ECONtribute
MARKETS & PUBLIC POLICY

ECONtribute
Discussion Paper No. 064

Social Capital: A Double-Edged Sword

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

February 2021

www.econtribute.de

UNIVERSITÄT BONN
Universität zu Köln

Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany´s Excellence Strategy – EXC 2126/1– 390838866 is gratefully acknowledged.

Social Capital: A Double-Edged Sword*

Harold L. Cole† Dirk Krueger‡ George J. Mailath§ Yena Park¶

February 1, 2021

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 which we term "social capital". We treat the contracting conditions of original and deviating coalitions symmetrically and show that higher social capital tightens incentive constraints since it facilitates both the formation of the original as well as a deviating coalition. As a consequence, although social capital facilitates the initial formation of coalitions, the extent of risk sharing in successfully formed coalitions is declining in the extent of social capital and equilibrium allocations might feature resource burning or utility burning: social capital is indeed a double-edged sword.

Keywords: Financial Coalition, Limited Enforcement, Risk Sharing, Coalition-Proof Equilibrium

JEL Classification Codes: E21, G22, D11, D91

*Cole, Krueger, and Mailath thank the National Science Foundation for research support (grants #SES-112354, 1260753, 1326781, 1559369, 1757084, 1851449). Earlier versions were circulated with the title "Coalition-Proof Risk Sharing Under Frictions."

†Department of Economics, University of Pennsylvania, and NBER.

‡Department of Economics, University of Pennsylvania, CEPR and NBER.

§Department of Economics, University of Pennsylvania, and RSE, Australian National University

¶Seoul National University

1 Introduction

A large literature in economics and political science argues that *social capital* is a critical determinant of the ability of communities to cooperate and that social capital differs systematically across cultures, countries, and time. In his influential book, Putnam describes social capital as the "social networks and the norms of reciprocity and trustworthiness that arise from them" (Putnam, 2000, p. 19). This description is consistent with a common view that agents behave cooperatively because they expect their cooperation to be reciprocated in the future. In other words, cooperation requires a *shared* belief in future cooperation. We interpret *social capital* as the ability to generate this shared belief and select efficient arrangements, and analyze its impact on the efficiency of social arrangements.

A critical feature of our modeling of social capital is that the notion of efficiency is *second best*, reflecting endogenously determined incentive constraints. Consequently, higher social capital need not imply greater social welfare. Social capital is a double-edged sword: agents in societies with more social capital are both more likely to enter into beneficial arrangements but also more likely to find ways to circumvent such arrangements when profitable opportunities arise to do so, by forming cooperative deviating coalitions with other members of the society at large, thus undermining the original arrangement.

To make this idea concrete, we study risk sharing in an infinite-horizon continuum economy with idiosyncratic income risk. By pooling income in each period, a coalition of agents can achieve higher ex ante utility for each agent. Such a cooperative agreement, however, requires currently rich agents to sacrifice current consumption. In the absence of commitment, the standard incentive device to induce cooperation by the rich agents is to exclude deviators from future insurance. But if agents were able to reach the original cooperative agreement, then there is also the possibility that rich agents deviate by leaving their current arrangement in the hope of, for example, replicating the current arrangement with other deviating rich agents.[1] Since we are interested in comparative statics with respect to social capital, we need a measure of the ability of a group of agents to reach cooperative agreements (based on a shared belief in future cooperation). We parameterize this measure in a stark fashion as the probability $\pi \in [0, 1]$ that a coalition can coordinate beliefs on the most efficient equilibrium. With complementary probability, the absence of belief coordination is permanent and there is no risk sharing; our results are robust to alternative assumptions (Section 9.1).

[1]This is the central difference between our work and that of Genicot and Ray (2003), who also study the formation and stability to joint deviations of risk-sharing coalitions. They study a finite population world, and restrict coalitional deviations to subcoalitions of the original group. Since smaller groups have less capacity to share risk, deviating subcoalitions cannot replicate the current arrangement.

Our parameterization of social capital accords well with the approach in the political science and sociology literature that defines social capital by its sources rather than its consequences, see e.g., Putnam (1993), Putnam (2001), Portes (1998), and Woolcock (1998).[2] In that literature, two approaches have been taken to measure social capital empirically, either by membership rates of organizations or by measured levels of trust.[3] Fukuyama (1995) argues that "trust" is fundamental to the formation of large corporations and hence a key component in explaining economic differences both across time and across countries. This argument is empirically tested by Knack and Keefer (1997) who find that while trust and cooperation are associated with stronger growth and investment, associational activity—measured by membership in groups—has no significant effect on economic performance. The premise of this paper that social capital is a double-edged sword and should accommodate its negative consequences has also been recently stressed in political science and sociology (see, for example, Portes and Landolt (1996), Putnam (2000), Woolcock (1998), Woolcock and Narayan (2000), Woolcock (2001)).

It is natural to require an equilibrium allocation to be robust to the possibility that any subset of agents could defect, not contribute in the current period and, with probability π, coordinate on an efficient equilibrium (which is similarly robust). But, as we will see, for high π, the robustness requirement can become *too* demanding, in the sense that *no* equilibrium allocation is robust in the sense just described. We therefore weaken the robustness requirement as follows: an allocation is *internally-incentive feasible* if it is robust to the possibility that a set of homogeneous agents (typically, but not always, the wealthy agents) could defect, not contribute in the current period and (with probability π) "reinitialize" risk-sharing using the *same* allocation.[4] An internally-incentive feasible allocation is a credible social norm or arrangement: A necessary condition for an allocation to be credible 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 *original*

[2]Fukuyama (2001, p. 7) makes precise his definition of social capital when he writes "While social capital has been given a number of different definitions, many of them refer to manifestations of social capital rather than to social capital itself. The definition I will use in this paper is: social capital is an instantiated informal norm that promotes co-operation between two or more individuals. ... By this definition, trust, networks, civil society, and the like, which have been associated with social capital, are all epiphenomenal, arising as a result of social capital but not constituting social capital itself."

[3]Membership rates in organizations (e.g. Putnam (1993)) is an input indicator since social capital can be accumulated through the associations and networks. Membership is a proxy for social capital in the same way that years of schooling are a proxy for human capital. Trust can be seen both as a component or consequence of social capital and the most commonly used trust indicator from the World Values Survey which measures trust to overall people in the society is likely to be an outcome of social capital. This trust indicator therefore likely reflects the mixed consequences of the double-edged effects of higher social capital.

[4]We argue below the we can restrict attention to deviations by homogeneous coalitions.

consumption plan (which is feasible for large coalitions). An allocation is an *equilibrium* if it is the ex ante utility maximizing internally-incentive feasible allocation.

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 equilibrium allocations.[5] Sections 5 and 6 describe this general indirect approach, which focuses on maximizing ex ante utility subject to exogenous outside options. For some parameters, there is a fixed point characterization of equilibrium relating the value of the outside options and the maximized value of ex ante utility. In that case, equilibrium allocations satisfy the stronger notion of robustness we described above: they are 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).

Proposition 3 is our central equilibrium characterization result. There is a threshold discount factor $\underline{\beta} < 1$, such that if the discount factor satisfies $\beta \leq \underline{\beta}$, the only equilibrium allocation is autarky, irrespective of the level of social capital. For $\beta > \underline{\beta}$, there is a critical value of social capital, $\overline{\pi}(\beta) \in (0, 1]$, such that for values of social capital below this threshold ($\pi \leq \overline{\pi}(\beta)$), the fixed point characterization applies, and the second-best allocation can be determined using standard techniques (this is done in Section 7.1). A larger value of π reduces the extent of risk-sharing and lowers expected utility from a successfully formed coalition, strictly so if the first-best, full insurance allocation cannot be sustained. Nonetheless, ex ante utility, the weighted sum of a successfully formed coalition (weight π) and an unsuccessful attempt at coordinating beliefs (weight $1 - \pi$), is strictly increasing in π.

As long as agents are not too patient, the critical value of social capital $\bar{\pi}(\beta)$ is strictly less than 1; indeed, as β approaches $\underline{\beta}$ from above, $\overline{\pi}(\beta)$ converges to 0. For high values of social capital ($\pi > \overline{\pi}(\beta)$), the value of the outside option is so attractive that equilibrium cannot satisfy the stronger notion of robustness discussed above. To prevent deviating coalitions forming, within original coalitions utility must be "burnt" (Section 7.3), either through introducing further inefficiencies in risk sharing or by burning resources. The need for utility burning is strictly increasing in π, and ex ante utility remains at its maximal sustainable level as π rises from $\overline{\pi}(\beta)$ to 1.

We proceed by placing our contribution in the literature in the next section. Following the theoretical analysis described above in Sections 3-7, Section 8 presents results for a computed example to convey the qualitative properties of the equilibrium and Section 9 discusses two

[5]Since ex ante utility is continuous, and the set of internally-incentive feasible allocations is compact (in the product topology), existence of equilibrium is immediate.

extensions, one to temporary delay of risk sharing after a failure to form a coalition, and the other to a production economy. Section 10 concludes.

