

Solving Dynamic Public Insurance Games with Endogenous Agent Distributions: Theory and Computational Approximation*

Timothy Kam and Ronald Stauber[†]

April 1, 2016

Abstract

We make two contributions in this paper. First, we extend the characterization of equilibrium payoff correspondences in history-dependent dynamic policy games to a class with endogenously heterogeneous private agents. In contrast to policy games involving representative agents, this extension has interesting consequences as it implies additional nonlinearity (i.e., bilinearity) between the game states (distributions) and continuation/promised values in the policymaker’s objective and incentive constraints. The second contribution of our paper is in addressing the computational challenges arising from this payoff-relevant nonlinearity. Exploiting the game’s structure, we propose implementable approximate bilinear programming formulations to construct estimates of the equilibrium value correspondence. Our approximation method respects the property of upper hemi-continuity in the target correspondence. We provide small-scale computational examples as proofs of concept.

Keywords: Heterogeneity; Dynamic games; Bilinear Programs; Approximation

JEL Classification: C61; C73; D86; E24; H21

*We are grateful to Árpád Ábrahám, Gaurab Aryal, Richard Brent, Begoña Domínguez, Zhigang Feng, Kenneth Judd, Tom Krebs, Felix Kübler, Paul Leopardi, Jonathan Ligon, Andy McLennan, Nicola Pavoni, Christopher Sleet and Tim Worrall for suggestions and comments. We also thank participants at the 2011 SAET Conference in Faro, Australasian Workshop on Macroeconomic Dynamics at the University of Queensland, and seminar participants at the Universities of Melbourne and Sydney, Monash University, and at the ANU Mathematical Sciences Institute (Computational Group). This paper previously circulated as “Worker Heterogeneity and Limited-commitment Public Insurance”.

[†]Research School of Economics, Australian National University, ACT 2601, Australia. Email: tcy.kam@gmail.com (T. Kam) and ronald.stauber@anu.edu.au (R. Stauber)

1 Introduction

The contribution of this paper is twofold. First, we extend the characterization of symmetric sequential equilibrium (SSE) value correspondences in history-dependent dynamic policy games [Phelan and Stacchetti, 2001] to a setting with endogenously heterogeneous private agents.

Second, the extended SSE value correspondence description has an interesting consequence: Agent heterogeneity implies that the domain of the equilibrium payoff correspondence—the relevant game state space—is now a set of distributions over agent types, and these distributions interact nonlinearly with agents’ continuation payoffs in the (utilitarian) policymaker’s objective and constraints. This payoff-relevant nonlinearity in our class of games poses a fresh computational challenge. We show that a practical way to tackle this is by exploiting the game’s structure, which gives rise to an approximate bilinear programming (BLP) formulation for the optimization problems characterizing SSE.¹ This provides a means to computing an approximate SSE payoff correspondence that preserves upper hemicontinuity of the true correspondence. We provide proof-of-concept evidence, in the form of small-scale numerical examples that can be solved on a desktop multi-core computer, and which should be scalable to massively distributed computing facilities.

In addition, we also characterize equilibrium under a class of fixed policies, showing that each of these arguably simple and operational fixed policies induces a unique corresponding long-run equilibrium, among which a socially optimal equilibrium is shown to exist. We show in a numerical example, that it is possible that some of these optimal long-run simple policies can attain a long run (welfare) outcome that can be sustainable as a particular SSE—i.e., one beginning from a game state consistent with the desired long run outcome. In contrast, an assumed variable commitment policy (also known as Ramsey policy) plan is shown not to be an SSE at all.

Theoretical extension and new computational methods. From a technical point of view, our results are based on the characterization of equilibrium payoff sets that was derived for repeated games by Abreu et al. [1990], and was extended to dynamic games with public state variables by Atkeson [1991], and with private state variables by Phelan and Stacchetti [2001]. The general setup we use draws mostly on Phelan and Stacchetti [2001], although in our model all the relevant state variables are public. However, we extend the framework

¹Standard linear programming problems involve the maximization of a linear objective function subject to linear constraints. A canonical bilinear programming problem has the form $\max_{x,y} \{c^T x + x^T Q y + d^T y\}$ subject to $Ax \leq a$ and $By \leq b$, where x, y, c, d, a and b are Euclidean vectors, and Q, A and B are matrices. Hence, such bilinear problems are linear in x for fixed y , and linear in y for fixed x .

of [Phelan and Stacchetti \[2001\]](#) by considering a game where the government as a large player faces a non-degenerate distribution over heterogeneous agents, who are small in the sense that they have no individual strategic impact on the game. Our main contribution is in constructing a corresponding heterogeneous-agent dynamic game model that can be analyzed in a tractable way using standard theoretical tools and techniques, and for which a computational approximation can be implemented. Furthermore, for the case where the government is able to commit to a fixed transfer policy, we show that this model allows a simple and novel characterization of existence and optimality of resulting steady state outcomes. The key to our approach is in using a finite-dimensional variable that captures the endogenously evolving distribution of private agent heterogeneity (drawn from a finite set) as a state variable that enters the government’s objective. This is in contrast to the government in [Phelan and Stacchetti \[2001\]](#), who is faced with a continuum of agents with identical characteristics, so their game essentially reduces to a game between the government and a representative agent. To make our model tractable—also with a view towards the computational implementation—we simplify one feature relative to the setting of [Phelan and Stacchetti \[2001\]](#), by not allowing private agents to individually accumulate capital. Since private agents are not homogeneous, the level of capital accumulation would be determined by their individual histories of actions and personal states, and not just by their current personal state, which would result in a need to represent combinations of the personal states and capital states using very complex, infinite-dimensional state variables.²

Applying the methods arising from [Abreu et al. \[1990\]](#) and the related literature to our class of dynamic games, we show that the set of Symmetric Sequential Equilibria (SSE) can be described by an equilibrium payoff correspondence. This correspondence maps the set of distributions over the private agents’ personal states—i.e., the set of game states—to payoff vectors summarizing the equilibrium payoffs derived by the government and the private agents. The SSE payoff correspondence can then be characterized as the largest fixed point of a correspondence-valued operator, and can be computed by recursively applying this operator. Computationally, we implement this operator based on the methods suggested by [Sleet and Yeltekin \[2000\]](#) and [Judd et al. \[2003\]](#). However, we show that in our setting this operator can be constructed using approximate bilinear programs (BLP). Our novel algorithm tractably extends standard linear programming (LP) based computational methods for dynamic games [e.g. [Feng, 2015](#)].³

²A previous version of this paper included government savings as an additional state variable. However, the resulting model does not significantly alter the analysis, while requiring additional notation.

³In turn, [Feng \[2015\]](#) extends the seminal work of [Judd et al. \[2003\]](#), who compute equilibrium payoff sets for repeated games. [Feng et al. \[2014\]](#) also use a similar method to fully describe and solve for incomplete-markets recursive equilibrium value correspondences.

The method of [Feng \[2015\]](#) discretizes the domain of an upper hemicontinuous SSE value correspondence of interest and utilizes LP formulations (of SSE optimization problems) to construct approximating outer and inner step correspondences for discrete slices of the true payoff correspondence.⁴ However, [Sleet and Yeltekin \[2000\]](#) had pointed out that it is no longer clear what one means by an outer- or an inner-approximating step correspondence in this case when the domain is discretized. This is because at each discretized state, the maximal and minimal levels of the “steps” only apply to that state, and, not to a continuous neighborhood of that state, and therefore, may not be outer- nor inner-bounding levels for nearby payoff sets. What is different in our limited-commitment policy problem, is that the policymaker’s objective (in terms of calculating worst punishment values) and incentive constraints will involve bilinear interaction terms between a game state (distribution) and a vector of continuation values for the small players.⁵ We exploit the game’s structure, which gives rise to an approximate bilinear programming formulation (BLP) for the optimization sub-problems characterizing SSE.⁶ This provides a means to computing an approximate SSE payoff correspondence that preserves upper hemicontinuity of the true correspondence. Thus, our proposed BLP approach does not run into the problem pointed out by [Sleet and Yeltekin \[2000\]](#).

An illustrative heterogenous-agent model. For concreteness, our main analysis is framed in the context of an unemployment-insurance game between a government and a continuum of private agents. However, our results can easily be applied to a wide range of dynamic public insurance games. A corresponding generalization of our model is discussed in Section 2.4. In the unemployment-insurance interpretation of our game, the private agents (“workers”) are distinguished by their heterogenous states of individual employment or unemployment

⁴[Feng \[2015\]](#) and [Feng et al. \[2014\]](#) define “step correspondences” in a different way from us and [Sleet and Yeltekin \[2000\]](#). For every slice of the true correspondence at each particular discrete state, they find a union of hypercubes that form the largest possible set contained within the true set (their inner approximation), or, a smallest union of hypercubes that cover the true set (outer approximation). The finite collection of these hypercube sets over all the discretized states are defined to be their respective inner- and outer-approximating “step correspondences.”

⁵The game state space is a probability simplex. This is in contrast to, for example, an interval, a square, or finite collections of intervals in known dynamic policy games [e.g. [Phelan and Stacchetti, 2001](#), [Feng, 2015](#)]. Moreover, in these well-known neoclassical growth model settings, the game state does not additionally interact in a nonlinear way with the representative private agent’s payoffs which are contingent on those states.

⁶BLP as a special class of nonlinear programs is well studied and there are known results for the existence of an “ ϵ -global” optimum, and for algorithms to find it [see e.g. [Horst and Tuy, 1996](#), [Bennett and Mangasarian, 1993](#), [McCormick, 1976](#)]. BLP is also widely used in the fields of operations research, chemical engineering and computational image processing [see e.g. [Nahapetyan and Pardalos, 2008](#), [Chandraker and Kriegman, 2008](#)]. More recently, BLP as a formulation to approximate large-scale Markov decision problems (MDP) have also been proposed [see [Petrik and Zilberstein, 2011](#)]. Our contribution here is related to [Petrik and Zilberstein \[2011\]](#) where the operator defining our equilibrium payoff correspondence has a similar spirit to the single decision maker’s Bellman operator in the MDP.

durations. Each worker is subject to repeated stochastic transitions between unemployment and employment, with transition probabilities determined jointly by costly and unobservable job search (or job retention) effort levels, and by the duration of the latest unemployment or employment spell. Since agents only receive wages when employed, the objective of the public unemployment insurance scheme is to provide a combination of wage taxes and unemployment benefits that maximizes the total social welfare by insuring the agents against income fluctuations, while at the same time providing incentives for the agents to exert a socially optimal level of job search or job retention effort. We define an agent’s “personal state” as the length of his latest unemployment or employment spell. Agents’ personal states then evolve according to a controlled Markov chain, whose transition probabilities depend on the agents’ equilibrium effort levels and on the length of their latest unemployment or employment spell, in the sense that a longer unemployment spell is associated with a lower probability of finding a job, conditional on a fixed job search effort, and similarly, a longer employment spell is associated with a lower probability of becoming unemployed, conditional on a fixed job retention effort.⁷ The equilibrium evolution of personal states induces an evolving heterogeneity in agents’ employment status and/or duration.

One restrictive aspect of the model we propose, which renders our analysis tractable, is that private agents are assumed to be anonymous, except for the identifiability of their personal states. As a consequence, at every given time period, the government will only be able to condition wage taxes and unemployment benefits on the agents’ personal states, as opposed to every agent’s individual infinite private employment history.⁸

We study three alternative notions of optimal policy in this heterogenous-agent environment. First, we analyze the case where the government lacks commitment to a social insurance plan. We characterize the private agent payoffs and social welfare levels that can be achieved in a sequential equilibrium of the resulting game as a function of the initial game state (i.e., the distribution of agents). We prove that the SSE payoff correspondence exists in the general policy game and that it can be found recursively. Second, we consider government commitment to some long-run, fixed tax-and-insure policy. This case is taken to be the extreme

⁷Although these assumptions yield a type of amnesia in transition probabilities that arises with any transition from unemployment to employment or vice versa, our modeling approach could in principle be extended to capture more realistic settings by defining a richer set of personal states that include longer (but still finite) memories of an agent’s employment status.

⁸This restriction is weakened by the fact that personal states do to some extent contain information about agents’ private employment histories, subject to the amnesia property of the agents’ transition probability functions. In addition, we believe that this assumption is quite realistic in as much as a legal framework for an unemployment insurance scheme will only be able to condition payouts on a finite number of circumstances. Note also that our analysis allows the government to condition wage taxes and unemployment benefits on the history of the distribution of personal states and on the history of the government budget.

opposite of the first case and it also has the interpretation of a class of simple policy rules.⁹ We prove that a unique steady state exists for every fixed policy, and that there exists an optimal fixed policy. Third, we study the standard public finance case with assumed government commitment to an ex-ante optimal Ramsey plan.¹⁰ We show that this is just a special restriction on the characterization of the SSE from the more general setting. In a computational example, we show that the expected social average discounted payoff under an optimal Ramsey policy cannot be sustained as an SSE. Finally, we show that an optimal fixed-policy steady state welfare outcome may not always be sustainable as a sequential equilibrium.

The illustrative framework we propose builds primarily on the literature initiated by [Shavell and Weiss \[1979\]](#) and extended by [Hopenhayn and Nicolini \[1997\]](#), which uses a recursive approach to characterize the optimal unemployment insurance contract as a solution to a repeated principal-agent problem. [Hopenhayn and Nicolini \[1997\]](#) consider the interaction between a principal, representing the insurance provider, and a single unemployed agent with unobservable job search effort and permanent employment opportunities, and derive the optimal combination of unemployment benefits and wage taxes for the case where the principal is able to commit to a contract. Since this problem is analyzed relative to a single agent, the determination of the optimal contract relies on an exogenously given promised utility level for the agent.¹¹

The main differences between our paper and the literature are twofold. First, the aggregate point of view of our model allows us to frame the principal’s objective as a welfare-maximization problem across a population of agents. This approach can be used to describe the expected utilities that can be attained across the different personal types of agents in equilibrium, given a budget balance requirement, without resorting to exogenously given promised utilities. Hence, our model can be used to determine “what can be achieved” by an unem-

⁹Simple policy rule exercises are commonly found in the monetary and fiscal policy literature in macroeconomics [see e.g., [Schmitt-Grohé and Uribe, 2007](#)].

¹⁰In this paper we also refer to this as the “variable commitment policy” regime. Hence “variable commitment policy” and “Ramsey policy” (with commitment) are to be taken as synonyms here.

¹¹The basic model of [Hopenhayn and Nicolini \[1997\]](#) has been extended in various directions. For example, [Kocherlakota \[2004\]](#) allows for the possibility of hidden savings, [Pavoni \[2007\]](#) requires the agent to receive a minimum level of discounted utility at every possible history, and [Pavoni and Violante \[2007\]](#) and [Pavoni \[2009\]](#) introduce human capital into the model, which depreciates during unemployment spells. Note that although our model does not explicitly consider human capital, our assumption that the probability of finding a job decreases during an unemployment spell, given a fixed job search effort, can implicitly be interpreted as the result of a depreciation in human capital.

The possibility of repeated unemployment spells, as assumed in our paper, is considered in [Zhao \[2001\]](#) and [Wang and Williamson \[1996, 2002\]](#). [Zhao \[2001\]](#) focuses on single agent problems and also introduces variations in earnings during employment spells. [Wang and Williamson \[1996, 2002\]](#) analyze an environment with overlapping generations of agents, which also allows for savings in their second paper, and characterize steady state allocations of expected utilities over a continuum of agents, subject to the requirement that each new labor-force entrant achieves a given level of promised expected utility.

ployment insurance program. Second, we provide a fully dynamic model that characterizes the payoffs that can be attained across agent types in any sequential equilibrium of the game we consider, for the case where the government cannot commit to an insurance policy, for the case where it can, and for steady states of our environment with government commitment. As a consequence, our model describes the aggregate evolution of personal states, and hence the evolution of the employment status across the population of agents, for the three environments we consider.

Comparison with “dual methods.” An alternative method to solve recursive incentive problems based on the dual problem corresponding to an initial optimal contracting problem, was developed by [Marcet and Marimon \[2011\]](#), and expanded by [Messner et al. \[2012\]](#) and [Messner et al. \[2013\]](#). While the corresponding methods are very useful in many settings, for our model, they do not provide an alternative to using the equilibrium-payoff-correspondence methods on which our analysis is based. For example, [Marcet and Marimon \[2011\]](#) assume a principal who can commit to a policy, and only derive the path corresponding to the resulting constrained-efficient policy—in contrast, our model assumes that the government cannot commit to a given policy, and derives all payoffs corresponding to some equilibrium outcome.

[Messner et al. \[2012\]](#) do show how problems where the principal cannot commit can be handled, but their approach relies on constraints that are additive in current actions and continuation payoffs, which is not the case with our hidden action game, where the agents’ effort levels are not observed by the principal and affect the transition probabilities to future personal states. The resulting separability between current rewards and future continuation payoffs assumed by [Messner et al. \[2012\]](#) allows a recursive characterization of the primal contracting problem, which is then dualized in order to derive an optimal solution. As a consequence, for problems without commitment, the non-deviation constraints corresponding to the endogenously determined space of equilibrium payoffs/promised utilities can be incorporated into the resulting recursive Bellman equation. This approach does not work in our model because the “effort optimality constraint” for the agents’ incentive problems—which presumably would be derived using a first-order approach—and the constraints that guarantee the necessary promised utilities, do not satisfy this separability. [Messner et al. \[2013\]](#) allow for such constraints to be non-separable, but their formulation does not include constraints that depend on the (endogenously determined) full set of equilibrium payoffs. Since they first derive the dual of the optimal contracting problem, and only then find a recursive representation, no obvious way seems to present itself that allows the constraints relating to limited commitment by a principal to be incorporated into this setting.

Roadmap. The remainder of this paper is organized as follows. In section 2, we outline the model environment. In section 3, we characterize the equilibria under the three policy environments we consider. In particular, in section 3.1, we define sequential equilibrium for our game and characterize the equilibrium payoff set as a function of relevant game states. In section 3.2, we consider the extreme case of the fixed policy environment. Here, we characterize steady states and discuss the nature of an optimal fixed policy. In section 3.3, we study the case of government commitment to ex-ante optimal Ramsey plans. We then discuss and explain our proposed BLP formulation for approximating the true upper-hemicontinuous equilibrium value correspondence in section 4. In section 5, we report our computational experience, and illustrate and compare the payoffs generated by equilibria under these three policy regimes. We conclude in section 6.

2 The Model

We begin by fixing ideas on the class of games we have in mind, using the public insurance game interpretation. The players of this game are given by a continuum $[0, 1]$ of heterogeneous private agents, and a single insurance fund planner, or in short, a government.

2.1 Transitions between private agent states

At each time period, every private agent is characterized by his *personal state*, which is drawn from the finite set $\mathcal{Z} := \{-N, \dots, -1, 1, \dots, M\}$, where $M, N \in \mathbb{Z}_+$. In this environment, positive states $j > 0$ denote an agent who has been working for j periods, and negative states $j < 0$ denote an agent who has been unemployed for $|j|$ periods. An agent in any employment state with duration $j > 0$ can only move to state $j+1$ or state -1 (i.e., one period unemployed), with the transition probabilities depending on an unobservable job retention effort exerted by the agent, and on his current state. Similarly, an agent in state $j < 0$ can only move to state $j-1$ or state 1 , with transition probabilities depending on an unobservable job search effort and on the current state of the agent. By assuming that the transition probabilities are constant after M periods for $j > 0$, or N periods for $j < 0$, we can use state M to characterize agents who have been employed for more than M periods, and state $-N$ to characterize agents who have been unemployed for more than N periods.

Let Y_t denote the state of an individual agent at time t , and let $a_t(j) \in [0, a^*] \subset \mathbb{R}_+$ denote effort exerted by an agent who is in state j in period t . Thus, in our example, effort $a_t(j)$ represents job-search intensity if $j < 0$, and work intensity if $j > 0$, so the upper bound a^* can be interpreted as a time constraint. The transition probabilities for private-agent states

are then characterized by a controlled Markov chain, where effort is the control variable. The transition probabilities of the corresponding Markov chains are defined for each state $j \in \mathcal{Z}$ by functions $p^j : [0, a^*] \rightarrow [0, 1]$, where

$$p^j(a) = \begin{cases} \mathbb{P}(Y_{t+1} = 1 \mid Y_t = j, a_t(j) = a), & \text{for } j < 0, \\ \mathbb{P}(Y_{t+1} = j + 1 \mid Y_t = j, a_t(j) = a), & \text{for } 0 < j < M, \\ \mathbb{P}(Y_{t+1} = j \mid Y_t = j, a_t(j) = a), & \text{for } j = M. \end{cases}$$

It follows that

$$1 - p^j(a) = \begin{cases} \mathbb{P}(Y_{t+1} = j \mid Y_t = j, a_t(j) = a), & \text{for } j = -N, \\ \mathbb{P}(Y_{t+1} = j - 1 \mid Y_t = j, a_t(j) = a), & \text{for } -N < j < 0, \\ \mathbb{P}(Y_{t+1} = -1 \mid Y_t = j, a_t(j) = a), & \text{for } j > 0. \end{cases}$$

We assume that for all $j \in \mathcal{Z}$, p^j is strictly increasing, strictly concave, and continuously differentiable, $p^j(0) = 0$ for $j < 0$ and $p^j(0) \geq 0$ for $j > 0$, and for every $a \in [0, a^*]$: (i) $p^j(a) < p^{j+1}(a)$; and, (ii) $dp^j(a)/da$ is increasing (decreasing) in j for $j < 0$ ($j > 0$). These assumptions provide a natural model for empirical conditional hazard rates of unemployment duration, for example.

