Social Insurance, Information Revelation, and Lack of Commitment*

Mikhail Golosov
Princeton

Luigi Iovino
Bocconi

March 2016

Abstract

We study the optimal provision of insurance against unobservable idiosyncratic shocks in a setting in which a benevolent government cannot commit. A continuum of agents and the government play an infinitely repeated game. Actions of the government are constrained by the threat of reverting to the worst perfect Bayesian equilibrium (PBE). We construct a recursive problem that characterizes the allocation of resources and the revelation of information on the Pareto frontier of the set of PBE. We show that the amount of information revealed by an agent depends on the continuation utility with which he enters the period. Agents who enter the period with low continuation utility reveal no information about their current shocks and receive no insurance. Agents who enter the period with high continuation utility reveal precise information about their current shocks and receive "second best" insurance as in economies with perfect commitment by the government.

*Golosov's email: golosov@princeton.edu. Iovino's email: luigi.iovino@unibocconi.it. We thank Mark Aguiar, Fernando Alvarez, Manuel Amador, V.V. Chari, Hugo Hopenhayn, Ramon Marimon, Stephen Morris, Nicola Pavoni, Chris Phelan, Ali Shourideh, Chris Sleet, Pierre Yared, Sevin Yeltekin, Ariel Zetlin-Jones for invaluable suggestions and all the participants at the seminars at Bocconi, Brown, Carnegie Mellon, Chicago Fed, Columbia, Duke, EIEF, Georgetown, HSE, MEDS, Minnesota, Norwegian Business School, NY Fed, NYU, Paris School of Economics, Penn State, Philadelphia Fed, Princeton, UCLA, University of Lausanne, University of Vienna, Washington University, the SED 2013, the SITE 2013, the ESSET 2013 meeting in Gerzensee, 12th Hydra Workshop on Dynamic Macroeconomics, 2014 Econometric Society meeting in Minneapolis, EFMPL Worskhop at NBER SI 2015. Golosov thanks the NSF for support and the EIEF for hospitality. Iovino thanks NYU Stern and NYU Econ for hospitality. We thank Sergii Kiiashko and Pedro Olea for excellent research assistance.

1 Introduction

The major insight of the normative public finance literature is that there are substantial benefits from using past and present information about individuals to provide them with insurance against risk and incentives to work. A common assumption of the normative literature is that the government is a benevolent social planner with perfect ability to commit. Commitment power typically implies that the more information the planner has, the more efficiently she can allocate resources.[1]

The political economy literature has long emphasized that such commitment may be difficult to achieve in practice.[2] Over time self-interested politicians and voters – whom we will broadly refer to as "the government" – are tempted to re-optimize and choose new policies. When the government cannot commit the benefits from more precise information are less clear. As governments become more informed, they may allocate resources more efficiently – as in the conventional normative analysis – but they may also be tempted to depart from the ex-ante desirable policies. The analysis of such environments is difficult because the main analytical tool to study private information economies – the Revelation Principle – fails when the decision maker cannot commit.

In this paper we study optimal resource allocation and information revelation in a simple model of social insurance – the unobservable taste shock environment of Atkeson and Lucas (1992). This environment, together with closely related models of Green (1987), Thomas and Worrall (1990), Phelan and Townsend (1991), provides theoretical foundation for a lot of recent work in macro and public finance.[3] The key departure from that literature is the assumption that resources are allocated by a government that, although benevolent, lacks commitment. We study how information revelation affects the incentives of the government and characterize the properties of the optimal insurance contract.

[1] The seminal work of Mirrlees (1971) started a large literature in public finance on taxation, redistribution and social insurance in the presence of private information about individuals' types. Well known work of Akerlof (1978) on "tagging" is another early example of how a benevolent government can use information about individuals to impove efficiency. For the surveys of the recent literature on social insurance and private informaiton see Golosov, Tsyvinski, and Werning (2006) and Kocherlakota (2010).

[2] There is a vast literature in political economy that studies frictions that policymakers face. For our purposes, work of Acemoglu (2003) and Besley and Coate (1998) is particularly relevant who argue that inefficiencies in a large class of politico-economic models can be traced back to the lack of commitment. Kydland and Prescott (1977) is the seminal contribution that was the first to analyze policy choices when the policymaker cannot commit.

[3] This set up and its extensions are used in a variety of applications, such as the design of unemployment and disability insurance (Hopenhayn and Nicolini (1997), Golosov and Tsyvinski (2006)), life cycle taxation (Farhi and Werning (2013), Golosov, Troshkin, and Tsyvinski (2016)), human capital policies (Stantcheva (2014)), firm dynamics (Clementi and Hopenhayn (2006)), military conflict (Yared (2010)), international borrowing and lending (Dovis (2009)).

Our economy is populated by a continuum of atomless agents/citizens who are subject to privately observed taste shocks and by a benevolent government that allocates an endowment so as to insure the citizens against these shocks. Agents transmit information about their shocks to the government by sending messages. The government uses these messages to form posterior beliefs about the realization of agents' types and to allocate resources. The main friction is that ex-post, upon acquiring information about agents' types, the government is tempted to allocate resources differently from what agents require ex-ante to reveal information. In particular, the more precise the information that is available to the government, the higher its payoff if it decides to re-allocate resources.

To highlight the main mechanism underlying our results, we begin the analysis of a simple two period economy in which individuals receive idiosyncratic shocks only in period 1. A benevolent utilitarian government makes pre-election promises about how to allocate resources across individuals. After agents communicate their information, the government can pay a cost to break its pre-election promises and choose new allocations. We characterize agents' and government's strategies in perfect Bayesian equilibria (PBE) that maximize the weighted average of lifetime utilities of all agents. We take these Pareto weights as exogenous in the two period economy, but they emerge naturally in the infinitely repeated game through the dynamic provision of incentives.

When the cost of breaking promises is infinite this problem is isomorphic to usual principal-agent models. In that case, standard Revelation Principle arguments apply and all agents reveal full information about their shocks and receive second best insurance. Full information revelation is no longer optimal if the cost of breaking promises is sufficiently low. To study equilibria in such settings we first show how to rank agents' reporting strategies by their informativeness. We then show that, at the optimum, the informativeness of the agents' reports is monotone in the agents' Pareto weights: agents with higher weights reveal more precise information and receive better insurance. In addition, if an agent's weight is sufficiently high, he reveals full information about his type and receives second best insurance. On the contrary, if an agent's weight is sufficiently low, he reveals no information and receives no insurance. All other agents reveal some but not all information about their shocks. We also identify a class of economies in which insurance and information revelation takes a simple rationing rule: the government allocates second best insurance contracts to a random subset of citizens while the remaining agents receive no insurance.

We extend our analysis to an infinitely repeated game between a continuum of agents who are subject to idiosyncratic taste shocks in each period and a benevolent government who lacks commitment. In the Pareto optimal equilibria government's actions are sustained by a threat of switching to the worst PBE, in which no information is revealed to the government. We show how to characterize the optimal information revelation and insurance recursively, with each agent's continuation utility on the equilibrium path serving as a state variable that summarizes his past history. As in the perfect commitment case of Atkeson and Lucas (1992), insurance against a high realization of the taste shock in the current period is provided by lowering agent's continuation utility. As agents experience different histories of shocks, there is a distribution of continuation utilities at any given period.

Similarly to the two period model, the agent's continuation utility at the beginning of the period determines his optimal information revelation. Under quite general conditions agents who enter the period with low continuation utilities reveal no information about the realization of their shocks in that period and receive no insurance. In contrast, under some additional assumptions on the utility function and the distribution of shocks, agents who enter the period with high continuation utilities reveal their private information fully and receive second best insurance.

The intuition for this result comes from comparing benefits and costs of revealing information to the government. The benefits come from the fact that more precise information about an agent's idiosyncratic shock allows the government to deliver any given continuation utility at a lower cost *on the equilibrium path*. These benefits depend on the agent's continuation utility; more precise information about agents who enter the period with higher continuation utilities saves more resources. The costs emerge because the government is tempted to deviate from the ex-ante optimal plan and to re-optimize. When the government deviates from its equilibrium strategies, it reneges on all past promises and allocates consumption only on the basis of its posterior beliefs about the agents' current types. Therefore, the payoff that the government receives off the equilibrium path depends only on the total amount of information that was revealed and not on the identity of the agent who reveals it. For this reason it is optimal that agents with higher continuation utilities on the equilibrium path reveal more precise information about their shocks.

The threat of switching to the worst equilibrium also prevents the emergence of the extreme inequality, known as immiseration, which is a common feature of environments with commitment. In the invariant distribution continuation utilities of agents exhibit mean-reversion and any agent whose continuation utility falls into the no-insurance region exits it in finite time. Moreover, in the invariant distribution there is generally an endogenous reflecting lower bound on agents' continuation utilities.

An important technical contribution of our paper is to derive a recursive formulation for an optimal insurance problem when the principal cannot commit. The main difficulty that we need to overcome is that the government's payoff after a deviation depends on the reports made by all the agents. Since the information revealed by any agent affects government's incentives to renege on the implicit promises made to all other agents, we cannot directly rely on standard recursive techniques that characterize optimal insurance by focusing on each history of past shocks in isolation from other histories. We make progress by constructing an upper bound for the value of deviation with some key properties. First, the value of this upper bound is weakly higher than the value of deviation for all reporting strategies of the agents. This property implies that, if we replace the true value of deviation with its upper bound, the incentive constraint for the government will be tighter. Second, the value of the upper bound coincides with the value of deviation if all agents play the best PBE. This property implies that the best PBE is also a solution to the modified problem. Finally, this upper bound can be represented as a history-by-history integral of functions that depend only on the current reporting strategy of a given agent and, thus, the modified problem can be written recursively. The Bellman equation that we derive resembles the standard problems in the recursive contract literature with two modifications: (i) agents are allowed to choose mixed rather than pure strategies over their reports and (ii) there is an extra term in the planner's objective function capturing the "temptation" costs of receiving more informative reports.

Our paper is related to a relatively small literature on mechanism design without commitment. Roberts (1984) was one of the first to explore the implications of lack of commitment for social insurance. He studied a dynamic economy in which types are private information but do not change over time. More recently, Sleet and Yeltekin (2006), Sleet and Yeltekin (2008), Acemoglu, Golosov, and Tsyvinski (2010), Farhi, Sleet, Werning, and Yeltekin (2012) all studied versions of dynamic economies with idiosyncratic shocks closely related to our economy but made various assumptions on commitment technology and shock processes to ensure that any information becomes obsolete once the government deviates. In contrast, the focus of our paper is on understanding incentives to reveal information and their interaction with the incentives of the government. Our results about efficient information revelation are also related to the insights on optimal monitoring in Aiyagari and Alvarez (1995). In their paper the government has commitment but can also use a costly monitoring technology to verify the agents' reports. They characterize how monitoring probabilities depend on the agents' promised values. Although our environment and theirs differ in many respects, they both share the same insight that more information should be revealed by those agents for whom efficiency gains from better information are the highest. Bisin and Rampini (2006) pointed out that in general it might be desirable to hide information from a benevolent government in a two period economy.

In a broader context our work is also related to Skreta (2006) and Skreta (2015), who builds on earlier work of Bester and Strausz (2001), Freixas, Guesnerie, and Tirole (1985), Laffont and Tirole (1988), to study the optimal auction design in the settings in which the principal cannot commit. Essentially all that work focuses on the interaction between a principal and one agent, while our focus is on the insurance provided to a large number of agents. Our work is also related to Shimer and Werning (2015), who study the design of trading mechanism without commitment, and Cole and Kocherlakota (2001), who study dynamic games with hidden actions and states.

The rest of the paper is organized as follows. Section 2 studies optimal insurance and information revelation in a two period model. Section 3 describes our baseline infinite period economy with i.i.d. shocks. Section 4 extends our analysis to Markov shocks.

2 Information revelation in a simple model

In this section we consider a simple model of social insurance where a policymaker's ability to commit to her promises is imperfect. Our environment is a two period version of the Atkeson and Lucas (1992) set up. This economy allows us to transparently illustrate the main results and explain the intuition behind them. The main steps in the analysis extend to more general dynamic economies we consider in Section 3.

The economy lasts for two periods and is populated by a continuum of agents of measure 1 with preferences given by

$$\theta \frac{c_1^{1-\rho}}{1-\rho} + \frac{c_2^{1-\rho}}{1-\rho} \tag{1}$$

for $\rho > 0$. These preferences are understood to be $\theta \ln c_1 + \ln c_2$ when $\rho = 1$. Here c_t is consumption in period t and θ is an idiosyncratic shock. We assume that $\theta \in \Theta = \{\theta_L, \theta_H\}$ with $\theta_H > \theta_L > 0$. The probability of θ is $\pi(\theta)$ and we normalize $\sum_\theta \pi(\theta)\theta = 1$. The idiosyncratic shocks are private information. Each agent belongs to one of the groups $i = 1, ..., I$ for some $I \geq 1$. The measure of agents in group i is denoted by ψ_i. Group membership is observable but does not affect preferences, shocks or endowments.

The economy has one unit of non-storable endowment in each period. It is allocated by a benevolent government whose preferences are given by the average utility of all agents. To allocate consumption the government collects information from agents about their idiosyncratic shocks. Agents transmit information by sending messages from a message space M, where M is a finite set with more than one element.[4] The government allocates consumption as a function of agents' reports. Our focus is on understanding properties of optimal information revelation when government's ability to commit is imperfect. It will be more convenient to think of resource allocations not in terms of consumption units c but in terms of utils $u = \frac{c^{1-\rho}}{1-\rho}$. The resource cost of providing u utils is $C(u) = [(1-\rho)u]^{1/(1-\rho)}$ for $\rho \neq 1$, $C(u) = \exp(u)$ for $\rho = 1$. Let $\underline{v}$ and $\bar{v}$ the the greatest lower bound and the least upper bound on u.[5]

Formally, we consider the following three stage game. In stage 1 the government makes initial promises $u^{pr}_{i,t} : M \to \mathbb{R}$ for all i, t where $u^{pr}_{i,t}(m)$ is the allocation in period t to agent in group i who reports message m. In stage 2 agents report their types using symmetric strategies $\sigma_i : \Theta \to \Delta(M)$. We use $\sigma_i(m|\theta)$ to denote the probability of reporting message m for an agent in group i who had shock θ. We use Σ to denote the space of such strategies. By the law of the large numbers, $\sigma_i(m|\theta)$ is also the measure of agents in group i with shock θ who report m to the government. Finally, in stage 3 the government chooses a resource allocation function $u_{i,t} : M \to \mathbb{R}$ for all i, t.

The expectation of any variable $x : M \times \Theta \to \mathbb{R}$ is denoted by $\mathbb{E}_\sigma x = \sum_{(m,\theta) \in M \times \Theta} x(m,\theta)\,\sigma(m|\theta)\,\pi(\theta)$. For any message m sent with positive probability (i.e. $\sigma(m|\theta) > 0$ for some θ) we analogously define $\mathbb{E}_\sigma[x|m]$ using Bayes' rule. We use boldface letters without subscripts to denote the entire collection of strategies for all agents and dates, e.g. $\mathbf{u} = \{u_{i,t}\}_{i,t}$. Feasibility dictates that $\mathbf{u}$ must satisfy

$$\sum_{i=1}^{I} \psi_i \mathbb{E}_{\sigma_i} C(u_{i,t}) \leq 1, \quad \text{for all } t. \tag{2}$$

If $\mathbf{u} - \mathbf{u}^{pr}$ is not equal to zero for any positive mass of agents, the government incurs a utility cost $\Upsilon \geq 0$. We focus on the Pareto frontier of the set of Perfect Bayesian Equilibria (PBE), which for shortness we call *best PBE,* i.e. PBE for which there are no other PBE that give higher lifetime expected utility to all groups, with strict inequality for at least one group.

Before proceeding we want to make several remarks about our set up. Our two period model can be interpreted as a simple model of social insurance provided by a politician whose ability to commit to her pre-election promises is imperfect. Probabilistic voting models along the lines of Lindbeck and Weibull (1987) naturally lead politicians to promise, before elections, to pursue policies that maximize a weighted average of groups' utilities.[6] After the politician is elected, she can break those promises at a cost Υ and pursue policies that maximize her own

[4] The finiteness assumption is made only to simplify the notation; our results extend direct to any set M.

[5] In particular, $\underline{v} = 0$ if $\rho < 1$ and $\underline{v} = -\infty$ if $\rho \geq 1$; $\bar{v} = \infty$ if $\rho \leq 1$ and $\bar{v} = 0$ if $\rho > 1$.

[6] See Song, Storesletten, and Zilibotti (2012), Farhi, Sleet, Werning, and Yeltekin (2012), Scheuer and Wolitzky (2014) for applications to dynamic settings.

objective function.[7] An important special case of our model is $I = 1$, which corresponds to a benevolent government that maximizes the utility of ex-ante identical agents. As we show below, it is easier to characterize the efficient equilibrium by starting with a more general economy with heterogeneity.

The structure of our two period economy also closely resembles that of infinitely repeated games which we consider later in the paper. In such games both the cost of reneging on (implicit) promises and the heterogeneity captured by the groups I emerge naturally. Trigger strategies in repeated games are used to support efficient allocations and our parameter Υ captures the cost of switching to the worst equilibrium if the government deviates from equilibrium strategies. Heterogeneity emerges in repeated games because the need to provide incentives to reveal information in previous periods implies that agents enter the current period with different expected lifetime utilities.

We characterize best PBE of this game using backward induction. First consider the welfare that the government can attain if it receives reports $\boldsymbol{\sigma} = \{\sigma_i\}_i$ in stage 3 and pays cost Υ to re-optimize. Since the government is benevolent, it maximizes the sum of the agents' expected utilities conditional on the information revealed by $\boldsymbol{\sigma}$. The optimal choice of the government in period 1 is the solution to

$$\tilde{W}(\boldsymbol{\sigma}) \equiv \max_{\{u_i\}_i} \sum_{i=1}^{I} \psi_i \mathbb{E}_{\sigma_i} \theta u_i \tag{3}$$

subject to

$$\sum_{i=1}^{I} \psi_i \mathbb{E}_{\sigma_i} C(u_i) \leq 1. \tag{4}$$

Since there are no shocks in period 2, all agents receive the same consumption allocation and we use $\underline{U}$ to denote welfare in period 2.[8]

It is not efficient to break pre-election promises and, therefore, in any best PBE $\mathbf{u} = \mathbf{u}^{pr}$ and $(\mathbf{u}, \boldsymbol{\sigma})$ satisfies

$$\sum_{i=1}^{I} \psi_i \mathbb{E}_{\sigma_i} [\theta u_{i,1} + u_{i,2}] \geq \tilde{W}(\boldsymbol{\sigma}) + \underline{U} - \Upsilon. \tag{5}$$

Agents' equilibrium reporting strategies satisfy

$$\mathbb{E}_{\sigma_i} [\theta u_{i,1} + u_{i,2}] \geq \mathbb{E}_{\sigma_i'} [\theta u_{i,1} + u_{i,2}] \text{ for all } i, \sigma_i'. \tag{6}$$

[7] The assumption that the politician's objective function is utilitarian is immaterial for our analysis and was made to be consistent with the assumptions we make in Section 3. Our analysis extends directly to situations where the politician weighs members of different groups differently, for example, by giving higher weights to members of special-interest groups or members of her own party.

[8] Since the per capita endowment is 1, $\underline{U}$ is equal to 1 if $\rho < 1$, to 0 if $\rho = 1$, and to -1 if $\rho > 1$.

To characterize best PBE it is sufficient to find $(\mathbf{u}^*, \boldsymbol{\sigma}^*)$ that maximize a weighted average of the agents' lifetime utilities subject to (2), (5), and (6). Let $\{\zeta_t\}_t$ be the Lagrange multipliers on (2). It is easy to verify that $(\mathbf{u}^*, \boldsymbol{\sigma}^*)$ can be written as a solution to a dual cost minimization problem

$$\min_{\mathbf{u},\boldsymbol{\sigma}} \sum_{i,t} \psi_i \zeta_t \mathbb{E}_{\sigma_i} C\left(u_{i,t}\right) \tag{7}$$

subject to (5), (6), and

$$v_i = \mathbb{E}_{\sigma_i}\left[\theta u_{i,1} + u_{i,2}\right] \text{ for all } i, \tag{8}$$

where v_i is the lifetime utility of agents in group i in a best PBE. Let $\mathbf{v} \equiv (v_1, ..., v_I)$ be a point on the Pareto frontier of the set of PBE.

The direct characterization of problem (7) is difficult because $\tilde{W}(\boldsymbol{\sigma})$ is potentially a complicated function of the reports of all agents. This captures the fact that the information revealed by agents in group i affects the incentives of the government to break its pre-election promises and choose new allocations for agents in all groups. An important intermediate step of our analysis, which is also central to our recursive characterization in Section 3, is to study a modified dual problem in which the decision to re-optimize can be written as a function that is separable in the reports of each group.