2 Literature Review

The nonexistence of equilibrium under the stronger robustness notion and the associated need for utility burning is a general phenomenon. 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 gift-exchange and related models (Carmichael and MacLeod, 1997, Kranton, 1996a,b, 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 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). That previous literature emphasized the difficulty of deterring opportunistic behavior in the setting of unilateral deviations. We extend this insight to the setting of coalitional deviations, suggesting that the difficulty is intrinsic to the incentive compatibility of many institutions.

Our symmetric treatment of the original and deviating coalitions in terms of available coalition members and feasible allocations underlies the stark difference between our results and the limited commitment literature in (macro-)economic theory, which assumes, explicitly or implicitly, that deviators have less opportunities than the originally formed coalition. Within this literature, in macroeconomics, Kehoe and Levine (1993) and Alvarez and Jermann (2000) characterize consumption allocations in a general equilibrium limited commitment 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 (2006) extend this literature to a risk sharing economy with as a continuum of households exactly of the form studied in this paper.[6] These papers share our focus on self-enforcing arrangements, but take the outside option as exogenously given, and equal to the autarkic allocation, which is essentially assuming $\pi = 1$ in the original coalition while $\pi = 0$ in the deviating coalition. *Given* this outside option, the qualitative properties of the equilibrium allocation in this work and our paper (when

[6] A strand of the sovereign debt literature also considers self-enforcing simple debt contracts because sovereigns cannot commit to repay, see e.g. Eaton and Gersovitz (1981) and Bulow and Rogoff (1989). Ábrahám and Laczó (2017) also analyze a limited commitment model with a private storage technology.

$\pi \leq \bar{\pi}(\beta)$) 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. Krueger and Uhlig (2006) assume that outside option is determined by the best insurance contract offered by a competing financial intermediary, who has long term commitment.[7] Hellwig and Lorenzoni (2009) endogenize the outside option by assuming that the only punishment for deviators is the denial of future credit (but they are allowed to save). These papers also define equilibrium as a fixed point, but unlike our paper, the nonexistence issue does not arise.[8] The central difference between that work and our paper is that they assume asymmetric contracting conditions between the original and deviating coalition while we assume exactly symmetric one. With an asymmetric treatment, there is more room for relaxing incentive constraints through adjustments of endogenous variables or exogenous parameters, which has asymmetric effects on the payoff within and outside the coalition. In Krueger and Uhlig (2006), higher return of storage makes the deviation more costly by losing the storage upon deviation, and in Hellwig and Lorenzoni (2009), endogenous lower interest rates make default less attractive by lowering the return of saving after default. With the symmetric treatment in our paper, however, the ex ante value and the deviation payoff are themselves related by a fixed point, which generates a strong feedback on risk sharing opportunities.

Also related are a set of papers analyzes how government social insurance policies (unemployment insurance, progressive taxation, disability insurance) impact the outside option and thus equilibrium private insurance, see e.g. Krueger and Perri (2011) and Park (2014). In this literature the outside option is endogenous from the perspective of the policy maker. There is also a related literature which endogenizes the outside option by assuming that private noncontingent intertemporal trades can be enforced and examines how this impacts on insurance (see, for example, Allen, 1985) and government taxation (see Farhi, Golosov, and Tsyvinski (2009)).[9]

Most papers on risk-sharing consider only unilateral deviations of individuals from the risk sharing arrangement, thereby limiting the extent of insurance that can be obtained after deviating. An exception to this is Genicot and Ray (2003), who study the formation and

[7] 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.

[8] More precisely, in Hellwig and Lorenzoni (2009), the fixed-point equilibrium is a fixed point in borrowing limits.

[9] In the context of private information, Cole and Kocherlakota (2001) endogenize the outside option, assuming hidden storage, and Golosov and Tsyvinski (2006) analyze optimal disability insurance.

stability to *joint deviations* of risk sharing coalitions in economies with finite populations.[10] 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.[11] Since larger coalitions are more prone to successful deviation, an optimal size of the original coalition emerges. This result stems from their 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 not only unilateral deviations by an individual, but to coalitional deviations. In contrast to Genicot and Ray (2003), however, we allow deviating coalition to have the same insurance capabilities as the original coalition.

Our comparative statics results with respect to social capital π imply that in societies with larger social capital risk sharing arrangements are more likely to form and ex-ante welfare is (weakly) higher, however ex post utility (conditional successfully forming a coalition) and risk-sharing within the formed organizations is lower.[12] This accords well with the differential evidence on the high degree of risk sharing in poor, rural village economies in developing countries (see e.g. Townsend (1994) or Ligon, Thomas, and Worrall (2002)) versus the relatively lower degree of risk-sharing in (see e.g. Attanasio and Davis (1996) or Altonji, Hayashi, and Kotlikoff (1997)).[13]

3 Model

3.1 The Environment: Income, Preferences and Technology

Time t is discrete and extends from period $t = 0$ to infinity. A unit measure of infinitely-lived agents face idiosyncratic income risk in each period. An agent in each period has low income $y = \ell > 0$ and high income $y = h > \ell$ with equal probability; we write $Y := \{\ell, h\}$. Income

[10] 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 an autarkic outside option.

[11] Genicot and Ray (2003) builds on the more abstract game theoretic literature on coalition deviations pioneered by Bernheim, Peleg, and Whinston (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," that is, $\pi = 1$.

[12] Once the outside option is binding and full risk-sharing is no longer possible.

[13] We do not contend that our mechanism limiting risk sharing is the only one consistent with this observation. The differential importance of private information about income could also explain these facts.

realizations are 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 $\overline{y} = \frac{1}{2}(\ell + h)$ aggregate income per capita, and 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 income, is

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

where ex ante autarky utility is $V^A := Eu(y)$. The first-best payoff obtained from consuming average income with certainty is

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

3.2 Coalition Formation and Deviation

In the initial period $t = 0$, agents attempt to form risk-sharing arrangements. 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. We assume that the size of the coalition does not restrict the size of deviating coalitions, since these can implicitly include members from outside the original coalition. Because future income risk is more effectively shared in large coalitions, all coalitions will be large (i.e composed of a continuum of individuals), both the initial and potential deviating coalitions. Note that the possibility of forming a new large coalition is most threatening to the original coalition.[14] A continuum population within a coalition simplifies our life in two ways. First, there is no aggregate income risk in any large coalition. Second, deviating

[14]In contrast, Genicot and Ray (2003) assume that deviations have to come from sub-coalitions and hence restricting the original size, while sacrificing some risk-share benefits, is optimal since it restricts the outside option for these sub-coalitions. Our symmetric treatment of initial and deviating coalitions with respect to the choice of size is a key factor in generating our results, as this binds the ex ante payoff to the original coalition and the outside option tightly together.

coalitions do not benefit from adding additional agents who were not part of the original agreement.

We do not model 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.

A risk-sharing agreement within a coalition is only reached if its members are confident that future cooperation is sustainable. This confidence requires significant *social capital* since the incentive compatibility of future cooperation depends on intertemporal incentives that themselves need to be incentive compatible. We model social capital in an admittedly crude fashion by assuming that any attempt to form a coalition succeeds with an exogenous probability $\pi \in [0, 1]$. If the attempt succeeds, then the coalition immediately implements a new risk-sharing agreement.[15] When a new (or deviating) coalitions fails to form (which happens with probability $1 - \pi$), agents receive their autarky payoff V^A.[16] 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.

3.3 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 histories y^t) 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

[15] If the deviation coalition is homogeneous, there is no risk sharing in the first period.

[16] 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. The case of a temporary delay in insurance (rather than permanent absence) is discussed in Section 9.1.

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).

4 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.

A 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.[17] 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 forming the new coalition, 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.

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

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

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

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 $y^t \in Y^t$,*

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

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.[18] 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,[19] the stronger requirement (which we discuss just after Proposition 2 in Section 5) 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.

[18] 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, in the thoery of repeated games, to Farrell and Maskin's (1989) notion of *weakly renegotiation proof* and Bernheim and Ray's (1989) notion of *internal consistency*. Note that these authors effectively assume $\pi = 1$.

[19] 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* by Farrell and Maskin (1989) or *strong consistency* by Bernheim and Ray (1989).

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 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. If a homogeneous large group with current income y does deviate, with probability π, the group is able to coordinate on future risk sharing (in the current period, agents consume y since all agents have identical current income), with payoff $(1-\beta)u(y)+\beta\mathbb{W}$. With probability $1-\pi$, there is no future risk sharing, and so $(1-\beta)u(y)+\beta\mathbb{F}$ is the expected payoff from deviating.

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. Moreover, as $\mathcal{C}$ is closed (in the product topology), the supremum is always attained and so equilibrium exists. The bulk of our analysis is concerned with the characterization of equilibrium.

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*

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

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, seceding 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 if

$$-(1-\beta)u'(h) + \frac{\beta}{2}[u'(l) - u'(h)] \geq 0 \implies \beta \geq \frac{2u'(h)}{u'(\ell) + u'(h)} > \frac{u'(h)}{u'(\ell)}. \tag{6}$$

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). Finally, it is worth noting that (4) was derived assuming $\pi = 1$ while (6) was derived assuming $\pi = 0$.

