Implications of Uncertainty for Optimal Policies*

Todd Lensman Maxim Troshkin

November 13, 2020

Abstract

We study the optimal policy implications of a broader view of uncertainty in otherwise conventional macro public finance environments with heterogeneous agents and private idiosyncratic shocks. We describe conditions under which broader uncertainty implies that it is optimal to periodically reform policies. Periodic reforms lead to simplified optimal policies that are not fully contingent on future shocks; at times they also lose dependence on the full history of past shocks. These simplified policies can be characterized without complete backward induction when the time horizon is finite. However, linear policies can be far from optimal. We also show that equilibria in decentralized versions of these economies are not generally efficient, implying a meaningful role for government provision of insurance, unlike in conventional environments with a narrower view of uncertainty.

Keywords: social insurance, redistribution, reforms, history dependence, efficiency of competitive equilibria, risk sharing, insurance, uncertainty, ambiguity, model misspecification, robustness

JEL: H21, H11, E62, D8

*Lensman: MIT, tlensman@mit.edu. Troshkin: University of Pennsylvania, Wharton, and University of Exeter, m.troshkin@exeter.ac.uk. We thank the editor and an anonymous referee for a number of constructive suggestions that streamlined our arguments. We benefited from comments by Marco Battaglini, Steve Coate, Lars Hansen, Narayana Kocherlakota, Asen Kochov, Alessandro Pavan, Florian Scheuer, Stefanie Stantcheva, Iván Werning, and the audiences at the NBER Summer Institute, Taxation Theory Conference, LAEF at UC Santa Barbara, and numerous seminars. We thank John Zhu for early discussions that led to a productive approach and key results. We gratefully acknowledge support from Cornell University's Robert S. Hatfield Fund.

"The future ain't what it used to be."
- attributed to Yogi Berra

1 Introduction

A sizable and growing literature shares the following approach to social insurance, redistribution, and normative questions in dynamic economies more generally:[1] Start with a friction, typically private information about idiosyncratic shocks, and characterize friction-constrained allocations that maximize an ex ante objective, typically social welfare. The optimal policies are then the ones that implement constrained-optimal allocations. Crucially, the policy designer and the agents in the economy are commonly assumed to know the data-generating process for the shocks ("rational expectations"). Three broad outcomes are closely associated with this assumption:

(i) policies are designed once (ex ante) and maintained forever;

(ii) optimal policies are generically complex (contingent on future shocks, history dependent, and significantly nonlinear); and

(iii) undistorted competitive equilibria are constrained-efficient, even with private idiosyncratic shocks.[2]

On the other hand, exact knowledge of the data-generating process is widely viewed as a strong assumption.[3] Moreover, real world policies are often far from (i) and (ii): Government policies, especially fiscal policies, are periodically reformed, at least somewhat insensitive to shocks, history independent, and often affine or piecewise linear. Can any of these real world properties be rationalized as optimal? This paper argues that they can, once we move away from the assumption of rational expectations and allow for uncertainty about the data-generating process. In addition, as a starting point for a normative approach, (iii) is restrictive because it implies that the only insurance role for government policy is to crowd out insurance provided by private markets. This paper also shows that (iii) may no longer limit the normative approach when the rational expectations assumption is relaxed.

[1]For examples of applications to fiscal policies see, e.g., Kocherlakota (2010). For an example of this approach to dynamic social insurance more broadly see, e.g., Williams (2011).

[2]See, e.g., Acemoglu and Simsek (2012), or in a classic moral hazard setting see Prescott and Townsend (1984).

[3]Growing evidence suggests that idiosyncratic shock distributions change significantly and often. For instance, recent administrative data evidence suggests that the distribution of pre-tax earnings in many developed countries has been going through dramatic changes that appear irregular, frequent, and apparently unanticipated by the governments (see, e.g., Piketty, Saez, and Zucman 2018). Related evidence suggests that people tend to hold distribution-incompatible beliefs about their future productivities, tax liabilities, etc. (see, e.g., Aghion, Akcigit, Lequien, and Stantcheva 2017). In addition, a large empirical literature documents substantial uncertainty about both macro and micro variables (see, e.g., Bloom 2014).

Our goal is to characterize broad properties of optimal policies that are robust with respect to incomplete knowledge of the stochastic process for shocks. To do so, we remove the assumption of exact knowledge of future distributions of shocks and instead allow a broader view of uncertainty. That is, the agents in the economy face both risk in the conventional sense of stochastic, heterogeneous skills, as well as (Knightian or model) uncertainty in the sense that agents entertain multiple possible distributions of future skills, commonly referred to as beliefs, models, or priors. The approach we take to modeling risk and uncertainty, with aversion to both, follows the approaches in macroeconomics and finance (e.g., Hansen and Sargent 2001, Epstein and Schneider 2003).[4]

We consider an otherwise conventional heterogeneous-agents environment with idiosyncratic and potentially persistent shocks to skills and beliefs. The data-generating process for shocks is arbitrary, and it is not known to anyone in the economy. Each period, every agent draws a shock and forms a set of distributions that he believes may represent the data-generating process for shocks in the next period. We impose a simple condition on these beliefs, and we keep the environment virtually agnostic about any learning that may map histories of observations and current distributions into updated beliefs about the future. To aid intuition, we present arguments using a recursive maxmin expected utility representation of preferences, with arbitrary belief updating rules.[5]

The economy has a government that seeks to provide social insurance against risk and uncertainty, as well as a degree of redistribution. It is constrained by the same uncertainty about the distribution of future shocks, so the government is not an abstract entity with perfect knowledge of the data-generating processes. Rather, the government is interpreted concretely as having at best the information that all of the agents in the economy have combined.

To make it transparent that the main force behind the results is uncertainty, we first develop them in a baseline finite-horizon environment with a finite number of agents, in which agents' idiosyncratic shocks are publicly observable (Section 2). We then extend our results to the setting in which shocks are privately observed by the agents (Section 3). To show that competitive equilibria are not generally efficient, we compare outcomes in this private-information economy to those in its decentralizations (Section 4). Finally, we further

[4]See also Hansen, Sargent, Turmuhambetova, and Williams (2006). More recent studies in those literatures have focused on showing that uncertainty can help explain the behavior of economic aggregates to a surprising degree, e.g., Bhandari, Borovička, and Ho (2017).

[5]The results, however, do not hinge on the maxmin representation and are readily generalizable to, e.g., dynamic Variational Preferences, and more generally would extend to a dynamic version of the Uncertainty Averse Preferences that unify many other specifications, including in particular the Multiplier Preferences of Hansen and Sargent (2001). See, e.g., Machina and Siniscalchi (2014) for a review of the links between these representations.

examine the simplicity of optimal policies by discussing conditions under which linear policies can be optimal (Section 5).

Briefly, the intuition for the main findings is as follows. First, we show that it is optimal for the government to periodically *reform*, i.e., that the (constrained) efficient level of welfare can be delivered with allocations that are redesigned as needed after the economy realizes a new set of shocks. Our condition on beliefs is that today each agent allows for the possibility of everyone believing tomorrow that some shocks will not be realized in the future. Such shocks can then be ignored by a non-paternalistic government when designing a welfare-maximizing policy, as long as this does not pose an issue for feasibility. If tomorrow agents' actual beliefs differ (i.e., a situation described by the epigraph takes place), then a Pareto improvement can be found. In other words, the government may find it possible in subsequent periods to improve everyone's welfare by redesigning the continuation allocations, i.e., by reforming. Every agent foresees this possibility but does not find it necessary to preempt it since their welfare weakly increases as a result.

By the same logic, broader uncertainty implies that optimal policies are *simplified* in the sense that they are not fully contingent on future shocks. Optimal allocations may also lose dependence on past shocks whenever a reform provides a big enough improvement to the status quo allocation. In addition, because they are periodically reformed, optimal allocations can be constructed by solving what we call a reform problem. The reform problem is recursive in nature, with the previously designed policy as a state variable, and provides an algorithm for characterizing the optimum period by period, without solving for a complete sequence of fully state-contingent policy functions. When the time horizon is finite, the optimality of reforms implies that optimal allocations can be characterized without the full backward induction ordinarily required. Despite these simplifications, restrictive assumptions are required for linear or even affine policies to be optimal, e.g., that agents supply labor inelastically and believe that idiosyncratic shocks are independently distributed.

Finally, broader uncertainty creates a potentially meaningful role for government provision of insurance. In decentralizations of the private-information economy, agents contract with competitive firms to provide labor and capital in exchange for consumption. We show that whether competitive equilibria are efficient depends crucially on the relationship between the firms' uncertainty and the government's uncertainty, as expressed through its feasibility constraint. If firms are more uncertain about the data-generating process than the government, they may be unwilling to provide the same degree of insurance as the government. In this sense, uncertainty may act as a friction that impedes the efficient operation of competitive markets. Nevertheless, by the same intuition as in the government's problem, any insurance that decentralized economies do provide can still be simplified and periodically reformed.

Related Literature

We contribute most directly to a literature that studies the design of social insurance and redistribution in dynamic economies (see, e.g., Kocherlakota 2010). It generally assumes exact knowledge of the data-generating process to characterize optimal policies when the government has few or no direct constraints on policy tools, but potentially faces informational or commitment frictions. A key lesson without uncertainty is that optimal dynamic policies are generically complex. For example, Farhi and Werning (2013) use the first-order approach to characterize complex dynamics of optimal income taxes over the life cycle. Recent contributions also introduce additional frictions or permit additional heterogeneity on the part of agents (e.g., Scheuer and Werning 2016, Stantcheva 2017, and Makris and Pavan 2018). However, many of these contributions also compute the welfare losses from restrictions on policy tools to argue that simpler policies can deliver welfare that closely approaches that of the optimal dynamic policies. Our results provide a theoretical foundation for simpler policies by showing that such policies can in fact be optimal when there is uncertainty about the data-generating process.

Parts of this broader literature focus on a government that has an inability to commit to its own policies and may seek to implement reforms. For example, Farhi, Sleet, Werning, and Yeltekin (2012) approach this by constraining the government ex ante to choosing policies that it will not seek to reform later. We instead study conditions under which periodic reforms are optimal even when the government has the ability to commit to a policy.

Closely related to the implementation of optimal policies, an important observation is that decentralized competitive equilibria result in (constrained) efficient allocations in many cases, even when agents have private information (e.g., Acemoglu and Simsek 2012). A common interpretation is that the only result of a government's social insurance policy is to crowd out insurance provided by private markets. Our results show that removing the assumption of exact knowledge of the data-generating process and permitting firms and the government to react heterogeneously to this uncertainty can overturn this conclusion, creating a potentially meaningful role for the government provision of insurance.

More broadly, a growing literature in macroeconomics has started to consider optimal fiscal and monetary policies when either the government or the agents are uncertain about the data-generating process (e.g., Hansen and Sargent 2012, Karantounias 2013, Benigno and Paciello 2014; see Barlevy 2011 for a review). Most related to this paper, Kocherlakota and Phelan (2009) consider an endowment shock economy in which the government, but not the agents, is uncertain about the data-generating process. They derive conditions under which the government cannot improve on the competitive equilibrium allocation. We consider optimal policies in a dynamic production economy in which no one is certain, with a potentially beneficial role for government intervention. Bhandari (2015) studies properties of optimal risk-sharing arrangements between uncertain agents, using the multiplier preferences of Hansen and Sargent (2001) to characterize the optimal consumption path and long-run inequality. Kocherlakota and Song (2018) reexamine mechanisms for the provision of a public good in economies where agents are uncertain about the distribution of private valuations. Similarly to our results, they find that uncertainty can lead to simple implementations of efficient policies.

A recent theoretical literature has also studied Pareto optimal allocations and optimal mechanisms in economies where the agents or the mechanism designer are uncertain about primitives of the environment (see Carroll 2019 for a review). A subset of this literature attempts to rationalize simple mechanisms observed in reality using uncertainty (e.g., Carroll 2015), and we similarly attempt to justify simple properties of real-world fiscal policies. Rigotti, Shannon, and Strzalecki (2008) and Strzalecki and Werner (2011) characterize conditions under which Pareto optimal allocations in static endowment economies feature full insurance against idiosyncratic risk. They show that full insurance obtains precisely when agents have at least one appropriately-defined subjective belief in common. We make a qualitatively different assumption on agents' beliefs to show that Pareto optimal allocations in a dynamic economy are insensitive to future shock realizations.

Methodologically, our formulation of the agents' preferences closely follows the recursive multiple-priors utility axiomatized by Epstein and Schneider (2003), and it includes a variation of Hansen and Sargent's (2001) constraint preferences as a special case. Our approach to characterizing optimal allocations for uncertainty averse agents follows Zhu (2016), who studies incomplete and affine contracts in the context of financial contracting. The formalism we use is distinct from but in the spirit of robust mechanism design (see, e.g., Bergemann and Morris 2013), in the sense that we seek to characterize policies that are robust with respect to misspecification of the environment and to determine if such policies display any inherent simplicity. In particular, we remove the assumption that the data-generating process is known to the government and the agents. However, the social choice functions we focus on depend on agents' beliefs so that the government is not paternalistic, evaluating the policies in line with how the agents do. The economies we study are also dynamic and allow for private information.

2 Uncertainty with Public Information

To make it clear that the main force behind the results is uncertainty, we first consider a conventional dynamic heterogeneous-agents economy with public information. The only unconventional element - in a sense the only friction in this baseline setup - is the assumption that no one in the economy knows with certainty the future distributions of idiosyncratic shocks. We define an efficient allocation as a solution to the problem of a government seeking to provide social insurance and a degree of redistribution. We then show that the government can achieve efficiency with policies that are both simplified and that can be constructed in a simplified way. Such policies are not fully contingent on future shocks, are periodically reformed, and lose dependence on the full history of past shocks.[6]

2.1 Baseline Setup

The economy exists in discrete time $t \in \{0, \ldots, T\}$ and is populated by a finite number of agents $i \in \{1, \ldots, N\}$.[7] At the beginning of $t = 0$, nature draws a sequence of idiosyncratic shocks $s_i^T \equiv (s_{i,0}, .., s_{i,T})$ for each agent and reveals $s_{i,t}$ to agent i at the beginning of t. Each *shock* $s_{i,t}$ is drawn from a finite set S, and we let $s_t \equiv (s_{1,t}, \ldots, s_{N,t})$ denote the *shock vector* in period t. The data-generating process for the shocks is arbitrary and is not known with certainty to anyone in the economy. In the baseline economy of this section, this lack of certainty is the only friction, so a shock becomes public information once it is revealed to an agent. Consequently, the vector of all agents' shock histories up to period t, $s^t \equiv (s_1^t, \ldots, s_N^t)$, is public information at the beginning of t. We call s^t the *state* of the economy at t.

The state s^t determines each agent's skill and beliefs. $\theta_{i,t}(s_{i,t})$ denotes agent i's idiosyncratic *skill*, so that if the amount of labor the agent exerts is $l_{i,t} \in [0, 1]$, then the effective labor supplied is $z_{i,t} \equiv \theta_{i,t} l_{i,t}$. Let $\Theta \subset \mathbb{R}_{++}$ denote the finite set of values that each skill $\theta_{i,t}$ can take, and let $\underline{\theta}$ and $\bar{\theta}$ denote the minimum and maximum elements of Θ respectively. Let $\theta_i^t \equiv (\theta_{i,0}, \ldots, \theta_{i,t})$ denote a period-t history of skills.

$\Pi_{i,t+1}(s^t)$ denotes agent i's set of subjective *beliefs* $\pi \in \Delta(S^N)$ over the next period's shock vector s_{t+1}. The set $\Pi_{i,t+1}(s^t)$ is a function of the state s^t, not just the shock $s_{i,t}$, because agent i observes s^t at the beginning of period t. As we describe below, this *belief set* represents agent i's uncertainty in state s^t about the data-generating process for shocks.

To examine properties common to social insurance and redistribution in general, without restricting attention to specific collections of policy tools, we focus on the allocations that a

[6]Without uncertainty about the data-generating process, the efficient allocation with public information does not display history dependence when the data-generating process is Markov. However, even then efficient capital investment can depend on the data-generating process in complex ways that preclude periodic reforms, especially when the process is persistent.

[7]In the main text, we assume that the time horizon T, the number of agents N, and the set of possible shocks are finite. Our results readily extend to infinite cases, as we discuss in Appendix B.

policy would deliver to agents: An *allocation*

$$C \equiv \left\{ c_t \left(s^t \right), z_t \left(s^t \right), k_{t+1} \left(s^t \right) \right\}_{t=0}^{T}$$

is a sequence of N-vector valued consumption, effective labor, and capital functions that depend on the state s^t. For example, in period t with state s^t, agent i consumes $c_{i,t} \left(s^t \right)$, supplies effective labor $z_{i,t} \left(s^t \right)$, and saves capital $k_{i,t+1} \left(s^t \right)$.[8]

Each agent's preferences over allocations are assumed to have a recursive representation with continuation utility

$$U_{i,t} \left(C \left| s^t \right. \right) \equiv u \left(c_{i,t} \left(s^t \right), \frac{z_{i,t} \left(s^t \right)}{\theta_{i,t} \left(s_{i,t} \right)} \right) + \beta \inf_{\pi \in \Pi_{i,t+1}(s^t)} \mathbb{E}_\pi \left[U_{i,t+1} \left(C \left| s^{t+1} \right. \right) \right]. \tag{1}$$

Here $\beta \in (0,1)$ is the subjective discount factor, the Bernoulli utility function u is concave and satisfies $u_c, -u_l > 0$, and $\mathbb{E}_\pi$ denotes an expectation with respect to the belief π. With this specification of preferences, agents are potentially averse to both risk originating from stochastic skills and uncertainty captured by multiple beliefs. The latter is usually interpreted as seeking to make choices that are robust with respect to the shock distribution: Instead of choosing what works best in a particular future scenario, agents choose what works decently in any scenario, which entails choosing what works best in the worst scenario.

Each agent is initially endowed with capital $k_0 > 0$. Output is produced using a constant returns to scale production function $f : \mathbb{R}_+^2 \to \mathbb{R}_+$ increasing in capital and effective labor. Given an initial state s_0, an allocation C is *feasible* if[9]

$$\sum_i \left[c_{i,t} \left(s^t \right) + k_{i,t+1} \left(s^t \right) \right] \leq \sum_i f \left(k_{i,t} \left(s^{t-1} \right), z_{i,t} \left(s^t \right) \right) \quad \forall t \geq 0, s^t \geq s_0. \tag{2}$$

This aggregate ex post feasibility constraint must hold for any state s^t that could follow s_0, because the agents may have different beliefs about the distribution (or even the support) of the state in each period.

The economy also has a government that seeks a degree of redistribution while providing social insurance. The government's problem is to maximize a weighted average of the agents' utilities, subject to feasibility and non-negativity of policy functions, where the weighting

[8]We could also consider *belief-free allocations* in which the consumption, effective labor, and capital functions depend on the state s^t only through the history of skills θ^t. Our results continue to apply when restricting to belief-free allocations. The same proofs apply with minor modifications, and in fact our key results (Propositions 1 and 3) hold under weaker conditions on beliefs than required below.

