NBER WORKING PAPER SERIES

COALITION-PROOF RISK SHARING UNDER FRICTIONS

Harold L. Cole
Dirk Krueger
George J. Mailath
Yena Park

Working Paper 26667
http://www.nber.org/papers/w26667

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2020

Cole, Krueger, and Mailath thank the National Science Foundation for research support (grants SES-112354, 1260753, 1326781, 1559369, 1757084, 1851449). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Coalition-Proof Risk Sharing Under Frictions
Harold L. Cole, Dirk Krueger, George J. Mailath, and Yena Park
NBER Working Paper No. 26667
January 2020
JEL No. D15,D16,E20

ABSTRACT

We analyze efficient risk-sharing arrangements when coalitions may deviate. Coalitions form to insure against idiosyncratic income risk. Self-enforcing contracts for both the original coalition and any deviating coalition rely on a belief in future cooperation, and we treat the contracting conditions of original and deviating coalitions symmetrically. We show that better belief coordination (higher social capital) tightens incentive constraints since it facilitates both the formation of the original as well as a deviating coalition. As a consequence, the payoff of successfully formed coalitions might be declining in the degree of belief coordination and equilibrium allocations might feature resource burning or utility burning.

Harold L. Cole
Economics Department
University of Pennsylvania
3718 Locust Walk
160 McNeil Building
Philadelphia, PA 19104
and NBER
colehl@sas.upenn.edu

George J. Mailath
Department of Economics
University of Pennsylvania
133 South 36th Street
Philadelphia, PA 19104-6297
and Research School of Economics,
Australian National University
gmailath@econ.upenn.edu

Dirk Krueger
Economics Department
University of Pennsylvania
The Ronald O. Perelman Center
for Political Science and Economics
133 South 36th Street
Philadelphia, PA 19104
and NBER
dkrueger@econ.upenn.edu

Yena Park
Department of Economics
University of Rochester
Rochester, NY 14627
parkye82@gmail.com

Contents

1 Introduction

All institutions must confront the challenge of deterring opportunistic behavior, and successful institutions survive because they are essentially self-enforcing agreements. By their nature, self-enforcing agreements are built on a *shared* belief in future cooperation. In the absence of such shared beliefs, people will behave opportunistically with accompanying efficiency losses. And a shared belief in future cooperation can be tenuous and difficult to achieve: While we do see many instances of successful Pareto-improving institutions, there are also many missed opportunities—situations where opportunistic behavior is not deterred because the required belief in future cooperation is absent.

We investigate the role of fragile belief coordination on future cooperation for the endogenous formation of self-enforcing risk-sharing institutions. It is common to model limited commitment in insurance settings as an outside option whose value is autarky (i.e., exclusion from the risk-sharing arrangement). But, in the absence of commitment, a major threat to any risk-sharing arrangement is that a wealthy subset of the original participants may defect from the arrangement and insure within the deviating coalition. Since insurance by the deviating coalition typically is more attractive than autarky (i.e., the value of the outside option is higher under insurance), such *coalitional* deviations will typically not be deterred in a risk-sharing arrangement designed considering only the autarkic outside option.

Just as the successful formation of a risk-sharing institution depends upon a shared belief in future cooperation, so does the successful defection by a wealthy coalition intending to internally insure. The possible inability of groups to share beliefs in future cooperation is a key friction limiting both initial and any deviating risk-sharing arrangements. We refer to the ability to coordinate beliefs on future cooperation as *social capital*, and parameterize it by the probability $\pi \in [0, 1]$ that a coalition can achieve such coordination of beliefs.

We undertake our analysis within a classical insurance environment with only idiosyncratic income risk and assume that the probability π determines the coordination likelihood for both the original coalition and any deviating coalition. Hence, social capital impacts both the ex-ante payoff directly and values of the outside option. In particular, while lower social capital has the potential to decrease ex ante welfare (since ex ante belief coordination on future cooperation is lower), it can also raise ex ante welfare if the defecting coalition's value of insurance decreases sufficiently. This natural feedback provides an explanation for why more developed societies do not always "solve" the risk-sharing puzzle. Greater social capital increases the extent to which Pareto improving arrangements are formed, but also tightens the constraints on such arrangements, lowering welfare conditional on formation.

We make both a methodological and substantive contribution. Methodologically, we introduce (in Sections 2 and 3) an equilibrium model of coalition formation and stability. Loosely, an equilibrium allocation is robust to the possibility that a subset of agents (typically, but not always, the wealthy agents) could defect, not contribute in the current period and "reinitialize" risk-sharing using the *same* allocation. This notion captures the idea that the allocation is a credible social norm or "self-enforcing" arrangement (i.e., belief in the arrangement should not be self-defeating). A critical feature of the equilibrium notion is that the value of the outside option is endogenous, depending upon the allocation. As a consequence, the constraint set for the program determining the equilibrium allocation is not convex, necessitating an indirect approach to characterizing efficient allocations.

Section 4 describes the indirect approach: For some parameters, there is a fixed point characterization of equilibrium. In that case, an equilibrium allocation satisfies a stronger notion of robustness: it is robust to the possibility that a subset of agents could defect, not contribute in the current period and "reinitialize" risk-sharing using *any* allocation (Proposition 2).

Substantively, we characterize the behavior of equilibrium risk sharing (and its value) as a function of social capital. Section 5 describes some critical comparative statics of the fixed point characterization. There is a critical value of social capital, $\bar{\pi} \in (0, 1)$, such that for medium to low values of the social capital ($\pi \leq \bar{\pi}$), the fixed point characterization applies, and the second-best allocation can be determined using standard techniques (Section 6.1). For high values of social capital ($\pi > \bar{\pi}$), the value of the outside option is sufficiently high that equilibrium cannot satisfy the stronger notion of robustness mentioned above and utility must be "burnt" (Section 6.3). In all cases, however, *ex ante* welfare is nondecreasing in social capital, though the amount of insurance provided may be strictly decreasing in social capital π (due to its impact on outside options).

The remainder of the paper then proceeds as follows: Section 7 presents results for an illustrative set of examples to convey the qualitative properties of the equilibrium. Finally, Section 8 describes the related literature, and Section 9 concludes.

2 Model

2.1 The Environment: Endowments, Preferences and Technology

Time t is discrete and extends from period $t = 0$ to infinity. A unit measure of infinitely-lived agents are endowed with stochastic streams of the non-storable consumption good (henceforth referred to interchangeably as endowment or income y). Each agent in each

period has low income $y = \ell > 0$ and high income $y = h > \ell$ with equal probability; we write $Y := \{\ell, h\}$. Agent's income is independent across both agents and time. As usual, we assume that in any positive measure (i.e., *large*) collection of agents (and thus the economy as a whole), there is no aggregate income risk. We denote by $\overline{y}$ the aggregate level of output per capita ($\overline{y} = \frac{1}{2}(\ell + h)$). We denote an individual's income history by y^t. The probability of income history y^t is denoted by $\Pr(y^t)$.

All individuals have identical preferences over consumption in periods $t \geq 1$ given by

$$(1-\beta)E\left\{\sum_{t=1}^{\infty}\beta^{t-1}u(c_t)\right\},$$

where the utility function is strictly increasing, strictly concave and satisfies the Inada conditions, and where we multiply period utility by $(1-\beta)$ to express period utility and lifetime utility in the same units. The *autarky payoff*, the payoff from consuming one's endowment, is therefore given by

$$V^A(y) := (1-\beta)u(y) + \beta Eu(y) =: (1-\beta)u(y) + \beta V^A,$$

so that the ex ante autarky utility is $V^A := Eu(y)$. The first-best payoff is

$$V^{FB} := u(\overline{y}),$$

obtained from consuming the average income with certainty.

2.2 Social Contracting, Coalition Formation and Coalition Deviations

In the initial period $t = 0$, agents attempt to form a risk-sharing arrangement to obtain insurance against idiosyncratic income risk. Any arrangement needs to be robust to the possibility of deviations, either by single agents or by coalitions of agents. Agents decide on deviations after learning their current income. The continual threat of deviations implies that any coalitional arrangement must itself be self-enforcing (against the possibility that some members may deviate after that coalition has been formed). Because future income risk is more effectively shared in large coalitions, the possibility of forming a new large coalition is most threatening to the original coalition.

We do not model attempted coalition formation and the associated decision to deviate as a noncooperative game. Rather, we take a cooperative game-theoretic approach and impose incentive constraints that ensure that such deviations are not profitable. This also

means that we do not need to specify the outcome for the remaining agents after a successful deviation.

We view the ability of a group to successfully form a coalition as a reflection of high *social efficiency* or *social capital*, since a coalition only forms if its members are confident that future cooperation is sustainable. This confidence requires significant social cohesion since the incentive compatibility of future cooperation depends on intertemporal incentives that themselves need to be incentive compatible. We model the degree of social cohesion in an admittedly crude fashion by assuming that any attempt to form a coalition succeeds with an exogenous probability π. When a new (or deviating) coalitions fails to form (which happens with probability $1-\pi$), agents receive their autarky payoff V^A (and, in this sense, permanently lose their social capital).[1] We also assume that once the option to attempt secession has been exercised, it cannot be undone.

Finally, we assume that the allocation within any newly formed coalition is determined by a social planning problem in which all members *initially* have equal weights and therefore are treated ex ante symmetrically.

2.3 Preliminary Analysis: The Coalitions

We now argue that without loss of generality, we can restrict attention to *large homogeneous* coalitions. The sufficiency of large coalitions follows from two observations. First, any finite coalition's per capita outcome can be replicated by a large coalition with the same initial output composition. Second, the large coalition improves on the original outcome since it has no aggregate randomness.

We can restrict attention to homogeneous (by income y) deviating coalitions because we assume the initial bargaining weight of each agent in a newly formed coalition is fixed and equal, and each agent's decision to join a newly formed coalition is irrevocable: If a coalition successfully forms, then consumptions will be equalized for all agents in the deviating coalition in the first period, and consumptions thereafter will depend upon the agent's realized history. This implies that an agent will prefer a coalition with high, rather than low, first period per capita income. Agents with the high income realization will therefore prefer to join a coalition composed only of other individuals with the high income realization (and so leaving low income agents to form a coalition without them).

[1]The precise specification after a deviating coalition fails to form is not important (though it does have implications for our quantitative analysis); it is important that the failure of an attempt to deviate is costly.

3 Equilibrium

An *allocation* for a coalition is a consumption plan c specifying, for all periods t, an agent's consumption $c(y^t)$ in period t for every possible sequence $y^t \in Y^t$ of individual income shocks. We assume, again without loss of generality, that individual consumption depends only on that agent's income history, independent of identity.

The initial coalition formed in period 0 faces an ex ante notion of feasibility since the member income levels are not known at the time of coalition formation.

Definition 1 *An allocation for a coalition c is* resource feasible *if*

$$\sum\nolimits_{y^t} c(y^t) \Pr(y^t) \leq \overline{y}, \quad \forall t \geq 1. \tag{1}$$

The lifetime utility from an arbitrary consumption allocation c is given by

$$W^0(c) := (1-\beta) \sum_{\tau=1}^{\infty} \sum_{y^\tau} \beta^{\tau-1} \Pr(y^\tau) u(c(y^\tau)).$$

In period 0, all agents are identical, and they will agree to follow any resource-feasible consumption plan c that maximizes $W^0(c)$, as long as they can be confident that the consumption plan will be followed in the future. The danger is that some coalition may find it optimal to leave the original arrangement and internally insure. A necessary condition for a consumption plan to be a credible social norm is that if all the agents do believe in it today, that it should not be the case that after some history, some large coalition finds it optimal to deviate, and after the deviating period follow the *same* consumption plan.[2] Phrased differently, suppose the grand coalition believes that the allocation $\tilde{c}$ is credible, but that a coalition after some history y^t with current income y_t receives strictly higher payoff from seceding, and if successful in coordinating beliefs, implementing $\tilde{c}$ from the next period. Such a history means that the grand coalition should not have believed in the credibility of the original allocation $\tilde{c}$, since it will *not* be implemented in its entirety. Accordingly, we are interested in allocations that are not subject to such a criticism.

For an arbitrary income history $y^t \in Y^t$, the *continuation* lifetime utility under the allocation is

$$W(y^t, c) := (1-\beta) u(c(y^t)) + (1-\beta) \sum_{\tau=1}^{\infty} \sum_{y^\tau} \beta^\tau \Pr(y^\tau) u(c(y^t y^\tau)),$$

[2]Since the coalition is large, the (per capita) resource-feasibility constraint faced by the coalition is identical to the (per capita) resource-feasibility constraint.

where $y^t y^\tau$ denotes the $t+\tau$-history that is the concatenation of t-period history y^t and the τ-period history y^τ.

Definition 2 *An allocation c is* internally-incentive feasible *if for all $t \geq 1$ and for all $y^t \in Y^t$,*

$$\begin{aligned} W(y^t, c) &\geq \pi\{(1-\beta)u(y_t) + \beta W^0(c)\} + (1-\pi)V^A(y_t) \\ &= (1-\beta)u(y_t) + \beta[\pi W^0(c) + (1-\pi)V^A] \end{aligned} \tag{2}$$

Let $\mathcal{C}$ denote the set of resource feasible and internally-incentive feasible allocations.

This is a weak notion of credibility when coalitional deviations are possible. For example, while the autarky allocation is trivially internally-incentive feasible, that allocation has lower utility than allocations with some insurance. The stability notion is "internal" in the sense that when evaluating the credibility of an allocation, agents only consider the possibility that if accepted, that allocation will also determine the outside for any deviating coalition.[3] Agents do not consider the possibility that the payoffs for a deviating coalition may determined by a different (possibly more attractive) allocation. As in the cooperative-game-theory and renegotiation-proof repeated-games literatures,[4] the stronger requirement (which we discuss just after Proposition 2 in Section 4) can lead to nonexistence of equilibrium.

The internal-incentive constraint (2) is the key friction that prevents full consumption insurance within a coalition.

Definition 3 *For given social capital π, an allocation c is an* equilibrium allocation *if it solves the program*

$$\max_{c \in \mathcal{C}} \; W^0(c).$$

Denote by $\mathbb{W} = \max_{c \in \mathcal{C}} \; W^0(c)$ the resulting optimal lifetime utility and by $\mathbb{F} = \pi\mathbb{W} + (1-\pi)V^A$ the associated ex ante (and so deviation continuation) utility.

An equilibrium allocation c is the best ex ante resource-feasible and internally-incentive-feasible allocation. Note that an equilibrium allocation maximizes ex ante utility given π, as well as the utility conditional on the agreement being reached. The value $\mathbb{W}$ is the maximum per capita value the grand coalition can achieve, given the credible threat that any group

[3] In this sense, the notion is similar to von Neumann and Morgenstern's (1944) *internal stability* notion; see the discussion in Greenberg (1990, Section 2.3). It is also similar to Farrell and Maskin's (1989) notion of *weakly renegotiation proof* in repeated games.

[4] For the former, the stronger analogous notion is von Neumann and Morgenstern's (1944) *external stability*; again see the discussion in Greenberg (1990, Section 2.3). For the latter, the analogous stronger notion is called *strongly renegotiation proof*; see Farrell and Maskin (1989).

of agents will deviate (and implement the same agreement) if the initial arrangement is not sufficiently generous to that group. Recall that if any group has an incentive to deviate, then a homogeneous large group does.

Since the autarkic allocation is trivially resource and internally-incentive feasible, the set of resource and internally-incentive-feasible allocations is nonempty, and so the supremum of $W^0(c)$ exists and is bounded above by $u(\overline{y})$, the utility of first-best insurance. We will show that in fact the supremum is always attained and so equilibrium exists.

Our first result (the proof is a straightforward calculation) is that first-best insurance is consistent with equilibrium only when social capital is not too large (and agents are sufficiently patient).

Proposition 1 *The first-best allocation is an equilibrium allocation if and only if*

$$\pi \leq \pi^{FB} := 1 - \frac{(1-\beta)[u(h) - V^{FB}]}{\beta\,[V^{FB} - V^{A}]} < 1. \tag{3}$$

Moreover, if

$$\beta < \beta^{FB} := \frac{u(h) - V^{FB}}{u(h) - V^{A}},$$

then $\pi^{FB} < 0$ *and full insurance is not an equilibrium for any level of social capital* π*.*

The requirement that social capital not be too large for full insurance should not be surprising. Under the first-best allocation, the currently h-income agents sacrifice current consumption to insure the currently ℓ-income agents. If π is close to one, seceeding and then immediately insuring within the deviating coalition incurs almost no loss in insurance and so secession is attractive.

Of more interest is the possibility of partial insurance in equilibrium, as illustrated by the next example. As in Krueger and Perri (2011), where the outside option is fixed, the lower bound on β in Example 1 turns out to be necessary for insurance as well (see Proposition 3.1 in the next section).

Example 1 Suppose $\beta u'(\ell) > u'(h)$, and consider the allocation

$$c_\varepsilon(y^t) = \begin{cases} h - \varepsilon, & y_t = h, \\ \ell + 2\varepsilon, & y_{t-1} = h, y_t = \ell, \\ \ell, & \text{otherwise.} \end{cases}$$

This allocation satisfies resource feasibility with equality in every period *except* the initial period, when ε resources are destroyed. We claim that for $\varepsilon > 0$ small, $c_\varepsilon \in \mathcal{C}$. Observe first that $W^0(c_\varepsilon) > V^A$ for ε small, and so this allocation does provide partial insurance.

A sufficient condition for $c_\varepsilon \in \mathcal{C}$ is

$$W(h, c_\varepsilon) \geq (1-\beta)u(h) + \beta W^0(c_\varepsilon). \tag{4}$$

This is the condition for internal-incentive feasibility when $\pi = 1$, which is stricter than internal-incentive feasibility for any $\pi < 1$ when $W^0(c_\varepsilon) > V^A$.

By deviating, an agent in the h-coalition gives up one period of 2ε insurance in the event that she has ℓ income in the next period (which occurs with probability $1/2$). So a sufficient condition for (4) to hold for ε small is that the marginal benefit of deviating be smaller than the marginal expected delayed cost,

$$(1-\beta)u'(h)\varepsilon \leq (1-\beta)\frac{\beta}{2}u'(\ell)2\varepsilon,$$

which reduces to the assumed bound on β.

★

Two features of Example 1 deserve mention. The first is that the initial period resource destruction plays a critical role in the internal-incentive feasibility of Example 1's allocation. In particular, if the ε resources sacrificed by the initial h-income agents is given to the initial ℓ-income agents (providing additional ex ante insurance), the resulting allocation is *not* internally-incentive feasible for high π (it is internally-incentive feasible for π close to 0); the proof of Lemma A.3 uses this property of the modified allocation.

The second is the time-varying nature of the insurance provided. When first-best insurance is not internally-incentive feasible, h-income agents optimally secede under the first-best allocation. To reduce this secession incentive, a natural modification is to consider simple allocations of the form

$$c_\zeta(y^t) := \begin{cases} h - \zeta, & y_t = h, \\ \ell + \zeta, & y_t = \ell. \end{cases} \tag{5}$$

For $\zeta = 0$, c_ζ is the autarkic allocation, while for $\zeta = h - \overline{y}$, c_ζ is the first-best allocation. While such an allocation can be internally-incentive feasible, it is less efficient in its provision of incentives. For example, for $\pi = 0$, c_ζ is only internally-incentive feasible for

$$\beta \geq \frac{2u'(h)}{u'(\ell) + u'(h)} > \frac{u'(h)}{u'(\ell)}.$$

The allocation in Example 1 achieves partial insurance without violating incentive feasibility for lower β by rewarding h-income agents through insurance: in exchange for giving up ε today, the allocation promises 2ε in insurance to any agent realizing ℓ tomorrow (while providing no insurance to agents who had realized ℓ previously and continue to realize ℓ).