For any function $w : \mathcal{Z} \rightarrow \mathbb{R}$, define $\mathbb{E}_{p^j(a)}[w] := \mathbb{E}[w(Y_{t+1}) \mid Y_t = j, a_t(j) = a]$. Denote the distribution over states in period t by λ_t , viewed as a row vector, and let $\Delta(\mathcal{Z})$ denote the set of distributions over \mathcal{Z} . We will consider symmetric equilibria where all agents having the same current personal state exert the same amount of effort. We refer to the effort levels as actions, and denote an action vector by $a_t \in \mathbb{R}^{\mathcal{Z}}$.¹² It follows that any fixed sequence of symmetric effort levels $\{a_t\}_{t=0}^\infty$ defines the evolution of the distribution of states, which will be deterministic in the aggregate, assuming that an appropriate law of large numbers holds [see Judd, 1985]. Given any fixed period t action vector a_t , denote the induced matrix of transition probabilities by $P(a_t)$. The period $t + 1$ distribution over personal states induced by a period t distribution λ_t and symmetric action vector a_t can then be computed as $\lambda_{t+1} = \lambda_t P(a_t)$.

2.2 The government and agent consumption insurance

For simplicity, suppose that all agents receive a constant per-period wage $m > 0$ during periods when they are in a state $j > 0$, and no wage in states $j < 0$. Agents have no opportunity to

¹²We use $\mathbb{R}^{\mathcal{Z}}$ to denote the set of real-valued functions with domain \mathcal{Z} . If the cardinality of \mathcal{Z} is finite, then $\mathbb{R}^{\mathcal{Z}}$ is the $|\mathcal{Z}|$ -dimensional Euclidean space. If $a \in \mathbb{R}^{\mathcal{Z}}$, we use $a(j)$ to denote the j -th coordinate of a , i.e., the action taken by an agent in state j .

save or borrow, and thus consume their entire income in each period. Agents in this world can only insure their consumption risk by participating in the public insurance scheme.

The government cannot observe effort levels, and can only observe agents' current personal states. The government is therefore restricted to state-contingent transfers, where taxes or insurance payouts are conditioned on the personal state of each agent. This requirement is equivalent to an anonymity assumption, in the sense that each agent's individual history of personal states is not observed by the government.¹³ The government can however observe the aggregate distribution over personal states at the beginning of each period t , and has a record of the evolution of these distributions $\lambda^t := (\lambda_0, \dots, \lambda_t)$. Thus, the government can choose its time t transfer policy as a function of the history λ^t . Since agents are distributed on a continuum, individual deviations have no effect on the evolution of the aggregate distribution over states, and therefore no effect on future government policies or other agents' actions.

We denote by $b_t(j)$ the period t net transfer that an agent in state j receives. A period t policy vector for the government is then given by a vector $b_t \in \mathbb{R}^Z$. A policy vector must satisfy $b_t(j) \geq -m$ for all $j > 0$, and $b_t(j) \geq 0$ for all $j < 0$. In addition, we assume that there is an exogenously given upper bound on the amount of benefits that the government can pay, i.e., there is a constant \bar{m} so that $b_t(j) \leq \bar{m}$ for all j and t . Letting $c_t^{b_t}(j)$ denote consumption in period t of an agent in state j , we get

$$c_t^{b_t}(j) = \begin{cases} m + b_t(j), & \text{if } j > 0, \\ b_t(j), & \text{if } j < 0. \end{cases}$$

The agents' common utility function is assumed to be separable in consumption and effort levels, so that the period t utility of an agent who is in state j can be expressed by $u(c_t^{b_t}(j)) - \phi(a_t(j))$, where u is non-negative, strictly increasing and concave and ϕ is non-negative, strictly increasing and strictly convex.

Each agent of type $j \in Z$ takes as given a sequence of policy vectors $\{b_t\}_t$, and chooses a sequence of actions $\{a_t(j)\}_t$ to maximize his own expected average discounted payoff (conditional on his current state j),

$$(1 - \delta) \mathbb{E}_{p^j(a(j))} \left\{ \sum_{t=0}^{\infty} \delta^t \left[u(c_t^{b_t}(j)) - \phi(a_t(j)) \right] \right\},$$

¹³The model could easily be extended to a more general framework, where transfers can be conditioned on any finite history of agent employment states. This could be achieved by redefining personal states to include a finite history of employment states. Doing this would yield a model where the insurance scheme provides a finer incentive structure, that does not only depend on the current employment state, but also on a finite number of previous employment states.

subject to aggregate law of motion $\lambda_{t+1} = \lambda_t P(a)$ and λ_0 given.

The government's period t net surplus from the insurance program is $-\sum_{j \in \mathcal{Z}} \lambda_t(j) b_t(j) \equiv -\lambda_t \cdot b_t$, where $\lambda_t \cdot b_t$ denotes the Euclidean inner product of the two vectors. Assuming that the government does not use alternative funds to subsidize the insurance scheme, feasibility of the transfer vector b_t then requires that $-\lambda_t \cdot b_t \geq 0$.¹⁴

The government's objective is to choose a sequence of policy vectors $\{b_t\}_t$ that maximizes the normalized expected discounted average utilities of the agents

$$\mathbb{E}_{\lambda_0} \left\{ (1 - \delta) \sum_{t=0}^{\infty} \delta^t \sum_{j \in \mathcal{Z}} \lambda_t(j) \left[u(c_t^{b_t}(j)) - \phi(a_t(j)) \right] \right\},$$

given the initial distribution over states λ_0 , and subject to the relevant constraints.

Depending on the value of λ_0 , we can interpret it as describing a situation where there is a recession, when unemployment is high, or a boom, when unemployment is low. Since the objective of the optimal public insurance scheme lies in maximizing the welfare over a population of agents, given the initial distribution λ_0 , we can determine the optimal welfare of any individual agent as a function of his initial private state endogenously. This is in contrast to principal-agent models where only a single agent is considered, and initial promised utilities are given exogenously, as in [Hopenhayn and Nicolini \[1997\]](#).

2.3 Game states, histories, and strategies

The model described above defines a dynamic game, where the continuation game at the beginning of each period t is characterized by λ_t . The distribution λ_t thus defines a *game state*, which changes over the course of play as a function of the players' actions. We restrict agents to use symmetric actions, and assume that an appropriate law of large numbers holds (as in [Judd \[1985\]](#) or [Uhlig \[1996\]](#)), in which case the game state evolves deterministically as a function of the agents' action vector a_t .¹⁵ The transition function for λ_t is then given by $\lambda_{t+1} = \lambda_t P(a_t)$, and the set of feasible government policies at period t is defined, as a function

¹⁴A per-period subsidy of $s \geq 0$ could easily be incorporated into the model by re-defining feasibility as $-\lambda_t \cdot b_t \geq -s$.

¹⁵To get a law of large numbers that guarantees a deterministic evolution of game states λ_t requires essentially a direct assumption that the probability measure that jointly governs the continuum of random variables describing the evolution of individual private-agent states satisfies such a law. [Judd \[1985\]](#) suggests that "economists assume that they have an extension of the Kolmogorov measure which satisfies the law of large numbers," which is common practice in the macroeconomics literature based on models with a continuum of agents.

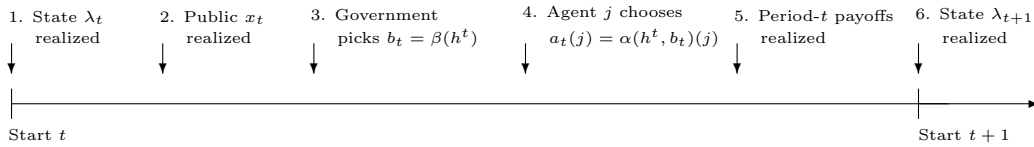
of the game state λ_t , by the following continuous compact- and convex-valued correspondence:

$$\mathcal{B}(\lambda_t) := \{b \in \mathbb{R}^Z \mid -\lambda_t \cdot b \geq 0, -m \leq b(j) \leq \bar{m} \text{ for } j > 0, 0 \leq b(j) \leq \bar{m} \text{ for } j < 0\}.$$

At the beginning of each period, all players can observe the realization of a public random variable $X_t \sim \text{i.i.d. } U[0, 1]$, and can condition their strategies on the history of past realizations of this variable, $x^t = (x_0, \dots, x_t)$. The use of such a public correlating device is a standard tool that convexifies the set of equilibrium payoffs [see, for example, [Mailath and Samuelson, 2006](#), [Judd et al., 2003](#), [Phelan and Stacchetti, 2001](#)]. Denote a public history at the beginning of period t by $h^t = (\lambda^t, x^t, b^{t-1})$, where $\lambda^t = (\lambda_0, \dots, \lambda_t)$ is the history of distributions over private agents' states, $x^t = (x_0, \dots, x_t)$ the realized history of the correlation variables, and $b^{t-1} = (b_0, \dots, b_{t-1})$ the policy history. A strategy for the government is a sequence of functions β_t , where each β_t maps a history h^t to a policy vector $b_t = \beta_t(h^t) \in \mathbb{R}^Z$. A symmetric strategy for the agents is a sequence of functions α_t , where each α_t maps a history (h^t, b_t) to an action vector $a_t = \alpha_t(h^t, b_t) \in \mathbb{R}^Z$. A strategy profile is defined by a pair $\sigma = (\beta, \alpha) := (\{\beta_t\}_t, \{\alpha_t\}_t)$. For any strategy profile σ and history h^t , we let $\sigma|_{h^t}$ denote the strategy induced by σ in the subgame following h^t .

Figure 1 summarizes the timing of information and actions in each period. The j -th component of $\alpha_t(h^t, b_t)$ specifies the action prescribed by the strategy α_t for an agent whose personal state Y_t is equal to j . Since agents are “small” players, individual unilateral deviations have no effect on the evolution of λ_t , and thus no effect on future play. Therefore, the private history of an agent's past actions and states will not affect his optimal action. We can thus ignore that agents could also condition their actions on their private histories.

Figure 1: Information and timing of actions



Any initial game state λ_0 and strategy profile σ recursively generate the following payoffs, where the subscript G denotes the government payoffs, i.e., the expected social welfare, and

a subscript j denotes the expected utility of an agent who starts out in state j :

$$V_G(\lambda_0, \sigma) = \sum_{j \in \mathcal{Z}} \lambda_0(j) V_j(\lambda_0, \sigma),$$

where

$$V_j(\lambda_0, \sigma) = \mathbb{E}_{\lambda_0, \sigma} \left\{ (1 - \delta) \sum_{t=0}^{\infty} \delta^t \left[u(c_t^{b_t}(Y_t)) - \phi(a_t(Y_t)) \right] \middle| Y_0 = j \right\}.$$

We use $V(\lambda_0, \sigma)$ to denote the vector containing the individual personal state-contingent expected payoffs and the government payoffs, so $V(\lambda_0, \sigma) \in \mathbb{R}^{\mathcal{Z} \cup \{G\}}$ for every (λ_0, σ) . Let $\overline{\mathcal{Z}} := \mathcal{Z} \cup \{G\}$. For any vector $v \in \mathbb{R}^{\overline{\mathcal{Z}}}$, we use $v(j)$ to denote the payoff of an agent in state j , and $v(G)$ to denote the government payoff.

2.4 A generalized public insurance game

Although our principal model is set up as an optimal unemployment insurance design problem, the framework we develop can be used to analyze other types of public insurance games. To define a corresponding general class of such games, we maintain the assumption of a continuum of private agents who are characterized at each period by a personal state j that is drawn from a finite set \mathcal{Z} , and must choose an unobservable action/effort variable a from a compact interval $[0, a^*]$. A generalized game is then defined by a family of Markovian transition matrices $P(a)$ for the private agents' personal state j , associated to effort levels a . The government chooses a policy that determines a transfer vector $b_t \in \mathbb{R}^{\mathcal{Z}}$ of period t transfers as a function of the agents' personal states, with the objective of maximizing the normalized discounted average utilities of the agents. Assuming that agents' per-period utilities are separable into a consumption/transfer component $u(j, b_t(j))$, which is determined by an agent's personal state j and corresponding government transfer $b_t(j)$ and replaces the consumption utility $u(c_t^{b_t}(j))$, and an effort component $\phi(a_t(j))$, where $a_t(j)$ is the effort exerted by a type- j agent in period t , our previous framework and subsequent analysis (including the computational results) can be applied almost directly to such a generalized model. With a finite state space \mathcal{Z} , game states can be defined by distribution vectors $\lambda_t \in \mathbb{R}^{\mathcal{Z}}$, and the bilinear interaction between game states and continuation payoffs, on which the computational approximation methods are based, still holds. Given appropriate boundedness and convexity assumption on u and ϕ , the principal generalization of our model thus results from the choice of the personal-state transition matrices $P(a)$. In order to apply the subsequent analysis, it only remains to check that the results that rely on specific properties of $P(a)$, which are mainly restricted to the steady state analysis associated with fixed transfer vectors, apply for a particular choice of

$P(a)$.

A potential alternative application would be the design of a public health insurance scheme. In this case, the private agents' effort levels could represent an index of lifestyle choices such as diet, exercise, and social habits. With suitable assumptions on payoffs and the Markov transition law on private-agent states, the agents' personal states could correspond to their levels of current health and thus determine their claim on, or contribution to, the public health insurance fund. The government policy would then contain a list of health insurance levies for the healthy and productive, and health insurance payouts for the infirmed.

3 Theoretical Analysis

In this section, we consider three policy environments. First, we study equilibria under the least restrictive assumption on the government's ability to commit in section 3.1. Here the government is not exogenously assumed to be able to commit. However, the requirements of our symmetric sequential equilibrium (SSE) in this game would encode endogenous incentives for the government to stick to its equilibrium strategy. Second, in section 3.2, we study the extreme opposite of this policy setting. Here, the government is assumed to be able to commit to a fixed policy vector. We provide a few useful characterizations relating to existence of optimal long-run equilibria, and, uniqueness of steady-state outcomes under any fixed policy. Third, we study the intermediate setting in section 3.3, where the government is also presumed to be able to commit, but to more flexible time-varying or Ramsey optimal policies. We show that the equilibrium characterization here is just a special case of the SSE characterization.

3.1 Limited commitment and sequential equilibria

In this section, we consider the case where the government is not able to commit to a policy sequence, and characterize the sequential equilibria of the corresponding game. Our restriction to symmetric strategies for the agents yields the following equilibrium definition:

Definition 1. A strategy profile σ is a *symmetric sequential equilibrium (SSE)* for λ_0 , if for all t , and all h^t accessible from λ_0 ,

- (i) $V_G(\lambda_t, \sigma|_{h^t}) \geq V_G(\lambda_t, \gamma, \alpha|_{h^t})$, where γ denotes any alternative government continuation strategy;
- (ii) For all j and b_t , $\alpha_t(h^t, b_t)(j)$ is an optimal action for a state j agent if the subsequent policies and agent actions are generated by the strategy $\sigma|_{(h^t, b_t)}$.

Since any individual agent is small, he cannot affect λ_{t+1} by changing his personal effort level. Thus, as long as α_t prescribes an optimal action for every private-agent state, we can ignore individual deviations by the agents. Thus, we only need to determine what happens after a deviation by the government. We follow [Abreu \[1988\]](#) and [Phelan and Stacchetti \[2001\]](#) and consider only extreme punishments that yield the lowest continuation equilibrium payoffs for the government.¹⁶ Since agents are small players, such punishments must generate optimal actions for the agents. Thus, in response to a deviation by the government, agents play an action profile \tilde{a} that generates a distribution over states $\lambda_{t+1} = \lambda_t P(\tilde{a})$, and an equilibrium continuation strategy for the government as a function of λ_{t+1} that yields the lowest feasible continuation equilibrium payoff for the government, and is such that \tilde{a} is optimal given the continuation equilibrium.¹⁷

We define the equilibrium value correspondence $\mathcal{V} : \Delta(\mathcal{Z}) \rightrightarrows \mathbb{R}^{\bar{\mathcal{Z}}}$ to be the set of payoff vectors that can be achieved in some SSE, as a function of the initial game state, i.e.,

$$\mathcal{V}(\lambda_0) = \{V(\lambda_0, \sigma) \mid \sigma \text{ is a SSE for } \lambda_0\}.$$

We show next that the correspondence \mathcal{V} can be characterized as the unique fixed point of a monotone set-valued operator, by adapting the approach of [Phelan and Stacchetti \[2001\]](#) and [Atkeson \[1991\]](#). These papers, in turn, extend the techniques of self-generation and factorization pioneered by [Abreu et al. \[1990\]](#) for repeated games. The proofs for our results follow standard methods. The consequence of these results is that in principle, the equilibrium correspondence can be computed recursively by applying the operator to a suitably defined initial correspondence.

Definition 2. Let $\mathcal{W} : \Delta(\mathcal{Z}) \rightrightarrows \mathbb{R}^{\bar{\mathcal{Z}}}$ be a compact- and convex-valued correspondence having the property that $w(G) = \sum_{j \in \mathcal{Z}} \lambda(j)w(j)$ for all $(\lambda, w) \in \text{graph}(\mathcal{W})$. A vector $(b, a, \lambda', w) \in \mathbb{R}^{\mathcal{Z}} \times \mathbb{R}^{\mathcal{Z}} \times \Delta(\mathcal{Z}) \times \mathbb{R}^{\bar{\mathcal{Z}}}$ is *consistent with respect to \mathcal{W} at λ* if

- (i) $b \in \mathcal{B}(\lambda)$
- (ii) $\lambda' = \lambda P(a)$;
- (iii) $w \in \mathcal{W}(\lambda')$;

¹⁶This assumption can be viewed as a modeling device that yields a simple characterization of all equilibrium payoffs—many of these equilibrium payoffs would also be obtainable based on less extreme punishments.

¹⁷Since we only consider symmetric equilibria, this also implies that all agents essentially coordinate on implementing a punishment strategy after a deviation by the government.

(iv) For all $j \in \mathcal{Z}$,

$$a(j) \in \operatorname{argmax}_{a'} \left\{ (1 - \delta) \left[u(c^b(j)) - \phi(a') \right] + \delta \mathbb{E}_{p^j(a')} [w] \right\}.$$

Definition 3. For $(\lambda, b) \in \Delta(\mathcal{Z}) \times \mathbb{R}^{\mathcal{Z}}$ such that $b \in \mathcal{B}(\lambda)$, let

$$\pi(\lambda, b) := \min_{(a', \lambda'', w')} \left[(1 - \delta) \sum_{j \in \mathcal{Z}} \lambda(j) [u(c^b(j)) - \phi(a'(j))] + \delta \sum_{j \in \mathcal{Z}} \lambda''(j) w'(j) \right],$$

subject to (b, a', λ'', w') is consistent with respect to \mathcal{W} at λ . Let $(\tilde{a}(\lambda, b), \tilde{\lambda}'(\lambda, b), \tilde{w}(\lambda, b))$ denote the solutions to the corresponding minimization problem. A vector $(b, a, \lambda', w) \in \mathbb{R}^{\mathcal{Z}} \times \mathbb{R}^{\mathcal{Z}} \times \Delta(\mathcal{Z}) \times \mathbb{R}^{\overline{\mathcal{Z}}}$ is said to be *admissible with respect to \mathcal{W} at λ* if

(i) (b, a, λ', w) is consistent with respect to \mathcal{W} at λ ;

(ii) $(1 - \delta) \sum_{j \in \mathcal{Z}} \lambda(j) [u(c^b(j)) - \phi(a(j))] + \delta \sum_{j \in \mathcal{Z}} \lambda'(j) w(j) \geq \sup_{b' \in \mathcal{B}(\lambda)} \pi(\lambda, b')$.

The payoff vector defined by an admissible vector (b, a, λ', w) at λ is given by

$$\begin{aligned} E_G(b, a, \lambda', w)(\lambda) &= (1 - \delta) \sum_{j \in \mathcal{Z}} \lambda(j) [u(c^b(j)) - \phi(a(j))] + \delta \sum_{j \in \mathcal{Z}} \lambda'(j) w(j), \text{ and} \\ E_j(b, a, \lambda', w)(\lambda) &= (1 - \delta) \left[u(c^b(j)) - \phi(a(j)) \right] + \delta \mathbb{E}_{p^j(a(j))} [w]. \end{aligned}$$

Note that $E_G(b, a, \lambda', w)(\lambda) = \sum_{j \in \mathcal{Z}} \lambda(j) E_j(b, a, \lambda', w)(\lambda)$. Let

$$\mathbf{B}(\mathcal{W})(\lambda) := \operatorname{co}\{E(b, a, \lambda', w)(\lambda) \mid (b, a, \lambda', w) \text{ is admissible with respect to } \mathcal{W} \text{ at } \lambda\},$$

where co denotes the convex hull of a set.

The following results, which yield the existence and uniqueness of a SSE payoff correspondence, are quite standard, and their proofs are relegated to Appendix A.