Suppose $(\mathbf{u}^*, \boldsymbol{\sigma}^*)$ is a best PBE that delivers lifetime utilities $\mathbf{v}$ to agents and let λ^w be the Lagrange multiplier on the feasibility constraint (4) when $\boldsymbol{\sigma} = \boldsymbol{\sigma}^*$. Define a function $W : \Sigma \to \mathbb{R}$ by

$$W(\sigma) \equiv \max_u \mathbb{E}_\sigma \left[\theta u - \lambda^w C(u) + \lambda^w\right]. \tag{9}$$

Since (3) is a convex maximization problem, $\tilde{W}(\boldsymbol{\sigma})$ can be written as (see Luenberger (1969), Theorem 1, p. 224))

$$\begin{aligned} \tilde{W}(\boldsymbol{\sigma}) &= \min_{\lambda \geq 0} \max_{\{u_i\}_i} \sum_{i=1}^{I} \psi_i \mathbb{E}_{\sigma_i}\left[\theta u_i - \lambda C(u_i) + \lambda\right] \\ &\leq \max_{\{u_i\}_i} \sum_{i=1}^{I} \psi_i \mathbb{E}_{\sigma_i}\left[\theta u_i - \lambda^w C(u_i) + \lambda^w\right] = \sum_{i=1}^{I} \psi_i W(\sigma_i), \end{aligned} \tag{10}$$

with equality if $\boldsymbol{\sigma} = \boldsymbol{\sigma}^*$. The *modified dual problem* is the cost minimization problem (7) in which (5) is replaced with

$$\sum_{i=1}^{I} \psi_i \mathbb{E}_{\sigma_i}\left[\theta u_{i,1} + u_{i,2}\right] \geq \sum_{i=1}^{I} \psi_i W(\sigma_i) + \underline{U} - \Upsilon. \tag{11}$$

Lemma 1 *Let $(\mathbf{u}^*, \boldsymbol{\sigma}^*)$ be a solution to the dual problem (7). Then $(\mathbf{u}^*, \boldsymbol{\sigma}^*)$ is also a solution to the modified dual.*

Proof. The constraint set is smaller in the modified dual due to (10). Since $(\mathbf{u}^*, \boldsymbol{\sigma}^*)$ is a solution to the dual and lies in a constraint set of the modified dual, it must be also a solution to the modified dual. ■

Function W plays an important role in our analysis. Before describing its properties we define uninformative and fully informative strategies. We say that σ is *uninformative* if $\mathbb{E}_\sigma [\theta|m] = 1$ for all m and *fully informative* if for each m sent with positive probability there is $\theta_j \in \Theta$ such that $\mathbb{E}_\sigma [\theta|m] = \theta_j$. We use Σ^{un} and Σ^{in} to denote the set of uninformative and informative strategies, and σ^{un} and σ^{in} to denote elements of Σ^{un} and Σ^{in}. All uninformative strategies have the same value of $W(\sigma^{un})$ and all informative strategies have the same value of $W\left(\sigma^{in}\right)$.

Lemma 2 *W is continuous, convex and achieves its minimum (maximum) if and only if σ is uninformative (fully informative). Its solution u^w satisfies $C'(u^w(m)) = \mathbb{E}_\sigma [\theta|m] / \lambda^w$ for all m sent with positive probability. The derivative $\frac{\partial W(\sigma)}{\partial \sigma'} \equiv \lim_{\alpha \downarrow 0} \frac{W((1-\alpha)\sigma + \alpha\sigma') - W(\sigma)}{\alpha}$ exists for all σ, σ'.*

Proof. In the Appendix. ■

The key advantage of studying the modified dual is that the separability of (11) allows a simple characterization of the optimal information revelation. Let $\chi \geq 0$ be the Lagrange multiplier on (11) and let $B(v_i, \sigma_i)$ be the set of (u_1, u_2) that satisfy (6) and (8) for given (v_i, σ_i). Then $(\mathbf{u}^*, \boldsymbol{\sigma}^*)$ is a solution to the Lagrangian

$$\mathcal{L} = \min_{\mathbf{u}, \boldsymbol{\sigma}} \sum_i \psi_i \left[\mathbb{E}_{\sigma_i} \sum_t \zeta_t C(u_{i,t}) + \chi W(\sigma_i) \right] \tag{12}$$

subject to $(u_{i,1}, u_{i,2}) \in B(v_i, \sigma_i)$.

The solution to the maximization problem (12) can be characterized in two steps. First, for any pair (v, σ) define

$$\kappa(v, \sigma) \equiv \min_{(u_1, u_2) \in B(v, \sigma)} \mathbb{E}_\sigma \sum_t \zeta_t C(u_t). \tag{13}$$

Function $\kappa(v, \sigma)$ is the resource cost of delivering utility v to an agent who plays reporting strategy σ. This problem reduces to a standard mechanism design problem when $\sigma \in \Sigma^{in}$. We call the solution to $\kappa\left(v, \sigma^{in}\right)$ the *second best insurance* that gives an agent utility v. In our settings κ captures the resource costs of delivering utility v *on the equilibrium path*, that is, if the government sticks to its pre-election promises. Function W, instead, captures the *off the equilibrium path* incentives to re-optimize.

The optimal reporting strategy of each agent depends on the following trade-off. More informative reporting strategies (which we formally define below) lower the cost of delivering v on the equilibrium path, but also increase the incentives for the government to re-optimize ex-post. The solution to this trade-off is captured by

$$k(v) = \min_{\sigma} \kappa(v, \sigma) + \chi W(\sigma), \tag{14}$$

which characterizes the optimal reporting strategy of an agent with utility v. The Lagrangian $\mathcal{L}$ satisfies $\mathcal{L} = \sum_i \psi_i k(v_i)$.

We are now ready to characterize the efficient information revelation. Since the constraint set $B(v, \sigma)$ is linear in (v, u_1, u_2) and C is homogeneous, function $\kappa(v, \sigma)$ takes the form $\kappa(v, \sigma) = d(\sigma) C(av)$ for some $d(\sigma) > 0$ and a constant $a > 0$.[9] This allows us to order all reporting strategies. We say that σ'' is *more informative* than σ', $\sigma'' \succeq \sigma'$, if $d(\sigma'') \leq d(\sigma')$. More informative strategies have a natural interpretation that they allow the government to deliver any given utility at a lower cost. Naturally, $\sigma^{in} \succeq \sigma \succeq \sigma^{un}$ for any σ. The central result of this section is the following proposition.

Proposition 1 *If $v_{i''} \geq v_{i'}$ then $\sigma^*_{i''} \succeq \sigma^*_{i'}$.*

Proof. The objective function in (14) has increasing differences in (σ, v) and, therefore, the result follows from Topkis (2011). ■

Function $\kappa(v, \sigma)$ satisfies a version of the single crossing property in a sense that $\kappa(\cdot, \sigma') - \kappa(\cdot, \sigma'')$ is increasing when $\sigma'' \succeq \sigma'$. The economic content of this result is that additional information about the idiosyncratic shock of a high-v agent saves more resources on the equilibrium path than additional information about the shock of a low-v agent. Since W does not depend on v, in equilibrium it is optimal that high-v agents reveal more information than low-v agents.

Figure 1 illustrates Proposition 1 graphically. Panel A plots κ for three different reporting strategies, σ^{in}, σ^{un} and $\sigma \notin \Sigma^{in} \cup \Sigma^{un}$. The resource gains from better information, $\kappa(v, \sigma^{un}) - \kappa(v, \sigma)$ and $\kappa(v, \sigma) - \kappa(v, \sigma^{in})$, monotonically increase in v, converge to zero as $v \to \underline{v}$ and diverge to infinity as $v \to \bar{v}$. Panel B adds the off the equilibrium cost of deviation assuming $\chi > 0$. Since $W(\sigma^{in}) > W(\sigma) > W(\sigma^{un})$ by Lemma 2, functions $\{\kappa(\cdot, \tilde{\sigma}) + \chi W(\tilde{\sigma})\}_{\tilde{\sigma} \in \{\sigma^{un}, \sigma^{in}, \sigma\}}$ must intersect, with less informative functions crossing more informative functions from below. The lower envelope of these functions characterizes the best reporting strategy for each v.

[9] In particular, $a = 1$ if $\rho \neq 1$ and $a = \frac{1}{2}$ if $\rho = 1$. Observe that if $u^*_t(m; v)$ is a solution to (13) for given v, then $u^*_t(m; v) = v u^*_t(m; 1)$ if $\rho < 1$, $u^*_t(m; v) = -v u^*_t(m; -1)$ if $\rho > 1$ and $u^*_t(m; v) = \frac{1}{2}v + u^*_t(m; 0)$ if $\rho = 1$.

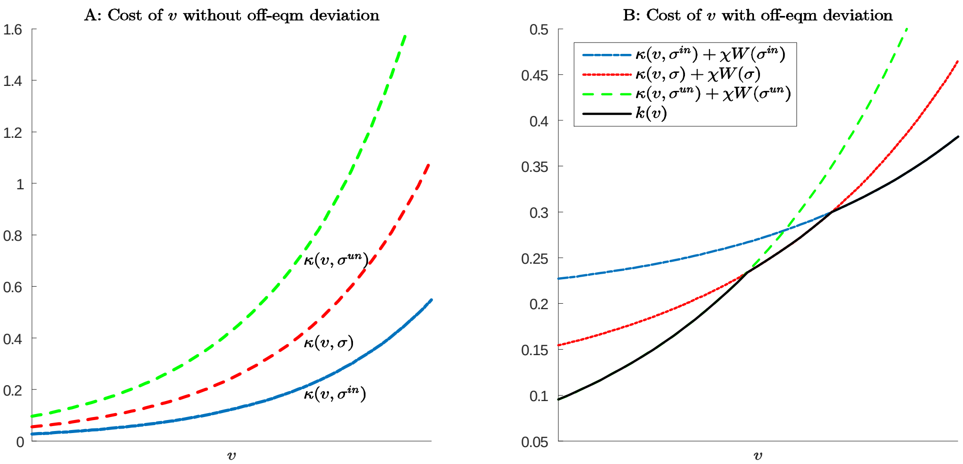

Figure 1: Panel A plots the resource costs needed to deliver utility v on the equilibrium path given any reporting strategy. Panel B adds the off the equilibrium cost of deviation.

This result illustrates the general principle behind optimal information revelation when the government cannot commit – those agents should reveal more information for whom the on the equilibrium path gains are high relative to the off the equilibrium path costs. In our setting the on the path gains are increasing in the agent's utility v (or, equivalently, in his Pareto weight) while the off path costs do not depend on v. This implies that the agents with higher weights should reveal more information.[10]

2.1 Information revelation and the provision of incentives

In this section we provide more insights about the strategies that agents use to report their information and the allocations they receive.

Lemma 3 *Any point on the Pareto frontier can be supported by reporting strategies such that each agent reports at most two messages with positive probability and for each group i at most one $\theta \in \Theta$ plays a mixed strategy.*

This lemma shows that we can restrict attention to simple strategies in which only one type θ randomizes between either pooling with the other type or separating from him. We can parameterize such strategies by a pair (j, s) where $j \in \{H, L\}$ is the identity of the type that

[10] This arguments extends directly to other environments. Suppose that, along the lines of the set up discussed in footnote 7, the politician is not benevolent but instead assigns weight $\tilde{\omega}_i$ to the utility of the members of group i. Consider the optimal information revelation in the best utilitarian equilibrium that maximizes the sum of utilities of all agents. Our arguments extend directly to this set up and show that agents who are valued more highly by the politician reveal less information on the equilibrium path.

separates and $s \in [0,1]$ is his probability of separation. In the appendix we show that Lemma 3 implies that the cost minimization problem if type L randomizes can be written as

$$\begin{aligned} \kappa^L(v,s) &= \min_{\{u_t(m_j)\}_{t\in\{1,2\},j\in\{H,L\}}} s\pi_L\left[\zeta_1 C(u_1(m_L)) + \zeta_2 C(u_2(m_L))\right] \\ &\quad + (\pi_H + (1-s)\pi_L)\left[\zeta_1 C(u_1(m_H)) + \zeta_2 C(u_2(m_H))\right] \end{aligned} \tag{15}$$

subject to

$$\begin{aligned} v &= s\pi_L\left[\theta_L u_1(m_L) + u_2(m_L)\right] + (1-s)\pi_L\left[\theta_L u_1(m_H) + u_2(m_H)\right] \\ &\quad + \pi_H\left[\theta_H u_1(m_H) + u_2(m_H)\right] \end{aligned} \tag{16}$$

and

$$\theta_L u_1(m_L) + u_2(m_L) = \theta_L u_1(m_H) + u_2(m_H). \tag{17}$$

The cost minimization problem if type H randomizes, $\kappa^H(v,s)$, is written analogously but (17) is replaced with

$$\theta_H u_1(m_L) + u_2(m_L) = \theta_H u_1(m_H) + u_2(m_H). \tag{18}$$

We define $W^L(s)$ and $W^H(s)$ similarly.

Let $\left\{u_t^j(m;v,s)\right\}_{m,t}$ be the solution to the minimization problem defined by $\kappa^j(v,s)$. Let $v^j(m_k;v,s) = \theta_k u_1^j(m_k;v,s) + u_2^j(m_k;v,s)$ be the utility received by type θ_k.

Proposition 2 *(a) Functions $\kappa^j(v,\cdot)$, $-W^j(\cdot)$, $-\left[u_2^j(m_L;v,\cdot) - u_2^j(m_H;v,\cdot)\right]$, $\left[u_1^j(m_L;v,\cdot) - u_1^j(m_H;v,\cdot)\right]$, $v^j(m_j;v,\cdot)$, $-v^j(m_{-j};v,\cdot)$ are all decreasing.*

(b) $\kappa^j(v,\cdot)$ is differentiable and its derivative takes a form $\frac{\partial}{\partial s}\kappa^j(v,s) = -b^j(s)C(av)$ for some $b^j(s)$. There exist strictly positive $\underline{\varepsilon},\bar{\varepsilon}$ such that $b^j(s) \in [\underline{\varepsilon},\bar{\varepsilon}]$ and $\frac{\partial}{\partial s}W^j(s) \in [\underline{\varepsilon},\bar{\varepsilon}]$ for all j,s.

Part (a) of Proposition 2 shows how the probability of separation is related to informativeness and insurance. Strategies with higher probability of separation are more informative (since $\kappa^j(v,\cdot)$ is decreasing and $W^j(\cdot)$ is increasing). More informative strategies save resources because they allow the government to provide better insurance ($u_1^j(m_H;v,\cdot) - u_1^j(m_L;v,\cdot)$ increases). The incentive compatibility is preserved by increasing $u_2^j(m_L;v,\cdot) - u_2^j(m_H;v,\cdot)$ as well. Contracts that incentivize agent j to separate with higher probability also lower that agent's utility in favor of the other agent ($v^j(m_j;v,\cdot)$ is decreasing, $v^j(m_{-j};v,\cdot)$ is increasing).

Part (b) of Proposition 2 characterizes marginal gains from more informative strategies on and off the equilibrium path, $\frac{\partial}{\partial s}\kappa^j(v,s)$ and $\frac{\partial}{\partial s}W^j(s)$. One important observation is that the marginal gain from better information is always *strictly* positive off the equilibrium path. To see this consider problem (15). The posterior beliefs of the government are bounded away from each other for any s since $\mathbb{E}_s[\theta|m_L] = \theta_L < 1 \leq \mathbb{E}_s[\theta|m_H]$. Thus any marginal increase in the informativeness of s yields a strictly positive gain. On the other hand, the marginal gain from better information on the equilibrium path, $\frac{\partial}{\partial s}\kappa^j(v,s)$, becomes unboundedly small as $v \to \underline{v}$ and unboundedly large as $v \to \bar{v}$. We then immediately get the following result.

Corollary 1 *For any point on the Pareto frontier, there are v^-, v^+, with $\underline{v} \leq v^- < v^+ < \bar{v}$ (with $\underline{v} < v^-$ if $\chi > 0$), such that if $v_i < v^-$ then σ_i^* is uninformative, and if $v_i > v^+$ then σ_i^* is fully informative.*

Proof. Since the difference $\kappa(v,\sigma^{un}) - \kappa\left(v,\sigma^{in}\right)$ goes to zero as $v \to \underline{v}$ and to infinity as $v \to \bar{v}$, while $W(\sigma^{un}) - W\left(\sigma^{in}\right)$ is bounded, full information revelation cannot be optimal for low values of v (as long as $\chi > 0$) and no information revelation cannot be optimal for high values of v. If any intermediate reporting strategy is optimal, by Lemma 3 it is equivalent to a strategy where only some type j randomizes between two messages. Since both κ^j and W^j are differentiable, the optimality condition can be written as $\frac{\partial}{\partial s}\kappa^j(v,s) = \chi\frac{\partial}{\partial s}W^j(s)$. The bounds in Proposition 2(b) rule out this possibility for sufficiently high and low v. ■

This proposition shows that for any point $\mathbf{v}$ there are some bounds $\{v^-, v^+\}$ so that if v_i is outside of these bounds then the optimal strategy is either uninformative or fully informative. These regions may or may not be empty depending on the point $\mathbf{v}$, although they are always nonempty provided that $\{v_i\}_i$ are sufficiently different from each other and $\Upsilon > 0$. As we shall see next, these regions play a key role once we consider a more general class of insurance mechanisms.

2.2 Stochastic mechanisms and rationing of insurance

So far we focused on deterministic mechanisms: all agents from the same group i were treated in the same way by the government and received allocations as a function of their group identity and their reports. Figure 1 suggests that such mechanisms may lead to a non-convex Pareto frontier. In such cases stochastic insurance mechanisms will further improve welfare. In this section we extend our analysis to such mechanisms.

Formally we consider the same environment as in the previous section but allow both the government and the agents to condition their strategies on the realization of an agent-specific, payoff-irrelevant variable z uniformly distributed on the set $Z = [0,1]$. We keep all the notation parallel to that in the previous section, but use bold letters to emphasize that the variable may depend on z. Thus $\mathbf{u}_{i,t}^{pr}, \mathbf{u}_{i,t} : M \times Z \to \mathbb{R}$ are promises and final allocations of the politician while $\boldsymbol{\sigma}_i : Z \times \Theta \to \mathbb{R}$ are the reporting strategies of the agents. The expectation for any variable $\mathbf{x} \in \Theta \times M \times Z$ is now defined as $\mathbb{E}_{\boldsymbol{\sigma}} \mathbf{x} \equiv \int_{\Theta \times M \times Z} \mathbf{x}(\theta, m, z)\, \pi(d\theta)\, \boldsymbol{\sigma}(dm|z,\theta)\, dz$.

Our analysis of this game proceeds with minimal changes. Same arguments to the ones used before show that the sustainability constraint for the politician can be written as

$$\sum_{i=1}^{I} \psi_i \mathbb{E}_{\boldsymbol{\sigma}_i} [\theta \mathbf{u}_{i,1} + \mathbf{u}_{i,2}] \geq \sum_{i=1}^{I} \psi_i \int_Z W(\boldsymbol{\sigma}_i(\cdot|z,\cdot))\, dz + \underline{U} - \Upsilon, \tag{19}$$

which is the stochastic analogue of (11). The equilibrium strategies are a solution to the Lagrangian

$$\mathcal{L}^{stoch} = \min_{\mathbf{u},\boldsymbol{\sigma}} \sum_i \psi_i \int_Z \left[\mathbb{E}_{\boldsymbol{\sigma}_i} \left[\sum_t \zeta_t C(\mathbf{u}_{i,t}) \middle| z \right] + \chi W(\boldsymbol{\sigma}_i(\cdot|z,\cdot)) \right] dz,$$

where $\mathbf{u}_i(\cdot, z)$ and $\boldsymbol{\sigma}_i(\cdot|z,\cdot)$ are subject to the incentive constraint (6) for all i and z, and the constraint

$$v_i = \int_Z \mathbb{E}_{\boldsymbol{\sigma}_i} [\theta \mathbf{u}_{i,1} + \mathbf{u}_{i,2} | z]\, dz \text{ for all } i.$$

Stochastic mechanisms improve welfare by relaxing constraint (8). For each realization of z, the optimal $\left\{ \mathbf{u}_{i,1}^*(\cdot, z), \mathbf{u}_{i,2}^*(\cdot, z), \boldsymbol{\sigma}_i^*(\cdot|z,\cdot) \right\}$ is a solution to problem (13) for some $\mathbf{v}_i(z)$ and the relationship between v_i and $\mathbf{v}_i(z)$ is given by $v_i = \int_Z \mathbf{v}_i(z)\, dz$. The value of the Lagrangian satisfies $\mathcal{L}^{stoch} = \sum_i k^{stoch}(v_i)\, \psi_i$ where $k^{stoch}(v)$ is the convex hull of $k(v)$ defined in (14).

The results of the previous section extend directly to stochastic mechanisms using the following notion of informativeness. Without loss of generality we assume that $\boldsymbol{\sigma}$ is increasing in z in a sense that $z'' \geq z'$ implies that $\boldsymbol{\sigma}(\cdot|z'',\cdot) \succeq \boldsymbol{\sigma}(\cdot|z',\cdot)$ and say that $\hat{\boldsymbol{\sigma}}$ is more informative than $\tilde{\boldsymbol{\sigma}}$ if $\hat{\boldsymbol{\sigma}}(\cdot|z,\cdot) \succeq \tilde{\boldsymbol{\sigma}}(\cdot|z,\cdot)$ for all z. We call $\boldsymbol{\sigma}$ fully informative (uninformative) if $\boldsymbol{\sigma}(\cdot|z,\cdot)$ is fully informative (uninformative) for all z. Since the optimal allocations for any $\tilde{\boldsymbol{\sigma}}(\cdot|z,\cdot)$ are the solution to (13) the analysis in Section 2.1 applies to stochastic mechanisms. The results of Proposition 1 and Corollary 1 extend directly as well. In particular, we have

Corollary 2 *Any best PBE is payoff-equivalent to a PBE with a property that $v_{i''} \geq v_{i'}$ implies $\boldsymbol{\sigma}_{i''}^* \succeq \boldsymbol{\sigma}_{i'}^*$. There exist v^-, v^+, with $\underline{v} \leq v^- < v^+ < \bar{v}$ (with $\underline{v} < v^-$ if $\chi > 0$), such that if $v_i < v^-$ then $\boldsymbol{\sigma}_i^*$ is uninformative, and if $v_i > v^+$ then $\boldsymbol{\sigma}_i^*$ is fully informative.*

The main new insight of this section is that stochastic mechanisms may lead to Pareto improvements and take a particularly simple form.