4 Equilibrium as a Fixed Point

Solving for equilibrium allocations is complicated by the nature of the internal-incentive-feasibility constraint. In particular, the set of internally-incentive-feasible allocations is not convex. This lack of convexity arises from the endogeneity of the outside option, i.e., the deviating coalition's payoff. Accordingly, we follow an indirect path that first solves for equilibrium via a fixed point argument for a subset of values of π, and then solves for equilibrium for the remaining values of π.

Recall that internal-incentive feasibility requires

$$W(y^t, c) \geq (1-\beta)u(y_t) + \beta[\pi W^0(c) + (1-\pi)V^A] \qquad \forall y^t \in \cup_\tau Y^\tau.$$

We begin by considering resource-feasible allocations that satisfy an exogenous version of this constraint, which we call the *incentive-feasibility constraint*,

$$W(y^t, c) \geq (1-\beta)u(y_t) + \beta F \qquad \forall y^t \in \cup_\tau Y^\tau. \tag{6}$$

For exogenous $F \in \mathbb{R}_+$, denote by $\mathcal{C}(F)$ the set of resource-feasible allocations satisfying (6). If F is too large, then $\mathcal{C}(F)$ will be empty. But if c is internally-incentive feasible, then $c \in \mathcal{C}(\pi W^0(c) + (1-\pi)V^A)$, and so the constraint set $\mathcal{C}(F) \neq \varnothing$ is non-empty for outside options $F \leq \pi W^0(c) + (1-\pi)V^A$.

When $\mathcal{C}(F) \neq \varnothing$, define

$$\mathbb{V}(F) := \max_{c \in \mathcal{C}(F)} W^0(c). \tag{7}$$

Social capital π does not appear in the maximization in (7). Instead, the exogenous value of the outside option F determines the optimal allocation and value.[5] But there is a connection. Since a deviating coalition only successfully coordinates after deviation with probability π, if F is the implied continuation value of the outside option for a deviating coalition, then,

[5]We discuss the connection with the earlier literature on efficient insurance under limited commitment with an exogenous outside option in a remark at the end of Section 6.1 and in the Related Literature Section 8.

for all $y \in Y$, the value of the outside option is determined by the mapping

$$\mathcal{T}(F;\pi) := \pi\mathbb{V}(F) + (1-\pi)V^A.$$

Proposition 2 *Suppose $F = \pi W^0(c^\dagger) + (1-\pi)V^A$ is a fixed point of $\mathcal{T}(\cdot;\pi)$ for some allocation $c^\dagger \in \mathcal{C}(F)$. Then $W^0(c^\dagger) = \mathbb{V}(F)$, $c^\dagger$ is an equilibrium allocation, and F is the ex ante value of the equilibrium.*

Proof. It is immediate that $W^0(c^\dagger) = \mathbb{V}(F)$ and that F is the ex ante value of the equilibrium if $c^\dagger$ is an equilibrium allocation. It remains to argue that $c^\dagger$ is an equilibrium allocation.

Since $c^\dagger \in \mathcal{C}(F)$, $c^\dagger$ is internally-incentive compatible. If $c^\dagger$ is not an equilibrium, there exists a resource and internally-incentive-compatible allocation c' with

$$W^0(c') > W^0(c^\dagger).$$

Then, for all $t \geq 1$ and $y^t \in Y^t$,

$$\begin{aligned} W(y^t, c') &\geq (1-\beta)u(y_t) + \beta[\pi W^0(c') + (1-\pi)V^A] \\ &> (1-\beta)u(y_t) + \beta[\pi W^0(c^\dagger) + (1-\pi)V^A] \\ &= (1-\beta)u(y_t) + \beta F, \end{aligned}$$

and so $c' \in \mathcal{C}(F)$, implying $W^0(c^\dagger)$ could not be a fixed point of $\mathcal{T}(\cdot;\pi)$. □

Proposition 2 indicates that equilibria exists for those π consistent with outside options that are fixed points of $\mathcal{T}(\,\cdot\,;\pi)$. But this is uninformative without a better understanding of the fixed points of $\mathcal{T}(\,\cdot\,;\pi)$ (which we provide in the next section).

The equilibrium nature of the fixed points of $\mathcal{T}(\cdot;\pi)$ deserves comment. The fixed points (when they exist) satisfy a *stronger* notion of credibility than that captured by internal-incentive feasibility. In particular, if $F = \pi W^0(c^\dagger) + (1-\pi)V^A$ is a fixed point of $\mathcal{T}(\,\cdot\,;\pi)$ for some allocation $c^\dagger \in \mathcal{C}(F)$, then it is robust to the threat of secession from any coalition when any seceding coalition is free to reoptimize subject only to the constraint that there may be further deviations by subcoalitions. As mentioned earlier, this is analogous to stronger notions of stability and renegotiation-proofness in game theory that are known to have nonexistence problems. Similarly, in our setting, there is no guaranteee that $\mathcal{T}(\,\cdot\,;\pi)$ will have a fixed point.

If a fixed point does exist, it is unique because $\mathbb{V}(F)$ and thus $\mathcal{T}(F;\pi)$ is weakly decreasing in F. The fixed point may fail to exist because the constraint set is not a "nice" function of

the parameter F, or the constraint set is empty for F in a relevant region. While Proposition 4 below (proved in Appendix A) assures us that the former is not an issue (the constraint set is a "nice" function of F), the constraint set *is* empty for large F (which will correspond to large π) and so a fixed point does not exist in that case. Define

$$\bar{F} := \sup\{F \mid \mathcal{C}(F) \neq \varnothing\}.$$

We can now state the main result of the paper (which summarizes the analysis to follow):

Proposition 3 *Equilibrium exists for all $\pi \in [0,1]$.*

1. *Suppose $\beta \leq u'(h)/u'(\ell)$. There is no risk sharing in equilibrium (i.e., autarky is the unique equilibrium).*

2. *Suppose $\beta > u'(h)/u'(\ell)$. Risk sharing does occur in equilibrium. There exists a value of π, $\bar{\pi} \in (0,1)$, such that*

   (a) *for $\pi \in [0, \bar{\pi}]$, equilibrium is unique and its ex ante value is strictly increasing in π, equaling $\bar{F} > V^A$ at $\bar{\pi}$, and*

   (b) *for $\pi \in (\bar{\pi}, 1]$, equilibrium allocations are not unique, but all have the same ex value of $\bar{F}$.*

Proof.

1. This is an implication of the machinery we develop to characterize $\bar{F}$, and is proved in Corollary 1 in Section 6.2.

2. (a) This is an immediate implication of Propositions 2 and 4 (which is in the next section).

   (b) This is Lemmas 1 and 2 in Section 6.3.

□

5 Understanding Equilibrium Values

We begin by studying the program (7) and the fixed points of $\mathcal{T}(\,\cdot\,;\pi)$. The proof of the following result is in Appendix A.

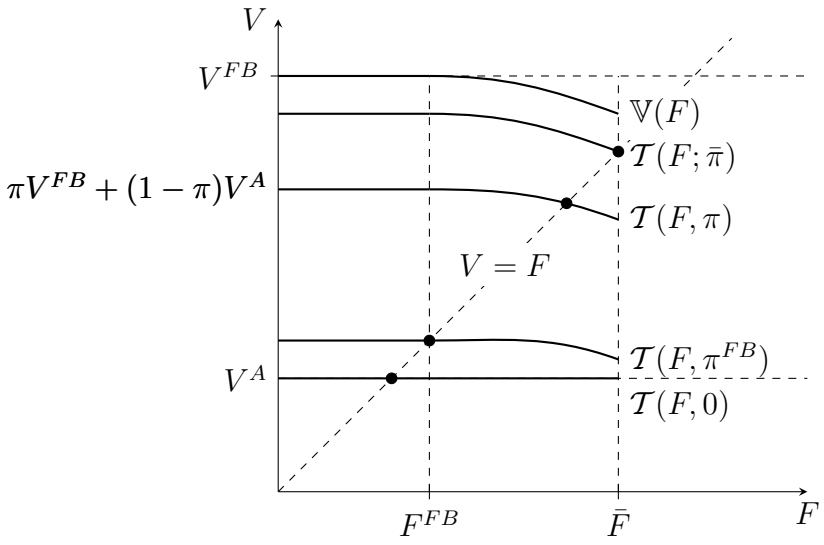

Figure 1: Determination of the fixed point of $\mathcal{T}(F;\pi) = \pi\mathbb{V}(F) + (1-\pi)V^A$ for different values of π. Drawn assuming $\beta > \beta^{FB}$, defined in Proposition 1 (Lemma A.2 verifies that in this case $F > F^{FB}$); if $\beta < \beta^{FB}$, then $F^{FB} < V^A$.

Proposition 4 *Suppose $\beta > u'(h)/u'(\ell)$.*

1. *$V^A < \bar{F}$.*

2. *$\mathcal{C}(\bar{F}) \neq \varnothing$.*

3. *For $F \leq \bar{F}$, the value of the problem (7), $\mathbb{V}(F)$, is continuous in F.*

4. *For all $F \in [V^A, \bar{F}]$, $F < \mathbb{V}(F)$.*

5. *Defining*

$$\bar{\pi} = \frac{\bar{F} - V^A}{\mathbb{V}(\bar{F}) - V^A} \in (0,1),$$

for all $\pi \in (0,\bar{\pi}]$, $\mathcal{T}(\,\cdot\,,\pi)$ has a fixed point. The value of this fixed point is increasing in π and

$$\bar{F} = \mathcal{T}(\bar{F};\bar{\pi}).$$

6. *For $\pi > \bar{\pi}$, $\mathcal{T}(\,\cdot\,,\pi)$ does not have a fixed point.*

Figure 1 presents the previous proposition graphically by plotting $\mathbb{V}(F)$ and $\mathcal{T}(F;\pi)$ against the value of the outside option F for various degrees of social capital π. At one extreme, $\pi = 0$ and we have $\mathcal{T}(F;0) = V^A$ and thus trivially $F = V^A$ is the unique fixed point for the outside option. In this case, for $\beta \geq \beta^{FB}$, Proposition 1 implies $\mathbb{V}(V^A) = V^{FB}$ and the allocation for the initial coalition would feature full insurance (but since $\pi = 0$, it

never successfully forms). From Proposition 1, full insurance remains the outcome for the successful coalition as long $\pi \leq \pi^{FB} < 1$. The associated largest deviation lifetime utility F^{FB} for which the full-insurance allocation can be sustained inside the initial coalition is given by

$$F^{FB} := \pi^{FB} V^{FB} + (1 - \pi^{FB}) V^A.$$

For $\pi \in (\pi^{FB}, \bar{\pi}]$, the value of the outside option F is determined as the fixed point of $\mathcal{T}(\,\cdot\,;\pi)$. The fixed point is larger than F^{FB}, and so the constraint (6) strictly binds at least for households with currently high income, implying the initial coalition cannot sustain first-best insurance (i.e., $\mathbb{V}(F) < V^{FB}$) and that the utility $\mathbb{V}(F)$ it delivers is strictly decreasing in F.

Proposition 4.4 implies that $\bar{\pi} < 1$. To see why $F < \mathbb{V}(F)$, suppose that for some $F > V^A$, we have $F = \mathbb{V}(F)$. But then that F is a fixed point of $\mathcal{T}(\,\cdot\,;1)$. In other words, by seceding, a rich coalition can guarantee a payoff of $(1-\beta)u(h) + \beta\mathbb{V}(F)$, implying that the optimal consumption allocation must promise these agents a current consumption of at least h. But then there is no insurance, contradicting $F > V^A$.

For $\pi > \bar{\pi}$,

$$\pi \mathbb{V}(\bar{F}) + (1-\pi) V^A > \bar{F}.$$

Since $\mathcal{C}(F)$ is empty for $F > \bar{F}$, this implies that $\mathcal{T}(\cdot;\pi)$ does not have a fixed point. However, this does *not* imply that there is no equilibrium (recall that the fixed point characterizes a stronger notion of incentive feasibility, and is only a sufficient condition for equilibrium in our setting).

Suppose $\mathbb{c}$ is an equilibrium allocation with value $W^0(\mathbb{c})$. Then it must satisfy

$$\mathbb{c} \in \mathcal{C}(\pi W^0(\mathbb{c}) + (1-\pi) V^A),$$

and so

$$\pi W^0(\mathbb{c}) + (1-\pi) V^A \leq \bar{F}. \tag{8}$$

Since $\pi > \bar{\pi}$, we have $W^0(\mathbb{c}) < \mathbb{V}(\bar{F})$, leading us to define:

Definition 4 *An equilibrium allocation* $\mathbb{c}$ burns utility *if*

$$W^0(\mathbb{c}) < \mathbb{V}(\bar{F}).$$

An equilibrium allocation maximizes ex ante utility (the left side of (8)). We show in Section 6.3 that equilibrium allocations in fact satisfy (8) with equality.

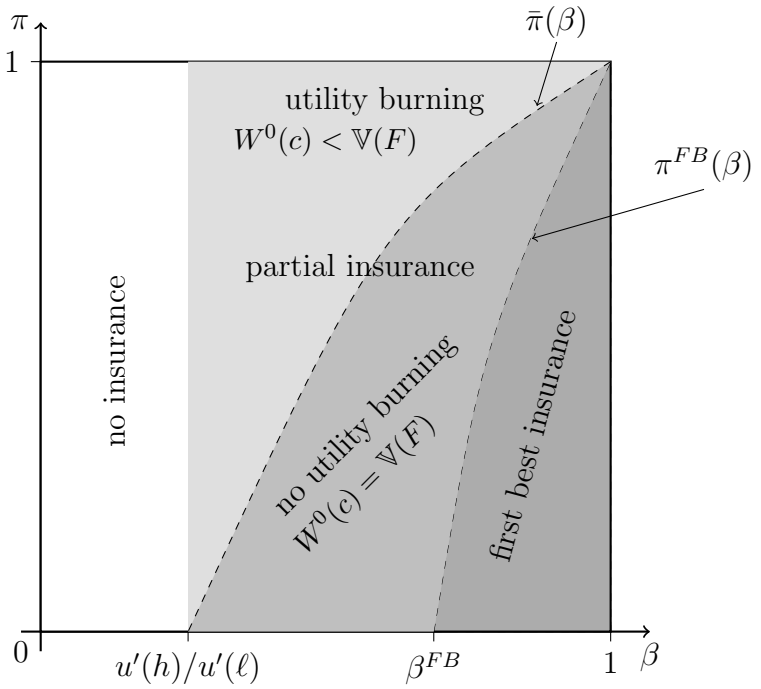

Figure 2: The insurance possibilities as a function of the discount factor and social capital. Equation (3) shows that $\pi^{FB}(\beta^{FB}) = 0$; see Lemma A.3 for the proof that $\bar{\pi}(\beta) = 0$ when $\beta = u'(h)/u'(\ell)$.

Our notation suppresses the dependence of $\bar{\pi}$ and π^{FB} on β, but it is worthwhile to clarify the relationship between β and π, which is illustrated in Figure 2. For $\pi \leq \pi^{FB}(\beta)$, a successfully formed coalition provides its members with full insurance and ex-ante utility is strictly increasing in social capital. For all $\pi \in (\pi^{FB}(\beta), \bar{\pi}(\beta)]$, $\mathcal{T}(\cdot, \pi)$ has a fixed point and its value (the value of ex ante utility) is strictly increasing in π. The associated allocation features partial insurance that gets worse with π, as does lifetime utility conditional on successfully forming the coalition. Finally, for $\pi > \bar{\pi}(\beta)$, $\mathcal{T}(\cdot; \pi)$ does not have a fixed point, the internal-feasibility constraint is binding in equilibrium, expected lifetime utility is fixed at $\bar{F}$ independent of π (since (8) holds with equality) and attained with an allocation that features utility burning and partial risk sharing.

6 Characterizing Equilibrium Allocations

From Proposition 1, if $\pi \leq \pi^{FB}$, the first-best allocation is consistent with equilibrium.

6.1 The case of no utility burning, $\pi \in (\pi^{FB}, \bar{\pi}]$

We now characterize the equilibrium allocations for intermediate values of π, that is, values of π that are consistent with a fixed point of $\mathcal{T}(\,\cdot\,;\pi)$ exceeding F^{FB}. We have already seen that this is equivalent to characterizing the allocations that maximize $W^0(c)$ subject to $c \in \mathcal{C}(F)$ for $F \in (F^{FB}, \bar{F}]$. This is a strictly concave problem, and so has a unique solution, that we denote by $\mathbb{c}$.

We first state some standard properties of the optimal allocation. The proofs (most of which are standard, though tedious, variational arguments) are left to Appendix B.

Proposition 5 *Suppose $\beta u'(\ell) > u'(h)$ and $F \in (F^{FB}, \bar{F}]$. The optimal allocation $\mathbb{c}$ has the following properties:*

1. *There exists $\delta_{t+1} < 1$ such that if incentive feasibility does not bind at y^{t+1}, then*

$$\frac{u'(\mathbb{c}(y^t))}{u'(\mathbb{c}(y^{t+1}))} = \delta_{t+1} \tag{9}$$

   and so

$$\mathbb{c}(y^t) > \mathbb{c}(y^{t+1}).$$

2. *Incentive feasibility binds at all $y^{t-1}h$, and so for all y^{t-1},*

$$W(y^{t-1}h, \mathbb{c}) = (1-\beta)u(h) + \beta F =: W^F(h), \tag{10}$$

   and for all y^{t-1} and $\hat{y}^{t-1}$,

$$\mathbb{c}(y^{t-1}h) = \mathbb{c}(\hat{y}^{t-1}h) =: \mathbb{c}_t(h).$$

3. *If incentive feasibility binds at some $y^{t-1}\ell$, then it binds at $y^{t-1}\ell\ell$.*

4. *If incentive feasibility binds at $y^t\ell$, then $\mathbb{c}(y^t\ell) = c_\ell(F)$, where $c_\ell(F) > \ell$ solves*

$$u(c_\ell(F)) = u(\ell) + \beta(F - V^A) > u(\ell),$$

   and for all y^t,

$$\mathbb{c}(y^t\ell) \geq c_\ell(F). \tag{11}$$

5. *Incentive feasibility does not bind in the initial period at ℓ, nor after any history of the form $y^t h\ell$.*

6. There is an L such that for $0 \leq k < L$ and all histories y^{t-1-k}, $\hat{y}^{t-1-k}$

$$\mathbb{c}(y^{t-1-k}h\ell^k) = \mathbb{c}(\hat{y}^{t-1-k}h\ell^k) := \mathbb{c}_t(h\ell^k),$$

and for $k \geq L$, $\mathbb{c}(y^{t-1-k}h\ell^k) = c_\ell(F)$.

Proposition 5 implies that the optimal allocation is a sequence of consumption ladders: the optimal consumption in any period is determined by the number of ℓ realizations after the last h realization with consumption falling after each additional ℓ realization until the consumption floor $c_\ell(F)$ is reached. Accordingly, with a slight abuse of notation, we write $c_{t+k}(h\ell^k)$ for the consumption in period $t+k$ after any history $y^{t-1-k}h\ell^k$.