5 Equilibrium as a Fixed Point

Characterizing 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*,

$$W(y^t, c) \geq (1-\beta)u(y_t) + \beta F \qquad \forall y^t \in \cup_\tau Y^\tau. \tag{7}$$

For exogenous $F \in \mathbb{R}_+$, denote by $\mathcal{C}(F)$ the set of resource-feasible allocations satisfying (7). 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

$$\mathbb{V}(F) := \max_{c \in \mathcal{C}(F)} W^0(c). \tag{8}$$

Social capital π does not appear in the maximization in (8). Instead, the exogenous value of the outside option F determines the optimal allocation and value. 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, 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 another 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 exist 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 guarantee 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 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

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

We can now state the main result of the paper (which summarizes the analysis to follow):[20]

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

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

2. *Suppose $\beta > \underline{\beta}$. Risk sharing does occur in equilibrium (and so autarky is not an equilibrium). There exists a value of π, $\overline{\pi}(\beta) \in (0,1]$, such that*

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

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

3. $\lim_{\beta \searrow \underline{\beta}} \overline{\pi}(\beta) = 0$.

We conjecture that $\overline{\pi} < 1$ for all $\beta \in (0,1)$ (and not just for β near $\underline{\beta}$, as guaranteed by Proposition 3.3). While we have not been able to prove this, all of our numerical examples have this property (we discuss this in more detail in Section 8).

Proof. Existence of equilibrium is immediate, as we discussed after Definition 3.

[20] To simplify notation, we occasionally leave the dependence on β of $\overline{\pi}$, $\overline{F}$, and similar functions implicit.

1. This is an implication of the machinery we develop to characterize $\overline{F}$, and is Corollary 1 in Section 7.2.

2. (a) This is an immediate implication of Propositions 2 and 4 (which is in the next section), and the strict concavity of the problem (8).

   (b) This is Lemmas 1 and 2 in Section 7.3.

3. This is also an implication of the machinery we develop to characterize $\overline{F}$, and is Corollary 2 in Section 7.2.

□

6 Understanding Equilibrium Values

We begin by studying the program (8) and the fixed points of $\mathcal{T}(\,\cdot\,;\pi)$. The proof of the following result is in Appendix A.

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

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

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

3. *For $F \leq \overline{F}$, the value of the problem (8), $\mathbb{V}(F)$, is continuous in F.*

4. *Defining*
$$\overline{\pi} := \min\left\{\frac{\overline{F} - V^A}{\mathbb{V}(\overline{F}) - V^A}, 1\right\},$$
for all $\pi \in (0, \overline{\pi}]$, $\mathcal{T}(\cdot, \pi)$ has a unique fixed point $F(\pi)$. The function $F(\cdot)$ is increasing in π. If $\overline{\pi} < 1$, $\overline{F} = F(\overline{\pi})$ and if $F(\overline{\pi}) < \overline{F}$, $\overline{\pi} = 1$.

5. *If $\overline{\pi} < 1$, for $\pi > \overline{\pi}$, $\mathcal{T}(\,\cdot\,,\pi)$ does not have a fixed point.*

Note that autarky is not a fixed point equilibrium when $\pi > \overline{\pi}$ (Proposition 4.5). Although autarky is internally-incentive feasible, it is dominated by a better allocation which is not internally-incentive feasible when a seceding coalition is free to reoptimize.

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 $\mathbb{V}(V^A) = V^{FB}$

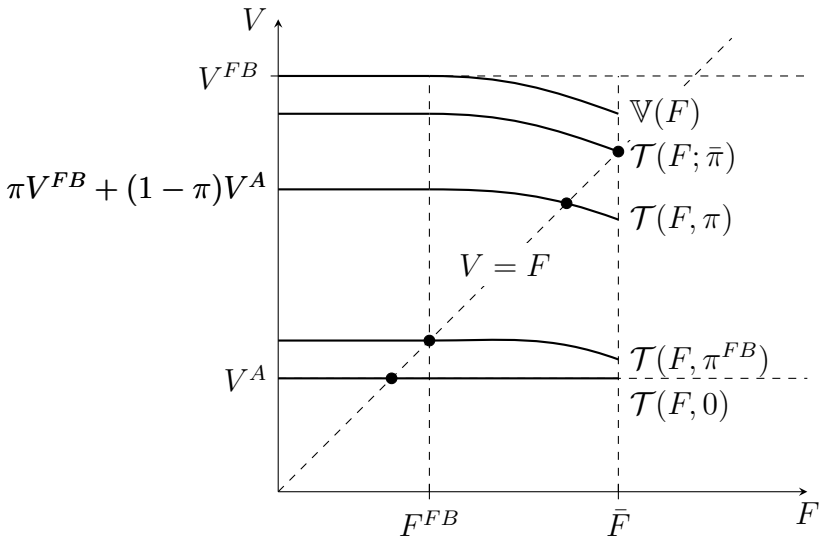

Figure 1: Determination of the fixed point of $\mathcal{T}(F;\pi) = \pi\mathbb{V}(F) + (1-\pi)V^A$ for different values of π. Drawn for $\beta > \beta^{FB}$, where β^{FB} is defined in Proposition 1 and assuming $\mathbb{V}(\overline{F}) > \overline{F}$ (Lemma A.2 verifies that $F > F^{FB}$ if $\beta^{FB}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

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 $\pi \in (\pi^{FB}, \overline{\pi}]$, 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 (7) 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.

Defining $\widetilde{F} := (1-\beta)V^A + \beta\overline{F}$, it is immediate that $\widetilde{F} < \mathbb{V}(\widetilde{F})$ from the observation that if (8) is solved by $\mathbb{c} \in \mathcal{C}(\overline{F})$, then the allocation

$$\mathbb{c}'(y^t) = \begin{cases} y_1, & \text{if } t = 1, \\ \mathbb{c}(y_2, \ldots, y_t), & \text{if } t \geq 2, \end{cases}$$

is an element of $\mathcal{C}(\widetilde{F})$. This in turn implies that for all $F \in [V^A, \widetilde{F}]$, $F < \mathbb{V}(F)$.

Suppose $\overline{\pi} < 1$. For $\pi > \overline{\pi}$,

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

Since $\mathcal{C}(F)$ is empty for $F > \overline{F}$, 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

$$\pi W^0(\mathbb{c}) + (1-\pi)V^A \leq \overline{F}. \tag{9}$$

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

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

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

An equilibrium allocation maximizes ex ante utility (the left side of (9)). We show in Section 7.3 that equilibrium allocations in fact satisfy (9) with equality.

Our notation suppresses the dependence of $\overline{\pi}$ 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 $\pi \in (\pi^{FB}(\beta), \overline{\pi}(\beta)]$, $\mathcal{T}(\,\cdot\,,\pi)$ 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 $\pi > \overline{\pi}(\beta)$, $\mathcal{T}(\cdot\,;\pi)$ does not have a fixed point, the internal-feasibility constraint is binding in equilibrium, expected lifetime utility is fixed at $\overline{F}$ independent of π (since (9) holds with equality) and attained with an allocation that features utility burning and partial risk sharing.[21]

[21] When $\pi = 1$, our model is directly comparable to the no-storage case of Krueger and Uhlig (2006)—where the return of the storage R is so low that the storage is not used. If $\beta \leq u'(h)/u'(\ell)$, our result is consistent with the autarkic fixed point equilibrium (with no storage usage) in Krueger and Uhlig (2006). However, they only considered $R > 1$ under which the storage is always used if $\beta > u'(h)/u'(\ell)$. That is, nonexistence issue does not arise in Krueger and Uhlig (2006) because of the restriction on the parameters (β, R).

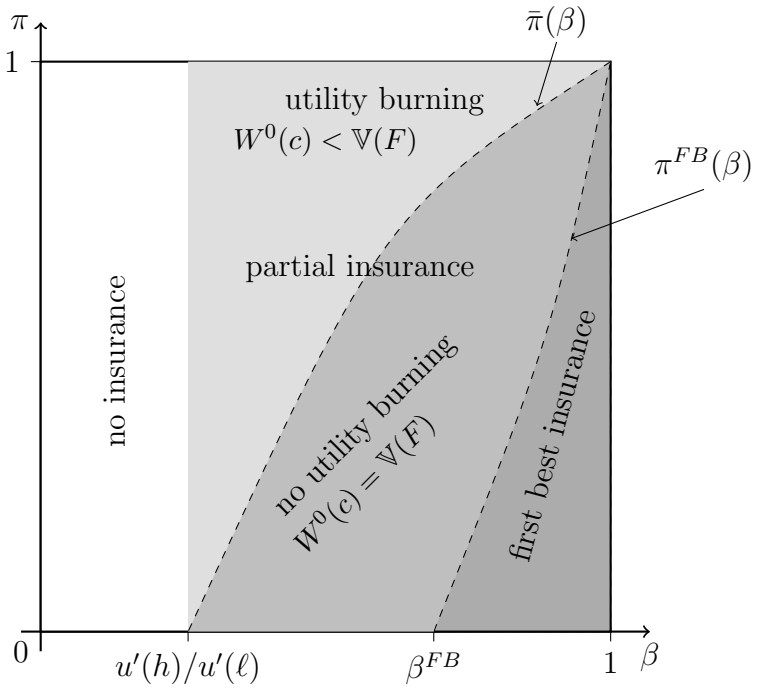

Figure 2: The insurance possibilities as a function of the discount factor and social capital. Equation (3) shows that $\pi^{FB}(\beta^{FB}) = 0$. By Corollary 2, $\overline{\pi}(\beta)$ converges to 0 as β tends to $\underline{\beta} := u'(h)/u'(\ell)$.

