends on the structure of the specific game, for agents in an infinite-horizon game with a Markov structure and a single ergodic set, participation constraints can be satisfied when agents are sufficiently patient.

In future work we plan to consider models of costly computation and communication. Some problems of computation costs can be considered as special cases of our model, by modeling an agent's computation of her type as information gathering.[19] We plan to extend our model to include the case where agents send reports from a restricted message space and so cannot fully reveal their information (e.g., because communication is costly), as in Fadel and Segal (2006).

We also intend to study the case of correlated values more closely. In static mechanism design, the case of independent types is viewed as the "hardest" case for mechanism design; with correlated types, efficiency with budget balance is achievable under generic conditions, as shown by d'Aspremont, Cremer, and Gerard-Varet (2003a,b). We expect that these results could be extended to dynamic mechanism design, but this would require different techniques than the ones we have used, and so we leave this extension for future research.[20]

References

[1] Akerlof, G. (1970), "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *The Quarterly Journal of Economics* 84(3), 488-500.

[2] Athey, S., and K. Bagwell (2004), "Collusion with Persistent Cost Shocks," mimeo, Stanford University, 2004.

[3] Athey, S., and D. Miller (2004), "Efficiency in repeated trade with hidden valuations," mimeo, Stanford University.

[19]See, e.g., Larson and Sandholm (2001) for such a model. Other kinds computational costs fall outside of our model, such as the agents' costly information acquisition about others' types, or the cost of computing the decisions and transfers by the mechanism.

[20]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 is examined by Obara and Rahman (2006).

[4] Athey, S. and I. Segal (2007), "Designing Efficient Mechanisms for Dynamic Bilateral Trading Games," *American Economic Review Papers and Proceedings*, forthcoming.

[5] Atkeson, A. and R. Lucas (1993), "On Efficient Distribution with Private Information," *Review of Economic Studies*, 59, 427-453.

[6] Arrow, K. (1979), "The property rights doctrine and demand revelation under incomplete information." In M Boskin, ed., *Economics and human welfare*. New York Academic Press.

[7] d'Aspremont, C., J. Cremer, and Gerard-Varet L. (2003a) "Balanced Bayesian Mechanisms." *Journal of Economic Theory*, 115, 385-396.

[8] d'Aspremont, C., J. Cremer, and Gerard-Varet L. (2003b), "Correlation, Independence, and Bayesian Incentives," *Social Choice and Welfare* 21(2), 281-310.

[9] d'Aspremont, C. and Gerard-Varet, L. (1979) "Incentives and incomplete information." *Journal of Public Economics*, 11:25–45.

[10] Bapna, A., and T. Weber (2005) "Efficient Dynamic Allocation with Uncertain Valuations," working paper, Stanford University

[11] Battaglini, M. (2003), "Optimality and Renegotiation in Dynamic Contracting," Mimeo, Princeton University.

[12] Battaglini, M. (2005), "Long-Term Contracting with Markovian Consumers," *American Economic Review* 95(3), pages 637-658.

[13] Behrends, Ehrhard (2000). *Introduction to Markov Chains, with Special Emphasis on Rapid Mixing*. Vieweg, 2000.

[14] Bergemann, Dirk, and Juuso Valimaki (2006), "Efficient Dynamic Auctions," Cowles Foundation Discussion Paper 1584

[15] Bester, H. and Strausz, R. (2001), "Contracting with Imperfect Commitment and the Revelation Principle: The Single Agent Case," *Econometrica*, 2001, 69 (4), pp. 1077–1098.

[16] Courty, P., and H. Li (2000), "Sequential Screening," *Review of Economic Studies*, 67(4), page 697-717.

[17] Cremer, J., Y. Spiegel, and C. Zheng (2003), "Optimal Selling Mechanisms with Costly Information Acquisition," IDEI Working Paper.

[18] Ericson, R. and A. Pakes (1995), "Markov-Perfect Industry Dynamics: A Framework for Empirical Work," *Review of Economic Studies* 62 (1), 53-82.