Definition 5 *A* period-t consumption ladder *is a finite sequence of consumptions, denoted $\left(\left(\mathbb{c}_{t+k}(h\ell^k)\right)_{k=0}^{L-1}, c_\ell(F)\right)$, specifying for each $k = 0, \ldots, L$, the consumption in period $t+k$ of an agent who had the income history $y^{t-1}h\ell^k$. A* stationary consumption ladder *is a finite sequence of consumptions, denoted $\left(c_*(h\ell^k)\right)_{k=0}^{L}$, specifying the consumption in any period t of an agent who had the income history $y^{t-k-1}h\ell^k$.*

We extend any finite consumption ladder to an infinite ladder (sequence) by setting $c_{t+k}(h\ell^k) := c_\ell(F)$ for $k \geq L$.

A period-t consumption ladder specifies the current and future consumption of an agent with current h-income and future ℓ-income realizations. If that agent again receives h in the subsequent period $t+k$, her consumption from period $t+k$ on is determined by a period $t+k$ consumption ladder. Consequently, the continuation lifetime utility of any agent with current h-income is determined by the details of the current and *future* consumption ladders.

The calculation of lifetime utility is simpler when the current and future ladders agree, i.e., for a stationary ladder. The lifetime utility of an agent with currently high income from a *stationary* ladder c_* is

$$\begin{aligned} W(h, c_*) &= (1-\beta)u(c_*(h)) \\ &\quad + \tfrac{\beta}{2}\left\{(1-\beta)u(c_*(h\ell)) + W(h, c_*)\right\} \\ &\quad + \left(\tfrac{\beta}{2}\right)^2\left\{(1-\beta)u(c_*(h\ell^2)) + W(h, c_*)\right\} \\ &\quad \vdots \\ &= (1-\beta)\sum_{k=0}^{\infty}(\tfrac{\beta}{2})^k u(c_*(h\ell^k)) + \tfrac{\beta}{2-\beta}W(h, c_*), \end{aligned}$$

and so, simplifying, we get

$$W(h, c_*) = \left(1 - \tfrac{\beta}{2}\right) \sum_{k=0}^{\infty} (\tfrac{\beta}{2})^k u(c_*(h\ell^k)). \tag{12}$$

The only income histories for which consumption is not specified by any consumption ladder have the form ℓ^k, and those consumptions are pinned down by resource feasibility, since in every period there is only one such history.

For $F > F^{FB}$, the optimal allocation provides maximal risk sharing consistent with incentive feasibility. Incentive feasibility always binds for h-income agents and sometimes for ℓ-income agents.

In order to deter an h-income agent from seceding, the optimal allocation does two things: First, it reduces her transfer to low-income individuals below the first-best level. And, second, the risk sharing offered is "front-loaded" so that ℓ-income agents who had more recently received a h realization receive more insurance than those who last received a h realization further in the past.

This front-loading (reflected in the declining consumption ladder) implies that eventually the consumption specified after a sufficiently long string of ℓ-realizations is determined by incentive feasibility for the ℓ-realization. The resulting lower bound on consumption, $c_\ell(F) > \ell$ reflects the following tradeoff: Seceding from $\mathbb{c}$ does mean that the agents give up some risk-sharing today, but the benefit is that in a new coalition tomorrow, any agent who receives another ℓ realization receives more generous risk sharing tomorrow (since incentive feasibility does not bind in the first period after ℓ by Proposition 5.5, $c_\ell(F) < \mathbb{c}(\ell)$).

Remark 1 When $\pi \leq \bar{\pi}$ and the equilibrium allocations are solutions to the fixed point of $\mathcal{T}(\,\cdot\,;\pi)$, the consumption allocations within the coalition can be decentralized as in Kehoe and Levine (1993). In the literature stimulated by Kehoe and Levine (1993), the outside option is taken to be autarky, but the key is that the efficient allocation is generated by optimizing against this option (see, for example, Chien and Lustig (2009), Alvarez and Jermann (2000)). This decentralization is one *within* a successfully formed coalition. The individual's optimization problem in this decentralization is to choose her consumption allocation so as to maximize her ex ante payoff subject to a single intertemporal budget constraint and a sequence of incentive constraints for each history state y^t. In the individual's present value budget constraint the price of a unit of consumption in her history state y^t is given by $\gamma_t Pr(y^t)$, where γ_t is the resource multiplier from the coalition's social planning problem given the outside options and hence corresponds to the individual-level present value constraint. In addition, the incentive compatibility constraints at the individual level are exactly

the incentive feasibility constraints in (6). Thus,the individual's problem is isomorphic to the Lagrangian problem for the coalition. Note that this is not the case when $\pi > \bar{\pi}$ since the coalition must respect a coalition-level constraint on the overall level of ex ante welfare. ♦

6.2 Characterizing $\bar{\pi}$

We now characterize $\bar{\pi}$, or equivalently, $\bar{F}$. It turns that $\bar{F}$ has a simple characterization as the maximum value of the outside option consistent with h-incentive feasibility. In particular, a specific stationary ladder attains this maximal sustainable deviation payoff $\bar{F}$. Using this property, we then argue that the associated equilibrium allocation also converges to this stationary ladder.

We are interested in the stationary ladder that maximizes ladder lifetime utility $W(h, c_*)$, given in (12), subject to incentive feasibility for ℓ realizations and resource feasibility. Recalling (11), this is

$$\mathbb{V}^*(h; F) := \max_{c \in \mathcal{C}_*(F)} W(h, c_*), \tag{13}$$

where $\mathcal{C}_*(F)$ is the set of infinite stationary ladders satisfying resource feasibility

$$\sum_{k=0}^{\infty} (\tfrac{1}{2})^{k+1} c_*(h\ell^k) \leq \overline{y} \tag{14}$$

and incentive feasibility

$$c_*(hy^k) \geq c_\ell(F) \text{ for all } k \geq 1. \tag{15}$$

In this problem, h-incentive feasibility does not appear as a constraint because we are maximizing the payoff of the h agents. Note also that resource feasibility is being imposed on the ladder, and so there is only one constraint. In contrast, resource feasibility was not imposed on any ladder in $\mathcal{C}(F)$, being imposed instead in each period.

The next proposition (proved in Appendix C) makes precise the sense in which $\bar{F}$ is the maximum value of the outside option consistent with h-incentive feasibility.

Proposition 6 *The set of resource and incentive feasible allocations $\mathcal{C}(F)$ is nonempty if and only if*

$$\mathbb{V}^*(h; F) \geq W^F(h),$$

where $W^F(h)$ is the deviation value of high-income individuals defined in (10). Moreover,

$$F = \bar{F} \iff \mathbb{V}^*(h; F) = W^F(h).$$

Corollary 1 *If $\beta u'(\ell) \leq u'(h)$, then*

$$\bar{F} = V^A.$$

Proof. Suppose $\bar{F} > V^A$. By Proposition 6, for all $F \in (V^A, \bar{F}]$,

$$\mathbb{V}^*(h, F) \geq (1-\beta)u(h) + \beta F. \tag{16}$$

But $\beta u'(\ell) \leq u'(h)$ implies that the autarkic consumption provides an upper bound for (13) and so

$$\begin{aligned}
\mathbb{V}^*(h, F) &\leq (1-\beta)u(h) + \frac{(1-\beta)\beta}{2-\beta}u(\ell) + \frac{\beta}{2-\beta}[(1-\beta)u(h) + \beta F] \\
&= (1-\beta)u(h) + \frac{\beta}{2-\beta}[(1-\beta)2V^A + \beta F] \\
&< (1-\beta)u(h) + \beta F, \qquad (17)
\end{aligned}$$

because

$$\begin{aligned}
&(1-\beta)2V^A + \beta F < (2-\beta)F \\
\iff \qquad &(1-\beta)2V^A < 2(1-\beta)F.
\end{aligned}$$

Since (17) contradicts (16), we must have $\bar{F} = V^A$. □

This corollary shows that under the specified condition the highest outside option that can be attained is autarky, and thus under this condition the only equilibrium is one without any insurance.

It remains to characterize the allocation that maximizes ex ante utility at $\bar{F}$ (the proof is in Appendix C)).

Proposition 7 *Suppose $\beta u'(\ell) > u'(h)$ and $F = \bar{F}$. The equilibrium allocation $\mathbb{c}$ converges to the unique solution to problem (13), $\bar{c}_*$, that is (where L is from Proposition 5.6),*

$$\lim_{t\to\infty} \mathbb{c}_t(h\ell^k) = \bar{c}_*(h\ell^k) \quad \textit{for any } k < L$$

and

$$\mathbb{c}_t(\ell^L) = c_\ell(\bar{F}).$$

Suppose u is CRRA, i.e., for some $\gamma \geq 0$,

$$u(c) = \begin{cases} \dfrac{c^{1-\gamma} - 1}{1-\gamma}, & \gamma \neq 1, \\ \log(c), & \gamma = 1. \end{cases} \tag{18}$$

Then convergence to the optimal stationary ladder (which is given by $\bar{g} = \beta^{1/\gamma}$) does not occur in finite time.

We have not been able to prove an analogous result to Proposition 7 when $F < \bar{F}$. Indeed, in these cases it is not obvious what the appropriate limiting stationary ladder is. Nonetheless, we can gain some insight by considering the following variant of our model: Assume (as we do in our computational exercises) that utility is CRRA, and suppose only coalitions with high income realizations can leave. In other words, ignore the ℓ-incentive feasibility, but maintain resource and h-incentive feasibility. Now, agents with a current ℓ realization never face a binding incentive constraint, and so in the optimal allocation, such individuals have consumptions that decay at a common rate. Of course, there is no floor on the consumption of such agents (beyond the feasibility floor of 0). This suggests that a stationary ladder of the form $c(h\ell^k) = c_h g^k$ will be optimal, for some value of g. The stationary resource constraint is then given by

$$\bar{y} = \frac{c_h}{2} \sum_{j=0}^{\infty} \left(\frac{g}{2}\right)^j = \frac{c_h}{2} \frac{1}{1 - \frac{g}{2}},$$

and so, $c_h = (2-g)\bar{y}$.

Consider the allocation in which the h agents are immediately put on the stationary ladder (and so after the history $y^{t-k-1}h\ell^k$ have consumption $c_h g^k$). In period t, agents with realizations ℓ^t receive the residual consumption

$$\bar{y} - \sum_{k=0}^{t-1} c_h g^k 2^{-k-1} = \bar{y} - \tfrac{1}{2} c_h \frac{1 - (g/2)^t}{1 - (g/2)} = \bar{y}(g/2)^t,$$

and since the mass of such agents is 2^{-t}, their per capita consumption is $\bar{y}g^t$. But this implies that the per capita consumption of the "residual" agents is declining at the *same* rate as agents with histories of the form $h\ell^k$, suggesting that the allocation in which the h agents are immediately put on the stationary ladder is in fact ex ante when the ℓ-incentive constraints are ignored.

It remains to pin down g, which is determined from the binding h-incentive-feasibility constraint for given F. So the ex ante value of the stationary ladder implied by that g is an upper bound for $W^0(\mathbb{c})$. A natural lower bound is given by the ex ante utility from putting the high income agents immediately on the stationary ladder with the consumption floor $c_\ell(F)$ and a binding h-incentive-feasibility constraint. The calculations in Section 7 suggest that these two bounds can be close.

6.3 The case of utility burning, $\pi > \bar{\pi}$

For high values of social capital ($\pi > \bar{\pi}$), equilibrium requires utility burning. While equilibrium must now impose additional inefficiencies, the precise nature of these inefficiencies is not determined. Rather, these inefficiencies are chosen to exactly offset the increase in social capital so that the ex ante value remains at $\bar{F}$.

We present two lemmas, illustrating two possible choices of inefficiencies due to either postponing risk sharing or burning resources. Denote by $\bar{\mathbb{c}}$ the optimal consumption for $F = \bar{F}$. The first lemma describes an equilibrium that postpones risk sharing.

Lemma 1 *Suppose $\pi > \bar{\pi}$. Denote the allocation specifying T periods of autarkic consumption followed by $\bar{\mathbb{c}}$ in a history independent manner by $c^{(T)}$. There exists $T(\pi)$ and $\alpha(\pi) \in [0,1]$ for which the convex combination*

$$c^{(\alpha(\pi))} := \alpha(\pi)c^{(T(\pi)-1)} + (1-\alpha(\pi))c^{(T(\pi))}.$$

is an equilibrium allocation, and the value of this allocation is $\bar{F}$.

The allocation $c^{(\alpha(\pi))}$ postpones risk sharing for $T(\pi) - 1$ periods and then provides intermediate risk sharing in future periods.

Proof. We first observe that if $c^{(T-1)} \in \mathcal{C}(\bar{F})$ and

$$\bar{F} \leq \pi[(1-\beta^{T-1})V^A + \beta^{T-1}V^*(\bar{F})] + (1-\pi)V^A,$$

then $c^{(T)} \in \mathcal{C}(\bar{F})$. This holds because

$$(1-\beta)u(y) + \beta W^0(c^{(T-1)}) \geq (1-\beta)u(y) + \beta\bar{F}.$$

Denote by $T(\pi)$ the unique value of T satisfying

$$\pi[(1-\beta^T)V^A + \beta^T V^*(\bar{F})] + (1-\pi)V^A < \bar{F} \leq \pi[(1-\beta^{T-1})V^A + \beta^{T-1}V^*(\bar{F})] + (1-\pi)V^A.$$

Since utility is concave, $c^{(\alpha)} \in \mathcal{C}(\bar{F})$ for all $\alpha \in [0,1]$. Moreover, $W^0(c^{(\alpha)})$ is continuous function of α, with

$$\pi W^0(c^{(0)}) + (1-\pi)V^A < \bar{F} \leq \pi W^0(c^{(1)}) + (1-\pi)V^A.$$

Thus, there exists $\alpha(\pi)$ such that

$$\pi W^0(c^{(\alpha(\pi))}) + (1-\pi)V^A = \bar{F},$$

and so $c^{(\alpha(\pi))}$ is an optimal consumption allocation for $\pi > \bar{\pi}$.

□

The next lemma describes an equilibrium that burns resources.

Lemma 2 *Define the consumption allocation $c^{[\alpha]}$ as follows:*

$$c^{[\alpha]}(y^t) = \begin{cases} \bar{\mathbb{c}}(y^t) & \text{if } y^t \neq \ell^t, \\ \alpha\bar{\mathbb{c}} + (1-\alpha)c_\ell(\bar{F}), & \text{if } y^t = \ell^t. \end{cases}$$

There exists $\alpha(\pi)$ for which $c^{[\alpha(\pi)]}$ is an equilibrium allocation whose value is $\bar{F}$.

Note that the consumption allocation $c^{[\alpha]}$ only differs from $\bar{\mathbb{c}}$ at histories ℓ^t. Moreover, since $\bar{\mathbb{c}}(\ell^t) = c_\ell(\bar{F})$ in finite time (Lemma B.9), $c^{[\alpha]}(y^t) = \bar{\mathbb{c}}(y^t)$ for $t \geq L$.

Proof. Since $\bar{\mathbb{c}}(\ell^t) \geq c_\ell(\bar{F})$ for all t, $c^{[\alpha]} \in \mathcal{C}(\bar{F})$.

Since the payoff to any agent receiving the income h in the initial period is the same as under $\bar{\mathbb{c}}$ and the h incentive feasibility constraint is always binding, the h payoff is given by

$$(1-\beta)u(h) + \beta\bar{F}.$$

The consumption $c_\ell(\bar{F})$ is determined by the requirement that the ℓ incentive feasibility constraint is binding, and so the payoff to any agent receiving the income ℓ in the initial period under $c^{[0]}$ is

$$(1-\beta)u(\ell) + \beta\bar{F}.$$

This implies

$$W^0(c^{[0]}) < \bar{F},$$

so that

$$\pi W^0(c^{[0]}) + (1-\pi)V^A < \bar{F} < \pi W^0(c^{[1]}) + (1-\pi)V^A.$$

Thus, there exists $\alpha(\pi)$ such that

$$\pi W^0(c^{[\alpha(\pi)]}) + (1-\pi)V^A = \bar{F},$$

and so $c^{[\alpha(\pi)]}$ is an optimal consumption allocation for $\pi > \bar{\pi}$.

□

7 Numerical Examples and Comparative Statics

In this section we illustrate the computation of equilibrium allocations, and present results for an illustrative set of examples to convey the qualitative properties of the equilibrium. Throughout we assume the CRRA period utility function (18). This functional form implies that equation (9) in Proposition 5 characterizing equilibrium allocations can be written as

$$\forall y^t, \mathbb{c}(y^t\ell) > c_\ell(F) \implies \frac{\mathbb{c}(y^t)^{-\gamma}}{\mathbb{c}(y^t\ell)^{-\gamma}} = \delta_{t+1}.$$

for some $\delta_{t+1} < 1$. Since $\delta_{t+1} < 1$, and defining $g_{t+1} := (\delta_{t+1})^{1/\gamma} < 1$,

$$\forall y^t, \mathbb{c}(y^t\ell) > c_\ell(F) \implies \mathbb{c}(y^t\ell) = g_{t+1}\mathbb{c}(y^t).$$

Thus, equilibrium allocations have the form of a sequence of consumption ladders (as in Definition 5), where the period t-ladder is determined by an initial consumption after the high income $y = h$ realization, $\mathbb{c}_t(h)$, and then a decreasing sequence of lower consumptions $g_{t+1}\mathbb{c}_t(h), g_{t+1}g_{t+2}\mathbb{c}_t(h), \ldots$, until the lower bound $c_\ell(F)$ is reached (after $L-1$ realizations of ℓ). Note that a stationary ladder has $g_t = g_{t+1} = g$. When $F = \bar{F}$ (equivalently, $\pi = \bar{\pi}$), the equilibrium allocation converges to the unique stationary ladder satisfying h-incentive feasibility, so that $g_t \to \bar{g} := \beta^{1/\gamma}$ (Proposition 7).

With these observations based on our theoretical results in hand, the computation of an equilibrium with associated outside option $F \in (V^A, \bar{F}]$ (and thus for social capital π associated with that outside option) proceeds as follows. The algorithm first computes a stationary consumption ladder and associated consumption decay rate g that satisfies the

h-incentive-feasibility constraint associated with F with equality (as well as the resource constraint and the ℓ-incentive-feasibility constraint with equality for those at the very bottom of the ladder).[6] The algorithm then determines the dynamic equilibrium consumption allocation under the assumption that it converges into the stationary ladder in finite (but potentially long) time. The key distinction between an arbitrary outside option F and $\bar{F}$ is that at the latter we know

1. the stationary decay rate $\bar{g}$,
2. the associated stationary ladder is unique, and
3. the dynamic equilibrium consumption allocation converges to the stationary ladder asymptotically.

We therefore focus on the $\bar{F}$ case in what follows.[7]

Figure 3 plots the dynamics of the equilibrium consumption allocation when the period utility function is logarithmic, $u(c) = \log(c)$, incomes are $(\ell, h) = (0.75, 1.25)$, and the discount factor is $\beta = .9$. Social capital is $\pi = \bar{\pi} = 0.41$ so that the value of the outside option is given by $F = \bar{F}$. Table 1 provides additional summary statistics for the allocation in this parameterization, as well as for alternative values of (β, γ) to display the comparative statics of the model with respect to its preference parameters (the values of $\bar{F}$ and $\bar{\pi}$ changes with (β, γ)).

From Figure 3 we observe that as the transition unfolds, consumption spreads out over time, and eventually converges to the stationary ladder, which for this parameterization has five consumption steps. Consumption insurance worsens over time but remains positive: for high income individuals the outside option is binding, but they consume substantially less than their income h (indicated by the upper dashed line) and thus provide insurance to low-income individuals. Initially low income individuals consume significantly more than their income (lower dashed line), and also significantly more than implied by a binding outside option, $c_\ell(\bar{F})$. Over time those with continuously low income see their consumption drift down until the outside option becomes binding, and $c = c_\ell(\bar{F})$. In the example this occurs in period four of the transition.