Lemma 1. $\mathcal{V}(\lambda_0)$ is a bounded subset of $\mathbb{R}^{\overline{\mathcal{Z}}}$ for every λ_0 . Furthermore, $\operatorname{graph}(\mathcal{V}) \subset \Delta(\mathcal{Z}) \times \mathbb{R}^{\overline{\mathcal{Z}}}$ is bounded.

Lemma 2. $\mathbf{B}(\mathcal{W}) \subset \mathbf{B}(\mathcal{W}')$ whenever $\mathcal{W} \subset \mathcal{W}'$.

Lemma 3. If \mathcal{W} has compact graph, then $\mathbf{B}(\mathcal{W})$ has compact graph.

Proposition 1 (Self-generation). *If $\mathcal{W}(\lambda) \subset \mathbf{B}(\mathcal{W})(\lambda)$ for all $\lambda \in \Delta(\mathcal{Z})$ and $\text{graph}(\mathcal{W})$ is bounded, then $\mathbf{B}(\mathcal{W}) \subset \mathcal{V}$.*

Proposition 2 (Factorization). *$\mathcal{V}(\lambda) \subset \mathbf{B}(\mathcal{V})(\lambda)$ for all λ , and \mathcal{V} has compact graph.*

Self-generation and factorization imply that \mathcal{V} is the largest fixed point of $\mathbf{B}(\mathcal{V})$, and that the recursive algorithm for the computation of value correspondences described in [Abreu et al. \[1990\]](#) and [Phelan and Stacchetti \[2001\]](#) can be used to compute \mathcal{V} . This algorithm starts with an initial correspondence \mathcal{W}_0 such that $\mathbf{B}(\mathcal{W}_0) \subset \mathcal{W}_0$ (the existence of such a correspondence follows from the arguments of Lemma 1), and inductively defines a decreasing (in the sense of set inclusion) sequence of non-empty correspondences \mathcal{W}_n , following Lemma 2. The set limit of this sequence is thus non-empty and satisfies self-generation, and is therefore equal to \mathcal{V} , as shown by the constructive argument in the proof of Proposition 1.

We call an equilibrium strategy corresponding to an initial state λ_0 an *optimal equilibrium* if it maximizes the government's payoff among all equilibria for λ_0 . Since \mathcal{V} is non-empty and compact-valued by the previous arguments, an optimal equilibrium exists for all λ_0 . Let $\mathcal{V}_G(\lambda)$ denote the projection of $\mathcal{V}(\lambda)$ onto the G coordinate, i.e., the government's component of the equilibrium payoff correspondence, and let $\bar{v}_G(\lambda) := \max_{v \in \mathcal{V}_G(\lambda)} v$. Then a strategy profile σ constitutes an optimal equilibrium if $V_G(\lambda_0, \sigma) = \bar{v}_G(\lambda_0)$.¹⁸

Note that if σ is an optimal equilibrium for the initial state λ_0 , then the continuation strategies $\sigma|_{(\lambda_0, x_0, \beta_0(\lambda_0, x_0), \alpha_0(\lambda_0, x_0))}$ need not yield optimal equilibrium strategies for the game described by an initial distribution given by $\lambda_1 = \lambda_0 P(\alpha_0(\lambda_0, x_0))$. This is because the agents' continuation payoffs corresponding to an optimal equilibrium for $\lambda_0 P(\alpha_0(\lambda_0, x_0))$ may prevent $\alpha_0(\lambda_0, x_0)$ from being optimal in the initial period, and a different effort vector may induce a second-period distribution that is different from $\lambda_0 P(\alpha_0(\lambda_0, x_0))$.

3.2 Steady states and fixed policies

We now consider the second policy environment. As an extreme limit of the model, it assumes that the government can commit, and that it commits to some fixed policy. A justification for this study might be that in reality, governments are institutionally restricted to simple policy rules. We define and characterize steady states, and analyze the evolution of the system when the government can commit to a fixed policy.

¹⁸[Phelan and Stacchetti \[2001\]](#) define *upper* and *lower boundaries* for equilibrium value correspondences, and define *best* and *worst equilibria* to be equilibria for which payoffs lie on the upper and respectively, lower boundary of the equilibrium value correspondence. Defining upper and lower boundaries in our model would involve maximizing, respectively minimizing government payoffs given an initial distribution and given a initial vector of agents' payoffs. Since the government payoff is just the expected value of the agents' payoffs using the initial distribution, it is uniquely defined by the specification of a payoff vector and initial distribution, and thus, the upper and lower boundaries would coincide.

Definition 4. A vector $(b, a, \lambda, v) \in \mathbb{R}^{\mathcal{Z}} \times \mathbb{R}^{\mathcal{Z}} \times \Delta(\mathcal{Z}) \times \mathbb{R}^{\overline{\mathcal{Z}}}$ is a *steady state* if:

- (i) $\lambda = \lambda P(a)$,
- (ii) For all $j \in \mathcal{Z}$, $a(j) \in \operatorname{argmax}_{a'} \{ (1 - \delta) [u(c^b(j)) - \phi(a')] + \delta \mathbb{E}_{p^j(a')} [v] \}$, and
 $v(j) = (1 - \delta) [u(c^b(j)) - \phi(a(j))] + \delta \mathbb{E}_{p^j(a(j))} [v]$,
- (iii) $v(G) = \sum_{j \in \mathcal{Z}} \lambda(j) v(j)$.

A steady state vector describes a path along which all variables characterizing our system are constant. In a steady state, agents are required to choose optimal effort levels, but no maximization of social welfare by the government is imposed. Each steady state does however define a unique corresponding level of social welfare, $v(G)$. Note that our definition does not require the government transfers to balance, in the sense that no restrictions are made on the corresponding per-period government budget surplus $-\lambda \cdot b$. Hence, a steady state is only feasible without additional subsidies as long as $-\lambda \cdot b \geq 0$. If this condition does not hold for a given steady state, a per-period subsidy $\lambda \cdot b$ is necessary to sustain the associated transfer vector.

We assumed that any vector of government transfers is bounded, so that $-m \leq b(j) \leq \overline{m}$ for $j > 0$, and $0 \leq b(j) \leq \overline{m}$ for $j < 0$, where \overline{m} is an exogenous constant. A transfer vector is *unconstrained feasible* if it satisfies these bounds. The following proposition shows that any unconstrained feasible transfer vector b defines a unique corresponding steady state vector.¹⁹

Proposition 3. *Given any unconstrained feasible transfer vector b , there exists a unique corresponding steady state vector (b, a^b, λ^b, v^b) .*

The elements of the steady state vector corresponding to a transfer vector b are continuous as a function of b :

Lemma 4. *v^b , a^b , and λ^b are continuous in b .*

Define a steady state (b, a^b, λ^b, v^b) to be an *optimal steady state* if it maximizes the corresponding social welfare $v^b(G)$ among all steady states for which b is unconstrained feasible and $-\lambda^b \cdot b \geq 0$. Thus, the steady state (b, a^b, λ^b, v^b) is optimal if b solves

$$\max_{b'} v^{b'}(G) \equiv \max_{b'} \sum_{j \in \mathcal{Z}} \lambda^{b'}(j) v^{b'}(j), \text{ subject to } b' \in \mathcal{B}(\lambda^{b'}).$$

Existence of an optimal steady state follows from an application of Weierstrass' Theorem, using Lemma 4:

Proposition 4. *An optimal steady state exists.*

¹⁹Note that the proofs for the following results are contained in Appendix A.

3.3 Variable commitment policies

In the previous section we showed that when the government can commit to a fixed transfer vector, the system will converge to a steady state. Note however that a government may wish to reconsider such a commitment, as a fixed policy will not, in general, maximize social welfare among all commitment strategies. We now briefly consider the case where the government can commit to any variable policy, i.e., it can commit to any strategy $\{\beta_t\}_t$, where each β_t maps the corresponding time t history $h^t = (\lambda^t, x^t, b^{t-1})$ to a policy vector b_t . Even though the strategy $\{\beta_t\}_t$ is fixed, the resulting policy vector b_t can vary over time, as opposed to the fixed policy vector associated with the steady states discussed in the previous section. We show that the set of payoff vectors that can be achieved through such a commitment policy can be characterized recursively, analogously to the case of SSE payoff vectors.

Commitment implies that the government will never deviate from the given strategy, and thus, that the policies specified by the strategy for histories that involve a deviation by the government will be irrelevant. Thus, without loss of generality, we can restrict our analysis to strategies that resort to autarky after any deviation by the government, where autarky is defined as a situation where the government makes no transfers, and agents choose actions that are optimal in the absence of government transfers. As before, individual deviations by the agents have no effect on the aggregate distribution of states, and can thus be ignored. Hence, we only need agents' strategies to be optimal given the evolution of states generated by these strategies and given the government's strategy. We can thus define an equilibrium with commitment as follows:

Definition 5. A strategy profile $\sigma = (\beta, \alpha)$ is a *symmetric sequential equilibrium with commitment (SSEC)* for λ_0 if for all t , h^t , and j , $\alpha_t(h^t, \beta_t(h^t))(j)$ is an optimal action for a state j agent if the subsequent policies and agent actions are generated by the strategy $\sigma|_{(h^t, b_t)}$.²⁰

To characterize the payoffs that can be achieved with commitment, we can use the same approach we used for the case without commitment. We start by defining a commitment value correspondence as the set of all payoff vectors that can be achieved with commitment:

$$\mathcal{V}^C(\lambda_0) = \{V(\lambda_0, \sigma) \mid \sigma \text{ is a SSEC for } \lambda_0\}.$$

²⁰Note that we assume that after any deviation by the government from the strategy β , both government and agents will play their autarky strategies, which will be optimal for the agents.

Given any compact- and convex-valued correspondence $\mathcal{W} : \Delta(\mathcal{Z}) \rightrightarrows \mathbb{R}^{\bar{\mathcal{Z}}}$ we define a map

$$\mathbf{B}^C(\mathcal{W})(\lambda) := \text{co}\{E(b, a, \lambda', w)(\lambda) \mid (b, a, \lambda', w) \text{ is consistent with respect to } \mathcal{W} \text{ at } \lambda\}.$$

The definition of $\mathbf{B}^C(\mathcal{W})$ just replaces admissibility with consistency in the definition of $\mathbf{B}(\mathcal{W})$. The results for the case with no commitment can then easily be adapted to prove self-generation and factorization for \mathcal{V}^C and \mathbf{B}^C . It follows that \mathcal{V}^C is the largest fixed point of \mathbf{B}^C . An analogous algorithm to the case of no commitment can then be used for the computation of the commitment value correspondence. An optimal commitment strategy for an initial state λ_0 will then maximize the government's payoff among all SSEC for λ_0 . As in the case without commitment, the compactness of the correspondence \mathcal{V}^C implies that such an optimal strategy exists.

Note also that every steady state defines a symmetric sequential equilibrium with commitment (SSEC). However, a constant strategy as defined by a steady state will only be a symmetric sequential equilibrium *without* commitment (SSE), if it also satisfies the government incentive constraint induced by the equilibrium value correspondence \mathcal{V} .

4 Approximation Methods and Computation

We approximate the equilibrium value correspondence \mathcal{V} using convex-valued step correspondences, as suggested by [Sleet and Yeltekin \[2000\]](#). Our computational contribution in this paper is in terms of constructing the levels of the step-correspondences using bilinear programming (BLP) formulations. A more detailed exposition of the descriptions in this section, along with implementation pseudocode and reference notes to actual codes are available from the authors' [online manual on computation](#) for this paper.²¹

We first describe the theoretical approach to approximating the SSE payoff correspondence in section 4.1. Then, in section 4.2, we discuss how this is implementable, and we explain our proposed BLP formulations.

4.1 Overview and conceptual preliminaries

To simplify the notation, let $D := \Delta(\mathcal{Z})$ denote the domain of any such value correspondence, i.e., the state space of our game. Note also that any given state λ and agent payoff vector $w \in \mathbb{R}^{\mathcal{Z}}$ determine a unique corresponding government payoff given by $v(G) := \lambda \cdot w$. We can

²¹Go to: http://people.anu.edu.au/timothy.kam/work/research/codes/_kam-stauber/_sphinx-doc/build/html/.

thus ignore the government payoff when defining the equilibrium value correspondences and their approximations, and restrict their codomain to \mathbb{R}^Z . Following [Sleet and Yeltekin \[2000\]](#), we take a partition of the state space D , and construct outer and inner approximations of the equilibrium value correspondence using step-correspondences that are constant on each partition element, and are described by convex polytopes as in [Judd, Yeltekin, and Conklin \[2003\]](#).

Let $\{Q_k \mid k = 1, \dots, K\}$ denote a partition of D , so $D = \bigcup_{k=1}^K Q_k$. As shown by [Sleet and Yeltekin \[2000\]](#), an upper hemicontinuous, compact- and convex-valued correspondence $\mathcal{W} : D \rightrightarrows \mathbb{R}^Z$ can be approximated by step-valued correspondences using the following procedures:

Letting

$$\omega_k^o(\lambda) := \begin{cases} \text{co} \bigcup_{\lambda' \in Q_k} \mathcal{W}(\lambda') & \text{if } \lambda \in Q_k, \\ \emptyset & \text{otherwise,} \end{cases}$$

the correspondence defined by $\mathcal{W}^o(\lambda) := \bigcup_k \omega_k^o(\lambda)$ defines an outer step-valued approximation of \mathcal{W} . Similarly, letting

$$\omega_k^i(\lambda) := \begin{cases} \bigcap_{\lambda' \in Q_k} \mathcal{W}(\lambda') & \text{if } \lambda \in Q_k, \\ \mathbb{R}^Z & \text{otherwise,} \end{cases}$$

the correspondence defined by $\mathcal{W}^i(\lambda) := \bigcap_k \omega_k^i(\lambda)$ defines an inner step-valued approximation of \mathcal{W} .²² Since the convex-valued approximations \mathcal{W}^o and \mathcal{W}^i are constant on each partition element Q_k , and there are finitely many partition elements, these approximations can be further approximated by constructing outer and inner approximations for the values $\omega_k^o(\lambda)$ and $\omega_k^i(\lambda)$ using convex polytopes as proposed by [Judd et al. \[2003\]](#). Let $\bar{\omega}_k^o(\lambda)$ and $\bar{\omega}_k^i(\lambda)$ denote the corresponding approximations, and let $\bar{\mathcal{W}}^o$ and $\bar{\mathcal{W}}^i$ denote the resulting correspondences. Then

$$\bar{\mathcal{W}}^i \subset \mathcal{W}^i \subset \mathcal{W} \subset \mathcal{W}^o \subset \bar{\mathcal{W}}^o.$$

Given the finiteness of the partition, and the fact that each such polytope is characterized by finitely many hyperplanes, the approximations $\bar{\mathcal{W}}^o$ and $\bar{\mathcal{W}}^i$ can be represented numerically. Clearly, the precision of these approximations increases as finer partitions are used or as the number of hyperplanes used to construct the corresponding polytopes is increased. For any correspondence $\mathcal{W} : D \rightrightarrows \mathbb{R}^Z$, define operators A^o and A^i by letting $A^o(\mathcal{W}) := \bar{\mathcal{W}}^o$ and $A^i(\mathcal{W}) := \bar{\mathcal{W}}^i$. Then, the operators A^o and A^i are monotone, as long as we use a fixed set of

²²Note that the use of a public correlation device in defining the SSE payoff correspondence implies that we can focus on the convex hull representations of approximants to the true SSE payoff correspondence. As is well-known in computational geometry, a convex hull object can be represented by finite data points.

directional vectors to construct the approximating convex polytopes. Thus, if $\mathcal{W} \subset \mathcal{W}'$, then $A^o(\mathcal{W}) \subset A^o(\mathcal{W}')$ and $A^i(\mathcal{W}) \subset A^i(\mathcal{W}')$.

However, given a correspondence $\mathcal{W} : D \rightrightarrows \mathbb{R}^{\mathcal{Z}}$, we want to find an outer (and subsequently, inner) approximation of $\mathbf{B}(\mathcal{W})$, for which the operators A^o and A^i are not directly applicable. For the computational exercise, we will specify \mathcal{W} as a step-valued correspondence, where each “step” is approximated by a convex polytope. For now however, let \mathcal{W} denote any correspondence. We show next how to construct an outer approximation of $\mathbf{B}(\mathcal{W})$, denoted by $\mathbf{B}^o(\mathcal{W})$, and described by a step-valued correspondence, where each “step” is defined by a convex polytope.

Since we will need to discretize the feasible sets of agent actions a_j and government policy vectors b for our numerical computation, let \tilde{A} and B denote the “discretized” finite sets of agent actions and government policy vectors, respectively, such that B only contains vectors b that are unconstrained feasible (so $-m \leq b(j) \leq \bar{m}$ for $j > 0$, and $0 \leq b(j) \leq \bar{m}$ for $j < 0$), and let $B(\lambda) := \{b \in B \mid -\lambda \cdot b \geq 0\}$.²³ Also, let $P^j(a_j)$ denote the vector of agent state transition probabilities for an agent in state j who chooses an action a_j , let $v_j(a, b) := u(c^b(j)) - \phi(a_j)$, and $v(a, b) := (v_j(a, b))_{j \in \mathcal{Z}}$. Consider the partition of D defined by sets Q_k , and let $H := \{h_1, \dots, h_L\} \subset \mathbb{S}^{|\mathcal{Z}|-1} := \{x \in \mathbb{R}^{\mathcal{Z}} : \|x\|_2 = 1\}$ denote a set of approximation subgradients, with h_l denoting an arbitrary element of H . For each $a \in \tilde{A}^{\mathcal{Z}} =: A$, define a correspondence $\tilde{\mathcal{W}} : D \times A \rightrightarrows \mathbb{R}^{\mathcal{Z}}$ by

$$\begin{aligned} \tilde{\mathcal{W}}(\lambda', a) := \{w \in \mathcal{W}(\lambda') \mid & \text{for all } j \in \mathcal{Z} \text{ and for all } a'_j \in \tilde{A}, \\ & \delta[P^j(a'_j) - P^j(a_j)] \cdot w \leq (1 - \delta)[\phi(a'_j) - \phi(a_j)]\}. \end{aligned}$$

Then $\tilde{\mathcal{W}}(\lambda', a)$ contains all continuation payoff vectors $w \in \mathcal{W}(\lambda')$ for which the action vector a is optimal. Note also that if $\mathcal{W}(\lambda')$ is defined as a convex polytope, then $\tilde{\mathcal{W}}(\lambda', a)$ is also a convex polytope, since the optimality conditions are given by a finite set of linear inequalities. Next, construct government punishment vectors $\tilde{\pi} = (\tilde{\pi}_k)_{k=1}^K$ and $\hat{\pi} = (\hat{\pi}_k)_{k=1}^K$, by letting

$$\pi(\lambda) := \max_{b \in B(\lambda)} \min_{a \in A} \min_{w \in \tilde{\mathcal{W}}(\lambda P(a), a)} \{(1 - \delta)\lambda \cdot v(a, b) + \delta[\lambda P(a)] \cdot w\},$$

and then defining $\tilde{\pi}_k := \min_{\lambda \in Q_k} \pi(\lambda)$ and $\hat{\pi}_k := \max_{\lambda \in Q_k} \pi(\lambda)$. The vectors $\tilde{\pi}$ and $\hat{\pi}$ yield punishments that are constant on each partition element Q_k —to get appropriate outer and inner approximations, we will use $\tilde{\pi}$ to compute the outer approximation, and $\hat{\pi}$ to compute

²³In theory, the government policy vector set B , and the agents' action profile set $A := \tilde{A}^{\mathcal{Z}}$ are continuous sets. On the computer, this would make the solution methods too complex as one has to deal with a family of very high-dimensional continuous optimization problems.

the inner approximation. Given a vector of agent actions a , a government policy vector b , and a vector of continuation payoffs w , the vector of agents' expected payoffs is defined by $E(a, b, w) := ((1 - \delta)v_j(a, b) + \delta P^j(a_j) \cdot w)_{j \in \mathcal{Z}}$.

Outer Approximation. We can now define the outer approximation $\mathbf{B}^o(\mathcal{W})$. For each $h_l \in H$ and each partition element Q_k , let

$$\begin{aligned} c_l^{o+}(k) &:= \max_{a \in A, b \in B} \max_{\lambda, w} \{h_l \cdot E(a, b, w)\}, \text{ subject to} \\ &\lambda \in Q_k \text{ and } \lambda \cdot b \leq 0, \\ &w \in \tilde{\mathcal{W}}(\lambda P(a), a), \\ &(1 - \delta)\lambda \cdot v(a, b) + \delta[\lambda P(a)] \cdot w \geq \tilde{\pi}_k, \end{aligned} \tag{1}$$

and define

$$\bar{\omega}_k^{o+}(\lambda) := \begin{cases} \bigcap_{l=1}^L \{z \in \mathbb{R}^{\mathcal{Z}} \mid h_l \cdot z \leq c_l^{o+}(k)\}, & \text{if } \lambda \in Q_k, \\ \emptyset, & \text{otherwise.} \end{cases}$$

Then $\mathbf{B}^o(\mathcal{W})(\lambda) = \bigcup_k \bar{\omega}_k^{o+}(\lambda)$. Clearly, \mathbf{B}^o is monotone, and $\mathbf{B}^o(\mathcal{W}) \supset \mathbf{B}(\mathcal{W})$. Thus, if $\mathcal{W} \supset \mathcal{V}$, then

$$\mathbf{B}^o(\cdots \mathbf{B}^o(\mathcal{W})) \supset \mathbf{B}(\cdots \mathbf{B}(\mathcal{W})) \supset \mathcal{V},$$

since

$$\mathbf{B}^o(\mathbf{B}^o(\mathcal{W})) \supset \mathbf{B}^o(\mathbf{B}(\mathcal{W})) \supset \mathbf{B}(\mathbf{B}(\mathcal{W})).$$

Therefore, repeatedly applying the operator \mathbf{B}^o to any initial set \mathcal{W} yields an outer approximation of the equilibrium value correspondence \mathcal{V} .