Proposition 3 *Suppose $\rho = 1$. There is an open set $D \subset \mathbb{R}^4_+$ such that if $\left(\{\theta_j, \pi(\theta_j)\}_j\right) \in D$ then for $v_i \in [v^-, v^+]$ the optimal strategies satisfy $\boldsymbol{\sigma}^*_i(\cdot|z, \cdot) = \sigma^{un}$, $v^*_i(z) = v^-$ if $z < \bar{z}_i$ and $\boldsymbol{\sigma}^*_i(\cdot|z, \cdot) = \sigma^{in}, v^*_i(z) = v^+$ if $z \geq \bar{z}_i$, where $\bar{z}_i = \frac{v^+ - v_i}{v^+ - v^-}$. The set D does not depend on the values of $\{\omega_i, \psi_i\}_i$ or Υ.*

The proof of this proposition is in the appendix. It shows that whether any strategy $\sigma \notin \{\sigma^{in}, \sigma^{un}\}$ is optimal depends only on the parameters $\{\theta_i, \pi(\theta_i)\}_i$, and not on any other variables, including the Lagrange multipliers in problem (13). It also provides sufficient conditions for $\{\theta_i, \pi(\theta_i)\}_i$ that ensure that partial pooling is never optimal.[11]

When the assumptions of Proposition 3 are satisfied, insurance provision takes a simple form. Only second best insurance, that requires full information revelation, is provided by the government but access to this insurance is limited. Low-v_i agents receive no insurance, high-v_i agents receive insurance with probability 1, while agents with intermediate values of v_i receive insurance allocated through a lottery. All agents in this intermediate range receive the same allocations if they win the lottery, but higher values of v_i imply better odds of winning the lottery. One natural interpretation of the lottery is that insurance is rationed.

Consider the implications of Proposition 3 for the case when there is no ex-ante heterogeneity across agents and the government maximizes the ex-ante utility of all citizens. This corresponds to $I = 1$ in our set up. Some information revelation is optimal in the best equilibrium for all $\Upsilon > 0$ but full information revelation is infeasible if Υ is not too high. Under the assumptions of Proposition 3 none of the agents plays a mixed reporting strategy in this case. Rather, agents are randomly assigned to two groups. Agents in the first group reveal full information about their shock and receive the second best insurance that gives them utility v^+. Agents in the second group reveal no information and receive no insurance obtaining utility of $v^- < v^+$.

Finally, in this section we characterized the efficient insurance arrangements when agents communicate directly with the government. This is a natural assumption in the context of many political economy environments. In the Supplementary material we extend our analysis to environments that involve a mediator, along the lines of Myerson (1982), and show that our main insights carry over to such economies.

[11] In fact, we cojecture that Proposition 3 is stronger as we have not been able to find parameters $\{\theta_i, \pi(\theta_i)\}_i$ for which it is *not* satisfied.

3 An infinitely repeated game

In this section we extend our analysis to infinitely repeated games. We consider a version of the Atkeson and Lucas (1992) environment in which insurance is provided by a benevolent government. Our main departure from that model is the assumption that the government cannot commit.

The economy is populated by a continuum of agents of total measure 1 and the government. There is an infinite number of periods, $t = 0, 1, 2, ...$ The economy is endowed with e units of a perishable good in each period. An agent's instantaneous utility from consuming c_t units of the good in period t is given by $\theta_t U(c_t)$ where $U : \mathbb{R}_+ \to \mathbb{R}$ is an increasing, strictly concave, continuously differentiable function. The utility function U satisfies Inada conditions $\lim_{c\to 0} U'(c) = \infty$ and $\lim_{c\to\infty} U'(c) = 0$ and it may be bounded or unbounded. Let $\bar{u} = \lim_{c\to\infty} U(c)$, $\underline{u} = \lim_{c\to 0} U(c)$ be the bounds (which may be infinite) of U. Let $C \equiv U^{-1}$ be the inverse of the utility function. All agents have a common discount factor β. Let $\bar{v} = \frac{\bar{u}}{1-\beta}$, $\underline{v} = \frac{\underline{u}}{1-\beta}$ be the bounds on the lifetime utility.

The taste shock θ_t takes values in a finite set Θ with cardinality $|\Theta|$. In this section we assume that θ_t are i.i.d. across agents and across time, but we relax this assumption in Section 4. Let $\pi(\theta) > 0$ be the probability of realization of $\theta \in \Theta$. We assume that $\theta_1 < ... < \theta_{|\Theta|}$ and normalize $\sum_{\theta\in\Theta} \pi(\theta)\theta = 1$. We use superscript t to denote a history of realizations of any variable up to period t, e.g. $\theta^t = (\theta_0, ..., \theta_t)$. Let $\pi_t(\theta^t)$ denote the probability of realization of history θ^t. We assume that types are private information. Each agent belongs to a group $v \in (\underline{v}, \bar{v})$ in period 0 ($[\underline{v}, \bar{v})$ if utility is bounded below) and the distribution of agents over $(\underline{v}, \bar{v})$ is denoted by ψ. For now we treat ψ as exogenous following Section 2, but in Section 3.3 we endogenize it when we consider properties of invariant distributions.

Consumption allocations are provided by the government, which is utilitarian but lacks commitment. Formally we consider an infinitely repeated game between the government and a continuum of agents along the lines of Chari and Kehoe (1990) and Chari and Kehoe (1993). Each period t is divided in two stages. In stage 1 agents transmit information to the government about their type using a message set M, which for simplicity we assume to be countable. Each agent sends a report $m_t \in M$ about the realization of his type using strategy $\boldsymbol{\sigma}_t$. The reports are a function of current and past realizations of shocks θ^t, current and past realizations of idiosyncratic sunspot variables z^t, past reports m^{t-1}, initial group identity v, and the history of government's actions that we describe below. Let $\breve{h}^t = (v, m^{t-1}, z^t)$ and $h^t = (v, m^t, z^t)$ be, respectively, the idiosyncratic histories of agents before and after they submit reports m_t, and let $\breve{H}^t$ and H^t be the spaces of all such histories. A reporting strategy $\boldsymbol{\sigma}_t$ induces a probability distribution over M denoted by $\boldsymbol{\sigma}_t\left(\cdot|\breve{h}^t,\theta^t\right)$, which also depends implicitly on the history of government's actions. We assume that the law of the large numbers holds and the aggregate distribution of histories h^t, denoted by μ_t, is given by[12]

$$\begin{aligned}
\mu_{-1}(v) &= \psi(v), \\
\mu_t\left(h^t\right) &= \mu_{t-1}\left(h^{t-1}\right)\Pr(z_t)\sum_{\theta^t\in\Theta^t}\pi_t\left(\theta^t\right)\boldsymbol{\sigma}_t\left(m_t|h^{t-1},z_t,\theta^t\right).
\end{aligned}$$

The triple H^t, its Borel sigma algebra, and μ_t is a probability space.

In stage 2 of each period the government chooses allocations. The allocations are measurable functions $\mathbf{u}_t$ from H^t into $(\underline{u},\bar{u})$ (into $[\underline{u},\bar{u})$ if U is bounded below) that satisfy the feasibility constraint. Using the shorthand notation $\mathbb{E}_{\boldsymbol{\sigma}}x_t=\int x_t d\mu_t$ for any measurable x_t, the feasibility constraint can be written as

$$\mathbb{E}_{\boldsymbol{\sigma}}C(\mathbf{u}_t)\leq e \text{ for all } t. \tag{20}$$

All variables defined above are also functions of aggregate histories. The aggregate histories include the distribution of reports, $\boldsymbol{\mu}=\{\mu_t\}_{t=0}^{\infty}$, and the distribution of allocations chosen by the government, $\mathbf{u}=\{\mathbf{u}_t\}_{t=0}^{\infty}$. The strategies of the agents and the government are restricted so that they take the same values for any two aggregate histories that differ for a measure zero of agents. Given this restriction the reporting strategy of any individual agent does not affect the aggregate allocations in the game.

A PBE consists of strategies of agents and the government and posterior beliefs such that, at each history of the game, each player chooses his best response given his posterior beliefs formulated using Bayes' rule. A best PBE is a PBE such that there is no other PBE that gives higher utility to a set of agents of measure 1, and strictly higher utility to a positive measure of agents. Without loss of generality we assume that v denotes the lifetime expected utility, or *payoff*, that the members of group v receive in a best PBE.

3.1 The recursive problem

Our definition of equilibrium implies that there is no aggregate uncertainty. Along the equilibrium path both the aggregate distribution of agents' reports $\boldsymbol{\mu}$ and the allocations $\mathbf{u}$ are

[12]Strictly speaking, since z_t is a continuous variable, μ_t is defined as follows. Let $\mu_{-1}=\psi$. Any Borel set A^t of H^t can be represented as a product $A^t=A^{t-1}\times B_m\times B_z$, where A^{t-1} is a Borel set of H^{t-1} and B_m,B_z are the m_t- and z-sections of some Borel set of $M_t\times Z$. Then μ_t is defined as

$$\mu_t\left(A^t\right)=\mu_{t-1}\left(A^{t-1}\right)\Pr(z_t\in B_z)\sum_{\theta^t}\pi_t\left(\theta^t\right)\boldsymbol{\sigma}_t\left(B_m|A^{t-1},B_z,\theta^t\right).$$

deterministic sequences. Following standard arguments, government's equilibrium strategies are supported by a threat to revert to a PBE that gives the government the lowest utility, which we call a worst PBE, if the government deviates. Next lemma constructs such an equilibrium.

Lemma 4 *In a worst PBE all agents report the same message for all histories* $\left(\breve{h}^t, \theta^t\right)$ *and the government allocates* $U(e)$ *independently of the agents' reports.*

Proof. Let $\boldsymbol{\sigma}^w$ be a reporting strategy in which the same message is reported for all histories, let $\mathbf{u}^w$ be the allocation rule that takes a constant value $U(e)$ for all h^t, and let the government's posterior beliefs be given by $\mathbb{E}_{\boldsymbol{\sigma}^w}\left[\theta|h^t\right] = 1$ for all h^t. It is easy to see that this triple is consistent with Bayes' rule and constitutes best responses of agents and the government to each other's strategies. Therefore it is a PBE. It gives the government payoff $\frac{U(e)}{1-\beta}$. Since the allocation $\mathbf{u}^w$ is feasible for any other reporting strategies of the agents, government's payoff must be at least $\frac{U(e)}{1-\beta}$ in any PBE. Therefore, the constructed equilibrium is a worst PBE. ■

Let $\boldsymbol{\sigma} = \{\boldsymbol{\sigma}_t\}_{t=0}^{\infty}$ be a reporting strategy and let $\boldsymbol{\mu}$ be the induced distribution of reports. The highest payoff that the government can achieve in period t is given by a function $\tilde{W}_t(\mu_t)$ defined by

$$\tilde{W}_t(\mu_t) = \max_{\mathbf{u}_t} \mathbb{E}_{\boldsymbol{\sigma}} \theta \mathbf{u}_t \tag{21}$$

subject to (20). Therefore the best response constraint of the government can be written as

$$\mathbb{E}_{\boldsymbol{\sigma}} \sum_{s=t}^{\infty} \beta^{s-t} \theta_s \mathbf{u}_s \geq \tilde{W}_t(\mu_t) + \frac{\beta}{1-\beta} U(e) \quad \text{for all } t. \tag{22}$$

Since each agent's report does not affect aggregate distributions, agents' incentive constraints are

$$\mathbb{E}_{\boldsymbol{\sigma}}\left[\sum_{t=0}^{\infty} \beta^t \theta_t \mathbf{u}_t \middle| v\right] \geq \mathbb{E}_{\boldsymbol{\sigma}'}\left[\sum_{t=0}^{\infty} \beta^t \theta_t \mathbf{u}_t \middle| v\right] \quad \text{for all } \boldsymbol{\sigma}', v. \tag{23}$$

Therefore, any best equilibrium is a solution to

$$\max_{\mathbf{u}, \boldsymbol{\sigma}} \mathbb{E}_{\boldsymbol{\sigma}} \sum_{t=0}^{\infty} \beta^t \theta_t \mathbf{u}_t \tag{24}$$

subject to (20), (22), (23), and

$$\mathbb{E}_{\boldsymbol{\sigma}}\left[\sum_{t=0}^{\infty} \beta^t \theta_t \mathbf{u}_t \middle| v\right] = v. \tag{25}$$

We start the analysis by simplifying strategies and allocations.

Lemma 5 *Any best PBE is payoff equivalent to a PBE in which $\boldsymbol{\sigma}_t$ is independent of θ^{t-1} and for which the following property holds: if there is some $w \in \mathbb{R}$ and histories h'^t, h''^t such that*

$$w = \mathbb{E}_{\boldsymbol{\sigma}} \left[\sum_{s=t}^{\infty} \beta^{s-t} \theta_s \mathbf{u}_s \middle| h'^t \right] = \mathbb{E}_{\boldsymbol{\sigma}} \left[\sum_{s=t}^{\infty} \beta^{s-t} \theta_s \mathbf{u}_s \middle| h''^t \right],$$

then $\boldsymbol{\sigma}_T \left(m | \breve{h}'^T, \theta_T \right) = \boldsymbol{\sigma}_T \left(m | \breve{h}''^T, \theta_T \right)$, $\mathbf{u}_T \left(\breve{h}'^T, m_T \right) = \mathbf{u}_T \left(\breve{h}''^T, m_T \right)$ for all $T > t$ where $\breve{h}'^T = \left(h'^t, z_{t+1}, m_{t+1}, ..., z_T \right)$, $\breve{h}''^T = \left(h''^t, z_{t+1}, m_{t+1}, ..., z_T \right)$ for some $(z_{t+1}, m_{t+1}, ..., z_T)$ and m_T.

This lemma is an intermediate step in our recursive characterization of best PBE. It shows that all the information required to characterize the agents' behavior after any period t can be summarized in a variable w that captures the agent's expected continuation payoff in period t along the equilibrium path.

Our analysis of optimal information revelation relies on the recursive formulation of problem (24). Let $(\mathbf{u}^*, \boldsymbol{\sigma}^*)$ be a best PBE and $\boldsymbol{\mu}^*$ be the distribution of histories induced by $\boldsymbol{\sigma}^*$. Let λ_t^w be the Lagrange multiplier on the feasibility constraint (20) in the maximization problem (21) when $\mu_t = \mu_t^*$. For any mapping $\sigma : \Theta \to \Delta(M)$ let $W_t(\sigma)$ be given by

$$W_t(\sigma) \equiv \max_{\{u(m)\}_{m \in M}} \sum_{(m,\theta) \in M \times \Theta} \left(\theta u(m) - \lambda_t^w C(u(m)) \right) \sigma(m|\theta) \pi(\theta) + \lambda_t^w e. \tag{26}$$

We use $\mathbb{E}_{\boldsymbol{\sigma}} W_t$ as a shorthand for $\int_{H^{t-1} \times Z} W_t \left(\boldsymbol{\sigma}_t \left(\cdot | h^{t-1}, z, \cdot \right) \right) dz d\mu_{t-1}$. The arguments of Lemma 1 immediately establish the following result.

Lemma 6 $(\mathbf{u}^*, \boldsymbol{\sigma}^*)$ *is a solution to the maximization problem (24) in which constraint (22) is replaced with*

$$\mathbb{E}_{\boldsymbol{\sigma}} \sum_{s=t}^{\infty} \beta^{s-t} \theta_s \mathbf{u}_s \geq \mathbb{E}_{\boldsymbol{\sigma}} W_t + \frac{\beta}{1-\beta} U(e) \text{ for all } t. \tag{27}$$

Lemma 6 allows us to form a Lagrangian to the constrained maximization problem and study it using recursive techniques along the lines of our analysis in Section 2. Let $\left\{ \beta^t \zeta_t^* \right\}_{t=0}^{\infty}$ and $\left\{ \beta^t \chi_t^* \right\}_{t=0}^{\infty}$ be the Lagrange multipliers on (20) and (27), respectively. The Lagrangian to the constrained maximization problem can be written as (see, e.g. Marcet and Marimon (2009) or Chapter 20.4 in Ljungqvist and Sargent (2012))

$$\mathcal{L} = \max_{\mathbf{u}, \boldsymbol{\sigma}} \mathbb{E}_{\boldsymbol{\sigma}} \sum_{t=0}^{\infty} \bar{\beta}_t \left[\theta_t \mathbf{u}_t - \zeta_t C(\mathbf{u}_t) - \chi_t W_t \right] \tag{28}$$

subject to (23) and (25), where $\bar{\beta}_t = \beta^t \left(1 + \sum_{s=0}^{t} \chi_s^*\right)$, $\zeta_t = \beta^t \zeta_t^* / \bar{\beta}_t$, and $\chi_t = \beta^t \chi_t^* / \bar{\beta}_t$. Let $\hat{\beta}_t \equiv \bar{\beta}_t / \bar{\beta}_{t-1}$ be the effective discount factor. By definition $\hat{\beta}_t \geq \beta$ with strict inequality if and only if constraint (27) binds in period t.

Problem (28) is the infinite period analogue of (12). Our analysis proceeds similarly to that in Section 2. Define

$$k_t(v) \equiv \frac{1}{\bar{\beta}_t} \max_{\mathbf{u},\boldsymbol{\sigma}} \mathbb{E}_{\boldsymbol{\sigma}} \left[\sum_{s=0}^{\infty} \bar{\beta}_{t+s} \left(\theta_s \mathbf{u}_s - \zeta_{t+s} C(\mathbf{u}_s) - \chi_{t+s} W_{t+s}\right) \middle| v \right] \tag{29}$$

subject to (23) and (25). The Lagrangian (28) satisfies $\mathcal{L} = \int \bar{\beta}_0 k_0(v)\, d\psi$.

Lemma 7 *W_t satisfies all the properties of Lemma 2.*

k_t is continuous, concave and differentiable with $\lim_{v\to\bar{v}} k_t'(v) = -\infty$. If utility is unbounded below then $\lim_{v\to\underline{v}} k_t'(v) = 1$; if utility is bounded below then $\lim_{v\to\underline{v}} k_t'(v) \geq 1$ with $\lim_{v\to\underline{v}} k_t'(v) = \infty$ if $\limsup \chi_t > 0$.

It is easy to write (29) recursively. Suppose $(\mathbf{u}^*, \boldsymbol{\sigma}^*)$ is a best PBE, which without loss of generality satisfies the properties of Lemma 5. For any h^{t-1} let $v = \mathbb{E}_{\boldsymbol{\sigma}^*}\left[\sum_{s=t}^{\infty} \beta^{s-t} \theta_s \mathbf{u}_s^* | h^{t-1}\right]$. Then $\left\{\mathbf{u}_t^*\left(h^{t-1}, m_t, z_t\right), \boldsymbol{\sigma}_t^*\left(m_t | h^{t-1}, z_t, \theta_t\right)\right\}_{m_t,\theta_t,z_t}$ is a solution to

$$k_t(v) = \max_{\mathbf{u},\mathbf{w},\boldsymbol{\sigma}} \int_Z \left[\sum_{\theta,m} \pi(\theta)\, \boldsymbol{\sigma}(m|z,\theta) \left[\theta \mathbf{u}(m,z) - \zeta_t C(\mathbf{u}(m,z)) + \hat{\beta}_{t+1} k_{t+1}(\mathbf{w}(m,z))\right] - \chi_t W(\boldsymbol{\sigma}(\cdot|z,\cdot)) \right] dz \tag{30}$$

subject to

$$v = \int_Z \sum_{\theta,m} \pi(\theta)\, \boldsymbol{\sigma}(m|z,\theta) \left[\theta \mathbf{u}(m,z) + \beta \mathbf{w}(m,z)\right] dz, \tag{31}$$

$$\sum_m \boldsymbol{\sigma}(m|z,\theta) \left[\theta \mathbf{u}(m,z) + \beta \mathbf{w}(m,z)\right] \geq \sum_m \boldsymbol{\sigma}'(m|z,\theta) \left[\theta \mathbf{u}(m,z) + \beta \mathbf{w}(m,z)\right] \text{ for all } z, \theta, \boldsymbol{\sigma}'. \tag{32}$$

To characterize the properties of efficient information revelation it is useful to separate the maximization problem (30) into two components. Let $\gamma_t(v) \equiv k_t'(v)$. For $\sigma : \Theta \to \Delta(M)$ and for any $x : M \times \Theta \to \mathbb{R}$ define $\mathbb{E}_\sigma x$ as in Section 2 and let

$$\kappa_t(v,\sigma) \equiv \max_{\{u(m),w(m)\}_{m\in M}} \mathbb{E}_\sigma \left[(1-\gamma_t(v))\, \theta u - \zeta_t C(u) + \hat{\beta}_{t+1} k_{t+1}(w) - \gamma_t(v)\, \beta w\right] + \gamma_t(v)\, v \tag{33}$$

subject to

$$\theta u(m_\theta) + \beta w(m_\theta) \geq \theta u(m) + \beta w(m) \text{ for all } \theta, m, \tag{34}$$

and

$$\sigma(m|\theta)\left[\left(\theta u\left(m_\theta\right)+\beta w\left(m_\theta\right)\right)-\left(\theta u(m)+\beta w(m)\right)\right]=0 \text{ for all } \theta, m, \tag{35}$$

where m_θ is any message such that $\sigma(m_\theta|\theta) > 0$. Let $(\mathbf{u}_v, \mathbf{w}_v, \boldsymbol{\sigma}_v)$ denote a solution to (30). Similarly, we use $(u_{v,\sigma}, w_{v,\sigma})$ to denote a solution to (33). The relationship between κ_t and k_t is given by the following lemma.

Lemma 8 *k_t satisfies*

$$k_t(v)=\max_{\sigma}\left\{\kappa_t(v,\sigma)-\chi_t W_t(\sigma)\right\}. \tag{36}$$

Moreover, $(\mathbf{u}_v(\cdot,z), \mathbf{w}_v(\cdot,z))$ is a solution to (33) with $\sigma=\boldsymbol{\sigma}_v(\cdot|z,\cdot)$ and $\boldsymbol{\sigma}_v(\cdot|z,\cdot)$ is a solution to (36) for all z.