The equilibrium allocation can generate high initial consumption insurance because the allocation does not inherit any implicit promises to past high income types. As time evolves, the consumption level of $c(\ell^t)$ declines as the burden of efficient smoothing of consumption to

[6] While there may be multiple stationary ladders satisfying the three constraints, each ladder is associated with a distinct value of g. Moreover, it is inefficient to converge to a stationary ladder with $g < \beta^{1/\gamma} = \bar{g}$ (Lemma D.1).

[7] The details of the computational procedure are described in Appendix D

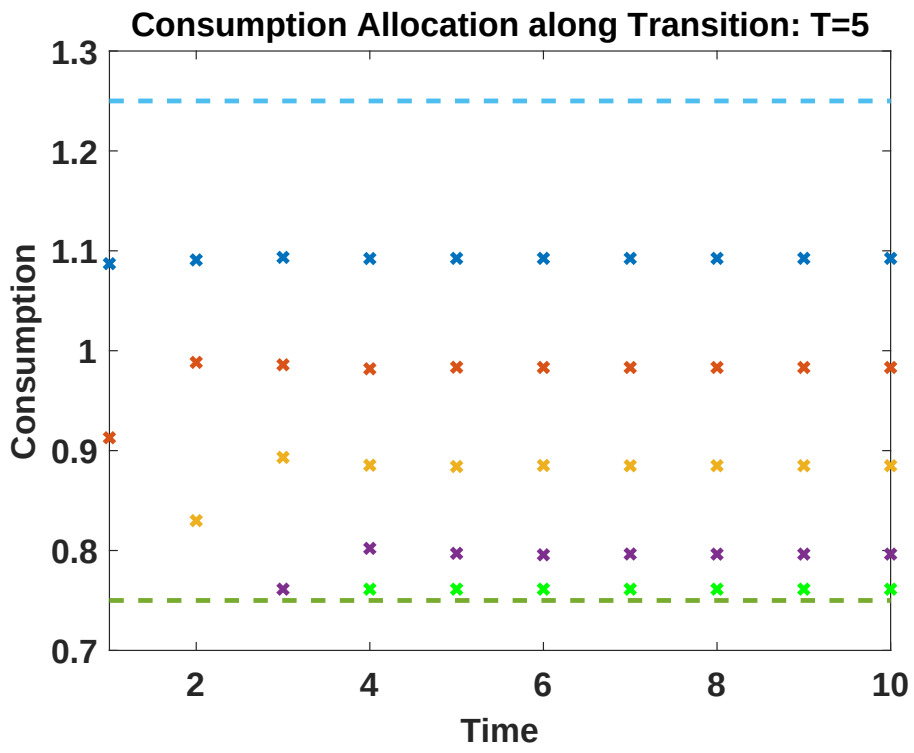

Figure 3: The consumption allocation along the transition, with $\ell = 0.75, h = 1.25, \beta = 0.9, \gamma = 1$.

past high income types makes consumption scarcer. The allocation also becomes statically inefficient since individuals with the same current income receive different consumption levels. Finally, the figure shows that although we do not force convergence to the stationary ladder until period 10 (the last period of the blending phase) in this example, effectively allocations have converged to the stationary ladder by period four of the transition. Expanding the length of the transition yields utility gains that are indistinguishable from zero. Thus, although theoretically convergence to the stationary ladder is only asymptotic, our numerical examples suggest that in practice convergence happens rapidly.

Table 1 contains summary statistics of equilibrium allocations along the transition for alternative parameterizations of the model. Focusing first on the benchmark case in the first column, we observe that the consumption allocation a coalition can implement improves significantly (worth 0.94% of consumption) on the outside option, by providing insurance to initially poor individuals, but also needs to leave significant insurance opportunities unexploited (worth 0.63% of consumption relative to full insurance). Insurance gets worse over time as expected period utility falls and consumption dispersion rises over time.[8] Comparing across parameterizations, as households become more patient (higher β) and more risk-averse

[8]We only display the first two periods, relative to the stationary ladder.

Table 1: Summary Statistics of the Transition

	$\gamma = 1$		$\gamma = 2$	
Statistic	$\beta = 0.9$	$\beta = 0.95$	$\beta = 0.9$	$\beta = 0.95$
$V^{FB}/V(\bar{F})$ in %	0.63%	0.22%	0.45%	0.12%
$V(\bar{F})/\bar{F}$ in %	0.94%	0.71%	1.24%	0.80%
$\bar{\pi}$	0.41	0.66	0.69	0.85
$c_\ell(\bar{F})$	0.761	0.767	0.776	0.782
c_h	1.092	1.049	1.050	1.025
Steps	5	8	7	12
$\frac{EU(c_1)}{EU(c_\infty)}$ in %	0.28%	0.11%	0.25%	0.07%
$\frac{EU(c_2)}{EU(c_\infty)}$ in %	0.05%	0.05%	0.11%	0.03%
$Var(c_\infty)$	0.01	0.004	0.004	0.001
$\frac{Var(c_1)}{Var(c_\infty)}$	0.62	0.55	0.55	0.52
$\frac{Var(c_2)}{Var(c_\infty)}$	0.94	0.80	0.81	0.77

Notes: Ratios of (lifetime) utilities are converted into consumption equivalent variation and give the percentage increase in consumption (uniform across all states or histories) required to equalize period (or lifetime) utility across the two alternatives. The first two lines measure the welfare loss from imperfect consumption insurance relative to full insurance, and the welfare gain of coalition allocations relative to the outside option. The second panel provides summary statistics of the stationary ladder, and the third and forth panels show how expected utility and consumption insurance declines over time.

(higher γ), the equilibrium allocations get closer to full insurance, but the gains from coalition risk sharing relative to the outside option become smaller. The stationary ladder has more steps and the support of the consumption distribution tightens. We also observe that increased patience (higher β), elevates the gains of coalition risk sharing (compared to the outside option) mostly through an improvement of the stationary ladder. An increase in risk aversion (larger γ), in contrast, leads to better risk sharing both because of an improved stationary ladder and longer initial insurance and thus slower convergence to the ladder.

8 Related Literature

There is a large literature using limited contract enforcement to rationalize incomplete insurance arrangements and imperfect risk allocations. In financial markets, Kehoe and Levine (1993) and Alvarez and Jermann (2000) characterize consumption allocations under limited commitment within a general equilibrium framework. In labor economics, Harris and Holmström (1982) and Thomas and Worrall (1988) study efficient long term-contracts between employers and employees under limited commitment. Kocherlakota (1996) models two-party

risk-sharing arrangements as a repeated game and Krueger and Perri (2011) extends this literature to a risk sharing economy with as a continuum of households exactly of the form studied in this paper. All these classic papers share our focus on self-enforcing arrangements, but take the outside option as exogenously given, and equal to the autarkic allocation. *Given* this outside option, the qualitative properties of the equilibrium allocation in this work and our paper are similar: high-income individuals receive high consumption to avoid defection, and consumption drifts down with low-income realizations until it hits a lower bound.

Building on these classic papers, a literature emerged that endogenizes the outside option by assuming that one side of the arrangement, typically a financial intermediary, has long term commitment, as in Krueger and Uhlig (2006).[9] A strand of the sovereign debt literature also considers self-enforcing simple debt contracts because sovereigns cannot commit to repay, see for example Eaton and Gersovitz (1981) and Bulow and Rogoff (1989). There is also a related literature which endogenizies the outside option by assuming that private non-contingent intertemporal trades can be enforced and examines how this impacts on insurance (see, for example, Allen, 1985) and government taxation (see Farhi et al. (2009)).[10]

Most papers on risk-sharing consider only unilateral deviations of individuals from the risk sharing arrangement, thereby explicitly or implicitly limiting the extent of insurance these individuals can obtain after deviating. An exception to this is Genicot and Ray (2003), who study the formation and stability to *joint deviations* of risk sharing coalitions in economies with finite populations. Bold and Broer (2018) estimate their model on Indian village data and find that stable risk sharing coalitions are typically small, and that the resulting consumption allocations accord better with the data than those generated by the standard limited commitment model with exogenous outside option, assumed again to be autarky.

In the finite population world of Genicot and Ray (2003), coalitions must be stable against deviations of smaller sub-coalitions of the original group, and the main purpose of the paper is to determine endogenously the *size* of stable coalitions. Since larger coalitions are more prone to successful deviation an optimal size of the original coalition emerges. This result stems from the assumption that the deviating coalition can only make an arrangement with the original coalition members, while in the formation of the original coalition, all members of the population could be considered as potential members. We share with this paper and with Bold and Broer (2018) the basic notion that risk-sharing coalitions must be immune to

[9] Phelan (1995) also endogenizes the outside option, and makes assumptions on the timing of the model that implies full commitment for one period. In his paper private information about income limits consumption insurance in his model.

[10] Within the context of a model with incomplete information, Cole and Kocherlakota (2001) endogenize the outside option, assuming hidden storage. Ábrahám and Laczó (2017) also analyze a limited commitment model with a private storage technology.

not only unilateral deviations by an individual, but to coalitional deviations. In contrast to Genicot and Ray (2003), the insurance capabilities of deviating continuum coalitions are no worse than in the original coalition, and so if any coalition is stable, the grand coalition is.

Genicot and Ray (2003) is closely related to the more abstract game theoretic literature on coalition deviations pioneered by Bernheim et al. (1987) and Greenberg (1990) (and extended/unified by Kahn and Mookherjee (1992, 1995) to infinite games and to adverse selection insurance economies in which agents have private information). This abstract literature shares with Genicot and Ray (2003) the assumption that coalition formation is "easy." Our view that coalition formation is "hard" and so the required nontrivial belief coordination on future behavior does not always arise is consistent with the recent work on strategic uncertainty. The classical insurance environment is an example of a repeated game (indeed, Mailath and Samuelson (2006) exposit Kocherlakota (1996) as an example of a repeated game). Repeated games have many equilibria, with efficient equilibria sustained by delicate intertemporal incentives that require belief coordination about continuation play. There is little work on the robustness of equilibria to shocks to belief coordination. The literature on private monitoring is motivated by a concern that agents may not have sufficient common knowledge about past histories in order to coordinate continuation play. Surprisingly, this literature has showed that efficiency is still consistent with equilibrium (see, for example, Hörner and Olszewski (2006)). Importantly, the constructions in this literature maintain the assumption of belief coordination on future behavior.

The large literature on global games (surveyed in Morris and Shin (2003)) and higher order beliefs (Rubinstein (1989) and Weinstein and Yildiz (2007)) is directly concerned with the difficulties in coordinating behavior. The global games literature has primarily focused on simple (and often static) coordination games. This permits a direct modeling of the source of difficulty in coordinating behavior. In order to focus on the implications of the belief coordination frictions, we model the friction in a reduced form manner. The literature on higher order beliefs in repeated games is thin (Weinstein and Yildiz (2013, 2016)), but the few results there confirm our intuition that coordination is no easier in a repeated game context than in a static one.

The use of utility and money burning at the beginning of the allocation is reminiscent of some efficiency wage (Shapiro and Stiglitz, 1984, MacLeod and Malcomson, 1989) and some gift-exchange and related models (Carmichael and MacLeod, 1997, Kranton, 1996b,a, Ghosh and Ray, 1996). In particular, the idea that if it is too easy to start a new relationship (worker-firm, principal-agency, partnership, etc) after opportunistic behavior (shirking for example), then it is impossible to deter opportunistic behavior. In order to deter deviations, it is therefore necessary to impose some form of tax or friction (such as delays in joining a

new firm, involuntary unemployment, or engaging in inefficient actions in the beginning of the new relationship, exchange of inefficient gifts).

9 Conclusion

Pareto improving activities that require trust because they have an intertemporal exchange aspect, are a major component of societies both primitive and modern. We have focused on one of the most basic of these, insurance, and parameterized social capital or trust in a very simple fashion through the ability to coordinate beliefs on the optimal outcome. Despite its starkness, the model paints a surprisingly rich picture of potential societies.

In societies with a low level of social capital many Pareto improving activities are not undertaken. However, when they are undertaken, they tend to be extensively exploited and a high level of insurance is achieved. This occurs because the low level of social capital implies that the outside options of the participants are low, and hence these outside options do not distort the allocations within the coalition. This implies that the insurance level is both intertemporally and statically efficient. Ex ante welfare rises linearly with social capital.

At medium levels of social capital, more Pareto improving activities are undertaken, but as a result, the outside options of participants are higher. These higher outside options bind increasingly on allocations attainable with higher social capital. As a result, efficient allocations must reward higher income. This leads to allocations being statically inefficient, but not intertemporally inefficient modulo the outside options. When this occurs, overall utility is declining over time as accumulated past promises of rewards for higher income build up. Consequently, the extent to which Pareto improving activities are exploited is declining, and so too is the payoff from forming a coalition. Ex ante welfare is still increasing in social capital, but now at a decreasing rate.

At high levels of social capital, most if not all of the Pareto improving activities are undertaken. However, the level of the outside options that can be attained has hit the maximum ($\bar{F}$ in our model). As a result, some form of waste must occur in the early phases of these arrangements because too high an ex ante payoff would carry the seeds of its own destruction and hence would not be internally incentive feasible. While the long-run arrangement is independent of social capital, the extent of short run money- or utility-burning is increasing in social capital, and the ex ante welfare is flat. Our arrangements feature explicit waste but this waste is front-loaded. In the long-run they are intertemporally, but not statically efficient, just as with medium levels of social capital.

References

Ábrahám, Árpád and Sarolta Laczó (2017), "Efficient risk sharing with limited commitment and storage." *Review of Economic Studies*, 85, 1389–1424. 27

Allen, Franklin (1985), "Repeated principal-agent relationships with lending and borrowing." *Economics Letters*, 17, 27–31. 27

Alvarez, Fernando and Urban J. Jermann (2000), "Efficiency, equilibrium, and asset pricing with risk of default." *Econometrica*, 68, 775–797. 17, 26

Bernheim, B. Douglas, Bezalel Peleg, and Michael D. Whinston (1987), "Coalition proof Nash equilibria i: Concepts." *Journal of Economic Theory*, 42, 1–12. 28

Bold, Tessa and Tobias Broer (2018), "Risk-sharing in village economies revisited." Stockholm University. 27

Bulow, Jeremy and Kenneth Rogoff (1989), "Sovereign debt: Is to forgive to forget?" *American Economic Review*, 79, 43–50. 27

Carmichael, H. Lorne and W. Bentley MacLeod (1997), "Gift giving and the evolution of cooperation." *International Economic Review*, 38, 485–509. 28

Chien, YiLi and Hanno Lustig (2009), "The market price of aggregate risk and the wealth distribution." *Review of Financial Studies*, 23, 1596–1650. 17

Cole, Harold L. and Narayana R. Kocherlakota (2001), "Efficient allocations with hidden income and hidden storage." *Review of Economic Studies*, 68, 523–542. 27

Eaton, Jonathan and Mark Gersovitz (1981), "Debt with potential repudiation: Theoretical and empirical analysis." *The Review of Economic Studies*, 48, 289–309. 27

Farhi, Emmanuel, Mikhail Golosov, and Aleh Tsyvinski (2009), "A theory of liquidity and regulation of financial intermediation." *Review of Economic Studies*, 76, 973–992. 27

Farrell, Joseph and Eric Maskin (1989), "Renegotiation in repeated games." *Games and Economic Behavior*, 1, 327–360. 6

Genicot, Garance and Debraj Ray (2003), "Group formation in risk-sharing arrangements." *Review of Economic Studies*, 70, 87–113. 27, 28

Ghosh, Parikshit and Debraj Ray (1996), "Cooperation in community interaction without information flows." *Review of Economic Studies*, 63, 491–519. 28

Greenberg, Joseph (1990), *The Theory of Social Situations: An Alternative Game-Theoretic Approach*. Cambridge University Press. 6, 28

Harris, Milton and Bengt Holmström (1982), "A Theory of Wage Dynamics." *The Review of Economic Studies*, 49, 315–333. 26

Hörner, Johannes and Wojciech Olszewski (2006), "The folk theorem for games with private almost-perfect monitoring." *Econometrica*, 74, 1499–1544. 28

Kahn, Charles M and Dilip Mookherjee (1992), "The good, the bad, and the ugly: Coalition proof equilibrium in infinite games." *Games and Economic Behavior*, 4, 101–121. 28

Kahn, Charles M. and Dilip Mookherjee (1995), "Coalition proof equilibrium in an adverse selection insurance economy." *Journal of Economic Theory*, 66, 113–138. 28

Kehoe, Timothy J. and David K. Levine (1993), "Debt-constrained asset markets." *Review of Economic Studies*, 60, 865–888. 17, 26

Kocherlakota, Narayana R. (1996), "Implications of efficient risk sharing without commitment." *Review of Economic Studies*, 63, 595–609. 26, 28

Kranton, Rachel E. (1996a), "The formation of cooperative relationships." *Journal of Law, Economics, and Organization*, 12, 214–233. 28

Kranton, Rachel E. (1996b), "Reciprocal exchange: A self-sustaining system." *American Economic Review*, 86, 830–851. 28

Krueger, Dirk and Fabrizio Perri (2011), "Public versus private risk sharing." *Journal of Economic Theory*, 146, 920–956. 7, 27

Krueger, Dirk and Harald Uhlig (2006), "Competitive risk sharing contracts with one-sided commitment." *Journal of Monetary Economics*, 53, 1661–1691. 27

MacLeod, W. Bentley and James M. Malcomson (1989), "Implicit contracts, incentive compatibility, and involuntary unemployment." *Econometrica*, 57, 447–480. 28

Mailath, George J. and Larry Samuelson (2006), *Repeated Games and Reputations: Long-Run Relationships*. Oxford University Press, New York, NY. 28

Morris, Stephen and Hyun Song Shin (2003), "Global games: Theory and applications." In *Advances in Economics and Econometrics (Proceedings of the Eighth World Congress of the Econometric Society)* (M. Dewatripont, L. Hansen, and S. Turnovsky, eds.), Cambridge University Press. 28

Phelan, Christopher (1995), "Repeated Moral Hazard and One-Sided Commitment." *Journal of Economic Theory*, 66, 488–506. 27

Rubinstein, Ariel (1989), "The electronic mail game: Strategic behavior under almost common knowledge." *American Economic Review*, 79, 385–391. 28

Shapiro, Carl and Joseph Stiglitz (1984), "Equilibrium unemployment as a worker discipline device." *American Economic Review*, 74, 433–444. 28

Thomas, Jonathan and Tim Worrall (1988), "Self-Enforcing Wage Contracts." *The Review of Economic Studies*, 55, 541–554. 26

von Neumann, John and Oskar Morgenstern (1944), *Theory of games and economic behavior*, 1st edition. Princeton University Press. 2nd edn 1947, 3rd edn 1953. 6

Weinstein, Jonathan and Muhamet Yildiz (2007), "A structure theorem for rationalizability with application to robust predictions of refinements." *Econometrica*, 75, 365–400. 28

Weinstein, Jonathan and Muhamet Yildiz (2013), "Robust predictions in infinite-horizon games—an unrefinable folk theorem." *Review of Economic Studies*, 80, 365–394. 28

Weinstein, Jonathan and Muhamet Yildiz (2016), "Reputation without commitment in finitely repeated games." *Theoretical Economics*, 11, 157–185. 28

Appendices

A Proofs for Section 5

We begin with a preliminary result.