Inner Approximation. To derive an inner approximation $\mathbf{B}^i(\mathcal{W})$, define for each $h_l \in H$ and each partition element Q_k ,

$$\begin{aligned} V_l^{i+}(k) &:= \min_{\lambda \in Q_k} \max_{a \in A, b \in B(\lambda)} \max_w \{h_l \cdot E(a, b, w)\}, \text{ subject to} \\ &w \in \tilde{\mathcal{W}}(\lambda P(a), a), \\ &(1 - \delta)\lambda \cdot v(a, b) + \delta[\lambda P(a)] \cdot w \geq \hat{\pi}_k, \end{aligned} \tag{2}$$

For each k and l , let $(a_l^*(k), b_l^*(k), w_l^*(k))$ denote the vector that attains the level $V_l^{i+}(k)$ above, and define

$$z_l^+(k) := E(a_l^*(k), b_l^*(k), w_l^*(k)).$$

The points $z_l^+(k)$ can be interpreted as a set of vertices $Z(k) := \{z_l^+(k) \mid l = 1, \dots, L\}$, whose convex hull $\text{co}(Z(k))$ can be represented by a finite set of subgradients $H' = \{h'_1, \dots, h'_{L'}\} \subset \mathbb{R}^{\mathcal{Z}}$ and corresponding constants $C'(k) = \{c'_l(k) \mid l = 1, \dots, L'\}$, such that

$$\text{co}(Z(k)) = \bigcap_{l=1}^{L'} \{z \in \mathbb{R}^{\mathcal{Z}} \mid h'_l \cdot z \leq c'_l(k)\}.$$

Letting

$$\bar{\omega}_k^{i+}(\lambda) := \begin{cases} \bigcap_{l=1}^{L'} \{z \in \mathbb{R}^{\mathcal{Z}} \mid h'_l \cdot z \leq c'_l(k)\}, & \text{if } \lambda \in Q_k, \\ \emptyset, & \text{otherwise,} \end{cases}$$

yields an inner approximation $\mathbf{B}^i(\mathcal{W})(\lambda) := \bigcup_k \bar{\omega}_k^{i+}(\lambda)$. By definition, \mathbf{B}^i is a monotone operator, and $\mathbf{B}^i(\mathcal{W}) \subset \mathbf{B}(\mathcal{W})$. It follows that

$$\mathbf{B}^i(\mathbf{B}^i(\mathcal{W})) \subset \mathbf{B}^i(\mathbf{B}(\mathcal{W})) \subset \mathbf{B}(\mathbf{B}(\mathcal{W})),$$

and hence that

$$\mathbf{B}^i(\dots \mathbf{B}^i(\mathcal{W})) \subset \mathbf{B}(\dots \mathbf{B}(\mathcal{W})).$$

Since $\bigcap_n \mathbf{B}^n(\mathcal{W}) = \mathcal{V}$, this implies that $\bigcap_n [\mathbf{B}^i]^n(\mathcal{W}) \subset \mathcal{V}$.

4.2 Implementation and the BLP formulations

For the computational implementation of outer and inner approximations, we start with a correspondence \mathcal{W} described as a step correspondence defined by convex polytopes, which except for the initial step, would arise from the previous iteration of the respective approximations. Such a correspondence \mathcal{W} is defined by a collection of “levels” $C(k) = \{c_l(k) \mid l = 1, \dots, L'\}$, for $k \in \{1, \dots, K\} =: \mathbf{K}$, and directional vectors $H' := \{h'_1, \dots, h'_{L'}\} \subset \mathbb{R}^{\mathcal{Z}}$, such that $\mathcal{W}(\lambda) = \bigcup_k \bar{\omega}_k(\lambda)$, with

$$\bar{\omega}_k(\lambda) := \begin{cases} \bigcap_{l=1}^{L'} \{z \in \mathbb{R}^{\mathcal{Z}} \mid h'_l \cdot z \leq c_l(k)\}, & \text{if } \lambda \in Q_k, \\ \emptyset, & \text{otherwise.} \end{cases}$$

Note that if such a $\bar{\omega}_k(\lambda)$ is the outcome of a previous outer approximation, then H' , the set of directional vectors on which the characterization of $\bar{\omega}_k(\lambda)$ is based, is equal to the initial set H used in the derivation of the approximation. However, if $\bar{\omega}_k(\lambda)$ is the outcome of a previous inner approximation, H' may be distinct from H , since the previously explained approach yields the vertices of the inner approximation. We will focus on the outer approximation in

describing our computational implementation, and hence consider the case where $H' = H$. (The corresponding inner approximation requires minor modifications and is discussed in more detail in our [online manual](#).)

We choose the set of directional vectors H as a subset of $\mathbb{S}^{|\mathcal{Z}|-1} := \{x \in \mathbb{R}^{\mathcal{Z}} \mid \|x\|_2 = 1\}$, the unit $(|\mathcal{Z}| - 1)$ -sphere, where the l_2 -norm is given by $\|x\|_2 = \left(\sum_{j=1}^{|\mathcal{Z}|} x_j^2\right)^{1/2}$. We do so by deterministically generating “evenly-spaced” points on $\mathbb{S}^{|\mathcal{Z}|-1}$.²⁴

4.2.1 State-space, transitions and intersections

Next, we define the partition $\{Q_k \mid k \in \mathbf{K}\}$ of the state space $D = \Delta(\mathcal{Z})$ as a collection of K equal-area $(|\mathcal{Z}| - 1)$ -simplices. (An example with $K = 16$ is depicted in Figure 2.)

[Figure 2 about here.]

For any action profile $a \in A$ and $k \in \mathbf{K}$, define

$$P(a)(Q_k) := \{\lambda' \in D \mid \lambda' = \lambda P(a) \text{ for some } \lambda \in Q_k\}.$$

Then $P(a)(Q_k)$ is a subset of D that intersects at least one partition element $Q_{k'}$ with $k' \in \mathbf{K}$. Denote the set of all indices $k' \in \mathbf{K}$ for which $P(a)(Q_k) \cap Q_{k'} \neq \emptyset$ by $\mathbf{I}(a, k)$, so

$$\mathbf{I}(a, k) := \{k' \in \mathbf{K} \mid P(a)(Q_k) \cap Q_{k'} \neq \emptyset\}.$$

Finally, for any $k' \in \mathbf{I}(a, k)$, denote the set of all $\lambda' \in P(a)(Q_k) \cap Q_{k'}$ by $\text{Poly}_{k'(a,k)}$, so

$$\text{Poly}_{k'(a,k)} := \{\lambda' \in D \mid \lambda' \in P(a)(Q_k) \cap Q_{k'}\}, \text{ and } \bigcup_{k' \in \mathbf{I}(a,k)} \text{Poly}_{k'(a,k)} = P(a)(Q_k).$$

Each $\text{Poly}_{k'(a,k)}$ is a polytope in D that can be characterized using a finite collection of linear inequalities. These linear inequalities can be represented using a matrix $M_{k'(a,k)}$ and a vector $d_{k'(a,k)}$, so that $\lambda' \in \text{Poly}_{k'(a,k)}$ if and only if $M_{k'(a,k)}(\lambda')^T \leq d_{k'(a,k)}$, where $(\lambda')^T$ denotes the transpose of the (row) vector λ' .

²⁴In our examples below, we can construct this easily using standard Cartesian-coordinate transforms of spherical coordinates to distribute $l = 1, \dots, L$ subgradients $h_l \in \mathbb{S}^2 \subset \mathbb{R}^3$. Given spherical coordinates (r, θ_l, φ_l) , $l = 1, \dots, L$, with $r = 1$, $\theta_l \in [0, \pi]$, $\varphi_l \in [0, 2\pi)$, each point is $h_l = (x_l, y_l, z_l) = (\sin \theta_l \cos \varphi_l, \sin \theta_l \sin \varphi_l, \cos \theta_l)$. As is well-known in computational geometry, the equal-area sphere partitioning algorithm of [Leopardi \[2006\]](#) can be used to define these points when the hypersphere is embedded in \mathbb{R}^m where $m \geq 4$. Thus, higher dimensional problems can also be tackled by this extension.

4.2.2 Intersecting sub-state-spaces and BLP formulations

Once we have partitioned the state space and enumerated all possible transitions from each partition element (and their resulting intersections with other partition elements), we can begin constructing bilinear program (BLP) formulations that will ultimately yield the levels of upper hemicontinuous step correspondences.²⁵ We break the discussion of this construct into two parts: computing punishment values using *separable* BLP formulations, and, computing the overall SSE approximation as a *nonseparable* BLP, using a stochastic ϵ -global optimization approach. We discuss the case of the outer approximation scheme below. (A corresponding inner approximation scheme can be done and this is discussed in our [online manual](#).)

Punishment values and separable BLP. The computation of the punishment level $\tilde{\pi}_k$ can be rewritten as

$$\begin{aligned} \tilde{\pi}_k = \max_{b \in B} \min_{a \in A} \min_{k' \in \mathbf{I}(a, k)} & \left[\min_{\lambda, w} \{ (1 - \delta)\lambda \cdot v(a, b) + \delta[\lambda P(a)] \cdot w \} \right], \\ \text{s.t. } & \lambda \in Q_k \text{ and } \lambda \cdot b \leq 0, \\ & M_{k'(a, k)}(\lambda P(a))^T \leq d_{k'(a, k)}, \\ & h_l \cdot w \leq c_l(k'), \quad \forall h_l \in H, \\ & \delta[P^j(a'_j) - P^j(a_j)] \cdot w \leq (1 - \delta)[\phi(a'_j) - \phi(a_j)], \quad \forall j \in \mathcal{Z}, \quad \forall a'_j \in \tilde{A}. \end{aligned}$$

Note that the four sets of constraints all correspond to the minimization problem over λ and w , which is thus a bilinear optimization problem where the constraints are separable in the two variables λ and w . Such *separable bilinear programs* are proven to have a global solution [see [Bennett and Mangasarian, 1993](#)]. Furthermore, successive approximation using branching-and-bounding methods—i.e., branching into subsets of the optimizer domain, then bounding the value function below by the solutions of linear programs on each subset of the function domain, and, above by the value from a local nonlinear optimizer—can be used to find the ϵ -global optimum [[McCormick, 1976](#), [Bennett and Mangasarian, 1993](#), [Horst and Tuy, 1996](#)].²⁶

Correspondence levels and stochastic BLP. Given the punishment level $\tilde{\pi}$, the computation of the levels $c_l^{o+}(k)$ corresponding to the outer approximation can be rewritten as

²⁵Recall that these correspondences are then used in the successive approximations of each candidate SSE operator, which in the outer-approximation case is some correspondence $\mathbf{B}^o(\mathcal{W})$.

²⁶A detailed discussion of how we do this is found in Sections 8 and 9 of our [online manual on computation](#). Also an inner-approximation counterpart is discussed there as well, and the corresponding punishment value is denoted by $\hat{\pi}_k$.

follows, for every $k \in \mathbf{K}$ and every $h_l \in H$:

$$\begin{aligned}
c_l^{o+}(k) &:= \max_{a \in A, b \in B} \max_{k' \in \mathbf{I}(a, k)} \left[\max_{\lambda, w} \{h_l \cdot E(a, b, w)\} \right], \\
\text{s.t. } &\lambda \in Q_k \text{ and } \lambda \cdot b \leq 0, \\
&M_{k'(a, k)}(\lambda P(a))^T \leq d_{k'(a, k)}, \\
&\tilde{h}_l \cdot w \leq c_l(k'), \quad \forall \tilde{h}_l \in H, \\
&\delta[P^j(a'_j) - P^j(a_j)] \cdot w \leq (1 - \delta)[\phi(a'_j) - \phi(a_j)], \quad \forall j \in \mathcal{Z}, \quad \forall a'_j \in \tilde{A}, \\
&(1 - \delta)\lambda \cdot v(a, b) + \delta[\lambda P(a)] \cdot w \geq \tilde{\pi}_k.
\end{aligned}$$

Note that given each $(a, b) \in A \times B$, and conditioning on each reachable continuation state-subspace indexed by $k' \in \mathbf{I}(a, k)$ from Q_k , the objective function of the maximization problem over λ and w is linear (in w). However, the last constraint is not separable in λ and w , but is still bilinear in (λ, w) .

Since there are no known properties of global optima for nonseparable BLPs, unlike the case of separable BLPs, we propose the following stochastic algorithm: First, we utilize a modified Markov-chain Monte Carlo sampler called the Hit-and-Run Algorithm, originally due to [Smith \[1986\]](#), to create large uniform samples of λ , one sample set on each partition element Q_k .²⁷ Second, we check that these sample realizations yields nonempty sets $B(\lambda)$, i.e., that there exist non-empty selections of λ and b satisfying government budget feasibility, at each partition element of the state space. Third, given each realization $\lambda \in Q_k$, solve the maximum problem over w . Notice that in this step, we are in the world of standard linear programs (LP). Fourth, given the values of each LP at each draw λ , $\text{LP}(\lambda)$, we maximize over λ . Since the sampling of λ is finite on the computer, the fourth step is a straightforward table-lookup problem. Fifth, we maximize over the tuples $\langle a, b, k'(a, k) \rangle$, which is again a table-lookup problem since $A \times B \times \mathbf{I}(a, k)$ is a finite set (table).²⁸

Summary. In this section, we have described the key steps of our computational approach (focusing on the case of BLP-outer approximations). We showed that a new type of computation problems arise in our setting where endogenous agent-type distributions interact in a bilinear way with their continuation payoffs, either in terms of the objective function when

²⁷This algorithm has a desirable property that it can (globally) reach any point in any arbitrarily given bounded set in \mathbb{R}^n in one step. That is, there is a positive probability of sampling from any neighborhood in that set. Moreover, it is proven by [Lovász \[1999\]](#) that the Hit-and-Run sampler converges fast (in probability) to a uniform distribution on convex bodies (viz. here our Q_k). [Lovász and Vempala \[2003\]](#) note that this algorithm is the fastest in practice.

²⁸More detailed discussion of this section can be found in Section 10 of our [online manual](#).

constructing punishment values, or in term of the incentive constraints, when constructing candidate SSE correspondence approximants. Our contribution here is to demonstrate how to attack these relative mild nonlinear problems (i.e. the BLPs) using both deterministic and stochastic global optimization tools common in mathematical programming and operations research.

5 Computational Experiments and Calculated Example

Based on our previously described computational approach, we can now report a few computational experiments and discuss one of the examples we have computed. The purpose here is to (i) demonstrate that such classes of dynamic policy games with agent heterogeneity can be computed, albeit at some time cost, on a desktop computer, and (ii) illustrate some qualitative features of the approximate equilibrium payoff correspondence, and compare equilibrium welfare levels induced by alternative policy regimes.

Parametrization. First we make some functional assumptions regarding the agents' common preference representation, and the job-search technology embedded in the assumed transition probability functions $p^j : [0, a^*] \rightarrow [0, 1]$, $j \in \mathcal{Z}$. We assume that the per-period payoff of each agent $j \in \mathcal{Z}$ is given by

$$u(c) - \phi(a) = \ln(c + \underline{c}) - \frac{a^{1+\tilde{\phi}}}{1+\tilde{\phi}},$$

where we set $\underline{c} = 0.01$, and, $\tilde{\phi} \geq 0$.

The transition probability functions satisfying our assumptions on p^j are, respectively,

$$p^j(a) = \begin{cases} \frac{1 - 1/\exp\{a(j)\}}{|j|} & \text{if } j < 0, \\ 1 - \frac{1}{j \exp\{a(j)\}} & \text{if } j > 0. \end{cases}$$

We set the common discount factor as $\delta = 1/(1.04)$, implying an annual frequency in the model's time observation. Without loss of generality we normalize the exogenous wage as $m = 1$. We set the preference parameter $\tilde{\phi} = 0$ so that the disutility of effort per period is linear in a . We have also experimented with alternative values of $\tilde{\phi} \in \{0.2, 0.5, 1\}$ and the results do not vary in qualitative terms. The upper bound on transfers is set as $\bar{m} = 0.5 \times m$. We let the duration bounds be $M = 1$ and $N = 2$. This has the interpretation that the government policies are institutionally restricted to a definition of the long-term unemployed as those in states $j \geq 2$, and treating them as an identical class of agents. Likewise, the

current example does not distinguish between agents' tenure of employment.

In theory, the government action set and the agents' action profile set are continuous sets. On the computer, this would make the solution methods too complex as one has to deal with a family of very high-dimensional continuous optimization problems. We thus discretize the set of individual agent effort levels to contain finitely possible actions, $\tilde{A} \subset (0, a_H)$, where a_H is computed as a natural upper bound on agent actions.²⁹ For each state j , we approximate the set of government transfers $b(j)$ using a set of two, three or six elements, so that we either have $|A| = |B| = 2^3 = 8$ (**Case 0**), $|A| = 2^3 = 8$ and $|B| = 3^3 = 27$ (**Case 1**), or we also consider a finer approximation of both action sets A and B with $|A| = 5^3 = 125$ and $|B| = 6^3 = 216$ (**Case 2**). These assumptions, respectively, can be interpreted as people having discrete efforts, e.g. measured in fixed units of time, and the government being institutionally restricted to a finite tax/transfer menu.³⁰ Finally, since the space of game states $D := \Delta(\mathcal{Z})$ is a probability simplex, we construct equal area partitions elements Q_k of this simplex.³¹ As in [Sleet and Yeltekin \[2000\]](#), our aim is to preserve the upper hemicontinuity property of the theoretical SSE payoff correspondence \mathcal{V} . Hence, we need to construct upper hemicontinuous step correspondences that take on constant steps over each continuous partition element Q_k of D . For manageable computation times, we fix $K = 16$ in our desktop-computer examples. Suppose $\hat{\mathcal{W}}$ is a current estimate of the SSE payoff correspondence \mathcal{V} . Our tolerance for convergence is given by $\|\hat{\mathcal{W}} - \mathbf{B}(\hat{\mathcal{W}})\| < 10^{-3}$, where $\|\cdot\|$ is given by a metric implied by the Hausdorff norm.

²⁹ From the model, we know the lowest possible action for each $a(j)$ is zero. Theoretically, A is continuous, and the agents' payoff functions are continuously differentiable with respect to effort a . Using this, we consider the following (rather generous) upper bound: For each current agent $j > 0$ who exerts some highest effort $a_H(j) \in (0, \infty)$, there is a probability $p^j(a_H(j))$ of receiving consumption $m > 0$ next period. An extreme upper bound on this would be to receive income m forever, with average discounted continuation value of $u(m)$. With probability $1 - p^j(a_H(j))$, he becomes unemployed ($j = -1$) next period and receives no consumption forever, but exerts the maximal effort of a_H forever. Thus $a_H(j) = \arg \max_a \{-(1 - \delta)\phi(a) + \delta p^j(a)u(m) + \delta(1 - p^j(a))[u(0) - \phi(a_H)]\}$. Since the derivative $a \mapsto -\phi'(a)$ is constant (or increasing) with respect to a , and the derivative $a \mapsto (p^j)'(a)$ is decreasing in a , and these are continuous in $a \in \mathbb{R}_+$, then there is a unique interior solution $a_H(i) \in (0, \infty)$. We then define $a_H = \max_{j \in \mathcal{Z}} \{a_H(j) : j > 0\}$. In our example, with $M = +1$, $a_H = a_H(+1)$ and given our parametrisation, $a_H \approx 4.5$.

³⁰This was also done in the original implementation of [Phelan and Stacchetti \[2001\]](#). It is also common practice in large scale computations of Markov decision processes to discretize the decision-maker's action sets [see e.g. [Petrik and Zilberstein, 2011](#)].

³¹Our method is readily scalable to larger M , N , K , $|A|$ and $|B|$, and, given its amenability to parallel computation, larger and more serious examples can be solved using either distributed supercomputing facilities or through cloud computing, which is currently beyond the scope of this paper. See also footnote [33](#).

Table 1: Convergence Experience for BLP-Outer Scheme

	$ B $	$ A $	Average Time ^a (hours)	Iterations ^b (n)
Case 0	$2^{ \mathcal{Z} } = 8$	$2^{ \mathcal{Z} } = 8$	1.30	216
Case 1	$3^{ \mathcal{Z} } = 27$	$2^{ \mathcal{Z} } = 8$	5.23	163
Case 2	$6^{ \mathcal{Z} } = 216$	$5^{ \mathcal{Z} } = 125$	18.12	48

Notes: All cases were set with $K = 16$ and $|\mathcal{Z}| = N + M \equiv 3$.

a. Per iteration of $\mathbf{B}^i(\mathcal{W})$.

b. To convergence at tolerance $\epsilon = 0.001$. Cases 1 and 2 use their respective prior case’s convergent result as initial guesses.