We use σ_v to denote a solution to (36). Problem (33) has a similar structure to the standard recursive characterization in dynamic contracting models with commitment (e.g., Atkeson and Lucas (1992), Farhi and Werning (2007), Sleet and Yeltekin (2006)), except that it allows agents to send noisy information about their type. When the principal (the government in our setting) cannot commit, more precise information carries costs, which are captured by $\chi_t W_t$. The optimal information revelation is characterized by (36). When the cost of information revelation is absent, $\chi_t = 0$, full information revelation is optimal, as in standard principal-agent models.

Before we proceed we comment on how the cardinality of the message set affects the payoffs in best PBE. Larger message sets weakly increase welfare because it is possible to replicate the payoffs of a smaller message set with a larger one. To see this, take any message space M' and consider an alternative message space M'' constructed by adding additional messages to M'. The government can always give the lowest payoff to any agent who reports a message $m \in M''\backslash M'$ and, thus, ensure that in equilibrium messages $m \in M''\backslash M'$ are not played. The next lemma shows that the highest welfare can be attained with a finite message space. Let M_Θ be a set with $2|\Theta|-2$ elements $m_1, ..., m_{2|\Theta|-2}$.

Lemma 9 *Any payoff of a best PBE in a game that uses message set M can be attained in a game that uses message set M_Θ.*

In particular, Lemma 9 implies that there is no gain from allowing the government to "pre-commit" to getting coarser information by choosing a smaller message set. For example, suppose we introduced a preliminary stage in every period of our dynamic game in which the government can choose the optimal message set. By Lemma 9, without loss of generality the government would simply choose M_Θ. For concreteness, we assume for the rest of the paper that $M = M_\Theta$.

3.2 Characterization

In this section we characterize properties of efficient information revelation. In addition to uninformative and fully informative strategies defined in Section 2 we say that strategy σ *reveals full information about type* j if there is $\tilde{M} \subset M_\Theta$ such that $\sigma\left(\tilde{M}|\theta_j\right) = 1$ and $\mathbb{E}_\sigma\left[\theta|\tilde{M}\right] = \theta_j$. The same definition applies to $\boldsymbol{\sigma}$ if $\boldsymbol{\sigma}(\cdot|z,\cdot)$ reveals full information about type j for all z. Some of our results require the following assumption.

Assumption 1 ***(decreasing absolute risk aversion)*** *U is twice continuously differentiable and*

$$\lim_{c\to\infty} U''(c)/U'(c) = 0. \tag{37}$$

We start our analysis by assuming that $|\Theta| = 2$. In this case many of the results of Section 2 extend directly. For example, the same arguments used in Lemma 3 show that we can focus on a message space containing only two messages with at most one type randomizing between them. Similarly, we can extend most of the comparative statics of Proposition 2(a). In particular, a higher probability of separation increases both $\kappa_t(v,\sigma)$ and $W_t(\sigma)$. Moreover, a higher probability of separation allows the government to provide better insurance ($u_{v,\sigma}(m_H) - u_{v,\sigma}(m_L)$ increases). The incentive compatibility is preserved by increasing $w_{v,\sigma}(m_L) - w_{v,\sigma}(m_H)$ as well. The next Proposition is the analogue of Corollaries 1 and 2.

Proposition 4 *Suppose* $|\Theta| = 2$.

(a). Suppose either utility is bounded below or $\chi_{t+1} > 0$. If $\chi_t > 0$ then there exists $v_t^- > \underline{v}$ such that $\sigma_v \in \Sigma^{un}$ and $\boldsymbol{\sigma}_v$ is uninformative for all $v \leq v_t^-$.

(b). If Assumption 1 is satisfied then there exists $v_t^+ < \bar{v}$ such that $\sigma_v \in \Sigma^{in}$ $\boldsymbol{\sigma}_v$ is fully informative for all $v \geq v_t^+$.

The proof of this proposition is in the appendix, here we sketch the main steps. Suppose that the probability of separation is interior, so that some type j plays $\sigma_v(m_j|\theta_j) \in (0,1)$. Consider an uninformative strategy $\sigma^{un}(m_{-j}|\theta_j) = \sigma^{un}(m_{-j}|\theta_{-j}) = 1$ and a fully informative strategy $\sigma^{in}(m_j|\theta_j) = \sigma^{in}(m_{-j}|\theta_{-j}) = 1$. Optimality of σ_v implies the following first order conditions:

$$\chi_t \frac{\partial W_t(\sigma_v)}{\partial \sigma^{un}} = \frac{\partial \kappa_t(v,\sigma_v)}{\partial \sigma^{un}}, \tag{38}$$

$$\chi_t \frac{\partial W_t(\sigma_v)}{\partial \sigma^{in}} = \frac{\partial \kappa_t(v,\sigma_v)}{\partial \sigma^{in}}. \tag{39}$$

The expressions on the left hand side of (38) and (39) capture the marginal cost off the equilibrium path from changing informativeness of agents' strategies. As in Proposition 2(b), the derivatives of W_t are finite and non-zero. The expressions on the right hand side of (38) and (39) capture the marginal gain on the equilibrium path from changing informativeness of agents' strategies. Under the assumptions of Proposition 4, the marginal gain of better information on path becomes arbitrarily small as v approaches $\underline{v}$ and arbitrarily large as v approaches $\bar{v}$. This implies that an interior probability of separation is suboptimal for low and high values of v, and that high-v agents play fully informative strategies while low-v agents play uninformative strategies.

Now consider the case with $|\Theta| > 2$. Part (a) of Proposition 4 extends to this case without any additional considerations. The reason for it is that the marginal gain of more information disappears as $v \to \underline{v}$ for any cardinality of Θ. The result of full information revelation for sufficiently high v requires extra assumptions. In general, when $|\Theta| > 2$ it might be optimal to bunch some types together and give them the same allocations even in the second best problem, that has no cost of information revelation. In such situations revealing full information is strictly suboptimal if $\chi_t > 0$. Part (b) provides sufficient conditions that ensure that it is not optimal to bunch type θ in the second best environment and show that under such conditions $\boldsymbol{\sigma}_v$ reveals full information about that type if v is sufficiently large.

Proposition 5 *For any* $|\Theta|$,

(a). Suppose either utility is bounded below or $\chi_{t+1} > 0$. *If* $\chi_t > 0$ *then there exists* $v_t^- > \underline{v}$ *such that* $\boldsymbol{\sigma}_v$ *is uninformative for all* $v \leq v_t^-$.

(b). If Assumption 1 is satisfied, then there exists $v_t^+ < \bar{v}$ *such that* $\boldsymbol{\sigma}_v$ *reveals full information about* θ_1 *for all* $v \geq v_t^+$. *If, in addition,* $\pi\left(\theta_{|\Theta|-1}\right)\left(\theta_{|\Theta|} - \theta_{|\Theta|-1}\right) > \left(\pi\left(\theta_{|\Theta|-1}\right) + \pi\left(\theta_{|\Theta|}\right)\right)\left(\theta_{|\Theta|-1} - \theta_{|\Theta|-2}\right)$ *and* $1 + \theta_1 - \theta_{|\Theta|} \geq 0$, *then* $\boldsymbol{\sigma}_v$ *reveals full information about* $\theta_{|\Theta|}$ *for all* $v \geq v_t^+$.

Moreover, if $U(c)$ *is CES with* $\rho \in (0,1)$ *and* θ *satisfies the no-bunching condition*

$$\pi\left(\theta_{n-1}\right)\left[\theta_n - \theta_{n-1}\right] - \left(\theta_{n-1} - \theta_{n-2}\right) \sum_{i=n-1}^{|\Theta|} \pi\left(\theta_i\right) \geq 0 \text{ for all } n > 2, \tag{40}$$

then $\boldsymbol{\sigma}_v$ *is fully informative for all* $v \geq v_t^+$.

3.3 Invariant distribution

In our analysis so far we took the initial distribution of utilities ψ as given. Any ψ is associated with Lagrange multipliers $\{\zeta_t, \chi_t, \lambda_t^w\}_{t=0}^{\infty}$ which, together with the Bellman equations (33) and (36), can be used to recover the equilibrium strategies that support ψ. Moreover, any ψ induces a sequence of distributions of continuation utilities of the agents, $v_t = \mathbb{E}_{\boldsymbol{\sigma}^*}\left[\sum_{s=0}^{\infty} \beta^s \theta_{t+s} \mathbf{u}_{t+s}^* \middle| h^{t-1}\right]$, which we denote by ψ_t with $\psi_0 = \psi$. We say that ψ is *invariant* if $\{\zeta_t, \chi_t, \lambda_t^w, \psi_t\}_{t=0}^{\infty}$ do not depend on t. This also implies that in an invariant distribution $\hat{\beta}_t = \beta / (1-\chi)$ does not depend on t.

Lemma 10 *In any invariant distribution $\chi > 0$, $\hat{\beta} > \beta$ and in each period a positive measure of agents does not play uninformative strategies. Continuation utilities are mean-reverting:*

$$\frac{\hat{\beta}}{\beta}\mathbb{E}_{\boldsymbol{\sigma}_v}\left[k'\left(\mathbf{w}_v\right)\right] = k'\left(v\right). \tag{41}$$

If $1+\theta_1-\theta_{|\Theta|} \geq 0$, then there is some $\underline{w} > \underline{v}$ such that $\mathbf{w}_v\left(m,z\right) \geq \underline{w}$ for all $v \in supp\left(\psi\right) \setminus \{\underline{v}\}$ with $\mathbf{w}_{\underline{w}}\left(m,z\right) > \underline{w}$ for some m, z.

This lemma shows that in any invariant distribution the sustainability constraint (22) binds. If it did not, our economy would be isomorphic to Atkeson and Lucas (1992). In that environment immiseration (the distribution that assigns mass 1 on $\underline{v}$) is the only feasible invariant distribution, but such distribution violates (27). The binding constraint (27) also implies that agents' continuation utilities exhibit mean-reversion (41); that in each period some agents reveal information to the government; and that, if any agent enters the region of continuation utilities in which it is optimal to reveal no information, he must exit it in finite time. Finally, as long as the dispersion of shocks is not too high, there is a reflecting lower bound $\underline{w}$ below which agents' continuation utilities do not fall if they start above it.[13]

Figure 2 illustrates the policy functions in an invariant distribution.[14] The optimal reporting strategy $\boldsymbol{\sigma}_v$ follows the same patterns as those in Proposition 3. $\boldsymbol{\sigma}_v\left(\cdot|z,\cdot\right)$ is either fully informative or uninformative for all realizations of (v, z). Agents reveal no information and receive no insurance with probability 1 for all $v \leq v^-$ (v^- is shown by the first dashed line in panel A) and reveal full information and receive the second best insurance with probability 1 for all $v \geq v^+$ (v^+ is shown by the second dashed line in panel A). Finally, insurance is rationed if $v \in (v^-, v^+)$. In this case the agent receives allocations associated with v^+ and reveals full information with probability $\frac{v^+-v}{v^+-v^-}$ and receives no insurance, reveals no information, and obtains utility v^- with probability $\frac{v-v^-}{v^+-v^-}$.

The typical dynamics of v_t in the invariant distribution can be seen from panel B. Consider an agent whose initial lifetime utility v_0 equals the lowest v in the support of the invariant distribution. The continuation utility of such agent initially grows deterministically over time.

[13] There exist invariant distributions that put a positive mass on $\underline{v}$, which is an absorbing state. The probability of reaching this point from any other point in the support of the invarinant point is zero.

[14] To compute this figure we set $U\left(c\right) = \ln\left(c\right)$, $\beta = 0.53$, $e = 1$ and $\Theta = \{0.8, 1.2\}$ with both shocks occuring with equal probability. These assumptions imply that $\lambda^w = 1$. To find an invariant distribution, we compute the stationary distribution implied policy functions to (33) and (36) and iterate on (ζ, χ) until the stationary distribution satisfies constraints (20) and (22).

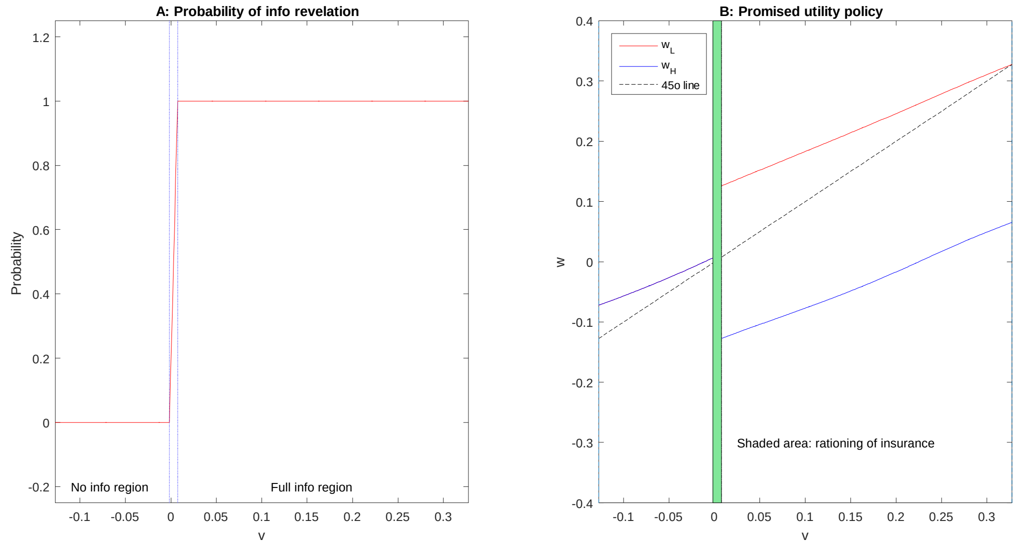

Figure 2: Policy functions in the invariant distribution. Panel A plots the probability with which agent with utility v reveals full info. There is no information revelation for $v \leq v^-$ (the first dashed line) and full information revelation for $v \geq v^+$ (the second dashed line). For $v \in \left(v^-, v^+\right)$ full information is revealed with probability $\frac{v^+ - v}{v^+ - v^-}$ and no information is revealed with probability $\frac{v - v^-}{v^+ - v^-}$. Panel B plots promised utilities $w_v\left(\theta\right)$.

It exits the no information regions in finite time and enters the region in which insurance is rationed. In this region v_t is delivered through a lottery and, depending on the outcome of such lottery, the agent receives either v^- or v^+. Finally, if v_t falls in the region where full information revelation is optimal, then next period it goes up if the agent reports θ_L (red line to the right of the shaded area) or goes down if he reports θ_H (blue line to the right of the shaded area). An agent with a string of θ_L reports stays in the full information region while a sufficiently long sequence of θ_H reports brings him back to the no information region.

4 Autocorrelated shocks

In this section we extend our analysis to first order Markov shocks. Let $\pi\left(\theta|\theta^-\right)$ denote the probability of realization of shock θ conditional on θ^- in the previous period. We assume that $\pi\left(\theta|\theta^-\right) > 0$ for all θ and θ^-. Let $\pi^t\left(\theta|\theta^-\right)$ and $\bar{E}_t\left(\theta^-\right) = \sum_\theta \theta\pi^t\left(\theta|\theta^-\right)$ be, respectively, the probability of realization of shock θ and the expected shock conditional on θ^- being realized t periods ago. The shock in period 0 is assumed to be drawn from some distribution $\bar{\pi} \in \Delta\left(\Theta\right)$. As in the i.i.d. case, each agent belongs to a group $v \in (\underline{v}, \bar{v})$ in period 0 ($[\underline{v}, \bar{v})$ if utility is bounded below) and the distribution of agents over $(\underline{v}, \bar{v})$ is denoted by ψ. In equilibrium members of group v receive lifetime expected utility v.

Many arguments when types are Markov are direct extensions of our previous analysis. We briefly lay out the arguments here and leave the details in Supplementary material. We assume throughout that agents are required to send messages from a finite message space M. Given a reporting strategy $\boldsymbol{\sigma}$, let $\mathbf{p}_t : H^t \to \Delta(\Theta)$ denote the government's belief about the agent's shock conditional on history h^t. These posteriors are generated recursively starting from $\mathbf{p}_0 = \bar{\pi}$ and using Bayes rule

$$\mathbf{p}_t\left(\theta|h^{t-1}, z_t, m_t\right) = \frac{\boldsymbol{\sigma}_t\left(m_t|h^{t-1}, z_t, \theta\right)\sum_{\theta^-}\pi\left(\theta|\theta^-\right)\mathbf{p}_{t-1}\left(\theta^-|h^{t-1}\right)}{\sum_{\theta,\theta^-}\boldsymbol{\sigma}_t\left(m_t|h^{t-1}, z_t, \theta\right)\pi\left(\theta|\theta^-\right)\mathbf{p}_{t-1}\left(\theta^-|h^{t-1}\right)}, \tag{42}$$

for all h^{t-1}, z_t, and m_t for which the expression is well-defined. For any $\mathbf{x} : H^t \times \Theta \to \mathbb{R}$, the expectation of $\mathbf{x}$ conditional on some history $h^{t-1} \in H^{t-1}$ and type $\theta^- \in \Theta$ is $\mathbb{E}_{\boldsymbol{\sigma}}\left[\mathbf{x}|h^{t-1}, \theta^-\right] = \int_{M\times Z}\sum_\theta \mathbf{x}\left(h^{t-1}, z, m, \theta\right)\boldsymbol{\sigma}_t\left(dm|z, \theta\right)\pi\left(\theta|\theta^-\right)dz$. Similarly, the unconditional expectation is $\mathbb{E}_{\boldsymbol{\sigma}}[\mathbf{x}] = \int_{H^{t-1}}\sum_{\theta^-}\mathbb{E}_{\boldsymbol{\sigma}}\left[\mathbf{x}|h, \theta^-\right]\mathbf{p}_{t-1}\left(\theta^-|h\right)d\mu_{t-1}$.

It is immediate to extend Lemma 4 to Markov shocks and show that in the worst equilibrium agents play uninformative strategies. Unlike the i.i.d. case, however, the payoff in that equilibrium depends on the beliefs of the government. The maximum payoff that the government can achieve in any period t is given by

$$\tilde{W}_t(\mu_t) \equiv \max_{\{\mathbf{u}_{t+s}(h)\}_{h\in H^t, s\geq 0}} \mathbb{E}_{\boldsymbol{\sigma}}\left[\sum_{s=0}^{\infty}\beta^s \bar{E}_s \mathbf{u}_{t+s}\right] \tag{43}$$

subject to the feasibility constraints (20).

Similarly to the i.i.d. case, we first bound $\tilde{W}_t(\mu_t)$ with a function that is linear in μ_{t-1}. Let $\left\{\lambda^w_{t,t+s}\right\}_{s=0}^{\infty}$ be the sequence of Lagrange multipliers on (20) in the maximization problem that defines $\tilde{W}_t(\mu^*_t)$. Let $\sigma : \Theta \to \Delta(M)$, the expectation of the random variable $x : M \times \Theta \to \mathbb{R}$ conditional on some $\theta^- \in \Theta$ is now $\mathbb{E}_\sigma\left[x|\theta^-\right] = \sum\limits_{m,\theta\in M\times\Theta} x(m, \theta)\,\sigma(m|\theta)\,\pi\left(\theta|\theta^-\right)$. For any $p \in \Delta(\Theta)$, let $W_t(\sigma, p)$ be defined as

$$W_t(\sigma, p) \equiv \max_{\{u_{t+s}(m)\}_{s\geq 0}} \sum_{\theta^-}\mathbb{E}_\sigma\left[\sum_{s=0}^{\infty}\beta^s\left(\bar{E}_s u_{t+s} - C(u_{t+s}) + \lambda^w_{t,t+s}e\right)\middle|\theta^-\right]p\left(\theta^-\right). \tag{44}$$

$W_t(\sigma, p)$ is the generalization of (26) to the Markov case and p represents the beliefs that the government holds about the agents' types in period $t-1$. $\tilde{W}_t(\mu_t)$ is bounded by

$$\tilde{W}_t(\mu_t) \leq \int_{H^{t-1}\times Z} W_t\left(\boldsymbol{\sigma}_t\left(\cdot|h^{t-1}, z, \cdot\right), \mathbf{p}_{t-1}\left(h^{t-1}\right)\right)dz\,d\mu_{t-1},$$

with equality if $\mu_t = \mu^*_t$, $\boldsymbol{\sigma}_t = \boldsymbol{\sigma}^*_t$, $\mathbf{p}_{t-1} = \mathbf{p}^*_{t-1}$, where $\{\mathbf{p}^*_t\}$ are the beliefs corresponding to $\boldsymbol{\sigma}^*$. This bound can then be used to replace the incentive constraint for the government with a constraint that is linear in μ_{t-1}.