[9]For any two states s^t and s^τ with $t \geq \tau$, we say that s^t *follows* s^τ, written $s^t \geq s^\tau$, if the first $\tau + 1$ components of s^t are s^τ. We use this ex post notion of feasibility for simplicity, but the results below readily generalize under less stringent feasibility constraints.

captures the redistribution motive and is given by a non-negative vector $\eta \in \mathbb{R}_+^N$. To focus on social insurance and redistribution, the agents are not allowed to leave the social contract designed by the government, i.e., they do not have access to an exogenous outside option. Given social welfare weights η and an initial state s_0, an allocation $C^*(s_0)$ is *efficient* if

$$C^*(s_0) \in \arg\max_{C} \sum_i \eta_i U_{i,0}(C \mid s_0), \tag{3}$$

subject to feasibility and non-negativity. Note that instead of an abstract entity with perfect knowledge of the data-generating process, the government concretely possesses the same information about realized shocks and the future state distribution as do all of the agents combined. Note also that C^* could depend in a complex way on the state in each period, and it is typically designed once and maintained forever. We assume that the government has full commitment power in the sense that if it designs an allocation at $t \geq 0$ and wishes to reform it at $\tau > t$, it must deliver the period-t continuation utility promised to agents at t.

2.2 Optimality of Periodic Reforms

We first show that under a suitable condition on agents' beliefs, the government can achieve efficiency with simplified policies that are not fully contingent on future shock realizations and are periodically reformed. Loosely speaking, the condition on beliefs is that today everyone allows for the possibility of everyone believing tomorrow that some of the states will not be realized the day after tomorrow. Such states then can be ignored to design a simplified policy without sacrificing efficiency, as long as this does not pose an issue for feasibility. If tomorrow the actual beliefs differ, then a Pareto improvement can be found.

For ease of exposition, we first illustrate the results in the case where everyone allows for the possibility of everyone believing tomorrow that the economy will follow its worst path; we then discuss how the results generalize in Section 2.4. To be more precise, assume there is a shock $\underline{s}$ such that $\theta_{i,t}(\underline{s}) = \underline{\theta}$ for any agent i and in any period t, and such that the agent believes everyone will realize $\underline{s}$ in $t+1$. When the distinction is clear from the context, we will let $\underline{s}$ denote the shock and the shock vector $(\underline{s}, \ldots, \underline{s})$.

To describe the condition on beliefs, assume that the set of shocks S is rich enough so that for any shock s, there is a shock $s'(s)$ that gives the same skill to the agent, but with a belief that everyone will receive $\underline{s}$ next period. That is, for any $t \leq T-2$, any state s^t, and any vector of shocks of the other agents $s_{-i,t+1}$, we have $\theta_{i,t+1}(s'(s)) = \theta_{i,t+1}(s)$ but $\Pi_{i,t+2}(s^t, (s'(s), s_{-i,t+1})) = \{\underline{\pi}\}$ for all i, where $\underline{\pi}(\underline{s}) = 1$. Notice that since an agent realizing the shock $\underline{s}$ believes that the shock will persist, we also have $\Pi_{i,t+3}(s^{t+1}, (\underline{s}, s_{-i,t+2})) = \{\underline{\pi}\}$.

We assume that s' is idempotent without loss of generality.

It will be convenient to let $S' \equiv s'(S) \subseteq S$ denote the subset of shocks of the form $s'(s)$, and let $S'^N \equiv s'\left(S^N\right) \subseteq S^N$ denote the subset of shock vectors found by applying s' to each element. Let $\pi' \equiv \pi \circ (s')^{-1}$ denote the pushforward measure obtained from a measure π under the mapping s', i.e., π' is the belief π "shifted" to place weight only on the subset of shock vectors S'^N. The condition on beliefs then can be stated

Assumption 1. *For any $t \leq T-2$, s^t, i, and $\pi \in \Pi_{i,t+1}(s^t)$, we also have $\pi' \in \Pi_{i,t+1}(s^t)$, where $\pi' = \pi \circ (s')^{-1}$.*

The requirement here is that, regardless of the beliefs at any t about $t+1$ skills, each agent considers the possibility of everyone believing at $t+1$ that some states cannot be realized in subsequent periods. Given the definition of s' in this case, these are the states in which some agent i realizes a skill $\theta > \underline{\theta}$.

We call an allocation C^t *simplified* at t if its policy functions depend on s_{t+1} only through $s'(s_{t+1})$ and if they do not depend on $s_{t+2}, \ldots, s_T$.[10]

Proposition 1. *In any period t, any feasible allocation C is weakly Pareto dominated by a simplified allocation C^t, i.e.,*

$$U_{i,t}\left(C^t \left| s^t\right.\right) \geq U_{i,t}\left(C \left| s^t\right.\right) \quad \forall i.$$

We describe the intuition here with $t=0$ and provide a detailed proof in Appendix A.1. By Assumption 1, at $t=0$ each agent believes that at $t=1$, regardless of the distribution of skills, it is possible to realize a state in which everyone is certain that the shock vector $\underline{s}$ will be realized at $t \geq 2$. Given an initial allocation C, the simplified allocation C^0 is defined by presuming that such a pessimistic $t=1$ state will obtain and that the shock vector $\underline{s}$ will be realized at $t \geq 2$. In particular, for any period-t state s^t, the C^0 allocation functions $c_t^0(s^t)$, $z_t^0(s^t)$, and $k_{t+1}^0(s^t)$ are set equal to the corresponding C allocation functions, but evaluated at the state $(s_0, s_1'(s_1), \underline{s}, \ldots, \underline{s})$ instead of s^t. C^0 is then feasible regardless of the realization of the state in each period. Moreover, since C^0 depends only on the (fixed) $t=0$ state and the vector of skills at $t=1$, for any state $s^t \neq (s_0, s_1'(s_1), \underline{s}, \ldots, \underline{s})$ in which agent i receives lower flow utility from C^0 than from C, Assumption 1 implies that agent i ignores s^t when comparing the two allocations. As a result, each agent weakly prefers C^0 to C.

To see why this argument generally fails when agents know the distribution of the states s^t (or more generally when agents' belief sets are singletons), note that Assumption 1 implies that each agent's $t=0$ continuation utility $U_{i,0}$ is not strictly increasing in his flow utility

[10] This definition also generalizes naturally to more general belief conditions, as we discuss in Section 2.4.

in each state s^t. Rather, because he is uncertain and is averse to this uncertainty, the agent ignores any state $s^t \neq (s_0, s'_1(s_1), \underline{s}, \ldots, \underline{s})$ in which his flow utility is higher than his flow utility in the corresponding state $(s_0, s'_1(s_1), \underline{s}, \ldots, \underline{s})$. By contrast, a Bayesian agent's $t = 0$ continuation utility is strictly increasing in his flow utility in every state in the support of his prior distribution, so lowering his flow utility in state $s^t \neq (s_0, s'_1(s_1), \underline{s}, \ldots, \underline{s})$ to that in state $(s_0, s'_1(s_1), \underline{s}, \ldots, \underline{s})$ will generally strictly lower his $t = 0$ continuation utility. Assumption 1 ensures that our agents are sufficiently uncertain so that this is not the case.

The relevance of Proposition 1 for optimal policy becomes apparent when applied to an efficient allocation C^*: An efficient allocation C^* can be implemented by a simplified allocation C^0, in the sense that C^0 delivers the same $t = 0$ continuation utility to each agent. However, if at $t = 1$ a shock vector outside of S'^N is realized, then the continuation allocation prescribed by C^0 may no longer be optimal. In that case, the government will seek to reform C^0 to a new allocation C^1 in order to raise $t = 1$ social welfare while continuing to deliver at least the $t = 0$ continuation utility promised to each agent under C^0. The same arguments as above imply that the reform allocation C^1 can be simplified at $t = 1$. In particular, Assumption 1 applies at $t = 1$, so the government can construct an optimal reform allocation C^1 by assuming that states of the form $s^t = (s^1, s'(s_2), \underline{s}, \ldots, \underline{s})$ will be realized at $t \geq 2$. By solving for such optimal reform allocations in each period, the government can construct a sequence of simplified allocations $\{C^t\}_{t=0}^T$ that implements C^*.

Constructing Optimal Reforms

Proposition 1 guarantees the existence of simplified, periodically-reformed allocations that implement the efficient allocation. More significantly, it also provides an algorithm that can be used to compute the simplified allocations $\{C^t\}_{t=0}^T$, without first solving for the efficient allocation C^*.

Corollary 1. *Optimal simplified allocations $\{C^t\}_{t=0}^T$ can be constructed period by period, without computing the fully state-contingent allocation C^*.*

In general, a simplified allocation C^{t-1} that is not fully state-contingent can be reformed to an optimal continuation allocation $C^*(s^t, C^{t-1})$, given by a solution to the *reform problem*

$$\max_{C} \sum_{i} \eta_i U_{i,t} \left(C \left| s^t \right. \right) \tag{4}$$

subject to non-negativity and

$$\sum_{i} \left[c_{i,\tau} \left(s^{\tau} \right) + k_{i,\tau+1} \left(s^{\tau} \right) \right] \leq \sum_{i} f \left(k_{i,\tau} \left(s^{\tau-1} \right), z_{i,\tau} \left(s^{\tau} \right) \right) \quad \forall \tau \geq t, s^{\tau} \geq s^t,$$

$$U_{i,t-1} \left(C_{t-1}^{t-1}, \left(C_{\tau} \right)_{\tau=t}^{T} \left| s^{t-1} \right. \right) \geq U_{i,t-1} \left(C^{t-1} \left| s^{t-1} \right. \right) \quad \forall i.$$

The second constraint is a form of promise-keeping (with no such constraint at $t = 0$). Here $\left(C_{t-1}^{t-1}, \left(C_{\tau} \right)_{\tau=t}^{T} \right)$ denotes the allocation that uses the C^{t-1} policy functions c_{t-1}^{t-1}, z_{t-1}^{t-1}, and k_t^{t-1} in period $t - 1$ and the C policy functions in periods $\tau \geq t$. The promise-keeping constraint appears because when the government chooses a policy at $t - 1$, it commits to delivering at least the $t - 1$ continuation utility promised to each agent, even after subsequent reforms. Proposition 1 implies that the government can choose a simplified allocation C^t to solve this reform problem, so this reform process constructs a sequence of optimal simplified allocations that implement the efficient allocation C^*. Note that the components of the C^{t-1} allocation from t onward serve as a fallback option in the case that the government cannot construct a better reform allocation at t. Even though the government has the ability to fully commit to maintaining an allocation forever, it may choose not to and instead design optimal simplified policies period by period, subsequently reforming them as necessary.

This algorithm for computing optimal simplified allocations period by period may entail substantial computational benefits relative to the case without uncertainty. When the data-generating process is known to the government and the agents, characterizing the efficient allocation for a given period requires full backward induction from the last period, considering all possible paths that the economy may follow. By contrast, with uncertainty it is not necessary to compute the fully-contingent efficient allocation C^*. Rather, the government must only compute simplified allocations in each period.

2.3 History Independence

The simplified allocations C^t that solve the reform problem above are simplified because they are not fully contingent on future shocks. In particular, the construction in the proof of Proposition 1 shows that they only depend on the period-t state s^t and the period-$t + 1$ skills $\theta_{i,t+1}$. We next show that they are also history independent in the sense that they lose full dependence on previous shocks whenever reforms provide an improvement to previously designed government policies. Specifically, whenever agents' belief sets $\Pi_{i,t+1} \left(s^t \right)$ do not depend on s^{t-1} and the government's promise-keeping constraints are slack at t, the optimal reform C^t is independent of the $t-1$ state s^{t-1}.

Proposition 2. *For any t at which $\Pi_{i,t+1}(s^t)$ does not depend on s^{t-1} for all i and the promise-keeping constraints in the reform problem (4) are slack, the optimal C^t is independent of s^{t-1} conditional on the distribution of capital $k_t^{t-1}(s^{t-1})$.*

Proof. When the promise-keeping constraints in problem (4) do not bind at t, the government must maximize an η-weighted average of agents' t continuation utilities, subject to feasibility at $\tau \geq t$. If all agents' belief sets $\Pi_{i,t+1}(s^t)$ do not depend on s^{t-1}, the continuation utility $U_{i,t}(C\,|s^t)$ does not depend on s^{t-1} other than through C, and similarly s^{t-1} is only relevant to the feasibility constraint through the period-t capital distribution $k_t^{t-1}(s^{t-1})$, which is fixed at the beginning of t. Thus to maximize its objective, the government will choose the optimal reformed C^t so that it does not depend on s^{t-1}, given fixed $k_t^{t-1}(s^{t-1})$. □

Agents' belief sets $\Pi_{i,t+1}(s^t)$ do not depend on s^{t-1} when, for example, they are derived by using Bayes' rule to update a fixed set of priors under which the shock vector follows a Markov process. The result also generalizes naturally to the case in which agents' belief sets $\Pi_{i,t+1}(s^t)$ depend only on shocks from period $\tau \leq t$ to period t, $(s_\tau, \ldots, s_t)$. In this case, the optimal C^t is independent of the $\tau-1$ state $s^{\tau-1}$ whenever a reform leads to the government's promise-keeping constraints being slack at t.

A benchmark case in which the allocation C^t will never depend on the $t-1$ state s^{t-1} for $t \geq 1$ arises when (i) agents' belief sets $\Pi_{i,t+1}(s^t)$ do not depend on s^{t-1} and (ii) C^t can be computed by backward induction for all t.[11] To see this, fix $t=0$, and start by solving for the optimal $t=1$ continuation allocation $C_1^0 \equiv \left\{c_t^0, z_t^0, k_{t+1}^0\right\}_{t=1}^T$. Given a $t=1$ shock vector $s_1' \in S'^N$ and a distribution of capital k_1^0, the continuation allocation $C_1^0(s_1', k_1^0)$ solves

$$\max_{C_1^0} \sum_i \eta_i U_{i,1}\left(C_1^0\,|s_1'\right)$$

subject to non-negativity and

$$\sum_i \left[c_{i,t}^0(s_1', \underline{s}, \ldots, \underline{s}) + k_{i,t+1}^0(s_1', \underline{s}, \ldots, \underline{s})\right] \leq \sum_i f\left(k_{i,t}^0(s_1', \underline{s}, \ldots, \underline{s}), z_{i,t}^0(s_1', \underline{s}, \ldots, \underline{s})\right) \quad \forall t \geq 1.$$

[11]It may not always be possible to construct optimal allocations using backward induction. For example, if beliefs are heterogeneous, it may be optimal in period t to promise an agent a larger share of consumption in some $t+1$ states than would be optimal when designing continuation allocations starting in period $t+1$. This complication arises because the government's preferences over allocations may not be dynamically consistent, even though the agents' preferences are.

Note that $U_{i,1}$ does not directly depend on s_0 because we assume that $\Pi_{i,2}(s^1)$ does not depend on s_0. The government solves this problem for each shock vector $s_1' \geq s_0$ and each capital distribution k_1^0, and the remaining policy functions c_0^0, z_0^0, and k_1^0 are then found by a similar optimization problem at $t = 0$, taking C_1^0 as given.

At $t = 1$, if a shock vector $s_1' \in S'^N$ is realized, then by construction the continuation allocation $C_1^0(s_1, k_1^0)$ is optimal. The government then chooses not to reform and sets $C^1 = C_1^0(s_1', k_1^0)$. If a shock vector $s_1 \notin S'^N$ is realized, then there exists an agent i who believes it possible that an agent j will realize a skill $\theta_{j,t} > \underline{\theta}$ in some period $t \geq 2$. In such a period-t state, the government can produce strictly greater output than if all agents realized skill $\underline{\theta}$ while delivering the same continuation utilities. The government would not reduce agents' $t = 1$ continuation utilities in response to one agent i receiving more optimistic beliefs than anticipated in period 1. As a result, C^1 must deliver weakly greater $t = 1$ continuation utility than C_1^0 to every agent regardless of the realized shock vector s_1, so any re-optimized C^1 must trivially satisfy the promise-keeping constraint. In particular, if the optimal simplified allocation C^t can be constructed by backward induction for all t, then promise-keeping constraints will never bind. The resulting C^t will then be independent of s^{t-1} for $t \geq 1$.

2.4 Discussion

Three points about our setting and results warrant further discussion. First, we briefly relate our economy's information structure and preferences to similar models of uncertainty aversion in decision theory and macroeconomics. Second, the optimality of periodic reforms may suggest that our results can be attributed to dynamic inconsistency of the agents' preferences.[12] We explain that this is not the case, and indeed the agents' preferences satisfy the standard notion of dynamic consistency used in the literature. Third, we discuss how the condition on beliefs can be weakened.

Information Structure and Preferences

The information structure and preferences described in Section 2.1 include many of those found in the literature as special cases. For example, the canonical setting of Epstein & Schneider (2003) is obtained by letting $S^{N(T+1)}$ denote the underlying state space and letting $(\mathcal{F}_t)_{t=0}^T$ denote the filtration generated by the stochastic process $(s^t)_{t=0}^T$. The recursive representation (1) is then precisely the one axiomatized by Epstein & Schneider (2003) when

[12]It is well-known that uncertainty (relaxing Savage's Sure-Thing Principle) creates the potential for dynamic inconsistency (see, e.g., Machina and Siniscalchi 2014).

allocations are permitted to be random, while the same representation was axiomatized with deterministic allocations by Kochov (2015).

In addition, while we remain agnostic about the belief sets $\Pi_{i,t+1}(s^t)$, one possibility is to generate these sets using the procedure described by Hansen and Sargent (2001): In each state s^t, each agent i constructs a statistical model $\pi^* \in \Delta(S^N)$ that describes what he thinks may be the true distribution of the next period's shock vector s_{t+1}. The agent distrusts this model and considers other models π that are "close to" π^* in the sense of a distance d on $\Delta(S^N)$, commonly taken to be relative entropy. Given a parameter $\epsilon \geq 0$ that governs how uncertain the agent is, his set of beliefs is $\Pi_{i,t+1}(s^t) \equiv \{\pi | \, d(\pi, \pi^*) \leq \epsilon\}$. If the distance d takes into account only the distributions of skills $\theta_{t+1} \equiv (\theta_{1,t+1}, \ldots, \theta_{N,t+1})$ implied by π and π^*, then Assumption 1 is easily satisfied. However, note that this construction of the agent's beliefs is generally not equivalent to selecting a prior $P^* \in \Delta(S^{N(T+1)})$ over the state space ex ante, using a distance d on $\Delta(S^{N(T+1)})$ to construct a family of multiple priors, and then allowing each agent to update the priors using Bayes' rule in each period. Unless the set of priors happens to satisfy the rectangularity property of Epstein & Schneider (2003), the resulting family of preferences will fail to be dynamically consistent and so will not have a recursive representation of the form (1).

Dynamic Consistency

The agents' preferences are defined recursively so that they satisfy the following notion of dynamic consistency: For any $t < T$ and any s^t, consider two allocations C and $\tilde{C}$ that coincide at t. Suppose that agent i weakly prefers $\tilde{C}$ to C at $t+1$ for any state $s^{t+1} \geq s^t$:

$$U_{i,t+1}\left(C \left| s^{t+1}\right.\right) \leq U_{i,t+1}\left(\tilde{C} \left| s^{t+1}\right.\right) \quad \forall s^{t+1} \geq s^t.$$

Then the agent's preferences are *dynamically consistent* in the sense that he will also weakly prefer $\tilde{C}$ to C at t:

$$U_{i,t}\left(C \left| s^t\right.\right) \leq U_{i,t}\left(\tilde{C} \left| s^t\right.\right).$$

This notion of dynamic consistency features prominently in the literature on dynamic preferences with uncertainty. For example, the axiomatizations of recursive Variational Preferences, recursive Maxmin Preferences, and recursive Smooth Ambiguity Preferences all require this property. Epstein and Schneider (2003) additionally show that dynamic consistency with recursive Maxmin Preferences is equivalent to a "rectangularity" condition on agents' prior distributions (see their Theorem 3.2), and it is immediate that the sets of priors induced by our belief sets $\Pi_{i,t+1}(s^t)$ satisfy that condition as well.