5.1 Computational Experience

We first report and discuss our computational experience, and then describe some qualitative results from examples. We would like to note that the purpose of this section is not to claim any economic or quantitative-policy implications from our numerical exercises. The task here is to demonstrate that our new BLP formulation for approximating the SSE payoff correspondence is solvable on a desktop computer for reasonably small examples.³² The method we propose—although a nonlinear (i.e., bilinear) programming formulation—yields an algorithm still naturally parallelizable [as in the LP formulations of Judd et al., 2003]. Therefore, our method is readily scalable so that more realistic versions of these problems can be solved in reasonable time on supercomputing clusters or on cloud computing servers with a MATLAB distributed-computing interface.³³

Table 1 reports three cases of our computational experience, focusing on the case of an outer approximation operator $\mathcal{W}^o := \mathbf{B}^o(\mathcal{W})$ described earlier.

We can see that computing these solutions on a desktop computer are nontrivial in terms of time, but these computational costs are not completely infeasible. This is because our method, which relies on the Abreu et al. [1990]-style equilibrium value set/correspondence operator, is not immune to the curse of dimensionality. For example, in Case 1 it takes up to about 35 days to obtain convergence of the sequence $\{(\mathbf{B}^o)^n(\mathcal{W}_0)\}_{n \geq 0}$. Granted the time cost, in our experience, the convergence appears to be “smooth,” as evidenced by the distance between successive approximations $\{(\mathbf{B}^o)^n(\mathcal{W}_0)\}_{n \geq 0}$: see figure 3. However, since

³²We rely on a MacPro 3,1 (desktop machine from circa 2008), equipped with two Quad-core Intel Xeon 2.8 GHz (E5462) processors, 16Gb RAM, operating on OS10.8 and using the GNU gcc4.3 compiler for the GLPK Linear Programming Toolkit.

³³This is currently available for users in the U.S. and Europe, but unfortunately at the time of writing, the authors in Australia have no access to Amazon’s EC2 via MATLAB. The authors also have no existing opportunities to access a supercomputing server.

our approximation method is easily parallelizable, this will help to reduce the burden of the curse, especially with large-scale cloud computing possibilities. Further, given a convergent approximation W^o , we can also apply an inner approximation, similar to what is proposed by Judd et al. [2003].³⁴ The inner approximating step correspondence \mathcal{W}^i is obtained by shrinking the outer approximation W^o arbitrarily and then applying the previously described inner approximation algorithm. The advantage of \mathcal{W}^i is that at every state $\lambda \in D$, we have $\mathcal{W}^i(\lambda) \subset \mathcal{V}(\lambda)$ —i.e. it contains only SSE values, whereas $W^o(\lambda) \supset \mathcal{V}(\lambda)$ may contain payoff vectors that are not strictly SSE values [see Judd et al., 2003, for more discussion].

[Figure 3 about here.]

5.2 Numerical examples

To aid visualization, consider figure 4, the limit W^o from Case 2 in table 1, in terms of its extreme points (vertices). Figure 5 shows (in various “slices”) the graph of an inner step correspondence, $\text{graph}(\mathcal{W}^i)$.³⁵ It approximates the true lower-dimensional projection of the SSE correspondence graph given by $\text{graph}(\{\mathcal{V}(\lambda) \setminus \{V_G(\lambda)\} \subset \mathbb{R}^Z : \lambda \in D\})$. Each panel of the figure represents the step correspondence “slice” defined over a particular partition element of the state space, $Q_k \in D$. These inner-approximating correspondence slices are constructed after we have obtained the outer approximation for \mathcal{V} . If we pick any initial state λ from some subset Q_k of the state space D , then for any point $w \in \mathcal{W}^i(\lambda)$, its corresponding government payoff is $v(G) = \lambda \cdot w$. Using this fact, we can easily back out information about the set of SSE sustainable social welfare values.

Consider first, the policy regime assuming that the government follows an optimal fixed steady state policy. Note that we defined optimality for this regime by only considering the long run welfare induced by each invariant distribution attained at the respective limit steady state. Hence this concept of optimality does not take into account the transition towards the optimal steady state. Why consider an optimal steady-state fixed policy? First, we proved that it exists and induces a unique distribution across agent types (and therefore welfare) for any initial distribution (Propositions 4 and 3 respectively). Second, from a prescriptive view, a simple policy strategy is easy to understand for policymakers. Figure 7a reports the optimal distribution λ^b of agents over the individual state space \mathcal{Z} . If state $j = -2$ is interpreted as the “long-term unemployed” agents, then this says that the optimal long run policy induces

³⁴More details can be found in our online manual.

³⁵Graphically, the rendering has been done using Delaunay triangulation of extreme points of the convex hull $\text{co}(\mathcal{W}^i)$. Note that the use of Delaunay triangularization here is not part of the solution algorithm, but only for graphical/visual rendering purposes in the figure.

a small proportion of the long-term unemployed. In the long run with this optimal fixed policy there is a majority of employed individuals. The fixed simple policy, which induced this distribution of agents, is one which is quite generous to the short term unemployed, and, there is quite a sizable transfer from the employed to the unemployed. Also note that the subsidy to the unemployed tapers off very quickly for the longer-term unemployed. The intuition for this mechanism here is quite simple, given the simplicity of the policy: While agents are employed, they have to be taxed to support the welfare system. However, when they are unemployed they are well-insured initially, but then in order to incentivize agents to not remain unemployed for longer terms, the benefit transfer goes to almost zero. Thus agents have to exert more effort to increase their probabilities of entering the employment state, and, if they are employed, they have to work hard to improve their probability of remaining employed.

This is shown in Figure 7b. Under the fixed optimal steady-state policy, agents in unemployment states exert quite high effort. This is shown in Figure 8. The observations here appear qualitatively unchanged as we allow the simple government policy to take on more possible combinations, i.e., as we increased the cardinality of the government action profile set B (up to $|B| = 10^3 \equiv 1000$ vectors).

[Figure 7 about here.]

Given the value correspondence approximations, we can now compare the maximal ex-ante social welfare of different policy regimes, approximated over every state-space partition element $Q_k \in D$, $k = 1, \dots, K$. That is, we can compare the case with commitment to variable policies (i.e. SSEC as described in Section 3.3), with the case of SSE with limited commitment by the government (see Section 3.1), and also with the optimal long run commitment policy equilibrium welfare (in Section 3.2).

Figure 6 sums up this exercise using the outer approximation W° .³⁶ On the horizontal axis, each point k indexes each continuous sub-state-space Q_k . Note that the horizontal axis has no particular ordering, so no meaning should be read into it. The approximate best-SSE (\square) and worst-SSE ($*$) average total payoffs, summarized in terms of the government payoff, defined over each element Q_k , give a sense of the size of the SSE correspondence in the dimension of the government payoff $v(G)$. Likewise, we can compute the SSEC best ex-ante payoff (\diamond) for the government over each Q_k . Finally, for each candidate government policy profile $b \in B$, we have a unique steady state distribution λ^b and government payoff $v_b(G)$.

³⁶The reason for this is that if all commitment equilibria (SSEC or optimal long run policy) cannot yield social welfare values that are in W° —i.e. they cannot be credibly supported by some SSE—then for sure, they are not in the true SSE value correspondence \mathcal{V} .

There is an optimal steady state payoff with associated steady state distribution that lies in the partition element Q_1 . Since this optimal steady state only considers the long run welfare attained at the limit invariant distribution, we represent the corresponding welfare level as a line in figure 6.

What this numerical example illustrates is that the welfare from ex-ante optimal commitment policy equilibria (SSEC) is not sustainable under any SSE, and that the optimal steady state welfare may be sustainable for some initial distribution of agents.³⁷ This provides a counterexample to the presumption that optimal tax and transfer policy analysis can be done by just assuming commitment (either to Ramsey policies or simple rules), a practice common in the public finance literature.

Note however, we do not make this claim universally as it may well be that in alternative model environments commitment may be credibly sustainable.³⁸ Thus, one should always check whether commitment in policymaking is a justifiable assumption. Our proposed method here provides a first step in dealing with the situation when commitment assumptions may be untenable in a heterogenous agent setting, and computing SSE values require more than standard LP methods. However, this example also illustrates that some simple long run optimal policy can be sustained as a particular SSE. In particular, this has to be some SSE beginning from an initial game state that coincides with the long run state under the desired optimal simple policy.

[Figure 8 about here.]

Finally, given the SSE value correspondence approximation W^i , we can recursively construct sample strategies. (We relegate the discussion on how the SSE sampling algorithm works to Appendix B.) Suppose we are interested in starting out at some given initial state λ_0 with some highest possible initial social value at $v_0(G) := \lambda_0 w_0$ given.

Figures 9, 10 and 11 depict a length- T sample of an SSE outcome that sustains the maximal social value $v_0(G)$ at λ_0 , where we set $T = 20$. We show these outcomes, respectively, in terms of the sequence of agent-type distributions (λ_t), government transfers (b_t), and, average discounted payoff profiles (w_t), beginning from the same initial game state as that of the optimal long-run fixed policy regimes. (Note that this is just an arbitrary initial game state we begin from.) This particular SSE outcome (see Figure 9) has a short transition to what appears to be a long run steady state distribution of agents, where most agents are employed

³⁷That the welfare under the optimal steady state policy equilibrium may be lower than some worst SSE is always a possibility. This is because the optimal steady state fixed policy is very restrictive by its simple nature.

³⁸For example, in a growth model environment with sufficient policy instruments [e.g. Domínguez, 2007a] or with institutional policy implementation lags [e.g. Domínguez, 2007b].

(93.5% being in state $j = +1$), with very few people being short-term employed (0.03% in state $j = -1$) and the rest being long-term unemployed (6.47% in state $j = -2$). Note that agents are still individually transiting between these states with positive probability each period in the long run.

Supporting this outcome is a sequence of time-varying government tax/transfer schedules (see Figure 10) that features some positive tax on the employed, with a tapered redistribution of transfers to the unemployed—i.e., the transfer to those who are one-period unemployed is sizable but this falls quickly for those longer-term unemployed, at each date t along the sample path. (Recall a similar, but stationary feature of the policy in the simple fixed-policy regime.) Each realized (best) SSE outcome requires enforcing a current tax/transfer schedule (b_t) and agent’s best response efforts (a_t) (not shown) with the correct incentives or promised future values (w_{t+1}) to each class of agents, whilst respecting the SSE-admissibility requirements (e.g., government incentive compatibility). At each date t , these incentives are allocated in such a way that yields the same ex-ante value w_t to each type of agent (and thus is attaining the initial best social welfare value $v_0(G)$).

[Figure 9 about here.]

[Figure 10 about here.]

[Figure 11 about here.]

6 Concluding Remarks

We have extended the seminal techniques of Abreu et al. [1990] and Phelan and Stacchetti [2001] to a game between a large player and a nondegenerate and endogenous distribution of small optimizing players. We have also illustrated that such a game can be solved using a bilinear programming generalization of the techniques proposed in Judd et al. [2003], combined with the method of Sleet and Yeltekin [2000]. The contribution of Sleet and Yeltekin [2000] preserves upper hemicontinuity of the true equilibrium payoff correspondence. We then showed how the BLP formulations can be implemented. As in standard LP approaches, our method is also scalable and parallelizable, and hence larger problems can be dealt with via large-scale distributed computing.

Our computational method can potentially be adapted to solve Ramsey taxation problems with limited commitment in Aiyagari [1994]-type environments.³⁹ Consider a fixed

³⁹Aiyagari [1994] is a standard neoclassical production general equilibrium economy, in which agents are ex-post heterogeneous as a result of idiosyncratic productivity shocks and their non-state-contingent asset

tax/transfer policy vector $b := (b(j))_j$ in this model, where each policy $b(j)$ is conditioned on some asset (or income) bracket/class $[\underline{y}_j, \bar{y}_j]$, and there is a finite collection of these classes that make up the entire individual asset (or income) space of agents. If a utilitarian government can commit to a fixed policy vector, then solving the policy equilibrium can be done using the standard approach of dynamic programming and constructing an ergodic competitive equilibrium distribution of agent types [see e.g. [Heer and Maussner, 2005](#), [Huggett, 1993](#)]. Likewise, if the government is a planner who is assumed to be able to commit to a fixed strategy. In both cases, it would suffice to track the competitive equilibrium’s relative price of capital as a one-dimensional sufficient state variable for the government policy problems. However, when the Ramsey planner cannot be trusted to commit, one has to keep track of the history of the distribution of agents in terms of their (expected) marginal valuations of individual saving/capital. The latter would be a generalization of what is a scalar auxiliary state variable in [Phelan and Stacchetti \[2001\]](#). In this setting, we can keep track of the (approximate) distribution over these heterogeneous marginal valuations over finite partitions of the individual capital state space. This idea would be a generalization of what we have done in this paper, at an even higher level of dimensionality, which in practice is still finite.⁴⁰ Moreover, because the planner cannot commit, defining the policymaker’s welfare criterion and also incentive constraints would also involve (computationally) an approximate bilinear formulation (in terms of the utilitarian planner’s expected continuation payoffs). Given that our proposed method is not free of the curse of dimensionality, this is beyond our desktop computational ability for now. However, it may not be impossible if one has access to a massively distributed computing facility or if one can port LP and therefore BLP methods to the GPGPU level of a machine.⁴¹ We leave this conjecture open for future work.

References

- Dilip Abreu. On the theory of infinitely repeated games with discounting. *Econometrica*, 56(2): 383–396, 1988. Cited on page(s): [\[14\]](#)
- Dilip Abreu, David Pearce, and Ennio Stacchetti. Toward a theory of discounted repeated games with imperfect monitoring. *Econometrica*, 58(5):1041–1063, 1990. Cited on page(s): [\[1\]](#), [\[2\]](#), [\[14\]](#), [\[16\]](#), [\[29\]](#), [\[33\]](#)
- S Rao Aiyagari. Uninsured idiosyncratic risk and aggregate saving. *The Quarterly Journal of Economics*, 109(3):659–84, 1994. Cited on page(s): [\[33\]](#)

accumulation decisions. In fact this would be akin to the setting of [Phelan and Stacchetti \[2001\]](#) where agents also face idiosyncratic shocks, where there is an ex-post distribution over individual agents’ asset/capital levels.

⁴⁰Computationally, in a heterogeneous agent model with private saving/capital, distributions over agent wealth levels are approximated by empirical distributions constructed over finite partitions of what is truly a bounded continuum support.

⁴¹GPGPU stands for general-purpose graphical processing units.

- Andrew Atkeson. International lending with moral hazard and risk of repudiation. *Econometrica*, 59(4):1069–1089, 1991. Cited on page(s): [1], [14]
- Kristin P. Bennett and Olvi L. Mangasarian. Bilinear separation of two sets in n -space. *Computational Optimization and Applications*, 2(3):207–227, 1993. Cited on page(s): [3], [25]
- Claude Berge. *Topological Spaces*. Dover Publications, 1997. originally published by Oliver & Boyd, Edinburgh and London, 1963. Cited on page(s): [40]
- M. Chandraker and D. Kriegman. Globally optimal bilinear programming for computer vision applications. In *Computer Vision and Pattern Recognition, CVPR 2008. IEEE Conference*, pages 1–8, 2008. doi: 10.1109/CVPR.2008.4587846. Cited on page(s): [3]
- Begoña Domínguez. Public debt and optimal taxes without commitment. *Journal of Economic Theory*, 135(1):159–170, 2007a. Cited on page(s): [32]
- Begoña Domínguez. On the time-consistency of optimal capital taxes. *Journal of Monetary Economics*, 54(3):686–705, 2007b. Cited on page(s): [32]
- Zhigang Feng. Time consistent optimal fiscal policy over the business cycle. *Quantitative Economics*, 6(1):189–221, 2015. Cited on page(s): [2], [3]
- Zhigang Feng, Jianjun Miao, Adrian Peralta-Alva, and Manuel Santos. Numerical simulation of non-optimal dynamic equilibrium models. *International Economic Review*, 55(1):83–110, 2014. Cited on page(s): [2], [3]
- M. K. Jr. Fort. Essential and non essential fixed points. *American Journal of Mathematics*, 72(2):315–322, 1950. Cited on page(s): [43]
- Burkhard Heer and Alfred Maussner. *Dynamic General Equilibrium Modelling: Computational Methods and Applications*. Springer, Heidelberg, 2005. Cited on page(s): [34]
- Hugo A. Hopenhayn and Juan Pablo Nicolini. Optimal unemployment insurance. *Journal of Political Economy*, 105(2):412–438, 1997. Cited on page(s): [5], [10]
- Reiner Horst and Hoang Tuy. *Global Optimization: Deterministic Approaches*. Springer, 1996. Cited on page(s): [3], [25]
- Mark Huggett. The risk-free rate in heterogeneous-agent incomplete-insurance economies. *Journal of Economic Dynamics and Control*, 17(5-6):953–969, 1993. Cited on page(s): [34]
- Kenneth Judd, Şevin Yeltekin, and James Conklin. Computing supergame equilibria. *Econometrica*, 71(4):1239–1254, 2003. Cited on page(s): [2], [11], [20], [29], [30], [33], [44]
- Kenneth L. Judd. The law of large numbers with a continuum of IID random variables. *Journal of Economic Theory*, 35(1):19–25, 1985. Cited on page(s): [8], [10]
- John G. Kemeny and J. Laurie Snell. *Finite Markov Chains*. Springer, 1983. originally published by Van Nostrand, 1960. Cited on page(s): [43]
- Narayana R. Kocherlakota. Figuring out the impact of hidden savings on optimal unemployment insurance. *Review of Economic Dynamics*, 7(3):541–554, 2004. Cited on page(s): [5]
- Paul Leopardi. A partition of the unit sphere into regions of equal area and small diameter. *Electronic Transactions on Numerical Analysis*, 25:309–327, 2006. Cited on page(s): [24]

- László Lovász. Hit-and-run mixes fast. *Mathematical Programming Ser. A*, 86:443–461, 1999. Cited on page(s): [26]
- László Lovász and Santosh Vempala. Hit-and-run is fast and fun. Technical Report MSR-TR-2003-05, Microsoft Research, 2003. Cited on page(s): [26]
- George Mailath and Larry Samuelson. *Repeated Games and Reputations: Long Run Relationships*. Oxford University Press, New York, 2006. Cited on page(s): [11], [42]
- Albert Marcet and Ramon Marimon. Recursive contracts. CEP Discussion Paper No. 1055, London School of Economics, 2011. Cited on page(s): [6]
- Garth P. McCormick. Computability of global solutions to factorable nonconvex problems: Part I – convex underestimating problems. *Mathematical Programming*, 10:147–175, 1976. Cited on page(s): [3], [25]
- Matthias Messner, Nicola Pavoni, and Christopher Sleet. Recursive methods for incentive problems. *Review of Economic Dynamics*, 15(4):501–525, 2012. Cited on page(s): [6]
- Matthias Messner, Nicola Pavoni, and Christopher Sleet. The dual approach to recursive optimization: Theory and examples. Working paper, 2013. Cited on page(s): [6]
- Artyom Nahapetyan and Panos Pardalos. Adaptive dynamic cost updating procedure for solving fixed charge network flow problems. *Computational Optimization and Applications*, 39(1):37–50, 2008. ISSN 0926-6003. doi: 10.1007/s10589-007-9060-x. Cited on page(s): [3]
- Nicola Pavoni. On optimal unemployment compensation. *Journal of Monetary Economics*, 54(6):1612–1630, 2007. Cited on page(s): [5]
- Nicola Pavoni. Optimal unemployment insurance, with human capital depreciation, and duration dependence. *International Economic Review*, 50(2):323–362, 2009. Cited on page(s): [5]
- Nicola Pavoni and G. L. Violante. Optimal welfare-to-work programs. *Review of Economic Studies*, 74(1):283–318, 2007. Cited on page(s): [5]
- Marek Petrik and Shlomo Zilberstein. Robust approximate bilinear programming for value function approximation. *Journal of Machine Learning Research*, 12:3027–3063, 2011. ISSN 1532-4435. Cited on page(s): [3], [28]
- Christopher Phelan and Ennio Stacchetti. Sequential equilibria in a Ramsey tax model. *Econometrica*, 69(6):1491–1518, 2001. Cited on page(s): [1], [2], [3], [11], [14], [16], [28], [33], [34]
- Martin L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley Series in Probability and Statistics. Wiley-Interscience, 2005. Cited on page(s): [42]
- Stephanie Schmitt-Grohé and Martín Uribe. Optimal simple and implementable monetary and fiscal rules. *Journal of Monetary Economics*, 54:1702–1725, 2007. Cited on page(s): [5]
- Stephen Shavell and Laurence Weiss. The optimal payment of unemployment insurance benefits over time. *Journal of Political Economy*, 87:1347–1362, 1979. Cited on page(s): [5]
- Christopher Sleet and Şevin Yeltekin. On the computation of value correspondences. KGMS-MEDS, Northwestern University, 2000. Cited on page(s): [2], [3], [19], [20], [28], [33]
- Robert L. Smith. Efficient Monte-Carlo procedures for generating points uniformly distributed over bounded regions. *Operations Research*, 32:1296–1308, 1986. Cited on page(s): [26]

- Harald Uhlig. A law of large numbers for large economies. *Economic Theory*, 8(1):41–50, 1996. Cited on page(s): [10]
- Cheng Wang and Stephen D. Williamson. Unemployment insurance with moral hazard in a dynamic economy. *Carnegie-Rochester Conference Series on Public Policy*, 44:1–41, 1996. Cited on page(s): [5]
- Cheng Wang and Stephen D. Williamson. Moral hazard, optimal unemployment insurance, and experience rating. *Journal of Monetary Economics*, 49(7):1337–1371, 2002. Cited on page(s): [5]
- Rui Zhao. The optimal unemployment insurance contract: Why a replacement ratio? Working paper, University of Illinois at Urbana-Champaign, 2001. Cited on page(s): [5]

APPENDIX

A Omitted Proofs

A.1 Proof of Lemma 1

Proof. Since the vectors b_t are assumed to be bounded, each agent's payoff is bounded above by $u(m + \bar{m})$, and therefore, so is the government's payoff. For the lower bound, note that the upper bound on effort a^* yields an associated lower bound on the cost of effort $-\phi(a^*)$, which determines a lower bound on the payoffs of the agents and of the government. Since \mathcal{V} has compact domain, it follows that $\text{graph}(\mathcal{V})$ is a bounded subset of $\Delta(\mathcal{Z}) \times \mathbb{R}^{\bar{\mathcal{Z}}}$. \square

A.2 Proof of Lemma 2

Proof. Given any λ , every vector that is consistent w.r.t. \mathcal{W} must also be consistent w.r.t. \mathcal{W}' . Thus, $\pi(\lambda, b)$ calculated w.r.t. \mathcal{W}' cannot exceed $\pi(\lambda, b)$ calculated w.r.t. \mathcal{W} , for every $b \in \mathcal{B}(\lambda)$. It follows that vectors that are admissible w.r.t. \mathcal{W} are also admissible w.r.t. \mathcal{W}' , which implies the result. \square

A.3 Proof of Lemma 3

Proof. If \mathcal{W} has compact graph, then the graph of $\mathbf{B}(\mathcal{W})$ is bounded since one-period utilities and continuation payoffs, which are drawn from \mathcal{W} , are bounded. Next, we show that the correspondence $\tilde{\mathbf{B}}$, defined by

$$\begin{aligned} \tilde{\mathbf{B}}(\mathcal{W})(\lambda) := \{E(b, a, \lambda', w)(\lambda) \mid (b, a, \lambda', w) \text{ is admissible with} \\ \text{respect to } \mathcal{W} \text{ at } \lambda\}, \end{aligned}$$

has closed graph, which implies the desired result, since the convex hull of a closed set is also closed.