Replacing the incentive constraint for the government, in turn, enables us to use Lagrangian methods and solve the problem recursively. In particular, we first define a Lagrangian and a value function $k_t\left(\overrightarrow{\mathbf{v}}, p\right)$, where now $\overrightarrow{\mathbf{v}} = \left(\overrightarrow{v}\left(\theta_1\right), ..., \overrightarrow{v}\left(\theta_{|\Theta|}\right)\right)$ is a vector of continuation utilities, which are the analogues of (28) and (29), respectively. We then rewrite the value function recursively extending the techniques of Fernandes and Phelan (2000). For any $\mathbf{x}: M \times \Theta \times Z \rightarrow \mathbb{R}$ let $\mathbb{E}_{\boldsymbol{\sigma}}\left[\mathbf{x}|\theta^-\right] = \int_Z \sum_{m,\theta} \pi\left(\theta|\theta^-\right) \boldsymbol{\sigma}\left(m|z,\theta\right) x\left(m,\theta,z\right) dz$. Also, let $\mathbf{u}, \boldsymbol{\sigma}, \mathbf{p}: M \times Z \rightarrow \mathbb{R}$ and $\overrightarrow{\mathbf{w}}: M \times Z \times \Theta \rightarrow \mathbb{R}$. The value function $k_t\left(\overrightarrow{\mathbf{v}}, p\right)$ satisfies

$$k_t\left(\overrightarrow{\mathbf{v}}, p\right) = \max_{\left(\mathbf{u}, \overrightarrow{\mathbf{w}}, \boldsymbol{\sigma}, p'\right)} \sum_{\theta^-} p\left(\theta^-\right) \mathbb{E}_{\boldsymbol{\sigma}}\left[\theta\mathbf{u} - \zeta_t C\left(\mathbf{u}\right) + \hat{\beta}_{t+1} k_{t+1}\left(\overrightarrow{\mathbf{w}}, \mathbf{p}'\right)|\theta^-\right] - \chi_t \int W_t\left(\boldsymbol{\sigma}\left(\cdot|z,\cdot\right), p\right) dz \tag{45}$$

subject to

$$\overrightarrow{v}\left(\theta^-\right) = \mathbb{E}_{\boldsymbol{\sigma}}\left[\theta\mathbf{u} + \beta\overrightarrow{\mathbf{w}}\left(\cdot,\cdot,\theta\right)|\theta^-\right] \text{ for all } \theta^-, \tag{46}$$

$$\mathbb{E}_{\boldsymbol{\sigma}}\left[\theta\mathbf{u} + \beta\overrightarrow{\mathbf{w}}\left(\cdot,\cdot,\theta\right)|\theta, z\right] \geq \mathbb{E}_{\boldsymbol{\sigma}'}\left[\theta\mathbf{u} + \beta\overrightarrow{\mathbf{w}}\left(\cdot,\cdot,\theta\right)|\theta, z\right] \text{ for all } z, \theta, \boldsymbol{\sigma}' \tag{47}$$

and

$$\mathbf{p}'\left(\theta|m,z\right)\left[\sum_{\theta,\theta^-} \boldsymbol{\sigma}\left(m|z,\theta\right) \pi\left(\theta|\theta^-\right) p\left(\theta^-\right)\right] = \boldsymbol{\sigma}\left(m|z,\theta\right) \sum_{\theta^-} \pi\left(\theta|\theta^-\right) p\left(\theta^-\right). \tag{48}$$

Constraints (46) and (47) are the analogues of (31) and (32) in the i.i.d. case. The key difference is that realization of shock θ in the current period affects the expected utility of an agent from the future consumption stream. Thus, the recursive formulation assigns a continuation utility for each possible realization of $\theta^- \in \Theta$. The probability measure p keeps track of the evolution of the posterior beliefs of the government. When $\chi_t = 0$ and agents play fully informative strategy, p assigns probability 1 to one of the values of Θ and the Bellman equation (45) simplifies to the recursive formulation of Fernandes and Phelan (2000). We conclude this section by a version of Proposition 5(a) for Markov shocks, which we prove in Supplementary material.

Proposition 6 *Suppose utility is bounded below (wlog by* 0*) and* $\chi_t > 0$*. Let* $\boldsymbol{\sigma}_{\overrightarrow{\mathbf{v}},p}$ *be a solution to (45), then* $\lim_{\overrightarrow{\mathbf{v}} \rightarrow \mathbf{0}} \Pr\left(\boldsymbol{\sigma}_{\overrightarrow{\mathbf{v}},p} \in \Sigma^{un}\right) = 1$*, uniformly in* p*.*

5 Final remarks

In this paper we took a step towards developing of theory of social insurance in a setting in which the principal cannot commit. We focused on the simplest version of no commitment that involves a direct, one-shot communication between the principal and the agents, and showed how such model can be incorporated into the standard recursive contracting framework with relatively few modifications. The natural extension of this approach is to incorporate it into richer models of social insurance cited in the introduction. This would allow one to explore how the allocations in best equilibria can be decentralized through a system of taxes and transfers, for example along the lines of Albanesi and Sleet (2006). Our methods should also be applicable to other principal-agent environments in which the principal interacts with a large number of agents and cannot commit, such as models of regulation, employer-employee relationships, bargaining and trading with private information.

References

Acemoglu, D. (2003): "Why not a political Coase theorem? Social conflict, commitment, and politics," *Journal of Comparative Economics*, 31(4), 620–652.

Acemoglu, D., M. Golosov, and A. Tsyvinski (2010): "Dynamic Mirrlees Taxation under Political Economy Constraints," *Review of Economic Studies*, 77(3), 841–881.

Aiyagari, S. R., and F. Alvarez (1995): "Efficient Dynamic Monitoring of Unemployment Insurance Claims," *Mimeo, University of Chicago*.

Akerlof, G. A. (1978): "The Economics of "Tagging" as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning," *The American Economic Review*, 68(1), pp. 8–19.

Albanesi, S., and C. Sleet (2006): "Dynamic optimal taxation with private information," *Review of Economic Studies*, 73(1), 1–30.

Atkeson, A., and R. E. Lucas (1992): "On Efficient Distribution with Private Information," *Review of Economic Studies*, 59(3), 427–453.

Besley, T., and S. Coate (1998): "Sources of Inefficiency in a Representative Democracy: A Dynamic Analysis," *American Economic Review*, 88(1), 139–56.

Bester, H., and R. Strausz (2001): "Contracting with Imperfect Commitment and the Revelation Principle: The Single Agent Case," *Econometrica*, 69(4), 1077–98.

Bisin, A., and A. Rampini (2006): "Markets as beneficial constraints on the government," *Journal of Public Economics*, 90(4-5), 601–629.

Chari, V. V., and P. J. Kehoe (1990): "Sustainable Plans," *Journal of Political Economy*, 98(4), 783–802.

——— (1993): "Sustainable Plans and Debt," *Journal of Economic Theory*, 61(2), 230–261.

Clementi, G. L., and H. A. Hopenhayn (2006): "A Theory of Financing Constraints and Firm Dynamics," *The Quarterly Journal of Economics*, 121(1), 229–265.

Cole, H. L., and N. Kocherlakota (2001): "Dynamic Games with Hidden Actions and Hidden States," *Journal of Economic Theory*, 98(1), 114 – 126.

Dovis, A. (2009): "Efficient Sovereign Default," *Mimeo, Penn State*.

Farhi, E., C. Sleet, I. Werning, and S. Yeltekin (2012): "Non-linear Capital Taxation Without Commitment," *Review of Economic Studies*, 79(4), 1469–1493.

Farhi, E., and I. Werning (2007): "Inequality and Social Discounting," *Journal of Political Economy*, 115(3), 365–402.

——— (2013): "Insurance and Taxation over the Life Cycle," *Review of Economic Studies*, 80(2), 596–635.

Fernandes, A., and C. Phelan (2000): "A Recursive Formulation for Repeated Agency with History Dependence," *Journal of Economic Theory*, 91(2), 223–247.

Freixas, X., R. Guesnerie, and J. Tirole (1985): "Planning under Incomplete Information and the Ratchet Effect," *Review of Economic Studies*, 52(2), 173–91.

Golosov, M., M. Troshkin, and A. Tsyvinski (2016): "Redistribution and Social Insurance," *American Economic Review*, 106(2), 359–86.

Golosov, M., and A. Tsyvinski (2006): "Designing optimal disability insurance: A case for asset testing," *Journal of Political Economy*, 114(2), 257–279.

Golosov, M., A. Tsyvinski, and I. Werning (2006): "New dynamic public finance: A user's guide," *NBER Macroeconomics Annual*, 21, 317–363.

Golosov, M., A. Tsyvinski, and N. Werquin (2013): "Recursive Contracts and Endogenously Incomplete Markets," *working paper*.

Green, E. J. (1987): "Lending and the Smoothing of Uninsurable Income," in *Contractual Arrangements for Intertemporal Trade*, ed. by E. C. Prescott, and N. Wallace. University of Minnesota Press., Minneapolis.

Hopenhayn, H. A., and J. P. Nicolini (1997): "Optimal Unemployment Insurance," *Journal of Political Economy*, 105(2), 412–38.

Kocherlakota, N. (2010): *The New Dynamic Public Finance*. Princeton University Press, USA.

Kydland, F. E., and E. C. Prescott (1977): "Rules Rather Than Discretion: The Inconsistency of Optimal Plans," *Journal of Political Economy, University of Chicago Press*, 85(3), 473–91.

Laffont, J.-J., and J. Tirole (1988): "The Dynamics of Incentive Contracts," *Econometrica*, 56(5), 1153–75.

Lindbeck, A., and J. W. Weibull (1987): "Balanced-Budget Redistribution as the Outcome of Political Competition," *Public choice*, 52(3), 273–297.

Ljungqvist, L., and T. J. Sargent (2012): *Recursive Macroeconomic Theory, Third Edition*. MIT Press.

Luenberger, D. (1969): *Optimization by Vector Space Methods*. Wiley-Interscience.

Marcet, A., and R. Marimon (2009): "Recursive Contracts," *Mimeo, European University Institute.*

Milgrom, P., and I. Segal (2002): "Envelope Theorems for Arbitrary Choice Sets," *Econometrica*, 70(2), 583–601.

Mirrlees, J. (1971): "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, 38(2), 175–208.

Myerson, R. B. (1982): "Optimal Coordination Mechanisms in Generalized Principal-Agent Problems," *Journal of Mathematical Economics*, 10(1), 67–81.

Phelan, C., and R. M. Townsend (1991): "Computing Multi-period, Information-Constrained Optima," *Review of Economic Studies, Wiley Blackwell*, 58(5), 853–81.

Roberts, K. (1984): "The Theoretical Limits of Redistribution," *Review of Economic Studies*, 51(2), 177–95.

Rockafellar, R. (1972): *Convex Analysis*, Princeton mathematical series. Princeton University Press.

Royden, H. (1988): *Real Analysis*, Mathematics and statistics. Macmillan.

Scheuer, F., and A. Wolitzky (2014): "Capital Taxation under Political Constraints," *NBER Working Paper 20043*.

Shimer, R., and I. Werning (2015): "Efficiency and Information Transmission in Bilateral Trading," Working Paper 21495, National Bureau of Economic Research.

Skreta, V. (2006): "Sequentially Optimal Mechanisms," *Review of Economic Studies*, 73(4), 1085–1111.

——— (2015): "Optimal auction design under non-commitment," *Journal of Economic Theory*, 159, Part B, 854 – 890.

Sleet, C., and S. Yeltekin (2006): "Credibility and endogenous societal discounting," *Review of Economic Dynamics*, 9(3), 410–437.

——— (2008): "Politically credible social insurance," *Journal of Monetary Economics*, 55(1), 129–151.

Song, Z., K. Storesletten, and F. Zilibotti (2012): "Rotten Parents and Disciplined Children: A Politico-Economic Theory of Public Expenditure and Debt," *Econometrica*, 80(6), 2785–2803.

Stantcheva, S. (2014): "Optimal Taxation and Human Capital Policies over the Life Cycle," *working paper*.

Thomas, J., and T. Worrall (1990): "Income Fluctuation and Asymmetric Information: An Example of a Repeated Principal-agent Problem," *Journal of Economic Theory*, 51(2), 367–390.

Topkis, D. (2011): *Supermodularity and Complementarity*, Frontiers of Economic Research. Princeton University Press.

Yared, P. (2010): "A dynamic theory of war and peace," *Journal of Economic Theory*, 145(5), 1921–1950.

6 Appendix

6.1 Proofs of Section 2

Proof of Lemma 2. It is immediate that the feasibility constraint (4) binds for any $\boldsymbol{\sigma}$ and therefore $\lambda^w > 0$. The first order condition to (9) gives

$$C'\left(u^w\left(m\right)\right) = \frac{1}{\lambda^w}\mathbb{E}_\sigma\left[\theta|m\right] \in \left[\frac{\theta_L}{\lambda^w}, \frac{\theta_H}{\lambda^w}\right] \tag{49}$$

for all m sent with positive probability, therefore such $u^w(m)$ lie in a compact set. Since messages sent with zero probability do not affect the value of W, $u^w(m)$ can be restricted to lie in a compact set for all m. The theorem of the maximum then implies that W is continuous.

For any σ', σ, and $\alpha \in [0,1]$ define $\sigma_\alpha = (1-\alpha)\sigma + \alpha\sigma'$.

$$\begin{aligned}
W\left(\sigma_\alpha\right) &= \max_u \alpha \sum_{m,\theta}\left[\theta u\left(m\right) - \lambda^w C\left(u\left(m\right)\right)\right]\sigma'\left(m|\theta\right)\pi\left(\theta\right) \\
&\quad + \left(1-\alpha\right)\sum_{m,\theta}\left[\theta u\left(m\right) - \lambda^w C\left(u\left(m\right)\right)\right]\sigma\left(m|\theta\right)\pi\left(\theta\right) + \lambda^w \\
&\leq \alpha \max_u \sum_{m,\theta}\left[\theta u\left(m\right) - \lambda^w C\left(u\left(m\right)\right)\right]\sigma'\left(m|\theta\right)\pi\left(\theta\right) \\
&\quad + \left(1-\alpha\right)\max_u \sum_{m,\theta}\left[\theta u\left(m\right) - \lambda^w C\left(u\left(m\right)\right)\right]\sigma\left(m|\theta\right)\pi\left(\theta\right) + \lambda^w \\
&= \alpha W\left(\sigma'\right) + \left(1-\alpha\right)W\left(\sigma\right),
\end{aligned}$$

which establishes convexity. Note that for any collection X of functions $x : M \to \mathbb{R}$, the family $\{\mathbb{E}_{\sigma_\alpha} x\}_{x\in X}$ is equidifferentiable at any $\alpha \in [0,1)$ since the expectation is linear in α. Therefore, the derivative $\frac{\partial W(\sigma)}{\partial \sigma'}$ exists by Theorem 3 in Milgrom and Segal (2002) and

$$\frac{\partial W\left(\sigma\right)}{\partial \sigma'} = \mathbb{E}_{\sigma'}\left[\theta u_0^w\left(m\right) - \lambda^w C\left(u_0^w\left(m\right)\right)\right] - \mathbb{E}_\sigma\left[\theta u_0^w\left(m\right) - \lambda^w C\left(u_0^w\left(m\right)\right)\right] \leq W\left(\sigma'\right) - W\left(\sigma\right), \tag{50}$$

where u_α^w is a solution to $W(\sigma_\alpha)$ for $\alpha > 0$ and $u_0^w(m) = \lim_{\alpha\to 0} u_\alpha^w(m)$.[15]

To show that $W(\sigma)$ achieves its minimum if and only if σ is uninformative, let u^{un} be the optimal allocation corresponding to an uninformative strategy, which without loss of generality

[15] The problem that defines W, (9), is strictly convex and, therefore, the solution $u_\alpha^w(m)$ is unique for each m sent with positive probability by σ_α. The definition of u_0^w pins down the values of $u_0^w(m)$ for which $\sum_\theta \sigma_\alpha(m|\theta)\pi(\theta) > 0$ for $\alpha > 0$ and $\lim_{\alpha\to 0}\sum_\theta \sigma_\alpha(m|\theta)\pi(\theta) = 0$.

satisfies $C'(u^{un}(m)) = 1/\lambda^w$ for all m. For any σ

$$\begin{aligned}
& W(\sigma) - W(\sigma^{un}) \\
= \ & \max_u \sum_{m,\theta} \left[(\theta u(m) - \lambda^w C(u(m))) - (\theta u^{un} - \lambda^w C(u^{un})) \right] \sigma(m|\theta) \pi(\theta) \\
= \ & \max_u \sum_m \left\{ (\mathbb{E}_\sigma[\theta|m] u(m) - \lambda^w C(u(m))) - (\mathbb{E}_\sigma[\theta|m] u^{un} - \lambda^w C(u^{un})) \right\} \left(\sum_\theta \sigma(m|\theta) \pi(\theta) \right).
\end{aligned}$$

The expression in curly bracket is non-negative, which implies that $W(\sigma) \geq W(\sigma^{un})$ for all σ. If $\sigma \notin \Sigma^{un}$ then $C'(u^w(m)) \neq 1/\lambda^w$ for at least one m sent with positive probability. For such m the expression in the curly brackets is strictly positive so that $W(\sigma) > W(\sigma^{un})$ if $\sigma \notin \Sigma^{un}$.

To show that $W(\sigma)$ achieves its maximum if and only if σ is fully informative, take any $\sigma \notin \Sigma^{in}$. By definition there must exist some message m sent with positive probability such that $\mathbb{E}_\sigma[\theta|m] \neq \theta_j$ for $j \in \{L, H\}$. By (49) the optimal allocation for such m satisfies $C'(u^w(m)) \neq \theta_j/\lambda^w$ for $j \in \{L, H\}$. Let $\hat{u}^{in}$ be the optimal allocation corresponding to some σ^{in}. It satisfied $C'\left(\hat{u}^{in}(m)\right) = \theta_j/\lambda^w$, $j \in \{L, H\}$ for all m sent with positive probability. By strict convexity of C the optimal allocation must be unique and therefore $W\left(\sigma^{in}\right) > W(\sigma)$ for all $\sigma \notin \Sigma^{in}$. ■

We first prove a preliminary result that is useful in the proof of both Lemma 3 and of the results in Section 3.2.

Lemma 11 *Any point on the Pareto frontier can be supported by reporting strategies such that all agents report one of three messages, with each $\theta \in \{\theta_L, \theta_H\}$ randomizing between at most two messages and with at most one message reported with positive probability by both θ.*

Proof. Fix any group i and let (u_1^*, u_2^*, σ^*) be a best equilibrium strategy for that group. We can partition M into four subset: (i) a subset M_L that consists of messages reported with positive probability by type θ_L and reporting which gives strictly lower utility to type θ_H, i.e. there exists a message $m \in M$ such that

$$\theta_H u_1^*(m) + u_2^*(m) > \theta_H u_1^*\left(m'\right) + u_2^*\left(m'\right) \text{ for all } m' \in M_L;$$

(ii) a subset M_H defined analogously for θ_H; (iii) a subset M_{HL} that consists of messages reported with positive probability by either θ_H or θ_L and for which

$$\theta u_1^*(m) + u_2^*(m) \geq \theta u_1^*\left(m'\right) + u_2^*\left(m'\right) \text{ for all } \theta \in \Theta, m \in M_{HL},\ m' \in M; \tag{51}$$

(iv) and a subset $M_\varnothing$ containing all other messages.

Consider the subset M_L. Bayes' rule implies $\mathbb{E}_{\sigma^*}[\theta|m] = \theta_L$ for any $m \in M_L$. If $(u_1^*(m), u_2^*(m))$ take the same values for all $m \in M_L$ then an alternative strategy of reporting any $m \in M_L$

with probability 1 gives the same allocations, the same payoff on equilibrium path, and by (49) also the same payoff off the equilibrium path. Thus, this alternative strategy is payoff equivalent to the original strategy. We now rule out the possibility that $(u_1^*(m'), u_2^*(m')) \neq (u_1^*(m''), u_2^*(m''))$ for some $m', m'' \in M_L$. Let $(u_1^\alpha, u_2^\alpha) = \alpha(u_1^*(m'), u_2^*(m')) + (1-\alpha)(u_1^*(m''), u_2^*(m''))$ for $\alpha \in (0,1)$. (u_1^α, u_2^α) gives the same utility to θ_L as $(u_1^*(m'), u_2^*(m'))$ and $(u_1^*(m''), u_2^*(m''))$ and strictly lower utility to type θ_H than any $m''' \in M_H \cup M_{HL}$. (u_1^*, u_2^*) must be a solution to the minimization problem (13) for (v_i, σ^*). Since the objective function in (13) is strictly convex, replacing $(u_1^*(m'), u_2^*(m'))$ and $(u_1^*(m''), u_2^*(m''))$ with (u_1^α, u_2^α) gives a strictly lower value of the objective function, contradicting the optimality of (u_1^*, u_2^*). Analogous arguments apply to M_H.

Consider the subset M_{HL} and suppose $\chi > 0$ (otherwise the result follows directly from the standard revelation principle). Condition (51) implies that $(u_1^*(m), u_2^*(m))$ takes the same values for all $m \in M_{HL}$. If $\mathbb{E}_{\sigma^*}[\theta|m]$ also takes the same value for any $m \in M_{HL}$ then we can replace the subset M_{HL} with one message. We next rule out that $\mathbb{E}_{\sigma^*}[\theta|m'] \neq \mathbb{E}_{\sigma^*}[\theta|m'']$ for some $m', m'' \in M_{HL}$.