Our agents' preferences also satisfy a slightly weaker property, but one that is arguably more relevant in macroeconomics and public finance (Hansen and Sargent 2006): Dynamic preferences are sufficiently consistent over time if a solution to a dynamic choice problem computed by backward induction is also optimal ex ante. This condition is strictly weaker than the notion of dynamic consistency above because it only implies consistency in preference orderings involving the ex ante optimal solution. For example, the constraint preferences of Hansen and Sargent (2001) do not satisfy the definition of dynamic consistency above, but they do satisfy the weaker property (see, e.g., Epstein and Schneider 2003, Section 5).

More General Belief Conditions

The condition on beliefs imposed by Assumption 1 essentially requires that, at $t = 0$, each agent believes that regardless of the skills realized at $t = 1$, everyone may be certain that all skills will equal $\underline{\theta}$ in subsequent periods. This "possibility of certainty" that the economy's worst-case scenario will obtain at $t \geq 2$ is likely a strong assumption. Moreover, it may (falsely) suggest that our results require the existence of states s^1 in which our uncertainty-averse agents entertain no uncertainty at all!

In Appendix A.2, we provide an example showing that this is not the case, and in fact periodic reforms can remain optimal under more realistic belief conditions. It is also simple to see how Proposition 1 generalizes when Assumption 1 is relaxed: Consider a relaxation requiring that at $t = 0$, each agent believes that regardless of the skills realized at $t = 1$, everyone may be certain that skills will remain below some value $\check{\theta} \in \Theta$ in subsequent periods. If $\check{\theta} > \underline{\theta}$, agents can remain uncertain at $t = 1$ about the distribution of future skills, though they are certain that all skills will satisfy $\theta_{i,t} \leq \check{\theta}$. Our arguments above then extend naturally to this case. As to be expected, the price paid for this relaxation is a greater degree of state dependence (equivalently, a smaller degree of simplification) in the simplified allocation C^0. Finally, we note that this relaxation and the example in the Appendix are meant to be suggestive rather than exhaustive; other assumptions on beliefs can lead to a lack of state dependence (e.g., only in a particular period) in Pareto optimal allocations.

3 Uncertainty with Private Information

For some applications, it is clearly too demanding to presume that uncertainty about the data-generating process is the only friction in the design of optimal policies. This section is devoted to relaxing the assumption of publicly observable skills and beliefs. We consider private shocks and show that the main results of Section 2 persist in economies constrained by uncertainty as well as private information.

3.1 Private Information Setup

Consider the following modification to the baseline setup in Section 2.1: Agents are privately informed about their shocks $s_{i,t}$, and after receiving them at the beginning of the period, they can make reports. Let $\hat{s}_{i,t}$ denote a *reported shock*, and let $\hat{s}_i^t \equiv (\hat{s}_{i,0}, \ldots, \hat{s}_{i,t})$ denote a history of reported shocks up to period t, called agent i's *reported type*; $s_i^t \equiv (s_{i,0}, \ldots, s_{i,t})$ denotes agent i's *type* in period t. Let $\hat{s}^t \equiv (\hat{s}_1^t, \ldots, \hat{s}_N^t)$ denote the *reported state* of the economy.

A *reporting strategy* is $\sigma_i \equiv \{\sigma_{i,t}\}_{t=0}^T$, where $\sigma_{i,t}$ maps a reported state $\hat{s}^{t-1}$ and an actual type s_i^t to a reported shock $\hat{s}_{i,t}$. Let Σ be the set of possible reporting strategies, and let $\sigma \equiv \{\sigma_i\}_{i=1}^N \in \Sigma^N$ be a strategy profile. The truth-telling strategy is denoted σ_i^*, and similarly $\sigma^* \equiv \{\sigma_i^*\}_{i=1}^N$ denotes the strategy profile in which all agents use the truth-telling strategy.

Each agent i's period-t skill $\theta_{i,t}(s_{i,t})$ continues to be a function of his current shock $s_{i,t}$, but we must adjust the definition of a belief π as well as the domain of the belief set mappings $\Pi_{i,t+1}$ to account for the incomplete information about other agents' shocks. A period-t belief $\pi \in \Delta\left(S^{(N-1)(t+1)+1}\right)$ for agent i is now a distribution over the other agents' current types s_{-i}^t and his own next period shock $s_{i,t+1}$. To allow for learning about the distribution of other agents' shocks over time, agent i's belief set $\Pi_{i,t+1}(\hat{s}^{t-1}, s_i^t, \sigma)$ is now a function of the reported state $\hat{s}^{t-1}$, his current type s_i^t, and the strategy profile σ. When there is no risk of confusion, we will suppress the dependence on σ for notational simplicity.

Given an allocation $C = \{c_t(\hat{s}^t), z_t(\hat{s}^t), k_{t+1}(\hat{s}^t)\}_{t=0}^T$, a reported state $\hat{s}^{t-1}$, a type s_i^t, and a strategy profile σ, agent i's period-t continuation utility is

$$U_{i,t}\big(C \,\big|\hat{s}^{t-1}, s_i^t\big)(\sigma) \equiv \inf_{\Pi_{i,t+1}\left(\hat{s}^{t-1}, s_i^t\right)} \mathbb{E}_\pi \left[W_{i,t}\big(C\,\big|\hat{s}^{t-1}, s^t, s_{i,t+1}\big)(\sigma)\right], \tag{5}$$

where

$$W_{i,t}\big(C\,\big|\hat{s}^{t-1}, s^t, s_{i,t+1}\big)(\sigma) \equiv u\left(c_{i,t}\left(\hat{s}^{t-1}, \sigma_t\left(\hat{s}^{t-1}, s^t\right)\right), \frac{z_{i,t}\left(\hat{s}^{t-1}, \sigma_t\left(\hat{s}^{t-1}, s^t\right)\right)}{\theta_{i,t}\left(s_{i,t}\right)}\right) + \beta U_{i,t+1}\big(C\,\big|\left(\hat{s}^{t-1}, \sigma_t\left(\hat{s}^{t-1}, s^t\right)\right), s_i^{t+1}\big)(\sigma). \tag{6}$$

In (6) we slightly abuse notation letting $\sigma_t(\hat{s}^{t-1}, s^t) \equiv (\sigma_{1,t}(\hat{s}^{t-1}, s_1^t), \ldots, \sigma_{N,t}(\hat{s}^{t-1}, s_N^t))$ denote the vector of all agents' period-t reports. Thus $(\hat{s}^{t-1}, \sigma_t(\hat{s}^{t-1}, s^t))$ denotes the reported state $\hat{s}^t$, given that the actual state realized at t is s^t and that the agents follow the strategy profile σ. With definitions (5) and (6), $W_{i,t}$ is the payoff of agent i after types are reported in period t, while $U_{i,t}$ gives the infimum of the expected payoff before reporting. In contrast to Section 2, period-t flow utility u now appears inside of the expectation $\mathbb{E}_\pi$ (and inside of the infimum) because the other agents' types s_{-i}^t are not publicly observable. As a result, agent i is potentially uncertain about the distribution of the other agents' current types as well as the distribution of future shocks, and he is averse to this uncertainty.

Given the information structure, we define an *equilibrium strategy profile* as a strategy profile $\sigma^e \in \Sigma^N$ such that[13]

$$U_{i,0}\left(C\,|s_{i,0}\right)(\sigma^e) \geq U_{i,0}\left(C\,|s_{i,0}\right)\left(\sigma^e_{-i}, \sigma_i\right) \quad \forall i, s_{i,0}, \sigma_i \in \Sigma.$$

Here $\left(\sigma^e_{-i}, \sigma_i\right)$ denotes the strategy profile in which agent i uses strategy σ_i and any agent $j \neq i$ uses his respective equilibrium strategy σ^e_j. Given the agents' preferences, this definition of an equilibrium strategy profile is natural: It implies that at $t = 0$, any agent i prefers to follow his equilibrium strategy σ^e_i in all future periods, given that all other agents $j \neq i$ also follow their equilibrium strategies σ^e_j. Finally, an allocation is *incentive-compatible* if σ^* is an equilibrium strategy profile.[14]

3.2 Constrained Efficiency

With private shocks, the government can no longer observe the $t = 0$ state s_0 before designing a policy. As a result, it cannot evaluate the objective in problem (3) unless it is allowed to form beliefs in each period about agents' types, which it can then use to take expectations over the unobservable state. We could expand the shock vector s_t to include a shock $s_{g,t}$ for the government, and define a belief set $\Pi_{g,t+1}$ that specifies the government's beliefs about the agents' types; another natural option is to choose one of the agents to serve as the government and use her beliefs to take expectations over the unobservable state in each period. We use this approach below, but the results generalize to any setting in which the government's objective is of the Bergson-Samuelson form.

At $t = 0$, an agent is chosen uniformly at random to serve as the government and design a social insurance policy. Denote the index of the governing agent by $g \in \{1, \ldots, N\}$. Given social welfare weights η and the governing agent's type $s_{g,0}$, an allocation $C^*\left(s_{g,0}\right)$ is

[13] A reported state $\hat{s}^{t-1}$ does not appear in the $t = 0$ continuation utility $U_{i,0}$ because agents have not yet made reports.

[14] For any C, Nash's existence theorem and the finiteness of T, N, and Σ imply the existence of an equilibrium profile in mixed strategies. If preferences satisfy "no-hedging" (i.e., the expectation for any mixed strategy appears outside of the infimum in $U_{i,0}$) and we allow the policy functions of an allocation C to map a reported state $\hat{s}^t$ to non-degenerate lotteries, a standard revelation principle justifies our focus on pure strategy equilibria. The results below are unaffected with this interpretation of an allocation.

constrained-efficient if

$$C^*(s_{g,0}) \in \arg\max_C \inf_{\Pi(s_{g,0})} \mathbb{E}_\pi \left[\sum_i \eta_i U_{i,0}(C\,|s_{i,0})(\sigma^*) \right] \tag{7}$$

subject to non-negativity and

$$\sum_i \left[c_{i,t}\left(\hat{s}^t\right) + k_{i,t+1}\left(\hat{s}^t\right) \right] \leq \sum_i f\left(k_{i,t}\left(\hat{s}^{t-1}\right), z_{i,t}\left(\hat{s}^t\right)\right) \quad \forall t, \hat{s}^t,$$

$$U_{i,0}(C\,|s_{i,0})(\sigma^*) \geq U_{i,0}(C\,|s_{i,0})\left(\sigma^*_{-i}, \sigma_i\right) \quad \forall i, s_{i,0}, \sigma_i \in \Sigma.$$

Similarly to Section 2, we assume that the government has full commitment power in the following sense: Once the government chooses an allocation at $t \geq 0$, it commits to delivering at least the period-t continuation utility promised to any truthfully-reporting agent, regardless of any possible future reforms; the government also commits to the incentives for truthful revelation, so that under any future reform, an agent who deviates from the truth-telling strategy at t cannot receive higher t continuation utility as a result of the reform.

3.3 Periodic Reforms with Private Information

We next demonstrate that an analog of Proposition 1 holds. As with public information, the condition on agents' beliefs is that today everyone allows for the possibility of everyone believing tomorrow that some of the states will not be realized the day after tomorrow. To simplify the exposition as before, we illustrate the results in the case where everyone allows for the possibility of everyone believing tomorrow that the economy will follow its worst path.

Belief Condition

We extend the belief condition from Section 2 to the private information setting, but here it is convenient to give a more technical description.

Assume again there is a shock $\underline{s}$ such that for any agent i, any period $t \leq T-2$, any period-$t+1$ reported state $\hat{s}^{t+1}$, any period-$t+1$ type s_i^{t+1}, and any strategy profile σ,

(i) $\theta_{i,t+2}(\underline{s}) = \underline{\theta}$; and

(ii) $\pi \in \Pi_{i,t+3}\left(\hat{s}^{t+1}, \left(s_i^{t+1}, \underline{s}\right), \sigma\right)$ implies $\pi\left(s_{j,t+2} = \underline{s}\ \forall j \neq i \text{ and } s_{i,t+3} = \underline{s}\right) = 1$.

As in Section 2, any agent who receives shock $\underline{s}$ realizes the worst skill $\underline{\theta}$ and believes he will receive shock $\underline{s}$ again in the next period. With privately observed shocks, we now require that the agent also assumes the remaining agents also realized shock $\underline{s}$. We continue to let

$\underline{s}$ denote the shock and the shock vector $(\underline{s},\ldots,\underline{s})$ when the distinction is clear from the context.

As before, we also assume that the set of shocks S is rich enough so that for any shock s, there is a shock $s'(s)$ that gives the same skill to the agent, but with a belief that everyone will receive $\underline{s}$ next period. That is, there exists a subset $S' \subseteq S$ and an idempotent mapping $s' : S \to S'$ such that for any agent i, any period $t \leq T-2$, any reported state $\hat{s}^t$, any type s_i^t, and any strategy profile σ,

(i) $\theta_{i,t+1}(s'(s)) = \theta_{i,t+1}(s)$; but

(ii) $\pi \in \Pi_{i,t+2}(\hat{s}^t, (s_i^t, s'(s)), \sigma)$ if and only if $\pi(s_{i,t+2} = \underline{s}) = 1$ and there exists $\tilde{\pi} \in \Pi_{i,t+2}(\hat{s}^t, (s_i^t, s), \sigma)$ such that $\pi|_{s_{-i,t+1}} = \tilde{\pi}|_{s_{-i,t+1}} \circ (s')^{-1}$.

Here $\pi|_{s_{-i,t+1}}$ denotes the marginal distribution of $s_{-i,t+1}$ under π. We also let $\pi \circ (s')^{-1}_{i,t+1}$ denote the pushforward measure of π when s' is applied only to $s_{i,t+1}$ and the remaining agents' shocks $s_{-i,t}$ are not shifted. The condition on beliefs is then

Assumption 2. *For any $t \leq T-2$, $\hat{s}^{t-1}$, s^t, σ, i, and $\pi \in \Pi_{i,t+1}(\hat{s}^{t-1}, s_i^t, \sigma)$, we also have $\pi' \in \Pi_{i,t+1}(\hat{s}^{t-1}, s_i^t, \sigma)$, where $\pi' = \pi \circ (s')^{-1}_{i,t+1}$.*

Similarly to Assumption 1, this requires that at the beginning of t, regardless of the reported state $\hat{s}^{t-1}$ or his own type s_i^t , any agent i considers the possibility of everyone being commonly certain at $t+1$ that skills will equal $\underline{\theta}$ in all subsequent periods. The agent believes this is possible regardless of his beliefs about the other agents' current shocks $s_{-i,t}$ or $t+1$ skills θ_{t+1}.

Notice that private information introduces an additional complication because implementable allocations C must not only be feasible, but incentive-compatible. Enforcing incentives for truth-telling is potentially a more difficult task for simplified allocations, because they depend less finely on future type reports. This is not a problem, however, for the allocations that satisfy a natural notion of monotonicity: We say that an allocation C is *weakly monotone* at t if for any agent i, any reported state $\hat{s}^{t-1}$, any type s_i^t, and any strategy $\sigma_i \in \Sigma$, the agent's continuation utility $U_{i,t}\left(C \,|\hat{s}^{t-1}, s_i^t\right)\left(\sigma^*_{-i}, \sigma_i\right)$ is weakly greater than his continuation utility when the other agents are certain to report shocks $\hat{s}_{-i,t+1} \in s'\left(S^{N-1}\right)$ at $t+1$ and the shock $\hat{s}_{-i,t+\tau} = \underline{s}$ for $\tau \geq 2$. Intuitively, since the agents have maxmin preferences, weak monotonicity ensures that any agent evaluates his utility of a deviation from truth-telling based on how the deviating strategy performs when the other agents realize shocks in S' at $t+1$ and the shock $\underline{s}$ in all subsequent periods.

Optimality of Periodic Reforms

Simplified, periodically-reformed policies are optimal with private information given the belief condition in Assumption 2 and weak monotonicity:

Proposition 3. *In any period t, any feasible, incentive-compatible, weakly monotone allocation C is weakly Pareto dominated by a feasible, incentive-compatible, weakly monotone simplified allocation C^t, i.e.,*

$$U_{i,t}\big(C^t \,\big|\hat{s}^{t-1}, s_i^t\big)\big(\sigma^*\big) \geq U_{i,t}\big(C \,\big|\hat{s}^{t-1}, s_i^t\big)\big(\sigma^*\big) \quad \forall i.$$

We again give an intuitive argument here and provide a detailed proof in Appendix A.4. Let $t = 0$, and consider an allocation C that is feasible, incentive-compatible, and weakly monotone. We can define a simplified allocation C^0 so that, after any period $t \geq 0$ reported state $\hat{s}^t$, C^0 allocates consumption, effective labor, and capital according to what the original allocation C prescribes in the reported state $(\hat{s}_0, s'(\hat{s}_1), \underline{s}, \ldots, \underline{s})$. Then C^0 is feasible, and by the same argument as in the public information case, Assumption 2 implies that all agents weakly prefer C^0 to C under the truth-telling strategy profile σ^*.

However, with privately observed shocks, we must additionally verify that the simplified allocation C^0 is incentive-compatible. We already know that each agent i weakly prefers C^0 to C under the truth-telling strategy profile, and C is incentive-compatible by assumption:

$$\begin{aligned} U_{i,0}\left(C^0 \,|s_{i,0}\right)(\sigma^*) &\geq U_{i,0}\left(C\,|s_{i,0}\right)(\sigma^*) \\ &\geq \max_{\sigma_i \in \Sigma} U_{i,0}\left(C\,|s_{i,0}\right)\left(\sigma^*_{-i}, \sigma_i\right). \end{aligned}$$

By weak monotonicity, agent i's utility from any deviation σ_i is weakly lower when the other agents are certain to report a shock vector $\hat{s}_{-i,1} \in S'^{N-1}$ at $t = 1$ and the shock vector $\underline{s}$ at $t \geq 2$. Moreover, his maximum utility from a deviation σ_i is weakly lower when he is constrained to deviations σ_i that report a shock $\hat{s}_{i,1} \in S'$ at $t = 1$ and the shock $\underline{s}$ at $t \geq 2$. But these are precisely the conditions maintained by the simplified allocation C^0, so we can lower bound agent i's maximum utility from deviating under C by his maximum utility from deviating under C^0:

$$\max_{\sigma_i \in \Sigma} U_{i,0}\left(C\,|s_{i,0}\right)\left(\sigma^*_{-i}, \sigma_i\right) \geq \max_{\sigma_i \in \Sigma} U_{i,0}\left(C^0\,|s_{i,0}\right)\left(\sigma^*_{-i}, \sigma_i\right).$$

This chain of inequalities implies that C^0 is incentive-compatible. Moreover, it implies that the inequality stated in the proposition actually holds with equality.