Consider a convergent sequence $(\lambda_n, v_n) \rightarrow (\lambda, v)$, such that $v_n \in \tilde{\mathbf{B}}(\mathcal{W})(\lambda_n)$ for all n . Then for each n , there exists a vector $(b_n, a_n, \lambda'_n, w_n)$ that is admissible with respect to \mathcal{W} at λ_n , and can be assumed without loss of generality to converge to a vector (b, a, λ', w) . Since $b_n \in \mathcal{B}(\lambda_n)$ and \mathcal{B} is a continuous correspondence, it follows that $b \in \mathcal{B}(\lambda)$. By continuity, $\lambda'_n = \lambda_n P(a_n)$ implies that $\lambda' = \lambda P(a)$. The strict convexity assumption we made imply that the arg max in the definition of consistency with respect to \mathcal{W} is unique, and hence continuous in continuation payoffs by the Maximum Theorem, so a must be the unique corresponding maximizer when the continuation payoffs are given by w . It follows that the vector (b, a, λ', w) is consistent with respect to \mathcal{W} at λ .

To complete the proof, it remains to show that (b, a, λ', w) is also admissible with respect to \mathcal{W} at λ . If we define $\bar{\pi}(\lambda) := \sup_{b' \in \mathcal{B}(\lambda)} \pi(\lambda, b')$, then this is equivalent to the requirement that

$$E_G(b, a, \lambda', w)(\lambda) \geq \bar{\pi}(\lambda).$$

The admissibility of $(b_n, a_n, \lambda'_n, w_n)$ with respect to \mathcal{W} at λ_n implies that

$$E_G(b_n, a_n, \lambda'_n, w_n)(\lambda_n) \geq \bar{\pi}(\lambda_n)$$

for all n , and the continuity of E_G that $E_G(b_n, a_n, \lambda'_n, w_n)(\lambda_n) \rightarrow E_G(b, a, \lambda', w)(\lambda)$. Admissibility of (b, a, λ', w) then follows if we can show that $\bar{\pi}$ is a lower semi-continuous function.⁴² To see this, note first that the minimization problem defining $\pi(\lambda, b)$ involves the minimization of a continuous function over a set of consistent vectors drawn from an upper hemi-continuous correspondence (where upper hemi-continuity follows from the arguments used above). Hence, Theorem 2 in Berge [1997, p.116] implies that π is a lower semi-continuous function. Then Theorem 1 in Berge [1997, p.115] implies that $\bar{\pi}$ must also be lower semi-continuous (since $\bar{\pi}$ is defined by maximizing a lower semi-continuous function over a set of vectors b' drawn from the continuous correspondence \mathcal{B}), and thus completes the proof. \square

A.4 Proof of Proposition 1

Proof. The proof proceeds by constructing, for every $\tau_0 \equiv (\lambda_0, v_0) \in \text{graph}(\mathbf{B}(\mathcal{W}))$, an equilibrium strategy σ^{τ_0} with $V(\lambda_0, \sigma^{\tau_0}) = v_0$. For every $\tau \equiv (\lambda, v) \in \text{graph}(\mathbf{B}(\mathcal{W}))$ and $x \in [0, 1]$, let $A(x, \tau) \in \mathbb{R}^{\mathcal{Z}} \times \mathbb{R}^{\mathcal{Z}} \times \Delta(\mathcal{Z}) \times \mathbb{R}^{\bar{\mathcal{Z}}}$ denote a vector that is admissible with respect to \mathcal{W} at λ and has the property that $\int_0^1 E(A(x, \tau))(\lambda) dx = v$, and use superscripts b, a, λ' and w for A to denote the corresponding components of the vector $A(x, \tau)$. By definition of $\mathbf{B}(\mathcal{W})$, such a function A exists for every $\tau \in \text{graph}(\mathbf{B}(\mathcal{W}))$ and $x \in [0, 1]$, and can be assumed to be measurable.

We define a strategy σ^{τ_0} recursively: At $t = 0$, let

$$\beta_0^{\tau_0}(\lambda_0, x_0) = A^b(x_0, \tau_0),$$

and

$$\alpha_0^{\tau_0}(\lambda_0, x_0, b_0) = \begin{cases} A^a(x_0, \tau_0), & \text{if } b_0 = \beta_0^{\tau_0}(\lambda_0, x_0), \\ \tilde{a}(\lambda_0, b_0), & \text{otherwise,} \end{cases}$$

⁴²A real-valued function f is lower semi-continuous (lsc) if $f(x) = \liminf_{y \rightarrow x} f(y)$, and is upper semi-continuous (usc) if $f(x) = \limsup_{y \rightarrow x} f(y)$

where the function \tilde{a} is defined in Definition 3. At $t = 1$, let

$$\beta_1^{\tau_0}(\lambda^1, x^1, b_0) = \begin{cases} A^b(x_1, A^{\lambda'}(x_0, \tau_0), A^w(x_0, \tau_0)), & \text{if } b_0 = \beta_0^{\tau_0}(\lambda_0, x_0), \\ A^b(x_1, \tilde{\lambda}'(\lambda_0, b_0), \tilde{w}(\lambda_0, b_0)), & \text{otherwise,} \end{cases}$$

and

$$\alpha_1^{\tau_0}(\lambda^1, x^1, b^1) = \begin{cases} A^a(x_1, A^{\lambda'}(x_0, \tau_0), A^w(x_0, \tau_0)), & \text{if } b_0 = \beta_0^{\tau_0}(\lambda_0, x_0) \text{ and} \\ & b_1 = \beta_1^{\tau_0}(\lambda^1, x^1, b_0), \\ \tilde{a}(\lambda_1, b_1), & \text{otherwise.} \end{cases}$$

Let $W_{t+1}(x^t, \lambda^t, b^t)$ denote the continuation payoffs after period t . Then

$$W_1(\lambda_0, x_0, b_0) = \begin{cases} A^w(x_0, \tau_0), & \text{if } b_0 = \beta_0^{\tau_0}(\lambda_0, x_0), \\ \tilde{w}(\lambda_0, b_0), & \text{otherwise,} \end{cases}$$

and

$$W_2(\lambda^1, x^1, b^1) = \begin{cases} A^w(x_1, A^{\lambda'}(x_0, \tau_0), A^w(x_0, \tau_0)), & \text{if } b_0 = \beta_0^{\tau_0}(\lambda_0, x_0) \text{ and} \\ & b_1 = \beta_1^{\tau_0}(\lambda^1, x^1, b_0), \\ \tilde{w}(\lambda_1, b_1), & \text{otherwise.} \end{cases}$$

We can continue in this way to recursively define continuation payoffs and corresponding strategies $\sigma^{\tau_0} = (\beta^{\tau_0}, \alpha^{\tau_0})$, to get

$$W_{t+1}(\lambda^t, x^t, b^t) = \begin{cases} A^w(x_t, \lambda_t, W_t(\lambda^{t-1}, x^{t-1}, b^{t-1})), & \text{if the government has} \\ & \text{never deviated,} \\ \tilde{w}(\lambda_t, b_t), & \text{otherwise,} \end{cases}$$

$$\beta_t^{\tau_0}(\lambda^t, x^t, b^{t-1}) = A^b(x_t, \lambda_t, W_t(\lambda^{t-1}, x^{t-1}, b^{t-1})), \text{ and}$$

$$\alpha_t^{\tau_0}(\lambda^t, x^t, b^t) = \begin{cases} A^a(x_t, \lambda_t, W_t(\lambda^{t-1}, x^{t-1}, b^{t-1})), & \text{if the government has} \\ & \text{never deviated,} \\ \tilde{a}(\lambda_t, b_t), & \text{otherwise.} \end{cases}$$

For any $\tau \equiv (\lambda, v) \in \text{graph}(\mathbf{B}(\mathcal{W}))$, we now show that $v = V(\lambda, \sigma^\tau)$. Since $v(G) = \sum_{j \in \mathcal{Z}} \lambda(j)v(j)$, and $V_G(\lambda, \sigma^\tau) = \sum_{j \in \mathcal{Z}} \lambda(j)V_j(\lambda, \sigma^\tau)$, it suffices to show that $v(j) = V_j(\lambda, \sigma^\tau)$ for all $j \in \mathcal{Z}$.

For every $\tau \in \text{graph}(\mathbf{B}(\mathcal{W}))$, let $b_x := A^b(x, \tau)$, $a_x := A^a(x, \tau)$, $\lambda'_x := A^{\lambda'}(x, \tau)$ and

$w_x := A^w(x, \tau)$. Denote continuation strategies following (λ, x, b_x) as $\sigma^\tau|_{\lambda, x, b_x} = \sigma^{\tau'}$. Then

$$v(j) = \int_0^1 \left\{ (1 - \delta) \left[u(c_x^{b_x}(j)) - \phi(a_x(j)) \right] + \delta \mathbb{E}_{p^j(a_x(j))} [w_x(i)] \right\} dx,$$

and

$$V_j(\lambda, \sigma^\tau) = \int_0^1 \left\{ (1 - \delta) \left[u(c_x^{b_x}(j)) - \phi(a_x(j)) \right] + \delta \mathbb{E}_{p^j(a_x(j))} [V_i(\lambda'_x, \sigma^{\tau'_x})] \right\} dx.$$

Subtracting the corresponding equations for any j , yields

$$|v(j) - V_j(\lambda, \sigma^\tau)| \leq \delta \sup_{(i, \lambda', w(i)) \in \text{graph}(\mathbf{B}(\mathcal{W}))} \left| w(i) - V_i(\lambda', \sigma^{\lambda', w}) \right|.$$

Since this equation must hold for all $(j, \lambda, v(j)) \in \text{graph}(\mathbf{B}(\mathcal{W}))$, and all payoffs are bounded by Lemma 1, we get

$$\sup_{(j, \lambda, v(j)) \in \text{graph}(\mathbf{B}(\mathcal{W}))} |v(j) - V_j(\lambda, \sigma^\tau)| \leq \delta \sup_{(i, \lambda', w(i)) \in \text{graph}(\mathbf{B}(\mathcal{W}))} \left| w(i) - V_i(\lambda', \sigma^{\lambda', w}) \right|.$$

Therefore, $v(j) = V_j(\lambda, \sigma^\tau)$ for all j and $\tau \equiv (\lambda, v) \in \text{graph}(\mathbf{B}(\mathcal{W}))$.

It follows that σ^{τ_0} defines a sequential equilibrium, once we show that there does not exist a profitable multi-round deviation for the government. But this is a consequence of a standard “one-shot deviation principle”, which can be derived along the lines of [Mailath and Samuelson \[2006\]](#). \square

A.5 Proof of Proposition 2

Proof. The closure of $\text{graph}(\mathcal{V})$, $\overline{\text{graph}(\mathcal{V})}$ defines a compact correspondence denoted by $\bar{\mathcal{V}}$. $\mathbf{B}(\bar{\mathcal{V}})$ is compact by Lemma 3. If $v \in \mathcal{V}(\lambda)$ for some λ , and $\sigma = (\beta, \alpha)$ is a corresponding equilibrium strategy, define $b(x) = \beta_0(\lambda, x)$, $a(x) = \alpha_0(\lambda, x, b(x))$, $\lambda'(x) = \lambda P(a(x))$, and $w(x) = V(\lambda'(x), \sigma|_{(\lambda, x, b(x))})$. Then for every $x \in [0, 1]$, $(b(x), a(x), \lambda'(x), w(x))$ is admissible with respect to \mathcal{V} at λ , and $v = \int_0^1 E(b(x), a(x), \lambda'(x), w(x)) dx$, which implies that $\mathcal{V} \subset \mathbf{B}(\bar{\mathcal{V}})$. Since $\mathbf{B}(\bar{\mathcal{V}})$ is compact, it follows that $\bar{\mathcal{V}} \subset \mathbf{B}(\bar{\mathcal{V}})$. Hence, $\bar{\mathcal{V}}$ is self-generating, so $\bar{\mathcal{V}} \subset \mathcal{V}$ and thus $\mathcal{V} = \mathbf{B}(\mathcal{V})$. \square

A.6 Proof of Proposition 3

Proof. Fix any feasible vector b . Given the assumptions we made regarding the transition probabilities $p^j(a)$ and the cost function $\phi(a)$, standard results from dynamic programming (see, for example, chapter 6 in [Puterman \[2005\]](#)) imply the existence of a unique value function $v^b : \mathcal{Z} \rightarrow \mathbb{R}$, and a unique corresponding vector of optimal actions a^b , that together satisfy

condition (ii) of Definition 4. Note that if the optimal effort level for any state j , $a^b(j)$, is *not* equal to zero, the uniqueness of $a^b(j)$ follows from the convexity assumption we made for ϕ and p^j .

We show next that the Markov chain defined by a^b must always have a unique ergodic set, which will also be regular.⁴³ If $a^b(-N) = 0$, $p^{-N}(a^b(-N)) = 0$, in which case $-N$ is absorbing and all other states are transient. This is a consequence of the assumption that $p^j(a^b(j)) < 1$ for all states j . Thus, $\{-N\}$ is a unique ergodic set. If $a^b(-N) > 0$, define \bar{j} to be the smallest positive state such that $p^{\bar{j}}(a^b(\bar{j})) = 0$, if such a state exists, and set $\bar{j} = M$ otherwise. Then $\{-N, \dots, -1, 1, \dots, \bar{j}\}$ is the unique ergodic set of the Markov chain. Since $p^{-N}(a^b(-N)) < 1$, it is also a regular ergodic set. Therefore, the Markov chain will have a unique invariant distribution λ^b , with $\lambda^b(j) > 0$ for all ergodic states j , and we can set $v^b(G) = \sum_{j \in \mathcal{Z}} \lambda^b(j) v^b(j)$. \square

A.7 Proof of Lemma 4

Proof. Let $T^b : \mathbb{R}^{\mathcal{Z}} \rightarrow \mathbb{R}^{\mathcal{Z}}$ denote the operator that defines the Bellman equation of the dynamic programming problem defined by the policy b . Thus, for every $v \in \mathbb{R}^{\mathcal{Z}}$,

$$T^b(v)(j) := \sup_{a'} \left\{ (1 - \delta) \left[u(c^b(j)) - \phi(a') \right] + \delta \mathbb{E}_{p^j(a')} [v] \right\}.$$

Standard results from dynamic programming imply that every T^b is a contraction mapping, and is therefore continuous and has a unique fixed point v^b . By Theorem 2 in Fort [1950], a unique fixed point must be an *essential fixed point*.⁴⁴ To show that v^b is continuous in b , it remains to show that for every $\varepsilon > 0$ there exists a $\zeta > 0$ such that $d^{sup}(T^b, T^{b'}) < \varepsilon$ whenever $d(b, b') < \zeta$, where d denotes the sup-metric on $\mathbb{R}^{\mathcal{Z}}$. But this follows from the continuity of the utility function u , and from the fact that given any vector $v \in \mathbb{R}^{\mathcal{Z}}$, we have $T^b(v)(j) - T^{b'}(v)(j) = (1 - \delta)[u(c^b(j)) - u(c^{b'}(j))]$.

Since v^b is continuous in b and the objective function in the corresponding Bellman equation has a unique solution for every state j , the Maximum Theorem implies that a^b is continuous in b .

λ^b is the unique fixed point of the linear map on $\Delta(\mathcal{Z})$ defined by $P(a^b)$, and is therefore an essential fixed point for every b . Since the functions p^j are continuous in actions, continuity of λ^b follows by a similar argument to the one used to prove the continuity of v^b . \square

⁴³See Kemeny and Snell [1983] for the terminology and results on finite Markov chains which are used in the remainder of the proof.

⁴⁴A fixed point y^* of a continuous function $f : Y \rightarrow Y$ defined on a compact metric space Y is an *essential fixed point*, if for every $\xi > 0$ there exists a $\varepsilon > 0$, such that every continuous function $g : Y \rightarrow Y$ with $d^{sup}(f, g) < \varepsilon$ has a fixed point y^{**} such that $d(y^*, y^{**}) < \xi$.

A.8 Proof of Proposition 4

Proof. Lemma 4 implies that λ^b and v^b are continuous in b , and hence so is the objective function. It remains to show that the constraint set $\{b' \mid b' \in \mathcal{B}(\lambda^{b'})\}$ is non-empty and compact. To show non-emptiness, set $\tilde{b}(j) = 0$ for $j < 0$ and $\tilde{b}(j) = -\varepsilon < 0$ for $j > 0$, with ε some small positive number. The proof of Proposition 3 shows that the optimal action profile corresponding to this \tilde{b} induces a Markov chain with a unique regular ergodic set of states, which is either equal to $\{-N\}$ or $\{-N, \dots, -1, 1, \dots, \bar{j}\}$, for some $\bar{j} > 0$. The resulting invariant distribution $\lambda^{\tilde{b}}$ clearly satisfies $-\lambda^{\tilde{b}} \cdot \tilde{b} \geq 0$, and hence \tilde{b} is an element of the constraint set.⁴⁵ To show compactness, note that since all elements of the constraint set must be unconstrained feasible, the constraint set is bounded. To see that it is also closed, let (b_n) be a sequence that is contained in the constraint set and converges to a vector b , which therefore must also be unconstrained feasible. Moreover, $-\lambda^{b_n} \cdot b_n \geq 0$ for all n , and by Lemma 4, $-\lambda^{b_n} \cdot b_n \rightarrow -\lambda^b \cdot b$, which implies that $-\lambda^b \cdot b \geq 0$. \square

B Simulating a sample SSE path

We adapt the inner-ray method of Judd et al. [2003] to construct sample outcome paths for SSE strategies. The pseudocode in Algorithm 2 summarizes the steps used in simulating random sample SSE paths of length $T + 1$, denoted by the strategy outcome pairs $\{(a_t, b_t)(\lambda_0, w_0)\}_{t=1}^T$, and, continuation state-promised-value pairs $\{(\lambda_{t+1}, w_{t+1})(\lambda_0, w_0)\}_{t=0}^T$, beginning from an initial game state and agents' value profile pair, (λ_0, w_0) . The simulations are possible given that we have found the limiting (inner approximant) correspondence $\mathcal{W}^i : D \rightrightarrows \mathbb{R}^Z$ that approximates the true but numerically unknown SSE payoff correspondence $\mathcal{V} : D \rightrightarrows \mathbb{R}^Z$, where $D := \cup_{k \in \mathbf{K}} Q_k$ is the game-state space, and, $\mathbf{K} := \{1, \dots, K\}$ is the index set of the finite-partition scheme of the game-state space D . (Recall that for each approximant $\mathcal{W}^i : D \rightrightarrows \mathbb{R}^Z$, there is a corresponding record of the punishment value function $\{\hat{\pi}_k : k \in \mathbf{K}\}$.) A slice of the inner approximant is the convex set $\mathcal{W}^i(\lambda) \equiv \mathcal{W}^i(Q_k)$ (if $\lambda \in Q_k$).