Lemma A.1

1. *$\mathcal{C}(F') \supset \mathcal{C}(F'')$ for $F' < F''$, and so $\mathcal{C}(F) \neq \varnothing$ for all $F \leq \bar{F}$.*

2. *$\mathcal{C}(F)$ is closed and convex for all $F \leq \bar{F}$.*

3. *$\mathcal{C}$ is a continuous correspondence at all $F \leq \bar{F}$ (at $\bar{F}$, the continuity is from the left).*

Proof.

1. This is immediate.

2. This is also immediate.

3. Since $\mathcal{C}$ is a decreasing correspondence in F, we need only show upper hemicontinuity from the left and lower hemicontinuity from the right. Upper hemicontinuity is immediate, since all the constraints are closed. Turning to lower hemicontinuity, we need to show that if $c \in \mathcal{C}(F)$ and $(F_k)_k$ is a sequence with $F_k \searrow F$, then there exists $c_k \in \mathcal{C}(F_k)$ with $c_k \to c$. Fix $c^\dagger \in \mathcal{C}(\bar{F})$. We now verify that for all k, there exists $\alpha_k \in [0,1]$ such that $\alpha_k c^\dagger + (1-\alpha_k)c \in \mathcal{C}(F_k)$ and $\alpha_k \to 0$.

   Fix k, and let $\alpha_k = (F_k - F)/(\bar{F} - F_k) > 0$. Then,

$$\begin{aligned} W(y^t, \alpha_k c^\dagger + (1-\alpha_k)c) &\geq \alpha_k W(y^t, c^\dagger) + (1-\alpha_k) W(y^t, c) \\ &\geq (1-\beta)u(y_t) + \alpha_k \beta \bar{F} + (1-\alpha_k)\beta F \\ &= (1-\beta)u(y_t) + \beta F_k, \end{aligned}$$

   and so incentive feasibility (6) is satisfied. Since (1) is trivially satisfied, we are done.

□

Proof of Proposition 4

1. Since, for ε small, the allocation in Example 1 is internally-incentive feasible for $\pi = 1$ and provides partial insurance, $\mathcal{C}(F) \neq \varnothing$ for some $F > V^A$, and so $\bar{F} > V^A$. This also shows that $\mathbb{V}(V^A) > V^A$.

2. Suppose $(F_k) \nearrow \bar{F}$ is a sequence satisfying $\mathcal{C}(F_k) \neq \varnothing$. Since the space of consumption allocations is sequentially compact (being the countable product of sequentially-compact spaces), we can assume there is a convergent sequence $(c_k)_k$, with $c_k \in \mathcal{C}(F_k)$ and limit c_∞. Since all the constraints defining $\mathcal{C}$ are closed (and continuous in F), the limit also satisfies these constraints (including (6) at $F = \bar{F}$), and so $c_\infty \in \mathcal{C}(\bar{F})$, and $\mathcal{C}(\bar{F}) \neq \varnothing$.

3. The continuity of $\mathbb{V}$ follows from the continuity of $\mathcal{C}$ (Lemma A.1) and the maximum theorem.

4. We now verify that for all $F \in (V^A, \bar{F}]$, $F < \mathbb{V}(F)$,[11] Example 1 having already demonstrated that $V^A < \mathbb{V}(V^A)$. Since $\mathbb{V}$ is continuous, it is enough to show that there is no $F \in (V^A, \bar{F}]$ for which $F = \mathbb{V}(F)$. The proof is by contradiction, so suppose there is such an F. Then $\mathbb{V}(F)$ is a fixed point of $T(\cdot; 1)$. Note first that $F = \mathbb{V}(F) = V^A$ is impossible (since if $F = V^A$, we immediately have $\mathbb{V}(F) > V^A$), and so $F = \mathbb{V}(F) > V^A$.

   For $\hat{y}^t \in Y^t$, denote by $c|_{\hat{y}^t}$ the consumption allocation

$$c|_{\hat{y}^t}(y^\tau) = c(\hat{y}^t y^\tau) \quad \forall y^\tau \in Y^\tau.$$

   Then,

$$W(y^t, c) = (1-\beta)u(c(y_t)) + \beta W^0(c|_{y^t}).$$

   Suppose $\mathbb{c}$ achieves $\mathbb{V}(F)$. Since $F = \mathbb{V}(F) = \max_{c \in \mathcal{C}(F)} W^0(c)$, and $\mathbb{c}|_{y^t} \in \mathcal{C}(F)$, we have $F \geq W^0(\mathbb{c}|_{y^t h})$ for all $y^t \in Y^t$, and so (using (6) in the first line)

$$\begin{aligned}(1-\beta)u(\mathbb{c}(y^t h)) + \beta F &\geq (1-\beta)u(h) + \beta F \\ \implies u(\mathbb{c}(y^t h)) \geq u(h) &\implies \mathbb{c}(y^t h) \geq h.\end{aligned}$$

   But then $F \leq V^A$, a contradiction.

5. The function $p : [V^A, \bar{F}] \to [0, \bar{\pi}]$ defined by

$$p(F) := \frac{F - V^A}{\mathbb{V}(F) - V^A}$$

   is strictly increasing, continuous, and onto (since $\mathbb{V}(V^A) > V^A$). It is straightforward to verify that for $\pi \in (0, \bar{\pi}]$, the fixed point is given by $\pi V^*(p^{-1}(\pi)) + (1-\pi)V^A$. It is

[11] If $\beta > \beta^{FB}$, the conclusion is immediate for $F \leq F^{FB}$, since $\mathbb{V}(F) = V^{FB}$ for such F.

also immediate that this is increasing in π and

$$\bar{F} = \mathcal{T}(\bar{F}; \bar{\pi}).$$

6. Finally, for $\pi > \bar{\pi}$, the required F is strictly greater than $\bar{F}$, implying that the constraint set is empty, and so there is no fixed point.

□

Lemma A.2 *If $\beta > \beta^{FB}$, then $\bar{F} > F^{FB}$.*

Proof. Recall the allocation c_ζ defined in (5):

$$c_\zeta(y^t) = \begin{cases} h - \zeta, & y_t = h, \\ \ell + \zeta, & y_t = \ell. \end{cases}$$

We now argue that there exists $\xi > 0$ such that for all $F \in (F^{FB}, F^{FB} + \xi]$, for

$$\zeta = \zeta^{FB} - 2\beta(F - F^{FB})/u'(\bar{y}), \tag{A.1}$$

we have $c_\zeta \in \mathcal{C}^\dagger(F)$, and so $\bar{F} > F^{FB}$.

Because marginal changes in ζ from ζ^{FB} result in only second losses to ex ante payoffs $(W^0(c_\zeta))$, we have

$$\begin{aligned} \frac{\partial W(h, c^{FB})}{\partial \zeta} &= -(1-\beta)u'(\bar{y}), \\ \text{and} \qquad \frac{\partial W(\ell, c^{FB})}{\partial \zeta} &= (1-\beta)u'(\bar{y}). \end{aligned}$$

Since

$$W(h, c^{FB}) = (1-\beta)u(h) + \beta F^{FB},$$

we have

$$W(\ell, c^{FB}) = W(h, c^{FB}) > (1-\beta)u(\ell) + \beta F^{FB}. \tag{A.2}$$

Since

$$\frac{\partial W(h, c^{FB})}{\partial \zeta} = -(1-\beta)u'(\bar{y}),$$

we have

$$\begin{aligned} W(h, c_\zeta) =& W(h, c^{FB}) - u'(\bar{y})(\zeta - \zeta^{FB}) + o((\zeta - \zeta^{FB})^2) \\ =& W(h, c^{FB}) + (\zeta^{FB} - \zeta)[u'(\bar{y}) + o((\zeta - \zeta^{FB})^2)/(\zeta - \zeta^{FB})]. \end{aligned}$$

For $\zeta^{FB} - \zeta < \xi'$, where $\xi' > 0$ is a sufficiently small constant, the magnitude of the last term is less than $u'(\bar{y})/2$, and so

$$W(h, c_\zeta) > W(h, c^{FB}) + (\zeta^{FB} - \zeta)u'(\bar{y})/2.$$

For $F = F^{FB} + (\zeta^{FB} - \zeta)u'(\bar{y})/(2\beta)$ (this is just a rewriting of (A.1)), we then have

$$W(h, c_\zeta) > (1 - \beta)u(h) + \beta F.$$

Moreover, there is $\xi'' > 0$, such that for $\zeta^{FB} - \zeta < \xi''$, the strict inequality on the ℓ-incentive constraint (A.2) is preserved:

$$W(\ell, c_\zeta) > (1 - \beta)u(\ell) + \beta F.$$

Setting

$$\xi := \min\{\xi', \xi''\}u'(\bar{y})/(2\beta)$$

completes the proof. □

Lemma A.3

$$\lim_{\beta \searrow u'(h)/u'(\ell)} \bar{\pi}(\beta) = 0.$$

Proof. Consider the allocation

$$c_\varepsilon^+(y^t) = \begin{cases} h - \varepsilon, & y_t = h, \\ \ell + 2\varepsilon, & y_{t-1} = h, y_t = \ell, \\ \ell, & y_{t-1} = y_t = \ell, \\ \ell + \varepsilon, & y_1 = \ell. \end{cases}$$

Then,

$$\begin{aligned} W(h, c_\varepsilon^+) &= (1-\beta)u(h-\varepsilon) + \tfrac{\beta}{2}(W(h, c_\varepsilon^+) + W(h\ell, c_\varepsilon^+)), \\ W(h\ell, c_\varepsilon^+) &= (1-\beta)u(\ell+2\varepsilon) + \tfrac{\beta}{2}(W(h, c_\varepsilon^+) + W(\ell\ell, c_\varepsilon^+)), \\ W(\ell\ell, c_\varepsilon^+) &= (1-\beta)u(\ell) + \tfrac{\beta}{2}(W(h, c_\varepsilon^+) + W(\ell\ell, c_\varepsilon^+)), \\ \text{and} \qquad W(\ell, c_\varepsilon^+) &= (1-\beta)u(\ell+\varepsilon) + \tfrac{\beta}{2}(W(h, c_\varepsilon^+) + W(\ell\ell, c_\varepsilon^+)). \end{aligned}$$

Hence,

$$W(\ell\ell, c_\varepsilon^+) = \frac{1}{2-\beta}\{2(1-\beta)u(\ell) + \beta W(h, c_\varepsilon^+)\}$$

and so

$$\begin{aligned} W(h\ell, c_\varepsilon^+) &= (1-\beta)u(\ell+2\varepsilon) + \frac{\beta}{2}\left\{W(h, c_\varepsilon^+) + \frac{1}{2-\beta}\{2(1-\beta)u(\ell) + \beta W(h, c_\varepsilon^+)\}\right\} \\ &= (1-\beta)\left\{u(\ell+2\varepsilon) + \frac{\beta}{2-\beta}u(\ell)\right\} + \frac{\beta}{(2-\beta)}W(h, c_\varepsilon^+). \end{aligned}$$

Thus,

$$\begin{aligned} W(h, c_\varepsilon^+) &= (1-\beta)u(h-\varepsilon) + \frac{\beta}{2}\left\{W(h, c_\varepsilon^+) + (1-\beta)\left\{u(\ell+2\varepsilon) + \frac{\beta}{2-\beta}u(\ell)\right\} + \frac{\beta}{(2-\beta)}W(h, c_\varepsilon^+)\right\} \\ &= (1-\beta)\left\{u(h-\varepsilon) + \frac{\beta}{2}u(\ell+2\varepsilon) + \frac{\beta^2}{2(2-\beta)}u(\ell)\right\} + \frac{\beta}{(2-\beta)}W(h, c_\varepsilon^+), \end{aligned}$$

which implies

$$2(1-\beta)W(h, c_\varepsilon^+) = (1-\beta)(2-\beta)\left\{u(h-\varepsilon) + \frac{\beta}{2}u(\ell+2\varepsilon) + \frac{\beta^2}{2(2-\beta)}u(\ell)\right\},$$

that is,

$$W(h, c_\varepsilon^+) = \frac{(2-\beta)}{2}u(h-\varepsilon) + \frac{\beta(2-\beta)}{4}u(\ell+2\varepsilon) + \frac{\beta^2}{4}u(\ell). \tag{A.3}$$

Note that $W(h, c_\varepsilon^+)$ is a strictly concave function of ε.

In order for c_ε^+ to be internally incentive feasible, we need

$$W(h, c_\varepsilon^+) \geq (1-\beta)u(h) + \pi\frac{\beta}{2}\left[W(h, c_\varepsilon^+) + W(\ell, c_\varepsilon^+)\right] + (1-\pi)\beta V^A.$$

For $\varepsilon = 0$, the above inequality holds with equality. Differentiating both sides with respect to ε and evaluating at $\varepsilon = 0$, yields the necessary condition

$$\frac{\partial}{\partial \varepsilon} W(h, c_\varepsilon^+) \geq \pi \frac{\beta}{2} \left[\frac{\partial}{\partial \varepsilon} W(h, c_\varepsilon^+) + \frac{\partial}{\partial \varepsilon} W(\ell, c_\varepsilon^+) \right],$$

that is,

$$(2 - \pi\beta) \frac{\partial}{\partial \varepsilon} W(h, c_\varepsilon^+) \geq \pi\beta \frac{\partial}{\partial \varepsilon} W(\ell, c_\varepsilon^+).$$

It is easy to verify that $\partial W(\ell, c_\varepsilon^+)/\partial \varepsilon > (1-\beta)u'(\ell)$, and so, differentiating (A.3), a necessary condition is

$$(2 - \pi\beta)(2 - \beta)(\beta u'(\ell) - u'(h)) \geq \pi\beta(1 - \beta)u'(\ell).$$

We now consider the implications of taking limits as β falls from above to $u'(h)/u'(\ell)$, so that we always have $\beta u'(\ell) > u'(h)$. For each $\beta > u'(h)/u'(\ell)$, set $\varepsilon > 0$ so that

$$u'(h - \varepsilon) = \beta u'(\ell + 2\varepsilon) \tag{A.4}$$

(this is the value of ε that maximizes $W(h, c_\varepsilon^+)$). Denote by $\hat{c}(\beta)$ the allocation c_ε^+ with ε satisfying (A.4). Note that if for any $\beta > u'(h)/u'(\ell)$, the allocation c_ε^+ is internally-incentive feasible, then so is $\hat{c}(\beta)$.

Denote by $\hat{\pi}(\beta)$ the maximum value of π for which the allocation $\hat{c}(\beta)$ is internally-incentive feasible. We then have

$$(2 - \hat{\pi}(\beta)\beta)(2 - \beta)(\beta u'(\ell) - u'(h)) \geq \hat{\pi}(\beta)\beta(1 - \beta)u'(\ell).$$

Hence, as $\beta u'(\ell) - u'(h) \searrow 0$, $\hat{\pi}(\beta) \to 0$.

Set (where β is included as an argument to make clear the dependence on β)

$$F(\beta) := \hat{\pi}(\beta) W^0(\hat{c}(\beta)) + (1 - \hat{\pi}(\beta)) V^A \leq \bar{F}(\beta).$$

Since $\hat{c}(\beta)$ is internally-incentive feasible,

$$\hat{c}(\beta) \in \mathcal{C}(F(\beta)) \neq \varnothing,$$

and so

$$\mathbb{V}(F(\beta)) \geq W^0(\hat{c}(\beta)),$$

and finally

$$\hat{\pi}(\beta) \geq \pi(F(\beta)) := \frac{F(\beta) - V^A}{\mathbb{V}(F(\beta)) - V^A} \geq \bar{\pi}(\beta).$$

Since $\hat{\pi}(\beta) \to 0$, $\bar{\pi}(\beta) \to 0$, as $\beta \searrow u'(h)/u'(\ell)$. □

B Proof of Proposition 5

We assume throughout this section that $F > F^{FB}$ and $\beta u'(\ell) > u'(h)$.

Lemma B.1 *There exists $\delta_{t+1} < 1$ such that if incentive feasibility does not bind at y^{t+1}, then*

$$\frac{u'(\mathbb{c}(y^t))}{u'(\mathbb{c}(y^{t+1}))} = \delta_{t+1}$$

and so

$$\mathbb{c}(y^t) > \mathbb{c}(y^{t+1}).$$

Proof. We first argue that if incentive feasibility does not bind at $\tilde{y}^{t+1}$, then for all $\hat{y}^{t+1}$

$$\frac{u'(\mathbb{c}(\tilde{y}^t))}{u'(\mathbb{c}(\tilde{y}^{t+1}))} \leq \frac{u'(\mathbb{c}(\hat{y}^t))}{u'(\mathbb{c}(\hat{y}^{t+1}))}. \tag{B.1}$$

Suppose not, so that (B.1) holds with a strict inequality in the reverse direction.

Define a new allocation $c^\dagger$ by setting

$$c^\dagger(y^\tau) = \begin{cases} \mathbb{c}(\tilde{y}^t) + \varepsilon, & y^\tau = \tilde{y}^t \\ \mathbb{c}(\hat{y}^t) - \varepsilon, & y^\tau = \hat{y}^t \\ \mathbb{c}(\tilde{y}^{t+1}) - \eta, & y^\tau = \tilde{y}^{t+1} \\ \mathbb{c}(\hat{y}^{t+1}) + \eta, & y^\tau = \hat{y}^{t+1} \\ \mathbb{c}(y^\tau), & \text{otherwise.} \end{cases}$$

Since $\Pr(\hat{y}^t) = \Pr(\tilde{y}^t)$ and $\Pr(\hat{y}^{t+1}) = \Pr(\tilde{y}^{t+1})$, the allocation $c^\dagger$ is resource feasible.

Choose $\eta = \eta(\varepsilon)$ so that

$$u(\mathbb{c}(\tilde{y}^t) + \varepsilon) + \frac{\beta}{2} u(\mathbb{c}(\tilde{y}^{t+1}) - \eta(\varepsilon)) = u(\mathbb{c}(\tilde{y}^t)) + \frac{\beta}{2} u(\mathbb{c}(\tilde{y}^{t+1}))$$

ensures that incentive feasibility is satisfied along the sequence $\tilde{y}^t$. For small η, it is also satisfied at $\tilde{y}^{t+1}$.

Differentiating with respect to ε and evaluating at $\varepsilon = 0$, we get

$$\eta'(0) = \frac{2u'(\mathbb{c}(\tilde{y}^t))}{\beta u'(\mathbb{c}(\tilde{y}^{t+1}))}.$$

At $\varepsilon = 0$, the derivative of

$$u(\mathbb{c}(\hat{y}^t) - \varepsilon) + \frac{\beta}{2} u(\mathbb{c}(\hat{y}^{t+1}) + \eta(\varepsilon))$$

is

$$\begin{aligned} -u'(\mathbb{c}(\hat{y}^t)) + \frac{\beta}{2} u'(\mathbb{c}(\hat{y}^{t+1}))\eta'(0) &= -u'(\mathbb{c}(\hat{y}^t)) + \frac{\beta}{2} u'(\mathbb{c}(\hat{y}^{t+1})) \frac{2u'(\mathbb{c}(\tilde{y}^t))}{\beta u'(\mathbb{c}(\tilde{y}^{t+1}))} \\ &= u'(\mathbb{c}(\hat{y}^{t+1})) \left\{ -\frac{u'(\mathbb{c}(\hat{y}^t))}{u'(\mathbb{c}(\hat{y}^{t+1}))} + \frac{u'(\mathbb{c}(\tilde{y}^t))}{u'(\mathbb{c}(\tilde{y}^{t+1}))} \right\} \\ &> 0. \end{aligned}$$

This implies that the values of the agents with histories $\hat{y}^t$ and $\hat{y}^{t+1}$ have increased, and so the ex ante value of $c^\dagger$ must exceed $\mathbb{c}$, contradicting the optimality of $\mathbb{c}$.