7 Characterizing Equilibrium Allocations

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

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

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 $F \in (F^{FB}, F(\overline{\pi})]$, where $F(\overline{\pi})$ is the fixed point associated with $\overline{\pi}$ (recall Proposition 4.4). 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 in Appendix B.

Proposition 5 *Suppose $\beta u'(\ell) > u'(h)$ and $F \in (F^{FB}, F(\overline{\pi})]$. 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*

$$\frac{u'(\mathbb{c}(y^t))}{u'(\mathbb{c}(y^{t+1}))} = \delta_{t+1} \tag{10}$$

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},*

$$W(y^{t-1}h, \mathbb{c}) = (1-\beta)u(h) + \beta F =: W^F(h), \tag{11}$$

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,

$$\mathbb{c}(y^t\ell) \geq c_\ell(F). \tag{12}$$

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

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

and so, simplifying, we get

$$W(h, c_*) = \left(1 - \tfrac{\beta}{2}\right)\sum_{k=0}^{\infty}(\tfrac{\beta}{2})^k u(c_*(h\ell^k)). \tag{13}$$

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 trade-off: 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 Our equilibrium definition determines allocations within a successfully formed coalition as the solution to a social planner problem with equal Pareto weights. When $\pi \leq \overline{\pi}$ and the equilibrium allocations are solutions to the fixed point of $\mathcal{T}(\,\cdot\,;\pi)$, these consumption allocations *within* a coalition can be decentralized as in Kehoe and Levine (1993).[22] 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 (7). Thus, the individual's problem is isomorphic to the Lagrangian problem for the coalition. Note that this is not the case when $\pi > \overline{\pi}$ since in that case the coalition must respect a binding coalition-level constraint on the overall level of ex ante welfare.

♦

7.2 Characterizing $\overline{\pi}$

We now characterize $\overline{\pi}$, or equivalently, $\overline{F}$. It turns that $\overline{F}$ 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 $\overline{F}$. Using this

[22] 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, or Alvarez and Jermann, 2000).

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 (13), subject to incentive feasibility for ℓ realizations and resource feasibility. Recalling (12), this is

$$\mathbb{V}^*(h; F) := \max_{c_* \in \mathcal{C}_*(F)} W(h, c_*), \tag{14}$$

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

$$\sum_{k=0}^{\infty} (\tfrac{1}{2})^{k+1} c_*(h\ell^k) \leq \overline{y} \tag{15}$$

and incentive feasibility

$$c_*(hy^k) \geq c_\ell(F) \text{ for all } k \geq 1. \tag{16}$$

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 $\overline{F}$ 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 (11). Moreover,

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

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

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

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

$$\mathbb{V}^*(h, F) \geq (1 - \beta)u(h) + \beta F. \tag{17}$$

But $\beta u'(\ell) \leq u'(h)$ implies that the autarkic consumption provides an upper bound for (14) and so (using (13))

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

contradicting (17). Hence, we must have $\overline{F} = V^A$. □

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. The next corollary (proved in Appendix C) confirms that we have continuity from the right.

Corollary 2

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

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

Proposition 7 *Suppose $\beta u'(\ell) > u'(h)$ and $F = \overline{F}$. The equilibrium allocation $\mathbb{c}$ converges to the unique solution to problem (14), $\bar{c}_*$, that is (where L is from Proposition 5.6),*

$$\begin{aligned}\lim_{t\to\infty} \mathbb{c}_t(h\ell^k) &= \bar{c}_*(h\ell^k) \quad \textit{for all } k < L \textit{ and}\\ \mathbb{c}_t(h\ell^k) &= c_\ell(\overline{F}) \quad \textit{for all } k \geq L \textit{ and } t > k.\end{aligned}$$

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

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

There exists $\beta_* \in (\underline{\beta}, 1)$, such that for all $\beta > \beta_*$, the 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 $F < \overline{F}$. 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. 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. 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

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

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. 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 8 suggest that these two bounds are close.

7.3 The case of utility burning, $\pi > \overline{\pi}$

For high values of social capital ($\pi > \overline{\pi}$), 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 $\overline{F}$.

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 $F = \overline{F}$. The first lemma describes an equilibrium that postpones risk sharing.

Lemma 1 *Suppose $\pi > \overline{\pi}$. 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 $\overline{F}$.

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 $c^{(T-1)} \in \mathcal{C}(\overline{F})$ and

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

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

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

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

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

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

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

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

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

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

□

The next lemma describes an equilibrium that burns resources.

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

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

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

Note that the consumption allocation $c^{[\alpha]}$ only differs from $\bar{\mathbb{c}}$ at histories ℓ^t. Moreover, since $\bar{\mathbb{c}}(\ell^t) = c_\ell(\overline{F})$ in finite time (Proposition 5.6), $c^{[\alpha]}(y^t) = \bar{\mathbb{c}}(y^t)$ for $t \geq L$.

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

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

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

The consumption $c_\ell(\overline{F})$ 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

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

This implies

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

so that

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

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

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

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

8 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 (10) 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 $F = \overline{F}$ (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 from our theoretical results in hand, the computation of an equilibrium with associated outside option $F \in (V^A, \overline{F}]$ (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).[23] The algorithm then determines the dynamic equilibrium consumption allocation imposing convergence to the stationary ladder in finite (but potentially long) time. The key distinction between an arbitrary outside option F and $\overline{F}$ is that at the latter we know a) the stationary decay rate $\bar{g}$, b) that the associated stationary ladder is unique, and c) that the dynamic equilibrium consumption allocation converges to the stationary ladder asymptotically. We therefore focus on the $\overline{F}$ case in what follows.[24]

Figure 3 plots the dynamics of the equilibrium consumption allocation with $u(c) = \log(c)$, incomes are $(\ell, h) = (0.75, 1.25)$, and the discount factor is $\beta = .9$. Social capital is $\pi = \overline{\pi} = 0.41$ so that the value of the outside option is given by $F = \overline{F}$. Table 1 provides additional summary statistics for the allocation in this parameterization, as well as for alternative

[23] 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).

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

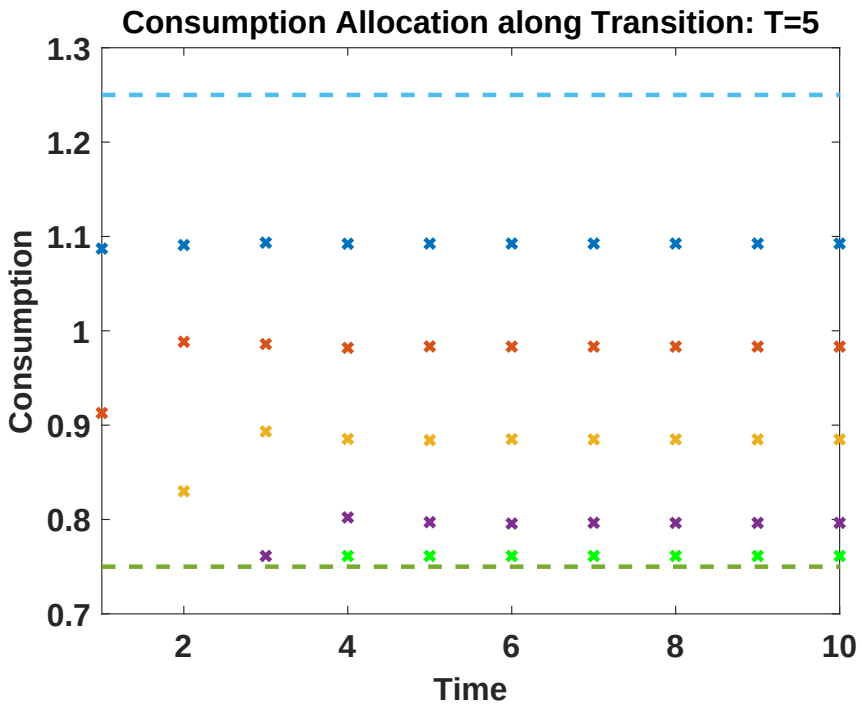

Figure 3: Consumption allocation along transition, with $\ell = 0.75$ (indicated by lower dashed horizontal line), $h = 1.25$ (upper dashed horizontal line), $\beta = 0.9$, $\pi = 0.41$, and $\gamma = 1$.

values of (β, γ) to display the comparative statics of the model with respect to its preference parameters (the values of $\overline{F}$ and $\overline{\pi}$ 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 more than implied by a binding outside option, $c_\ell(\overline{F})$. Over time those with continuously low income see their consumption drift down until the outside option binds and $c = c_\ell(\overline{F})$. 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 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