Just as in the public information case, applying Proposition 3 to the government's problem

(7) implies that any constrained-efficient allocation C^* can be implemented by a simplified allocation C^0. However, in the presence of private information, the reform process becomes more complex. We assume that the government remains committed at $t = 1$ to the utility promises it made at $t = 0$ as well as the incentives it provided for truthful revelation. As a result, when seeking to implement a reform C^1, it must satisfy appropriate promise-keeping and threat-keeping constraints. The *promise-keeping* constraint requires that the government deliver at least as much $t = 0$ utility to every agent when the reform C^1 is substituted for the $t = 0$ allocation C^0 in periods $t \geq 1$:

$$U_{i,0}\big(C_0^0, \big(C_\tau^1\big)_{\tau=1}^T \,|s_{i,0}\big)\big(\sigma^*\big) \geq U_{i,0}\big(C^0 \,|s_{i,0}\big)\big(\sigma^*\big) \quad \forall i, s_{i,0}.$$

This is equivalent to the promise-keeping constraint in the public information reform problem (4). The *threat-keeping* constraint requires that the government punish $t = 0$ deviations from truth-telling at least as much with the reform C^1 as with the allocation C^0:

$$\begin{aligned} &U_{i,0}\big(C_0^0, \big(C_\tau^1\big)_{\tau=1}^T \,|s_{i,0}\big)\big(\sigma_{-i,0}^*, \sigma_{i,0}, \big(\sigma_\tau^*\big)_{\tau=1}^T\big) \\ &\qquad \leq U_{i,0}\big(C^0 \,|s_{i,0}\big)\big(\sigma_{-i,0}^*, \sigma_{i,0}, \big(\sigma_\tau^*\big)_{\tau=1}^T\big) \quad \forall i, s_{i,0}, \sigma_{i,0}. \end{aligned}$$

This constraint has no analog in the public information reform problem (4) because it pertains directly to how the government incentivizes truthful revelation. Along with the $t = 1$ and $t = 0$ incentive-compatibility constraints, the threat-keeping constraint ensures that every agent weakly prefers the truth-telling equilibrium at $t = 0$, even if C^1 is substituted for C^0 in the future. Next, we describe how the government constructs optimal reforms while respecting these new constraints.

Constructing Optimal Reforms

As in the public information case, the arguments above provide an algorithm for computing the simplified allocations $\{C^t\}_{t=0}^T$, without first solving for the constrained-efficient allocation C^*. Rather, the sequence can be constructed by solving for an optimal reform in each period.

Corollary 2. *Optimal simplified allocations $\{C^t\}_{t=0}^T$ can be constructed period by period, without computing the fully state-contingent allocation C^*.*

In general, at any $t \geq 0$ the government seeks to design a reform that maximizes period-t social welfare. Suppose the allocation was last reformed in period $r < t$. Given the reported state $\hat{s}^{t-1}$ and the governing agent's type s_g^t, an optimal reform $C^*\left(\hat{s}^{t-1}, s_g^t, C^r\right)$ is a solution to the reform problem

$$\max_{C} \inf_{\Pi_{g,t+1}\left(\hat{s}^{t-1}, s_g^t\right)} \mathbb{E}_\pi \left[\sum_i \eta_i U_{i,t}\left(C \left| \hat{s}^{t-1}, s_i^t\right.\right)\left(\sigma^*\right) \right] \tag{8}$$

subject to non-negativity and

$$\sum_i \left[c_{i,\tau}\left(\hat{s}^\tau\right) + k_{i,\tau+1}\left(\hat{s}^\tau\right)\right] \leq \sum_i f\left(k_{i,\tau}\left(\hat{s}^{\tau-1}\right), z_{i,\tau}\left(\hat{s}^\tau\right)\right) \quad \forall \tau \geq t, \hat{s}^\tau \geq \hat{s}^{t-1},$$

$$U_{i,t}\left(C \left| \hat{s}^{t-1}, s_i^t\right.\right)\left(\sigma^*\right) \geq U_{i,t}\left(C \left| \hat{s}^{t-1}, s_i^t\right.\right)\left(\sigma_{-i}^*, \sigma_i\right) \quad \forall i, s_i^t, \sigma_i \in \Sigma,$$

$$U_{i,r}\left(\left(C_\tau^r\right)_{\tau=r}^{t-1}, \left(C_\tau\right)_{\tau=t}^{T} \left| \hat{s}^{r-1}, s_i^r\right.\right)\left(\sigma^*\right) \geq U_{i,r}\left(C^r \left| \hat{s}^{r-1}, s_i^r\right.\right)\left(\sigma^*\right) \quad \forall i, s_i^r$$

$$U_{i,r}\left(\left(C_\tau^r\right)_{\tau=r}^{t-1}, \left(C_\tau\right)_{\tau=t}^{T} \left| \hat{s}^{r-1}, s_i^r\right.\right)\left(\left(\sigma_{-i,\tau}^*, \sigma_{i,\tau}\right)_{\tau=r}^{t-1}, \left(\sigma_\tau^*\right)_{\tau=t}^{T}\right)$$
$$\leq U_{i,r}\left(C^r \left| \hat{s}^{r-1}, s_i^r\right.\right)\left(\left(\sigma_{-i,\tau}^*, \sigma_{i,\tau}\right)_{\tau=r}^{t-1}, \left(\sigma_\tau^*\right)_{\tau=t}^{T}\right) \quad \forall i, s_i^r, \left(\sigma_{i,\tau}\right)_{\tau=r}^{t-1},$$

Note that the threat-keeping constraint in this problem involves punishments for multi-period deviations from truth-telling because the allocation may not have been reformed in the previous period. When $t = 0$, there are no promise-keeping or threat-keeping constraints.

If there exists an allocation in the constraint set and the optimal allocation is weakly monotone, arguments analogous to those in the proof of Proposition 3 imply that the government can choose a simplified allocation C^t to solve this reform problem (see Appendix A.5 for details). If at any $t > 0$ the constraint set is empty,[15] then the government sets $C^t \equiv C^{t-1}$ and picks the equilibrium σ^e of C^t with respect to the period-t utility functions $U_{i,t}$ that maximizes period-t social welfare

$$\inf_{\Pi_{g,t+1}\left(\hat{s}^{t-1}, s_g^t\right)} \mathbb{E}_\pi \left[\sum_i \eta_i U_{i,t}\left(C^t \left| \hat{s}^{t-1}, s_i^t\right.\right)\left(\sigma^e\right) \right].$$

This process then constructs a sequence of simplified allocations $\{C^t\}_{t=0}^T$ that implements the constrained-efficient allocation C^*, without first solving for C^*.

History Independence

Similarly to the public information case in Section 2.3, the structure of the reform problem (8) characterizes conditions under which optimal policies may lose full dependence on the history of past reports $\hat{s}^{t-1}$. In particular, suppose that at t, each agent i's belief set $\Pi_{i,t+1}\left(\hat{s}^{t-1}, s_i^t, \sigma\right)$ does not depend on $\hat{s}^{t-2}$.[16] If a reform at t can provide an improvement to previously designed

[15] If C^r is a simplified allocation, the constraint set must be non-empty at either $r+1$ or $r+2$ because C^r does not depend on reports made in any period $t \geq r+2$.

[16] For example, this will hold if the agents followed the truth-telling strategy profile σ^* at $t-1$ and if the belief sets $\Pi_{i,t+1}\left(\hat{s}^{t-1}, s_i^t, \sigma\right)$ are derived by using Bayes' rule to update a fixed set of priors under which each

government policies, i.e., if the promise- and threat-keeping constraints in the reform problem (8) do not bind at t, then the optimal reform C^t will not depend on the $t-2$ reported state $\hat{s}^{t-2}$. The reasoning for this is analogous to that in Section 2.3: If $\Pi_{i,t+1}\left(\hat{s}^{t-1}, s_i^t, \sigma\right)$ does not depend on $\hat{s}^{t-2}$ for each agent i, then the objective function and the incentive-compatibility constraint depend on $\hat{s}^{t-2}$ only through C, and the feasibility constraint depends on $\hat{s}^{t-2}$ only through C and the (fixed) period-t capital distribution $k_t^{t-1}\left(\hat{s}^{t-1}\right)$. Because of this, the government will choose C^t so that it does not depend on $\hat{s}^{t-2}$, conditional on the capital distribution $k_t^{t-1}\left(\hat{s}^{t-1}\right)$.

With private information, such improvement and consequent loss of full history dependence require that the threat-keeping constraint be non-binding at t in addition to the promise-keeping constraint. Since the threat-keeping constraint does not appear in the public information reform problem (4), policies may lose history dependence less frequently when agents have private information about shocks $s_{i,t}$. This is because the government must provide incentives for truthful revelation in the current period, so it will generally condition an agent's future consumption and effective labor on his current report. The government also commits to maintaining these incentives even after subsequent reforms, and this may require that reforms maintain some dependence on the history of shocks. In contrast, with no uncertainty, loss of history dependence is generically suboptimal and can confer large welfare losses over policies with full history dependence (see, e.g., Kapička 2017).

4 (In)Efficiency of Competitive Equilibria

In conventional, rational expectations versions of the economies above, competitive equilibria result in constrained-efficient allocations even when types are privately known to agents.[17] A common interpretation is that government-provided social insurance only crowds out insurance provided by private markets (e.g., Acemoglu and Simsek 2012). In this section, we discuss conditions under which this efficiency result survives in the presence of broader uncertainty, and conditions under which it breaks. The deciding factor is the relationship between (i) the firms' uncertainty and (ii) the government's uncertainty as expressed through its feasibility constraints: Whenever firms entertain a greater degree of uncertainty about the shock process than the government does, competitive equilibria may fail to provide efficient

agent's shock follows a Markov process.

[17]This holds when allocations are designed ex ante, or equivalently when at $t=0$ agents are homogeneous and have observable shocks. The rational expectations versions of these economies closely resemble the "moral hazard" economies in Prescott and Townsend (1984), which they show can be decentralized efficiently. This should be distinguished from their "adverse selection" economies, in which competitive equilibria are generically inefficient (see, e.g., Bisin and Gottardi 2006).

insurance. In other words, a broader view of uncertainty and resulting belief heterogeneity may create a meaningful role for the government provision of insurance.

Nevertheless, we also show that the results of previous sections persist in that insurance in a decentralized economy can be provided with sequences of simplified allocations that are periodically reformed.

4.1 Decentralization

Consider once again the economy of Section 3 in which agents are privately informed about their shocks $s_{i,t}$. As before, we illustrate our arguments in a simple special case and then discuss their generality. For direct contrast with rational expectations economies where competitive equilibria are known to be efficient, we consider an analogue of the "private-information labor market" economy of Prescott and Townsend (1984), modified to introduce broader uncertainty:[18] There are two periods $t \in \{0, 1\}$, with period 0 interpreted as an ex ante period in which agents are homogeneous and have publicly observable preferences. At $t = 0$ each agent receives the same publicly observed shock. Since this shock is fixed and observable, we suppress it in the notation below. Period-0 consumption c_0, effective labor z_0, and capital investment k_1 are also taken to be fixed and equal across agents. Shocks remain private information at $t = 1$, so an allocation $C = \{c_1(s_1), z_1(s_1)\}$ is incentive-compatible if it induces truthful revelation at $t = 1$:[19]

$$U_{i,1}(C|\, s_{i,1})(\sigma^*) \geq U_{i,1}(C|\, s_{i,1})\left(\sigma^*_{-i}, \sigma_i\right) \quad \forall i, s_{i,1}, \sigma_i \in \Sigma.$$

To ensure that they are homogeneous at $t = 0$, all agents must share a common belief set $\Pi_{A,1} \subseteq \Delta(S)$ that describes each agent's beliefs about his $t = 1$ shock. Similarly, at $t = 1$ the skill mappings $\theta_{i,1}$ and the belief mappings $\Pi_{i,2}$ are the same across agents.[20] We assume that the belief set $\Pi_{i,2}(s) \subseteq \Delta\left(S^{N-1}\right)$ for $s \in S$ is *exchangeable* in the sense that for any belief $\pi \in \Pi_{i,2}(s)$, we also have $\pi \circ (s^p)^{-1} \in \Pi_{i,2}(s)$ for any permutation p of the indices $\{1, \ldots, N\} \setminus \{i\}$, where $s^p(s_{-i}) \equiv \left(s_{p(j),1}\right)_{j \neq i}$.

To decentralize this economy, we introduce identical competitive firms that offer insurance contracts (allocations) to the agents. At $t = 0$, a (representative) firm forms an exchangeable set of beliefs $\Pi_{F,1} \subseteq \Delta\left(S^N\right)$ about the agents' $t = 1$ shocks that it uses to evaluate shock-contingent streams of profits. In particular, at $t = 0$ the firm assigns a given stream of profits

[18]Strictly speaking, Prescott and Townsend (1984) show efficiency in an economy with a representative agent; the arguments below also apply in this case.

[19]The notion of incentive-compatibility here is consistent with that in Section 3, because we have simply relabeled the first period in which agents make reports.

[20]Consistent with Section 3, $\Pi_{i,2}(s)$ here is a set of distribution over the other agents' period-1 types $s_{-i,1}$.

$d(s_1)$ the value

$$V(d) \equiv \inf_{\pi \in \Pi_{F,1}} \mathbb{E}_{\pi}\left[d(s_1)\right]. \tag{9}$$

With this objective, the firm is not averse to risk, but it is potentially averse to uncertainty about the distribution of shocks at $t = 1$.

At $t = 0$, the firm chooses an allocation C to maximize its objective, subject to incentive-compatibility at $t = 1$ and the constraint that it provide equilibrium reservation utility $\underline{U}$ to each agent:

$$\begin{aligned} &\max_{C} V(d) \qquad\qquad\qquad\qquad (10)\\ &\text{subject to non-negativity and}\\ &\sum_i c_{i,1}(\hat{s}_1) + d(\hat{s}_1) = \sum_i f(k_{i,1}, z_{i,1}(\hat{s}_1)) \quad \forall \hat{s}_1,\\ &U_{i,1}(C|\, s_{i,1})(\sigma^*) \geq U_{i,1}(C|\, s_{i,1})\left(\sigma^*_{-i}, \sigma_i\right) \quad \forall i, s_{i,1}, \sigma_i \in \Sigma,\\ &U_{i,0}(C)(\sigma^*) \geq \underline{U} \quad \forall i. \end{aligned}$$

A *competitive equilibrium* is an allocation C^e and reservation utility $\underline{U}$ such that

(i) C^e solves the firm's problem (10) and delivers non-negative value to the firm;

(ii) each agent i contracts with the firm offering the best allocation and receives utility $U_{i,0}(C^e)(\sigma^*) = \underline{U}$; and

(iii) the feasibility constraints (2) hold for $t = 1$.

Consistent with Section 3, an allocation C^* is constrained-efficient if

$$C^* \in \arg\max_{C} \sum_i U_{i,0}(C)(\sigma^*),$$

subject to non-negativity, feasibility, and incentive-compatibility. Note that we restrict to equal Pareto weights $\eta_i = 1$ because agents are homogeneous at $t = 0$.

4.2 (In)Efficiency

To ground our discussion, it is useful to begin with an example of the economy above in which competitive equilibria continue to be efficient in the presence of broader uncertainty: Let $\Pi_{F,1} = \Delta\left(S^N\right)$ so that firms are maximally uncertain, and suppose toward a contradiction that a competitive equilibrium is inefficient, i.e., $U_{i,0}(C^*)(\sigma^*) > \underline{U}$. Then C^* must be in the constraint set of the firm's problem. By reducing each agent i's consumption after each shock vector s_1 by a sufficiently small amount, the firm can modify C^* so that it obtains positive profits $d(s_1) > 0$ for each shock vector while respecting the incentive-compatibility and reservation utility constraints. This implies that the equilibrium allocation C^e does not solve the firm's problem, a contradiction.

Efficiency survives in this example because a familiar duality between the government's problem and the firm's problem is maintained. This is apparent if we note that, since $\Pi_{F,1} = \Delta\left(S^N\right)$, the feasibility constraints (2) observed by the government can be written

$$\inf_{\Pi_{F,1}} \mathbb{E}_\pi \left\{ \sum_i \left[f\left(k_{i,1}, z_{i,1}(s_1)\right) - c_{i,1}(s_1) \right] \right\} \geq 0.$$

The left side of this inequality is precisely the value that the firm derives from the allocation $C = \{c_1(s_1), z_1(s_1)\}$, which must equal zero in equilibrium. Intuitively, agents must obtain at least utility $U_{i,0}(C^*)$ in equilibrium due to competitive pressures; they cannot obtain more, because the symmetry between the firm's objective and the government's feasibility constraint would imply that the equilibrium allocation is feasible for the government and delivers greater utility than the constrained-efficient allocation.

This suggests that competitive equilibria may not be efficient when this duality between the government's problem and the firm's problem is broken. In an economy with uncertainty about the distribution of agents' shocks, this happens naturally whenever firms entertain a greater or lesser degree of uncertainty than the government. For a sharp example, suppose that firms are still maximally uncertain, with $\Pi_{F,1} = \Delta\left(S^N\right)$. Suppose, however, that the agents and the government are Bayesian, with a single belief $\pi^* \in \Delta\left(S^N\right)$ at $t = 0$ about the distribution of the $t = 1$ shock vector s_1. Then a natural treatment of feasibility, consistent with the approach in rational expectation economies, is to require that the government maintain positive expected net resources under π^*:

$$\mathbb{E}_{\pi^*} \left\{ \sum_i \left[f\left(k_{i,1}, z_{i,1}(s_1)\right) - c_{i,1}(s_1) \right] \right\} \geq 0. \tag{11}$$

This constraint is weaker than the more stringent pointwise constraints (2). When the pointwise constraints are replaced with (11) in the government's problem and in the definition of a competitive equilibrium, competitive equilibria do not generally deliver the constrained-efficient level of utility:

Proposition 4. *Competitive equilibria with uncertainty may not be efficient.*

To see this, suppose that π^* is such that the agents' shocks are independently and identically distributed, and that each agent realizes skills $\underline{\theta}$ and $\bar{\theta}$ with equal probability. The constrained-efficient allocation C^* under the weakened feasibility constraint (11) will generally feature a deficit for the state in which each agent realizes skill $\bar{\theta}$. However, such an allocation would deliver negative value to the firm under the objective (9), so it cannot be sustained in a competitive equilibrium. As a result, any competitive equilibrium must be inefficient.

More broadly, this discussion shows that if the government and firms form heterogeneous beliefs in response to uncertainty about the data-generating process, the government and competitive markets may implement different outcomes. In this sense, broader uncertainty can be viewed as a friction that prevents the proper functioning of competitive markets. In particular, whenever the firms entertain a greater degree of uncertainty about the shock process than the government, competitive equilibria may fail to provide the efficient amount of insurance.[21]

4.3 Periodic Reforms in Equilibrium

Even with the potential for inefficiency, agents may still obtain some degree of insurance in a decentralized economy. We next argue that any insurance provided by competitive firms can be obtained with simplified allocations that are periodically reformed.

Consider a dynamic extension of the decentralized economy described above: Firms and agents continue to contract at $t = 0$, but at the end of each period t firms may also trade a risk-free bond b in zero net supply that pays one unit of consumption at $t+1$. Let $q_t\left(\hat{s}^t, \sigma\right)$ denote the equilibrium price of the risk-free bond after reported history $\hat{s}^t$ when agents follow the strategy profile σ. Extending the firm's problem from Section 4.1, the firm chooses an allocation C and bond purchases $\left\{b_t\left(\hat{s}^t\right)\right\}_{t=0}^T$ to maximize its period-1 value $V_{i,1}$. Given a reported state-contingent stream of profits $D \equiv \{d_t\}_{t=0}^T$, a $t-1$ reported state $\hat{s}^{t-1}$, and a strategy profile σ for the agents, the firm's period-t continuation value is

$$V_t\left(D|\, \hat{s}^{t-1}\right)(\sigma) \equiv \inf_{\Pi_{F,t}(\hat{s}^{t-1},\sigma)} \mathbb{E}_\pi \Big[d_t\left(\hat{s}^{t-1}, \sigma_t\left(\hat{s}^{t-1}, s^t\right)\right) \\ + q_t\left(\hat{s}^{t-1}, \sigma_t\left(\hat{s}^{t-1}, s^t\right), \sigma\right) V_{t+1}\left(D|\, \hat{s}^{t-1}, \sigma_t\left(\hat{s}^{t-1}, s^t\right)\right)(\sigma)\Big],$$

where $\Pi_{F,t}\left(\hat{s}^{t-1}, \sigma\right) \subseteq \Delta\left(S^{N(t+1)}\right)$ is the firm's set of beliefs about the period-t state s^t.