Hence, (B.1) must hold as written. If incentive feasibility also does not bind at $\hat{y}^{t+1}$, then the weak inequality holds as an equality.

We now argue that if incentive feasibility does not hold at $\tilde{y}^{t+1}$, then

$$\frac{u'(\mathbb{c}(\tilde{y}^t))}{u'(\mathbb{c}(\tilde{y}^{t+1}))} < 1.$$

If not, then for all histories,

$$\frac{u'(\mathbb{c}(y^t))}{u'(\mathbb{c}(y^{t+1}))} \geq 1.$$

But this implies for all y^{t+1},

$$\mathbb{c}(y^t) \leq \mathbb{c}(y^{t+1}).$$

This is only consistent with resource feasibility if $\mathbb{c}(y^t) = \mathbb{c}(y^{t+1})$, which implies $\mathbb{c}$ is the first best allocation. But $F > F^{FB}$ precludes the first best allocation as an equilibrium.

□

Lemma B.2 *At the optimal allocation $\mathbb{c}$, if the incentive constraint binds at $\tilde{y}^t$ and $\hat{y}^t$ with $\tilde{y}_t = \hat{y}_t$, then*

$$\mathbb{c}(\tilde{y}^t) = \mathbb{c}(\hat{y}^t).$$

Proof. Suppose not. then the incentive constraint binds at two histories $\tilde{y}^t$ and $\hat{y}^t$ with $\tilde{y}_t = \hat{y}_t$, and

$$\mathbb{c}(\tilde{y}^t) \neq \mathbb{c}(\hat{y}^t).$$

Define a new consumption allocation $c^\dagger$ as follows:

$$c^\dagger(y^\tau) = \begin{cases} \frac{1}{2}\mathbb{c}(\hat{y}^t{}_t y^\tau) + \frac{1}{2}\mathbb{c}(\tilde{y}^t{}_t y^\tau), & \tau \geq t, {}^t y^\tau = \tilde{y}^t, \hat{y}^t, \\ \mathbb{c}(y^\tau), & \text{otherwise,} \end{cases}$$

where ${}_t y^\tau$ is the last $\tau - t$ periods of the income history y^t (so that $y^\tau = {}^t y^\tau {}_t y^\tau$)). Since $\Pr(\tilde{y}^t) = \Pr(\hat{y}^t)$, $c^\dagger$ satisfies (1).

Moreover, the incentive constraints are satisfied at all histories:

1. For $\tau < t$, since the incentive constraints bind at two histories $\tilde{y}^t$ and $\hat{y}^t$ with $\tilde{y}_t = \hat{y}_t$, $W(\tilde{y}^t, \mathbb{c}) = W(\hat{y}^t, \mathbb{c})$, and so $W(y^t, c^\dagger) \geq W(y^t, \mathbb{c})$ for all y^t (with equality holding for $y^t \notin \{\tilde{y}^t, \hat{y}^t\}$). Hence,

$$\begin{aligned} W(y^\tau, c^\dagger) &= (1-\beta)u(\mathbb{c}(y^\tau)) + (1-\beta)\sum_{r=1}^{t-\tau-1} \beta^r \sum_{y^r} \Pr(y^r) u(\mathbb{c}(y^\tau y^r)) \\ &\qquad + \beta^{t-\tau} \sum_{y^t} \Pr(y^t) W(y^t, c^\dagger) \\ &\geq W(y^\tau, \mathbb{c}). \end{aligned}$$

2. For $\tau \geq t$, the concavity of u implies

$$W(y^t, c^\dagger) \geq \min\{W(\hat{y}^t{}_t y^\tau, \mathbb{c}), W(\tilde{y}^t{}_t y^\tau, \mathbb{c})\} \geq W^F(y_\tau).$$

Finally, concavity implies $W^0(c^\dagger) > W^0(\mathbb{c})$, which is impossible, since $\mathbb{c}$ is by assumption optimal. □

Lemma B.3 *In the optimal allocation, incentive feasibility binds at all y^t for which $y_t = h$, and so for all y^{t-1},*

$$W(y^{t-1}h, \mathbb{c}) = W^F(h) := (1-\beta)u(h) + \beta F.$$

Proof. Since $F > F^{FB}$,

$$(1-\beta)u(\bar{y}) + \beta V^{FB} < W^F(h),$$

and so

$$(1-\beta)u(\bar{y}) + \beta\mathbb{V}(F) < W^F(h).$$

Thus, incentive feasibility at $\hat{y}^{t-1}h$ requires $\mathbb{c}(\hat{y}^{t-1}h) > \bar{y}$. Suppose

$$W(\hat{y}^{t-1}h, \mathbb{c}) > W^F(h).$$

Define

$$c^\varepsilon(y^\tau) = \begin{cases} \mathbb{c}(\hat{y}^{t-1}h) - \varepsilon, & y^\tau = \hat{y}^{t-1}h, \\ \mathbb{c}(\hat{y}^{t-1}\ell) + \varepsilon, & y^\tau = \hat{y}^{t-1}\ell, \\ \mathbb{c}(y^t), & \text{otherwise.} \end{cases}$$

Since h and ℓ are equally likely, c^ε satisfies resource feasibility. For sufficiently small $\varepsilon > 0$, c^ε satisfies inetrnal-incentive feasibility, and so we have a contradiction (since c^ε has higher ex ante utility than $\mathbb{c}$). Thus, the incentive constraint binds at all $\hat{y}^t$ for which $\hat{y}_t = h$. □

Lemma B.4 *For all $\tilde{y}^{t-1}, \hat{y}^{t-1}$,*

$$\mathbb{c}(\tilde{y}^{t-1}) \geq \mathbb{c}(\hat{y}^{t-1}) \implies \mathbb{c}(\tilde{y}^{t-1}y) \geq \mathbb{c}(\hat{y}^{t-1}y) \text{ and } W(\tilde{y}^{t-1}\ell, \mathbb{c}) \geq W(\hat{y}^{t-1}\ell, \mathbb{c}).$$

Proof. Lemmas B.2 and B.3, imply

$$\mathbb{c}(\tilde{y}^{t-1}h) = \mathbb{c}(\hat{y}^{t-1}h) \qquad \forall \tilde{y}^{t-1}, \hat{y}^{t-1}.$$

Suppose now, en route to a contradiction that there are two histories $\tilde{y}^{t-1}$ and $\hat{y}^{t-1}$ such that

$$\mathbb{c}(\tilde{y}^{t-1}) \geq \mathbb{c}(\hat{y}^{t-1}) \text{ and } \mathbb{c}(\tilde{y}^{t-1}\ell) < \mathbb{c}(\hat{y}^{t-1}\ell).$$

The idea is to construct a dominating consumption allocation by moving consumption from the relatively high-consumption histories to the low-consumption histories. For any small $\varepsilon > 0$, define $\eta(\varepsilon)$ as the value η solving

$$u(\mathbb{c}(\tilde{y}^{t-1}) - \eta) + \tfrac{\beta}{2}u(\mathbb{c}(\tilde{y}^{t-1}\ell) + \varepsilon) = u(\mathbb{c}(\tilde{y}^{t-1})) + \tfrac{\beta}{2}u(\mathbb{c}(\tilde{y}^{t-1}\ell)),$$

and define a new consumption allocation as

$$c^{\varepsilon}(y^{\tau}) = \begin{cases} \mathbb{c}(y^{\tau}) - \eta(\varepsilon), & y^{\tau} = \tilde{y}^{t-1}, \\ \mathbb{c}(y^{\tau}) + \eta(\varepsilon), & y^{\tau} = \hat{y}^{t-1}, \\ \mathbb{c}(y^{\tau}) + \varepsilon, & y^{\tau} = \tilde{y}^{t-1}\ell, \\ \mathbb{c}(y^{\tau}) - \varepsilon, & y^{\tau} = \hat{y}^{t-1}\ell, \\ \mathbb{c}(y^{\tau}), & \text{otherwise.} \end{cases}$$

From the concavity of u, $u'(\mathbb{c}(\tilde{y}^{t-1})) \leq u'(\mathbb{c}(\hat{y}^{t-1}))$ and

$$\xi := u'(\mathbb{c}(\tilde{y}^{t-1}\ell)) - u'(\mathbb{c}(\hat{y}^{t-1}\ell)) > 0.$$

Moreover, the function η is $\mathcal{C}^1$ with $\eta'(0) > 0$. Then we have (where each function o_j, for $j = 1, \ldots, 4$ satisfies $o_j(\varepsilon)/\varepsilon \to 0$ as $\varepsilon \to 0$),

$$\begin{aligned} \tfrac{\beta}{2}\{u(\mathbb{c}(\hat{y}^{t-1}\ell)) - u(\mathbb{c}(\hat{y}^{t-1}\ell) - \varepsilon)\} &= \tfrac{\beta}{2}\{u'(\mathbb{c}(\hat{y}^{t-1}\ell))\varepsilon + o_1(\varepsilon)\} \\ &= \tfrac{\beta}{2}\{u'(\mathbb{c}(\tilde{y}^{t-1}\ell))\varepsilon - \xi\varepsilon + o_1(\varepsilon)\} \\ &= \tfrac{\beta}{2}\{u(\mathbb{c}(\tilde{y}^{t-1}\ell) + \varepsilon) - u(\mathbb{c}(\tilde{y}^{t-1}\ell)) - \xi\varepsilon\} + o_2(\varepsilon) \\ &= u(\mathbb{c}(\tilde{y}^{t-1})) - u(\mathbb{c}(\tilde{y}^{t-1}) - \eta(\varepsilon)) - \tfrac{\beta}{2}\xi\varepsilon + o_2(\varepsilon) \\ &= u'(\mathbb{c}(\tilde{y}^{t-1}))\eta(\varepsilon) - \tfrac{\beta}{2}\xi\varepsilon + o_3(\varepsilon) \\ &\leq u'(\mathbb{c}(\hat{y}^{t-1}))\eta(\varepsilon) - \tfrac{\beta}{2}\xi\varepsilon + o_3(\varepsilon) \\ &= u(\mathbb{c}(\hat{y}^{t-1}) + \eta(\varepsilon)) - u(\mathbb{c}(\hat{y}^{t-1})) - \tfrac{\beta}{2}\xi\varepsilon + o_4(\varepsilon). \end{aligned}$$

Rearranging,

$$u(\mathbb{c}(\hat{y}^{t-1})) + \frac{\beta}{2}u(\mathbb{c}(\hat{y}^{t-1}\ell)) + \tfrac{\beta}{2}\xi\varepsilon \leq u(\mathbb{c}(\hat{y}^{t-1}) + \eta(\varepsilon)) + \frac{\beta}{2}u(\mathbb{c}(\hat{y}^{t-1}\ell) - \varepsilon) + o_4(\varepsilon),$$

and so, if $\varepsilon > 0$ is sufficiently small that

$$|o_4(\varepsilon)| < \tfrac{\beta}{2}\xi\varepsilon,$$

we have

$$u(\mathbb{c}(\hat{y}^{t-1})) + \frac{\beta}{2}u(\mathbb{c}(\hat{y}^{t-1}\ell)) < u(\mathbb{c}(\hat{y}^{t-1}) + \eta(\varepsilon)) + \frac{\beta}{2}u(\mathbb{c}(\hat{y}^{t-1}\ell) - \varepsilon).$$

Since $\mathbb{c}(y^{\tau}) \leq c^{\varepsilon}(y^{\tau})$ for all y^{τ}, with a strict inequality on one positive-measure history, $\mathbb{c}$ cannot have been optimal.

The inequality on continuation values then immediately follows from the following calculation: For any y^t, denote by $y^t\ell^k$ the history formed by adding k periods of ℓ after y^t (so that $y^t\ell^0 = y^t$). Then,

$$\begin{aligned} W(y^t, \mathbb{c}) &= (1-\beta)u(\mathbb{c}(y^t)) + \tfrac{\beta}{2}\{W^F(h) + W(y^t\ell, \mathbb{c})\} \\ &= (1-\beta)\sum_{k=0}^{\infty}(\tfrac{\beta}{2})^k u(\mathbb{c}(y^t\ell^k)) + \tfrac{\beta}{2-\beta}W^F(h). \end{aligned} \tag{B.2}$$

□

Lemma B.5 *If incentive feasibility at $y^t\ell$ is binding, then for all $\hat{y}^t$,*

$$\mathbb{c}(y^t\ell) \leq \mathbb{c}(\hat{y}^t\ell).$$

Proof. Suppose

$$\mathbb{c}(y^t\ell) > \mathbb{c}(\hat{y}^t\ell).$$

Then, from Lemma B.4,

$$\begin{aligned} u(\mathbb{c}(y^t\ell)) + \tfrac{\beta}{2}\{W^F(h) + W(y^t\ell\ell, \mathbb{c})\} &> \mathbb{c}(\hat{y}^t\ell) + \tfrac{\beta}{2}\{W^F(h) + W(\hat{y}^t\ell\ell, \mathbb{c})\} \\ &\geq W^F(\ell), \end{aligned}$$

which is impossible if incentive feasibility binds at $y^t\ell$. □

Lemma B.6 *Suppose incentive feasibility binds at some $y^{t-1}\ell$ in an optimal allocation. Then incentive feasibility binds at $y^{t-1}\ell\ell$.*

Proof. Suppose incentive feasibility binds at $y^{t-1}\ell$ but not at $y^{t-1}\ell^2$. Then

$$\begin{aligned} u(\mathbb{c}(y^{t-1}\ell)) + \tfrac{\beta}{2}\{W^F(h) + W(y^{t-1}\ell^2, \mathbb{c})\} &= W^F(\ell), \\ W(y^{t-1}\ell^2, \mathbb{c}) = u(\mathbb{c}(y^{t-1}\ell^2)) + \tfrac{\beta}{2}\{W^F(h) + W(y^{t-1}\ell^3, \mathbb{c})\} &> W^F(\ell), \end{aligned}$$

and (because the last incentive feasibility constraint is not binding)

$$\mathbb{c}(y^{t-1}\ell) > \mathbb{c}(y^{t-1}\ell^2).$$

Since

$$u(\mathbb{c}(y^{t-1}\ell)) > u(\mathbb{c}(y^{t-1}\ell^2)),$$

we therefore have (because incentive feasibility is binding at $y^{t-1}\ell$)

$$W(y^{t-1}\ell^3, \mathbb{c}) > W(y^{t-1}\ell^2, \mathbb{c}) > W^F(\ell), \tag{B.3}$$

and incentive feasibility is also not binding at $y^{t-1}\ell^3$. This implies

$$\mathbb{c}(y^{t-1}\ell) > \mathbb{c}(y^{t-1}\ell^3),$$

and so

$$W(y^{t-1}\ell^4, \mathbb{c}) > W(y^{t-1}\ell^2, \mathbb{c}) > W^F(\ell).$$

Repeated applications of this argument shows that incentive feasibility is not binding for any history $y^{t-1}\ell^r$, $r \geq 2$, and so $(\mathbb{c}(y^{t-1}\ell^r))_{r\geq 1}$ is a monotonically declining sequence. Hence, from (B.2), so is $(W(y^{t-1}\ell^r, \mathbb{c}))_{r\geq 1}$. But this contradicts (B.3). □

Lemma B.7 *If incentive feasibility binds at $y^t\ell$, then*

$$\mathbb{c}(y^t\ell) = c_\ell,$$

where $c_\ell > \ell$ is the unique consumption satisfying

$$u(c_\ell) = u(\ell) + \beta(F - V^A) > u(\ell).$$

Note that c_ℓ is an increasing function of F, so that for $F > F^{FB}$ (i.e., for $\pi > \pi^{FB}$) but arbitrarily close, c_ℓ is bounded away from $\overline{y}$.

Proof. Since incentive binds at $y^t\ell$ (and so at $y^t y_\ell^2$), we have

$$(1-\beta)u(\mathbb{c}(y^t\ell)) + \tfrac{\beta}{2}\{W^F(h) + W^F(\ell)\} = W^F(\ell).$$

Rearranging and dividing by $(1-\beta)$ yields

$$u(\mathbb{c}(y^t\ell)) = (1 - \tfrac{\beta}{2})u(\ell) - \tfrac{\beta}{2}u(h) + \beta F,$$

which is the displayed equation (recall that $V^A = Eu(y)$). □

Lemma B.8 *Incentive feasibility does not bind in the initial period at ℓ, nor after any history of the form $y^t h\ell$.*

Proof. If the incentive feasibility binds in the initial period, then

$$\begin{aligned}\mathbb{V}(F) &= \tfrac{1}{2}(1-\beta)\{u(h)+u(\ell)\}+\beta F\\ &= (1-\beta)V^A+\beta F\\ &\implies \mathbb{V}(F)<F,\end{aligned}$$

contradicting Proposition 4.

Suppose now incentive feasibility binds after a history of the form $y^t h\ell$. Since incentive feasibility always binds after any realization of h, we have

$$\begin{aligned}(1-\beta)u(h)+\beta F &= (1-\beta)u(\mathbb{c}(y^t h))+\beta\{(1-\beta)V^A+\beta F\}\\ \implies (1-\beta)u(h) &= (1-\beta)u(\mathbb{c}(y^t h))-\beta(1-\beta)(F-V^A)\\ \implies u(\mathbb{c}(y^t h)) &= u(h)+\beta(F-V^A)\\ \implies \mathbb{c}(y^t h) &> h,\end{aligned}$$

which is ruled out by resource feasibility and $\mathbb{c}(y^{t+1}) \geq c_\ell > \ell$. □

Lemma B.9 *Suppose $\pi > \pi^{FB}$. In the optimal allocation, there exists L such that the incentive constraint binds at any history of the form $y^t\ell^L$.*

Proof. Lemma B.4 implies that optimal consumption in any period is determined by the number of ℓ realizations after the last h realization. From Lemma B.6, once the ℓ incentive constraint binds, it continues to bind after each subsequent ℓ realization.

We need to prove that the number of ℓ realizations before the ℓ incentive constrain binds is bounded as we vary the period in which h is realized.

We prove by contradiction: Suppose there is a subsequence $(t_n)_n$ of periods with the property that if h is realized in that period, the number of ℓ realizations before the ℓ-incentive constraint binds goes to ∞. Without loss of generality, assume there are at least n realizations of ℓ after h in period t_n before the ℓ-incentive constraint binds.

For each t_n, $(\mathbb{c}(y^{t_n-1}h\ell^k))_{k=1}^n$ is monotonically declining in k, is bounded above by h, and below by ℓ. Hence, for all $\varepsilon>0$, the number of periods in the interval $\{t_n+1, t_n+2, \ldots, t_n+n\}$ for which $\delta_t < 1-\varepsilon$ is less than $(h-\ell)/\varepsilon$, a bound independent of n, the number of periods in the interval. That is, asymptotically, the fraction of periods in which $\delta_t \in (1-\varepsilon, 1)$ converges to one. This implies that for all T, there exists t such that $\delta_\tau \in (1-\varepsilon,1)$ for all $\tau = t, t+1, \ldots, t+T$.

By choosing ε sufficiently small, for such t, resource feasibility implies $\mathbb{c}(y^{t+T}h)$ is arbitrarily close to $\overline{y}$ (since for many k, $\mathbb{c}(y^{t+T-k}h\ell^k)$ will be close to $\mathbb{c}(y^{t+T-k}h)$, which is no smaller than $\overline{y}$).