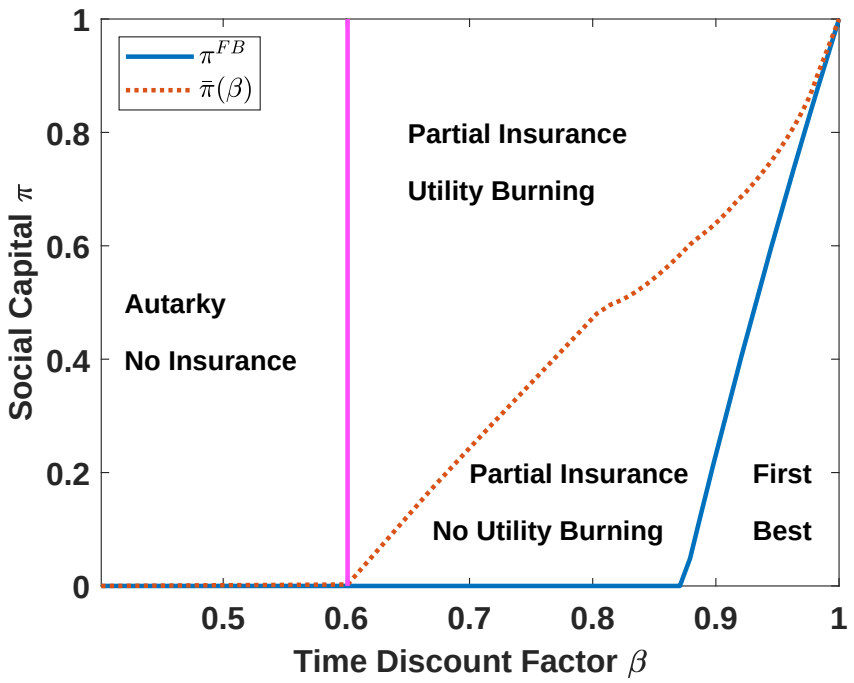

Figure 4: Insurance possibilities as a function of (β, π), for $h = 1.25$, $\ell = 0.75$, $u = \log$.

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 examples suggest that numerically convergence occurs very rapidly.

Figure 4 plots, for the same utility function and possible values of income, the computed counterpart of Figure 2. It demonstrates that for discount factors $\beta > \underline{\beta}$ the equilibrium changes qualitatively as social capital π increases. Take $\beta = 0.9$; for low values $\pi \leq \pi^{FB}$ full insurance can be sustained, for intermediate values $\pi \in (\pi^{FB}, \bar{\pi}]$ there is partial risk sharing but no utility burning, and for $\pi > \bar{\pi}$ the equilibrium requires utility burning. Importantly, the numerical example shows that for all $\beta < 1$, the threshold $\overline{\pi}(\beta)$ is always less than one, a feature that we have robustly found through many parameterizations we have explored.

Table 1 contains summary statistics of equilibrium allocations along the transition for alternative parameterizations of the model. Focus 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

Statistic	$\gamma = 1$		$\gamma = 2$	
	$\beta = 0.9$	$\beta = 0.95$	$\beta = 0.9$	$\beta = 0.95$
$V^{FB}/V(\overline{F})$ in %	0.63%	0.22%	0.45%	0.12%
$V(\overline{F})/\overline{F}$ in %	0.94%	0.71%	1.24%	0.80%
$\overline{\pi}$	0.41	0.66	0.69	0.85
$c_\ell(\overline{F})$	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

Table 1: Summary Statistics of the Transition

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.

over time as expected period utility falls and consumption dispersion rises over time.[25] As households become more patient (higher β) and more risk-averse (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.

9 Model Extensions

In this section we discuss two extensions to our simple model. In the first we consider a more general model of temporary delays to agreement after an initial failure to successfully form a coalition. In the second we extend out model to allow for production.

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

9.1 Temporary Delay

We have assumed that a deviating coalition succeeds with probability π and is in permanent autarky with complementary probability. We now assume that a failure to form a coalition is followed by $T \geq 1$ periods of autarky before another attempt can be made (so that if $T = 1$, a new attempt can be made in the next period after a failure). Under this assumption, after a deviation, coalition formation always eventually occurs. For fixed social capital π a reduction of T increases the outside option. We now argue that the extension with a delay of T is equivalent to our original model with social capital

$$\pi^\dagger := \frac{\pi}{1-(1-\pi)\beta^T}.$$

Suppose $\mathbb{c}^\dagger$ is an equilibrium allocation in the model with T-period delay. Then, the value of the outside option after deviating satisfies

$$W^d = \pi W^0(\mathbb{c}^\dagger) + (1-\pi)[(1-\beta^T)V^A + \beta^T W^d],$$

that is,

$$W^d = \pi^\dagger W^0(\mathbb{c}^\dagger) + (1-\pi^\dagger)V^A.$$

It is easy to verify that since $\mathbb{c}^\dagger$ is an equilibrium allocation in the model with T-period delay, it must also be an equilibrium allocation in our original model for social capital $\pi^\dagger$.

With finite exclusion, all agents are eventually in a risk-sharing arrangement, irrespective of the level of social capital. However, the level of risk-sharing is declining in social capital, which accords well with the empirical literature that finds differences in consumption risk-sharing across developing and developed countries (recall our discussion in the Introduction and Related Literature sections).

9.2 Risk Sharing and Production

We now briefly discuss how to extend our model to a production economy where output is produced and consumption is allocated within coalitions we will call production clubs, or firms for short. Output y_t produced by individual at time t depends upon idiosyncratic productivity $e_t \in E = \{e_\ell, e_h\}$ and labor effort l_t.

$$y_t = e_t l_t$$

Individual preferences are given by

$$(1-\beta)\mathbb{E}\Big\{\sum_{t=1}^{\infty}\beta^t U(c_t,l_t)\Big\},$$

and labor effort is bounded by the unit interval, so $l_t \in [0,1]$. All other aspects of the environment are the same as in the endowment economy studied thus far.

As before, risk-sharing incentives lead to continuum-sized firms being efficient, just as in our endowment economy. Since this implies that there is no aggregate output risk within a firm, an allocation within a continuum-sized firm are sequences of consumption and labor effort, both functions of the individual productivity history, $\{c_t(e^t), l_t(e^t)\}$.

In the special case in which labor is inelastically supplied at 1, and preferences are separable in consumption and labor, the equilibrium of our model becomes essentially the same as in the endowment case, with endowment income $y \in \{\ell, h\}$ replaced by production income $y \in \{e_\ell \times 1, e_h \times 1\}$. This is the content of the next proposition.

Proposition 8 *Suppose flow utility $U(c_t,l_t) = u(c_t) - v(l_t)$ is separable between consumption and labor, $(e_\ell, e_h) = (\ell, h)$, and $u'(y_h)y_\ell \geq v'(1)$.*

1. *There exists an equilibrium with a consumption allocation that is identical to that in the endowment economy with $c(e^t) = c(y^t)$ and labor equal to $l_t(e^t) = 1$.*

2. *The equilibrium payoff to forming a firm is the same as in the coalition payoff in the endowment economy, net of the cost of labor effort:*

$$(1-\beta)\mathbb{E}\Big\{\sum_{t=1}^{\infty}\beta^t[u(c_t(e^t)) - v(l_t(e^t))\Big\} = W^0(c) - v(1).$$

3. *The largest probability of successfully forming a firm for which there is a fixed point equilibrium is still $\overline{\pi}$ from the endowment economy, however the associated highest feasible outside option is $\overline{F} - v(1)$.*

This proposition follows from the fact that the rankings of consumption sequences is unaffected by subtracting a constant labor cost in each period. For $\pi > \overline{\pi}$ utility-burning needs to occur in equilibrium, and while this can be done just as in the endowment case, richer possibilities involving the labor allocation emerge in the production economy.

The key to the previous proposition is that the within-firm consumption-labor allocation can be solved sequentially. In a first step the optimal labor allocation is determined, and in a second step the consumption risk-sharing allocation is chosen, taking as given the stochastic

income process from the first stage. For a general utility function where labor is interior both consumption and labor are determined jointly.