A competitive equilibrium is then defined as in Section 4.1, extended to include the market-clearing condition for $b_t\left(\hat{s}^t\right)$ and the baseline feasibility constraints (2) for all t and all $\hat{s}^t$. If we again maintain Assumption 2, as well as its analogue for the firm's beliefs

[21] Interestingly, a symmetric argument to the one given above suggests that if the firms entertain less uncertainty than the government, competitive equilibria may deliver *greater* utility to agents than the constrained-efficient allocation.

$\Pi_{F,t}\left(\hat{s}^{t-1},\sigma\right)$, a decentralized version of Proposition 3 holds:

Proposition 5. *For any set of prices* $\left\{q\left(\hat{s}^t,\sigma^*\right)\right\}_{t=0}^{T-1}$, *any weakly monotone competitive equilibrium allocation* C^e *with associated profit stream* D^e *is weakly Pareto dominated by a weakly monotone simplified allocation* C^0 *with associated profit stream* D^0, *i.e.,*

$$U_{i,0}\left(C^0\left|s_0\right.\right)(\sigma^*) \geq U_{i,0}\left(C^e\left|s_0\right.\right)(\sigma^*) \quad \forall s_0, \forall i,$$
$$V_0\left(D^0\right)(\sigma^*) \geq V_0\left(D^e\right)(\sigma^*).$$

The intuition for the result and the construction of the simplified allocation C^0 are essentially the same as in Proposition 3.

5 Linearity

Finally, we consider whether even simpler linear or affine policies can be optimal.[22] In particular, we ask when an optimal simplified allocation C^t can be implemented with fiscal policies that are affine in individual income, $f\left(k_{i,t},z_{i,t}\right)$. We argue that this is not generally the case: One must place strong assumptions on agents' beliefs and on allocations for affine policy functions to be optimal, and there are substantive limitations to generalizations. For example, risk aversion provides a significant roadblock, as does elastically supplied labor. When linearity does persist, we show that it is with respect to an agent's skill shock, which is generally not equivalent to linearity in income.

It suffices to consider again the baseline public information economy of Section 2. We maintain Assumption 1, and we additionally assume the following:

Assumption 3. *For any* $t \leq T-1$, s^t, i *and* $\pi \in \Pi_{i,t+1}\left(s^t\right)$, *there also exists* $\tilde{\pi} \in \Pi_{i,t+1}\left(s^t\right)$, *such that*

1. $\theta_{i,t+1}$ *and* $\theta_{-i,t+1}$ *are independent under* $\tilde{\pi}$;
2. $\tilde{\pi}$ *and* π *imply the same marginal distribution of* $\theta_{-i,t+1}$;
3. *the distribution of* $\theta_{i,t+1}$ *under* $\tilde{\pi}$ *places weight only on* $\left\{\underline{\theta},\bar{\theta}\right\}$ *to satisfy* $\mathbb{E}_{\tilde{\pi}}\left[\theta_{i,t+1}\right] = \mathbb{E}_{\pi}\left[\theta_{i,t+1}\right]$.

[22] This linearity conjecture is motivated in part by the findings that, under some conditions, uncertainty can lead to linearity in financial contracting (see, e.g., Carroll 2015, Zhu 2016). A further, applied motivation is strongly suggested by a vast Ramsey optimal taxation literature that typically starts with an ad hoc restriction to linear or affine policy instruments.

For example, this assumption is satisfied if at each t, each agent i considers all beliefs π such that his $t+1$ skill $\theta_{i,t+1}$ is independently distributed from the other agents' skills $\theta_{-i,t+1}$ and such that the expectation of $\theta_{i,t+1}$ under π equals some fixed value $\tilde{\theta}$. We require independence to ensure that agent i's beliefs about other agents' skills do not change how he evaluates expectations with respect to his own skill. This is essential to the arguments, because the linearity conjecture concerns policies or continuation utilities that are linear with respect to agent i's skill, holding other agents' skills fixed. In addition, we note that by contrast with Assumption 1, the arguments below do not generalize immediately under weaker versions of Assumption 3.

We maintain Assumption 1, so Proposition 1 implies that the efficient allocation C^* can be implemented by a sequence of simplified allocations $\{C^t\}_{t=0}^T$. For the arguments below it is also crucial to note that in the period-t allocation C^t, the consumption and effective labor functions $\left\{c_{i,\tau}^t, z_{i,\tau}^t\right\}_{\tau=t}^T$ for each agent i depend only on s^t and $\theta_{t+1} \equiv (\theta_{1,t+1}, \ldots, \theta_{N,t+1})$. Recall that this holds because C^t can be constructed by assuming that all agents will realize a shock $s' \in S'$ at $\tau = t+1$, implying the degenerate belief set $\Pi_{i,t+2}(s^t, s'(s_{t+1})) = \{\underline{\pi}\}$, and the shock $\underline{s}$ at $\tau \geq t+2$.

Let us now focus on conditions under which the solution to the period-t reform problem (4) features linear (affine) policy functions or linearity in agents' flow utilities with respect to their skills. We address the latter in detail, and at the end of the section we describe the additional assumptions needed for linear policy functions to be optimal. To state the result, we begin by examining the solution $C^0(s_0)$ to the government's $t=0$ problem and considering how to modify it so that agent i's $t=1$ continuation utility is affine in his skill shock $\theta_{i,1}$.

To make the argument as clear as possible, we illustrate it in the context of Figure 1. Holding $(s_0, \theta_{-i,1})$ fixed, let $u_{i,1}^0(\theta_{i,1}) \equiv u\left(c_{i,1}^0(s_0, \theta_1), z_{i,1}^0(s_0, \theta_1)/\theta_{i,1}\right)$ denote agent i's $t=1$ flow utility, the solid line in Figure 1. This function is typically not affine and so does not coincide with its secant line from $\theta_{i,1} = \underline{\theta}$ to $\theta_{i,1} = \bar{\theta}$, the dot-dashed line in Figure 1. Define new policy functions $\hat{c}_{i,1}, \hat{z}_{i,1}$ such that for any fixed $(s_0, \theta_{-i,1})$, $\hat{u}_{i,1}(\theta_{i,1}) \equiv u(\hat{c}_{i,1}, \hat{z}_{i,1}/\theta_{i,1})$ is equal to $u_{i,1}^0(\theta_{i,1})$ for $\theta_{i,1} \in \left\{\underline{\theta}, \bar{\theta}\right\}$ but is affine over Θ. That is, $\hat{u}_{i,1}(\theta_{i,1})$ is the function of the secant line of $u_{i,1}^0(\theta_{i,1})$ from $\theta_{i,1} = \underline{\theta}$ to $\theta_{i,1} = \bar{\theta}$, and its graph is given by the dot-dashed line in Figure 1.

If feasible, the government would like to deliver utility $\hat{u}_{i,1}(\theta_{i,1})$ wherever $u_{i,1}^0(\theta_{i,1})$ falls below its secant line. In particular, the government would like to use the policy functions

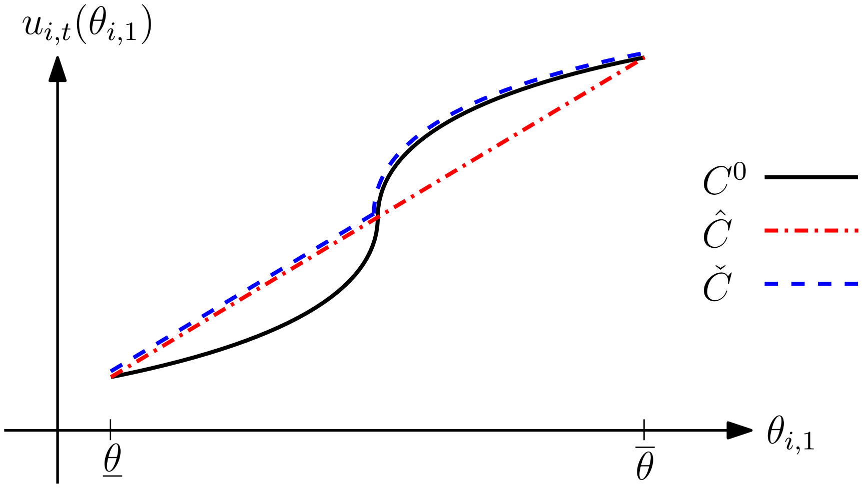

Figure 1: Linearity in agents' utilities with respect to their skills.

$\check{c}_{i,1}, \check{z}_{i,1}$ defined by

$$\left(\check{c}_{i,1}\left(s_0,\theta_1\right),\check{z}_{i,1}\left(s_0,\theta_1\right)\right)=\begin{cases}\left(\hat{c}_{i,1}\left(s_0,\theta_1\right),\hat{z}_{i,1}\left(s_0,\theta_1\right)\right) & \text{if } \hat{u}_{i,1}\left(\theta_{i,1}\right)\geq u^0_{i,1}\left(\theta_{i,1}\right), \\ \left(c^0_{i,1}\left(s_0,\theta_1\right),z^0_{i,1}\left(s_0,\theta_1\right)\right) & \text{else.}\end{cases}$$

The agent's utility $\check{u}_{i,1}\left(\theta_{i,1}\right)\equiv u\left(\check{c}_{i,1},\check{z}_{i,1}/\theta_{i,1}\right)$ under these new policy functions is then the maximum of the utilities $\hat{u}_{i,1}\left(\theta_{i,1}\right)$ and $u^0_{i,1}\left(\theta_{i,1}\right)$, given by the dashed line in Figure 1.

For $t>1$, we define $\hat{c}_{i,t}, \hat{z}_{i,t}$ and $\check{c}_{i,t}, \check{z}_{i,t}$ similarly to that above: Agent i's period t flow utility under allocation C^0 is given by $u\left(c^0_{i,t}\left(s_0,\theta_1\right),z^0_{i,t}\left(s_0,\theta_1\right)/\underline{\theta}\right)$, so we let $\hat{c}_{i,t}, \hat{z}_{i,t}$ be such that for fixed $\left(s_0,\theta_{-i,1}\right)$, $u\left(\hat{c}_{i,1},\hat{z}_{i,1}/\underline{\theta}\right)$ is the function of the secant line of $u\left(c^0_{i,1},z^0_{i,1}/\underline{\theta}\right)$ from $\theta_{i,1}=\underline{\theta}$ to $\theta_{i,1}=\bar{\theta}$. We then define $\check{c}_{i,t}, \check{z}_{i,t}$ analogously to how $\check{c}_{i,1}, \check{z}_{i,1}$ are defined, delivering the maximum of the period-t flow utilities under the other two allocations. Finally, at $t=0$ let

$$\left\{\check{c}_{i,0},\check{z}_{i,0}\right\}=\left\{\hat{c}_{i,0},\hat{z}_{i,0}\right\}=\left\{c^0_{i,0},z^0_{i,0}\right\}.$$

Feasibility constraints may preclude the government from using the policy functions $\left\{\check{c}_{i,t},\check{z}_{i,t}\right\}_{t=0}^{T}$, but as Figure 1 indicates, these are clearly preferred by agent i to $\left\{c^0_{i,t},z^0_{i,t}\right\}_{t=0}^{T}$ and $\left\{\hat{c}_{i,t},\hat{z}_{i,t}\right\}_{t=0}^{T}$.

We next show that the agent is actually indifferent between the policy functions $\left\{\check{c}_{i,t},\check{z}_{i,t}\right\}_{t=0}^{T}$ and $\left\{\hat{c}_{i,t},\hat{z}_{i,t}\right\}_{t=0}^{T}$.

Proposition 6. *Let* $\hat{C}_i\equiv\left\{\hat{c}_{i,t},\hat{z}_{i,t},k^0_{i,t+1}\right\}_{t=0}^{T}$, $\check{C}_i\equiv\left\{\check{c}_{i,t},\check{z}_{i,t},k^0_{i,t+1}\right\}_{t=0}^{T}$. *Then*

$$U_{i,0}\left(\hat{C}_i\left|s_0\right.\right)=U_{i,0}\left(\check{C}_i\left|s_0\right.\right).$$

We provide a proof in Appendix A.6, but the idea can be seen in Figure 1. When considered as a function of $\theta_{i,1}$ with fixed $(s_0, \theta_{-i,1})$, agent i's flow utility at $t \geq 1$ under $\check{C}_i$ (the dashed line) lies weakly above its secant line from from $\theta_{i,1} = \underline{\theta}$ to $\theta_{i,1} = \bar{\theta}$ (the dot-dashed line). The "worst" belief distributions $\tilde{\pi} \in \Pi_{i,1}(s_0)$, and thus the ones considered by the agent when evaluating his $t = 1$ continuation utility, are then those that are supported on $\{\underline{\theta}, \bar{\theta}\}$. However, this only holds because of the independence of $\theta_{i,1}$ and $\theta_{-i,1}$ under each $\pi \in \Pi_{i,1}(s_0)$. If independence is violated and the distribution of $\theta_{i,1}$ conditional on $\theta_{-i,1}$ changes with $\theta_{-i,1}$, agent i's expected $t = 1$ utility with the belief π may instead be weakly lower than his expected $t = 1$ utility with the corresponding belief $\tilde{\pi}$.

Since agent i only considers distributions of the form $\tilde{\pi} \in \Pi_{i,1}(s_0)$ when evaluating his expected $t = 1$ utility, it is easy to see that he is indifferent between $\check{C}_i$ and $\hat{C}_i$. Under $\hat{C}_i$, the agent's $t \geq 1$ flow utility is affine in $\theta_{i,1}$, and it coincides with the flow utility under $\check{C}_i$ when $\theta_{i,1} \in \{\underline{\theta}, \bar{\theta}\}$. Distributions of the form $\tilde{\pi}$ place weight only on events with $\theta_{i,1} \in \{\underline{\theta}, \bar{\theta}\}$, so agent i's expected $t = 1$ utility is the same under $\check{C}_i$ and $\hat{C}_i$. These allocations also give him the same $t = 0$ flow utility, so the agent is indifferent between $\check{C}_i$ and $\hat{C}_i$.

Proposition 6 implies that the government weakly prefers the allocation $\hat{C}_i$, for which agent i's $t \geq 1$ flow utility is affine in $\theta_{i,1}$, to C_i^0. The government will always seek to design $t = 0$ allocations in which each agent's flow utility at $t \geq 1$ is affine in his own $t = 1$ skill. However, the assumptions needed to prove this result are so strong that, on the basis of the argument above, affine policies cannot be expected to be optimal in practice. For example, the proof of Proposition 6 makes heavy use of the independence condition in Assumption 3, but it is not clear why an uncertain agent would restrict his beliefs to product distributions. Another serious issue regards feasibility: We can clearly assume without loss of generality that the feasibility constraints in the $t = 0$ government's problem will bind, so the intermediate allocation $\check{C}_i$ constructed above is almost certain to be infeasible. In this case, the allocation $\hat{C}_i$ may also be infeasible.

Moreover, to obtain a more concrete property we must make the additional assumption that labor supply is inelastic. In particular, suppose that at $t = 0$, the government is constrained so that at each $t \geq 1$, agent i will exert some fixed amount of labor $\bar{l}_{i,t}(s_0)$. If this is the case, then agent i's $t = 1$ skill $\theta_{i,1}$ affects his $t \geq 1$ flow utility only through consumption, and we can use the same methods as above to show that the government weakly prefers consumption functions that are affine in an agent's own skill. The following proposition collects these observations:

Proposition 7. *Fix i. If $t \geq 1$ labor supply is inelastic, agent i weakly prefers* $\left\{c_{i,0}^0, \hat{c}_{i,t}\right\}_{t=1}^T$

to $\left\{c^0_{i,t}\right\}_{t=0}^T$, *where* $\hat{c}_{i,t}$ $t \geq 1$ *is the affine consumption function given by*

$$\hat{c}_{i,t}\left(s_0, \theta_1\right) \equiv \frac{\bar{\theta} - \theta_{i,1}}{\bar{\theta} - \underline{\theta}} c^0_{i,t}\left(s_0, \left(\underline{\theta}, \theta_{-i,1}\right)\right) + \frac{\theta_{i,1} - \underline{\theta}}{\bar{\theta} - \underline{\theta}} c^0_{i,t}\left(s_0, \left(\bar{\theta}, \theta_{-i,1}\right)\right).$$

6 Concluding Remarks

This paper studied the optimal policy implications of uncertainty about the distribution of shocks in an otherwise conventional dynamic economy, and characterized general properties of policies that are robust with respect to such uncertainty. We described conditions under which optimal policies are simplified in the sense that they are not fully contingent on future shocks, lose dependence on the full history of past shocks, and are reformed periodically, consistent with what is commonly observed in reality. We argued, however, that restrictive assumptions are required for linear policies to be optimal. In contrast to rational expectations environments, decentralized versions of these economies are not generally efficient, implying a potentially meaningful role for the government provision of insurance.

While this paper focused on social insurance and fiscal policies, the above insights are applicable to risk-sharing environments more broadly. We believe that the paper also opens a number of interesting questions for future research. One significant possible extension is to environments where exogenous lack of commitment on the part of the agents is a salient friction. Applications include wage contracting between a firm and its workers who are free to walk away from long-term contracts, informal insurance arrangements in village economies, as well as other contexts in development economics. State-run (rather than federally-run) risk-sharing programs are another example of broader risk-sharing environments where enforcement of contracts may be difficult given relatively low costs of moving to a different state.

One approach to such situations is to characterize contracts that are self enforcing, i.e., that provide incentives to stay within the contract. When uncertainty about the distribution of payoff-relevant variables of the kind discussed in this paper is also present, our methods can be used to characterize properties of optimal contracts. In particular, self-enforcement constraints with an exogenously specified outside option can be included in our optimal reform problems, and by solving such reform problems in each period, a sequence of optimal, simplified, self-enforcing allocations can be characterized without the need to construct a complete constrained-efficient allocation. This property may provide novel insights about optimal self-enforcing contracts in environments with uncertainty.

References

[1] Acemoglu, Daron and Alp Simsek, "Moral Hazard and Efficiency in General Equilibrium with Anonymous Trading," working paper, August 2012.

[2] Aghion, Philippe, Ufuk Akcigit, Matthieu Lequien, and Stefanie Stantcheva, "Tax Simplicity and Heterogeneous Learning," NBER Working Paper, No. 24049, November 2017.

[3] Barlevy, Gadi, "Robustness and Macreoconomic Policy," Annual Reviews of Economics 3: 1-14, 2011.

[4] Benigno, Pierpaolo, and Luigi Paciello, "Monetary Policy, Doubts and Asset Prices," Journal of Monetary Economics 64: 85-98.

[5] Bhandari, Anmol, "Doubts, Asymmetries, and Insurance," working paper, September 2015.

[6] Bhandari, Anmol, Jaroslav Borovička, and Paul Ho, "Survey data and subjective beliefs in business cycle models," working paper, November 2017.

[7] Bisin, Alberto, and Piero Gottardi, "Competitive Equilibria with Asymmetric Information," Journal of Economic Theory 87 (1): 1-48, 1999.

[8] Bisin, Alberto, and Piero Gottardi, "Efficient Competitive Equilibria with Adverse Selection," Journal of Political Economy 114 (3): 485-516, 2006.