Since $\pi > \pi^{FB}$ and so $F > F^{FB}$, the incentive constraint for $y^t h$ is violated. □

C Proofs for Section 6.2

Proof of Proposition 6 The outside option F only affects $\mathbb{V}^*(h;F)$ through c_ℓ (which is a strictly increasing function of F, and so makes the constraints strictly more demanding). Hence, $\mathbb{V}^*(h;F)$ is strictly decreasing function of F. It remains to prove that $V^*(h;F) = W^F(h)$ at $\bar{F}$.

If $\mathbb{c}_*$ is the stationary ladder yielding $\mathbb{V}^*(h;F)$, define an allocation as follows

$$c^F(y^t) := \begin{cases} \mathbb{c}_*(h), & y_t = h, \\ \mathbb{c}_*(h\ell^\tau), & y^t = y^{t-\tau-1}h\ell^\tau, \\ \hat{c}(\ell^t), & y^t = \ell^t, \end{cases} \tag{C.1}$$

where

$$\Pr(\ell^t)\hat{c}(\ell^t) = \overline{y} - \sum\nolimits_{y^t \neq \ell^t} \Pr(y^t)c^F(y^t).$$

By construction, c^F satisfies resource feasibility, and incentive feasibility for any history ending in a realization of ℓ (since $\mathbb{c}_*$ satisfies (14), $\hat{c}(\ell^t) \geq c_\ell$).

If $\mathbb{V}^*(h;F) \geq W^F(h)$, then the incentive constraint on $y^t h$ is satisfied under c^F for all y^t. Hence, $c^F \in \mathcal{C}(F)$, and so $F \leq \bar{F}$.

Suppose $\mathbb{V}^*(h;F) > W^F(h)$. A marginal increase in F preserves the inequality and so $F < \bar{F}$.

Finally, we prove that if $F \leq \bar{F}$, then $\mathbb{V}^*(h;F) \geq W^F(h)$. We do this by proving that if $\mathcal{C}(F)$ is nonempty, then it implies a feasible *stationary* ladder of the form (C.1). We construct the stationary ladder by time averaging over histories that have the same number of y realizations after an h realization.

Suppose $\mathbb{c} \in \mathcal{C}(F)$ is optimal. From Lemmas B.8 and B.9, there exists $L \geq 2$ such that for all $\tau \geq L$, the ℓ incentive constraint binds at any history of the form $y^t\ell^\tau$. Moreover,

$$\mathbb{c}(y^t\ell^\tau) = c_\ell, \qquad \forall \tau \geq L. \tag{C.2}$$

For $K \geq 1$, define the ladder $(c_\tau^K)_{k=0}^{T^\ddagger}$ (recall that $\Pr(y^t) = 2^{-t}$):

$$c_k^K = \begin{cases} \frac{1}{K+1} \sum_{t=L}^{L+K} \sum_{y^{t-k-1}} (\frac{1}{2})^{t-k-1} \mathbb{c}(y^{t-k-1} h \ell^k), & 0 \leq k < L \\ \frac{1}{K+1} \sum_{t=L}^{L+K} \sum_{y^{t-L}} (\frac{1}{2})^{t-T^\dagger} \mathbb{c}(y^{t-L} \ell^L), & k = L \end{cases}$$
$$= \begin{cases} \frac{1}{K+1} \sum_{t=L}^{L+K} \sum_{y^{t-k-1}} (\frac{1}{2})^{t-k-1} \mathbb{c}(y^{t-k-1} h \ell^k), & 0 \leq k < L \\ c_\ell, & k = L. \end{cases}$$

We claim that $(c_k^K)_k$ satisfies (14) (where we set $c_k^K = c_\ell$ for $k > L$):

$$\sum_{k=0}^{\infty} (\tfrac{1}{2})^{k+1} c_k^K = \sum_{k=0}^{L-1} (\tfrac{1}{2})^{k+1} \frac{1}{K+1} \sum_{t=L}^{L+K} \sum_{y^{t-k-1}} (\tfrac{1}{2})^{t-k-1} \mathbb{c}(y^{t-k-1} h \ell^k) + (\tfrac{1}{2})^L c_\ell$$
$$= \frac{1}{K+1} \sum_{t=L}^{L+K} \left\{ \sum_{k=0}^{L-1} (\tfrac{1}{2})^{k+1} \sum_{y^{t-k-1}} (\tfrac{1}{2})^{t-k-1} \mathbb{c}(y^{t-k-1} h \ell^k) + (\tfrac{1}{2})^L c_\ell \right\}$$
$$= \frac{1}{K+1} \sum_{t=L}^{L+K} \left\{ \sum_{k=0}^{L-1} (\tfrac{1}{2})^{k+1} \sum_{y^{t-k-1}} (\tfrac{1}{2})^{t-k-1} \mathbb{c}(y^{t-k-1} h \ell^k) + (\tfrac{1}{2})^L c_\ell \right\}$$
$$= \frac{1}{K+1} \sum_{t=L}^{L+K} \left\{ \sum_{y^t \neq y^{t-L} \ell^L} \Pr(y^t) \mathbb{c}(y^t) + (\tfrac{1}{2})^L c_\ell \right\}.$$

But (C.2) implies

$$(\tfrac{1}{2})^L c_\ell = \sum_{y^t = y^{t-L} \ell^L} \Pr(y^t) \mathbb{c}(y^t)$$

and so

$$\sum_{k=0}^{\infty} (\tfrac{1}{2})^{k+1} c_k^K = \frac{1}{K+1} \sum_{t=L}^{L+K} E\mathbb{c}(y^t) \leq \overline{y}.$$

Since $\mathbb{c}(y^t - 1, \ell) \geq c_\ell$, it is immediate that c_k^K satisfies (15).

We now show that accumulation points c^* of $(c_k^K)_k)_K$ satisfy $W(h, c^*) \geq W^F(h)$. We do this by considering time averages of continuation utilities for large K.

Since $\mathbb{c}$ is incentive feasible, for all y^t,

$$W^F(h) \leq W(y^t h, \mathbb{c}).$$

Consequently,

$$\begin{aligned}
W^F(h) &\leq \frac{1}{K+1}\sum_{t=L}^{K+L}\sum_{y^{t-1}}(\tfrac{1}{2})^{t-1}W(y^{t-1}h,\mathbb{c}) \\
&= \frac{1}{K+1}\sum_{t=L}^{K+L}\sum_{y^{t-1}}(\tfrac{1}{2})^{t-1}(1-\beta)\sum_{k=0}^{\infty}(\tfrac{\beta}{2})^k u(\mathbb{c}(y^{t-1}h\ell^k)) + \tfrac{\beta}{2-\beta}W^F(h) \\
&= (1-\beta)\sum_{k=0}^{\infty}(\tfrac{\beta}{2})^k\frac{1}{K+1}\sum_{t=L}^{K+L}\sum_{y^{t-1}}(\tfrac{1}{2})^{t-1}u(\mathbb{c}(y^{t-1}h\ell^k)) + \tfrac{\beta}{2-\beta}W^F(h).
\end{aligned}$$

Now

$$\frac{1}{K+1}\sum_{t=L}^{K+L}\sum_{y^{t-1}}(\tfrac{1}{2})^{t-1}u(\mathbb{c}(y^{t-1}h\ell^k)) \leq u\left(\tfrac{1}{K+1}\sum_{t=L}^{K+L}\sum_{y^{t-1}}(\tfrac{1}{2})^{t-1}\mathbb{c}(y^{t-1}h\ell^k)\right),$$

and so

$$W^F(h) \leq (1-\beta)\sum_{k=0}^{\infty}(\tfrac{\beta}{2})^k u\left(\frac{1}{K+1}\sum_{t=L}^{K+L}\sum_{y^{t-1}}(\tfrac{1}{2})^{t-1}\mathbb{c}(y^{t-1}h\ell^k)\right) + \tfrac{\beta}{2-\beta}W^F(h). \tag{C.3}$$

If the arguments of the utility function were c_k^K (which they are not), we would be done, since then the expression on the right hand side is simply $W(h, c^K)$.

However, we are almost done, since the discrepancy can be made arbitrarily small. For $k < L < K$, we have

$$\begin{aligned}
&c_k^K - \frac{1}{K+1}\sum_{t=L}^{K+L}\sum_{y^{t-1}}(\tfrac{1}{2})^{t-1}\mathbb{c}(y^{t-1}h\ell^k) \\
&\quad= \frac{1}{K+1}\sum_{t=L}^{L+K}\sum_{y^{t-k-1}}(\tfrac{1}{2})^{t-k-1}\mathbb{c}(y^{t-k-1}h\ell^k) - \frac{1}{K+1}\sum_{t=L}^{K+L}\sum_{y^{t-1}}(\tfrac{1}{2})^{t-1}\mathbb{c}(y^{t-1}h\ell^k) \\
&\quad= \frac{1}{K+1}\sum_{t=L}^{L+K}\sum_{y^{t-k-1}}(\tfrac{1}{2})^{t-k-1}\mathbb{c}(y^{t-k-1}h\ell^k) - \frac{1}{K+1}\sum_{t=L+k}^{K+L+k}\sum_{y^{t-k-1}}(\tfrac{1}{2})^{t-k-1}\mathbb{c}(y^{t-k-1}h\ell^k) \\
&\quad= \frac{1}{K+1}\sum_{t=L}^{L+k-1}\sum_{y^{t-k-1}}(\tfrac{1}{2})^{t-k-1}\mathbb{c}(y^{t-k-1}h\ell^k) - \frac{1}{K+1}\sum_{t=L+K+1}^{K+L+k}\sum_{y^{t-k-1}}(\tfrac{1}{2})^{t-k-1}\mathbb{c}(y^{t-k-1}h\ell^k).
\end{aligned}$$

The magnitude of this expression is bounded above by $kh/(K+1)$. An identical argument shows that we have the bound of $Lh/(K+1)$ for the divergence of c_L^K.

Using (C.2), we can rewrite (C.3) as

$$W^F(h) \leq (1-\beta)\sum_{k=0}^{L}(\tfrac{\beta}{2})^k u\left(\frac{1}{K+1}\sum_{t=L}^{K+L}\sum_{y^{t-1}}(\tfrac{1}{2})^{t-1}\mathbb{c}(y^{t-1}h\ell^k)\right) + \frac{2(1-\beta)}{(2-\beta)}\left(\frac{\beta}{2}\right)^{L+1}u(c_\ell) + \frac{\beta}{2-\beta}W^F(h). \quad \text{(C.4)}$$

For all $\varepsilon > 0$, there exists K^ε such that if $K > K^\varepsilon$, for all $k = 0, \ldots, L$ the upper bound of $Lh/(K+1)$ on consumption divergences is sufficiently small that the right side of (C.4) is within ε of $W(h, c^K)$, implying

$$W^F(h) \leq W(h, c^K) + \varepsilon.$$

Since $(c_k^K)_k \in [0,h]^L$, a closed and bounded set, the sequence $((c_k^K)_k)_K$ has a convergent subsequence with limit $(c_k^*)_k$. Moreover,

$$W^F(h) \leq W(h, c^*).$$

□

The proof of Proposition 7 is broken into several lemmas.

Lemma C.1 *Suppose $F = \bar{F}$. The equilibrium allocation $\mathbb{c}$ converges to the unique solution to problem* (13)*, $\bar{\mathbb{c}}_*$, that is,*

$$\lim_{t\to\infty}\mathbb{c}_t(h\ell^k) = \bar{\mathbb{c}}_*(h\ell^k) \quad \textit{for any } k < L$$

and

$$\mathbb{c}_t(\ell^L) = \bar{\mathbb{c}}_*(\ell^L) = c_\ell(\bar{F}).$$

Proof. Resource feasibility in period t is

$$\sum_{k=0}^{L-1} 2^{-k+1}\mathbb{c}_t(h\ell^k) + 2^{-L}c_\ell(F) \leq \overline{y}. \quad \text{(C.5)}$$

We denote the period-t consumption ladder by (since by Proposition 5, we can ignore the history before the last realization of h)

$$(\mathbb{c}_{t+k}(h\ell^k)_{k=0}^{L-1}, c_\ell).$$

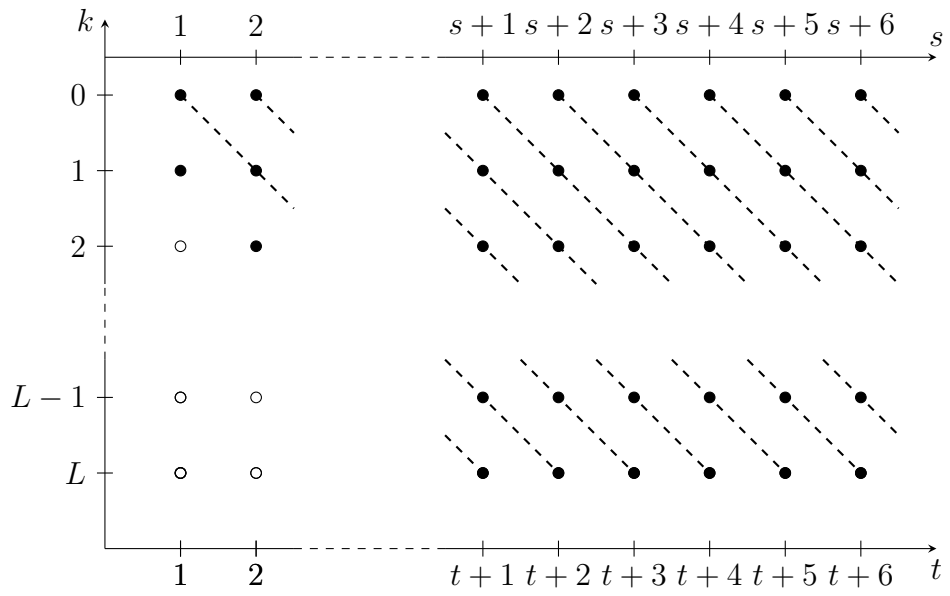

Figure C.1: Illustrating the reindexing for the proof (the reindexing omits $k = L$, since $c_t(\ell^L) = c_\ell(F)$ is determined by Lemma B.7). The ladder-s resource constraints sum over the diagonal dashed lines, while the period-t resource constraints sum vertically. Since there is at most one realization of ℓ in any history in period 1, k can only equal 0 or 1; similarly, in period 2, $k \leq 2$.

Summing inequality (C.5) over periods $1, \ldots, T$, and rearranging to sum over ladders rather than periods (see Figure C.1), for $T \geq L$, gives

$$\begin{aligned}
0 &\geq \sum_{t=1}^{T} \sum_{k=0}^{L-1} 2^{-k+1} \mathbb{c}_t(h\ell^k) + T2^{-L} c_\ell(F) - T\overline{y} \\
&= \sum_{s=2-L}^{0} \sum_{k=1-s}^{L-1} 2^{-k+1} \mathbb{c}_{s+k}(h\ell^k) + \sum_{s=1}^{T-L+1} \sum_{k=0}^{L-1} 2^{-k+1} \mathbb{c}_{s+k}(h\ell^k) \\
&\quad + \sum_{s=T-L+2}^{T} \sum_{k=0}^{T-s} 2^{-k+1} \mathbb{c}_{s+k}(h\ell^k) + T2^{-L} c_\ell(F) - T\overline{y} \\
&\geq \sum_{s=1}^{T-L+1} \sum_{k=0}^{L-1} 2^{-k+1} \mathbb{c}_{s+k}(h\ell^k) + T2^{-L} c_\ell(F) - T\overline{y}.
\end{aligned}$$

Since the consumption ladder that yields $\mathbb{V}^*(h, \bar{F})$ is unique, and since the h-incentive constraint is satisfied in every period, we must have, for all t,

$$\sum_{k=0}^{L-1} 2^{-k+1} \mathbb{c}_{t+k}(h\ell^k) + 2^{-L} c_\ell(F) \geq \overline{y}, \tag{C.6}$$

with equality holding if and only if the consumption ladder equals $\bar{\mathbb{c}}_*$, the unique solution to problem (13).

We now argue that

$$\lim_{s\to\infty} \sum_{k=0}^{L-1} 2^{-k+1}\mathbb{c}_{s+k}(h\ell^k) + 2^{-L}c_\ell(F) = \overline{y}.$$

The proof is by contradiction. If not, inequality (C.6) implies there exists an $\varepsilon > 0$ such that

$$\sum_{k=0}^{L-1} 2^{-k+1}\mathbb{c}_{s+k}(h\ell^k) + 2^{-L}c_\ell(F) - \overline{y} > \varepsilon \tag{C.7}$$

for infinitely many values of s. Let S denote the infinite set of values of s for which (C.7) holds, and define the function $h(T) := |S \cap \{s \leq T - L + 1\}|$. Observe that $h(T) \to +\infty$ as $T \to \infty$. Then,

$$\begin{aligned}
0 &\geq \sum_{s=1}^{T-L+1} \sum_{k=0}^{L-1} 2^{-k+1}\mathbb{c}_{s+k}(h\ell^k) + T2^{-L}c_\ell(F) - T\overline{y} \\
&\geq (T-L)(\overline{y} - 2^{-L}c_\ell(F)) + \varepsilon h(T) + T2^{-L}c_\ell(F) - T\overline{y} \\
&= \varepsilon h(T) + L(2^{-L}c_\ell(F) - \overline{y}),
\end{aligned}$$

which is impossible for large T.