We can describe the facets of each polytope approximant $\mathcal{W}^i(Q_k)$ by linear inequalities $G_k w \leq c_k^g$, for each state-space partition element $Q_k \subset D$, indexed by $k \in \mathbf{K}$. Let $\{(G_k, c_k^g) : k \in \mathbf{K}\}$ be the family of all such facets over all partition elements. When we denote $w \in \mathcal{W}^i(Q_k)$ in the algorithms below, this implies the implementable set of linear constraints of the form: $G_k w \leq c_k^g$. Other notations used follow the discussion earlier in Section 4.

It is useful to first define an algorithmic function R that will be recurrently used below.

⁴⁵Note that in fact, $-\lambda \cdot \tilde{b} \geq 0$ for any λ .

First we compute a set of centroids (reference points), $\{w_k^o\}_{k \in \mathbf{K}}$ where each centroid w_k^o is the center of mass of a (convex) polytope slice $\mathcal{W}^i(Q_k)$. Then we construct, for a given $k \in \mathbf{K}$, a finite number of rays $\theta_m := (x - w_k^o) / \|x - w_k^o\| \in \Theta$, where $x \in \mathbb{R}^Z$. For any θ_m , $\lambda_t \in Q_k \leftrightarrow k$, $\hat{\pi}_k$ and feasible action sets $\tilde{A} \times B_k \ni (a, b)$, where $B_k := \{b \in B : b \in B(\lambda_t)\}$, we can evaluate the function R defined in Algorithm 1 below.

Algorithm 1 (the R function) has the purpose of finding the scalar length $\ell := L(a, b; k) > 0$ that would extend a particular ray θ_m in order to yield some continuation payoff vector $w := w(\theta_m, \ell, a, b)$ on the boundary of the payoff set $\mathcal{W}^i(Q_k)$ that enforces each action-policy profile (a, b) , such that (a, b, w) is admissible with respect to the SSE payoff correspondence. It may turn out to be that for an arbitrary direction/ray θ_m , the set $\{L(a, b; k) : (a, b) \in A \times \tilde{B}\}$ is empty. However, since we are dealing with a convex-valued and non-empty SSE payoff correspondence, there is bound to be some non-empty solution in some directions $\theta_{m'} \in \Theta$, given a dense enough set of spherical codes $\Theta \subset \mathbb{R}^Z$. Given a current initial state and payoff vector, by searching across the set Θ and evaluating the function R for each direction $\theta \in \Theta$, we would have reconstructed the set of extreme points of $\mathcal{W}^i(Q_k)$ by the inner-ray method. Along with each extreme point, we can store its corresponding pure-strategy action-policy-promised-value profile (a, b, w) . These extreme points may be used if randomization between them is required to support an arbitrarily given initial payoff vector w_0 . (We explain this next.)

ALGORITHM 1. The R function

Input: $(\theta_m, \lambda_t, k, \hat{\pi}_k, \tilde{A}, B_k, \mathcal{W}^i, w_k^o)$
Output: Set of lengths: $\{L^m(a, b) : (a, b) \in \tilde{A} \times B_k\}$

```

1 for  $(a, b) \in \tilde{A} \times B_k$  do
2    $k' \leftarrow \mathbf{I}(a, k)$                                      // Index of continuation state (see Section 4.2.1)
3    $u_t \leftarrow (1 - \delta)u \circ \phi(a, b)$                        // Normalized current payoff
4    $L(a, b) = \max_{\ell > 0, w} \left\{ \ell : \right.$ 
5     (i)  $w = \delta^{-1} [w_k^o + \ell \theta_m - u_t]$                    // Promises under possible randomization
6     (ii)  $w \in \mathcal{W}^i(Q_{k'})$                                    // Consistency w.r.t.  $\mathcal{W}^i$  at  $k'$ 
7     (iii)  $\lambda_t [u_t + \delta w] \geq \hat{\pi}_k$                          // Government incentive constraint
8     (iv)  $\delta [P^j(a'_j) - P^j(a_j)] w \leq (1 - \delta) [\phi(a'_j) - \phi(a_j)], \forall j \in \mathcal{Z}, a' \in \tilde{A} \}$  // Agents
9   if  $\nexists w$  satisfying (i)-(iv) then
10    |  $L(a, b) = -\infty$ 
11   end
12 end

```

Given the function R in Algorithm 1, now we can define the recursion that simulates a sample SSE outcome. This is detailed as a pseudocode in Algorithm 2. We summarize these details verbally as follows.

1. At date t , given initial payoff profile w_t and distribution λ_t :
 - Find the state-space partition element $Q_k \ni \lambda_t$.
 - Given Q_k find the appropriate slice of the inner-approximation of the SSE payoff correspondence $W^i(Q_k)$ and then define the ray, $\theta_t = (w_t - w_k^o) / \|w_t - w_k^o\|$, from the set's centroid w_k^o to the point w_t .
 - Find the scalar l_t^* that supports w_t with some admissible vector (a, b, w) .

This step is comprised of Lines 1-12 in Algorithm 1.

2. Check if $l_t^* = 1$, which implies that w_t itself is an extreme point supported by some admissible pure-strategy action profile (a_t, b_t) together with continuation payoffs w_{t+1} . (Lines 13-17 in Algorithm 1.)
3. If not, then, w_t is either interior or on the boundary of the set $W^i(Q_k)$, and w_t is a convex combination of the extreme points of $W^i(Q_k)$.
 - Since $W^i(Q_k)$ is represented by a convex polytope, approximate it by rays $\theta_m \in \Theta$.
 - For each ray θ_m , compute its extreme point and supporting pure-strategy action profile. Record these admissible vertices as the list $\vec{z} \equiv (\vec{\theta}, \vec{\ell}; w_k^o)$, and their corresponding pure-strategy action profiles in the lists of profiles, \vec{a} and \vec{b} . (Lines 18-34 in Algorithm 1.)
 - Find the (endogenous) set of weights $\vec{\rho}$ that would yield w_t as a weighted average of these extreme points in \vec{z} . This is a problem of finding a feasible vector $\vec{\rho}$ that satisfies a system of simultaneous equations with inequality constraints (since $\vec{\rho}$ must be a probability distribution over the vertices). This can be handled using a standard linear-programming simplex algorithm to find the feasible vector(s). Note that $\vec{\rho}$ may not be unique and we have to pick one arbitrarily. (See Line 35.)
 - Draw a uniform random number ν , and relative to the (cumulative sum of these) randomization weights in $\vec{\rho}$, select a vertex from the collection \vec{z} , where its supporting pure-strategy admissible selection is (a_t, b_t, w_{t+1}) . (See Lines 37-48 in Algorithm 1.)
 - Either Step 2 or Step 3 generates the outcome (a_t, b_t, w_{t+1}) . Given a_t , the next period game state is $\lambda_{t+1} = \lambda_t P(a_t)$.
 - Set $t + 1$ as t and repeat from Step 1 again until $t = T$, where T is the chosen horizon for simulating the SSE outcome path.

ALGORITHM 2. Simulating SSE given correspondence \mathcal{W}^i .

Input: $(\lambda_0, w_0, T), \mathcal{W}^i$
Output: SSE sample $\{(a_t, b_t)(\lambda_0, w_0)\}_{t=0}^T; \{(\lambda_{t+1}, w_{t+1})(\lambda_0, w_0)\}_{t=0}^T$

```

1  set
2  |    $t \leftarrow 0$ 
3  end
4  while  $t \leq T$  do
5  |   set
6  |   |    $k \leftarrow Q_k \ni \lambda_t$  // Current state-space partition element
7  |   |    $B_k \leftarrow \{b \in B : b \in B(\lambda_t)\}$  // Feasible policies
8  |   |    $w_k^o \leftarrow \text{CENTROID}(\mathcal{W}^i(Q_k))$  // Centroid reference point
9  |   |    $\theta_t \leftarrow (w_t - w_k^o) / \|w_t - w_k^o\|$  // Ray to initial total payoff
10 |  end
11 |   $\{L^t(a, b)\} \leftarrow R(\theta_t, \lambda_t, k, \hat{\pi}_k, A, B_k, \mathcal{W}^i, w_k^o)$  // Check if  $w_t$  extreme point: See Algorithm 1
12 |   $\ell_t^* \leftarrow \max_{b \in B_k} \left\{ \max_{a \in \bar{A}} \{L^t(a, b)\} \right\}$ 
13 |  if  $\ell_t^* = 1$  then
14 |   |   //  $w_t$  is an extreme point
15 |   |    $a_t \leftarrow \arg \max_{a \in \bar{A}} \{L^t(a, b)\}$ 
16 |   |    $b_t \leftarrow \arg \max_{b \in B_k} \{L^t(a_t, b)\}$ 
17 |   |    $w_{t+1} \leftarrow \delta^{-1} [(w_k^o + \ell_t^* \theta_t - (1 - \delta)u \circ \phi(a_t, b_t)]$ 
18 |  else
19 |   |   // Randomize over other extreme points
20 |   |   for  $\theta_m \in \Theta$  do
21 |   |   |    $\{L^m(a, b)\} \leftarrow R(\theta_m, \lambda_t, k, \hat{\pi}_k, A, B_k, \mathcal{W}^i, w_k^o)$  // Candidate extreme points: See Algorithm 1
22 |   |   |    $\ell_m^* \leftarrow \max_{b \in B_k} \left\{ \max_{a \in \bar{A}} \{L^m(a, b)\} \right\}$ 
23 |   |   |   if  $\ell_m^* > 0$  then
24 |   |   |   |    $a_{m^*} \leftarrow \arg \max_{a \in \bar{A}} \{L^m(a, b)\}$ 
25 |   |   |   |    $b_{m^*} \leftarrow \arg \max_{b \in B_k} \{L^m(a_{m^*}, b)\}$ 
26 |   |   |   |    $\theta_{m^*} \leftarrow \theta_m$  // Store and index admissible extreme points
27 |   |   |   |    $\ell_{m^*}^* \leftarrow \ell_m^*$ 
28 |   |   |   end
29 |   |   end
30 |   |   set
31 |   |   |    $(\vec{a}, \vec{b}, \vec{\theta}, \vec{\ell}) \leftarrow \{(a_{m^*}, b_{m^*}, \theta_{m^*}, \ell_{m^*}^*) : \ell_{m^*}^* > 0\}$  // Actions and rays supporting extreme points
32 |   |   |    $\vec{z} \leftarrow \{z_{m^*} : z_{m^*} = w_k^o + \ell_{m^*}^* \theta_{m^*} \text{ and } \ell_{m^*}^* > 0\}_{m^*=1}^{M^*}$  // Extreme points neighboring  $w_t$ 
33 |   |   |    $M^* \leftarrow |\vec{\theta}|$  // Number of extreme points
34 |   |   end
35 |   |    $\vec{\rho} \leftarrow \{\rho_m : w_t = \sum_{m=1}^{M^*} \rho_m z_m, \sum_{m=1}^{M^*} \rho_m = 1, \rho_m \in [0, 1], \forall m = 1, \dots, M^*\}$ 
36 |   |   // Endogenous weights on extreme points
37 |   |    $S \leftarrow \emptyset$  and  $N^* \leftarrow |\{\rho_{m^*} \in \vec{\rho} : \rho_{m^*} \in (0, 1]\}|$ 
38 |   |   for  $\rho_{m^*} \in \vec{\rho}$  do
39 |   |   |    $S \leftarrow S \cup \{\sum_{i \leq m^*} \rho_i\}$  // Cumulative sums over ordered  $\vec{\rho}$  elements
40 |   |   end
41 |   |   Draw  $\nu \sim U[0, 1]$  //  $w_t$  not extreme point: Randomize over vertices
42 |   |   if  $S_{n-1} < \nu \leq S_n$  then
43 |   |   |   set
44 |   |   |   |    $(\vec{a}_n, \vec{b}_n, \vec{\theta}_n, \vec{\ell}_n) \leftarrow n$ 
45 |   |   |   |    $(a_t, b_t) \leftarrow (\vec{a}_n, \vec{b}_n)$  // Ex-post selection
46 |   |   |   |    $w_{t+1} \leftarrow \delta^{-1} [(w_k^o + \vec{\ell}_n \vec{\theta}_n - (1 - \delta)u \circ \phi(a_t, b_t)]$ 
47 |   |   |   end
48 |   |   end
49 |  end
50 |  set
51 |  |    $\lambda_{t+1} \leftarrow \lambda_t P(a_t)$ 
52 |  |    $t \leftarrow t + 1$ 
53 |  end
54 end

```

FIGURES IN THE PAPER

Figure 2: Example partition of state space $D = \cup_{k=1}^K Q_k$, $K = 16$

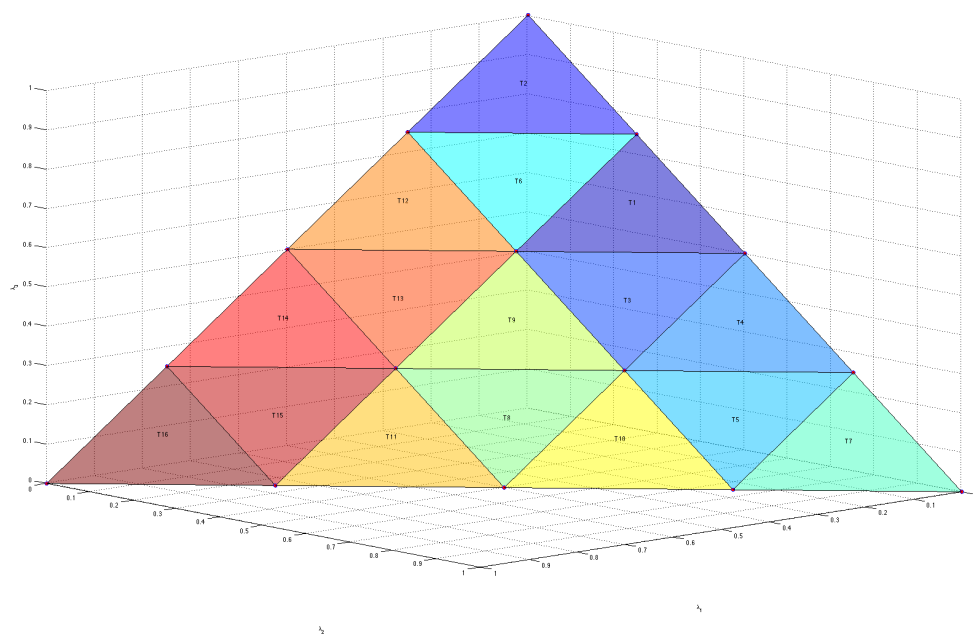


Figure 3: Convergence experience in Case 1 with norm $\|\mathbf{B}_n^o(\mathcal{W}_0) - \mathbf{B}_{n-1}^o(\mathcal{W}_0)\| \rightarrow 0, n \rightarrow \infty$. The horizontal axis measures the number of iterations n and the vertical axis is the norm between successive approximating step correspondences.

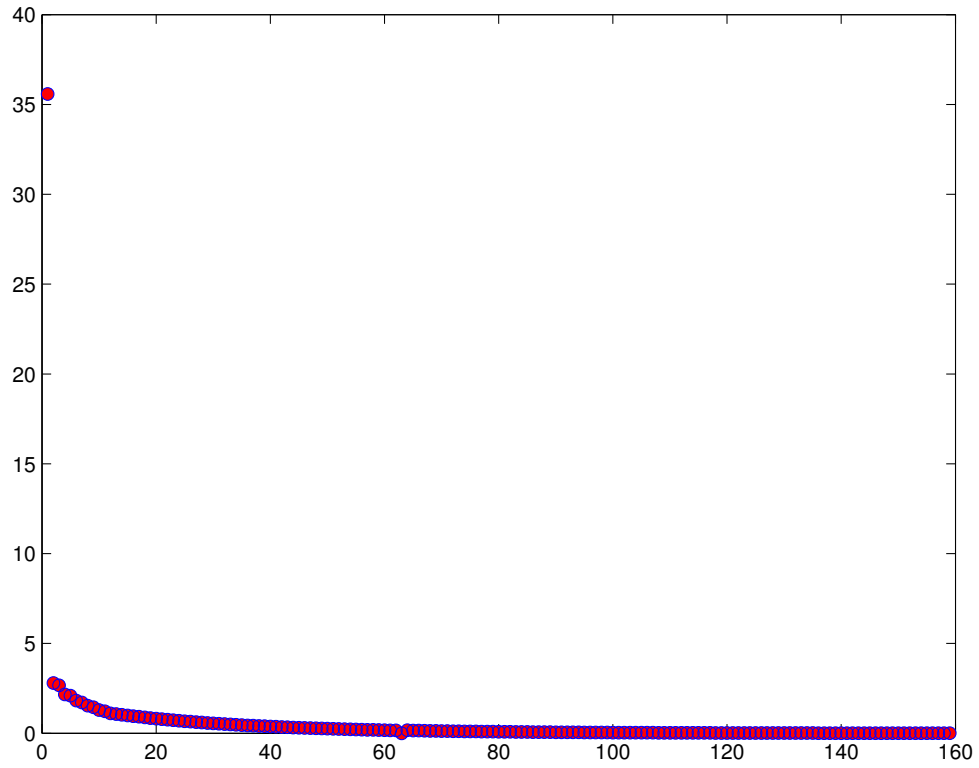


Figure 4: Extreme points of outer step correspondence approximation of \mathcal{W}^o , projected onto the agents payoff vector space. Each panel is a “slice” of the step correspondence defined over a state-space partition element Q_k . The dots (\cdot) are numerically calculated centroids of each convex polytope.

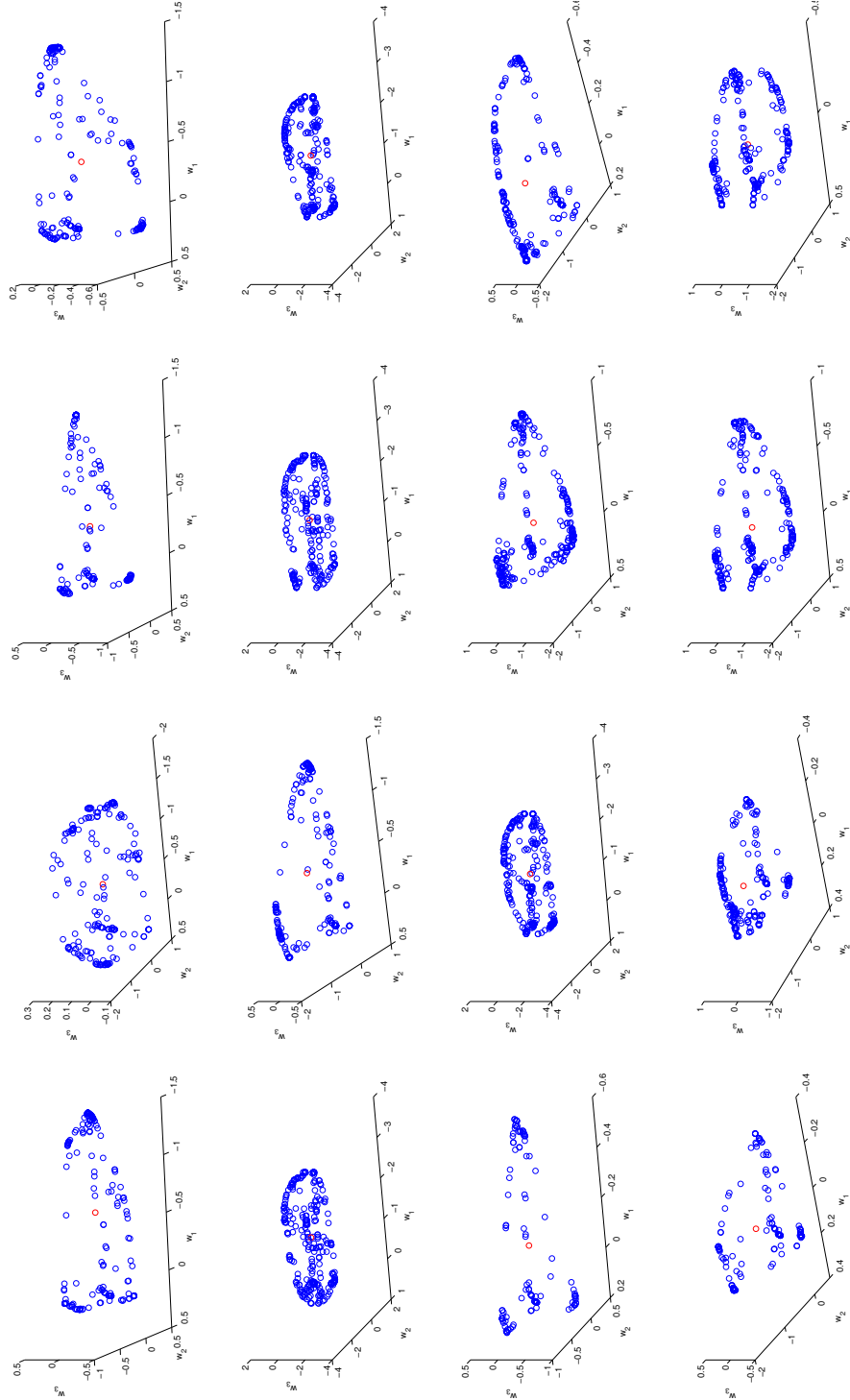


Figure 5: Convex hull of inner step correspondence approximation of \mathcal{W}^i , projected onto the agents payoff vector space. Each panel is a “slice” of the step correspondence defined over a state-space partition element Q_k . The dots (•) are numerically calculated centroids of each convex polytope.

