Efficient Mechanisms with Dynamic Populations and Dynamic Types*

Ruggiero Cavallo
SEAS, University of Pennsylvania,
Philadelphia PA, USA
cavallo@cis.upenn.edu

David C. Parkes
SEAS, Harvard University
Cambridge MA, USA
parkes@eecs.harvard.edu

Satinder Singh
Computer Science and Engineering
University of Michigan
Ann Arbor MI, USA
baveja@umich.edu

First version: June 5, 2007, Revised: December 19, 2008
This version: July 7, 2009

Abstract

We consider the truthful implementation of an efficient decision policy when agents have dynamic type and are periodically-inaccessible, with agents unable to report information or make payments while inaccessible. This concept of inaccessibility includes a model of arrival-departure dynamics as a special case. We generalize the *dynamic VCG* mechanism [Bergemann and Välimäki, 2008] to this environment, achieving within-period *ex post* incentive compatibility for agents with the same communication constraints as the center. In doing so, we offer a new proof of the correctness of the dynamic VCG mechanism, emphasizing its position within a family of dynamic Groves mechanisms. In considering the special case of an arrival-departure model with dynamic type, we obtain a mechanism that is efficient and within-period *ex post* incentive compatible for arrival processes in which future arrivals are conditionally independent of past arrivals given the actions of the center. The mechanism is shown to be payoff equivalent at arrival for agents with static types to the *online VCG* mechanism [Parkes and Singh, 2003], which satisfies a stronger *ex post* participation constraint than the dynamic VCG mechanism in such domains. In closing, we highlight a structural difficulty in extending the dynamic VCG mechanism to achieve an *ex post* no deficit, efficient mechanism in an environment with dynamic types and interdependent type transitions.

*This paper subsumes a previous version entitled "Efficient Online Mechanisms for Persistent, Periodically Inaccessible Self-Interested Agents", dating from June, 2007. Thanks to seminar participants at Dagstuhl, Stonybrook, GAMES, Cornell, Yale, EPFL, NYU Stern, Aarhus, and Stanford for useful feedback. Thanks also to Sven Seuken for comments on an earlier version.

1 Introduction

Mechanism design addresses social choice problems involving multiple self-interested agents, each with private information relevant to the decision and a utility function on outcomes. The problem is to design a communication game such that there is a non-cooperative equilibrium in which the outcome satisfies a desired set of properties. Mechanism design has mainly focused on static problems with a fixed population of agents, each with fixed private information, and a fixed set of outcomes. But many problems of interest are dynamic; e.g., a repeated allocation problem in which agents are learning their value for a resource, or selling a time-sensitive good such as a theater ticket to impatient buyers arriving that arrive at different times.

In extending mechanism design to dynamic and uncertain environments, most models consider either a dynamic population with static, private information [Lavi and Nisan, 2004; Parkes and Singh, 2003] or a persistent population of agents with dynamic, privately revealed information [Athey and Segal, 2007; Cavallo *et al.*, 2006; Bergemann and Välimäki, 2008]. The former problems have been described as those of online auctions or online mechanism design (emphasizing the analog to *online* algorithm problems of computer science and operations research), while the latter have been described as problems of *dynamic* mechanism design.

In this paper, we unify these models and allow for both dynamic populations and dynamic types by considering the possibility that agents can transition between accessible and inaccessible local states. While inaccessible, an agent is unable to communicate with a central planner or receive payments but remains important because it continues to have values for decisions, both now and in the future. In addition to modeling domains in which agents become disconnected from a mechanism because of faulty technology or reasons of limited attention, the generalization to periodically-inaccessible agents is helpful because it provides insight into the design of an efficient mechanisms for arrival-departure models. Models of dynamic arrival and departure are captured by restricting inaccessibility to a single cycle, so that an agent is first inaccessible, then "arrives" and is accessible according to a probabilistic model known to the center, and then "departs" and is permanently inaccessible. Upon departure, an agent has no value for any further actions and thus will no longer be pivotal to the decisions of an efficient policy.

The focus in this paper is on the truthful implementation of the efficient policy in such domains, and we consider a private values model and, for the most part, independent type transitions. We show that this independent transition requirement translates into an independence requirement on the arrival process in a domain with arrival-departure dynamics. We first review the case of persistent agents with dynamic type, and offer an independent proof of the incentive-compatibility of the dynamic VCG mechanism [Bergemann and Välimäki, 2008] (referred to as the "dynamic pivot" mechanism in their paper), that emphasizes its position within a family of dynamic Groves mechanisms. We then present our general model, in which agents may also become periodically inaccessible, and extend the dynamic VCG mechanism by maintaining beliefs about the type of inaccessible agents. We obtain the *dynamic VCG# mechanism*, which retains the same within-period *ex post* incentive compatibility property of the dynamic VCG mechanism when the information that accessible agents have about inaccessible agents is restricted in the same way as the center through failure of communication. An additional requirement is that agents may not be inaccessible forever, in order to allow for transfers to reflect decisions made while an agent is inaccessible.

In restricting to an arrival-departure model, but still retaining dynamic types, we recover the dynamic VCG mechanism, suitably modified to select the actions of a policy that is efficient given that the center has a correct probabilistic model of the arrival process. The dynamic VCG mechanism is within-period *ex post* incentive compatible and efficient, for an arrival process in which the types of new agents are conditionally independent of past arrivals, given the actions of the mechanism. Under a further restriction to *static types*, so that an agent gains no additional private information subsequent to its arrival, we establish the payoff equivalence upon arrival between the dynamic VCG mechanism and the *online VCG* mechanism [Parkes and Singh, 2003]. In doing so, we unify within a single model these separate threads in the recent literature on dynamic mechanism design. The online VCG mechanism may be preferred over the dynamic VCG mechanism in this dynamic population, static type environment because it satisfies *ex post* participation where the dynamic VCG mechanism satisfies only within-period *ex post* participation.

In closing, we return to a domain with persistent and accessible agents and dynamic type, allowing here for interdependent type transitions. As in Mezzetti [2004] (for a static problem) and Athey and Segal [2007] (for the dynamic problem), we emphasize the importance of private value when conditioned on state transitions and also that the information externality present in another agent's type is subsumed by private information realized to an agent subsequent to an action by the center. Both requirements are easily modeled as a simple generalization to our framework, where the state transition of an agent can depend on the private state of other agents and an agent's value for an action depends on its current state and its next state. Our contribution is to highlight a structural difficulty in extending the construction of the dynamic VCG mechanism to achieve a mechanism with transfers that are *ex post* no deficit.

1.1 Related work

We cateogorize the related work by the kind of environment considered. First, we review related work for an environment with a persistent population and dynamic type. This includes the efficient mechanisms of Athey and Segal [2007], Bergemann and Välimäki [2008] and Cavallo et al. [2006], along with some mechanisms that have been developed for special cases. Then we review related work for an environment with a dymamic population but static type. This includes the efficient mechanism of Parkes and Singh [2003], as well as a number of mechanisms for special cases including both revenue-maximizing mechanisms and prior-free mechanisms that are analyzed within a worst-case framework. Finally, we mention a recent algorithmic development inspired by an extension to the model in this paper, in which agents can take local actions while inaccessible.

Persistent population, dynamic type. Athey and Segal [2007] obtain a Bayes-Nash incentive compatible, efficient and budget-balanced mechanism for a persistent-population, dynamic type environment with private values and independent type transitions. The mechanism extends the expected externality mechanism [Arrow, 1979; d'Aspermont and Gérard-Varet, 1979] to a dynamic environment. The main limitation is that it is only able to provide for *ex ante* participation constraints, although the authors also characterize a sufficient conditions for *interim* participation in an infinite horizon setting with sufficiently patient agents. Athey and Segal [2007] also present the team mechanism (or dynamic Groves) mechanism for interdependent type transitions, but do not present a dynamic VCG (or pivot) mechanism.

Bergemann and Välimäki [2008] obtain the dynamic VCG mechanism for the same environment, with a persistent population, independent type transitions, and private values. The dynamic VCG mechanism truthfully implements an efficient policy, and provides within-period *ex post* incentive compatibility and within-period *ex post* individual-rationality. The dynamic VCG mechanism is also *ex post* no deficit in economic environments without positive externalities (e.g., one-sided auctions or social-choice problems.) The authors establish that the mechanism is the unique mechanism with these properties amongst those that additionally satisfy an *efficient exit* property, which requires that no transfers should occur once an agent's reports are no longer pivotal.[1] Applications are given to a scheduling problem, and also to a problem with Bayesian learning by agents, modeled as a multi-armed bandits auction.

Cavallo et al. [2006] independently develop a different variation on a dynamic VCG mechanism for the same environment, with a persistent population, independent type transitions, and private values. The mechanism modifies the team mechanism with a single *ex ante* charge-back term whose flow value can be smoothed over the course of the mechanism. As such, it is a dynamic Groves mechanism and within-period *ex post* incentive compatible. The main limitation is that it provides only *ex ante* participation, while the dynamic VCG mechanism developed in Bergemann and Välimäki [2008] provides within-period *ex post* participation. The authors also provide an application to a multi-armed bandits auction, which provides a simple model of learning by doing and has a tractable planning problem via an index policy.[2]

[1] The uniqueness result holds for persistent population, dynamic type environments but not for dynamic population, static type environments, in which incentive-compatibility constraints only need to bind upon arrival. For example, the online VCG mechanism [Parkes and Singh, 2003] also satisfies these properties in such an environment but has different flow payoffs forward from states other than an agent's arrival.

[2] Bapna and Weber [2005] also study this multi-armed bandits auction but are unable to identify

Cremer et al. [2009] independently develop a special case of the Bergemann and Välimäki [2008] dynamic VCG mechanism for an application with one-time type transitions, modeling costly information acquisition by agents. In an application in which each agent has instead a general, sequential process for costly value refinement, Cavallo and Parkes [2008] apply the dynamic VCG mechanism (subsequent to its development in Bergemann and Välimäki [2008]), obtaining a reduction to a multi-armed bandits auction problem. In handling Larson and Sandholm's problem of "strategic deliberation," in which an agent also has technology for costly deliberation about the value of *another* agent, these authors also extend the dynamic VCG mechanism in order to retain incentive compatibility despite the limited interdependence that this introduces into agent type transitions.

Pavan et al. [Pavan *et al.*, 2009] develop necessary and sufficient conditions for Bayes-Nash incentive compatibility for a persistent population, dynamic type environment with independent type transitions and interdependent valuations. In addition to obtaining a general revenue-equivalence result for dynamic environments, the analysis is applied to develop optimal dynamic mechanisms for environments in which agent type transitions are modeled as an auto-regressive process. The main limitation is that the characterization only provides *ex ante* participation.

Cavallo [2008] characterizes conditions under which dynamic Groves mechanisms are unique amongst efficient, within-period *ex post* incentive-compatible mechanisms, and for which the dynamic VCG mechanism is revenue maximizing amongst this class. A redistribution mechanism is also introduced for a multi-armed bandits auction, redistributing payment streams in the dynamic VCG mechanism to reduce the budget surplus without running at a deficit or losing incentive compatibility. This adapts to dynamic environments the framework of the redistribution for Groves mechanisms in static environments [Bailey, 1997; Cavallo, 2006].

An earlier literature develops dynamic mechanisms for persistent agents with *time-separable* types. For example, Atkeson and Lucas [1992] consider a continuum population in which agents receive new i.i.d. types each period, and characterize incentive-compatible distribution policies for a time-sensitive good. Athey et al. [2004] also adopt a dynamic mechanism design approach in analyzing the equilibrium behavior of two competing firms, each with a private cost sampled i.i.d. in each period. In a single agent, non time-separable problem, Courty and Li [2000] consider a problem with two periods and one persistent agent, designing an optimal mechanism in which an agent learns a distribution about its value for a good in period one and before its value is realized in period two.[3]

a first-best mechanism, instead providing bounds on the equilibrium behavior in a non incentive-compatible mechanism.

[3] Also related is the literature on dynamic contracting models, where the focus is on the role of commitment in limiting what a principal can achieve in problems with moral hazard. Freixas et al. [1985] obtain a characterization of the second-best due to lack of commitment for an agent with static type. For an agent with a dynamic but time-separable type, Wang [1995] studies the dynamic allocation of a perishable good between two firms and Levin [2003] considers contracting under incomplete information without the ability to commit even to payments based on observable

Dynamic population, static type. Parkes and Singh [2003] obtain the online VCG mechanism for an environment with a dynamic population and static type, with a known probabilistic model of the arrival process and agents with private values on sequences of decisions.[4] An agent's valuation function completely realized upon its arrival and can misreport its arrival period and type. The online VCG mechanism is within-period *ex post* incentive compatible and efficient, collecting a single payment from an agent in its *commitment period*, which is the first period in which all decisions with respect to the agent's value are determined.[5] The mechanism is *ex post* individual-rational and *ex ante* no deficit. The mechanism is developed in the context of a finite time horizon, but it is a simple matter to generalize the mechanism to an environment with an infinite time horizon and discounting, as we do in this paper.[6] Mierendorff [2008] specializes the online VCG mechanism to an application with a single item to sell to buyers that arrive with a value schedule for receiving the item in different time periods, obtaining a mechanism that is efficient and both *ex post* no deficit and *ex post* individual-rational by reconfiguring the payment flows.

In an early contribution, Dolan [1978] develops an online VCG mechanism for a scheduling domain with Poisson arrivals and agents with different delay costs. No consideration is giving to temporal strategies, and agents are only able to misreport their cost of delay. On the other hand, the author is able to characterize the efficient policy and anticipates the recent developments in the literature, by proposing to charge an agent the expected externality it imposes on the system upon its arrival. Dolan [1978] also observes that the dynamic mechanism will not be dominant strategy incentive compatible because it requires agreement about the probabilistic model of the arrival process.

Looking at recent applications, Gershkov and Moldovanu [2008a; 2008b] study efficient and revenue-optimal mechanisms for a problem with commonly ranked, distinct items to allocate by a deadline to buyers that arrive with unit demand. Subsequent to this, these authors also consider a model in which the center is learning

outcomes. Battaglini [2005] considers an agent with a Markovian type dynamic (and thus non time-separable) and shows that commitment is no longer necessary to implement a first-best contract.

[4]The phrasing "online" VCG comes from the analogy to online algorithms in computer science, in which decisions are made sequentially as new parts of the input are revealed.

[5]The authors establish that the mechanism is Bayes-Nash incentive compatibility. In this paper we emphasize that the online VCG mechanism achieves this stronger, *ex post* incentive-compatibility property. We also emphasize the requirement that the arrival process satisfy a independence property, with new arrivals conditionally independent of past arrivals given decisions by the mechanism.

[6]Earlier, Friedman and Parkes [2003] had developed a Bayes-Nash incentive-compatible, online VCG mechanism in which all payments are delayed until the final time period. The proof technique of Parkes and Singh [2003] establishes that the flow payoff to an agent in the online VCG mechanism is the same as in this delayed VCG mechanism. Parkes et al. [2004] slightly generalize the incentive analysis of the online VCG mechanism to allow for an ϵ-efficient policy, obtaining an ϵ-Bayes Nash incentive compatible and ϵ-efficient mechanism, and present simulation results for an application to a scheduling problem, in which the mechanism is coupled with a sparse-sampling algorithm to compute approximately-efficient decisions in each period and estimate the expected marginal product of an agent in order to determine payments.

about the arrival process, identifying a way in which interdependence is introduced into the problem [Gershkov and Moldovanu, 2008c] and developing sufficient conditions on the environment that allows for first-best policies, under Bayesian and non-Bayesian learning paradigms [Gershkov and Moldovanu, 2009].[7] Dizdar et al. [2009] study a dynamic knapsack auction, with a finite number of items to allocate by some deadline to impatient agents that each demand some number of items. The authors develop conditions under which the first-best is possible for efficiency and revenue optimality. Constantin and Parkes [2009] study a generalization in which buyers are patient, with an arrival time and departure time (or deadline), and address the computational difficulty of computing efficient allocation decisions by adopting an online stochastic combinatorial optimization algorithm that is automatically modified to make it dominant-strategy implementable.