Fix any $\hat{m} \in M_{HL}$. Consider an alternative strategy σ' that coincides with σ^* except for $\sigma'(\hat{m}|\theta) = \sum_{m \in M_{HL}} \sigma^*(m|\theta)$, $\sigma'(m'|\theta) = 0$, for all $m' \in M_{HL}$, $m' \neq \hat{m}$, and all θ. The strategy profile $\sigma^\alpha = (1-\alpha)\sigma' + \alpha\sigma^*$ satisfies (6) and (8) for any $\alpha \in [0,1]$. Since $(u_1^*(m), u_2^*(m))$ takes the same value for all $m \in M_{HL}$, the optimality condition for (14) can be written as $\frac{\partial W(\sigma^*)}{\partial \sigma'} \geq 0$, where the derivative exists by Lemma 9. From (50),

$$
\begin{aligned}
0 \quad \leq \quad & \frac{\partial W(\sigma^*)}{\partial \sigma'} = \sum_{\theta, m \in M_{HL}} (\theta u^w(\hat{m}) - \lambda^w C(u^w(\hat{m})))\sigma'(m|\theta)\pi(\theta) \\
& - \sum_{\theta, m \in M_{HL}} (\theta u^w(m) - \lambda^w C(u^w(m)))\sigma^*(m|\theta)\pi(\theta) \\
= \quad & (\mathbb{E}_{\sigma'}[\theta|\hat{m}]\, u^w(\hat{m}) - \lambda^w C(u^w(\hat{m}))) \sum_{\theta} \sigma'(\hat{m}|\theta)\pi(\theta) \\
& - \sum_{m \in M_{HL}} (\mathbb{E}_{\sigma^*}[\theta|m]\, u^w(m) - \lambda^w C(u^w(m))) \left(\sum_{\theta} \sigma^*(m|\theta)\pi(\theta) \right).
\end{aligned}
$$

By construction,

$$
\mathbb{E}_{\sigma'}[\theta|\hat{m}] = \frac{\sum_\theta \theta\sigma'(\hat{m}|\theta)\pi(\theta)}{\sum_\theta \sigma'(\hat{m}|\theta)\pi(\theta)} = \frac{\sum_{\theta, m \in M_{HL}} \theta\sigma^*(m|\theta)\pi(\theta)}{\sum_{\theta, m \in M_{HL}} \sigma^*(m|\theta)\pi(\theta)} = \mathbb{E}_{\sigma^*}[\theta|M_{HL}].
$$

Therefore, the expression above can be re-written as

$$
\mathbb{E}_{\sigma^*}[\theta|M_{HL}]\, u^w(\hat{m}) - \lambda^w C(u^w(\hat{m})) \geq \mathbb{E}_{\sigma^*}[(\mathbb{E}_{\sigma^*}[\theta|m]\, u^w(m) - \lambda^w C(u^w(m)))|M_{HL}].
$$

The expression on the right hand side does not depend on m. The expression on the left hand side holds for all $\hat{m} \in M_{HL}$. Multiply both sides by $\frac{\sum_\theta \sigma^*(\hat{m}|\theta)\pi(\theta)}{\sum_{\theta, m\in M_{HL}} \sigma^*(m|\theta)\pi(\theta)}$ and sum across all $\hat{m} \in M_{HL}$ to get

$$\mathbb{E}_{\sigma^*}[\theta|M_{HL}]\,\mathbb{E}_{\sigma^*}[u^w(m)|M_{HL}] \geq \mathbb{E}_{\sigma^*}[(\mathbb{E}_{\sigma^*}[\theta|m]\,u^w(m))|M_{HL}],$$

which implies that $cov\left(\mathbb{E}_{\sigma^*}[\theta|m], u^w(m)\right) \leq 0$. On the other hand, (49) implies that $u^w(m)$ is monotonically increasing in $\mathbb{E}_{\sigma^*}[\theta|m]$, thus, $cov\left(\mathbb{E}_{\sigma^*}[\theta|m], u^w(m)\right) \geq 0$. The two conditions can be satisfied only if $\mathbb{E}_{\sigma^*}[\theta|m]$ takes the same values for all $m \in M_{HL}$.

Finally, any messages that are sent with zero probability can be dropped, so $M_\varnothing$ can be eliminated from the message set. Thus, it is enough that M has at most three messages, one for each subsets M_L, M_H and M_{HL}. If any of the subsets M_L, M_H or M_{HL} is empty, we can add additional messages reported with zero probability, which proves the statement of the lemma. ■

Proof of Lemma 3. By Lemma 11, we can restrict attention to a message space that consists of 3 messages $M = \{m_L, m_H, m_{HL}\}$, in which type θ_L randomizes between m_L and m_{HL} and type θ_H randomizes between m_H and m_{HL}. We show in this lemma that it is suboptimal to have interior reporting probabilities for both types and $(u_1^*(m_L), u_2^*(m_L)) \neq (u_1^*(m_H), u_2^*(m_H)) \neq (u_1^*(m_{HL}), u_2^*(m_{HL}))$. Given the arguments of Lemma 11 this is sufficient to establish that M can be restricted to two messages. We assume $\chi > 0$, otherwise the result is trivial.

We argue by contradiction. Suppose $\sigma^*(m_j|\theta_j), \sigma^*(m_{HL}|\theta_j) \in (0,1)$ for $j \in \{H, L\}$. Consider strategy $\sigma_{[s]}$ defined by $\sigma_{[s]}(m_{HL}|\theta_j) = (1-s)\,\sigma^*(m_{HL}|\theta_j)$, $\sigma_{[s]}(m_j|\theta_j) = 1 - \sigma_{[s]}(m_{HL}|\theta_j)$ for all j, for $s \in [-\varepsilon, 1]$, for small $\varepsilon > 0$. Since $\sigma^*(m_{HL}|\theta_j) < 1$ for all j, there exist $\varepsilon > 0$ for which $\sigma_{[s]}$ is a well-defined reporting strategy. Let $f(v,s) \equiv \kappa\left(v, \sigma_{[s]}\right)$ and $g(s) = \chi W\left(\sigma_{[s]}\right)$.

Since type j reports m_j and m_{HL} for all j and $s < 1$, we can write

$$f(v,s) = \min_{\{u_t\}_t} \mathbb{E}_{\sigma_{[s]}} \sum_t \zeta_t C(u_t)$$

subject to, for each $j \in \{L, H\}$ and $-j \in \{L, H\}$ with $-j \neq j$,

$$\begin{aligned} \theta_j u_1(m_j) + u_2(m_j) &= \theta_j u_1(m_{HL}) + u_2(m_{HL}), && (52)\\ \theta_j u_1(m_j) + u_2(m_j) &\geq \theta_j u_1(m_{-j}) + u_2(m_{-j}), \\ v &= \sum_j \pi(\theta_j)\left[\theta_j u_1(m_j) + u_2(m_j)\right]. \end{aligned}$$

Let $u_{t,[s]}$ be a solution to this problem as a function of s. Note that $\left\{u_{t,[0]}\right\}_t = \{u_t^*\}_t$ and that $\left\{u_{t,[1]}\right\}$ is the optimal solution to a fully informative strategy. $f(v,s)$ is differentiable in s (see the proof of Lemma 2) with

$$\begin{aligned}\frac{\partial}{\partial s} f(v,s) &= \pi(\theta_L)\sigma^*(m_{HL}|\theta_L)\left[\sum_t \zeta_t C\left(u_{t,[s]}(m_L)\right) - \sum_t \zeta_t C\left(u_{t,[s]}(m_{HL})\right)\right] \\ &\quad + \pi(\theta_H)\sigma^*(m_{HL}|\theta_H)\left[\sum_t \zeta_t C\left(u_{t,[s]}(m_H)\right) - \sum_t \zeta_t C\left(u_{t,[s]}(m_{HL})\right)\right].\end{aligned}$$

Similar considerations imply

$$\begin{aligned}\frac{\partial}{\partial s} g(s) &= \chi\pi(\theta_L)\sigma^*(m_{HL}|\theta_L)\left[\begin{array}{c}\left(\theta_L u^w_{[s]}(m_L) - \lambda^w C\left(u^w_{[s]}(m_L)\right)\right) \\ -\left(\theta_L u^w_{[s]}(m_{HL}) - \lambda^w C\left(u^w_{[s]}(m_{HL})\right)\right)\end{array}\right] \\ &\quad + \chi\pi(\theta_H)\sigma^*(m_{HL}|\theta_H)\left[\begin{array}{c}\left(\theta_H u^w_{[s]}(m_H) - \lambda^w C\left(u^w_{[s]}(m_H)\right)\right) \\ -\left(\theta_H u^w_{[s]}(m_{HL}) - \lambda^w C\left(u^w_{[s]}(m_{HL})\right)\right)\end{array}\right].\end{aligned}$$

Note that

$$f(v,s)+g(s) = \sum_{j\in\{H,L\}} \pi(\theta_j)\left[\sum_t \zeta_t C\left(u_{t,[s]}(m_j)\right) + \chi\left\{\theta_j u^w_{[s]}(m_j) - \lambda^w C\left(u^w_{[s]}(m_j)\right)\right\}\right] - \frac{\partial}{\partial s}\left[f(v,s)+g(s)\right].$$

If σ^* is optimal, then $\frac{\partial}{\partial s}\left[f(v,s)+g(s)\right]\big|_{s=0} = 0$ and, therefore,

$$\begin{aligned}&\sum_{j\in\{H,L\}} \pi(\theta_j)\left[\sum_t \zeta_t C\left(u_{t,[0]}(m_j)\right) + \chi\left\{\theta_j u^w_{[0]}(m_j) - \lambda^w C\left(u^w_{[0]}(m_j)\right)\right\}\right] = f(v,0)+g(0) \\ &\leq f(v,1)+g(1) = \sum_{j\in\{H,L\}} \pi(\theta_j)\left[\sum_t \zeta_t C\left(u_{t,[1]}(m_j)\right) + \chi\left\{\theta_j u^w_{[1]}(m_j) - \lambda^w C\left(u^w_{[1]}(m_j)\right)\right\}\right],\end{aligned}$$

where the inequality follows from the fact that $s = 0$ minimizes $f(v,s)+g(s)$. From (49) $u^w_{[s]}(m_j) = C'^{-1}\left(\frac{\theta_j}{\lambda^w}\right)$ for all s, which implies that

$$\sum_{j\in\{H,L\},t} \pi(\theta_j)\zeta_t C\left(u_{t,[0]}(m_j)\right) \leq \sum_{j\in\{H,L\},t} \pi(\theta_j)\zeta_t C\left(u_{t,[1]}(m_j)\right). \tag{53}$$

On the other hand, $\left\{u_{t,[1]}\right\}_t$ is the unique solution (up to measure 0 messages) that minimizes the right hand side of (53) subject to (52) and, therefore,

$$\sum_{j\in\{H,L\},t} \pi(\theta_j)\zeta_t C\left(u_{t,[0]}(m_j)\right) \geq \sum_{j\in\{H,L\},t} \pi(\theta_j)\zeta_t C\left(u_{t,[1]}(m_j)\right). \tag{54}$$

Incentive compatibility implies that $\theta_j u_{1,[1]}(m_j) + u_{2,[1]}(m_j) = \theta_j u_{1,[1]}(m_{-j}) + u_{2,[1]}(m_{-j})$ for some j and $-j$ with $-j \neq j$. Since $\theta_j u_{1,[0]}(m_j) + u_{2,[0]}(m_j) > \theta_j u_{1,[0]}(m_{-j}) + u_{2,[0]}(m_{-j})$ for all $j, -j$ with $-j \neq j$ by assumption, inequality (54) must be strict, establishing a contradiction. ■

Lemma 12 *Suppose only type θ_j plays a mixed strategy for some v, j. Then the optimal allocation given this reporting strategy is characterized by the solution to $\kappa^j(v,s)$.*

Proof. By Lemma 3 it is enough to consider only two messages and at most one type randomizing between them. Suppose θ_L randomizes, the constraint set defined by (52) reduces to

$$\begin{aligned} \theta_L u_1(m_L) + u_2(m_L) &= \theta_L u_1(m_H) + u_2(m_H), \\ \theta_H u_1(m_H) + u_2(m_H) &\geq \theta_H u_1(m_L) + u_2(m_L), \\ v &= \sum_j \pi(\theta_j)\left[\theta_j u_1(m_j) + u_2(m_j)\right]. \end{aligned} \tag{55}$$

The constraint set defined by (16), (17) is larger than the one defined by (55). We therefore want to show that any solution to (15) satisfies (55). We assume $v > \underline{v}$ since otherwise the result is trivial.

Consider a relaxed minimization problem (15) in which we replace (17) with

$$\theta_L u_1(m_L) + u_2(m_L) \geq \theta_L u_1(m_H) + u_2(m_H). \tag{56}$$

In the relaxed problem constraint (56) binds. The solution $\left\{u_t^{RL}(m_k)\right\}_{k\in\{H,L\},t}$ is unique and satisfies $u_1^{RL}(m_H) > u_1^{RL}(m_L)$. We want to show that it is incentive compatible for θ_H to report m_H. Suppose not, so that

$$\theta_H u_1^{RL}(m_H) + u_2^{RL}(m_H) < \theta_H u_1^{RL}(m_L) + u_2^{RL}(m_L).$$

Sum with (56) and re-arrange to show $(\theta_H - \theta_L) u_1^{RL}(m_H) < (\theta_H - \theta_L) u_1^{RL}(m_L)$, which is a contradiction.

If θ_H randomizes we follow the same steps but replace (18) with

$$\theta_H u_1(m_H) + u_2(m_H) \leq \theta_H u_1(m_L) + u_2(m_L).$$

■

Proof of Proposition 2. (a). We show this result for $j = L$, the other case is similar. We can use (16) and (17) to express $u_1(m_L), u_2(m_L), u_1(m_H)$ as functions of w, Δ:

$$u_1(m_H) = v - w, u_2(m_H) = w, u_1(m_L) = v - w - \frac{\Delta}{\theta_L}, u_2(m_L) = w + \Delta. \tag{57}$$

Using these definitions, write $\kappa^L(v,s)$ as

$$\kappa^L(v,s) = \min_{w,\Delta}(1 - s\pi_L)\left[\underbrace{\zeta_1 C(v-w) + \zeta_2 C(w)}_{\equiv g(w)}\right] + s\pi_L\left[\underbrace{\zeta_1 C\left(v - w - \frac{\Delta}{\theta_L}\right) + \zeta_2 C(w+\Delta)}_{\equiv f(w,\Delta)}\right]. \tag{58}$$

The optimality conditions for Δ and w are, respectively,

$$-\frac{1}{\theta_L}\zeta_1 C'\left(v-w-\frac{\Delta}{\theta_L}\right)+\zeta_2 C'(w+\Delta)=0, \tag{59}$$

$$(1-s\pi_L)\left[-\zeta_1 C'(v-w)+\zeta_2 C'(w)\right]+s\pi_L\left[-\zeta_1 C'\left(v-w-\frac{\Delta}{\theta_L}\right)+\zeta_2 C'(w+\Delta)\right]=0. \tag{60}$$

These conditions imply that in the optimum, (w^*,Δ^*), we have $f_\Delta(w^*,\Delta^*)=0$ and $g_w(w^*)\le 0$ where f_Δ and g_w denote (partial) derivatives. Moreover,

$$\begin{aligned} f_\Delta(w^*,0) &= -\frac{1}{\theta_L}\zeta_1 C'(v-w^*)+\zeta_2 C'(w^*) \\ &\le -\zeta_1 C'(v-w^*)+\zeta_2 C'(w^*)=g_w(w^*)\le 0. \end{aligned}$$

Strict convexity of $f(w^*,\cdot)$ then implies that $f(w^*,\Delta^*)\le f(w^*,0)$ and $\Delta^*\ge 0$. The latter gives $u_1^*(m_H)\ge u_1^*(m_L), u_2^*(m_L)\ge u_2^*(m_H)$.

To prove that $\kappa^L(v,\cdot)$ is decreasing we first show that it is differentiable. Observe that constraint (16) can equivalently be replaced with

$$v=\pi_L\left[\theta_L u_1(m_L)+u_2(m_L)\right]+\pi_H\left[\theta_H u_1(m_H)+u_2(m_H)\right]. \tag{61}$$

Since (17) and (61) do not depend on s, differentiability follows from the envelope theorem of Milgrom and Segal (2002). Using the definition of f

$$\frac{\partial}{\partial s}\kappa^L(v,s)=f(w^*,\Delta^*)-f(w^*,0)\le 0, \tag{62}$$

so that $\kappa^L(v,\cdot)$ is decreasing. Analogous arguments applied to W^L show that $\frac{\partial}{\partial s}W^L(s)\ge 0$.

Let (w^{**},Δ^{**}) and (w^*,Δ^*) be the solutions for $s^{**}\ge s^*$. We must have

$$f(w^{**},\Delta^{**})\le f(w^*,\Delta^*),\ g(w^{**})\ge g(w^*),$$

otherwise they cannot be solutions to (58). Since g is strictly convex with $g_w(w^*), g_w(w^{**})\le 0$, $g(w^{**})\ge g(w^*)$ implies $w^{**}\le w^*$. It also implies $u_1^{**}(m_H)\ge u_1^*(m_H)$ from (57).

We want to show that $\Delta^{**}\ge\Delta^*$. Suppose $\Delta^{**}<\Delta^*$. Then $C'(w^{**}+\Delta^{**})<C'(w^*+\Delta^*)$ and, therefore, (59) implies $C'\left(v-w^{**}-\frac{\Delta^{**}}{\theta_L}\right)<C'\left(v-w^*-\frac{\Delta^*}{\theta_L}\right)$. This implies a contradiction

$$0>\frac{1}{\theta_L}(\Delta^{**}-\Delta^*)>w^*-w^{**}\ge 0.$$

Thus $\Delta^{**}\ge\Delta^*$ and therefore $u_2^{**}(m_L)-u_2^{**}(m_H)\ge u_2^*(m_L)-u_2^*(m_H)$ and $u_1^{**}(m_H)-u_1^{**}(m_L)\ge u_1^*(m_H)-u_1^*(m_L)$.

We next show that $u_2^{**}(m_L) \leq u_2^*(m_L)$. Suppose $u_2^{**}(m_L) > u_2^*(m_L)$, which is equivalent to $w^{**}+\Delta^{**} > w^*+\Delta^*$ from (57). Then (59) implies $C'\left(v - w^{**} - \frac{\Delta^{**}}{\theta_L}\right) > C'\left(v - w^* - \frac{\Delta^*}{\theta_L}\right)$. Substituting for $u_2(m_L)$ we get

$$-u_2^{**}(m_L) - \frac{1-\theta_L}{\theta_L}\Delta^{**} > -u_2^*(m_L) - \frac{1-\theta_L}{\theta_L}\Delta^*.$$

This implies a contradiction

$$0 \leq \frac{1-\theta_L}{\theta_L}(\Delta^{**}-\Delta^*) < u_2^*(m_L) - u_2^{**}(m_L) < 0.$$

$u_2^{**}(m_L) \leq u_2^*(m_L)$ implies $C'\left(v - w^{**} - \frac{\Delta^{**}}{\theta_L}\right) \leq C'\left(v - w^* - \frac{\Delta^*}{\theta_L}\right)$, which is equivalent to $u_1^{**}(m_L) \leq u_1^*(m_L)$ by (57).

Utility of type θ_L is $\theta_L u_1(m_L) + u_2(m_L)$. It decreases in s since we showed that both $u_1(m_L)$ and $u_2(m_L)$ decrease in s. Since the weighted sum of utilities of the two types is constant by (61), the utility of θ_H increases in s.

Similar arguments applied to κ^H and W^H establish that functions $\kappa^H(v,\cdot)$, $-W^H(\cdot)$, $u_1^H(m_H;v,\cdot)$, $u_1^H(m_L;v,\cdot)$, $u_2^H(m_H;v,\cdot)$, $-u_2^H(m_L;v,\cdot)$, $-\left[u_2^H(m_L;v,\cdot) - u_2^H(m_H;v,\cdot)\right]$, $\left[u_1^H(m_L;v,\cdot) - u_1^H(m_H;v,\cdot)\right]$, $-v^H(m_L;v,\cdot)$, $v^H(m_H;v,\cdot)$ are all decreasing.

(b). We proved differentiability of $\kappa^L(v,\cdot)$ in part (a). The same arguments prove that $\kappa^H(v,\cdot)$ is differentiable. To show that $\frac{\partial}{\partial s}\kappa^j(v,s) = -b^j(s)C(av)$ for some $b^j(s) > 0$, let $\left(w_{v,s}^*, \Delta_{v,s}^*\right)$ be a solution to (58) for (v,s). Homogeneity of C implies that $\left(w_{v,s}^*, \Delta_{v,s}^*\right) = v \cdot \left(w_{1,s}^*, \Delta_{1,s}^*\right)$ if $\rho \in (0,1)$, $\left(w_{v,s}^*, \Delta_{v,s}^*\right) = \left(-\frac{1}{2}v + w_{0,s}^*, \Delta_{0,s}^*\right)$ if $\rho = 1$ and $\left(w_{v,s}^*, \Delta_{v,s}^*\right) = -v \cdot \left(w_{-1,s}^*, \Delta_{-1,s}^*\right)$ if $\rho > 1$. Then (62) and the functional form of C establishes that $\frac{\partial}{\partial s}\kappa^j(v,s) = -b^j(s)C(av)$. Since $\kappa^j(v,\cdot)$ is decreasing by part (a), $b^j(s) \geq 0$ for all s. We next show that $b^j(\cdot)$ is bounded away from zero and bounded above.

Fix any $v > \underline{v}$. We show that $\frac{\partial}{\partial s}\kappa^L(v,\cdot)$ is in a compact set bounded away from zero, which, given the previous result, is sufficient to establish the bounds on $b^L(s)$ stated in the proposition (the arguments for $\frac{\partial}{\partial s}\kappa^H(v,\cdot)$ are analogous). Equation (59) defines Δ as an implicit continuous function of w. Then (60) shows that $w_{v,s}^*$ lies in a compact set which can be chosen independently of s, and therefore $\Delta_{v,s}^*$ also lies in a compact set independent of s. Also observe that $\Delta = 0$ cannot satisfy (59) and (60). Therefore $\Delta_{v,s}^* > 0$ for all $s \in [0,1]$ and hence bounded away from zero. Then (62) establishes that $\frac{\partial}{\partial s}\kappa^L(v,\cdot)$ is in a compact set bounded away from 0.