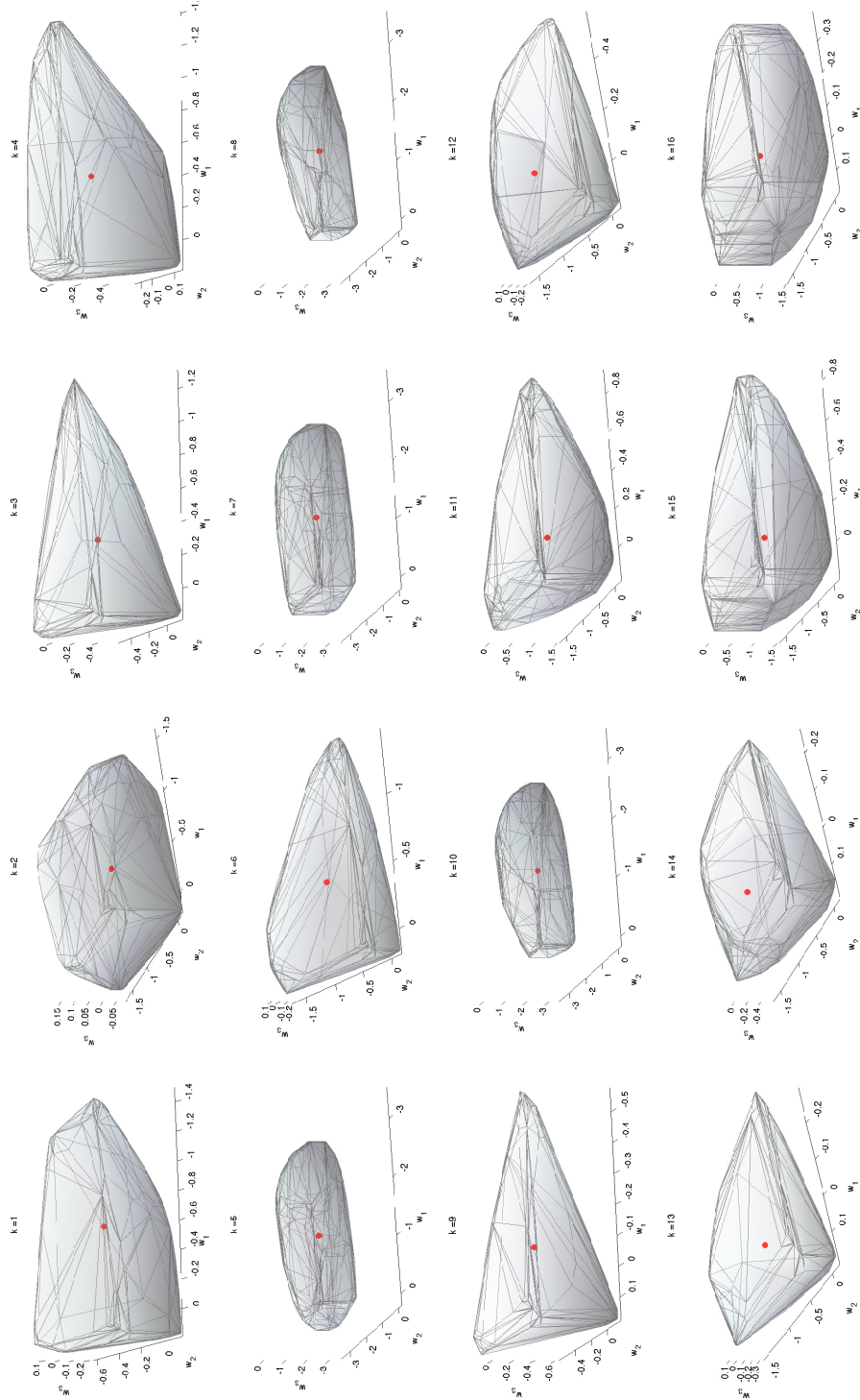


Figure 6: SSE vs. Ramsey vs. Optimal Steady State Welfare over initial state-space partition elements. The horizontal axis indexes the state-space partition elements Q_k in no particular order. The vertical axis measure welfare in terms of average expected discounted payoffs.

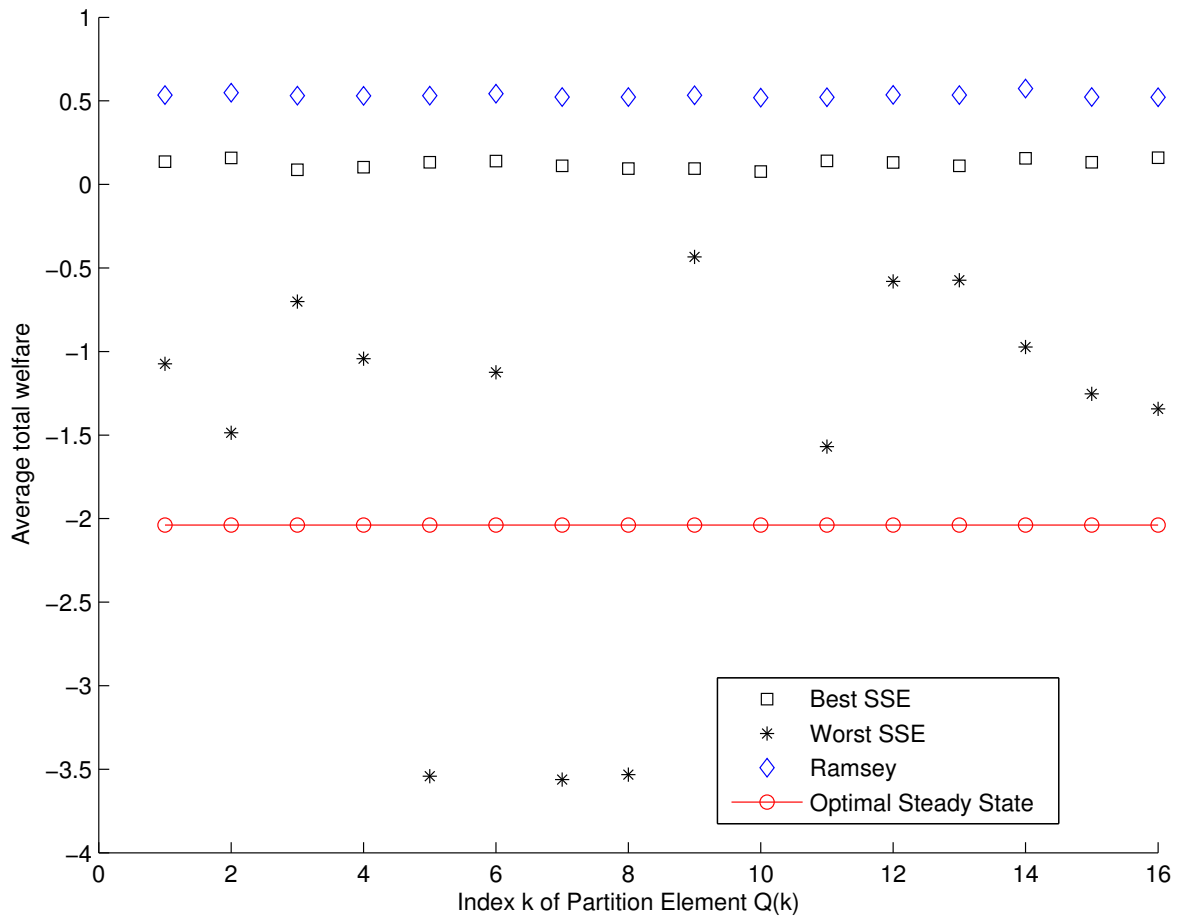
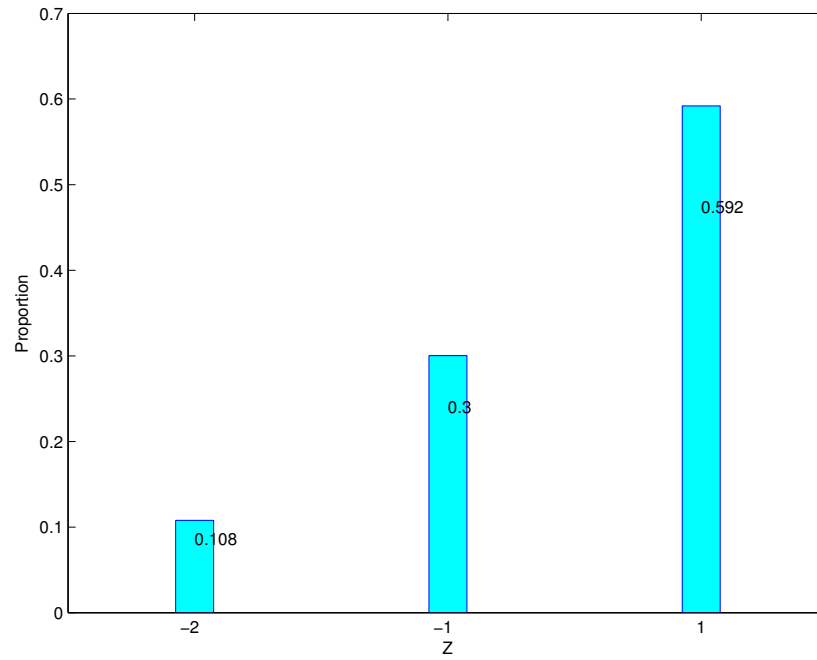


Figure 7: Optimal Steady-state commitment outcomes.

(a) Distribution of agents



(b) Policy

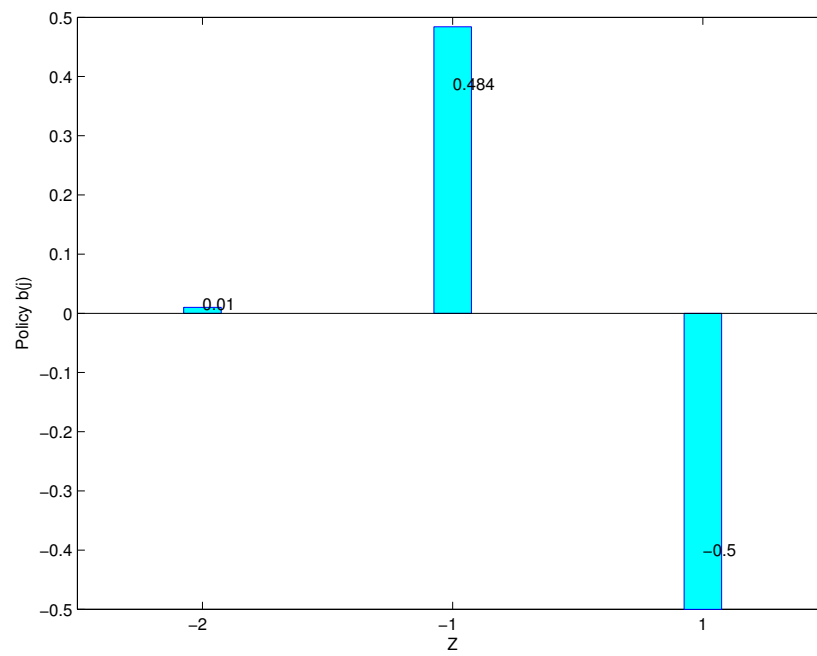
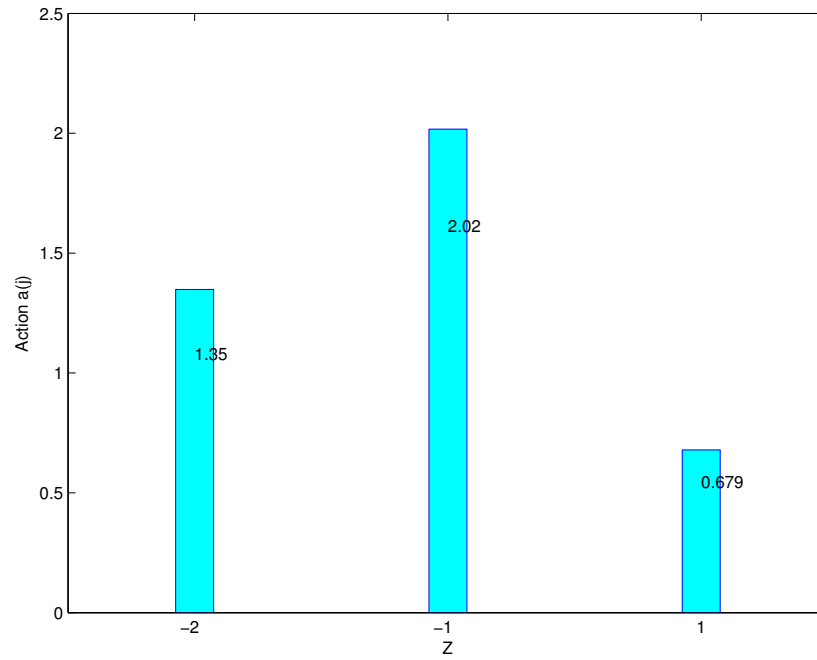


Figure 8: Optimal Steady-state commitment outcomes.

(a) Agent effort levels



(b) Agent average discounted payoffs

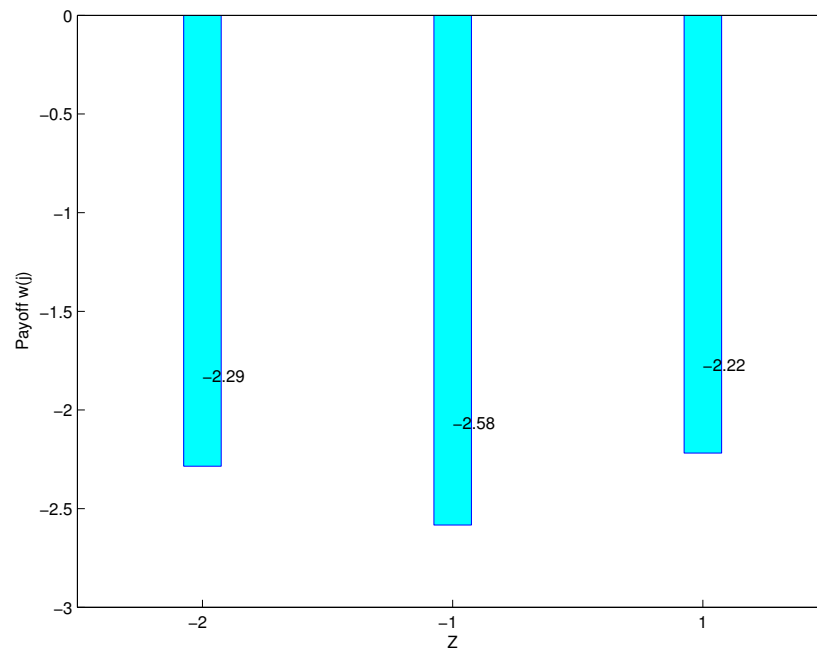


Figure 9: Sample SSE path for induced agent distributions beginning from the optimal fixed-policy's long-run state, where SSE supports the best SSE average discounted social payoff.

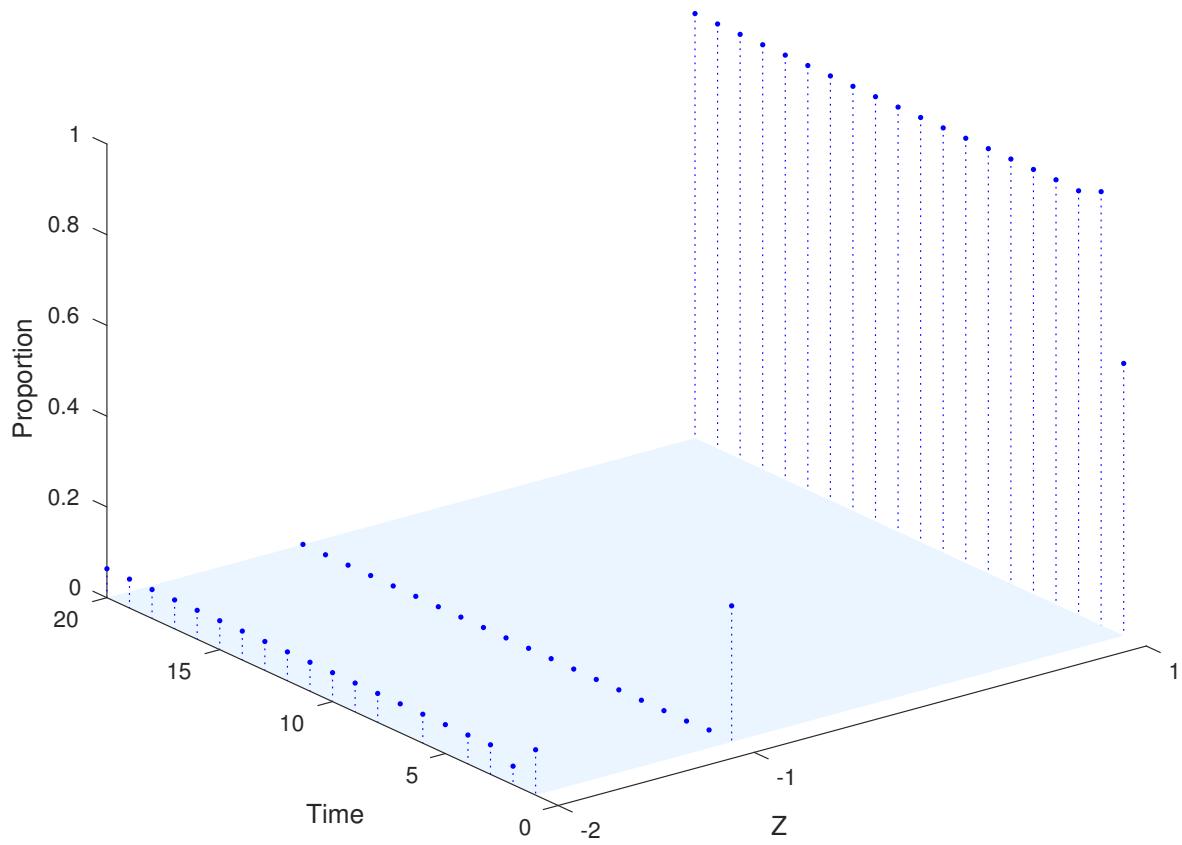


Figure 10: Sample SSE path of transfer policies beginning from the optimal fixed-policy's long-run state, but supporting the best SSE average discounted social payoff.

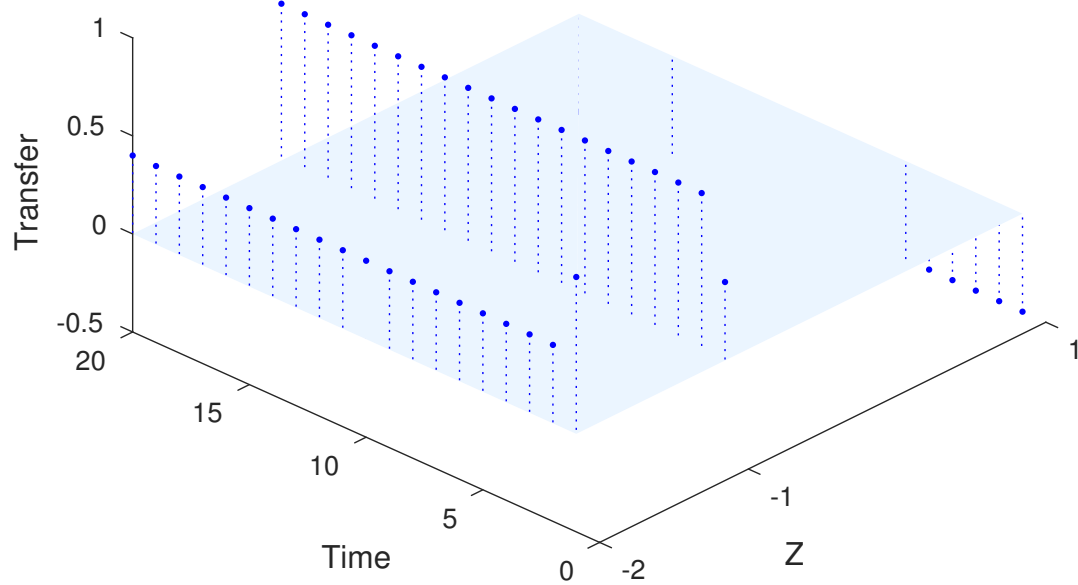


Figure 11: Sample SSE path of agents' averaged discounted payoffs (ADP) beginning from the optimal fixed-policy's long-run state, but supporting the best SSE average discounted social payoff.