There is also by now a rich literature on revenue-optimal, dynamic mechanisms. Vulcano et al. [2002] develop revenue-optimal auctions in a standard unit-demand model with i.i.d. agent values, impatient agents and a finite time horizon. In models with discounting and infinite time horizons, Gallien [2003] develops a revenue-optimal auction for a continuous time model, while Said [2009] establishes the revenue-optimality of a dynamic VCG mechanism defined on virtual valuations for a setting with discrete time periods and an expiring good, one unit of which must be allocated in each period. One of the few developments of an indirect mechanism is also provided by Said [2009], via a sequence of ascending-price auctions. In a finite-time horizon model with patient agents that arrive with a deadline and are indifferent between receiving an item in any period between arrival and departure, Pai and Vohra [2008] obtain sufficient conditions under which the revenue-optimal auction is implementable and Mierendorff [2009] studies the irregular case for allocating a single item over two time periods, developing the revenue-optimal mechanism.[8]

A sequence of papers adopt *worst-case* rather than Bayesian analysis for the design of online mechanisms. The objective is to develop mechanisms that can do well relative to what would be possible in a static problem with all type information available in the first period, and whatever the actual (dynamic) realization of agent types. Initiating this line of research, Lavi and Nisan [2004] provide a worst-case analysis for an online auction in which a number of identical goods are sold by some deadline to agents with marginal-decreasing values for each additional good. Ng et al. [2003] develop a strategyproof online auction for a model of patient agents in application to non-preemptive scheduling, in which agents need an allocation of

[7] Riley and Zeckhauser [1983] had earlier considered the special case of selling one item to a sequence of buyers that arrive, one per period, in which the seller seeks to maximize revenue while learning about the buyer type distribution.

[8] Also related is a literature that studies the problem of a monopolist selling multiple items of a durable good via posted price mechanisms to dynamic arrivals of impatient, unit-demand agents. For continuous time models, revenue-optimal price schedules are developed in a series of papers by Kincaid and Darling [1963], Gallego and Van Ryzin [1994] for an exponential demand model, and McAfee and te Velde [2008] for a Pareto demand model. Board [2008] develops revenue-optimal pricing schedules with buyers that arrive over time and discount the future, in a discrete time setting with time-varying demand.

resources by some period in time. Porter [2004] provides a worst-case analysis with respect to efficiency in a related, preemptive scheduling problem. Hajiaghayi et al. [2005] improve these results for the special case of agents with unit-length jobs. All of these mechanisms are dominant-strategy incentive compatible for restricted type misreports, in which agents that cannot report early arrivals and the result of scheduling a job can be held until an agent's reported departure. Hajiaghayi et al. [2005] also develop a characterization of dominant-strategy implementable policies under different assumptions on available misreports.[9][10]

For durable goods, Hajiaghayi et al. [2004] consider patient unit-demand buyers and develop an adaptive, dominant-strategy incentive compatible online mechanism as a variation on the classical *secretary problem* for the problem of allocating a single item by a deadline. Juda and Parkes [2009] develop a dominant-strategy incentive-compatible auction for allocating an uncertain supply of distinct items to a dynamic population of buyers with deadlines and general (combinatorial) valuations, and provide an empirical analysis of its efficiency and revenue in application to a market on eBay for LCD monitors. Bredin and Parkes [2007] develop a dominant-strategy incentive-compatible double auction for dynamic allocation between arriving buyers and sellers with deadlines, providing a dynamic generalization of the McAfee [1992] trade-reduction auction rule.

Dynamic population, Dynamic type. Subsequent to this present paper, Seuken et al. [2008] extend the partially-inaccessible model to allow for agents with private actions. In such an environment, the planner should provide agents with *emergency polices*, to be used by an agent to select local actions while inaccessible. The authors provide a team mechanism but are unable to develop a dynamic VCG mechanism, and focus instead on computational issues in computing these emergency policies, making connections with the established literature within AI on decentralized Markov decision processes.

2 A Fixed Population with Dynamic Types

Consider an environment with a fixed set $I = \{1, \dots, n\}$ of agents able to communicate with a central decision-maker ("the center"). Actions are selected by a

[9]The condition in Theorem 6 in Hajiaghayi et al. [2005] was subsequently shown to be sufficient but not necessary; see Theorem 1.25 in Parkes [2007] for a clarification.

[10]A number of variations on these scheduling problems have been considered: e.g., Cole et al. [2008] require "prompt payments" that are collected immediately upon an allocation of resources; Lavi and Nisan [2005] adopt a set-Nash analysis of a sequence of ascending-price auctions in a problem in which late misreports of departure are possible; and Lavi and Segev [2008] adopt an analysis in undominated strategies for a variation in which agents are patient until the final time period upon arrival. Mahdian and Saberi [2006] study a variation in which the buyer population is fixed but the supply of expiring items is online and uncertain. Babaioff et al. [2009] add the additional requirement of prompt payments and provide a hybrid analysis, that is worst-case with respect to agent valuations but average-case with respect to a probabilistic model of supply and requires dominant-strategy incentive compatibility with regard to any possible supply realization, precluding the use of an online VCG mechanism.

center in each of a sequence of discrete time periods, perhaps infinite. The center chooses an action $a \in A$ from the set of possible actions A in each period. Each agent $i \in I$ has a private type that consists of *state*, $s_i \in S_i$ in state space S_i, a *stochastic transition function*

$$\tau_i : S_i \times A \rightarrow S_i, \tag{1}$$

such that for all $s_i \in S_i$ and $a \in A$, $\sum_{s'_i \in S_i} Pr(\tau_i(s_i, a) = s'_i) = 1$ (for probability function Pr), and a *reward function*

$$r_i : S_i \times A \rightarrow \mathcal{R}, \tag{2}$$

which defines the reward $r_i(s_i, a)$ to an agent when the center takes action a given state s_i. This is a *private-values* environment, because an agent's reward $r_i(s_i, a)$ depends only on its own state. Moreover, type transitions are independent because $\tau_i(s_i, a)$ is conditionally independent of other agents' types, when conditioned on the actions taken by the center.[11] Collectively, we refer to an agent's type, $t_i = (s_i, \tau_i, r_i) \in T_i$, as a *dynamic type* when the transition function is stochastic, and thus new information is privately revealed to an agent over time. For a deterministic transition function this is equivalent to a standard (static) type, describing a value for any possible sequence of actions, and an agent can completely describe its value for any sequence of actions in a single report.

We associate the center with state, $s_0 \in S_0$ and stochastic transition function $\tau_0 : S_0 \times A \rightarrow S_0$ to model the feasible actions available in each period. In an auction problem, this state can include information about which items are already allocated, for example, with $A(s_0) \subseteq A$ to denote the feasible allocations in the current period. To keep our presentation simple we shall simply adopt fixed action set A in what follows, but $A(s_0)$ can be easily substituted.[12]

Let $t_0 \in T_0$ define (s_0, τ_0), and $T = T_0 \times T_1 \times \ldots \times \ldots T_n$ denote the type space with $t \in T$ a joint type profile. Similarly, let $S = S_0 \times S_1 \times \ldots \times S_n$ denote the state space. Let $\tau(s, a) = (\tau_0(s_0, a), \tau_1(s_1, a), \ldots, \tau_n(s_n, a)) \in S$, and define $\tau(t, a)$ analogously. Let $t_{-i} = (t_0, \ldots, t_{i-1}, t_{i+1}, \ldots, t_n)$ denote the type profile without agent i, and s_{-i} the state profile without agent i.

We assume discount factor $0 < \gamma \leq 1$, that is common to all agents, and let K denote

[11] Without inaccessibility, and thus precluding a setting with agent arrivals and departures, there are no additional technical difficulties in also allowing for serially correlated types, wherein the transitions $\tau_i(s_i^k, a, z^k) \in S_i$ and rewards $r_i(s_i^k, a, z^k)$ also depend on an exogenous random process $z^0, z^1, \ldots$ observable to all agents. Serially correlated types are considered, for example, in, Athey and Segal [2007] and Cavallo [2008]. But this presents a problem with inaccessible agents (e.g., dynamic arrival-departures) because it presents an information externality [Gershkov and Moldovanu, 2008c].

[12] The center may also have reward $r_0(s_0, a)$ for action a, for example to model the value of leaving a resource unallocated or to indicate through a large negative reward that action a is infeasible given s_0 (e.g., perhaps all resources have been allocated). This presents no additional technical difficulty.

a time horizon, perhaps infinite so that the decision periods are $\ell \in \{0, 1, \dots, K\}$.[13] A decision policy, $\pi : T \to A$ defines an action in every type profile $t \in T$. Fixing τ and reward functions $r_1, \dots, r_n$, then value function,

$$V_i(t^\ell, \pi) = \mathbb{E}_{t^k}\left[\sum_{k=\ell}^{K} \gamma^{k-\ell}\, r_i^k(t_i^k, \pi(s^k)) \mid t^\ell, \pi\right], \tag{3}$$

denotes the expected discounted reward to agent i, or *flow value*, given policy π and type profile t^ℓ, with $t^k = \tau(t^{k-1}, \pi(t^{k-1}))$ for $k > \ell$. Let $r(t^\ell, a)$ denote $\sum_{i \in I} r_i(s_i^\ell, a)$, with $V(t^\ell, \pi) = \sum_{i \in I} V_i(t^\ell, \pi)$. An *efficient decision policy*, π^*, solves,

$$\pi^* \in \arg\max_{\pi \in \Pi} V(t, \pi), \quad \forall t \in T, \tag{4}$$

where Π is the space of feasible policies. Let $V_{-i}(t^\ell, \pi) = \sum_{j \neq i} V_j(t^\ell, \pi)$, with π^*_{-i} to denote a decision policy that is efficient for the other agents, i.e., $\pi^*_{-i} \in \arg\max_{\pi \in \Pi} V_{-i}(t, \pi), \forall t \in T$.

From a modeling perspective, we have defined a *loosely-coupled multi-agent Markov Decision Process* (MDP), in which each agent has a local state, stochastic transition function that is independent of that of other agents (conditioned on actions), and reward function; see Cavallo et al. [2006]. Note that the only dynamic component of type is the state, the reward and transition functions remain invariant. The joint MDP is coupled through the single action that is taken in each period, which depends in turn on the state of the center and thus can depend on past actions.

An incentive-compatible mechanism will implement the efficient joint policy even though each agent can choose to misreport its state, transition function, and reward function.

2.1 A Dynamic Mechanism

A *dynamic mechanism*, $M = (\pi, T)$, is defined by a decision policy $\pi : T \to A$, and a transfer policy $x = \{x_1, \dots, x_n\}$, with $x_i : T \to \mathcal{R}$, for all $i \in I$.[14] Given type profile t^ℓ in period ℓ, the mechanism selects action $a^\ell = \pi(t^\ell) \in A$ and makes transfer $x_i(t^\ell) \in \mathcal{R}$ to each agent.

Each agent can make a report about its type in each period. An agent's strategy can depend on prior reports by itself and by other agents. Let $h^\ell \in \mathcal{H}^\ell$ denote a

[13] All the results in this paper can be developed with simple modification for a finite time horizon problem without discounting.

[14] Myerson [1986] gives a general revelation principle for dynamic communication games; see also Athey and Segal [2007]. The revelation principle does not hold directly in environments with communication constraints such as those studied here; e.g., an indirect mechanism can have more implementation power because it can prevent agents making incredible claims via a direct report in some earlier period that they will still be accessible in a later period (see Parkes [2007] for an example). But given that we obtain the truthful implementation of efficient policies with direct mechanisms we will not consider indirect mechanisms further in this paper.

sequence of type reports sent to the center up to and including period ℓ, from the set of possible reports $\mathcal{H}^\ell$. A strategy

$$\sigma_i^\ell : \mathcal{H}^{\ell-1} \times T_i \to T_i, \tag{5}$$

defines the report made by agent i given this history and given its current state. Let $\sigma^\ell(h^{\ell-1}, t) = (t_0, \sigma_1^\ell(h^{\ell-1}, t_1), \ldots, \sigma_n(h^{\ell-1}, t_n)) \in T$, denote the reported type profile in period ℓ when the true type is t, given history $h^{\ell-1} \in \mathcal{H}^{\ell-1}$ and strategy profile $\sigma = (\sigma_1, \ldots, \sigma_n)$. We write the (true) expected discounted value (or *flow value*) to agent i given strategy σ_i when agents $\neq i$ are truthful, as

$$V_i(h^{\ell-1}, t^\ell, \pi, \sigma_i) = \mathbb{E}_{s_\sigma^k, a_\sigma^k}\left[\sum_{k=\ell}^{K} \gamma^{k-\ell}\, r_i^k(s_{\sigma,i}^k, a_\sigma^k) \,|\, h^{\ell-1}, t^\ell, \pi, \sigma_i\right], \tag{6}$$

where s_σ^k denotes the (true) state profile in period k and a_σ^k the action taken by mechanism policy π, both induced by agent i's strategy σ_i. Letting t_σ^k denote the state profile in period k, then we have $s_\sigma^{k+1} = \tau(s_\sigma^k, a_\sigma^k)$, $a_\sigma^{k+1} = \pi(\sigma_i^k(h^{k-1}, t_{\sigma,i}^k), t_{\sigma,-i}^k)$, and $h^k = (h^{k-1}, (\sigma_i^k(h^{k-1}, t_{\sigma,i}^k), t_{\sigma,-i}^k))$. We similarly define the expected discounted sum of transfers, or *flow transfer*, to agent i forward from period ℓ, given that it adopts strategy σ_i and other agents are truthful, as

$$X_i(h^{\ell-1}, t^\ell, \pi, \sigma_i) = \mathbb{E}_{t_\sigma^k}\left[\sum_{k=\ell}^{K} \gamma^{k-\ell} x_i(\sigma_i^k(h^{k-1}, t_{\sigma,i}^k), t_{\sigma,-i}^k) \,|\, h^{\ell-1}, t^\ell, \pi, \sigma_i\right], \tag{7}$$

where t_σ^k is the type profile induced in period k by strategy σ_i.

Agents have quasi-linear utility functions, so that an agent's expected discounted utility (or *flow payoff*), given joint type t^ℓ and strategy σ_i is just

$$V_i(h^{\ell-1}, t^\ell, \pi, \sigma_i) + X_i(h^{\ell-1}, t^\ell, \pi, \sigma_i). \tag{8}$$

Let σ_i^* denote a truthful strategy, with $\sigma_i^*(h^{\ell-1}, t_i^\ell) = t_i^\ell$ for all ℓ, all $h^{\ell-1}$ and all t_i^ℓ. Following Bergemann and Välimäki [2008] and Athey and Segal [2007], we define a within-period *ex post* incentive-compatible mechanism:

Definition 1 (within-period ex post incentive compatible)**.** *A dynamic mechanism, $M = (\pi, x)$, is within-period ex post incentive-compatible if, for all times ℓ, for any agent $i \in I$, for any type profile $t^\ell \in T$, for any history $h^{\ell-1}$, and for all $\sigma_i' \neq \sigma_i^*$,*

$$V_i(h^{\ell-1}, t^\ell, \pi, \sigma_i^*) + X_i(h^{\ell-1}, t^\ell, \pi, \sigma_i^*) \geq V_i(h^{\ell-1}, t^\ell, \pi, \sigma_i') + X_i(h^{\ell-1}, t^\ell, \pi, \sigma_i'), \tag{9}$$

so that agent i maximizes its expected discounted payoff from truthful strategy σ_i when other agents are truthful.