[9] Bloom, Nicholas, "Fluctuations in Uncertainty," Journal of Economic Perspectives 28 (2):153-176, 2014.

[10] Carroll, Gabriel, "Robustness and Linear Contracts," American Economic Review, 105 (2), 536-63, 2015.

[11] Carroll, Gabriel, "Robustness in Mechanism Design and Contracting," Annual Reviews of Economics 11: 139-166, 2019.

[12] Epstein, Larry, and Martin Schneider, "Recursive Multiple Priors," Journal of Economic Theory, 113 (1): 32-50, 2003.

[13] Farhi, Emmanual, Christopher Sleet, Iván Werning, and Sevin Yeltekin, "Non-linear Capital Taxation without Commitment," Review of Economic Studies 79: 1469–1493, 2012.

[14] Farhi, Emmanuel and Iván Werning, “Insurance and Taxation over the Life Cycle,” Review of Economic Studies 80 (2): 596-635, 2013.

[15] Hansen, Lars P. and Thomas J. Sargent, “Robust Control and Model Uncertainty,” American Economic Review 91 (2): 60-66, 2001.

[16] Hansen, Lars P. and Thomas J. Sargent, “Time Inconsistency of Robust Control?” November 2006.

[17] Hansen, Lars P. and Thomas J. Sargent, “Three Types of Ambiguity,” Journal of Monetary Economics 59 (5): 422-445, 2012.

[18] Hansen, Lars P., Thomas J. Sargent, Gauhar A. Turmuhambetova, and Noah Williams, “Robust Control and Model Misspecification,” Journal of Economic Theory 128 (1):45-90, 2006.

[19] Kapička, Marek, “Quantifying the Welfare Gains from History Dependent Income Taxation,” working paper, August 2017.

[20] Karantounias, Anastasios G., “Managing Pessimistic Expectations and Fiscal Policy,” Theoretical Economics 8 (1): 193-231, 2013.

[21] Kocherlakota, Narayana R., “The New Dynamic Public Finance,” Princeton University Press, 2010.

[22] Kocherlakota, Narayana R. and Christopher Phelan, “On the Robustness of Laissez-Faire,” Journal of Economic Theory 144: 2372-2387, 2009.

[23] Kocherlakota, Narayana R. and Yangwei Song, “Public Goods with Ambiguity in Large Economies,” working paper, February 2018.

[24] Kochov, Asen, “Time and No Lotteries: An Axiomatization of Maxmin Expected Utility,” Econometrica 83 (1): 239-262, 2015.

[25] Machina, Mark J. and Marciano Siniscalchi, "Ambiguity and Ambiguity Aversion," Handbook of the Economics of Risk and Uncertainty, Vol. 1, North-Holland, 729-807, 2014.

[26] Makris, Miltiadis and Alessandro Pavan, “Taxation under Learning-by-Doing,” forthcoming in the Journal of Political Economy, 2019.

[27] Piketty, Thomas, Emmanuel Saez, and Gabriel Zucman, "Distributional National Accounts: Methods and Estimates for the United States," Quarterly Journal of Economics 133 (2): 553-609, 2018.

[28] Prescott, Edward C. and Robert M. Townsend, “Pareto Optima and Competitive Equilibria with Adverse Selection and Moral Hazard”. Econometrica 52 (1):21-45, 1984.

[29] Rigotti, Luca, Chris Shannon, and Tomasz Strzalecki, “Subjective Beliefs and Ex Ante Trade,” Econometrica 76 (5): 1167-1190, 2008.

[30] Scheuer, Florian and Iván Werning, “Mirrlees meets Diamond-Mirrlees: Simplifying Nonlinear Income Taxation,” working paper, September 2016.

[31] Stantcheva, Stefanie, “Optimal Taxation and Human Capital Policies over the Life Cycle,” Journal of Political Economy 125 (6): 1931-1990, 2017.

[32] Strzalecki, Tomasz and Jan Werner, “Efficient Allocations under Ambiguity,” Journal of Economic Theory 146 (3): 1173-1194, 2011.

[33] Williams, Noah, “Persistent Private Information,” Econometrica 79 (4): 1233-1275, 2011.

[34] Zhu, John Y., “Renegotiation of Dynamically Incomplete Contracts,” working paper, November 2016.

Appendix

A Proofs and Additional Details

A.1 Proof of Proposition 1

We will construct a $t=0$ simplified allocation $C^0 = \{c_t^0, z_t^0, k_t^0\}_{t=0}^T$ such that each agent weakly prefers C^0 to C. First define

$$\left\{c_0^0, z_0^0, k_1^0\right\} \equiv \{c_0, z_0, k_1\}.$$

Thus C^0 and C coincide at $t=0$. For $t \geq 1$, define

$$\begin{aligned} c_t^0\left(s^t\right) &\equiv c_t\left(s_0, s'\left(s_1\right), \underline{s}, \ldots, \underline{s}\right), \\ z_t^0\left(s^t\right) &\equiv z_t\left(s_0, s'\left(s_1\right), \underline{s}, \ldots, \underline{s}\right), \\ k_{t+1}^0\left(s^t\right) &\equiv k_{t+1}\left(s_0, s'\left(s_1\right), \underline{s}, \ldots, \underline{s}\right). \end{aligned}$$

At $t \geq 1$, regardless of the realized state s^t, C^0 allocates consumption, effective labor, and capital as if a shock vector of the form $s' \in S'^N$ were realized at $t=1$ and the shock vector $\underline{s}$ were realized in every subsequent period.

To see that each agent i weakly prefers C^0 to C, fix an initial state s_0 and let $\Pi'_{i,1}(s_0) \subseteq \Pi_{i,1}(s_0)$ be all beliefs of the form $\pi' = \pi \circ (s')^{-1}$. Assumption 1 implies that $\Pi'_{i,1}(s_0)$ is non-empty, and we have

$$\begin{aligned} U_{i,0}\left(C \left| s_0\right.\right) &= u\left(c_{i,0}\left(s_0\right), \frac{z_{i,0}\left(s_0\right)}{\theta_{i,0}\left(s_0\right)}\right) + \beta \inf_{\Pi_{i,1}(s_0)} \mathbb{E}_\pi\left[U_{i,1}\left(C \left| s^1\right.\right)\right] \\ &\leq u\left(c_{i,0}\left(s_0\right), \frac{z_{i,0}\left(s_0\right)}{\theta_{i,0}\left(s_0\right)}\right) + \beta \inf_{\Pi'_{i,1}(s_0)} \mathbb{E}_{\pi'}\left[U_{i,1}\left(C \left| s^1\right.\right)\right] \\ &= u\left(c_{i,0}^0\left(s_0\right), \frac{z_{i,0}^0\left(s_0\right)}{\theta_{i,0}\left(s_0\right)}\right) + \beta \inf_{\Pi'_{i,1}(s_0)} \mathbb{E}_{\pi'}\left[U_{i,1}\left(C^0 \left| s^1\right.\right)\right] \\ &\leq u\left(c_{i,0}^0\left(s_0\right), \frac{z_{i,0}^0\left(s_0\right)}{\theta_{i,0}\left(s_0\right)}\right) + \beta \inf_{\Pi_{i,1}(s_0)} \mathbb{E}_{\pi}\left[U_{i,1}\left(C^0 \left| s^1\right.\right)\right] \\ &= U_{i,0}\left(C^0 \left| s_0\right.\right). \end{aligned}$$

The key step is the fourth line, which makes substantial use of Assumption 1: By definition, the $t \geq 1$ policy functions $\left\{c_t^0, z_t^0, k_{t+1}^0\right\}_{t=1}^T$ in C^0 depend only on the $t=0$ shock vector s_0 and the shock vector $s'(s_1)$ corresponding to each $t=1$ shock vector s_1. Since s' is idempotent, any belief $\pi \in \Pi_{i,1}(s_0)$ and its corresponding belief $\pi' = \pi \circ (s')^{-1}$ have the same implied distribution of the shock vectors $s'(s_1)$, and hence the same implied distribution of future consumption and effective labor allocated to each agent. However, under the distribution π, it is possible that agent i could realize a skill $\theta_{i,t} \neq \underline{\theta}$ at some $t \geq 2$. With this realization, the agent's period-t flow utility would be strictly higher than if he realized the skill $\underline{\theta}$. Given the recursive definition of $U_{i,1}$, this implies that the expected $t=1$ continuation utility under π' is weakly lower than the expected $t=1$ continuation utility under π:

$$\mathbb{E}_{\pi'}\left[U_{i,1}\left(C^0 \left| s^1\right.\right)\right] \leq \mathbb{E}_{\pi}\left[U_{i,1}\left(C^0 \left| s^1\right.\right)\right].$$

Assumption 1 implies that for each belief $\pi \in \Pi_{i,1}(s_0)$, there exists a corresponding belief $\pi' \in \Pi'_{i,1}(s_0)$, so we can take infima on both sides to find

$$\inf_{\Pi'_{i,1}(s_0)} \mathbb{E}_{\pi'}\left[U_{i,1}\left(C^0 \left| s^1\right.\right)\right] \leq \inf_{\Pi_{i,1}(s_0)} \mathbb{E}_{\pi}\left[U_{i,1}\left(C^0 \left| s^1\right.\right)\right],$$

and the fourth line then follows. □

A.2 More General Belief Conditions Example

Suppose there are three periods $t \in \{0,1,2\}$, three potential shocks $S = \{\underline{s}, \dot{s}, \ddot{s}\}$, and three potential skills $\Theta = \{\underline{\theta}, \check{\theta}, \bar{\theta}\}$, where $\underline{\theta} < \check{\theta} < \bar{\theta}$. At $t=2$, the skill mappings $\theta_{i,2}(s)$ are defined by

$$\theta_{i,2}(s) = \begin{cases} \underline{\theta} & \text{if } s = \underline{s}, \\ \check{\theta} & \text{if } s = \dot{s}, \\ \bar{\theta} & \text{if } s = \ddot{s}. \end{cases}$$

At $t=1$, the skill mappings $\theta_{i,1}(s)$ can be chosen arbitrarily, though we require $\theta_{i,1}(\ddot{s}) = \theta_{i,1}(\dot{s})$. Define the mapping $\underline{s}_2 : S \to S$ by

$$\underline{s}_2(s) = \begin{cases} \underline{s} & \text{if } s = \underline{s}, \\ \dot{s} & \text{if } s = \dot{s}, \\ \dot{s} & \text{if } s = \ddot{s}. \end{cases}$$

At $t=1$, the belief sets $\Pi_{i,2}(s_0, (\ddot{s}, s_{-i,1}))$ can be chosen arbitrarily, though for simplicity we will assume that $\Pi_{i,2}$ is constant with respect to $s_{-i,1}$. We will assume that

$$\Pi_{i,2}(s_0, (\underline{s}, s_{-i,1})) = \Pi_{i,2}(s_0, (\dot{s}, s_{-i,1})) = \underline{\Pi}_{i,2},$$

where $\underline{\Pi}_{i,2}$ consists precisely of beliefs of the form $\underline{\pi} = \pi \circ \underline{s}_2^{-1}$ for $\pi \in \Pi_{i,2}(s_0, (\ddot{s}, s_{-i,1}))$. Here $\underline{\pi}$ assigns probabilities according to

$$\underline{\pi}(s_{A,2}, s_{B,2}) = \pi\left(\underline{s}_2^{-1}(s_{A,2}, s_{B,2})\right).$$

Thus $\underline{\Pi}_{i,2}$ is defined such that $\underline{\pi} \in \underline{\Pi}_{i,2}$ if and only if $\underline{\pi}$ is supported on $\underline{s}_2(S^2)$ and there exists a belief $\pi \in \Pi_{i,2}(s_0, (\ddot{s}, s_{-i,1}))$ such that the distribution of $\underline{s}_2(s_2)$ under π is $\underline{\pi}$. Finally, define the mapping $s_1' : S \to S$ by

$$s_1'(s) = \begin{cases} \underline{s} & \text{if } s = \underline{s}, \\ \dot{s} & \text{if } s = \dot{s}, \\ \dot{s} & \text{if } s = \ddot{s}. \end{cases}$$

Then $\Pi_{i,1}(s_0)$ can be chosen arbitrarily, subject to the following restriction: For any $\pi \in \Pi_{i,1}(s_0)$, we also have $\pi' \in \Pi_{i,1}(s_0)$, where $\pi' = \pi \circ (s_1')^{-1}$. Thus π' is the distribution supported on $s_1'(S^2)$ such that the distribution of $s_1'(s_1)$ under π is π'.

For an interpretation of this setup, we will see below that the mappings s_1' and $\underline{s}_2$ identify the subset of states that simplified allocations C^0 must take into account. In this economy, an allocation C^0 is *simplified* if it depends only on the shifted state $(s_0, s_1'(s_1), \underline{s}_2(s_2))$ rather than the full state s^2. This definition strictly generalizes the one given in Section 2, and we can recover the original definition of a simplified allocation by taking $\underline{s}_2$ to be the constant map $\underline{s}_2(s) \equiv \underline{s}$. By considering a non-constant map $\underline{s}_2$, we can avoid assuming that agents may receive a degenerate belief set $\{\underline{\pi}\}$ at $t = 1$. Indeed, at $t = 0$ agents no longer have to consider the possibility that they will be certain at $t = 1$ that a particular shock vector will be realized at $t = 2$. Rather, our assumption on beliefs $\Pi_{i,1}(s_0)$ requires that agents consider the possibility that they will be certain at $t = 1$ that the shock vector at $t = 2$ will fall in a subset $\underline{s}_2(S^2) \subseteq S^2$. The more technical aspect of our assumption states that after a "pessimistic" $t = 1$ shock vector $s_1'(s_1)$, each agent's beliefs $\Pi_{i,2}(s_0, s_1'(s_1))$ are essentially his beliefs $\Pi_{i,2}(s_0, s_1)$ after the shock s_1, but shifted so that they place full support on $\underline{s}_2(S^2)$. The significance of this condition will be clear below when we demonstrate the optimality of simplified allocations.

With this setup, the following version of Proposition 1 holds: Any allocation C is weakly Pareto-dominated by another allocation C^0 that is simplified in the sense that the allocation functions depend only on $(s_0, s_1'(s_1), \underline{s}_2(s_2))$. To see this, we begin by defining C^0 from C.

At $t = 0$, keep the allocation unchanged, and set

$$\left\{c_0^0 (s_0), z_0^0 (s_0), k_1^0 (s_0)\right\} \equiv \{c_0 (s_0), z_0 (s_0), k_1 (s_0)\}.$$

At $t = 1$, define C^0 such that if the state (s_0, s_1) is realized at $t = 1$, each agent is allocated according to what C would have prescribed if the state $(s_0, s_1' (s_1))$ were realized. For example, c_1^0 satisfies

$$c_1^0 (s_0, s_1) \equiv c_1 (s_0, s_1' (s_1)).$$

Finally, at $t = 2$, define C^0 such that if the state (s_0, s_1, s_2) is realized at $t = 2$, each agent is allocated according to what C would have prescribed if the state $(s_0, s_1' (s_1), \underline{s}_2 (s_2))$ were realized. In this case, c_2^0 satisfies

$$c_2^0 (s_0, s_1, s_2) \equiv c_2 (s_0, s_1' (s_1), \underline{s}_2 (s_2)).$$

We claim that each agent i weakly prefers C^0 to C at $t = 0$. To see this, note that since C^0 and C coincide at $t = 0$, it suffices to show

$$\inf_{\Pi_{i,1}(s_0)} \mathbb{E}_\pi \left[U_{i,1} \left(C| \, s^1\right)\right] \leq \inf_{\Pi_{i,1}(s_0)} \mathbb{E}_\pi \left[U_{i,1} \left(C^0 \middle| \, s^1\right)\right].$$

Since agent i is averse to uncertainty, we can bound the left side above by the expected continuation utility under the subset $\Pi'_{i,1} (s_0) \subseteq \Pi_{i,1} (s_0)$ of beliefs of the form π':

$$\inf_{\Pi_{i,1}(s_0)} \mathbb{E}_\pi \left[U_{i,1} \left(C| \, s^1\right)\right] \leq \inf_{\Pi'_{i,1}(s_0)} \mathbb{E}_{\pi'} \left[U_{i,1} \left(C| \, s^1\right)\right].$$

Under any belief π', agent i assumes that only shock vectors in $s_1' (S^2)$ can be realized at $t = 1$ and only shock vectors in $\underline{s}_2 (S^2)$ can be realized at $t = 2$. However, C^0 is defined so that it coincides with C along such paths. As a result, we can replace C with C^0 when evaluating expected $t = 1$ continuation utility using the belief π':

$$\inf_{\Pi'_{i,1}(s_0)} \mathbb{E}_{\pi'} \left[U_{i,1} \left(C| \, s^1\right)\right] = \inf_{\Pi'_{i,1}(s_0)} \mathbb{E}_{\pi'} \left[U_{i,1} \left(C^0 \middle| \, s^1\right)\right].$$

To complete the argument, we must show that we can replace $\Pi'_{i,1} (s_0)$ in the infimum on the right side by the superset $\Pi_{i,1} (s_0)$ without lowering the value of the infimum. We begin by noting that for any $t = 1$ shock vector s_1, agent i's $t = 1$ continuation utility in state (s_0, s_1)

is weakly greater than his continuation utility in state $(s_0, s_1'(s_1))$. This follows from

$$\begin{aligned}
U_{i,1}\left(C^0\middle| s_0, s_1\right) &= u\left(c^0_{i,1}(s_0,s_1), \frac{z^0_{i,1}(s_0,s_1)}{\theta_{i,1}(s_{i,1})}\right) + \inf_{\Pi_{i,2}(s_0,s_1)} \mathbb{E}_\pi\left[U_{i,2}\left(C^0\middle| s_0,s_1,s_2\right)\right] \\
&= u\left(c^0_{i,1}(s_0,s_1'(s_1)), \frac{z^0_{i,1}(s_0,s_1'(s_1))}{\theta_{i,1}(s_1'(s_{i,1}))}\right) + \inf_{\Pi_{i,2}(s_0,s_1)} \mathbb{E}_\pi\left[U_{i,2}\left(C^0\middle| s_0,s'(s_1),s_2\right)\right] \\
&\geq u\left(c^0_{i,1}(s_0,s_1'(s_1)), \frac{z^0_{i,1}(s_0,s_1'(s_1))}{\theta_{i,1}(s_1'(s_{i,1}))}\right) + \inf_{\Pi_{i,2}\left(s_0,s_1'(s_1)\right)} \mathbb{E}_\pi\left[U_{i,2}\left(C^0\middle| s_0,s'(s_1),s_2\right)\right] \\
&= U_{i,1}\left(C^0\middle| s_0, s_1'(s_1)\right).
\end{aligned}$$

The first equality holds because the definition of s_1' implies $\theta_{i,1}(s_{i,1}) = \theta_{i,1}(s_1'(s_{i,1}))$, and by construction the $t \geq 1$ allocation functions $c^0_{i,t}$ and $z^0_{i,t}$ are invariant under the mapping s_1'. The inequality holds because the $t = 2$ allocation functions $c^0_{i,2}$ and $z^0_{i,2}$ depend only on $(s_0, s_1'(s_1), \underline{s}_2(s_2))$, the map $\underline{s}_2$ is idempotent, and the agent's beliefs are such that $\underline{\pi} \in \Pi_{i,2}(s_0, s_1'(s_1))$ if and only if $\underline{\pi}$ is supported on $\underline{s}_2(S^2)$ and there exists a belief $\pi \in \Pi_{i,2}(s_0,s_1)$ such that the distribution of $\underline{s}_2(s_2)$ under π is $\underline{\pi}$. We have an inequality here instead of an equality because under a belief $\pi \in \Pi_{i,2}(s_0,s_1)$, agent i may receive skill $\theta_{i,2}(s_{i,2}) = \bar{\theta}$ and achieve a greater $t = 2$ continuation utility. Since this inequality holds for each shock vector s_1, we must have

$$\inf_{\Pi_{i,1}(s_0)} \mathbb{E}_\pi\left[U_{i,1}\left(C^0\middle| s^1\right)\right] \geq \inf_{\Pi_{i,1}(s_0)} \mathbb{E}_\pi\left[U_{i,1}\left(C^0\middle| s_0, s_1'(s_1)\right)\right].$$

But by assumption, for any belief $\pi \in \Pi_{i,1}(s_0)$, there exists another belief $\pi' \in \Pi'_{i,1}(s_0)$ such that

$$\begin{aligned}
\mathbb{E}_\pi\left[U_{i,1}\left(C^0\middle| s_0, s_1'(s_1)\right)\right] &= \mathbb{E}_{\pi'}\left[U_{i,1}\left(C^0\middle| s_0, s_1'(s_1)\right)\right] \\
&= \mathbb{E}_{\pi'}\left[U_{i,1}\left(C^0\middle| s_0, s_1\right)\right].
\end{aligned}$$

The second equality holds because s_1' is idempotent. Using the previous inequality and the fact that $\Pi'_{i,1}(s_0) \subseteq \Pi_{i,1}(s_0)$, we have

$$\inf_{\Pi'_{i,1}(s_0)} \mathbb{E}_{\pi'}\left[U_{i,1}\left(C^0\middle| s^1\right)\right] = \inf_{\Pi_{i,1}(s_0)} \mathbb{E}_\pi\left[U_{i,1}\left(C^0\middle| s^1\right)\right].$$

As a result, we can conclude that agent i weakly prefers C^0 to C at $t = 0$.