An exception are utility functions without income effects on labor supply. For example, suppose households have Greenwood, Hercowitz, and Huffman (1988) preferences of the form

$$U(c,l) = \frac{1}{1-\gamma}\left\{c - \Psi\frac{l^{1+\theta}}{1+\theta}\right\}^{1-\gamma}$$

then the optimal labor allocation is determined by $l_t(e^t) = (e_t/\Psi)^{1/\theta}$ if Ψ is sufficiently large relative to e_t so that $l_t(e^t) < 1$. Now idiosyncratic income is given as $y(e^t) = \frac{e_t^{1+1/\theta}}{\Psi^{1/\theta}}$ and is efficiently shared within the firm as before, leading to a consumption allocation similar to the endowment economy. However, now we need to adjust the payoffs to take account of the differential labor utility costs. For example, the decay condition (10) in Proposition 5 becomes

$$\frac{\left(c(e^t) - (e_t/\Psi)^{(1+\theta)/\theta}\right)^{-\gamma}}{\left(c(e^{t+1}) - (e_{t+1}/\Psi)^{(1+\theta)/\theta)}\right)^{-\gamma}} = \delta_{t+1}.$$

Finally, it is easy to accommodate the notion that firms can realize increasing returns to scale, up to a point, in the size of its workforce, and that the production coalitions we model partially form not only for risk sharing purposes, but also for production efficiency purposes. Suppose that individual output within a firm is now given by

$$y_t = ze_tl_t$$

where $z = z(x)$ is a positive and weakly increasing function of the size x of the workers of the firm, with $z(x) = 1$ for $x \geq X$. That is, for firms larger than size $X < \infty$, which include those with an infinite number or a continuum of members, $z(x) = 1$. When $z(0) < 1$, then producing in autarky involves not only a loss in consumption smoothing but also a reduction in productivity. This again leads to a consumption allocation that has the same characteristics as in the endowment economy, but with a reduction in the value of autarky. With period utility that is separable and CRRA in consumption the utility from autarky is scaled to $u(z(0))V^A(y)$.[26] Scaling down the utility from autarky raises π^{FB} and $\overline{\pi}$, the social

[26]If the disutility of labor such that it is always efficient to supply a unit of labor in autarky for all levels of idiosyncratic productivity, then this simply shifts down the autarky payoff in the production economy relative to the endowment economy and is given by

$$(1-\beta)[u(z(0)y) - v(1)] + \beta\mathbb{E}_{y'}[u(z(0)y') - v(1)] = u(z(0))V^A(y) - v(1).$$

capital at which first-best insurance can be sustained and the threshold social capital for which the fixed-point equilibrium exists and utility burning is unnecessary. Thus, while the qualitative features of the analysis are unaffected by productivity benefits of large coalitions, quantitatively such production coalitions can provide better insurance when formed.

Our model of production clubs can qualitatively account for a number of well known features of the data. In the context of the literature on social capital, Fukuyama (1995, p. 309, 312) asserts that while "there continues to be a steady proliferation of interest groups of all sorts in American life ... communities of shared values whose members are willing to subordinate their private interests for the sake of larger goals of the community ... have become rarer." This is consistent with the prediction of our model that more coalitions forming goes hand in hand with shallower cooperation within coalitions. On the issue of risk sharing within a firm, Guiso, Pistaferri, and Schivardi (2005) find that while temporary shocks are well-insured, permanent ones are not. This is consistent with our model, since a permanent shock to a worker's income would rescale their outside option and hence lead to a permanently different consumption ladder.[27]

10 Conclusion

In this paper we have proposed a model in which social capital facilitates the formation of efficiency-enhancing risk-sharing or production coalitions as well as coalitional deviations from these original arrangements. The symmetric treatment of initial and deviating coalitions, both with respect to the allocation chosen and the composition of the group, ties together tightly the ex ante payoff and the outside option. This tight link implies that as our notion of social capital, π, increases, these two payoffs rise together. The strength of this effect eventually becomes so large (as π rises towards 1) that the standard notion of equilibrium as a fixed point in the value of forming a coalition familiar from the limited commitment literature ceases to exist. We propose an expanded notion of equilibrium which encompasses fixed point equilibria when they exist. The double-edged aspect of π, making it easier to form both initial and deviating coalitions, leads to the differential impact of a higher π on ex ante utility (which is weakly increasing), and ex post utility conditional on formation as well as the steady state distribution of continuation payoffs (which are weakly decreasing in π). Moreover, at high degrees of social capital, when fixed-point equilibria cease to exist, our expanded equilibrium concept exhibits utility burning as necessary feature.

[27]With homothetic preferences, a permanent multiplicative shock to productivity for a (positive measure) subset of agents would simply scale these agents' consumption allocation by the permanent shock, since these agents with the positive shock can always secede and guarantee themselves the scaled consumption process.

The comparative statics with respect to π exhibit three regions. With a low probability of forming a coalition, ex ante welfare is linearly increasing in π and conditional on coalition formation, members receive complete insurance. At an intermediate range ex ante welfare is increasing in π but at a decreasing rate and conditional on coalition formation, insurance is incomplete and declining in π. Allocations feature wasteful inequality but are intertemporally efficient. With high levels of social capital, ex ante welfare is flat in π, and allocations feature significant inefficiencies, manifested in utility or resource burning within a coalition to prevent defections. In a nutshell, an increase in π enables groups to more readily *trust each other* by agreeing on Pareto improving exchanges but at the same time making this *trust shallower*.

References

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

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

Altonji, Joseph G., Fumio Hayashi, and Laurence J. Kotlikoff (1997), "Parental altruism and inter vivos transfers: Theory and evidence." *Journal of Political Economy*, 105, 1121–1166.

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

Attanasio, Orazio and Steven J. Davis (1996), "Relative wage movements and the distribution of consumption." *Journal of Political Economy*, 104, 1227–1262.

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

Bernheim, B. Douglas and Debraj Ray (1989), "Collective dynamic consistency in repeated games." *Games and Economic Behavior*, 1, 295–326.

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

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

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

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

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

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

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

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

Fukuyama, Francis (1995), *Trust: The social virtues and the creation of prosperity*. Free Press, New York.

Fukuyama, Francis (2001), "Social capital, civil society and development." *Third World Quarterly*, 22, 7–20.

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

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

Golosov, Mikhael and Aleh Tsyvinski (2006), "Designing optimal disability insurance: A case for asset testing." *Journal of Political Economy*, 114, 257–279.

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

Greenwood, Jeremy, Zvi Hercowitz, and Gregory W. Huffman (1988), "Investment, capacity utilization, and the real business cycle." *American Economic Review*, 402–417.

Guiso, Luigi, Luigi Pistaferri, and Fabiano Schivardi (2005), "Insurance within the firm." *Journal of Political Economy*, 113, 1054–1087.

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

Hellwig, Christian and Guido Lorenzoni (2009), "Bubbles and self-enforcing debt." *Econometrica*, 77, 1137–1164.

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.

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

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

Knack, Stephen and Philip Keefer (1997), "Does social capital have an economic payoff? a cross-country investigation." *Quarterly Journal of Economics*, 112, 1251–1288.

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

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

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

Krueger, Dirk and Fabrizio Perri (2006), "Does income inequality lead to consumption inequality? Evidence and theory." *Review of Economic Studies*, 73, 163–193.

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

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

Ligon, Ethan, Jonathan P. Thomas, and Tim Worrall (2002), "Informal insurance arrangements with limited commitment: Theory and evidence from village economies." *Review of Economic Studies*, 69, 209–244.

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

Park, Yena (2014), "Optimal taxation in a limited commitment economy." *Review of Economic Studies*, 81, 884–918.

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

Portes, Alejandro (1998), "Social capital: Its origins and applications in modern sociology." *Annual Review of Sociology*, 24, 1–24.

Portes, Alejandro and Patricia Landolt (1996), "The downside of social capital." *The American Prospect*, 26, 18–22.

Putnam, Robert D. (1993), "What makes democracy work?" *National Civic Review*, 82, 101–107.

Putnam, Robert D. (2000), *Bowling Alone: The Collapse and Revival of American Community*. Simon and Schuster.

Putnam, Robert D. (2001), "Social capital: Measurement and consequences." *Canadian Journal of Policy Research*, 2, 41–51.

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

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

Townsend, Robert M. (1994), "Risk and insurance in village india." *Econometrica*, 62, 539–591.

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

Woolcock, Michael (1998), "Social capital and economic development: Toward a theoretical synthesis and policy framework." *Theory and Society*, 27, 151–208.

Woolcock, Michael (2001), "The place of social capital in understanding social and economic outcomes." *Canadian Journal of Policy Research*, 2, 11–17.

Woolcock, Michael and Deepa Narayan (2000), "Social capital: Implications for development theory, research, and policy." *The World Bank Research Observer*, 15, 225–249.

Appendices

A Proofs for Section 6

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 $F \leq \overline{F}$.*
2. *$\mathcal{C}(F)$ is closed and convex for all $F \leq \overline{F}$.*
3. *$\mathcal{C}$ is a continuous correspondence at all $F \leq \overline{F}$ (at $\overline{F}$, 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 $c^\dagger \in \mathcal{C}(\overline{F})$. 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 $\alpha_k = (F_k - F)/(\overline{F} - F_k) > 0$. Then,

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

   and so incentive feasibility (7) 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 $\overline{F} > V^A$. This also shows that $\mathbb{V}(V^A) > V^A$.

2. Suppose $(F_k) \nearrow \overline{F}$ 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 (7) at $F = \overline{F}$), and so $c_\infty \in \mathcal{C}(\overline{F})$, and $\mathcal{C}(\overline{F}) \neq \varnothing$.

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

4. The function $p : [V^A, \overline{F}] \to [0, \overline{\pi}]$ 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 $\pi \in (0, \overline{\pi}]$, the fixed point is given by $F(\pi) := \pi V^*(p^{-1}(\pi)) + (1-\pi)V^A$. The remaining claims are immediate.

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

□

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

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

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

where $\zeta^{FB} = h - \bar{y}$, we have $c_\zeta \in \mathcal{C}(F)$, and so $\overline{F} > F^{FB}$.