In a within-period *ex post* incentive-compatible (w.p. EPIC) mechanism, truthful revelation of state, transition and reward function is the best-response of an agent, regardless of the current types, so long as all agents follow the equilibrium strategy in current and future play.[15]

For the total expected value to agent j (or flow value), given agent i's strategy σ_i and that other agents are truthful, we have,

$$V_j(h^{\ell-1}, t^\ell, \pi, \sigma_i) = \mathbb{E}_{s^k_\sigma, a^k_\sigma}\left[\sum_{k=\ell}^{K} \gamma^{k-\ell}\, r^k_j(s^k_{\sigma,j}, a^k_\sigma) \,|\, h^{\ell-1}, t^\ell, \pi, \sigma_i\right], \tag{10}$$

where s^k_σ and a^k_σ are the state profiles and actions induced in period k when agent i adopts strategy σ_i. Let $V_{-i}(h^{\ell-1}, t^\ell, \pi, \sigma_i) = \sum_{j\neq i} V_j(h^{\ell-1}, t^\ell, \pi, \sigma_i)$ denote the total flow value to agents $\neq i$. We have $V(h^{\ell-1}, t^\ell, \pi^*, \sigma^*_i) = \sum_{j\in I} V_j(h^{\ell-1}, t^\ell, \pi^*, \sigma^*_i) = V(t^\ell, \pi^*)$ when agent i is truthful.

Lemma 1. *A dynamic mechanism (π, x) in an environment with a fixed, accessible population and dynamic type is efficient and w.p. EPIC, if it is a dynamic Groves mechanism, which requires that:*

i) policy π is efficient with respect to the reported type profile,

ii) each agent i's expected discounted transfer given type profile t^ℓ, strategy σ_i, history $h^{\ell-1}$, and given that agents $\neq i$ follow a truthful strategy in period ℓ and forward, is $V_{-i}(h^{\ell-1}, t^\ell, \pi, \sigma_i) - C_i(t^\ell_{-i})$, where $C_i(t^\ell_{-i})$ is a quantity that is independent of agent i's own strategy in this period and forward.

Proof. Let π^* denote the efficient policy associated with the (true) social planner's problem. Fix period ℓ, and agent i, and suppose agents $\neq i$ are truthful in this period and forward. Assume for contradiction that w.p. EPIC fails. Then by properties (i) and (ii), there must be some strategy $\sigma_i \neq \sigma^*_i$, history $h^{\ell-1}$, and type profile t^ℓ for which,

$$\begin{aligned} V_i(h^{\ell-1}, t^\ell, \pi, \sigma_i) + [V_{-i}(h^{\ell-1}, t^\ell, \pi, \sigma_i) - C_i(t^\ell_{-i})] &> \\ V_i(t^\ell, \pi^*) + [V_{-i}(t^\ell, \pi^*) - C_i(t^\ell_{-i})], \end{aligned} \tag{11}$$

where the terms on the RHS follow from the efficiency of the policy when agent i is truthful. By the principle of one deviation, it is sufficient to consider a strategy σ_i in which agent i submits a misreport in only the current type profile t^ℓ. Let $\widehat{t}^\ell_i = \sigma_i(h^{\ell-1}, t^\ell_i)$ denote this type report. But now, we can construct policy π' from π^* by setting π' equal to π^* in every type profile except for t^ℓ, where we define $\pi'(t^\ell) = \pi(\widehat{t}^\ell_i, t^\ell_{-i})$. We have $V(t^\ell, \pi') = V(h^{\ell-1}, t^\ell, \pi, \sigma_i) > V(t^\ell, \pi^*)$, and a contradiction. □

[15] Truthful revelation in a within-period EPIC mechanism is also a Markov-perfect Bayesian equilibrium, since it is also a best-response for any beliefs an agent may have over the types of other agents.

In a dynamic Groves mechanism, each agent's payoff is aligned with the total flow value to all agents, and w.p. EPIC is achieved because the policy will be efficient when agent i is truthful, given that all other agents are truthful. The second term in the flow transfer, $C_i(t^\ell_{-i})$, can depend on the current type of other agents, and thus on an agent's strategy in previous periods, but must be independent of agent i's strategy in this period and forward.

2.2 The Dynamic-VCG Mechanism

The *dynamic VCG mechanism* [Bergemann and Välimäki, 2008] is w.p. EPIC and efficient, and also *ex post* no deficit in economic environments without positive externalities, such as social choice and one-sided auction problems.

Definition 2 (Dynamic-VCG mechanism)**.** *In the dynamic VCG mechanism, each agent reports (perhaps untruthfully) its type $t^\ell_i = (s^\ell_i, \tau_i, r_i)$ in each period ℓ, and the mechanism selects action $a^\ell = \pi^*(s^\ell)$ for the policy π^* that is efficient given reported types. Each agent i receives a transfer:*

$$x_i(t^\ell) = r_{-i}(s^\ell, a^\ell) + \gamma \mathbb{E}_{t^{\ell+1}} \left[V_{-i}(t^{\ell+1}, \pi^*_{-i}) \mid t^\ell, a^\ell \right] - V_{-i}(t^\ell, \pi^*_{-i}), \qquad (12)$$

in each period ℓ, where π^*_{-i} is the policy that is efficient for the agents without i given reported types.

The expression $\mathbb{E}_{t^{\ell+1}}[V_{-i}(t^{\ell+1}, \pi^*_{-i}) \mid t^\ell, a^\ell]$ is the expected optimal flow value to agents $\neq i$ forward from period $\ell+1$, given the current (reported) type profile *and given that action a^ℓ is taken in period ℓ*, where the expectation is taken over possible next type profiles. The transfer to every agent in each period is its "flow marginal contribution," i.e., the positive impact that i has on the ability for the *other* agents to obtain value in the current time-step. The first two terms reflect the value to the other agents in the current period together with the expected future value these agents would receive under the efficient policy to agents $\neq i$ in future periods. The final term reflects the optimal flow value these other agents could receive from the current period forward, adopt the efficient policy to these other agents in this current period as well.

To establish the incentive compatibility of the dynamic VCG mechanism we show that it belongs to the family of dynamic Groves mechanisms.[16]

Theorem 1. *[Bergemann and Välimäki, 2008] The dynamic VCG mechanism is efficient and w.p. EPIC in an environment with a fixed, accessible population, dynamic type, independent type transitions and private values.*

Proof. Property (i) in Lemma 1 holds for the dynamic VCG mechanism by construction. Let π^* denote this efficient policy. Now fix some agent i, strategy σ_i, history

[16] We provide the proof for a finite number of periods, K. A simple limiting argument extends the result to allow for an infinite number of periods, with the summation of Eq. (15) and (16) yielding an additional term $\gamma^{K+1-\ell}\mathbb{E}_{t^{K+1}_\sigma}[V_{-i}(\hat{s}^{K+1}_\sigma, \pi^*_{-i})]$, which tends to 0 as $K \to \infty$.

$h^{\ell-1}$, and current type profile t^ℓ. Adopt notation t^k_σ and s^k_σ to denote the type and state profiles in period k given strategy σ_i, and $\widehat{t}^k_\sigma = (\widehat{t}^k_{\sigma,i}, t^k_{\sigma,-i})$ and $\widehat{s}^k_\sigma = (\widehat{s}^k_{\sigma,i}, s^k_{\sigma,-i})$ to denote the reported type and state profiles. Distributions on $s^k_\sigma, \widehat{t}^k_\sigma$ and $\widehat{s}^k_\sigma$ are all induced by distributions on t^k_σ. Let a^k_σ denote the action taken in period k.

To establish property (ii) for the flow transfer, observe that the first term in the dynamic VCG transfer provides flow transfer $\mathbb{E}_{t^k_\sigma}[\sum_{k=\ell}^K \gamma^{k-\ell} r_{-i}(\widehat{s}^k_\sigma, \pi^*(\widehat{t}^k_\sigma)] = \mathbb{E}_{t^k_\sigma}[\sum_{k=\ell}^K \gamma^{k-\ell} r_{-i}(s^k_\sigma, \pi^*(\widehat{t}^k_\sigma)] = V_{-i}(h^{\ell-1}, t^\ell_\sigma, \pi^*, \sigma_i) = V_{-i}(h^{\ell-1}, t^\ell, \pi^*, \sigma_i)$, where the first equality holds because of private values and the second because the strategy forward from period ℓ does not affect the current type profile t^ℓ but only its report. The second term in the dynamic VCG transfer provides flow transfer

$$\gamma \; \mathbb{E}_{t^k_\sigma, a^k_\sigma} \left[\sum_{k=\ell}^{K} \gamma^{k-\ell} \; \mathbb{E}_{t'} \left[V_{-i}(t', \pi^*_{-i}) | \widehat{t}^k_\sigma, a^k_\sigma \right] \right], \tag{13}$$

where $t' \in T$ is the next type profile given reported types $\widehat{t}^k_\sigma$ and action a^k_σ. Focusing on the inner expectation, we have

$$\mathbb{E}_{t'}\Big[V_{-i}(t', \pi^*_{-i}) | \widehat{t}^k_\sigma, a^k_\sigma\Big] = \mathbb{E}_{t^{k+1}_\sigma}\Big[V_{-i}((t''_i, t^{k+1}_{\sigma,-i}), \pi^*_{-i})\Big] = \mathbb{E}_{t^{k+1}_\sigma}\Big[V_{-i}(\widehat{t}^{k+1}_\sigma, \pi^*_{-i})\Big], \tag{14}$$

where $t''_i \in T_i$ is any type of agent i, and the first equality holds because of the private values and transition independence. This expectation can then be taken with respect to the actual distribution on types induced by the mechanism, and we can simplify Eq. (13) as,

$$\gamma \; \mathbb{E}_{t^k_\sigma} \left[\sum_{k=\ell}^{K} \gamma^{k-\ell} \; V_{-i}(\widehat{t}^{k+1}_\sigma, \pi^*_{-i}) \right] \tag{15}$$

Finally, the third term in the dynamic VCG transfer provides flow transfer,

$$-V_{-i}(\widehat{t}^\ell_\sigma, \pi^*_{-i}) - \gamma \mathbb{E}_{t^k_\sigma} \left[\sum_{k=\ell}^{K} \gamma^{k-\ell} V_{-i}(\widehat{t}^{k+1}_\sigma, \pi^*_{-i}) \right], \tag{16}$$

where we use $V_{-i}(\widehat{t}^{K+1}, \pi^*_{-i}) = 0$. Summing this with Eq. (15), the expectations cancel and we obtain $-V_{-i}(\widehat{t}^\ell_\sigma, \pi^*_{-i}) = -V_{-i}(t^\ell_\sigma, \pi^*_{-i}) = -V_{-i}(t^\ell, \pi^*_{-i})$, where the first equality holds because of private values and transition independence. This completes the proof, because $-V_{-i}(t^\ell, \pi^*_{-i})$ is independent of agent i's strategy forward form this period. □

The dynamic VCG mechanism is within period *ex post* individual rational, with agent i's flow payoff forward from any type profile t^ℓ just $V(t^\ell, \pi^*) - V_{-i}(t^\ell, \pi^*_{-i})$ and non-negative.[17] An agent will have non-negative payoff in expectation forward from any period, as long as other agents follow the equilibrium play in the future.

[17] We have $V(t, \pi^*) \geq V(t_{-i}, \pi^*_{-i})$, for all type profiles t because of the private values model and since the feasible actions $A(s)$ in a state are independent (conditioned on earlier actions) of the private types of agents. This precludes, for example, environments in which the mere presence of an agent can block the selection of certain alternatives.

We now give a simple auction example to illustrate the mechanism. The center has a single item to allocate, and in each period can either allocate the item to one of the agents or wait. Each agent receives a reward that depends on its current state and the action of the center. In the figures, nodes represent states and transitions by directed edges annotated with a probability ($p = 0.x$) where there is non-determinism and rewards (in **bold**) where there is a non-zero reward. Terminal states are denoted $\rightarrow 1$ or $\rightarrow 2$ to indicate that the item is allocated to agent 1 or 2 respectively.

Example 1. *Consider the simple two agent example portrayed in Figure 1. Assume discount factor $\gamma = 1$ (i.e., no discounting). The efficient policy allocates to agent 1 in state BE, to agent 2 in states $\{CG, CH\}$, and makes no allocation in states $\{AD, BF\}$. Because of the special structure of this domain, the VCG payment to agent i with type s^ℓ is $-V^*_{-i}(t^\ell_{-i})$ when it is allocated the item, because $r_{-i}(s^\ell_{-i}, \pi^*(t^\ell)) + \gamma\mathbb{E}[V_{-i}(\tau(t^\ell_{-i}, \pi^*(t^\ell)))] = 0$ since the other agent cannot get the item. In other cases, the payments are always 0 except when the presence of agent i with a particular type precludes the other agent from being allocated immediately. In such a case, the payment is the cost (if any) of a delay in the decision. To be concrete, consider (true) state BE. If agent 2 reports E ("low value") agent 1 is allocated the item and receives payment -4; agent 2's payoff is 0. If agent 2 reports F ("high value") its payment is -6 (the externality it imposes on agent 1). Continuing, the (true) next state after BE must be CG. Whether agent 2 reports G or H it will be allocated the item (achieving value 4) and its payment will be -2, for a net payoff over the two time-steps of $-6 + 4 - 2 = -4$. The deviation has had a mal effect for agent 2. The up-shot is that agent 2 is best-off truthfully reporting states E and F when they occur, and the center gains the information it needs to determine whether or not to allocate to agent 1 in period 1.*

3 Dynamic Types and Periodic Inaccessibility

We consider now an environment with a fixed population of agents that are *periodically-inaccessible* by the center. This inaccessibility is an endogenous property that is private to an agent and can depend on the actions of the center. To be inaccessible means that an agent cannot send messages to the center, or receive or make payments. On the other hand, an inaccessible agent can continue to undergo type transitions and receive value in a way that depends on actions taken by the center. For example, the center can continue to allocate resources based on the estimated needs of an agent and these can lead to type transitions. Whether or not an agent transitions to an inaccessible state can also depend on both the action taken and the agent's current type.

This model of periodic inaccessibility applies, for instance, to environments in which an agent might periodically lose contact with the center due to faulty or limited communication, or because of bounded attention where an agent needs to periodically attend to other decisions or apply full attention to utilizing a resource just assigned

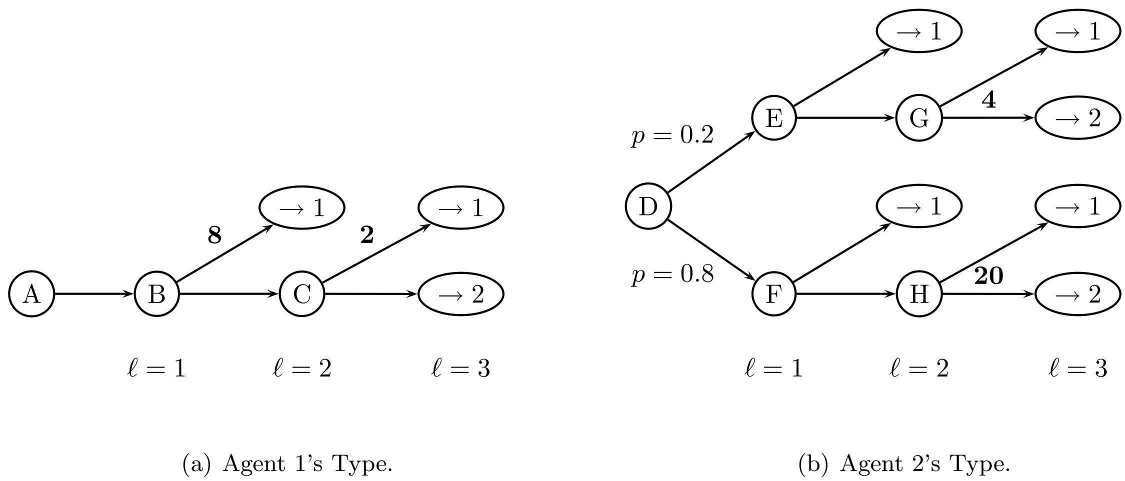


Figure 1: Two-agent, 3 time-step problem with a single item to allocate. Initial joint state is AD. Decisions ({allocate to 1, allocate to 2, don't allocate}) are implicit in the state transitions. Agent 1's type has deterministic transitions, while agent 2's type has non-determinism only in the first period.

by the center. We will see in Section 4 that it also generalizes earlier arrival-departure models of online mechanism design, in which an agent is inaccessible and then accessible and then inaccessible again.