A.3 Details of Proposition 3 Assumptions

It will be useful to give a technical description of weak monotonicity. Given a strategy $\sigma_i \in \Sigma$, let σ_i^t denote the strategy defined by

$$\sigma_{i,\tau}^t\left(\hat{s}^{\tau-1}, s_i^\tau\right) \equiv \begin{cases} \sigma_{i,t}\left(\hat{s}^{t-1}, s_i^t\right) & \tau \leq t, \\ \sigma_{i,t+1}\left(\hat{s}^t, \left(s_i^t, s'\left(s_{i,t+1}\right)\right)\right) & \tau = t+1, \\ \underline{s} & \tau \geq t+2. \end{cases} \tag{12}$$

Thus σ_i^t coincides with σ_i through period t. At $t+1$, an agent using reporting strategy σ_i^t reports as if he realized the shock $s'\left(s_{i,t+1}\right)$ and if he were using the reporting strategy σ_i. In period $\tau \geq t+2$, the strategy σ_i^t always reports shock $\underline{s}$.

With this definition, an allocation C is weakly monotone at t if for any agent i, any $t-1$ reported state $\hat{s}^{t-1}$, any type s_i^t, and any strategy $\sigma_i \in \Sigma$,

$$U_{i,t}\left(C \left| \hat{s}^{t-1}, s_i^t\right.\right)\left(\sigma_{-i}^*, \sigma_i\right) \geq U_{i,t}\left(C \left| \hat{s}^{t-1}, s_i^t\right.\right)\left(\sigma_{-i}^{*t}, \sigma_i\right)$$

This condition implies that if the $-i$ agents report truthfully, then regardless of the reporting strategy used by agent i, his continuation utility at $t+1$ is weakly greater than his continuation utility when the $-i$ agents report shocks in S' at $t+1$ and the shock $\underline{s}$ at $\tau \geq t+2$.

A.4 Proof of Proposition 3

We will construct a $t=0$ simplified allocation $C^0 = \{c_t^0, z_t^0, k_t^0\}_{t=0}^T$ that is feasible, incentive-compatible, and is weakly preferred by each agent to the original allocation C. Let

$$\left\{c_0^0, z_0^0, k_1^0\right\} \equiv \left\{c_0, z_0, k_1\right\},$$

so that C^0 and C coincide at $t=0$. For $t \geq 1$, define

$$\begin{aligned} c_t^0\left(\hat{s}^t\right) &\equiv c_t\left(\hat{s}_0, s'\left(\hat{s}_1\right), \underline{s}, \ldots, \underline{s}\right), \\ z_t^0\left(\hat{s}^t\right) &\equiv z_t\left(\hat{s}_0, s'\left(\hat{s}_1\right), \underline{s}, \ldots, \underline{s}\right), \\ k_{t+1}^0\left(\hat{s}^t\right) &\equiv k_{t+1}\left(\hat{s}_0, s'\left(\hat{s}_1\right), \underline{s}, \ldots, \underline{s}\right). \end{aligned}$$

Thus at $t \geq 1$, C^0 allocates consumption, effective labor, and capital as if a shock vector of the form $s' \in s'\left(S^N\right)$ was reported at $t=1$ and the shock vector $\underline{s}$ was reported in every subsequent period.

Under the truth-telling strategy profile σ^*, all agents weakly prefer C^0 to C. To see this, fix any initial state s_0 and any i. Let $\Pi'_{i,1}(s_{i,0}) \subseteq \Pi_{i,1}(s_{i,0})$ be the set of all beliefs of the form $\pi' = \pi \circ (s')^{-1}_{i,1}$ for $\pi \in \Pi_{i,1}(s_{i,0})$. Assumption 2 implies that $\Pi'_{i,1}(s_{i,0})$ is non-empty, and we have

$$\begin{aligned} U_{i,0}(C\,|s_{i,0})(\sigma^*) &= \inf_{\Pi_{i,1}(s_{i,0})} \mathbb{E}_{\pi}\left[W_{i,0}(C\,|s_0, s_{i,1})(\sigma^*)\right] \\ &\leq \inf_{\Pi'_{i,1}(s_{i,0})} \mathbb{E}_{\pi'}\left[W_{i,0}(C\,|s_0, s_{i,1})(\sigma^*)\right] \\ &= \inf_{\Pi'_{i,1}(s_{i,0})} \mathbb{E}_{\pi'}\left[W_{i,0}\left(C^0\,|s_0, s_{i,1}\right)(\sigma^*)\right] \\ &\leq \inf_{\Pi_{i,1}(s_{i,0})} \mathbb{E}_{\pi}\left[W_{i,0}\left(C^0\,|s_0, s_{i,1}\right)(\sigma^*)\right] \\ &= U_{i,0}\left(C^0\,|s_{i,0}\right)(\sigma^*). \end{aligned} \tag{13}$$

The key step is again the fourth line, and it holds by essentially the same argument as in the public information case: Let $\pi' \in \Pi'_{i,1}(s_{i,0})$, and let $\pi \in \Pi_{i,1}(s_{i,0})$ be the corresponding belief such that $\pi' = \pi \circ (s')^{-1}_{i,1}$. It suffices to prove

$$\mathbb{E}_{\pi}\left[W_{i,0}\left(C^0\,|s_0\right)(\sigma^*)\right] \geq \mathbb{E}_{\pi'}\left[W_{i,0}\left(C^0\,|s_0\right)(\sigma^*)\right].$$

But since π and π' have the same marginal distribution over $s_{-i,0}$, this holds if and only if

$$\begin{aligned} \mathbb{E}_{\pi}\left[U_{i,1}\left(C^0\left|s_0, s_i^1\right.\right)(\sigma^*)\right] &\geq \mathbb{E}_{\pi'}\left[U_{i,1}\left(C^0\left|s_0, s_i^1\right.\right)(\sigma^*)\right] \\ &= \mathbb{E}_{\pi}\left[U_{i,1}\left(C^0\,|s_0, (s_{i,0}, s'(s_{i,1}))\right)(\sigma^*)\right]. \end{aligned}$$

The second line holds because π' is the pushforward belief of π under the mapping $(s_{-i,0}, s_{i,1}) \mapsto (s_{-i,0}, s'(s_{i,1}))$. It then suffices to prove

$$U_{i,1}\left(C^0\left|s_0, s_i^1\right.\right)(\sigma^*) \geq U_{i,1}\left(C^0\,|s_0, (s_{i,0}, s'(s_{i,1}))\right)(\sigma^*) \quad \forall s_{i,1} \in S.$$

Using the recursive definition of $U_{i,0}$, this inequality can be stated

$$\begin{aligned} &\inf_{\Pi_{i,2}\left(s_0, s_i^1\right)} \mathbb{E}_{\pi}\left[W_{i,1}\left(C^0\left|s_0, s^1, s_{i,2}\right.\right)(\sigma^*)\right] \\ &\geq \inf_{\Pi_{i,2}(s_0, (s_{i,0}, s'(s_{i,1})))} \mathbb{E}_{\pi}\left[W_{i,1}\left(C^0\,|s_0, (s_{-i,1}, (s_{i,0}, s'(s_{i,1})), s_{i,2})\right)(\sigma^*)\right]. \end{aligned} \tag{14}$$

By definition, the $t \geq 1$ policy functions $\left\{c_t^0, z_t^0, k_{t+1}^0\right\}_{t=1}^T$ in C^0 depend only on the $t = 0$ reported shock vector $\hat{s}_0$ and the shock vector $s'(\hat{s}_1)$ corresponding to each $t = 1$ reported shock vector $\hat{s}_1$. Moreover, Assumption 2 implies that $\pi_2 \in \Pi_{i,2}(s_0, (s_{i,0}, s'(s_{i,1})))$ if and only if $\pi_2(s_{i,2} = \underline{s}) = 1$ and there exists $\tilde{\pi}_2 \in \Pi_{i,2}(s_0, s_i^1)$ such that $\pi_2|_{s_{-i,1}} = \tilde{\pi}_2|_{s_{-i,1}} \circ (s')^{-1}$. Since s' is idempotent, π_2 and $\tilde{\pi}_2$ have the same implied distribution of the $-i$ agents' shifted shock vector $s'(s_{-i,1})$, and hence the same distribution of $t \geq 1$ consumption and effective labor allocated to agent i. However, under the distribution $\tilde{\pi}_2$, it is possible that agent i could realize a skill $\theta_{i,t} \neq \underline{\theta}$ at some $t \geq 2$. With this realization, the agent's period-t flow utility would be strictly higher than if he realized the skill $\underline{\theta}$. Given the recursive definition of $W_{i,1}$, this implies that (14) holds, and thus that (13) holds.

It is clear that C^0 satisfies non-negativity and feasibility, and we now demonstrate that C^0 is incentive-compatible. To see this, note that (13) and the incentive-compatibility of C respectively imply

$$\begin{aligned} U_{i,0}\left(C^0 \left| s_{i,0}\right.\right)(\sigma^*) &\geq U_{i,0}\left(C \left| s_{i,0}\right.\right)(\sigma^*) \\ &\geq \max_{\sigma_i \in \Sigma} U_{i,0}\left(C \left| s_{i,0}\right.\right)\left(\sigma^*_{-i}, \sigma_i\right). \end{aligned}$$

Since C is assumed weakly monotone at $t = 0$, we can lower bound the second term on the right by replacing σ^*_{-i} with σ^{*0}_{-i}:

$$\max_{\sigma_i \in \Sigma} U_{i,0}\left(C \left| s_{i,0}\right.\right)\left(\sigma^*_{-i}, \sigma_i\right) \geq \max_{\sigma_i \in \Sigma} U_{i,0}\left(C \left| s_{i,0}\right.\right)\left(\sigma^{*0}_{-i}, \sigma_i\right).$$

We also have that $\sigma_i^0 \in \Sigma$ for any $\sigma_i \in \Sigma$, so we can lower bound the right side by replacing σ_i with σ_i^0:

$$\max_{\sigma_i \in \Sigma} U_{i,0}\left(C \left| s_{i,0}\right.\right)\left(\sigma^{*0}_{-i}, \sigma_i\right) \geq \max_{\sigma_i \in \Sigma} U_{i,0}\left(C \left| s_{i,0}\right.\right)\left(\sigma^{*0}_{-i}, \sigma_i^0\right).$$

But C and C^0 coincide for $t \geq 0$ reported states of the form $(\hat{s}_0, s'(\hat{s}_1), \underline{s}, \ldots, \underline{s})$, so we can replace C with C^0 to find the equality

$$\max_{\sigma_i \in \Sigma} U_{i,0}\left(C \left| s_{i,0}\right.\right)\left(\sigma^{*0}_{-i}, \sigma_i^0\right) = \max_{\sigma_i \in \Sigma} U_{i,0}\left(C^0 \left| s_{i,0}\right.\right)\left(\sigma^*_{-i}, \sigma_i\right).$$

This sequence of relations implies that C^0 is incentive-compatible. □

A.5 Optimal Reform Details

Proposition 3 implies that the government can solve its $t = 0$ problem (7) using a simplified allocation C^0. As discussed in Section 3.3, since C^0 is not fully contingent on agents' reports at $t = 1$, the government may seek to implement a reform C^1. In particular, given the simplified allocation C^0 designed at $t = 0$, the truthfully-reported $t = 0$ state s_0, and the

$t = 1$ type s_g^1 of the governing agent, the government seeks to solve the problem

$$\max_{C} \inf_{\Pi_{g,2}\left(\hat{s}_0, s_g^1\right)} \mathbb{E}_\pi \left[\sum_i \eta_i U_{i,1}\left(C \left| s_0, s_i^1 \right.\right)\left(\sigma^*\right) \right] \tag{15}$$

$$\begin{aligned}
&\text{subject to non-negativity and} \\
&\sum_i \left[c_{i,t}\left(\hat{s}^t\right) + k_{i,t+1}\left(\hat{s}^t\right) \right] \leq \sum_i f\left(k_{i,t}\left(\hat{s}^{t-1}\right), z_{i,t}\left(\hat{s}^t\right)\right) \quad \forall t \geq 1, \hat{s}^t \geq \hat{s}_0, \\
&U_{i,1}\left(C \left| \hat{s}_0, s_i^1 \right.\right)\left(\sigma^*\right) \geq U_{i,1}\left(C \left| \hat{s}_0, s_i^1 \right.\right)\left(\sigma_{-i}^*, \sigma_i\right) \quad \forall i, s_i^1, \sigma_i \in \Sigma, \\
&U_{i,0}\left(C_0^0, \left(C_\tau\right)_{\tau=1}^T \left| s_{i,0} \right.\right)\left(\sigma^*\right) \geq U_{i,0}\left(C^0 \left| s_{i,0} \right.\right)\left(\sigma^*\right) \quad \forall i, s_{i,0} \\
&U_{i,0}\left(C_0^0, \left(C_\tau\right)_{\tau=1}^T \left| s_{i,0} \right.\right)\left(\sigma_{-i,0}^*, \sigma_{i,0}, \left(\sigma_\tau^*\right)_{\tau=1}^T\right) \\
&\qquad \leq U_{i,0}\left(C^0 \left| s_{i,0} \right.\right)\left(\sigma_{-i,0}^*, \sigma_{i,0}, \left(\sigma_\tau^*\right)_{\tau=1}^T\right) \quad \forall i, s_{i,0}, \sigma_{i,0}.
\end{aligned}$$

If the constraint set is non-empty, then the government chooses the allocation that solves problem (15). In this case, we can use arguments similar to those above to show that a simplified allocation is in the solution set.

In particular, let $C^{*1} \equiv C^*\left(\hat{s}_0, s_g^1, C^0\right)$ solve problem (15), and suppose that it is weakly monotone at $t = 1$. Define the simplified allocation $C^1 = \{c_t^1, z_t^1, k_t^1\}_{t=1}^T$ analogously to how C^0 was defined from C^*: Let

$$\left\{c_1^1, z_1^1, k_2^1\right\} \equiv \left\{c_1^{*1}, z_1^{*1}, k_2^{*1}\right\},$$

so that C^1 and C^{*1} coincide at $t = 1$. For $t \geq 2$, define

$$\begin{aligned}
c_t^1\left(\hat{s}^t\right) &\equiv c_t^{*1}\left(\hat{s}^1, s'\left(\hat{s}_2\right), \underline{s}, \ldots, \underline{s}\right), \\
z_t^1\left(\hat{s}^t\right) &\equiv z_t^{*1}\left(\hat{s}^1, s'\left(\hat{s}_2\right), \underline{s}, \ldots, \underline{s}\right), \\
k_{t+1}^1\left(\hat{s}^t\right) &\equiv k_{t+1}^{*1}\left(\hat{s}^1, s'\left(\hat{s}_2\right), \underline{s}, \ldots, \underline{s}\right).
\end{aligned}$$

The simplified allocation C^1 clearly satisfies non-negativity and feasibility, and since we assume that C^{*1} is weakly monotone at $t = 1$, the same argument as for C^0 implies that C^1 is incentive-compatible at $t = 1$. To see that C^1 satisfies the promise-keeping constraint (the third constraint in problem 15), note first that the $t = 1$ analogue of (13) implies that for all i, all $\hat{s}_0$, and all s_i^1,

$$U_{i,1}\left(C^{*1} \left| \hat{s}_0, s_i^1 \right.\right)\left(\sigma^*\right) \leq U_{i,1}\left(C^1 \left| \hat{s}_0, s_i^1 \right.\right)\left(\sigma^*\right). \tag{16}$$

Using this inequality and the definition of the utility function $U_{i,0}$, we then find that C^1

satisfies the promise-keeping constraint:

$$\begin{aligned}
U_{i,0}\big(C_0^0, \big(C_\tau^1\big)_{\tau=1}^T \,|s_{i,0}\big)\big(\sigma^*\big) &= \inf_{\Pi_{i,1}(s_{i,0})} \mathbb{E}_\pi \left[u\left(c_0^0\left(s_0\right), \frac{z_0^0\left(s_0\right)}{\theta_{i,0}\left(s_0\right)}\right) + \beta U_{i,1}\big(C^1 \left|s_0, s_i^1\right.\big)\big(\sigma^*\big)\right]. \\
&\geq \inf_{\Pi_{i,1}(s_{i,0})} \mathbb{E}_\pi \left[u\left(c_0^0\left(s_0\right), \frac{z_0^0\left(s_0\right)}{\theta_{i,0}\left(s_0\right)}\right) + \beta U_{i,1}\big(C^{*1} \left|s_0, s_i^1\right.\big)\big(\sigma^*\big)\right] \\
&= U_{i,0}\big(C_0^0, \big(C_\tau^{*1}\big)_{\tau=1}^T \,|s_{i,0}\big)\big(\sigma^*\big) \\
&\geq U_{i,0}\big(C^0 \,|s_{i,0}\big)\big(\sigma^*\big).
\end{aligned}$$

The final inequality holds because C^{*1} satisfies the promise-keeping constraint by assumption.