The envelope theorem gives

$$\frac{\partial}{\partial s}W^L(s) = \pi_L\left\{\left[\theta_L u_s^w(m_L) - \lambda^w C(u_s^w(m_L))\right] - \left[\theta_L u_s^w(m_H) - \lambda^w C(u_s^w(m_H))\right]\right\}. \quad (63)$$

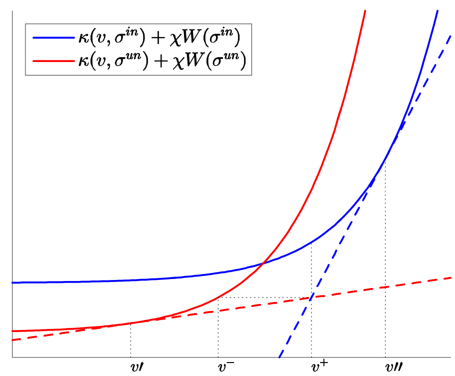

Figure 3: Bounds for the convex hull

Since $u_s^w(m)$ are in a compact set by Lemma 9, $\frac{\partial}{\partial s}W^L(\cdot)$ is bounded. We next show that it is bounded away from 0. We have $\mathbb{E}_s[\theta|m_H] \geq 1$ and $\mathbb{E}_s[\theta|m_L] = \theta_L$ for all $s > 0$ and therefore $u_s^w(m_H) \geq C'^{-1}\left(\frac{1}{\lambda^w}\right)$, $u_s^w(m_L) = C'^{-1}\left(\frac{\theta_L}{\lambda^w}\right)$ from (49) for $s > 0$. By Theorem 3 in Milgrom and Segal (2002), $u_0^w(m_L) = \lim_{s\to 0} u_s^w(m_L) = C'^{-1}\left(\frac{\theta_L}{\lambda^w}\right)$ and $u_0^w(m_H) = C'^{-1}\left(\frac{1}{\lambda^w}\right)$. Thus $u_s^w(m_H) - u_s^w(m_L)$ is bounded away from 0 for all s and hence the expression in the curly brackets in (63) is bounded away from 0 for all s. ■

Proof of Corollary 2. We first prove the second part of the corollary, showing in the process that the convex hull of $k(v)$ is well defined. Suppose that $\chi > 0$ (otherwise, the result is trivial). From Corollary 1 there are two thresholds, $v^- > -\infty$ and $v^+ > v^-$, such that $k(v) = \kappa(v, \sigma^{un}) + \chi W(\sigma^{un})$ for $v \leq v^-$ and $k(v) = \kappa(v, \sigma^{in}) + \chi W(\sigma^{in})$ for $v \geq v^+$ and $\kappa(\cdot, \sigma^{un})$ and $\kappa(\cdot, \sigma^{in})$ are strictly convex. Let v' and v'' be given by the unique solutions to (i) $k(v') - k(v^-) = k'(v')(v' - v^+)$ and $v' \leq v^-$; and (ii) $k(v'') - k(v^-) = k'(v'')(v'' - v^+)$ and $v'' \geq v^+$, respectively. Figure 3 illustrates how v' and v'' are constructed.

By construction the two dashed lines intersect at the point $(v^+, k(v^-))$ and are tangent to $k(v)$ at v' and v'', respectively. Note that the shape of $k(v)$ for $v \leq v^-$ and $v \geq v^+$ guarantees the existence of such v' and v''. Let V be the set of points above the two solid lines together with the points above the dashed lines when $v' \leq v \leq v''$. Formally, $V = \{(v, y) : y \geq k(v)$, for $v \leq v'$, $y \geq k(v^-) + k'(v')(v - v^+)$, for $v' < v \leq v^+$, $y \geq k(v^-) + k'(v'')(v - v^+)$, for $v^+ < v \leq v''$, $y \geq k(v)$, for $v > v''\}$. V is convex and, since $k(v^-) \leq k(v) \leq k(v^+)$, the set V contains the set $\{(v, y) : y \geq k(v)\}$. Since the convex hull is the intersection of all the convex sets containing $\{(v, y) : y \geq k(v)\}$, then the convex hull of k, k^{co}, must be a subset of V with $k^{co}(v) = k(v)$ for $v \leq v'$ and $v \geq v''$. Since k is strictly convex in those regions, so is $k^{co}(v)$

and no randomization is done for $v \le v'$ and $v \ge v''$.

We now show that any point on the Pareto frontier can be supported with strategies in which agents with higher v play more informative strategies. Let $(\mathbf{u}^*, \hat{\boldsymbol{\sigma}})$ be a best PBE. Take $v_i > v_j$ and suppose $v_i = \sum_{s=1}^{I} p_s'' \hat{v}_s$ and $v_j = \sum_{s=1}^{I} p_s' \hat{v}_s$, for some finite set of points $\hat{v}_1 < ... < \hat{v}_I$, $I > 1$,[16] with $v^- \le \hat{v}_s \le v^+$, for all s, where v^-, v^+ are defined in part (a). To simplify notation, let σ_s be the solution to (14) corresponding to $\hat{v}_s$. (The arguments extend with minor modifications if i and j play different strategies for some $\hat{v}_k$ with $p_k'', p_k' > 0$.) By Proposition 1, if $t > s$, then $\sigma_t \succeq \sigma_s$. By Lemma 21 in Supplementary material, we can find $\tilde{p}'', \tilde{p}' \in \Delta(\{\hat{v}_1, ..., \hat{v}_I\})$ with the following properties: (i) $\tilde{p}''$ FOSD $\tilde{p}'$, that is, $\sum_{s=1}^{k} \tilde{p}_s'' \le \sum_{s=1}^{k} \tilde{p}_s'$ for all $k \le I$; (ii) $\tilde{p}_s'' + \tilde{p}_s' = p_s'' + p_s'$ for all s; (iii) $\sum_{s=1}^{I} \tilde{p}_s'' \hat{v}_s = v_i$ and $\sum_{s=1}^{I} \tilde{p}_s' \hat{v}_s = v_j$. Let $\boldsymbol{\sigma}_i^*(\cdot|z, \cdot) = \sigma_k$, for $\sum_{s=1}^{k-1} \tilde{p}_{s-1}'' \le z < \sum_{s=1}^{k} \tilde{p}_s''$, for $k = 1, ..., I$, with $\sum_{s=1}^{-1} \tilde{p}_s'' = 0$, and define $\boldsymbol{\sigma}_j^*(\cdot|z, \cdot)$ analogously using $\tilde{p}'$. Property (i) implies $\boldsymbol{\sigma}_i^*(\cdot|z, \cdot) \succeq \boldsymbol{\sigma}_j^*(\cdot|z, \cdot)$ for all z. Property (ii) implies that $(\mathbf{u}^*, \boldsymbol{\sigma}^*)$ satisfies (2) and (19) and, thus, it is a best PBE. ■

Proof of Proposition 3. To simplify notation, let $K^j(v, s) = \kappa^j(v, s) + \chi W^j(s)$. We first derive sufficient conditions that ensure that the convex hull of k is obtained by randomizing between $K^L(v^-, 0)$ and $K^L(v^+, 1)$ for some v^-, v^+ and show that when $\rho = 1$ these sufficient conditions do not depend on multipliers $(\zeta_1, \zeta_2, \chi, \lambda^w)$. Then we verify that they hold for an open set of $\left(\{\theta_j, \pi(\theta_j)\}_j\right)$. By the arguments in the text when $\rho = 1$ we can write $\kappa^j(v, s) = d^j(s) \exp\left(\frac{v}{2}\right)$. We assume that $\chi > 0$, otherwise the result is immediate (in this case for any $\left(\{\theta_j, \pi(\theta_j)\}_j\right)$ the unique optimal reporting strategy is fully informative).

Sufficient conditions

We define a convex hull of the functions $d^L(0) \exp\left(\frac{v}{2}\right) + \chi W^L(0)$ and $d^L(1) \exp\left(\frac{v}{2}\right) + \chi W^L(1)$. Since $d^L(0) > d^L(1)$ and $W^L(0) < W^L(1)$, it is described by $\tilde{k}(v) = d^L(0) \exp\left(\frac{v}{2}\right) + \chi W^L(0)$ for $v \le v^-$, $\tilde{k}(v) = d^L(1) \exp\left(\frac{v}{2}\right) + \chi W^L(1)$ for $v \ge v^+$ and $\tilde{k}(v) = Av + B$ for $v \in [v^-, v^+]$ for some (v^-, v^+, A, B) that satisfy

$$\begin{aligned} d^L(0) \exp\left(\frac{v^-}{2}\right) + \chi W^L(0) &= Av^- + B, \\ d^L(1) \exp\left(\frac{v^+}{2}\right) + \chi W^L(1) &= Av^+ + B, \\ \frac{1}{2} d^L(0) \exp\left(\frac{v^-}{2}\right) &= A, \\ \frac{1}{2} d^L(1) \exp\left(\frac{v^+}{2}\right) &= A. \end{aligned}$$

[16] For simplicity, we assume that v_i, v_j are delivered with only a finite number of points. All the proofs extend immediately to a countable set of points by letting $I = \infty$.

We can solve this system for the four variables (v^-, v^+, A, B) as a function of $\left(d^L(0), d^L(1), W^L(0), W^L(1)\right)$:

$$\begin{aligned} v^- &= 2\ln\left(\frac{2A}{d^L(0)}\right), v^+ = 2\ln\left(\frac{2A}{d^L(1)}\right), \\ A &= \frac{1}{2}\chi\frac{W^L(1) - W^L(0)}{\ln(d^L(0)) - \ln(d^L(1))}, \\ B &= 2A - 2A\ln\left(\frac{2A}{d^L(0)}\right) + \chi W^L(0). \end{aligned} \tag{64}$$

Claim. If $K^j(v,s) \geq Av + B$ for all v then $K^j(v,s) \geq \tilde{k}(v)$ for all v.

Proof of the claim. By construction, $K^j(v,s) \geq \tilde{k}(v)$ for $v \in [v^-, v^+]$, we need to verify $K^j(v,s) \geq \tilde{k}(v)$ for $v > v^+$ and $v > v^+$. Suppose $K^j(v,s) < \tilde{k}(v)$ for some $v < v^-$. We have $d^j(s) \geq d^L(1)$ for all $s \in [0,1], j \in \{H, L\}$ because in problem (13) for $\sigma = \sigma^{in}$ only the incentive constraint for θ_L type binds (guess and verify or see Atkeson and Lucas (1992)). Therefore any function $K^j(v,s) = d^j(s)\exp(v/2) + \chi W^j(s)$ intersects $K^L(v,1)$ at most once from below. But then $K^j(v,s)$ must also intersect line $Av + B$, a contradiction. Analogous arguments apply for $v < v^-$. ■

Given this claim, we find sufficient conditions to ensure that $K^j(v,s) \geq Av + B$ for all v, j, s. Clearly, if we hold $d^j(s)$ fixed and change $W^j(s)$, this equation will be satisfied for high $W^j(s)$. Let's find a cut-off $\bar{W}_s^j$ so that this inequality holds. $\bar{W}_s^j$ should be such that $Av + B$ is also a lower envelope for $d^j(s)\exp(v/2) + \chi\bar{W}_s^j$ all v, j, s. Therefore, for any (j,s) there must exist $\left(\hat{v}_s^j, \bar{W}_s^j\right)$ such that

$$\begin{aligned} d^j(s)\exp\left(\frac{\hat{v}_s^j}{2}\right) + \chi\bar{W}_s^j &= A\hat{v}_s^j + B, \\ \frac{1}{2}d^j(s)\exp\left(\frac{\hat{v}_s^j}{2}\right) &= A. \end{aligned}$$

This gives $\chi\bar{W}_s^j = 2A\ln\left(\frac{2A}{d^j(s)}\right) - 2A + B$. Any $K^j(v,s) \geq \tilde{k}(v)$ if $\chi W^j(s) > \chi\bar{W}_s^j$ or, using (64),

$$\begin{aligned} \chi W^j(s) &\geq 2A\ln\left(\frac{2A}{d^j(s)}\right) - 2A + B \\ &= \chi\left[\left(W^L(1) - W^L(0)\right)\frac{\ln\left(d^L(0)\right) - \ln\left(d^j(s)\right)}{\ln(d^L(0)) - \ln(d^L(1))} + W^L(0)\right] \end{aligned}$$

or

$$W^j(s) \geq \left(W^L(1) - W^L(0)\right)\frac{\ln\left(d^L(0)\right) - \ln\left(d^j(s)\right)}{\ln(d^L(0)) - \ln(d^L(1))} + W^L(0). \tag{65}$$

This inequality does not depend on χ. Also $W^j(s)$ can be written only as a function of (j,s) since (49) implies that $\lambda^w = 1$ when $\rho = 1$.

We next show that $d^j(s)$ are independent of (ζ_1, ζ_2) and therefore whether equation (65) is satisfied depends only on s and $\left(\{\theta_j, \pi(\theta_j)\}_j\right)$. We consider the case with $j = L$, the other is analogous. Using the homogeneity properties of $C(\cdot)$, we can rewrite condition (59) as

$$\Delta = \frac{\theta_L}{1+\theta_L} \ln\left(\frac{\zeta_1}{\theta_L \zeta_2}\right) - 2\frac{\theta_L}{1+\theta_L} w.$$

Plugging this back into (58) gives

$$d^L(s) = \min_w (1 - s\pi_L)\left[\zeta_1 \exp(-w) + \zeta_2 \exp(w)\right] + s\pi_L \zeta_1^{\frac{\theta_L}{1+\theta_L}} \zeta_2^{\frac{1}{1+\theta_L}} \phi \exp\left(\frac{1-\theta_L}{1+\theta_L} w\right) \quad (66)$$

where $\phi \equiv \theta_L^{\frac{1}{1+\theta_L}} + \theta_L^{-\frac{\theta_L}{1+\theta_L}}$. Also, condition (60) implies

$$(1 - s\pi_L)\left[-\zeta_1 \exp(-w) + \zeta_2 \exp(w)\right] + s\pi_L \zeta_1^{\frac{\theta_L}{1+\theta_L}} \zeta_2^{\frac{1}{1+\theta_L}} \phi \frac{1-\theta_L}{1+\theta_L} \exp\left(\frac{1-\theta_L}{1+\theta_L} w\right) = 0$$

or, dividing by ζ_1,

$$(1 - s\pi_L)\left[-\exp(-w) + \frac{\zeta_2}{\zeta_1}\exp(w)\right] + s\pi_L \left(\frac{\zeta_2}{\zeta_1}\right)^{\frac{1}{1+\theta_L}} \phi \frac{1-\theta_L}{1+\theta_L} \exp\left(\frac{1-\theta_L}{1+\theta_L} w\right) = 0.$$

If we let $\hat{w} \equiv w + \frac{1}{2}\ln\frac{\zeta_2}{\zeta_1}$, the latter becomes

$$(1 - s\pi_L)\left[-\exp(-\hat{w}) + \exp(\hat{w})\right] + s\pi_L \phi \frac{1-\theta_L}{1+\theta_L} \exp\left(\frac{1-\theta_L}{1+\theta_L} \hat{w}\right) = 0. \quad (67)$$

and, thus, the optimal $\hat{w}^*$, which is the solution to (67), is independent of (ζ_1, ζ_2). Plugging this back into (66) gives

$$d^L(s) = \zeta_1^{\frac{1}{2}} \zeta_2^{\frac{1}{2}} \bar{d}^L(s)$$

where $\bar{d}^L(s) \equiv (1 - s\pi_L)\left[\exp(-\hat{w}^*) + \exp(\hat{w}^*)\right] + s\pi_L \phi \exp\left(\frac{1-\theta_L}{1+\theta_L}\hat{w}^*\right)$ is independent of (ζ_1, ζ_2). Therefore, condition (65) can be restated as

$$\underbrace{W^j(s) - W^L(0) - \left(W^L(1) - W^L(0)\right)\frac{\ln\left(\bar{d}^L(0)\right) - \ln\left(\bar{d}^j(s)\right)}{\ln\left(\bar{d}^L(0)\right) - \ln\left(\bar{d}^L(1)\right)}}_{\equiv r^j(s)} \geq 0 \text{ for all } s, j. \quad (68)$$

A sufficient condition for any interior reporting strategy to be suboptimal is $r^j(s) > 0$ for all s, j. This condition depends only on $\left(\{\theta_j, \pi(\theta_j)\}_j\right)$.

Verifying (68)

We have $0 = r^L(0) = r^L(1) < r^H(1)$. To establish (68) it is sufficient to verify that $r^j(\cdot)$ is either increasing or hump-shaped. Figure 4 plots derivatives of $r^j(\cdot)$ for $(\theta_L, \pi_L) = (0.4, 0.5)$. They are strictly positive for r^H; and strictly positive at $s = 0$ and change sign only once for r^L, which ensures that r^H is increasing while r^L is hump-shaped. By the theorem of the maximum the solution to (66) is continuous in (θ_L, π_L) which ensures that there exists an open set of parameters around $(\theta_L, \pi_L) = (0.4, 0.5)$ for which (68) is satisfied. ■

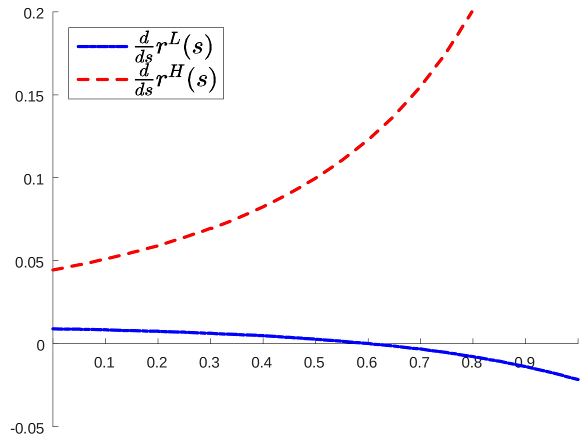

Figure 4: The derivatives of $r^j(\cdot)$, $j \in \{L, H\}$.

6.2 Proofs of Section 3.1

Proof of Lemma 5. For any $\breve{h}^t \in \breve{H}^t$, define strategy $\boldsymbol{\sigma}'$ by

$$\boldsymbol{\sigma}'_t\left(\cdot\left|\breve{h}^t, \left(\hat{\theta}^{t-1}, \theta_t\right)\right.\right) = \sum_{\theta^{t-1}} \boldsymbol{\sigma}_t\left(\cdot\left|\breve{h}^t, \left(\theta^{t-1}, \theta_t\right)\right.\right) \pi_{t-1}\left(\theta^{t-1}\right),$$

for all $\hat{\theta}^{t-1}$. By construction, $\boldsymbol{\sigma}'_t\left(\cdot\left|\breve{h}^t, \left(\hat{\theta}^{t-1}, \theta_t\right)\right.\right) = \boldsymbol{\sigma}'_t\left(\cdot\left|\breve{h}^t, \left(\tilde{\theta}^{t-1}, \theta_t\right)\right.\right)$ for all $\tilde{\theta}^{t-1}, \hat{\theta}^{t-1}$. Since any agent with a history $\left(\breve{h}^t, \left(\tilde{\theta}^{t-1}, \theta_t\right)\right)$ can replicate the strategy of the agent with a history $\left(\breve{h}^t, \left(\hat{\theta}^{t-1}, \theta_t\right)\right)$ and achieve the same payoff as that agent, and $\boldsymbol{\sigma}_t\left(\cdot\left|\breve{h}^t, \left(\hat{\theta}^{t-1}, \theta_t\right)\right.\right)$ is the optimal choice of the agent with history $\left(\breve{h}^t, \left(\hat{\theta}^{t-1}, \theta_t\right)\right)$, the new strategy $\boldsymbol{\sigma}'$ satisfies the agents' best response constraint (23). The strategy $\boldsymbol{\sigma}'$ induces distributions $\boldsymbol{\mu}'$ which satisfy $\mu'_t = \mu_t$ for all aggregate histories, hence, the feasibility constraint (20) is still satisfied if agents play $\boldsymbol{\sigma}'$. Finally, after any history $h^t \in H^t$, the posterior beliefs are the same, $\mathbb{E}_{\boldsymbol{\sigma}'}\left[\theta_t | h^t\right] = \mathbb{E}_{\boldsymbol{\sigma}}\left[\theta_t | h^t\right]$.

For simplicity we assume that $\mu_t\left(h'^t\right), \mu_t\left(h''^t\right) > 0$. Let $\alpha = \mu_t\left(h'^t\right) / \left(\mu_t\left(h'^t\right) + \mu_t\left(h''^t\right)\right)$ and define $\phi' : [0, \alpha] \to [0, 1]$ by $\phi'(z) = z/\alpha$ and $\phi'' : (\alpha, 1] \to [0, 1]$ by $\phi''(z) = (z - \alpha)/(1 - \alpha)$. Define a new strategy and allocations $(\boldsymbol{\sigma}', \mathbf{u}')$ for all $T \geq 1$, $h^t \in \left\{h'^t, h''^t\right\}$, θ^{t+T} as

$$\begin{aligned}
\mathbf{u}'_{t+T}\left(h^t, z_{t+1}, m_{t+1}, ..., m_{t+T}\right) &= \mathbf{u}^*_{t+T}\left(h'^t, \phi'\left(z_{t+1}\right), m_{t+1}, ..., m_{t+T}\right), \\
\boldsymbol{\sigma}'_{t+T}\left(\cdot | h^t, z_{t+1}, m_{t+1}, ..., z_{t+T}; \theta^{t+T}\right) &= \boldsymbol{\sigma}^*_{t+T}\left(\cdot | h''^t, \phi'\left(z_{t+1}\right), m_{t+1}, ..., z_{t+T}; \theta^{t+T}\right)
\end{aligned}$$

if $z_{t+1} \leq \alpha$ and

$$\begin{aligned}
\mathbf{u}'_{t+T}\left(h^t, z_{t+1}, m_{t+1}, ..., m_{t+T}\right) &= \mathbf{u}^*_{t+T}\left(h''^t, \phi''\left(z_{t+1}\right), m_{t+1}, ..., m_{t+T}\right), \\
\boldsymbol{\sigma}'_{t+T}\left(\cdot|h^t, z_{t+1}, m_{t+1}, ..., z_{t+T}; \theta^{t+T}\right) &= \boldsymbol{\sigma}^*_{t+T}\left(\cdot|h'''^t, \phi''\left(z_{t+1}\right), m_{t+1}, ..., z_{t+T}; \theta^{t+T}\right)
\end{aligned}$$

if $z_{t+1} > \alpha$ and $\mathbf{u}'_s = \mathbf{u}_s, \boldsymbol{\sigma}'_s = \boldsymbol{\sigma}_s$ for all other histories and periods s. Agents with histories h'^t, h''^t could have replicated each other strategies after period t, so they must be indifferent between them. The strategy $\boldsymbol{\sigma}'$ gives them the same utility for all histories following $\left\{h'^t, h''^t\right\}$ leaving all other histories unchanged, therefore it is incentive compatible, i.e. satisfies (23). The strategy profile $\boldsymbol{\sigma}'$ induces $\boldsymbol{\mu}'$, which assigns the same probability to any realization of $\mathbf{u}$ as $\boldsymbol{\mu}$, therefore, the feasibility constraint (20) is satisfied. Finally, $\mathbb{E}_{\boldsymbol{\sigma}}\left[\theta_t|h^t\right] = \mathbb{E}_{\boldsymbol{\sigma}'}\left[\theta_t|h^t\right]$ for all $h^t \in H^t$, hence, (22) is satisfied. Therefore $(\mathbf{u}', \boldsymbol{\sigma}')$ is a PBE which is payoff equivalent to $(\mathbf{u}, \boldsymbol{\sigma})$. ■

Proof of Lemma 7. *Properties of* W_t. The arguments in the proof of Lemma 2 extend immediately to W_t.