Efficiency is constrained by the communication constraints, which preclude a social planner from knowing the private type of an agent while an agent is inaccessible. Moreover, inaccessible agents have no opportunity to misreport their type to the center, and therefore incentive-compatibility requirements will also be modified so that they need hold only while an agent is accessible.

We introduce the dymamic VCG# mechanism, which is *communication-restricted* w.p. EPIC and efficient. The mechanism is incentive compatible for an accessible agent that is restricted to making the same inferences about inaccessible agent type as the center; i.e., accessible agents are subject to the same communication restrictions with inaccessible agents as the center. The dynamic VCG# mechanism traces beliefs about the type of inaccessible agents, and collects transfers when an agent becomes accessible to "catch-up" for missed transfers. To achieve commuication-restricted w.p. EPIC we require that an inaccessible agent must eventually become accessible again, in order to undergo these catch-up transfers. This requirement can be dropped for the arrival-departure special case because inaccessible agents are never pivotal.

To gain some intuition, we can consider a simple mechanism in which the decision policy is defined in a way that simply ignores inaccessible agents, and selects the action in each period that is efficient as though the population consists of only accessible agents. In addition, suppose the mechanism makes a payment to each accessible agents in every period that is equal to the (reported) value obtained by the other accessible agents in the period. The following example shows that such a simple mechanism is not incentive-compatible for this environment.

Example 2. *Consider again the problem in Figure 1, but now extended with the following periodic-inaccessibility dynamics: with very high probability agent 1 is accessible in all periods; agent 2 is definitely inaccessible in period 0, but will become accessible in periods 1 or 2 or, with negligible probability $\varepsilon > 0$, not at all. In the simple mechanism described above, if agent 2 is not accessible in period 1, then agent 1 should pretend to be inaccessible to avoid receiving the item. It is likely that agent 2 will become accessible in period 2, be allocated the item, and obtain a higher (expected) value than agent 1's value, which agent 1 would then benefit from via a payment from the center.*

To extend the dynamic VCG mechanism to this environment, an efficient policy must consider the impact of actions on both inaccessible and accessible agents. For example, not only might inaccessible agents have value for the allocation of resources while being unable to communicate, or make transfers, with the center, but inaccessible agents might also have high value for resources when they are again accessible and thus the center should retain the option to allocate to such an agent.

Consider first the view point of a social planner and put aside incentive considerations. To model the inaccessibility of agents, define the *planner's belief state* (or simply *belief state*) for agent i, $\omega_i^\ell \in \Omega_i = \Delta(S_i)$, to define a probability distribution on the state $s_i^\ell \in S_i$ of agent i in period ℓ. For an agent that is accessible in period ℓ, the belief state ω_i^ℓ assigns probability 1 to the agent's state s_i^ℓ. For an inaccessible agent, the belief state is derived by using Bayes rule from i's state in the last period it was accessible and the sequence of actions forward from that period.[18] We will assume that every agent is always accessible at time 0. This can be relaxed by assuming a prior on agent types.

Let $\omega^\ell = (s_0^\ell, \omega_1^\ell, \ldots, \omega_n^\ell) \in \Omega = S_0 \times \Omega_1 \times \ldots \Omega_n$, denote the belief state profile, where we again include here state $s_0^\ell \in S_0$ for the center to allow the set of feasible actions to depend on past actions. The reward and transition functions can be extended to this setting in the natural way. Let $\tilde{r}_i(\omega_i^\ell, a) = \mathbb{E}_{s_i^\ell}\left[r_i(s_i^\ell, a) \,|\, \omega_i^\ell\right]$ denote the expected reward to an agent given the distribution on possible states, with $\tilde{r}(\omega^\ell, a) = \sum_i \tilde{r}_i(\omega_i^\ell, a)$. Let $\tilde{\tau}_i : \Omega_i \times A \to \Omega_i$ extend the stochastic transition function from states to belief states, so that $\tilde{\tau}_i(\omega_i^\ell, a)$ simply transitions to the belief state corresponding to the agent's state $s_i^{\ell+1}$ as induced by $\tau_i(s_i^\ell, a)$ when the agent is accessible, and is otherwise defined according to Bayes rule when agent i remains inaccessible. Altogether, this defines a *belief type*, $bt_i = (\omega_i, \tau_i, r_i) \in BT_i$, with joint belief type space, $BT = (T_0 \times BT_1 \times \ldots \times BT_n)$.

Note that $\tilde{r}_i(\omega_i^l, a)$ and $\tilde{\tau}_i(\omega_i^l, a)$ retain conditional independence of the belief states of other agents given the action of the center. Let $\tilde{\tau} : \Omega \times A \to \Omega$, with $\tilde{\tau} =$

[18] For example, consider agent 2 from Figure 1. If the agent is inaccessible in period 1, then ω_2^1 assigns probabilities 0.2 and 0.8 to states E and F, respectively.

$(\tau_0, \tilde{\tau}_1, \ldots, \tilde{\tau}_n)$ and $\tau_0 : s_0 \times A \to s_0$, define the joint transition function. The reward function and transition functions are also extended to belief types in the natural way. Given a decision policy, $\tilde{\pi} : BT \to A$, then

$$\tilde{V}(bt^\ell, \tilde{\pi}) = \mathbb{E}_{\omega^k}\left[\sum_{k=\ell}^{K} \gamma^{k-\ell}\ \tilde{r}(\omega^k, \tilde{\pi}(bt^k)) \mid bt^\ell, \tilde{\pi}\right], \tag{17}$$

denotes the expected total discounted reward given policy $\tilde{\pi}$ and belief type profile bt^ℓ, with analogous variants for $\tilde{V}_i$ and $\tilde{V}_{-i}$. The efficient policy $\tilde{\pi}^*$ solves $\tilde{\pi}^* \in \arg\max_{\tilde{\pi}\in\tilde{\Pi}} \tilde{V}(bt^\ell, \tilde{\pi})$, $\forall bt^\ell \in BT$, where $\tilde{\Pi}$ denotes the space of feasible policies. The efficient policy tracks the distribution on possible states of inaccessible agents and selects actions to maximize expected, discounted value.[19] Let $\tilde{\pi}^*_{-i}$ denote the policy that is efficient for agents $j \neq i$.

3.1 Extending the Dynamic-VCG Mechanism

A dynamic mechanism $M = (\tilde{\pi}, \tilde{x})$ in this environment is defined by a decision policy $\tilde{\pi} : BT \to A$ and a transfer policy $\tilde{x} : BT \to \mathcal{R}^n$ (with $\tilde{x}_i(bt^\ell) = 0$ when agent i is not accessible.) Given belief type profile $bt^\ell \in BT$ in period ℓ, the mechanism selects action $a^\ell = \tilde{\pi}(bt^\ell)$ and makes transfers $\tilde{x}_i(bt^\ell) \in \mathcal{R}$ to each agent i that is accessible.

For an accessible agent, strategy $\sigma_i(h^{\ell-1}, t_i) \in \{T_i \cup \phi\}$ denotes its claim about type, where ϕ indicates no report is made and the agent pretends to be inaccessible and $h^{\ell-1} \in \mathcal{H}^{\ell-1}$ is the history of reports up to and including period ℓ. Let $\mathcal{A}(\ell)$ and $\mathcal{NA}(\ell)$ denote the accessible and inaccessible agents in period ℓ. Only accessible agents, $i \in A$, can submit $\sigma_i(h^{\ell-1}, t_i) \neq \phi$. Let $\breve{bt}^\ell$ denote a *partially-truthful belief type profile*, which assigns probability 1 to the true type of accessible agents, with the belief type for each inaccessible agent determined by applying Bayesian updates given the most recent type report received from the agent and subsequent actions by the center.[20]

Fix an accessible agent i in period ℓ, and suppose agents $\neq i$ are truthful going forward from period ℓ. Let a^k_σ denote the action selected by the mechanism in period k given strategy σ_i by agent i. Let ω^k_σ, bt^k_σ denote the belief state and belief type given strategy σ_i, where this is the true belief state and type except for agents that are inaccessible in period ℓ and still inaccessible, wherein it is the belief state

[19] Given that the accessibility of an agent can also depend on actions, such a policy will inherently factor in considerations about the value-of-information by taking an action that will make an agent accessible and thus collapse the belief type for such an agent to its actual type. For a computational approach to solve these belief-state MDP problems, see the survey of algorithms for Partially Observable MDPs in Kaelbling [1996].

[20] Given that all agents are assumed to be accessible in period 0, we can allow the mechanism to further constrain agent strategies so that every agent must report a type in the first period. This ensures that the belief type profile maintained by the mechanism is always well-defined. In Section 4 we can dispose of this requirement because the arrival process, known to the center, plays the role of making sure that belief types are well defined.

and type consistent with $\breve{bt}^{\ell}$. The true expected discounted value to accessible agent i with the same beliefs as the center about inaccessible agents and knowledge of the types of accessible agents (i.e., its *subjective flow value*), given strategy σ_i and that other agents are truthful going forward from period ℓ, is

$$\tilde{V}_i(h^{\ell-1}, \breve{bt}^{\ell}, \tilde{\pi}, \sigma_i) = \mathbb{E}_{\omega_\sigma^k, a_\sigma^k}\left[\sum_{k=\ell}^{K} \gamma^{k-\ell}\ \tilde{r}_i^k(\omega_{\sigma,i}^k, a_\sigma^k) \,|\, h^{\ell-1}, \breve{bt}^{\ell}, \tilde{\pi}, \sigma_i\right], \tag{18}$$

where $\omega_\sigma^{k+1} = \tilde{\tau}(\omega_\sigma^k, a_\sigma^k)$ and $a_\sigma^k = \tilde{\pi}(\Gamma_i^k(\sigma_i(h^{k-1}, t_{\sigma,i}^k)), bt_{\sigma,-i}^k)$, with $\Gamma_i^k : \{T \times \phi\} \to BT_i$ returning a belief type with probability 1 assigned to $\widehat{t}_i = \sigma_i(h^{k-1}, t_{\sigma,i}^k)$ if $\widehat{t}_i \neq \phi$ and with Bayesian updates on the belief type in period $k-1$ given action a_σ^{k-1} otherwise. History $h^k = (h^{k-1}, (\sigma_i(h^{k-1}, t_{\sigma,i}^k), bt_{\sigma,-i}^k))$. We can similarly define the expected discounted transfer to agent i (i.e., its *subjective flow transfer*), as $\tilde{X}_i(h^{\ell-1}, \breve{bt}^{\ell}, \tilde{\pi}, \sigma_i)$. Let σ_i^* denote the truthful strategy in this environment, i.e. reporting the true type whenever an agent is accessible.

Definition 3 (communication-restricted w.p. ex post incentive compatible)**.** *Dynamic mechanism $(\tilde{\pi}, \tilde{x})$ in an environment with periodic inaccessibility is communication-restricted within-period ex post incentive-compatible if, for all times ℓ, for any accessible agent $i \in \mathcal{A}(\ell)$, for any partially-truthful belief-type profile $\breve{bt}^{\ell}$ induced by the true type of accessible agents and most recent report of inaccessible agents, for any history $h^{\ell-1}$, and for all $\sigma_i' \neq \sigma_i^*$,*

$$\tilde{V}_i(h^{\ell-1}, \breve{bt}^{\ell}, \tilde{\pi}, \sigma_i^*) + \tilde{X}_i(h^{\ell-1}, \breve{bt}^{\ell}, \tilde{\pi}, \sigma_i^*) \geq \tilde{V}_i(h^{\ell-1}, \breve{bt}^{\ell}, \tilde{\pi}, \sigma_i') + \tilde{X}_i(h^{\ell-1}, \breve{bt}^{\ell}, \tilde{\pi}, \sigma_i'), \tag{19}$$

so that agent i maximizes its subjective flow payoff given $\breve{bt}^{\ell}$ and truthful strategy σ_i^*, given that all other agents are truthful going forward.

A communication-restricted within-period EPIC (c.r.w.p. EPIC) dynamic mechanism places no incentive-compatibility constraints for agents that are inaccessible because they cannot make any reports. For accessible agents, then regardless of the current type of accessible agents, an agent's best-response is to be truthful given that other agents are also truthful going forward. This equilibrium concept is "communication restricted" in the sense that an accessible agent is assumed to be restricted to the same communication constraints as the center: it cannot know the true state of an inaccessible agent.[21]

[21] This is not a w.p. ex post equilibrium with respect to the belief an agent would hold about the type of an inaccessible agent if the agent had known the true type of the accessible agent in the last period in which it was accessible. Thus, this is a slight weakening of w.p. EPIC in the case of inaccessible agents that played off equilibrium in their last accessible period. But it is stronger than a Markov-perfect equilibrium because of its ex post characteristic with respect to the type of accessible agents and reduces to the standard w.p. EPIC definition without periodically-inaccessible agents.

Let $\tilde{V}_j(h^{\ell-1}, \breve{bt}^{\ell}, \tilde{\pi}, \sigma_i)$ denote the subjective flow value to agent j given partially-truthful belief type profile $\breve{bt}^{\ell}$, and given strategy σ_i from agent i, with other agents reporting true types when accessible; this is the subjective view of agent i of agent j's expected discounted value, given that i can only adopt belief type $\breve{bt}_j^{\ell}$ for an agent j that is currently inaccessible. Let $\tilde{V}_{-i}(h^{\ell-1}, \breve{bt}^{\ell}, \tilde{\pi}, \sigma_i) = \sum_{j \neq i} \tilde{V}_j(h^{\ell-1}, \breve{bt}^{\ell}, \tilde{\pi}, \sigma_i)$, and $\tilde{V}(h^{\ell-1}, \breve{bt}^{\ell}, \tilde{\pi}, \sigma_i) = \sum_{j \in I} \tilde{V}_j(h^{\ell-1}, \breve{bt}^{\ell}, \tilde{\pi}, \sigma_i)$.

We can now state an analogue of Lemma 1, characterizing the set of dynamic Groves mechanisms for this environment.

Lemma 2. *A dynamic mechanism $(\tilde{\pi}, \tilde{x})$ in an environment with a fixed population of periodically-inaccessible agents is efficient and c.r.w.p. EPIC, if it is a dynamic Groves mechanism, which requires that:*

i) the policy $\tilde{\pi}$ is efficient with respect to reported belief types,

ii) each accessible agent i's expected discounted transfer given any partially-truthful belief type profile, $\breve{bt}^{\ell}$, induced by the true type of accessible agents and most recent report of inaccessible agents, for any history $h^{\ell-1}$ and any σ_i, and given that agents $\neq i$ follow a truthful strategy in this period and forward, is $\tilde{V}_{-i}(h^{\ell-1}, \breve{bt}^{\ell}, \tilde{\pi}, \sigma_i) - C_i(\breve{bt}_{-i}^{\ell})$, where $C_i(\breve{bt}_{-i}^{\ell})$ is a quantity independent of agent i's strategy in this period and forward.