Since the resource constraint is satisfied by the period-s ladder asymptotically, the sequence of ladders must converge to the unique solution to problem (13) (if not, there is a subsequence converging to a different ladder limit also satisfying the resource and incentive constraints, which is impossible). □

Lemma C.2 *If utility is CRRA, the equilibrium consumption allocation for $F = \bar{F}$ does not start immediately on the stationary ladder, that is, it is not given by* (C.1) *for $c = \bar{\mathbb{c}}_*$.*

Proof. When utility is CRRA with coefficient γ, solving (13) for the optimal stationary ladder gives, for $g = \beta^{1/\gamma} < 1$,

$$\bar{\mathbb{c}}_*(h\ell^{k+1}) = g\bar{\mathbb{c}}_*(h\ell^k) \tag{C.8}$$

when the incentive constraint is not binding on $h\ell^{k+1}$. To ease notation, define (where $\hat{c}$ is defined in (C.1))

$$\bar{c}_h := \bar{\mathbb{c}}_*(h), \quad c_t := \hat{c}(\ell^t), \quad \text{and} \quad \bar{c}_\ell := c_\ell(\bar{F}).$$

Let L be the length of the ladder, so that

$$g^{L-1}\bar{c}_h > \bar{c}_\ell \geq g^L \bar{c}_h.$$

The ladder resource constraint is

$$\sum_{k=0}^{L-1} 2^{-(k+1)} g^k \bar{c}_h + 2^{-L}\bar{c}_\ell = \bar{y}.$$

From (C.1),

$$2^{-t}c_t = \bar{y} - \sum_{k=0}^{t-1} 2^{-(k+1)} g^k \bar{c}_h,$$

and so

$$2^{-t}c_t = \sum_{k=t}^{L-1} 2^{-(k+1)} g^k \bar{c}_h + 2^{-L}\bar{c}_\ell.$$

Then,

$$\begin{aligned}
2^{-t-1}c_{t+1} &= \sum_{k=t+1}^{L-1} 2^{-(k+1)} g^k \bar{c}_h + 2^{-L}\bar{c}_\ell \\
&= \frac{g}{2} \sum_{k=t}^{L-2} 2^{-(k+1)} g^k \bar{c}_h + 2^{-L}\bar{c}_\ell \\
&= \frac{g}{2} \left\{ \sum_{k=t}^{L-1} 2^{-(k+1)} g^k \bar{c}_h - 2^{-L} g^{L-1} \bar{c}_h \right\} + 2^{-L}\bar{c}_\ell \\
&= \frac{g}{2} \left\{ 2^{-t}c_t - 2^{-L}\bar{c}_\ell - 2^{-L} g^{L-1} \bar{c}_h \right\} + 2^{-L}\bar{c}_\ell.
\end{aligned}$$

Hence,

$$c_{t+1} = g c_t + 2^t \left\{ -g 2^{-L}\bar{c}_\ell - 2^{-L} g^L \bar{c}_h + 2^{-L+1}\bar{c}_\ell \right\}.$$

Finally, since

$$-g 2^{-L}\bar{c}_\ell - 2^{-L} g^L \bar{c}_h + 2^{-L+1}\bar{c}_\ell = 2^{-L}\bar{c}_\ell (1-g) + 2^{-L}(\bar{c}_\ell - g^L \bar{c}_h) > 0,$$

we have

$$c_{t+1} > g c_t,$$

and so

$$u'(c_t) > \beta u'(c_{t+1}). \tag{C.9}$$

Consider now the following local change:

$$\begin{aligned}\hat{c}_1(h) &= \bar{c}_h - \varepsilon,\ \hat{c}_1(\ell) = c_1 + \varepsilon,\\ \hat{c}_2(h\ell) &= g\bar{c}_h + \eta(\varepsilon),\ \text{and } \hat{c}_2(\ell^2) = c_2 - \eta(\varepsilon),\end{aligned}$$

where η satisfies

$$u(\bar{c}_h) + \tfrac{\beta}{2}u(g\bar{c}_h) = u(\bar{c}_h - \varepsilon) + \tfrac{\beta}{2}u(g\bar{c}_h + \eta(\varepsilon)).$$

From the implicit function theorem and (C.8),

$$\eta'(0) = \frac{2}{\beta}\frac{u'(\bar{c}_h)}{u'(g\bar{c}_h)} = 2.$$

The impact on payoff to the low income agents is

$$u(c_1 + \varepsilon) + \tfrac{\beta}{2}u(c_2 - \eta(\varepsilon)),$$

which has slope at $\varepsilon = 0$ of

$$u'(c_1) - \tfrac{\beta}{2}u'(c_2)\eta'(0) = u'(c_1) - \beta u'(c_2),$$

which is strictly positive from (C.9). This implies the local change is ex ante welfare improving over the stationary ladder. □

Lemma C.3 *If utility is CRRA, the equilibrium consumption allocation for $F = \bar{F}$ does not reach the stationary ladder $\bar{\mathbb{c}}_*$ in finite time, that is, for all T, there exists $t > T$ and $k < L$, for which*

$$\mathbb{c}_t(h\ell^k) \neq \bar{\mathbb{c}}_*(h\ell^k).$$

Proof. Suppose not, that is, suppose there exists some T such that for all $t > T$ and $k < L$,

$$\mathbb{c}_t(h\ell^k) = \bar{\mathbb{c}}_*(h\ell^k).$$

We first claim that

$$\mathbb{c}_T(h) = \bar{\mathbb{c}}_*(h).$$

This is true because the h-incentive-feasibility constraint just binds on the agents who received an h income realization in period T, and their consumptions in all future periods are determined by the stationary ladder $\bar{\mathbb{c}}_*$.

Since the ℓ-incentive-feasibility constraint is not binding in period $T+1$, the consumption decay g_{T+1} is given by

$$g_{T+1} = \frac{\mathbb{c}_{T+1}(h\ell)}{\mathbb{c}_T(h)} = \frac{\bar{\mathbb{c}}_*(h\ell)}{\bar{\mathbb{c}}_*(h)} = \beta^{1/\gamma} =: g,$$

where the first equality is (9), the second is from the claim just proved, and the third comes from the CRRA assumption.

The same consumption decay applies in period T at all histories at the ℓ-incentive-feasibility constraint is not binding, and so we have

$$\mathbb{c}_T(h\ell^k) = g^{-1}\mathbb{c}_{T+1}(h\ell^{k+1}) \quad \text{for all } k < L-1,$$

and so

$$\mathbb{c}_T(h\ell^k) = \bar{\mathbb{c}}_*(h\ell^k) \quad \text{for all } k < L-1,$$

The h-incentive-feasibility constraint just binds on the agents who received an h-income realization in period $T-1$, and since their consumptions in all future periods are determined by the stationary ladder $\bar{\mathbb{c}}_*$, current consumption must equal $\bar{\mathbb{c}}_*(h)$. But this implies that the stationary consumption decay also applies in period $T-1$. Proceeding in this way, we conclude that the consumption for the initial h-realization agents must be $\bar{\mathbb{c}}_*(h)$. But this is impossible by Lemma C.2, and so we have a contradiction. □

D Appendix for Section 7

In this section we provide the details of how we compute equilibria in Section 7 of the main text. Section D.1 describes how to compute a stationary ladder that delivers an outside option $F \in (V^A, \bar{F})$. Section D.2 describes how to determine the value of $\bar{F}$ together with the stationary ladder attaining it. Finally, Section D.3 describes the calculation of an entire dynamic equilibrium consumption allocation converging to a stationary ladder.

D.1 Stationary Ladder

For a fixed F, a stationary ladder $c_* = (c_*(h), gc_*(h), g^2c_*(h), \ldots, c_\ell)$ that satisfies resource feasibility and h-incentive feasibility with equality (as well as ℓ-incentive feasibility) is fully

characterized by the upper and lower bound of consumption $(c_*(h), c_\ell)$, the decay rate g and the length of the ladder L. These values, all functions of a given $F \in (V^A, \bar{F})$, are calculated as follows:

1. Determine the consumption floor $c_\ell = c_\ell(F)$ from Proposition 5.4, i.e.,

$$u(c_\ell(F)) = u(\ell) + \beta\left(F - V^A\right)$$

and recall (10), which defines the value of the outside option for the high income agents as

$$W^F(h) := (1-\beta)u(h) + \beta F.$$

2. The ladder is then determined by three equations in three unknowns $c_*(h), g, L$ from

$$L = \max\left\{k : g^{k-1}c_*(h) > c_\ell(F)\right\}, \tag{D.1}$$

$$\frac{1}{2}\sum_{t=0}^{L-1}\left(\frac{1}{2}\right)^t c_*(h)g^t + \left(\frac{1}{2}\right)^L c_\ell(F) = \bar{y}, \tag{D.2}$$

and (using $W(h, c_*) = W^F(h)$ in (12))

$$W^F(h) = \left(1 - \frac{\beta}{2}\right)\left[\sum_{k=0}^{L-1}\left(\frac{\beta}{2}\right)^k u(c_*(h)g^k)\right] + \left(\frac{\beta}{2}\right)^L u(c_\ell(F)). \tag{D.3}$$

This system of equations can be reduced to one non-linear equation in one unknown $g \in [\ell/h, 1]$. Use equation (D.1) to solve for $L(g, c_*(h))$ and then equation (D.2) to solve for $c_*(h)$ and insert into (D.3) to obtain one equation in the unknown decay rate g. The result is a stationary ladder summarized by $(c_*(h)(F), g(F), L(F))$ as a function of the outside option F.

In general the stationary ladder associated with an outside option F need not be unique, although it is for $F = \bar{F}$, as we have seen in Section 6.2. To better understand the potential multiplicity of stationary ladders, instead of calculating the consumption decay rate g (and the associated $(c_*(h), L)$) as a function of F, we can in step 2 above reverse the order and calculate, for a given stationary consumption decay rate $g \in (\ell/h, 1)$, the outside option $F(g)$ associated with this g.

Numerically, we find that the mapping $F(\cdot)$ is hump-shaped with a maximum at $\bar{g} = \beta^{1/\gamma} < 1$ that delivers the maximum value $\bar{F}$. The reason for the hump-shape of $F(\cdot)$ is as follows. Start at $g = 1$, and thus a constant consumption allocation with full insurance, and

now lower g infinitesimally. Individuals with current income $y = h$ strictly prefer a more front loaded consumption allocation even though it entails more consumption risk in the future. As g initially falls from $g = 1$, both $W(h, c_*)$ and $c_*(h)$ increase, which in turn leads the fixed point $F(g)$ to increase as g falls. At $g = \beta^{1/\gamma}$ the optimal front loading is attained from the perspective of the current h types; by reducing g further the associated increased future consumption risk more than offsets the higher current consumption $c_*(h)$ chosen to satisfy the resource constraint. Thus $W(h, c_*)$ and $F(g)$ decline as g falls beyond $g = \beta^{1/\gamma}$.

We cannot prove that $F(g)$ is hump-shaped in g but always found this to be the case in our examples. This implies, in particular, that for $F < \bar{F}$ there are two associated stationary ladders that deliver the same outside option F, one with little risk sharing ($g < \bar{g}$) and one with more risk sharing ($g > \bar{g}$). Since the algorithm for computing a dynamic equilibrium is based on the convergence of the allocation to a stationary ladder, it is important to know which ladder to pick, for a given $F < \bar{F}$. The following lemma is informative for this choice.

Lemma D.1 *No equilibrium allocation converges to a stationary ladder with decay $g < \beta^{1/\gamma}$.*

Proof. Suppose an equilibrium allocation for some F converges to a stationary ladder. It is immediate that the stationary ladder cannot be Pareto dominated by another stationary ladder. We now argue that any stationary ladder c_* with $g < \beta^{1/\gamma}$ is Pareto dominated by another ladder stationary (with the same number of steps), which proves the lemma.

Since $c_*(h\ell)/c_*(h) = g$,

$$\frac{c_*(h\ell)^{-\gamma}}{c_*(h)^{-\gamma}} = \frac{u'(c_*(h\ell))}{u'(c_*(h))} > \frac{1}{\beta}. \tag{D.4}$$

Define a new stationary ladder as

$$c_*^\epsilon(h\ell^k) = \begin{cases} c_*(h) - \epsilon, & k = 0, \\ c_*(h\ell) + \eta(\epsilon), & k = 1, \\ c_*(h\ell^k), & k = 2, \ldots, L-1, \\ c_\ell(F) + 2^L \cdot \left(\frac{1}{2}\epsilon - \frac{1}{2^2}\eta(\epsilon)\right), & k = L, \end{cases}$$

where $\eta(\epsilon)$ satisfies

$$u(c_*(h) - \epsilon) + \frac{\beta}{2} u(c_*(h\ell) + \eta(\epsilon)) = u(c_*(h)) + \frac{\beta}{2} u(c_*(h\ell))). \tag{D.5}$$

The new stationary ladder c_*^ϵ satisfies the resource constraint because the change in the aggregate consumption is $-\frac{1}{2}\epsilon + \frac{1}{2^2}\eta(\epsilon) + \frac{1}{2^L} \cdot 2^L \cdot \left(\frac{1}{2}\epsilon - \frac{1}{2^2}\eta(\epsilon)\right) = 0$.

Applying the envelope theorem to (D.5) and using (D.4), we have

$$\eta'(0) = \frac{2u'(c_*(h))}{\beta u'(c_*(h\ell))} < 2.$$

Since $\eta(0) = 0$, for small $\epsilon > 0$, $\frac{1}{2}\epsilon - \frac{1}{2^2}\eta(\epsilon) > 0$, and so $c_*^\epsilon(h\ell^L) > c_\ell(F)$ and so c_*^ϵ satisfies ℓ-incentive feasibility. With (D.5), this also implies c_*^ϵ satisfies h-incentive feasibility.

Finally, c_*^ϵ clearly Pareto dominates c_* □

D.2 Determination of the Outside Option $\bar{F}$

To determine $\bar{F}$ we proceed as follows: At $F = \bar{F}$, Proposition 6 implies that there is a unique stationary ladder satisfying h-incentive feasibility and this ladder solves (13), so we know that the consumption decay rate is given by

$$g(\bar{F}) = \beta^{1/\gamma}.$$

In effect, $\bar{F}$ is the peak of the $F(\cdot)$ map discussed above, and is reached at $g = \bar{g}$. Since the value of $\bar{F}$ itself is unknown, we have to determine the lower consumption floor $c_\ell = c_\ell(\bar{F})$ jointly with $\bar{F}$, $c_*(h)$, and L. The relevant equations, with $g = g(\bar{F}) = \beta^{1/\gamma}$ are

$$u(c_\ell) = u(\ell) + \beta\left(\bar{F} - V^A\right), \tag{D.6}$$

$$\bar{y} = \frac{1}{2}\sum_{t=0}^{L-1}\left(\frac{1}{2}\right)^t c_*(h)g^t + \left(\frac{1}{2}\right)^L c_\ell, \tag{D.7}$$

$$L = \max\{k : g^{k-1}c_*(h) > c_\ell\}, \quad \text{and} \tag{D.8}$$

$$(1-\beta)u(h) + \beta\bar{F} = \left(1 - \frac{\beta}{2}\right)\left[\sum_{k=0}^{L-1}\left(\frac{\beta}{2}\right)^k u(c_*(h)g^k)\right] + \left(\frac{\beta}{2}\right)^L u(c_\ell). \tag{D.9}$$

The algorithm to determine $\bar{F}$ is then a slightly modified version of the procedure from the previous subsection, with $\bar{F}$ replacing g as the unknown to be computed (and identical to the computations we do when solving for $F(g)$ for a gven $g \neq \bar{g}$.)

1. Guess $\bar{F} \in (V^A, V^{FB})$.
2. For a given $\bar{F}$:
   (a) Solve for c_ℓ from (D.6).
   (b) Jointly solve for $(c_*(h), L)$ from (D.7) and (D.8).

(c) Calculate the right side of (D.9).

3. Solve $\bar{F}$ such that (D.9) holds.

D.3 Computation of the Transition

As discussed in the main text, the computational procedure solves for the equilibrium allocation, imposing the stationary ladder from an exogenously specified period T. We now describe the computation of the allocations for fixed T and fixed outside option $F \leq \bar{F}$. We take as given the stationary ladder associated with F, summarized by $(c_*(h)(F), g(F), L(F))$, including the lifetime utilities $V_{i,\infty}(F)$, as described in the previous two subsections.[12] As described in the main text, the algorithm calculates consumption in three phases.

In the first $t \leq T$ periods the algorithm picks time-varying consumption of individuals with currently high income (and so have binding incentive constraints), $(\mathbb{c}_t(h))_{t=1}^T$ and uses the resource constraints and the fact that individuals without binding constraints have common consumption decay rates (or consume the lower bound consume $c_\ell(F)$) to pin down the remainder of the consumption allocation. In a second phase, from $t = T+1, ..., T+L(F)$ the allocation blends into the stationary ladder: all individuals with high income consume according to the stationary ladder, and all households with low income drift down from consumption in the previous period at a common (but time-varying) decay rate g_t.[13] Finally, for all $t > T + L(F)$, the allocation coincides with the stationary ladder. More precisely, the algorithm works as follows:

1. Guess $(\mathbb{c}_t(h))_{t=1}^T \in (\bar{y}, h)^T$.

2. Calculate the consumption allocation implied by this guess, imposing the characterization of an equilibrium allocation: the h-incentive-feasibility constraint binds in every period, and all agents with low income either have non-binding constraints and their consumption decays at a common rate or they consume c_ℓ. The implied consumption allocations $(c_{i,t})_{i=0}^t$ for all $t = 1, \ldots, T, T+1, \ldots, T+L(F)$, are calculated as follows, where i again indicates the position on the consumption ladder:

[12] The only part that distinguishes the calculations for $F < \bar{F}$ and $F = \bar{F}$ is the calculation of the stationary ladder(s), and in case of $F < \bar{F}$, the selection of the right ladder.

[13] Similar arguments to those proving Proposition 5.1 show that this property must also hold for constrained optimal allocations.

(a) Set

$$\begin{aligned} & c_{0,t} = \mathbb{c}_t(h) \text{ for } t = 1, \ldots, T, \\ \text{and} \quad & c_{0,t} = c_*(h)(F) \text{ for } t = T+1, \ldots, T+L(F). \end{aligned}$$

(b) For $t = 1$, determine $c_{1,1}$ from

$$\frac{1}{2}\left[c_{0,1} + c_{1,1}\right] = \bar{y}.$$

(c) For $t = 2, \ldots, T$, determine the consumption decay rates $(g_t)_{t=2}^T$ recursively (beginning with $t = 2$) as follows:

The consumption decay g_t solves

$$\frac{1}{2}\sum_{i=0}^{t-1}\left(\frac{1}{2}\right)^i c_{i,t} + \left(\frac{1}{2}\right)^t c_{t,t} = \bar{y},$$

where for all $i = 1, \ldots, t$,

$$c_{i,t} = \max\{g_t c_{i-1,t-1}, c_\ell(F)\}.$$

For each t, g_t is determined by one equation. The equations are solved forward in time since the allocations $\{c_{i,t}\}$ require knowledge of allocations $\{c_{i-1,t-1}\}$.

(d) For $t = T+1, \ldots, T+L(F)$, part of the consumption allocations are on the stationary ladder. For each $t = T+1, \ldots, T+L(F)$, the consumption decay g_t solves

$$\frac{1}{2}\sum_{i=0}^{t-1}\left(\frac{1}{2}\right)^i c_{i,t} + \left(\frac{1}{2}\right)^t c_{t,t} = \bar{y},$$

where

$$c_{i,t} = \begin{cases} g^i c_h(F), & \text{for } i = 1, \ldots, t-T-1, \\ \max\{g_t c_{i-1,t-1}, c_\ell(F)\}, & \text{for } i = t-T, \ldots, t. \end{cases}$$

3. For a given guess $(\mathbb{c}_t(h))_{t=1}^T$, the previous step delivers the entire allocation $(c_{i,t})_{i=0}^t$ for periods $t = 1, \ldots, T, T+1, \ldots, T+L(F)$. From date $t = T+L(F)+1$ on the consumption allocation coincides, by assumption, with the stationary ladder. Now we need to determine $(\mathbb{c}_t(h))_{t=1}^T$. These values must yield a consumption allocation that delivers the outside option $W^F(h)$ for all $t = 1, \ldots, T$. Construct the lifetime utility in period t after the history $y^{t-1-i}h\ell^i$, $V_{i,t}$, from the consumption allocation computed

in the previous step. This can be done recursively, going backward in time. Lifetime utilities are given by, for each $t = T + L, \ldots, 1$ (working backwards in time) and all $i = 0, \ldots, t$,

$$V_{i,t} = (1 - \beta)u(c_{i,t}) + \frac{\beta}{2}\left[V_{0,t+1} + V_{i+1,t+1}.\right]$$

Note that these calculations are the same before and in the blended phase, because $V_{0,t}$ is a function of $V_{i,t+i}$ for $i = 1, \ldots, L$, with $V_{L,T+L} = (1 - \beta)u(\ell) + \beta F$ and $t \leq T + L$. The only role the consumptions from the stationary ladder play is in step 2 above in determining $c_{i,t}$ via resource feasibility.

Finally we need to check whether the entry consumption levels $(\mathbb{c}_t(h))_{t=1}^T$ are such that the resulting consumption allocation hits the outside option for each $t = 1, \ldots, T$

$$V_{0,t} = (1 - \beta)u(h) + \beta F.$$

If yes, we are done. If not, go back to step 1 and adjust the guess for $(\mathbb{c}_t(h))_{t=1}^T$.