[19] Fadel, R., and I. Segal (2006), "The Communication Cost of Selfishness," working paper, Stanford University.

[20] Friedman, E., and D. C. Parkes (2003), "Pricing WiFi at Starbucks–Issues in Online Mechanism Design," *Fourth ACM Conference on Electronic Commerce*, pp. 240-241.

[21] Fudenberg, D., D. Levine, and E. Maskin, "The Folk Theorem with Imperfect Public Information," *Econometrica*, 62, 997-1039.

[22] Fudenberg, D. and J. Tirole (1991a), *Game Theory*, MIT Press, Cambridge, MA.

[23] Fudenberg, D. and J. Tirole (1991b), "Perfect Bayesian Equilibrium and Sequential Equilibrium," *Journal of Economic Theory*, 53, 236-260.

[24] Gershkov, A., and B. Szentes (2004), "Optimal Voting Scheme with Costly Information Acquisition," mimeo, University of Chicago.

[25] Jehiel, P. and B. Moldovanu (2001), "Efficient Design with Interdependent Valuations," *Econometrica* 69: 1237-1259.

[26] Kohlberg, E., and J.-F. Mertens (1986), "On the strategic stability of equilibria," *Econometrica* 54:1003-1008.

[27] Krishna, V. and J. Morgan (2004), "Contracting for Information under Imperfect Commitment," Mimeo, Berkeley.

[28] Larson, K. and T. Sandholm (2001), "Costly Valuation Computation in Auctions." In *Proceedings of the Theoretical Aspects of Reasoning about Knowledge (TARK)*, pp. 169-182.

[29] Levin, J. (2003), "Relational Incentive Contracts," *American Economic Review*, June.

[30] Mailath, G. and A. Postlewaite (1990), "Asymmetric Information Bargaining Problems with Many Agents," *Review of Economic Studies*, 57, 351-360.

[31] Mezzetti, C. (2004), "Mechanism Design with Interdependent Valuations: Efficiency," *Econometrica* 72(5): 1617.

[32] Miller, D. (2004), "The dynamic cost of ex post incentive compatibility in repeated games of private information," Mimeo, UCSD.

[33] Myerson, R. (1986), "Multistage Games with Communication," *Econometrica*, 54, 323-358.

[34] Myerson, R. (1991), *Game Theory: Analysis of Conflict*, Harvard University Press.

[35] Obara, I., and D. Rahman (2006), "Approximate Implementability with Ex Post Budget Balance," working paper, UCLA Department of Economics.

[36] Pollard, D. (2002), *A User's Guide to Measure-Theoretic Probability*, Cambridge University Press.

[37] Proschan, M.A., and B. Presnell (1998), "Expect the Unexpected from Conditional Expectation," *The American Statistician*, 52(3), 248-252.

[38] Puterman, M. (1994), *Markov Decision Processes*. Wiley: Hoboken, NJ.

[39] Rayo, L. (2003), "Relational Team Incentives and Ownership," Mimeo, Chicago GSB.

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

[41] Smorodinsky, R., and M. Tennenholtz (2003), "Overcoming Free Riding in Multi-Party Computations-The Anonymous Case," Mimeo, Technion.

[42] Stokey, N. and R. Lucas (1989), *Recursive Methods in Economic Dynamics.*

[43] Wang, Cheng (1995), "Dynamic Insurance with Private Information and Balanced Budgets," *Review of Economic Studies* ; 62.4, October, pp. 577-95.

8. Appendix

Lemma 2. *A functional* $F : \prod_{t=1}^{\infty} Z_t \to \mathbb{R}$ *is LCI with some* $\in (0;1)$ *if and only if* $F(z) = \sum_{t=0}^{\infty} {}^{t} f_t(z^t)$ *for some uniformly bounded sequence of functions* $\{f_t : Z^t \to \mathbb{R}\}_{t=0}^{\infty}$.