Proof. Let $\tilde{\pi}^*$ denote the efficient policy in the MDP associated with the (true) social planner's problem on belief states. Fix period ℓ, and accessible agent i, and suppose agents $\neq i$ are truthful in this period and forward. Assume for contradiction that c.r.w.p. EPIC fails. Then by properties (i) and (ii), there must be some strategy $\sigma_i \neq \sigma_i^*$, history $h^{\ell-1}$, and partially-truthful belief type profile $\breve{bt}^{\ell}$, for which,

$$\tilde{V}(h^{\ell-1}, \breve{bt}^{\ell}, \tilde{\pi}, \sigma_i) - C_i(\breve{bt}_{-i}^{\ell}) > \tilde{V}(bt^{\ell}, \tilde{\pi}^*) - C_i(\breve{bt}_{-i}^{\ell}), \tag{20}$$

where the first term on the right-hand side follows from the efficiency of the policy, given that the other agents are truthful. By the principle of one deviation, we can consider a strategy σ_i in which agent i misreports its type in only the current belief type profile, $\breve{bt}^{\ell}$. Let $\widehat{bt}_i^{\ell} = \Gamma_i^{\ell}(\sigma_i(h^{\ell-1}, t_i^{\ell}))$ denote this reported belief type; i.e., assigning probability 1 to its reported type if $\sigma_i(h^{\ell-1}, t_i^{\ell}) \neq \phi$ and using Bayesian updates on the belief type from period $\ell - 1$ otherwise. But now, we can construct policy $\tilde{\pi}'$ from $\tilde{\pi}^*$ by setting $\tilde{\pi}'$ equal to $\tilde{\pi}^*$ in every belief type profile except for bt^{ℓ}, where we define $\tilde{\pi}'(bt^{\ell}) = \tilde{\pi}(\widehat{bt}_i^{\ell}, \breve{bt}_{-i}^{\ell})$. We have $\tilde{V}(bt^{\ell}, \tilde{\pi}') = \tilde{V}(h^{\ell-1}, \breve{bt}^{\ell}, \tilde{\pi}, \sigma_i) > \tilde{V}(bt^{\ell}, \tilde{\pi}^*)$, and a contradiction with the efficiency of $\tilde{\pi}^*$. ∎

The appropriate dynamic VCG mechanism in this environment adopts belief types in place of types, and monitors the payments that an inaccessible agent would make if it could be charged while inaccessible, collecting a discount-adjusted equivalent amount when an agent becomes accessible:

Definition 4 (Dynamic-VCG# for periodic inaccessibility)**.** *In each period ℓ, each accessible agent can report (perhaps untruthfully) type $t^\ell = (s_i^\ell, \tau_i, r_i)$, and the mechanism adopts into its current belief state ω^ℓ such reports along with an updated belief state for agents for which no new report is received, and selects action $a^\ell = \tilde{\pi}^*(\omega^\ell)$ where $\tilde{\pi}^*$ is the policy that is efficient given the belief-type profile induced by reports. Each agent $i \in I$ that makes a report in period ℓ, receives a transfer:*

$$\tilde{x}_i^{\#}(bt^\ell) = \sum_{k=\ell-\delta(\ell)}^{\ell} \frac{\tilde{x}_i(bt^k)}{\gamma^{\ell-k}}, \text{ where} \tag{21}$$

$$\tilde{x}_i(bt^k) = \tilde{r}_{-i}(\omega^k, a^k) + \gamma \; \mathbb{E}_{bt'}\Big[\tilde{V}_{-i}(bt', \tilde{\pi}^*_{-i}) \mid bt^k, a^k\Big] - \tilde{V}_{-i}(bt^k, \tilde{\pi}^*_{-i}), \tag{22}$$

where $\tilde{\pi}^*_{-i}$ is the policy that is efficient to agents $j \neq i$ given the belief-type profile induced by their reports, a^k is the action selected by the mechanism in subsequent period k, and $\delta(\ell) \geq 0$ is the number of successive periods prior to period ℓ that agent i reported inaccessibility.

The term $\mathbb{E}_{bt'}[\tilde{V}_{-i}(bt', \tilde{\pi}^*_{-i}) | bt^k, a^k]$ is the expected optimal flow value to agents $\neq i$ forward from period $k+1$, under the policy efficient for those agents, given the current (perhaps untruthful) belief type bt^k and given action a^k in period k. The transfer policy makes the mechanism payoff equivalent for accessible agents to a mechanism in which transfers $\tilde{x}_i(bt^\ell)$ can be made directly in every period, irrespective of whether or not an agent is accessible. In order to establish c.r.w.p. incentive compatibility we require the following assumption:

Assumption 1. *An agent that is inaccessible will not remain so for all future periods and must report itself as accessible in some future period.*

Thus, this is a model of periodically but not persistently inaccessible agents.[22] With this assumption we can establish the incentive-compatibility of this mechanism:

Theorem 2. *The dynamic VCG# mechanism is efficient and c.r.w.p. EPIC for a fixed population of periodically-inaccessible agents under Assumption 1 and for private values and independent type transitions.*

Proof. Property (i) in Lemma 2 holds for the dynamic VCG# mechanism by construction. Now fix some accessible agent i, strategy σ_i, history $h^{\ell-1}$, partially-truthful belief type profile $\breve{bt}^\ell$ induced by the true type of accessible agents and

[22]To justify the further requirement that an agent that eventually becomes accessible will also eventually report itself as such (and thus be subject to transfers), we could in addition require an agent to post a bond that is only returned over time, but only if the agent reports itself as accessible. We shall not model this aspect explicitly, however.

most recent report of inaccessible agents. In establishing property (ii) of Lemma 2, consider first a simplified mechanism in which transfer $\tilde{x}_i(\breve{bt}^{\ell})$ can be made directly, in every period. By making notational substitution into the proof of Theorem 1, and noting that private values and independent type transitions are retained, one can directly establish that the flow transfer to agent i would then be $\tilde{V}_{-i}(h^{\ell-1}, \breve{bt}^{\ell}, \tilde{\pi}^*, \sigma_i) - \tilde{V}_{-i}(\breve{bt}^{\ell}, \tilde{\pi}^*_{-i})$, where $\tilde{\pi}^*$ and $\tilde{\pi}^*_{-i}$ are the efficient policies with all agents and without agent i respectively, and the second term is independent of agent i's strategy in period ℓ and forward.

Left to show is that the expected discounted transfer to an accessible agent i is equivalent, for any strategy of agent i, to that in this intermediate mechanism in which the transfers are made in every period. We need to establish that,

$$\mathbb{E}_{bt^k_\sigma}\Big[\sum_{k=\ell}^{K} \gamma^{k-\ell}\, \tilde{x}_i(\Gamma^k_i(\sigma^k_i(h^{k-1}, t^k_{\sigma,i}), bt^k_{\sigma,-i}))\,\Big|\, h^{\ell-1}, \breve{bt}^{\ell}, \sigma_i, \tilde{\pi}^*\Big] =$$

$$\mathbb{E}_{bt^k_\sigma}\Big[\sum_{\substack{k=\ell \\ k\in H(\sigma_i, bt^k_{\sigma,i})}}^{K} \gamma^{k-\ell}\, x^{\#}_i(\Gamma^k_i(\sigma_i(h^{k-1}, t^k_{\sigma,i}), bt^k_{\sigma,-i}))\,\Big|\, h^{\ell-1}, \breve{bt}^{\ell}, \sigma_i, \tilde{\pi}^*\Big], \tag{23}$$

where the expectation is taken with respect to the distribution on belief types bt^k_σ induced by agent i's strategy. The summation on the RHS restricts to time-periods $k \in H(\sigma_i, bt^k_{\sigma,i})$ in which agent i reports a non-null type to the center (and thus receives a transfer). To establish this equivalence, we show the stronger property of equivalence for any *realization sequence* of belief types $\boldsymbol{bt}^{\ell}_{\sigma}, \ldots, \boldsymbol{bt}^{K}_{\sigma}$ given strategy σ_i. For such a realized sequence of belief types, let $F(\sigma_i, \boldsymbol{bt}^k_{\sigma,i})$ isolate the time periods in which agent i *first* becomes accessible again, after one or more periods of (reported) inaccessibility. Expanding the definition of transfer $x^{\#}_i$, we have for the RHS of Eq. (23):

$$\sum_{\substack{k=\ell \\ k\in H(\sigma_i, \boldsymbol{bt}^k_{\sigma,i})\setminus F(\sigma_i, \boldsymbol{bt}^k_{\sigma,i})}}^{K} \gamma^{k-\ell}\, \tilde{x}_i(\Gamma^k_i(\sigma_i(\boldsymbol{t}^k_{\sigma,i}), \boldsymbol{bt}^k_{-i}) \;+ \sum_{\substack{k'=\ell \\ k'\in F(\sigma_i, \boldsymbol{bt}^k_{\sigma,i})}}^{K} \gamma^{k'-\ell} \sum_{k=k'-\delta(k)}^{k'} \frac{\tilde{x}_i(\Gamma^k_i(\sigma_i(\boldsymbol{t}^k_{\sigma,i}), \boldsymbol{bt}^k_{-i})}{\gamma^{k'-k}},$$

$$= \sum_{k=\ell}^{K} \gamma^{k-\ell}\, \tilde{x}_i(\Gamma^k_i(\sigma_i(\boldsymbol{t}^k_{\sigma,i}), \boldsymbol{bt}^k_{\sigma,-i}), \tag{24}$$

where $\delta(k) > 0$ is the number of contiguous periods the agent was reported to be inaccessible before becoming accessible again given this realization sequence, and the periods are grouped into those in which the agent is accessible but not first accessible, and those in which the agent is first accessible. To keep the notation simple we drop the dependence of strategy σ_i on history. Eq. (24) follows from simple algebra, coupled with Assumption 1. This completes the proof. $\square$

Example 3. *Consider again the problem in Figure 1, in which there is a single item to allocate, and no discount factor. Assume agent 1 is truthful. First consider* *what would happen in a mechanism in which transfers can be avoided by pretending to be inaccessible. In this case, if agent 2 is truthful and accessible in period 1, and in state E, then agent 1 would be allocated the item in period 1 and agent 2's payoff would be zero. But by pretending to be inaccessible in period 1, the policy will delay making an allocation until period 2 because* $8 < (0.2)4 + (0.8)20 = 16.8$ *(ignoring ε). Both agents' payments in period 1 will be zero (agent 2's because it is inaccessible). Agent 2 can then report state G in period 2, receive the item, and make a payment of -2 for net payoff $4-2 = 2$. But under the dynamic VCG# mechanism the manipulation goes away. Agent 2 can no longer benefit from pretending to be inaccessible, because it will face a payment of $-6-2$ if it makes itself accessible in period 2.*

4 A Dynamic Population: Arrivals and Departures

We now specialize the environment with periodic inaccessibility and use it to model a problem in which there are dynamic agent arrivals and departures. As a motivating example, imagine a family that arrives to New York and is interested in buying theater tickets for multiple shows during a week; they may also update their value for future tickets in response to attending earlier shows or observing other unpredictable events such as the weather. Once the family leaves town at the end of the week, the actions of the mechanism are no longer relevant and have no ongoing value.

This is a special case of the periodic-inaccessibility model, in which each agent transitions through a single cycle from inaccessible, to accessible and back to inaccessible, and without any value for actions while inaccessible. It will be possible to drop the use of catch-up transfers, and dispense with Assumption 1 (that agents will become accessible again in the future) because inaccessible agents are not pivotal in an arrival-departure model. With this, we recover a suitable generalization of the dynamic VCG mechanism to this environment, and identify the importance of a conditional independence property on the arrival process.

Upon a further restriction to static types, modeling dynamic auction problems, we are in an environment in which the efficient policy can be implemented by the *online VCG* mechanism [Parkes and Singh, 2003]. We show that the online VCG mechanism is payoff equivalent to the dynamic VCG mechanism in this environment upon an agent's arrival. This is sufficient to achieve incentive compatibility, which only imposes constraints on the payoff provided by a mechanism in an agent's arrival period in this environment. Moving to static types also breaks the uniqueness of the dynamic VCG mechanism, allowing a multiplicity of w.p. EPIC and efficient mechanisms that satisfy the efficient exit condition of Bergemann and Välimäki [2008]. In particular, the flow payoffs are different under the online VCG and dynamic VCG mechanisms, and the online VCG mechanism satisfies *ex post* participation constraints where the dynamic VCG provides only within-period *ex post* participation. If agents are impatient, with decisions made immediately upon arrival, then the two mechanisms are identical.

4.1 The Environment: Modeling a Stochastic Arrival Process

The agent arrival process is characterized by a stochastic process $z^0, z^1, \ldots \in Z$, with the next arrival state z^ℓ defined by probability function,

$$g(z^\ell | z^0, \ldots, z^{\ell-1}, a^0, \ldots, a^{\ell-1}) \in [0,1], \tag{25}$$

such that $\sum_{z^\ell \in Z^\ell} g(z^\ell | z^0, \ldots, z^{\ell-1}, a^0, \ldots, a^{\ell-1}) = 1$. Given state z^ℓ, this defines a corresponding multi-set, $T(z^\ell)$, defining the types of agents that arrive in period ℓ. For such an agent i, then we refer to type $t_i^\ell = (s_i^\ell, \tau_i, r_i) \in T(z^\ell)$ as its *arrival type*. Upon arrival, an agent can then undergo stochastic type transitions and receive rewards as is familiar from the fixed population models.

The social planner's problem can again be formulated as an MDP, with joint state profile $s^\ell = (s_*^\ell, s_0^\ell, \{s_i^\ell | i \in \bigcup_{t' \leq t} I(s_*^{t'})\}) \in S$, and where state $s_*^\ell \in S_*$ is introduced to model the arrival process, with stochastic transition function $\tau_* : S_* \times A \rightarrow S_*$.[23] Let $I(s_*^\ell) \subseteq I = \{1, 2, \ldots, \infty\}$ define the set of agents that arrive in period ℓ, given state s_*^ℓ, with an agent $i \in I(s_*^\ell)$ that arrives in period ℓ receiving arrival type $(s_i^\ell, \tau_i, r_i) = T_i(s_*^\ell)$. State $s_0^\ell \in S_0$ continues to model the decision-making problem of the center; e.g., it allows for the dependence of feasible actions on previous actions to be modeled. Each agent may at some point transition to being inaccessible (to model its departure), from which point we require that its reward will be zero for all actions for all future time.

Taken together, agent types and the arrival process define a joint MDP and we can adopt $\pi^* : T \rightarrow A$ and $V(t^\ell, \pi^*)$ to denote the efficient policy and its total expected discounted value respectively, with transition and reward functions extended to types in the natural way.

As a problem of dynamic mechanism design, agents are free to misreport their types, including delaying a report of an agent's arrival or announcing an early departure. While the arrival process, modeled through transition function τ_*, is known to the center, the state $s_*^\ell \in S_*$, which defines the arrival types in period ℓ, is private to agents and unobservable by the center.

The problem can be cast within the fixed population, periodically-inaccessible model of Section 3 by allowing for a potentially infinite set of agents $I = \{1, 2, \ldots, \infty\}$, all initially inaccessible, and with the arrival process defining a belief type (known to the center) for an agent until its arrival. Thus, this problem fits within the framework of problems solvable via the dynamic VCG# mechanism. Moreover, the special arrival-departure structure of this problem simplifies the mechanism description and also its analysis:

- No catch-up transfer is required upon an agent's arrival because a dynamic Groves mechanism only needs to align the payoff of an agent that is accessible,

[23] The dependence of $g(z^\ell | z^0, \ldots, z^{\ell-1}, a^0, \ldots, a^{\ell-1})$ on the history of prior arrivals and actions can be captured within Markovian dynamics by defining an appropriate state space S_*.

and agents are never accessible before they arrive.

- No catch-up transfer is required after an agent's departure because a departed agent is never pivotal (by definition, since a departed agent is required to have zero value for all subsequent actions by the center), and thus the transfers tracked by the dynamic VCG# mechanism for an agent that has departed are always exactly zero.

Moreover, we will not need the Assumption 1 because an agent's catch-up transfer is always zero after departure in any case and thus no subsequent transfer is needed to align incentives. These observations in hand, the dynamic VCG# mechanism reduces to the dynamic VCG mechanism (suitably generalized to a population with a known dynamic arrival process):

Definition 5 (Dynamic-VCG mechanism for a dynamic population)**.** *Each agent that is present in period ℓ can report (perhaps untruthfully) its type $t_i^\ell = (s_i^\ell, \tau_i, r_i)$. The mechanism tracks the state of the arrival process based on reports, and selects action $a^\ell = \pi^*(s^\ell)$ for the policy π^* that is efficient given reported types and the arrival process. Each agent i that makes a report in period ℓ receives a transfer:*

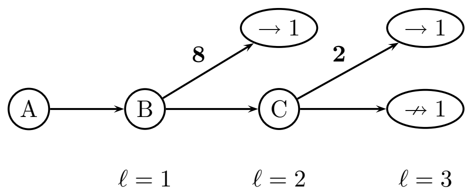
$$x_i(t^\ell) = r_{-i}(s^\ell, a^\ell) + \gamma \mathbb{E}_{t'}\Big[V_{-i}(t', \pi^*_{-i}) \mid t^\ell, a^\ell\Big] - V_{-i}(t^\ell, \pi^*_{-i}), \tag{26}$$

where π^*_{-i} is the policy that is efficient for the agents without i given the reported type profile and the arrival process.