By the definition of F^{FB},

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

and so

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

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

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

and so

$$\begin{aligned} W(h, c_\zeta) =& W(h, c^{FB}) - (1-\beta)u'(\bar{y})(\zeta - \zeta^{FB}) + o((\zeta - \zeta^{FB})^2) \\ =& W(h, c^{FB}) + (\zeta^{FB} - \zeta)[(1-\beta)u'(\bar{y}) + o((\zeta - \zeta^{FB})^2)/(\zeta - \zeta^{FB})]. \end{aligned}$$

For $\zeta^{FB} - \zeta < \xi'$, where $\xi' > 0$ is a sufficiently small constant, the magnitude of the last term is less than $(1-\beta)u'(\bar{y})/2$, and so

$$W(h, c_\zeta) > W(h, c^{FB}) + (\zeta^{FB} - \zeta)(1-\beta)u'(\bar{y})/2.$$

For $F = F^{FB} + (\zeta^{FB} - \zeta)(1-\beta)u'(\bar{y})/(2\beta)$ (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. □

In the next lemma, an allocation is π-internally-incentive feasible if it satisfies the internal-incentive feasibility constraint (2) at the value π.

Lemma A.3 *Define the allocation $c_{\varepsilon,\alpha}$ by*

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

Define $\underline{\beta} := u'(h)/u'(\ell)$. For all $\pi > 0$, there exists $\eta > 0$, such that for all $\beta \in [\underline{\beta}, \underline{\beta}+\eta]$, all $\varepsilon' \in (0,\eta)$, and all $\alpha \in [1,2]$, the allocation $c_{\varepsilon,\alpha}$ is not π-internally-incentive feasible.

Proof. We first calculate the values of the allocation $c_{\varepsilon,\alpha}$ after different histories, where we simplify notation by writing $c_h = h-\varepsilon$, $c'_\ell = \ell+\alpha\varepsilon$, and $c''_\ell = \ell+(2-\alpha)\varepsilon$:

$$\begin{aligned} W(h, c_{\varepsilon,\alpha}) &= (1-\beta)u(c_h) + \tfrac{\beta}{2}(W(h, c_{\varepsilon,\alpha}) + W(h\ell, c_{\varepsilon,\alpha})), \\ W(h\ell, c_{\varepsilon,\alpha}) &= (1-\beta)u(c'_\ell) + \tfrac{\beta}{2}(W(h, c_{\varepsilon,\alpha}) + W(\ell\ell, c_{\varepsilon,\alpha})), \\ W(\ell\ell, c_{\varepsilon,\alpha}) &= (1-\beta)u(c''_\ell) + \tfrac{\beta}{2}(W(h, c_{\varepsilon,\alpha}) + W(\ell\ell, c_{\varepsilon,\alpha})), \\ \text{and} \qquad W(\ell, c_{\varepsilon,\alpha}) &= (1-\beta)u(2\bar{y}-c_h) + \tfrac{\beta}{2}(W(h, c_{\varepsilon,\alpha}) + W(\ell\ell, c_{\varepsilon,\alpha})). \end{aligned}$$

Hence,

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

and so

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

Thus,

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

which implies

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

that is,

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

A necessary condition for $c_{\varepsilon,\alpha}$ to be π-internally-incentive feasible is

$$f(\varepsilon;\beta) := W(h, c_{\varepsilon,\alpha}) - (1-\beta)u(h) - \frac{\beta}{2}\overline{\pi}\left[W(h, c_{\varepsilon,\alpha}) + W(\ell, c_{\varepsilon,\alpha})\right] - \beta(1-\overline{\pi})V^A \geq 0.$$

Note that for all β, $f(0;\beta) = 0$. We now argue that there exists $\eta > 0$, such that for all $\beta \in [\underline{\beta}, \underline{\beta} + \eta]$ and all $\varepsilon' \in (0, \eta)$, $\partial f(\varepsilon';\beta)/\partial\varepsilon < 0$, implying

$$f(\varepsilon';\beta) < 0 \qquad \forall \beta \in [\underline{\beta}, \underline{\beta} + \eta],\ \varepsilon' \in (0, \eta).$$

Recalling our definition of $c_{\varepsilon,\alpha}$ and differentiating (A.3) with respect to ε,

$$\frac{\partial}{\partial\varepsilon}W(h, c_{\varepsilon,\alpha}) = \frac{1}{4}\left\{-2(2-\beta)u'(c_h) + \beta(2-\beta)\alpha u'(c'_\ell) + \beta^2(2-\alpha)u'(c''_\ell)\right\}.$$

Evaluating this expression at $\beta = \underline{\beta} = u'(h)/u'(\ell)$ and $\varepsilon = 0$ yields (for any $\alpha \in [1, 2]$)

$$\tfrac{1}{4}u'(\ell)\{-2(2-\underline{\beta})\underline{\beta} + \underline{\beta}(2-\underline{\beta})\alpha + \underline{\beta}^2(2-\alpha)\} \leq 0,$$

and so there exists η' such that for all $\beta \in [\underline{\beta}, \underline{\beta} + \eta']$ and all $\varepsilon' \in (0, \eta')$,

$$\frac{\partial}{\partial\varepsilon}W(h, c_{\varepsilon,\alpha}) < \frac{\pi\underline{\beta}}{(2-\pi\underline{\beta})}\frac{1}{3}[u'(\ell) - u'(h)].$$

Turning to $W(\ell, c_{\varepsilon,\alpha})$, we have

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

Note that $(1-\beta)u'(\ell+\varepsilon) = u'(\ell) - u'(h)$ for $\beta = \underline{\beta}$ and $\varepsilon = 0$. Moreover, the term in $\{\cdot\}$ is nonnegative at $\beta = \underline{\beta}$ and $\varepsilon = 0$. Thus, there exists η'' such that for all $\beta \in [\underline{\beta}, \underline{\beta} + \eta'']$ and all $\varepsilon' \in (0, \eta'')$,

$$\frac{\partial}{\partial\varepsilon}W(\ell, c_{\varepsilon,\alpha}) > \frac{2}{3}[u'(\ell) - u'(h)]. \tag{A.4}$$

Taking $\eta' = \min\{\eta', \eta''\}$, we thus have for all $\beta \in [\underline{\beta}, \underline{\beta} + \eta]$ and all $\varepsilon' \in (0, \eta)$,

$$\partial f(\varepsilon';\beta)/\partial\varepsilon < 0.$$

This implies that for the specified bounds on β and ε, the allocation $c_{\varepsilon,\alpha}$ for *any* value of $\alpha \in [1,2]$ is not π-internally incentive feasible.

□

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 internal-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 incentive feasibility binds in the initial period, then

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

which is impossible, since $F \leq F(\overline{\pi})$ implies $\mathbb{V}(F) \geq F$.

Suppose 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 period t_n, the number of ℓ realizations before the ℓ-incentive constraint binds goes to ∞ as n goes to infinity. 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 $(\log \ell - \log h)/\log(1-\varepsilon)$, 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 7.2

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 $\overline{F}$.

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

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

where $\hat{c}(\ell^t)$ satisfies

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

By construction, c^F satisfies resource feasibility, and incentive feasibility for any history ending in a realization of ℓ (since $\mathbb{c}_*$ satisfies (15), $\hat{c}(\ell^t) \geq c_\ell$).

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 $F \leq \overline{F}$.

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

Finally, we prove that if $F \leq \overline{F}$, then $\mathbb{V}^*(h;F) \geq W^F(h)$. We do this by proving that if $\mathcal{C}(F)$ is nonempty, then there is 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 Lemma 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 $y^t\ell^\tau$ and so, from Lemma B.7,

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

For $M \geq 1$, define the ladder $(c_k^M)_{k=0}^L$ (recall that $\Pr(y^t) = 2^{-t}$):

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

We claim that $(c_k^M)_k$ satisfies (15) (where we set $c_k^M = c_\ell$ for $k > L$):

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

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

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

Since $\mathbb{c}(y^{t-1}\ell) \geq c_\ell$, it is immediate that c_k^M satisfies (16). Thus, for each M, $c^M \in \mathcal{C}_*(F)$.

Since $(c_k^M)_k \in [0,h]^L$, a closed and bounded set, the sequence $\left((c_k^M)_k\right)_M$ has a convergent subsequence with limit $(c_k^*)_k$. We now argue that $W(h, c^*) \geq W^F(h)$, completing the proof of the Proposition.

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

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

Consequently, taking time averages

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

Since u is strictly concave,

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

and so

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

If the arguments of the utility function were c_k^M (which they are not), the proof would be done without the need to pass to the limit, since then the expression on the right hand side is simply $W(h,c^M)$.