Proof of Lemma 2: "If": Letting $M = \sup_{i;t;z^t \in Z^t} \left| f\left(z^t\right)\right| < \infty$, we can write

$$|F(z) - F(y)| \leq \sum_{t=\inf\{\ :z_\tau \neq y_\tau\}}^{\infty} {}^{t} \left| f_t(z^t) - f_t(y^t)\right| \leq \sum_{t=\inf\{\ :z_\tau \neq u_\tau\}}^{\infty} {}^{t} \cdot 2M$$
$$= \frac{2M}{1-} {}^{\inf\{\ :z_\tau \neq y_\tau\}} = \frac{2M}{1-} \quad (z;y):$$

"Only if": fix $\bar{z} \in Z$ and take $f_0 = F(\bar{z})$ and $f_t\left(z^t\right) = {}^{-t}\left[F\left(z^t;\bar{z}^{-t}\right) - F\left(z^{t-1};\bar{z}^{-(t-1)}\right)\right]$ for $t \geq 1$. Then we have $\left|f_t\left(z^t\right)\right| \leq {}^{-t}C \cdot \quad \left(\left(z^t;\bar{z}^{-t}\right);\left(z^{t-1};\bar{z}^{-(t-1)}\right)\right) \leq C$ for $t \geq 1$. Furthermore,

$$\left|\sum_{t=0}^{T} {}^{t} f_t(z^t) - F(z)\right| = \left|F\left(z^T;\bar{z}^{-T}\right) - F(z)\right| \leq C \quad \left(\left(z^T;\bar{z}^{-T}\right);z\right) \leq C \ {}^{T} \to 0 \text{ as } T \to \infty. \blacksquare$$

Proof of Lemma 1: It suffices to establish a bound for any given pair $;\ ' \in \bar{\Theta}$, since the state space $\bar{\Theta}$ is finite. For the "transient" states $' \in \bar{\Theta} \backslash S$, the expression is bounded since by definition $\sum_{t=0}^{\infty} {}^{t} p^{(t)}_{;\ '} < \infty$ and $\ _{'} = 0$. For the "ergodic" states $' \in S$, note that all such states have the same period $n \geq 1$, and let $G \subset S$ denote the cyclically moving subset containing $'$. Let $t_0 < n$ be the first $t \geq 0$ s.t. $p^{(t)}_{;'} > 0$. Then $p^{(t)}_{;'} > 0$ only for $t = t_0 + rd$, for $r = 0; 1; :::$. Furthermore, the n-step transition process with the transition matrix $p^{(n)}_{;'}$ behaves like an irreducible aperiodic chain on G, with the invariant probabilities described by $d \;\; _{'}$ on $' \in G$. Applying Lucas-Stokey (1989, Theorem 11.4), the n-step process converges geometrically to its invariant distribution, i.e., there exists $\;\; < 1$ such that $\left| p^{(t_0 + rd)}_{;'} - d \;\; _{'} \right| \leq \;\; ^{r}$. Therefore,

$$
\begin{aligned}
\left| \sum_{t=0}^{\infty} \;\; ^{t} \left(p^{(t)}_{;'} - \;\; _{'} \right) \right| & \leq \sum_{t=0}^{\infty} \;\; ^{t_0 + rd} \left| p^{(t_0 + rd)}_{;'} - d \;\; _{'} \right| + \;\; _{'} \left| n \sum_{t=0}^{\infty} \;\; ^{t_0 + rd} - \sum_{t=0}^{\infty} \;\; ^{t} \right| \\
& \leq \sum_{r=0}^{\infty} \;\; ^{r} + \sum_{r=0}^{\infty} \;\; ^{rd} \left| d \;\; ^{k_0} - \sum_{t=0}^{n-1} \;\; ^{t} \right| \\
& \leq \frac{1}{1 - } + \frac{1}{1 - \;\; ^{n}} n \left(1 - \;\; ^{n-1} \right) \leq \frac{1}{1 - } + n: \blacksquare
\end{aligned}
$$