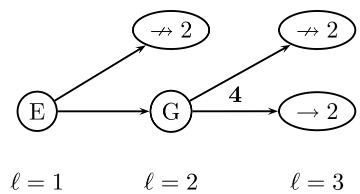
The term $\mathbb{E}_{t'}[V_{-i}(t', \pi^*_{-i}) \mid t^\ell, a^\ell]$ is the expected flow value to agents except i, under the efficient policy to these agents, forward from the next period given action a^ℓ in period ℓ and the type profile t^l. But without a restriction on the arrival process, the dynamic VCG mechanism is not incentive-compatible with dynamic populations, as illustrated in the following example.

Example 4. *Figure 2 depicts a variation on the previous problem, in which there are now four possible arrival types. Define an arrival process such that a single agent of type 1 always arrives in period 0, while at most one agent among types 2, 3, or 4 can ever arrive, and it is very likely that a type 4 agent will arrive in period 2. If an agent of type 2 arrives in period 1, then it will hide and claim to be inaccessible. The efficient policy will wait to allocate the resource because it is likely that a type 4 agent will arrive in the next period. In period 2, the type 2 agent can truthfully report state G (claiming to be a type 3 agent that just arrived), and will be allocated the item and have to make a payment of 2. This causes an efficiency loss because the item should have been allocated to agent 1 in period 1.*

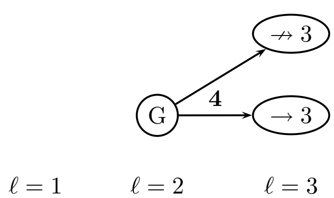
Lemma 2 continues to hold in this environment, but the earlier proof that the dynamic VCG# mechanism is a dynamic Groves mechanism fails because the requirement of independent type transitions is not satisfied by a general arrival process g in



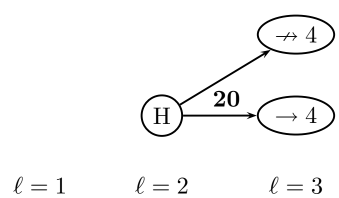
(a) Type 1.



(b) Type 2.



(c) Type 3.



(d) Type 4.

Figure 2: An example with a dynamic agent population. One type 1 agent arrives in period 1, and at most one agent of types 2, 3, or 4 will arrive (in periods 2 or 3, depending on their type).

which new agent arrival types may depend on previous arrival types. When modeling such an arrival process within the periodic-inaccessibility model, this translates into a dependency in which the belief state transition of an agent that is currently inaccessible and has not arrived may depend on the state (and thus arrival type) of another agent.

We require the following conditional independence property in the arrival process.

Assumption 2 (CIA property)**.** *An arrival process is satisfies the conditional-independence on arrivals (CIA) property when the arrival process,*

$$g(z^\ell | z^0, \ldots, z^{\ell-1}, a^0, \ldots, a^{\ell-1}) = g(z^\ell | a^0, \ldots, a^{\ell-1}), \tag{27}$$

so that future agent arrival types are independent of past arrivals, conditioned on the history of actions.

With this in place, the local type dynamics of each inaccessible agent (modeling the arrival process) are independent of the states of other agents and the incentive compatiblity of the dynamic VCG# mechanism is recovered. Under this CIA property, future arrivals can depend on the actions of the center; e.g., types may be "censored" because all good resources are already sold or affected via marketing decisions within the remit of the mechanism. The assumption allows, for example, an environment in which whether or not a high type arrives in period $\ell + 1$ depends on whether or not an item is allocated in the current period. But it does not allow environments in which exactly one high type will arrive, and therefore if a high type does not arrive in period ℓ then it is more likely that a high type will arrive in period $\ell + 1$, and vice versa.[24]

Theorem 3. *The dynamic VCG mechanism for a dynamic population with arrivals and departures, and allowing dynamic types, is efficient and w.p. EPIC given private values, independent type transitions, the CIA property, and a center with a correct model of agent arrivals.*

Proof. Interpreting the requirements of Lemma 2 in this environment, we need (i) that the policy π followed by the mechanism is efficient with respect to reported types and the center's model of the arrival process, and (ii) the flow transfer *to an accessible agent i* forward from any type profile t^ℓ, for any history $h^{\ell-1}$ and any σ_i, and given that agents $\neq i$ follow a truthful strategy in this period and forward and the center's model of the arrival process is correct, is $V_{-i}(h^{\ell-1}, t^\ell, \pi, \sigma_i) - C_i(t^\ell_{-i})$, where $C_i(t^\ell_{-i})$ is a quantity independent of agent i's strategy in this period and forward. Property (i) holds by construction. To establish property (ii), consider a modified dynamic VCG mechanism in which the transfer in Eq. (26) is collected in every period. Given this, the flow transfer to an agent forward from some period ℓ in which it is accessible is $V_{-i}(h^{\ell-1}, t^\ell, \pi^*, \sigma_i) - V_{-i}(t^\ell, \pi^*_{-i})$, where π^* and

[24] Example 4 is precluded by the assumption because whether an agent of types 3 or 4 arrives must now be independent of the arrival of an agent of type 2.

π^*_{-i} are the efficient policy for all agents and without agent i respectively, by notational substitution into the proof of Theorem 1, and observing that the independent type transitions property continues to hold because of the CIA property. This satisfies property (ii), and thus would be a dynamic Groves mechanism, with term $V_{-i}(t^\ell, \pi^*_{-i})$ agent-independent because of the CIA property. Now, the transfer in Eq. (26) would be zero in any period before an agent reports its arrival because no agent is pivotal before it reports its arrival because of the CIA property. In addition, the transfer in Eq. (26) would be zero in any period after an agent's reported departure because agents have no value for any actions in any period after departure. This completes the proof, because the flow transfer forward from an agent's arrival by collecting transfers only during an agent's reported arrival-departure interval is the same as when transfers can be collected in every period. Note that c.r.w.p. EPIC is equivalent to w.p. EPIC in this environment, because reports by agents that are now inaccessible are irrelevant. Such an agent must have now reported its departure, in which case it is no longer pivotal and its last report before becoming inaccessible is irrelevant. $\square$

The w.p. EPIC property provides that truthful reporting is optimal for an agent, whatever the current type profile and whatever the arrival process, as long as other agents are truthful forward from the current period and the center's model of the arrival process is correct. It must be common knowledge that the center has a correct model of the arrival process, although the arrival process itself does not need to be common knowledge.

4.2 Specializing to Dynamic Arrivals but Static Types

An interesting special case is to a dynamic population of agents, each of which has a *fixed* valuation for different sequences of actions; i.e., agents arrive with a private, but known valuation function on sequences of actions and no uncertainty. This models, for example, *dynamic auction problems* in which an agent can describe upon arrival a valuation function over any possible sequence of allocation decisions. This setting includes combinatorial auction problems, with substitutes and complements valuations on sets of distinct items.

In the language of this paper, we say that each agent has a *static type*; i.e., an agent is able to completely describe its private information with a single report upon arrival into the environment. Given the MDP preference model, this static type requirement corresponds to agents with a transition function $\tau_i : S_i \times A \to S_i$ that is restricted to be *deterministic*; the MDP local to an agent becomes a finite state automaton. Given a single report of a type $\theta_i^\ell = (s_i^\ell, \tau_i, r_i)$, the state of the agent, and thus its value for subsequent actions, can then be inferred by an observer. Figure 3 depicts MDP representations of some example static types.

We are interested in establishing a connection between the dynamic VCG mechanism and the online VCG mechanism [Parkes and Singh, 2003] in this environment.

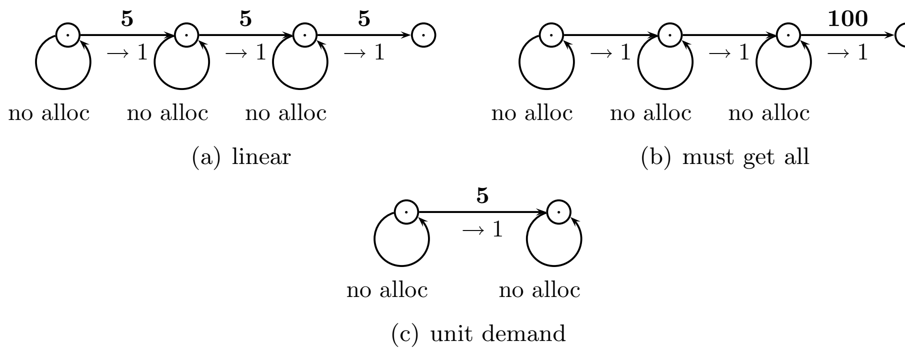


Figure 3: Examples of MDP representations of static types in a problem with 1 item to allocate in each of 3 periods. In (a) the agent obtains value 5 upon each allocation of an item; in (b) the agent obtains value 100 only if allocated all items; in (c) the agent obtains value 5 for any one item.

To make progress, consider a modification to the rules of dynamic mechanisms to insist that an agent makes only a single type report. We refer to such a mechanism as a *report-once* dynamic mechanism. Given this, we can specialize w.p. EPIC so that it only needs to hold in an agent's arrival period. For this, define a *partially-truthful type profile*, $\breve{t}^\ell$, which combines earlier reports received from agents with the current type of agents that either arrive in period ℓ or are already present but did not yet submit a report to the mechanism. We adopt standard notation from earlier in the paper, for example with $V_i(h^{\ell-1}, \breve{t}^\ell, \pi, \sigma_i)$ to denote the flow value to agent i under strategy σ_i given knowledge of $\breve{t}^\ell$.

Definition 6 (within-period ex post incentive compatible)**.** *A report-once dynamic mechanism, $m = (\pi, x)$, is within-period ex post incentive-compatible in the dynamic population, static type environment if, for any period ℓ, and any agent i arriving in period ℓ, any partially-truthful type $\breve{t}^\ell \in \mathcal{T}$ induced by earlier reports and the true types of agents arriving in period ℓ, for any history $h^{\ell-1}$, and for all $\sigma'_i \neq \sigma^*_i$,*

$$V_i(h^{\ell-1}, \breve{t}^\ell, \pi, \sigma^*_i) + X_i(h^{\ell-1}, \breve{t}^\ell, \pi, \sigma^*_i) \geq V_i(h^{\ell-1}, \breve{t}^\ell, \pi, \sigma'_i) + X_i(h^{\ell-1}, \breve{t}^\ell, \pi, \sigma'_i), \quad (28)$$

so that agent i maximizes its expected discounted payoff with truthful strategy σ^*_i given that all other agents are truthful forward from this period.

The new thing here is that the incentive-compatibility constraint need only be satisfied in the arrival period of an agent. Given this, the equilibrium strategy of an agent is to be truthful in its arrival period, from which point forward the agent has no additional strategic decisions. The equilibrium is robust, in that truthful reporting is a best-response for an agent whatever the reported types up until this period and whatever the types of agents arriving in this period, as long as other agents are truthful in the future and the center has a correct probabilistic model of agent arrivals. As a consequence, a dynamic Groves mechanism in this environment need only align an agent's flow transfer with the flow value to the other agents in an agent's arrival period:

Lemma 3. *A report-once dynamic mechanism (π, x) in an environment with a dynamic population and static type is efficient and w.p. EPIC, if it is a dynamic Groves mechanism, which requires that:*

i) policy π is efficient for the correct arrival process,

ii) the expected discounted transfer to any agent i that arrives in period ℓ, given any partially-truthful type profile $\breve{t}^\ell$, history $h^{\ell-1}$, policy π, and strategy σ_i, is $V_{-i}(h^{\ell-1}, \breve{t}^\ell, \pi, \sigma_i) - C_i(\breve{t}^\ell_{-i})$, where $C_i(\breve{t}^\ell_{-i})$ is a quantity that is independent of agent i's own strategy in this period and forward.

The proof of this lemma follows the same outline as earlier: suppose that the mechanism is not w.p. EPIC, and establish a contradiction with the efficiency of the policy by showing that the center could instead adopt the actions taken given agent i's misreports and obtain a policy with greater total expected, discounted value. We now define the online VCG mechanism[25] for this environment:

Definition 7 (online VCG mechanism [Parkes and Singh, 2003])**.** *In the online VCG mechanism, each agent can make a single report (perhaps untruthfully) about its type $t_i^\ell = (s_i^l, \tau_i, r_i)$ in some period ℓ between arrival and departure. The mechanism selects action $a^\ell = \pi^*(s^\ell)$ for the policy π^* that is efficient given reported types and a correct model of the arrival process. Each agent i that makes a report in period ℓ ("arrival") receives transfer $x_i(t^\ell) = -r_i(s^\ell, a^\ell) + \left[V(t^\ell, \pi^*) - V_{-i}(t^\ell, \pi^*_{-i})\right]$, where π^*_{-i} is the policy that is efficient for agents $\neq i$. Other agents that are accessible ("not departed") receive transfer $x_i(t^\ell) = -r_i(s^\ell, a^\ell)$.*

Theorem 4. *The online VCG mechanism and the dynamic VCG mechanism are payoff equivalent upon arrival for a dynamic population with departures and arrivals and static types, and are efficient and w.p. EPIC given private values, the CIA property, and a center with a correct model of agent arrivals.*

Proof. The dynamic VCG mechanism, suitably modified so that agents can only submit a single type report, satisfies the properties of Lemma 3 as a corollary to Theorem 3. Moreover, each agent has flow transfer of $V_{-i}(h^{\ell-1}, \breve{t}^\ell, \pi^*, \sigma_i) - V_{-i}(\breve{t}^\ell, \pi^*_{-i})$ upon its arrival.

For the online VCG mechanism, property (i) holds by construction. Let π^* denote this efficient policy. Now fix some agent i arriving in period ℓ, strategy σ_i, history

[25] The online VCG mechanism was originally proposed by Parkes and Singh [2003] in a setting without discounting, with a transfer that is zero except upon departure, when an agent's transfer is $-\sum_{k=\underline{t}}^{\overline{t}} r_i(\hat{s}_i^k, \pi^*(\hat{s}^k)) + V(\hat{s}^{\underline{t}}, \pi^*) - V_{-i}(\hat{s}^{\overline{t}}, \pi^*_{-i})$ where $\underline{t}$ is the reported arrival period and $\overline{t}$ is the reported departure period. The online VCG mechanism presented here is one of many simple variants that provide the analogous payoffs in a setting with discounting.

$h^{\ell-1}$, and partially-truthful type profile $\check{t}^\ell$, representing previous reports and the current true types of agents that arrive in period ℓ or have yet to report. Assume agents $\neq i$ are truthful going forward. Adopt notation t^k_σ, and s^k_σ to denote the type profile and state profile in period $k \geq \ell$ given strategy σ_i, and let $\widehat{t}^k_\sigma = (\widehat{t}^k_{\sigma,i}, t^k_{\sigma,-i})$ and $\widehat{s}^k_\sigma = (\widehat{s}^k_{\sigma,i}, s^k_{\sigma,-i})$ denote the reported type and state profiles. Let a^k_σ denote the action taken in period k, π^*_σ denote the optimal policy induced by π^* given strategy σ_i, and $k(\sigma)$ denote the period in which agent i announces its arrival.