Finally, we show that the simplified allocation C^1 also satisfies the threat-keeping constraint (the fourth constraint in problem 15). Given the allocation $\left(C_0^0, (C_t^1)_{t=1}^T\right)$, the strategy profile $\left(\sigma_{-i,0}^*, \sigma_{i,0}, (\sigma_\tau^*)_{\tau=1}^T\right)$, and the $t = 0$ type $s_{i,0}$, consider the $t = 0$ utility of agent i:

$$\begin{aligned}
&U_{i,0}\big(C_0^0, \big(C_\tau^1\big)_{\tau=1}^T \,|s_{i,0}\big)\big(\sigma_{-i,0}^*, \sigma_{i,0}, \big(\sigma_\tau^*\big)_{\tau=1}^T\big) \\
&= \inf_{\Pi_{i,1}(s_{i,0})} \mathbb{E}_\pi \left[u\left(c_0^0\left(s_{-i,0}, \sigma_{i,0}\left(s_{i,0}\right)\right), \frac{z_0^0\left(s_{-i,0}, \sigma_{i,0}\left(s_{i,0}\right)\right)}{\theta_{i,0}\left(s_0\right)}\right)\right. \\
&\qquad\qquad \left. + \beta U_{i,1}\big(C^1 \left|\big(s_{-i,0}, \sigma_{i,0}\big(s_{i,0}\big)\big), s_i^1\right.\big)\big(\sigma^*\big)\right].
\end{aligned} \tag{17}$$

The argument for the incentive-compatibility of C^{*1} implies that inequality (16) holds with equality, so we can replace C^1 with C^{*1} in the argument of $U_{i,1}$ to find that (17) is equal to

$$\begin{aligned}
&\inf_{\Pi_{i,1}(s_{i,0})} \mathbb{E}_\pi \left[u\left(c_0^0\left(s_{-i,0}, \sigma_{i,0}\left(s_{i,0}\right)\right), \frac{z_0^0\left(s_{-i,0}, \sigma_{i,0}\left(s_{i,0}\right)\right)}{\theta_{i,0}\left(s_0\right)}\right)\right. \\
&\qquad\qquad \left. + \beta U_{i,1}\big(C^{*1} \left|\big(s_{-i,0}, \sigma_{i,0}\big(s_{i,0}\big)\big), s_i^1\right.\big)\big(\sigma^*\big)\right]. \\
&= U_{i,0}\big(C_0^0, \big(C_\tau^{*1}\big)_{\tau=1}^T \,|s_{i,0}\big)\big(\sigma_{-i,0}^*, \sigma_{i,0}, \big(\sigma_\tau^*\big)_{\tau=1}^T\big).
\end{aligned} \tag{18}$$

But C^{*1} satisfies the threat-keeping constraint by assumption, so

$$\begin{aligned}
&U_{i,0}\big(C_0^0, \big(C_\tau^{*1}\big)_{\tau=1}^T \,|s_{i,0}\big)\big(\sigma_{-i,0}^*, \sigma_{i,0}, \big(\sigma_\tau^*\big)_{\tau=1}^T\big) \\
&\leq U_{i,0}\big(C^0 \,|s_{i,0}\big)\big(\sigma_{-i,0}^*, \sigma_{i,0}, \big(\sigma_\tau^*\big)_{\tau=1}^T\big).
\end{aligned} \tag{19}$$

Expressions (17)-(19) then imply that C^1 also satisfies the threat-keeping constraint. We have shown that the simplified allocation C^1 is in the constraint set of problem (15), so inequality (16) implies that it is also a solution to problem (15).

A.6 Proof of Proposition 6

Given allocation $\check{C}_i$, agent i's $t=0$ continuation utility is given by

$$
\begin{aligned}
& U_{i,0}\left(\check{C}_i \,|s_0\right) \\
& = u\left(\check{c}_{i,0}\left(s_0\right), \frac{\check{z}_{i,0}\left(s_0\right)}{\theta_{i,0}}\right) \\
& \quad + \inf_{\Pi_{i,1}(s_0)} \mathbb{E}_\pi\left[\beta u\left(\check{c}_{i,1}\left(s_0,\theta_1\right), \frac{\check{z}_{i,1}\left(s_0,\theta_1\right)}{\theta_{i,1}}\right) + \sum_{t=2}^{T} \beta^t u\left(\check{c}_{i,t}\left(s_0,\theta_1\right), \frac{\check{z}_{i,t}\left(s_0,\theta_1\right)}{\underline{\theta}}\right)\right].
\end{aligned}
$$

For notational simplicity, define

$$
\check{v}\left(s_0,\theta_1\right) \equiv \beta u\left(\check{c}_{i,1}\left(s_0,\theta_1\right), \frac{\check{z}_{i,1}\left(s_0,\theta_1\right)}{\theta_{i,1}}\right) + \sum_{t=2}^{T} \beta^t u\left(\check{c}_{i,t}\left(s_0,\theta_1\right), \frac{\check{z}_{i,t}\left(s_0,\theta_1\right)}{\underline{\theta}}\right).
$$

Define $\hat{v}\left(s_0,\theta_1\right)$ similarly. By Assumption 3, we know that for each $\pi \in \Pi_{i,1}\left(s_0\right)$, there exists $\tilde{\pi} \in \Pi_{i,1}\left(s_0\right)$ such that π and $\tilde{\pi}$ have the same marginal distribution over $\theta_{-i,1}$ conditional on s_0, but the distribution of $\theta_{i,1}$ under $\tilde{\pi}$ places weight only on $\underline{\theta}$ and $\bar{\theta}$ so as to satisfy $\mathbb{E}_{\tilde{\pi}}\left[\theta_{i,1}\right] = \mathbb{E}_\pi\left[\theta_{i,1}\right]$. Let $\tilde{\Pi}_{i,1}\left(s_0\right) \subset \Pi_{i,1}\left(s_0\right)$ denote the subset of all distributions of the form $\tilde{\pi}$. By set inclusion, the following inequalities are immediate:

$$
\inf_{\Pi_{i,1}(s_0)} \mathbb{E}_\pi\left[\check{v}\left(s_0,\theta_1\right)\right] \leq \inf_{\tilde{\Pi}_{i,1}(s_0)} \mathbb{E}_{\tilde{\pi}}\left[\check{v}\left(s_0,\theta_1\right)\right], \tag{20}
$$

$$
\inf_{\Pi_{i,1}(s_0)} \mathbb{E}_\pi\left[\hat{v}\left(s_0,\theta_1\right)\right] \leq \inf_{\tilde{\Pi}_{i,1}(s_0)} \mathbb{E}_{\tilde{\pi}}\left[\hat{v}\left(s_0,\theta_1\right)\right]. \tag{21}
$$

To see that the opposite inequalities also hold, let $p\left(s_0\right) \in [0,1]$ be such that

$$
p\left(s_0\right)\underline{\theta} + \left(1-p\left(s_0\right)\right)\bar{\theta} = \mathbb{E}_\pi\left[\theta_{i,1}\right].
$$

By the definition of $\check{C}_i$, $\check{v}\left(s_0,\theta_{-i,1},\cdot\right)$ lies weakly above its secant line between $\underline{\theta}$ and $\bar{\theta}$, so by the independence of $\theta_{-i,1}$ and $\theta_{i,1}$ under π and $\tilde{\pi}$, we have

$$
\begin{aligned}
\mathbb{E}_\pi\left[\check{v}\left(s_0,\theta_1\right)\right] &= \mathbb{E}_\pi\left[\mathbb{E}_\pi\left[\check{v}\left(s_0,\theta_1\right)\middle|\,\theta_{-i,1}\right]\right] \\
&\geq \mathbb{E}_\pi\left[p\left(s_0\right)\check{v}\left(s_0,\theta_{-i,1},\underline{\theta}\right) + \left(1-p\left(s_0\right)\right)\check{v}\left(s_0,\theta_{-i,1},\bar{\theta}\right)\right] \\
&= \mathbb{E}_{\tilde{\pi}}\left[p\left(s_0\right)\check{v}\left(s_0,\theta_{-i,1},\underline{\theta}\right) + \left(1-p\left(s_0\right)\right)\check{v}\left(s_0,\theta_{-i,1},\bar{\theta}\right)\right] \\
&= \mathbb{E}_{\tilde{\pi}}\left[\check{v}\left(s_0,\theta_1\right)\right].
\end{aligned}
$$

A similar set of calculations applies to $\hat{v}(s_0, \theta_1)$. By taking infima on both sides, we find that equality must hold in (20) and (21).

Now since $\check{v}(s_0, \theta_{-i,1}, \cdot)$ and $\hat{v}(s_0, \theta_{-i,1}, \cdot)$ coincide on the endpoints of Θ, the independence of $\theta_{i,1}$ and $\theta_{-i,1}$ imply

$$\begin{aligned}
\mathbb{E}_{\tilde{\pi}}\left[\check{v}(s_0, \theta_1)\right] &= \mathbb{E}_{\tilde{\pi}}\left[p(s_0)\,\check{v}(s_0, \theta_{-i,1}, \underline{\theta}) + (1 - p(s_0))\,\check{v}\left(s_0, \theta_{-i,1}, \bar{\theta}\right)\right] \\
&= \mathbb{E}_{\tilde{\pi}}\left[p(s_0)\,\hat{v}(s_0, \theta_{-i,1}, \underline{\theta}) + (1 - p(s_0))\,\hat{v}\left(s_0, \theta_{-i,1}, \bar{\theta}\right)\right] \\
&= \mathbb{E}_{\tilde{\pi}}\left[\hat{v}(s_0, \theta_1)\right].
\end{aligned}$$

Taking infima over $\tilde{\Pi}_{i,1}(s_0)$ on both sides, we find

$$\inf_{\tilde{\Pi}_{i,1}(s_0)} \mathbb{E}_{\tilde{\pi}}\left[\check{v}(s_0, \theta_1)\right] = \inf_{\tilde{\Pi}_{i,1}(s_0)} \mathbb{E}_{\tilde{\pi}}\left[\hat{v}(s_0, \theta_1)\right]. \tag{22}$$

Since equality holds in (20) and (21), (22) then implies

$$\inf_{\Pi_{i,1}(s_0)} \mathbb{E}_{\pi}\left[\check{v}(s_0, \theta_1)\right] = \inf_{\Pi_{i,1}(s_0)} \mathbb{E}_{\pi}\left[\hat{v}(s_0, \theta_1)\right].$$

Thus agent i is indifferent between $\check{C}_i$ and $\hat{C}_i$. □

B Infinite Time Horizon, Continuum of Agents

This section describes an extension of the model in Section 2 to the case in which the time horizon is infinite ($T = \infty$) and there is a continuum of agents. To avoid difficulties in using backward induction to define certain belief sets, the model formulation will be sequential. For simplicity, we also assume that each shock $s_{i,t}$ can be identified with the corresponding skill $\theta_{i,t}$.

Identify each agent with a real number in the interval $I_1 = [0, 1]$, and equip I_1 with the Borel sigma algebra $\mathcal{B}$ and Lebesgue measure λ. Let $\Theta \subset \mathbb{R}_{++}$ be the set of skill shocks that are possible in each period, which we will suppose is compact and connected for simplicity. As before, let $\underline{\theta}$ and $\bar{\theta}$ denote the minimum and maximum elements of Θ, respectively. Give $\Theta^{\mathbb{N}_0}$ the sigma algebra generated by finite rectangles, i.e., sets of the form

$$(a_1, b_1] \times (a_2, b_2] \times \ldots \times (a_n, b_n] \times \Theta \times \Theta \times \ldots$$

with $\underline{\theta} \leq a_k \leq b_k \leq \bar{\theta}$, $k = 1, \ldots, n$, and $n \in \mathbb{N}$.

Let $(\Omega, \mathcal{F}, \mu)$ be a measure space, where $\omega \in \Omega$ denotes a possible state of the economy and μ is the distribution of states. Give $I_1 \times \Omega$ the product sigma algebra $\mathcal{B} \otimes \mathcal{F}$ with the product measure $\varphi \equiv \lambda \otimes \mu$. Let $\theta^\infty : I_1 \times \Omega \to \Theta^{\mathbb{N}_0}$ be $\mathcal{B} \otimes \mathcal{F}$-measurable. For any $i \in I_1$, $\theta^\infty(i, \omega)$ denotes agent i's infinite sequence of skill shocks if the state is ω. For notational simplicity, we will write $\theta_i^\infty \equiv \theta^\infty(i, \cdot)$. Define θ^t as the projection of θ^∞ to Θ^{t+1}, and let $\mathcal{F}_t \equiv \sigma(\theta^t) \subset \mathcal{B} \otimes \mathcal{F}$ be the sigma algebra generated by θ^t. Then $(\mathcal{F}_t)_{t=0}^\infty$ is a filtration on $\mathcal{B} \otimes \mathcal{F}$ to which the shock history process $(\theta^t)_{t=0}^\infty$ is adapted.

Let $C \equiv \{c_{i,t}(\theta^t), z_{i,t}(\theta^t), k_{i,t+1}(\theta^t)\}_{t=0, i \in I_1}^\infty$ denote an allocation, and let $f : \mathbb{R}_+^2 \to \mathbb{R}_+$ be a constant returns-to-scale production function that is increasing in capital and effective labor. We say that C is *feasible* if for all t and all θ^t,

$$\int_{I_1} c_{i,t}\left(\theta^t\right) + k_{i,t+1}\left(\theta^t\right) d\lambda \leq \int_{I_1} f\left(k_{i,t}\left(\theta^{t-1}\right), z_{i,t}\left(\theta^t\right)\right) d\lambda.$$

For each agent $i \in I_1$, Let $P_i \subset \Delta(I_1 \times \Omega, \mathcal{B} \otimes \mathcal{F})$ be a non-empty set of prior distributions on $(I_1 \times \Omega, \mathcal{B} \otimes \mathcal{F})$ that represent agent i's beliefs. We assume that each distribution $p_i \in P_i$ is a product measure of the form $\lambda \otimes \mu_i$ for some $\mu_i \in \Delta(\Omega, \mathcal{F})$. Unlike in the main text, for simplicity we will assume here that that agents update their beliefs using Bayes's rule prior-by-prior. Given an allocation C, and a state θ^t, agent i's t continuation utility is given by

$$U_{i,t}\left(C \left| \theta^t\right.\right) := \inf_{p_i \in P_i} \mathbb{E}_{p_i}\left[\sum_{\tau=t}^{\infty} \beta^{\tau-t} u\left(c_{i,\tau}\left(\theta^\tau\right), \frac{z_{i,\tau}\left(\theta^\tau\right)}{\theta_{i,\tau}}\right) \middle| \theta^t\right],$$

where $\beta \in (0,1)$. With social welfare weights $\eta \in \Delta(I_1, \mathcal{B})$ and an initial state θ_0, an *efficient* allocation $C^*(\theta_0)$ is given by

$$C^*(\theta_0) \in \arg\max_C \int_{I_1} U_{i,0}(C | \theta_0)\, d\eta,$$

subject to feasibility and non-negativity.

The analog of Assumption 1 in this setup is

Assumption 4. *For any $t \geq 0$, θ^t, i, and $p_i \in P_i$, there exists $p_i' \in P_i$ such that*

$$p_i'\left(\cdot \middle| \theta^t\right)\Big|_{\mathcal{F}_{t+1}} = p_i\left(\cdot \middle| \theta^t\right)\Big|_{\mathcal{F}_{t+1}},$$

but

$$p_i'\left(\left(\theta^\infty\right)^{-1}\left(\left\{\ell \in \Theta^{\mathbb{N}_0} : \ell_\tau = \underline{\theta},\ \tau \geq t+2\right\}\right)\right) = 1.$$

This assumption implies that in any period t and for any belief $p_i \in P_i$, there exists another belief $p_i' \in P_i$ such that p_i and p_i' imply the same distribution of the $t+1$ state θ^{+1}

conditional on θ^t. However, under p'_i, almost surely almost all agents will realize the shock $\underline{\theta}$ at $\tau \geq t+2$.

To prove an analogue of Proposition 1, we must make an additional assumption. Since the policy functions in an allocation depend fully on the state θ^t, it is possible that changing the shock histories of agents in a set of λ-measure zero could alter an agent i's allocation. This property is inconsistent with Assumption 4, which ensures only that a full measure of agents realize the shock $\underline{\theta}$ at $\tau \geq t+2$ under the distribution p'_i. We will thus assume that the policy functions do not distinguish between any two t states θ^t and $\tilde{\theta}^t$ such that $\theta^t_i \neq \tilde{\theta}^t_i$ for i in a set of λ-measure zero.[23]

Given the assumptions above, periodically-reformed policies are optimal:

Proposition 8. *In any period t, any feasible allocation C is weakly Pareto dominated by a simplified allocation C^t, i.e.,*

$$U_{i,t}\left(C^t \left|\theta^t\right.\right) \geq U_{i,t}\left(C \left|\theta^t\right.\right) \quad \forall i.$$

Proof. Define C^0 to coincide with C at $t=0$ and at $t=1$. Then let

$$c^0_t\left(\theta^t\right) \equiv c_t\left(\theta^1, \left(\underline{\theta}_\tau\right)^{t-2}_{\tau=2}\right).$$

Here $\underline{\theta}_\tau : I_1 \times \Omega \to \Theta$ satisfies $\underline{\theta}_\tau(i,\omega) = \underline{\theta}$ for all $(i,\omega) \in I_1 \times \Omega$. Define z_t and k_{t+1} similarly. Let $P'_i \subseteq P_i$ denote all distributions of the form $p_i \in P_i$. Then for all i,

$$\begin{aligned}
U_{i,0}\left(C \left|\theta^0\right.\right) &= \inf_{P_i} \mathbb{E}_{p_i}\left[\sum_{t=0}^{\infty} \beta^t u\left(c_{i,t}\left(\theta^t\right), \frac{z_{i,t}\left(\theta^t\right)}{\theta_{i,t}}\right)\middle| \theta^0\right] \\
&\leq \inf_{P'_i} \mathbb{E}_{p'_i}\left[\sum_{t=0}^{\infty} \beta^t u\left(c_{i,t}\left(\theta^t\right), \frac{z_{i,t}\left(\theta^t\right)}{\theta_{i,t}}\right)\middle| \theta^0\right] \\
&= \inf_{P'_i} \mathbb{E}_{p'_i}\left[\sum_{t=0}^{\infty} \beta^t u\left(c^0_{i,t}\left(\theta^t\right), \frac{z^0_{i,t}\left(\theta^t\right)}{\theta_{i,t}}\right)\middle| \theta^0\right] \\
&\leq \inf_{P_i} \mathbb{E}_{p_i}\left[\sum_{t=0}^{\infty} \beta^t u\left(c^0_{i,t}\left(\theta^t\right), \frac{z^0_{i,t}\left(\theta^t\right)}{\theta_{i,t}}\right)\middle| \theta^0\right] \\
&= U_{i,0}\left(C^0 \left|\theta^0\right.\right).
\end{aligned}$$

[23]This assumption is without loss of generality when the social welfare measure η is absolutely continuous with respect to Lebesgue measure. Alternatively, a natural way to enforce this assumption is to constrain policy functions to depend only on an agent's shock history θ^t_i as well as the distribution of shock histories in the economy. To define the information structure appropriately, random measures should be used to model the agent's beliefs about the distribution of shock histories that will be observed in subsequent periods.

The third line holds because C^0 and C^* coincide when almost every agent realizes the shock $\underline{\theta}$ at $t \geq 2$. The last line holds because c_t^0 and z_t^0 only depend on θ^1 for $t \geq 2$, and agent i can potentially realize a shock $\theta_{i,t} \neq \underline{\theta}$ at some $t \geq 2$ under a distribution $p_i \in P_i \setminus P_i'$. Hence all agents weakly prefer C^0 to C^*, and the feasibility of C^0 follows from that of C^*. By iterating this process in each period, we have the result. □

With Proposition 8, we find that even when the time horizon is infinite and there is a continuum of measureless agents, the efficient allocation can be implemented by a sequence of simplified allocations $\{C^t\}_{t=0}^{\infty}$ that are reformed after each period. As in Section 2.2, each allocation C^t displays limited state dependence at $\tau \geq t+2$, and an argument analogous to that in Section 2.3 implies that they are history independent whenever the government's promise-keeping constraints are slack.