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

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

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

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

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

For all $\varepsilon > 0$, there exists M_1^ε such that if $M > M_1^\varepsilon$, for all $k = 0, \ldots, L$ the upper bound of $Lh/(M+1)$ on consumption divergences is sufficiently small that the right side of (C.4) is within ε of $W(h, c^M)$, implying $W^F(h) < W(h, c^M) + \varepsilon$. Moreover, there exists M_2^ε such that for all $M > M_2^\varepsilon$, $|W(h, c^*) - W(h, c^M)| < \varepsilon$. So, for $M > \max\{M_1^\varepsilon, M_2^\varepsilon\}$, we have

$$\begin{aligned} W^F(h) &< W(h, c^*) - W(h, c^*) + W(h, c^M) + \varepsilon \\ &< W(h, c^*) + 2\varepsilon. \end{aligned}$$

Since this holds for all $\varepsilon > 0$, we have

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

completing the proof. □

Proof of Corollary 2. We first argue that, for β larger than but near $\underline{\beta}$, the stationary ladder solving (14) for $F = \overline{F}$ has length 2 (which will allow us to use Lemma A.3): If the ladder is 3 or longer, then the consumption lower bound after realizations ℓ and $\ell\ell$ is not binding, and so

$$u'(\bar{c}_*(h)) = \beta u'(\bar{c}_*(h\ell)) = \beta^2 u'(\bar{c}_*(h\ell\ell)).$$

But for β close to $\underline{\beta}$, $c_*(h\ell)$ and $c_*(h\ell\ell)$ are both close to ℓ, and so $u'(\bar{c}_*(h\ell)) \approx u'(\bar{c}_*(h\ell\ell))$, implying β is close to 1, a contradiction.

Let $c^{\overline{F}}$ denote the allocation defined in (C.1) using the stationary ladder for $\overline{F}$. While we do not explicitly indicate the dependence of $\overline{F}$ and so $c^{\overline{F}}$ on β, both objects will vary with β: For β close to $\underline{\beta}$, the allocation $c^{\overline{F}}$ is given by $c_{\varepsilon,\alpha}$, the allocation defined in Lemma A.3, for an appropriate choice of ε and $\alpha \in [1, 2]$. Moreover, ε converges to 0 as β tends to $\underline{\beta}$.

For each $\pi > 0$, denote by $\eta(\pi) > 0$ the η-bound identified in Lemma A.3 (note that $\eta(1)$ is a nondecreasing function of π). There then exists $\eta'''(\pi) > 0$ such that for $\beta \in$

$[\underline{\beta}, \underline{\beta} + \eta'''(\pi)]$, the ε associated with $c^{\overline{F}}$ is smaller than $\eta(\pi)$. This implies that for $\beta \in [\underline{\beta}, \underline{\beta} + \min\{\eta(\pi), \eta'''(\pi)\}]$, $c^{\overline{F}}$ cannot be π-internally-incentive feasible.

We first prove that $\overline{\pi}(\beta) < 1$ for β close to $\underline{\beta}$. For if not, then for β close to $\underline{\beta}$,

$$\mathbb{V}(\overline{F}) \leq \overline{F}.$$

But this implies $c^{\overline{F}}$ is π-internally incentive feasible for $\pi = 1$ (and so for any smaller π):

$$\begin{aligned} W(y, c^{\overline{F}}) &\geq (1-\beta)u(y) + \beta\overline{F} \\ &\geq (1-\beta)u(y) + \beta\mathbb{V}(\overline{F}) \\ &\geq (1-\beta)u(y) + \beta W^0(c^{\overline{F}}), \end{aligned}$$

which we have just seen is impossible for all $\beta \in [\underline{\beta}, \underline{\beta} + \min\{\eta(1), \eta'''(1)\}]$.

If

$$\mathbb{V}(\overline{F}) > \overline{F},$$

then $\overline{\pi}(\beta) < 1$, and we again have that the allocation $c^{\overline{F}}$ is $\overline{\pi}(\beta)$-internally-incentive efficient:

$$\begin{aligned} W(y, c^{\overline{F}}) &\geq (1-\beta)u(y) + \beta\overline{F} \\ &= (1-\beta)u(y) + \beta\{\overline{\pi}\mathbb{V}(\overline{F}) + (1-\overline{\pi})V^A\} \\ &\geq (1-\beta)u(y) + \beta\{\overline{\pi}W^0(c^{\overline{F}}) + (1-\overline{\pi})V^A\}. \end{aligned}$$

This completes the argument, since for any fixed $\pi > 0$, for β sufficiently close to $\underline{\beta}$, $c^{\overline{F}}$ is not π-internally-incentive feasible. □

The proof of Proposition 7 is broken into several lemmas.

Lemma C.1 *Suppose $F = \overline{F}$. The equilibrium allocation $\mathbb{c}$ converges to the unique solution to problem (14), $\bar{\mathbb{c}}_*$, 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

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

Proof. Resource feasibility in period t is

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

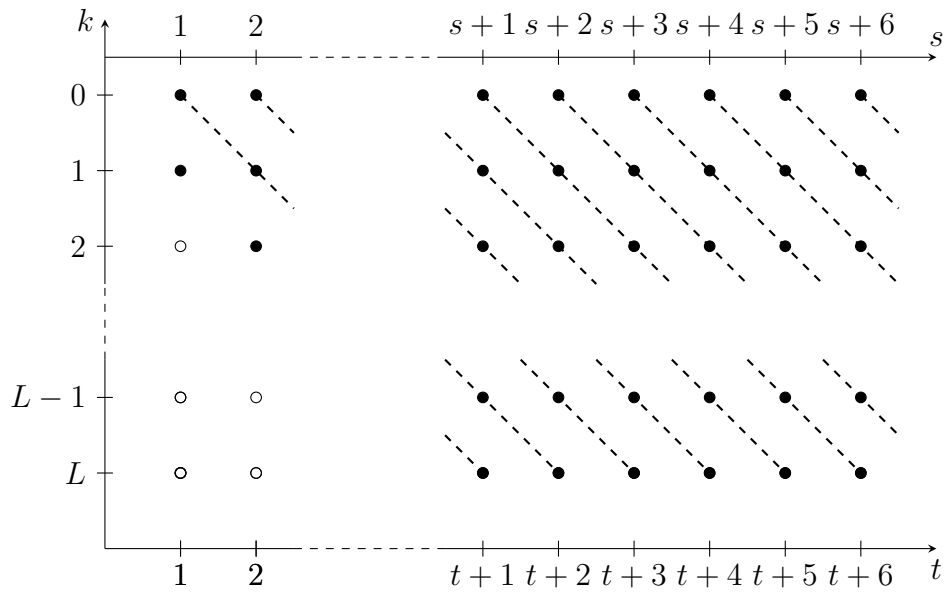

Figure C.1: Illustrating the re-indexing for the proof (the re-indexing omits $k = L$, since $c_t(\ell^L) = c_\ell(F)$ is determined by Lemma B.7). The ladder-s resource constraints sum over the diagonal dashed lines, while the period-t resource constraints sum vertically. Since there is at most one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

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

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

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 $\mathbb{V}^*(h, \overline{F})$ 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 (14).

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 (14) (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 *Suppose utility is CRRA. There exists $\beta_* \in (\underline{\beta}, 1)$, such that for all $\beta > \beta_*$, the equilibrium consumption allocation for $F = \overline{F}$ 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 (14) 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))

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

Let L be the length of the ladder, so that

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

There exists $\beta_* \in (\underline{\beta}, 1)$, such that for all $\beta > \beta_*$, $L \geq 3$ (since $c_\ell(\overline{F})$ is bounded away from $\bar{y}$).

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,

$$c_{t+1} = gc_t + 2^t \left\{ -g2^{-L}\bar{c}_\ell - 2^{-L}g^L\bar{c}_h + 2^{-L+1}\bar{c}_\ell \right\}.$$

Finally, since

$$-g2^{-L}\bar{c}_\ell - 2^{-L}g^L\bar{c}_h + 2^{-L+1}\bar{c}_\ell = 2^{-L}\bar{c}_\ell(1-g) + 2^{-L}(\bar{c}_\ell - g^L\bar{c}_h) > 0,$$

we have

$$c_{t+1} > gc_t,$$

and so

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

Consider now the following local change (which is feasible, since $L \geq 3$):

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

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 $F = \overline{F}$ 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 (10), 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 8

In this section we provide the details of how we compute equilibria in Section 8 of the main text. Section D.1 describes how to compute a stationary ladder that delivers an outside option $F \in (V^A, \overline{F})$. Section D.2 describes how to determine the value of $\overline{F}$ 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 $F \in (V^A, \overline{F})$, 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 (11), 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 (13))

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

   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 $F = \overline{F}$, as we have seen in Section 7.2. 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 $\overline{F}$. 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 $F < \overline{F}$ 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 $F < \overline{F}$. 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 $\overline{F}$

To determine $\overline{F}$ we proceed as follows: At $F = \overline{F}$, Proposition 6 implies that there is a unique stationary ladder satisfying h-incentive feasibility and this ladder solves (14), so we know that the consumption decay rate is given by

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

In effect, $\overline{F}$ is the peak of the $F(\cdot)$ map discussed above, and is reached at $g = \bar{g}$. Since the value of $\overline{F}$ itself is unknown, we have to determine the lower consumption floor $c_\ell = c_\ell(\overline{F})$

jointly with $\overline{F}$, $c_*(h)$, and L. The relevant equations, with $g = g(\overline{F}) = \beta^{1/\gamma}$ are

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

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

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

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

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

1. Guess $\overline{F} \in (V^A, V^{FB})$.
2. For a given $\overline{F}$:
   (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 $\overline{F}$ 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 $F \leq \overline{F}$. 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.[28] 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 $t = T+1, ..., T+L(F)$

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

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.[29] 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:

   (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 $\{c_{i-1,t-1}\}$.

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

(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 $(\mathbb{c}_t(h))_{t=1}^T$.