The expected discounted transfer to agent i forward from arrival period ℓ given strategy σ_i is:

$$\mathbb{E}_{t^k_\sigma,\ k(\sigma)}\left[\gamma^{k(\sigma)-\ell}\left(V_\sigma(\widehat{t}^{k(\sigma)}_\sigma, \pi^*_\sigma) - V_{-i}(\widehat{t}^{k(\sigma)}_\sigma, \pi^*_{-i}) - \sum_{k'=k(\sigma)}^{K} \widehat{r}_i(\widehat{s}^{k'}_\sigma, a^{k'}_\sigma)\right)\right], \tag{29}$$

where $\widehat{r}_i$ denotes agent i's reported reward function under strategy σ_i, and we adopt V_σ in place of V because the expected value of policy π^*_σ depends on the reported reward and transition function, and thus strategy, of agent i. Subtracting the flow reward to agent i from the first term, Eq. (29) reduces to:

$$\mathbb{E}_{t^k_\sigma,\ k(\sigma)}\Big[\gamma^{k(\sigma)-\ell} V_{-i}(\widehat{t}^{k(\sigma)}_\sigma, \pi^*_\sigma)\Big] \ - \ \mathbb{E}_{t^k_\sigma,\ k(\sigma)}\Big[\gamma^{k(\sigma)-\ell} V_{-i}(\widehat{t}^{k(\sigma)}_\sigma, \pi^*_{-i})\Big], \tag{30}$$

Now, adding the flow value to all agents except i under the actions followed by the mechanism in periods from ℓ to $k(\sigma)-1$, that is

$$\mathbb{E}_{t^k_\sigma,\ k(\sigma)}\left[\sum_{k'=\ell}^{k(\sigma)-1} \gamma^{k'-\ell} r_{-i}(\widehat{s}^{k'}_\sigma, a^{k'}_\sigma)\right], \tag{31}$$

to the first term in Eq. (30), we obtain $\mathbb{E}_{t^k_\sigma,\ k(\sigma)}[\sum_{k'=\ell}^{k(\sigma)-1} \gamma^{k'-\ell} r_{-i}(\widehat{s}^{k'}_\sigma, a^{k'}_\sigma)] + \mathbb{E}_{t^k_\sigma,\ k(\sigma)}[\gamma^{k(\sigma)-\ell} V_{-i}(\widehat{t}^{k(\sigma)}_\sigma, \pi^*_\sigma)] = V_{-i}(\widehat{t}^\ell_\sigma, \pi^*_\sigma) = V_{-i}(t^\ell_\sigma, \pi^*_\sigma) = V_{-i}(\check{t}^\ell, \pi^*_\sigma) = V_{-i}(h^{\ell-1}, \check{t}^\ell, \pi^*, \sigma_i)$, where the second equality is by private values and transition independence. Now, subtracting Eq. (31) from the second term in Eq. (30) we obtain,

$$-\mathbb{E}_{t^k_\sigma,\ k(\sigma)}\left[\sum_{k'=\ell}^{k(\sigma)-1} \gamma^{k'-\ell} r_{-i}(\widehat{s}^{k'}_\sigma, a^{k'}_\sigma)\right] - \ \mathbb{E}_{t^k_\sigma,\ k(\sigma)}\Big[\gamma^{k(\sigma)-\ell} V_{-i}(\widehat{t}^{k(\sigma)}_\sigma, \pi^*_{-i})\Big], \tag{32}$$

and this is equal to $-V_{-i}(\widehat{t}^\ell_\sigma, \pi^*_{-i})$ since $a^{k'}_\sigma = \pi^*_{-i}(\widehat{t}^{k'}_\sigma)$ in periods k' before agent i's arrival because of the conditional-independence property of the arrival process. Finally, we have $-V_{-i}(\widehat{t}^\ell_\sigma, \pi^*_{-i}) = -V_{-i}(t^\ell_\sigma, \pi^*_{-i}) = -V_{-i}(\check{t}^\ell, \pi^*_{-i})$ where the first equality is by private values and transition independence. This completes the proof, since $-V_{-i}(\check{t}^\ell, \pi^*_{-i})$ is independent of agent i's strategy, and the flow transfer is $V_{-i}(h^{\ell-1}, \check{t}^\ell, \pi^*, \sigma_i) - V_{-i}(\check{t}^\ell, \pi^*_{-i})$ upon its arrival. $\square$

	Agent 1		Agent 3	
Online VCG	Win	0.145	Win	$v_3 - 0.7$
	Lose	0.145	Lose	0
Dynamic VCG	Win	$0.6 - v_3$	Win	$v_3 - 0.7$
	Lose	-0.1	Lose	0

Table 1: Utility to agents 1 and 3 in online VCG and dynamic VCG in Example 5.

The online VCG and dynamic VCG mechanisms are both w.p. EPIC and efficient in this environment. But the online VCG mechanism provides *ex post* individual-rationality whereas the dynamic VCG is only w.p. *ex post* individual rational. In the online VCG mechanism, an agent's total discounted payoff from reporting its true type, for all type profiles and all strategies of other agents, is equal to the estimated marginal product it contributes to the system at arrival and non-negative in social choice and resource allocation problems.

Example 5. *Consider a problem with two periods, one item to allocate, and three agents. In period 1, there is an agent with value \$0.7 and arrival-departure (1,2) so it is patient. Another agent has value \$0.6 and is arrival-departure (1,1) so it is impatient. In period 2, there will be an agent with value* $v_3 \sim U(0,1)$. *Suppose there is no discounting. The efficient policy is to wait until period 2 and then allocate agent 1 if* $v_3 \leq 0.7$ *and agent 3 otherwise.*

(a) Online VCG. We have $V(s^1, \pi^*) = (0.7)(0.7) + (0.3)(0.85) = 0.745$ *and* $V_{-1}(s^1, \pi^*) = 0.6$. *If agent 1 wins, then its total transfer is -0.7 + (0.745-0.6)=-0.555. If agent 1 loses, then its total transfer is 0+(0.745-0.6)=0.145. Agent 2 makes no transfer because it is not pivotal. Agent 3 receives no transfer if it loses, and its transfer is* $-v_3 + (v_3 - 0.7) = -0.7$ *if it wins.*

(b) Dynamic VCG. We have $\mathbb{E}_{s'}[V_{-1}(s', \pi^*_{-1})|\mathit{don't\ allocate}] = 0.5$ *(the expected value to other agents in period 2 given that the item is not allocated in period 1), and* $V_{-1}(s^1, \pi^*_{-1}) = 0.6$. *In period 1, agent 1's transfer is 0 + (0.5-0.6)= -0.1. In period 2, if agent 1 wins and* $v_3 \leq 0.7$ *then its transfer is* $0 + (0 - v_3)$*; otherwise, if agent 1 loses then its transfer is 0. Agent 2 makes no transfer because it is not pivotal. Agent 3's transfer is 0 in period 1 because it has not arrived. If agent 3 wins, then it receives transfer* $0 + (0 - 0.7) = -0.7$ *and has no transfer if it loses.*

The utility to agents 1 and 3 is tabulated in Table 1. Although the expected utility is the same to both agents in the mechanisms, in the dynamic VCG agent 1's utility is negative when it wins and $v_3 \in (0.6, 0.7]$ *and when it loses. The online VCG mechanism is* ex post *individual-rational.*

On the other hand, while the online VCG and dynamic VCG mechanisms have the same *ex ante* flow payment, the dynamic VCG mechanism is *ex post* no deficit in economic environments in without positive externalities (e.g., social choice and one-sided auction problems.) So, the two mechanisms exhibit a tradeoff between achieving *ex post* individual-rationality and *ex post* no deficit.

	Period 2 arrivals	Agent 1 transfer	Other agents	Total transfer
Online VCG	0	$(-2\epsilon+\epsilon^2)100$	0	$(-2\epsilon+\epsilon^2)100$
	1	$100+(-2\epsilon+\epsilon^2)100$	-100	$(-2\epsilon+\epsilon^2)100$
	2	$100+(-2\epsilon+\epsilon^2)100$	0	$100+(-2\epsilon+\epsilon^2)100$
Dynamic VCG	0	0	0	0
	1	0	-100	-100
	2	0	0	0

Table 2: Transfers in the online VCG and dynamic VCG mechanisms in Example 6.

Example 6. *Consider a problem with two periods, and two units of an item to allocate. In period 1, there is a patient agent with value \$100 and arrival-departure (1,2). In period 2, two agents might arrive, each with low probability $\epsilon > 0$ of arriving and with value \$150 for one unit. There is no discounting. The efficient policy is to wait until period 2 and allocate to agent 1 unless one or both of the high value agents arrive, in which case they are allocated.*

(a) Online VCG. We have $V(s^1, \pi^*) = (1 - 2\epsilon + \epsilon^2)(100) + 2\epsilon(150) + \epsilon^2(300)$ *and* $V_{-1}(s^1, \pi^*) = 2\epsilon(150) + \epsilon^2(300)$*. If agent 1 wins, then its total transfer is* $-100 + ((1 - 2\epsilon + \epsilon^2)(100) + 2\epsilon(150) + \epsilon^2(300)) - (2\epsilon(150) + \epsilon^2(300)) = (-2\epsilon + \epsilon^2)(100)$*. If agent 1 loses, then its total transfer is* $100 + (-2\epsilon + \epsilon^2)(100)$*. If one agent arrives in period 2, it wins and its transfer is* $-150 + (150 - 100) = -100$*. If two agents arrive in period 2, they each win and have transfer* $-150 + (300 - 150) = 0$*.*

(b) Dynamic VCG. If agent 1 wins or loses, its transfer is 0. If one agent arrives in period 2, then it wins and its transfer is $0 + (0 - 100) = -100$*. If two agents arrive in period 2, then they each win and have transfer* $150 + (0 - 150) = 0$*.*

In Table 2 we summarize the transfers that occur in each mechanism, depending on the number of agents to arrive in period 2. Whereas the expected transfers are the same in both mechanisms, the online VCG mechanism incurs a deficit when 2 agents arrive in period 2 of \$100 as $\epsilon \to 0$*. On the other hand, the dynamic VCG mechanism always runs without a deficit.*

5 Dynamic, Interdependent Types

We have assumed until now a private values model and independent type transitions, with this latter requirement translating into an independence requirement on the arrival process for an environment with a dynamic population. But many interesting environments exhibit interdependent type transitions. For example, consider an oil field that is being explored in parallel by multiple firms, each of which is competing for future drilling rights. The value that one firm has may depend on the information available to other firms about the quality of the oil field, as represented by their local states. In Internet advertising, the value of firm A for advertising to a particular demographic might be expected to depend on what another, similar firm B learns about its conversions in advertising to the same user demographic.

In static problems, implementing the first best (efficient) allocation is impossible for generic, multi-dimensional value interdependence [Jehiel and Moldovanu, 2001; Dasgupta and Maskin, 2000]. Similarly, it is impossible to implement efficient policies in interdependent dynamic problems in which there are constraints on when transfers are possible [Gershkov and Moldovanu, 2008c]. For this reason, our focus here is on an environment of persistent, always accessible agents. We retain private valuations, but model type interdependence through interdependent type transitions:

$$s_i \;:\; S_1 \times \ldots \times S_n \times A \to S_i \tag{33}$$
$$r_i \;:\; S_i \times A \times S_i \to \mathcal{R} \tag{34}$$

The next state $\tau_i(s,a) \in S_i$ can depend on the state of other agents. Whereas we earlier had reward $r_i(s_i,a) \in \mathcal{R}$, we now allow reward $r_i(s_i,a,s_i') \in \mathcal{R}$ where $s_i' = \tau_i(s,a)$.[26] Thus, the interdependence occurs through the effect of one agent's current state on another agent's next state. But agent i's value is privately realized by agent i given this state transition. For example, this model allows the reward received by agent 2 for the allocation of a resource to depend in arbitrary ways on the private state of agent 1.

Mezzetti [2004] provides a w.p. EPIC mechanism for a single-period version of this problem by using a second-stage of reports. In period 0, given type profile $t^0 = (s^0,\tau,r)$, then action $a^0 \in A$ causes agent i to realize private transition $s_i^1 = \tau_i(s^0,a)$ and receive private reward $r_i(s_i^0,a,s_i^1)$ (still in in period 0) that depends on its next state s_i^1. Mezzetti's types have deterministic transitions but this is unnecessary and they can be dynamic types. In two-step mechanism, each agent makes a report, $\widehat{t}_i^0$, about its type and the action a that maximizes the expected total value to all agents given this report is selected. Having transitioned into a new state, each agent i then makes a subsequent report $\widehat{r}_i$ about its received reward. The reward to an agent depends on its private state transition, which depends in turn on the *true* type profile t^0. By making a subsequent transfer of $\sum_{j \neq i} \widehat{r}_j$, each agent receives as a transfer the total value received by other agents and the mechanism is a Groves mechanism and w.p. EPIC.

We generalize Mezzetti's observation to a multi-period, dynamic environment with interdependent types. Just as in Mezzetti [2004], an agent in our model knows its own value immediately *after* an action is selected by the center. This precludes models of "prestige" goods, such as a watch, where the realized value of an allocation (even after it has been made) may in fact depend on the earlier claims by other agents about their values. The temporal extent of any informational externality in

[26] One can think about the earlier use of $r_i(s_i,a)$ as equivalent to the expected reward to agent i for reward function $r_i(s_i,a,s_i')$, given private state transitions $\tau_i(s_i,a)$.

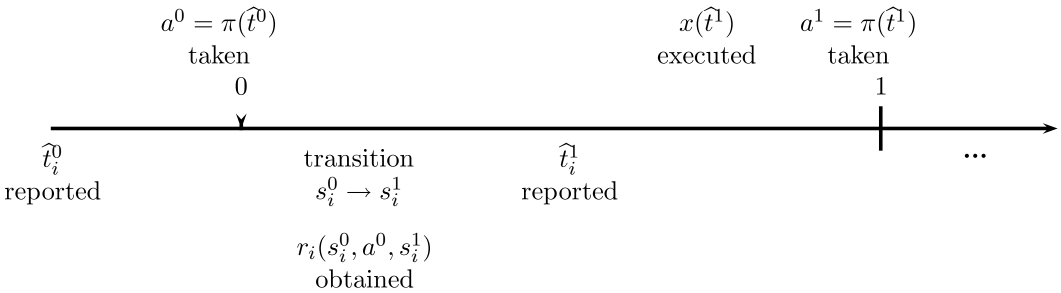


Figure 4: An illustration of the timing of a mechanism implemented in the interdependent-dynamic framework. The reward and transfers received in the same period are discounted in the same way.

the current period lasts only until the next period.[27] Figure 4 illustrates the timing of reports, actions, transitions, rewards and transfers in a dynamic mechanism in this interdependent type environment. Action $a^\ell = \pi(\widehat{t}^\ell)$ triggers the start of period ℓ. Each agent then learns about its state transition $s_i^{\ell+1} = \tau_i(s^\ell, a)$ and obtains reward $r_i(s_i^\ell, a, s_i^{\ell+1})$. Having received reports $\widehat{t}^{\ell+1}$ about each agent's type for the next period, the center makes transfer $x_i(\widehat{t}^{\ell+1})$. The value of this transfer to agent i accrues in period ℓ, along with its received reward $r_i(s_i^\ell, a, s_i^{\ell+1})$. The next period triggers with action $a^{\ell+1} = \pi(\widehat{t}^{\ell+1})$.

Lemma 1, which establishes the w.p. EPIC of dynamic Groves mechanisms, continues to hold (with the same proof) for this interdependent type environment. All that is required is to redefine,

$$V_j(h^{\ell-1}, t^\ell, \pi, \sigma_i) = \mathbb{E}_{t_\sigma^k,\ a_\sigma^k}\left[\sum_{k=\ell}^{K} \gamma^{k-\ell}\left(\sum_{j \neq i} r_j^k(s_j^k, a_\sigma^k, s_j^{k+1})\right) \mid h^{\ell-1}, t^\ell, \pi, \sigma_i\right], \quad (35)$$

so that this correctly reflects the true flow value to agents $\neq i$ given that agent i follows strategy σ_i forward from period ℓ. Consider now a naive interdependent-type, dynamic Groves mechanism, that in each period,

- selects action a^ℓ at the start of period ℓ according to the policy that is efficient given agent type reports,
- receives reports $t^{\ell+1}$ (perhaps untruthful) about the type profile that agents will adopt in the next period, and
- makes transfer $x_i(t^\ell, t^{\ell+1}) = \sum_{j \neq i} r_j(s_j^\ell, a^\ell, s_j^{l+1})$ to each agent.

[27] Athey and Segal [2007] also propose a dynamic team mechanism that can handle interdependent transitions with private values. Our contribution is to make a structural observation about the additional difficulty in obtaining a dynamic VCG mechanism that does incur a deficit to the mechanism while retaining w.p. *ex post* individual rationality.

Theorem 5. *The interdependent-type, dynamic Groves mechanism is efficient and w.p. EPIC for a fixed agent population with dynamic, interdependent type but private valuations.*

The equilibrium in this mechanism is less fragile than Mezzetti's single period mechanism because the type report of an agent in period ℓ (about its type in period $\ell+1$) is used not only to align the incentives of other agents but also affects the subsequent action selected by the mechanism. Each agent therefore has strict incentives to report its true type in each period, except in just the very last period in a problem with a finite time horizon.

The particular dynamic Groves mechanism (or team mechanism) presented above still has serious budgetary problems. One can impose an *ex ante* charge on agents and achieve a mechanism that is *ex ante* no-deficit and *ex ante* individual-rational, but this would give up on w.p. *ex post* individual-rationality. To get around this, we would need to introduce a chargeback transfer that is applied adaptively, in each period, in order to obtain no deficit without losing w.p. *ex post* individual-rationality. One might, for example, think to redefine the transfer as,

$$x_i(t^\ell, t^{\ell+1}) = r_{-i}(s^\ell, a^\ell, s^{\ell+1}) + \gamma \mathbb{E}_{t'_{-i}}\Big[V_{-i}((\boldsymbol{t}_i^{\ell+1}, t'_{-i}), \pi^*_{-i}) | \boldsymbol{t}_i^l, t_{-i}^\ell, a^\ell\Big] \\ - V_{-i}((\boldsymbol{t}_i^\ell, t_{-i}^\ell), \pi^*_{-i}), \tag{36}$$

where $\boldsymbol{t}_i^k$, for $k \in \{0, 1, \dots\}$ is an *ex ante* fixed sequence of type profiles for agent i that are independent of agent i's reports. The second expression,

$$\mathbb{E}_{t'_{-i}}\Big[V_{-i}((\boldsymbol{t}_i^{\ell+1}, t'_{-i}), \pi^*_{-i}) | \boldsymbol{t}_i^l, t_{-i}^\ell, a^\ell\Big], \tag{37}$$

is the estimated flow value to agents other than i forward from period $\ell+1$ given action a^ℓ and given type profile $(\boldsymbol{t}_i^l, t_{-i}^\ell)$ in the current period. Looking back to the proof of w.p. EPIC for the dynamic VCG mechanism in the private values model, we required that Eq. (13) reduced to Eq. (15) and canceled with the flow transfer terms in the next period and forward that occur due to the third component of the transfer rule. Using the same trick here, then the flow transfer forward from period ℓ would be $V_{-i}((\boldsymbol{t}_i^\ell, t_{-i}^\ell), \pi^*_{-i})$ and independent of agent i's strategy as required for w.p. EPIC. But whereas the next state distribution in Eq. (14) is the same (in equilibrium) as that actually traversed due to the actions taken by the mechanism, this is not the case for the next-type distribution in Eq. (37) because the state transitions to the agents $\neq i$ are conditioned on $\boldsymbol{t}_i^\ell$ rather than agent i's actual type, t_i^ℓ. It does not seem likely, to us, that it will be possible to find a suitable modification to the transfer terms of dynamic VCG to achieve w.p. EPIC along with no deficit and w.p. *ex post* individual rationality in this dynamic, interdependent type setting.

6 Conclusions

In this paper we have extended the dynamic VCG mechanism [Bergemann and Välimäki, 2008] to environments in which agents are periodically-inaccessible and used this extension to derive a mechanism that is w.p. EPIC and efficient for problems with dynamic populations and dynamic agent type. We require private values and type transitions that are conditionally independent of the type of other agents, when conditioned on actions by the center. Similarly, agent arrivals must be conditionally-independent of earlier arrivals, when conditioned on actions of the center. This generalizes existing results, which offered a truthful implementation either for a static population with dynamic type or a dynamic population with static type. For the special case of dynamic population and static type, we show that the dynamic VCG mechanism is payoff equivalent upon arrival to the online VCG mechanism [Parkes and Singh, 2003], which is *ex post* individual-rational where the dynamic VCG mechanism is only within-period *ex post* individual-rational. Finally, we observe a structural difficulty in extending dynamic VCG mechanisms to environments in which agents have interdependent type transitions.

References

[Arrow, 1979] Kenneth J. Arrow. The property rights doctrine and demand revelation under incomplete information. In M. Boskin, editor, *Economics and Human Welfare*. Academic Press, 1979.

[Athey and Segal, 2007] Susan Athey and Ilya Segal. An efficient dynamic mechanism. Working paper, Stanford University. Earlier version circulated in 2004., 2007.

[Athey *et al.*, 2004] Susan Athey, Kyle Bagwell, and Chris Sanchirico. Collusion and price rigidity. *The Review of Economic Studies*, 71:317–349, 2004.

[Atkeson and Lucas, 1992] Andrew Atkeson and Robert E Lucas. On efficient distribution with private information. *The Review of Economic Studies*, 59:427–453, 1992.

[Babaioff *et al.*, 2009] Moshe Babaioff, Liad Blumrosen, and Aaron Roth. Auctions with online supply. In *Fifth Workshop on Ad Auctions*, 2009.

[Bailey, 1997] Martin J. Bailey. The demand revealing process: To distribute the surplus. *Public Choice*, 91:107–126, 1997.

[Bapna and Weber, 2005] Abbishek Bapna and Thomas A Weber. Efficient dynamic allocation with uncertain valuations. Working paper, Stanford University, 2005.

[Battaglini, 2005] Marco Battaglini. Long-term contracting with Markovian consumers. *American Economic Review*, 95:637–658, 2005.

[Bergemann and Välimäki, 2008] Dirk Bergemann and Juuso Välimäki. The dynamic pivot mechanism. Cowles Foundation Discussion Paper 1672, Earlier version circulated as "Efficient Dynamic Auctions", 2006, 2008.

[Board, 2008] Simon Board. Durable-goods monopoly with varying demand. *The Review of Economic Studies*, 75:391–413, 2008.

[Bredin *et al.*, 2007] Jonathan Bredin, David C. Parkes, and Quang Duong. Chain: A dynamic double auction framework for matching patient agents. *Journal of Artificial Intelligence Research*, 30:133–179, 2007. Earlier version appeared in Proc. UAI'05.

[Cavallo and Parkes, 2008] Ruggiero Cavallo and David C. Parkes. Efficient metadeliberation auctions. In *Proc. 23rd AAAI Conference on Artificial Intelligence (AAAI'08)*, pages 50–56, Chicago, IL, 2008.

[Cavallo *et al.*, 2006] Ruggiero Cavallo, David C. Parkes, and Satinder Singh. Optimal coordinated planning amongst self-interested agents with private state. In *Proceedings of the Twenty-second Annual Conference on Uncertainty in Artificial Intelligence (UAI'06)*, pages 55–62, 2006.

[Cavallo, 2006] Ruggiero Cavallo. Optimal decision-making with minimal waste: Strategyproof redistribution of VCG payments. In *Proc. of the 5th Int. Joint Conf. on Autonomous Agents and Multi Agent Systems (AAMAS'06)*, Hakodate, Japan, 2006.

[Cavallo, 2008] Ruggiero Cavallo. Efficiency and redistribution in dynamic mechanism design. In *Proceedings of the 9th ACM Conference on Electronic Commerce (EC-08)*, 2008.

[Cole *et al.*, 2008] R. Cole, S. Dobzinski, and L. Fleischer. Prompt mechanisms for online auctions. In *Symposium on Algorithmic Game Theory (SAGT)*, pages 170–181, 2008.

[Constantin and Parkes, 2009] Florin Constantin and David C. Parkes. Self-Correcting Sampling-Based Dynamic Multi-Unit Auctions. In *10th ACM Electronic Commerce Conference (EC'09)*, 2009. To appear.

[Courty and Li, 2000] Pascal Courty and Hao Li. Sequential screeing. *The Review of Economic Studies*, 67:697–717, 2000.

[Crémer *et al.*, 2009] Jacques Crémer, Yossi Spiegel, and Charles Z. Zheng. Auctions with costly information acquisition. *Economic Theory*, 38:41–72, 2009.

[Dasgupta and Maskin, 2000] Partha Dasgupta and Eric Maskin. Efficient auctions. *The Quarterly Journal of Economics*, 115(2):341–388, 2000.

[d'Aspermont and Gérard-Varet, 1979] C. d'Aspermont and L.A. Gérard-Varet. Incentives and incomplete information. *Journal of Public Economics*, 11:25–45, 1979.

[Dizdar *et al.*, 2009] Deniz Dizdar, Alex Gershkov, and Benny Moldovanu. Revenue maximization in the dynamic knapsack problem. Working paper, University of Bonn, 2009.

[Dolan, 1978] Robert J Dolan. Incentive mechainsms for priority queuing problem. *Bell Journal of Economics*, 9:421–436, 1978.

[Freixas *et al.*, 1985] Xavier Freixas, Roger Guesnerie, and Jean Tirole. Planning under incomplete information and the ratchet effect. *The Review of Economic Studies*, 52:173–191, 1985.

[Friedman and Parkes, 2003] E. Friedman and D. C. Parkes. Pricing WiFi at Starbucks– issues in online mechanism design. In *Proc. Fourth ACM Conference on Electronic Commerce (EC'03)*, pages 240–241, 2003.

[Gallego and van Ryzin, 1994] Guillermo Gallego and Garrett van Ryzin. Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management Science*, 40:999–1020, 1994.

[Gallien, 2003] Jeremie Gallien. Dynamic mechanism design for online commerce. *Operations Research*, 54:291–310, 2003.

[Gershkov and Moldovanu, 2008a] Alex Gershkov and Benny Moldovanu. Dynamic revenue maximization with heterogeneous objects: A mechanism design approach. *American Economic Journal: Microeconomics*, 2008. To appear.

[Gershkov and Moldovanu, 2008b] Alex Gershkov and Benny Moldovanu. Efficient sequential assignment with incomplete information. *Games and Economic Behavior*, 2008. To appear.

[Gershkov and Moldovanu, 2008c] Alex Gershkov and Benny Moldovanu. Learning about the future and dynamic efficiency. *American Economic Review*, 2008. To appear.

[Gershkov and Moldovanu, 2009] Alex Gershkov and Benny Moldovanu. Optimal search, learning and implementation. Working paper, University of Bonn, 2009.

[Hajiaghayi *et al.*, 2004] Mohammad T. Hajiaghayi, Robert Kleinberg, and David C. Parkes. Adaptive limited-supply online auctions. In *Proc. ACM Conf. on Electronic Commerce*, pages 71–80, 2004.

[Hajiaghayi *et al.*, 2005] Mohammad T. Hajiaghayi, Robert Kleinberg, Mohammad Mahdian, and David C. Parkes. Online auctions with re-usable goods. In *Proc. ACM Conf. on Electronic Commerce*, pages 165–174, 2005.

[Jehiel and Moldovanu, 2001] Philippe Jehiel and Benny Moldovanu. Efficient design with interdependent valuations. *Econometrica*, 69:1237–1259, 2001.

[Juda and Parkes, 2009] Adam I. Juda and David C. Parkes. An options-based solution to the sequential auction problem. *Artificial Intelligence*, 173:876–899, 2009. Early version at AMEC workshop (2004) and ACMEC (2006).

[Kaelbling *et al.*, 1996] Leslie Pack Kaelbling, Michael L. Littman, and Andrew W. Moore. Reinforcement learning: A survey. *Journal of Artificial Intelligence Research*, 4:237–285, 1996.

[Kincaid and Darling, 1963] Wilfred M. Kincaid and Donald A. Darling. An inventory pricing problem. *Journal of Mathematical Analysis and Applications*, 7:183–208, 1963.

[Lavi and Nisan, 2004] Ron Lavi and Noam Nisan. Competitive analysis of incentive compatible on-line auctions. *Theoretical Computer Science*, 310:159–180, 2004. Earlier version in ACMEC 2000.

[Lavi and Nisan, 2005] Ron Lavi and Noam Nisan. Online ascending auctions for gradually expiring goods. In *Proc. of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1146–1155, 2005.

[Lavi and Segev, 2008] R. Lavi and E. Segev. Efficiency levels in sequential auctions with dynamic arrivals. Working paper, The Technion, 2008.

[Levin, 2003] J. Levin. Relational incentive constraints. *Americn Economic Review*, 93:835–857, 2003.

[Mahdian and Saberi, 2006] M. Mahdian and A. Saberi. Multi-unit auctions with unknown supply. In *Proc. ACM conference on Electronic Commerce (EC)*, pages 243–249, 2006.

[McAfee and te Velde, 2008] R. Preston McAfee and Vera te Velde. Dynamic pricing with constant demand elasticity. *Production and Operations Management*, 17:432–438, 2008.

[McAfee, 1992] R. Preston McAfee. A dominant strategy double auction. *Journal of Economic Theory*, 56:434–450, 1992.

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

[Mierendorff, 2008] Konrad Mierendorff. Efficient intertemporal auction. Working paper, University of Bonn, 2008.

[Mierendorff, 2009] Konrad Mierendorff. Optimal dynamic mechanism design with deadlines. Working paper, University of Bonn, 2009.

[Myerson, 1986] Roger Myerson. Multistage games with communication. *Econometrica*, 54(2):323–358, 1986.

[Ng *et al.*, 2003] Chaki Ng, David C. Parkes, and Margo Seltzer. Virtual Worlds: Fast and Strategyproof Auctions for Dynamic Resource Allocation. In *Proc. Fourth ACM Conf. on Electronic Commerce (EC'03)*, pages 238–239, 2003.

[Pai and Vohra, 2008] Mallesh Pai and Rakesh Vohra. Optimal dynamic auctions. Working paper, MEDS Department, Kellogg School of Management, 2008.

[Parkes and Singh, 2003] David C. Parkes and Satinder Singh. An MDP-based approach to Online Mechanism Design. In *Proc. 17th Annual Conf. on Neural Information Processing Systems (NIPS'03)*, 2003.

[Parkes *et al.*, 2004] David C. Parkes, Satinder Singh, and Dimah Yanovsky. Approximately efficient online mechanism design. In *Proc. 18th Annual Conf. on Neural Information Processing Systems (NIPS'04)*, 2004.

[Parkes, 2007] David C Parkes. On-line mechanisms. In Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay Vazirani, editors, *Algorithmic Game Theory*, chapter 16. Cambridge University Press, 2007.

[Pavan *et al.*, 2009] Alessandro Pavan, Ilya Segal, and Jusso Toikka. Dynamic mechanism design: Incentive compatibility, profit maximization and information disclosure. Working paper, Northwestern University, 2009.

[Porter, 2004] Ryan Porter. Mechanism design for online real-time scheduling. In *Proc. ACM Conf. on Electronic Commerce (EC'04)*, pages 61–70, 2004.

[Riley and Zeckhauser, 1983] John Riley and Richard Zeckhauser. Optimal selling strategies: When to haggle, when to hold firm. *The Quarterly Journal of Economics*, 98:267–289, 1983.

[Said, 2009] Maher Said. Auctions with dynamic populations: Efficiency and revenue maximization. Working paper, Yale University, 2009.

[Seuken *et al.*, 2008] Sven Seuken, Ruggiero Cavallo, and David C. Parkes. Partially-synchronized DEC-MDPs in dynamic mechanism design. In *Proc. 23rd National Conference on Arti*

cial Intelligence (AAAI'08), pages 162–169, 2008.

[Vulcano *et al.*, 2002] G. Vulcano, Garret van Ryzin, and C. Maglaras. Optimal dynamic auctions for revenue management. *Management Science*, 48:1388–1407, 2002.

[Wang, 1995] Cheng Wang. Dynamic insurance with private information and balanced budgets. *The Review of Economic Studies*, 62:577–595, 1995.