Properties of k_t. To prove concavity let $(\mathbf{u}', \boldsymbol{\sigma}')$ and $(\mathbf{u}'', \boldsymbol{\sigma}'')$ be solutions to (29) for some $v' < v''$. For any $v \in (v', v'')$ choose α such that $v = \alpha v' + (1-\alpha) v''$. Let $(\hat{\mathbf{u}}, \hat{\boldsymbol{\sigma}})$ be such that $\hat{\mathbf{u}}_t(m,z) = \mathbf{u}'_t(m, z/\alpha)$, $\hat{\boldsymbol{\sigma}}_t(m,z) = \boldsymbol{\sigma}_t(m, z/\alpha)$, if $z \leq \alpha$, and $\hat{\mathbf{u}}_t(m,z) = \mathbf{u}'_t(m, (z-\alpha)/(1-\alpha))$, $\hat{\boldsymbol{\sigma}}_t(m,z) = \boldsymbol{\sigma}_t(m,(z-\alpha)/(1-\alpha))$, if $z > \alpha$. The pair $(\hat{\mathbf{u}}, \hat{\boldsymbol{\sigma}})$ satisfies (23) and (25) for v. Therefore, $k_t(v) \geq \alpha k_t(v') + (1-\alpha) k_t(v'')$.

Concavity of k_t implies continuity on $(\underline{v}, \bar{v})$. To show that the continuity extends to $\underline{v}$ suppose without loss of generality that $\underline{v} = 0$. Define

$$k^*_t(v) = \max_{\mathbf{u}, \boldsymbol{\sigma}} \mathbb{E}_{\boldsymbol{\sigma}}\left[\sum_{s=0}^{\infty} \frac{\bar{\beta}_{t+s}}{\bar{\beta}_t}\left(\theta_s \mathbf{u}_s - \zeta_{t+s} C\left(\mathbf{u}_s\right)\right)\middle| v\right] - \sum_{s=0}^{\infty} \frac{\bar{\beta}_{t+s}}{\bar{\beta}_t} \chi_{t+s} W_{t+s}\left(\sigma^{un}\right)$$

subject to (25). $k^*_t(\cdot)$ is a continuous function with $k^*_t(v) \geq k_t(v)$. At $v = \underline{v}$ its solution sets $\mathbf{u}_s(h^s) = U(0)$ for all h^s. This allocation together with an uninformative reporting strategy satisfies (23) and therefore $k^*_t(\underline{v}) = k_t(\underline{v})$. This establishes continuity of k_t at $\underline{v}$.

To show differentiability first consider unbounded utility functions. Fix an interior v_0, let $(\mathbf{u}_{v_0}, \boldsymbol{\sigma}_{v_0})$ be the solution to $k_t(v_0)$ and consider the alternative pair $\hat{\mathbf{u}}$ such that $\hat{\mathbf{u}}_t = \mathbf{u}_{v_0,t} + v - v_0$, $\hat{\mathbf{u}}_{t+s} = \mathbf{u}_{v_0,t+s}$, for all $s > 0$. The pair $(\hat{\mathbf{u}}, \boldsymbol{\sigma}_{v_0})$ satisfies (23), delivers v, and has value

$$\begin{aligned}
V_t(v) &= \mathbb{E}_{\boldsymbol{\sigma}_{v_0}}\left[\left(\theta_t\left(\mathbf{u}_{v_0,t} + v - v_0\right) - \zeta_t C\left(\mathbf{u}_{v_0,t} + v - v_0\right) - \chi_t W_t\right)\middle| v\right] \\
&\quad + \frac{1}{\bar{\beta}_t} \mathbb{E}_{\boldsymbol{\sigma}_{v_0}}\left[\sum_{s=1}^{\infty} \bar{\beta}_{t+s}\left(\theta_{t+s}\mathbf{u}_{v_0,t+s} - \zeta_{t+s} C\left(\mathbf{u}_{v_0,t+s}\right) - \chi_{t+s} W_{t+s}\right)\middle| v\right].
\end{aligned}$$

Clearly, $k_t(v) \geq V_t(v)$ with equality at v_0. Also, $V_t(v)$ is concave and continuously differentiable.[17] Thus, from Benveniste-Scheinkman theorem we have that $k_t(v)$ is differentiable and

$$k_t'(v_0) = V_t'(v_0) = 1 - \zeta_t \mathbb{E}_{\boldsymbol{\sigma}_{v_0}}\left[C'(\mathbf{u}_{v_0,t})\right]. \tag{69}$$

Note that if k_t is twice differentiable, it also implies that

$$0 \geq k_t''(v_0) \geq -\zeta_t \mathbb{E}_{\boldsymbol{\sigma}_{v_0}}\left[C''(\mathbf{u}_{v_0,t})\right]. \tag{70}$$

If utility is bounded below (without loss of generality by 0) but not above, we can follow analogous steps as above using the pair $(\hat{\mathbf{u}}, \boldsymbol{\sigma}_{v_0})$ such that $\hat{\mathbf{u}}_{t+s} = \frac{v}{v_0}\mathbf{u}_{v_0,t+s}$, for all $s > 0$, and $\hat{\boldsymbol{\sigma}} = \boldsymbol{\sigma}_{v_0}$. A symmetric argument works for the case where utility is bounded above but not below. Finally, when utility is bounded we can construct a function V_t separately for $v \leq v_0$ and $v > v_0$.

We next establish the value of derivatives $k_t'(v)$ in the limits. Define a function

$$\bar{K}_t(v) = \max_{\mathbf{u},\boldsymbol{\sigma}} \mathbb{E}_{\boldsymbol{\sigma}} \sum_{t=0}^{\infty} \beta^t \left[\theta_t \mathbf{u}_t - \zeta_t C(\mathbf{u}_t)\right]$$

subject to (25). We first show that $k_t(v) \leq \bar{K}_t(v) + const$. Let $\upsilon_t = \max_{\theta\in\Theta, c\geq 0}\left[\theta U(c) - \zeta_t c\right]$ and let $(\mathbf{u}_v, \boldsymbol{\sigma}_v)$ be a solution to (29). Then

$$\begin{aligned}
& k_t(v) - \sum_{s=0}^{\infty} \frac{\bar{\beta}_{t+s}}{\bar{\beta}_t}\upsilon_{t+s} = \mathbb{E}_{\boldsymbol{\sigma}_v} \sum_{s=0}^{\infty} \frac{\bar{\beta}_{t+s}}{\bar{\beta}_t}\left[\theta_{t+s}\mathbf{u}_{v,t+s} - \zeta_{t+s}C(\mathbf{u}_{v,t+s}) - \upsilon_{t+s}\right] - \mathbb{E}_{\boldsymbol{\sigma}_v} \sum_{s=0}^{\infty} \frac{\bar{\beta}_{t+s}}{\bar{\beta}_t}\chi_{t+s}W_{t+s} \\
\leq\ & \mathbb{E}_{\boldsymbol{\sigma}_v} \sum_{s=0}^{\infty} \beta^s \left[\theta_{t+s}\mathbf{u}_{v,t+s} - \zeta_{t+s}C(\mathbf{u}_{v,t+s}) - \upsilon_{t+s}\right] - \sum_{s=0}^{\infty} \frac{\bar{\beta}_{t+s}}{\bar{\beta}_t}\chi_{t+s}W_{t+s}(\sigma^{un}) \\
\leq\ & \bar{K}_t(v) - \sum_{s=0}^{\infty} \beta^s \upsilon_{t+s} - \sum_{s=0}^{\infty} \frac{\bar{\beta}_{t+s}}{\bar{\beta}_t}\chi_{t+s}W_{t+s}(\sigma^{un}),
\end{aligned}$$

where the first inequality follows from the fact that the expression in square brackets is negative and $\bar{\beta}_s/\bar{\beta}_t \geq \beta^{s-t}$ and the second inequality follows from the fact that $\bar{K}_t(v)$ maximizes $\mathbb{E}_{\boldsymbol{\sigma}} \sum_{s=0}^{\infty} \beta^s \left[\theta_{t+s}\mathbf{u}_{t+s} - \zeta_{t+s}C(\mathbf{u}_{t+s})\right]$ without incentive constraints. Since $k_t(v) \leq \bar{K}_t(v) + const$ and $\bar{K}_t(v)$ is concave, $\lim_{v\to\bar{v}} k_t'(v) \leq \lim_{v\to\bar{v}} \bar{K}_t'(v) = -\infty$ and, if utility is unbounded below, $\lim_{v\to\underline{v}} k_t'(v) \geq \lim_{v\to\underline{v}} \bar{K}_t'(v) = 1$. Since $k_t'(v) < 1$ if utility is unbounded below from (69), we have $\lim_{v\to\underline{v}} k_t'(v) = 1$.

[17] The latter comes from Leibniz's theorem since $f(u,v) = \theta_t(\mathbf{u}_{v_0,t} + v - v_0) - \zeta_t C(\mathbf{u}_{v_0,t} + v - v_0)$ is a Carathéodory function (continuous in v and measurable in u) which is locally uniformly integrably bounded because, for each v, there is a neighborhood U_v and a positive number B such that $|f(u,\hat{v})| \leq B$, for all $\hat{v} \in U_v$. Finally, f_v' is continuous and also locally uniformly integrably bounded.

If utility is bounded below, constraint $u \geq 0$ may bind. Let

$$\underline{K}_t(v) = \frac{1}{\bar{\beta}_t} \max_{\mathbf{u}} \sum_{s=0}^{\infty} \bar{\beta}_{t+s} \left[\mathbf{u}_{t+s} - \zeta_{t+s} C(\mathbf{u}_{t+s}) \right] - \sum_{s=0}^{\infty} \frac{\bar{\beta}_{t+s}}{\bar{\beta}_t} \chi_{t+s} W_{t+s}(\sigma^{un})$$

subject to

$$\sum_{s=0}^{\infty} \beta^s \mathbf{u}_{t+s} = v. \tag{71}$$

$k_t(v) \geq \underline{K}_t(v)$ for all v with $k_t(\underline{v}) = \underline{K}_t(\underline{v})$. Since $\underline{K}'_t(\underline{v}) \geq 1$, we have $k'_t(\underline{v}) \geq 1$.

It remains show that $\limsup \chi_t > 0$ implies $\lim_{v \to \underline{v}} \underline{K}'_t(v) = \infty$ and therefore $\lim_{v \to \underline{v}} \underline{k}'_t(v) = \infty$. Let $\underline{\gamma}_t(v)$ be the Lagrange multiplier on (71). The first order condition for $\mathbf{u}_{t+s}$, $s \geq 0$, is

$$1 - \zeta_{t+s} C'(\mathbf{u}_{v,t+s}) \leq \frac{\beta^s}{\bar{\beta}_{t+s}/\bar{\beta}_t} \underline{\gamma}_t(v). \tag{72}$$

Suppose that $\limsup \chi_t > 0$ but $\underline{\gamma}_t(v) = \underline{K}'_t(\underline{v}) < \infty$. If $\limsup \chi_t > 0$, then $\frac{\beta^s}{\bar{\beta}_{t+s}/\bar{\beta}_t} \to 0$ and there is some $T > t$ such that $1 - \frac{\beta^T}{\bar{\beta}_T/\bar{\beta}_t} \underline{\gamma}_t(v) > 0$. For such T the optimality condition (72) is satisfied only for $\mathbf{u}_{v,T} > 0$. This is impossible since $\lim_{v \to \underline{v}} \mathbf{u}_{v,t} = 0$ for all t. ■

Proof of Lemma 8. Let $\mathbf{X}(\boldsymbol{\sigma})$ be a set of $(\mathbf{u}, \mathbf{w})$ that satisfy (32) and $X(\sigma)$ be a set of (u, w) that satisfy (34) and (35). Observe that $(\mathbf{u}, \mathbf{w}) \in \mathbf{X}(\boldsymbol{\sigma})$ if and only if $(\mathbf{u}(\cdot, z), \mathbf{w}(\cdot, z)) \in X(\boldsymbol{\sigma}(\cdot|z, \cdot))$ for all z. From Luenberger (1969) (page 236, problem 7) we can form a Lagrangian

$$\begin{aligned} k_t(v) &= \max_{\substack{\mathbf{u},\mathbf{w},\boldsymbol{\sigma} \\ (\mathbf{u},\mathbf{w}) \in \mathbf{X}(\boldsymbol{\sigma})}} \int_Z \Bigg[\sum_{\theta,m} \pi(\theta) \boldsymbol{\sigma}(m|z,\theta) \left[(1 - \gamma_t(v)) \theta \mathbf{u}(m,z) - \zeta_t C(\mathbf{u}(m,z)) \right. \\ &\quad \left. + \hat{\beta}_{t+1} k_{t+1}(\mathbf{w}(m,z)) - \beta \gamma_t(v) \mathbf{w}(m,z) \right] - \chi_t W(\boldsymbol{\sigma}(\cdot|z,\cdot)) + \gamma_t(v) v \Bigg] dz \\ &= \max_{\boldsymbol{\sigma}} \int_Z \max_{\substack{(\mathbf{u}(\cdot,z),\mathbf{w}(\cdot,z)) \\ (\mathbf{u}(\cdot,z),\mathbf{w}(\cdot,z)) \in X(\boldsymbol{\sigma}(\cdot|z,\cdot))}} \Bigg[\sum_{\theta,m} \pi(\theta) \boldsymbol{\sigma}(m|z,\theta) \left[(1 - \gamma_t(v)) \theta \mathbf{u}(m,z) - \zeta_t C(\mathbf{u}(m,z)) \right. \\ &\quad \left. + \hat{\beta}_{t+1} k_{t+1}(\mathbf{w}(m,z)) - \beta \gamma_t(v) \mathbf{w}(m,z) \right] - \chi_t W(\boldsymbol{\sigma}(\cdot|z,\cdot)) + \gamma_t(v) v \Bigg] dz. \end{aligned}$$

Benveniste and Scheinkman arguments applied to the first maximum establish that $k'_t(v) = \gamma_t(v)$. Therefore $k_t(v) = \max_{\sigma} \{\kappa_t(v, \sigma) - \chi_t W_t(\sigma)\}$. ■

Proof of Lemma 9. The arguments in the text establish this lemma for $|M| < |M_\Theta|$, here we extend them to $|M| > |M_\Theta|$. The proof follows similar steps to those in the proof of Lemma 11. First, we argue that the incentive constraints (34) and (35) imply that we can partition any message space M into $2|\Theta|$ subsets: $|\Theta|$ subsets M_j of messages that are reported with positive probability by type θ_j and give the highest utility only to type θ_j; $|\Theta| - 1$ subsets $M_{j,j+1}$ of messages that are reported with positive probability by either θ_j or θ_{j+1} and give the highest utility to both θ_j and θ_{j+1}; and a subset $M_\varnothing$ of messages that are not sent with positive probability by any type (we omit subscript t to simplify the notation). To see that these subsets are enough to partition M, suppose m is sent with positive probability by θ_i and gives the highest utility to some other type θ_j, $j > i+1$. For any m' that gives the highest utility to θ_{i+1} we have

$$\theta_{i+1} u\left(m'\right) + \beta w\left(m'\right) \geq \theta_{i+1} u\left(m\right) + \beta w\left(m\right)$$

and

$$\begin{aligned} \theta_i u\left(m\right) + \beta w\left(m\right) &\geq \theta_i u\left(m'\right) + \beta w\left(m'\right), \\ \theta_j u\left(m\right) + \beta w\left(m\right) &\geq \theta_j u\left(m'\right) + \beta w\left(m'\right). \end{aligned}$$

The sum of the first the second inequalities implies $u\left(m'\right) \geq u\left(m\right)$, the sum of the first and the third inequalities implies $u\left(m'\right) \leq u\left(m\right)$, therefore, $u\left(m'\right) = u\left(m\right)$ and m gives the highest utility also to θ_{i+1}, thus, $m \in M_{i,i+1}$.

The same arguments as in the proof of Lemma 11 imply that it is without loss of generality to choose a message space $\bar{M}$ with $2\left|\Theta\right| - 1$ messages: one message for each subset M_j, $j = 1, ..., \left|\Theta\right|$; one message for each subset $M_{j,j+1}$, $j = 1, ..., \left|\Theta\right| - 1$; and no messages in $M_\varnothing$. To further restrict the message space note that, if a message is played with zero probability, we can always remove it from the message set. If instead all messages in $\bar{M}$ are played with positive probability, we can define an alternative strategy $\sigma_{[s]}$ such that $\sigma_{[s]}\left(m|\theta\right) = \left(1-s\right)\sigma_v\left(m|\theta\right)$, $\theta \in \Theta$, $m \in M_{i,i+1}$, $i = 1, ..., \left|\Theta\right| - 1$, $\sigma_{[s]}\left(m_i|\theta_i\right) = 1 - \sum_{m \neq m_i} \sigma_{[s]}\left(m|\theta_i\right)$, $m_i \in M_i$, $i = 1, ..., \left|\Theta\right|$, and adapt the arguments in the proof of Lemma 11 to restrict the message space to $2\left|\Theta\right| - 2$ messages. ■

6.3 Proofs of Sections 3.2 and 3.3

We first introduce some notation. For given v and σ, let $M_{v,\sigma}\left(\theta\right) \subset M_\Theta$ be the set of all messages that give type θ the highest utility. This set is uniquely defined only up to the set of messages that are sent with positive probability, so $M_{v,\sigma}\left(\theta\right)$ refers to any of such sets. Let $M_{v,\sigma} \equiv \cup_\theta M_{v,\sigma}\left(\theta\right)$. Also, let Σ^+ denote the set of strategies such that there are non-constant $\left\{u\left(m\right), w\left(m\right)\right\}_m$ that satisfy constraints (34) and (35).

Observe that if $m \in M_{v,\sigma}\left(\theta\right)$ and $m' \in M_{v,\sigma}\left(\theta'\right)$, with $\theta > \theta'$, then combining $\theta u\left(m\right) + \beta w\left(m\right) \geq \theta u\left(m'\right) + \beta w\left(m'\right)$ with $\theta' u\left(m'\right) + \beta w\left(m'\right) \geq \theta' u\left(m\right) + \beta w\left(m\right)$ gives $u\left(m\right) \geq u\left(m'\right)$ and $w\left(m\right) \leq w\left(m'\right)$. Thus, if we denote by m_1 any message in $M_{v,\sigma}$ such that $u\left(m_1\right) \leq u\left(m\right)$

for all $m \in M_{v,\sigma}$, we can always order messages in $M_{v,\sigma}$ as

$$u(m_1) \leq ... \leq u\left(m_{|M_{v,\sigma}|}\right), \quad w(m_1) \geq ... \geq w\left(m_{|M_{v,\sigma}|}\right). \tag{73}$$

Lemma 13 *If $\gamma_t(v) > 1$, then $u_{v,\sigma}(m) = 0$ and $w_{v,\sigma}(m) = \hat{w}$ for all m sent with positive probability where either $\hat{w} = \underline{v}$ or $\hat{w}$ satisfies $\frac{\hat{\beta}_{t+1}}{\beta} k'_{t+1}(\hat{w}) = \gamma_t(v)$. If $\gamma_t(v) \leq 1$ then*

$$\frac{\hat{\beta}_{t+1}}{\beta} \mathbb{E}_\sigma\left[k'_{t+1}(w_{v,\sigma})\right] \leq \gamma_t(v) = 1 - \zeta_t \mathbb{E}_\sigma\left[C'(u_{v,\sigma})\right] \tag{74}$$

with equality if $w_{v,\sigma}(m)$ is interior for all m sent with positive probability, and

$$(1 - \gamma_t(v))\,\theta_1 \leq \zeta_t C'(u_{v,\sigma}(m)) \leq (1 - \gamma_t(v))\,\theta_{|\Theta|}, \tag{75}$$

$$\underline{\varrho}\left[1 - \gamma_t(v)\right] + \left(1 - \frac{\beta}{\hat{\beta}_{t+1}}\right) \leq 1 - k'_{t+1}(w_{v,\sigma}(m)) \leq \bar{\varrho}\left[1 - \gamma_t(v)\right] + \left(1 - \frac{\beta}{\hat{\beta}_{t+1}}\right) \tag{76}$$

for $\bar{\varrho} = \frac{\beta}{\hat{\beta}_{t+1}} \frac{1+\theta_{|\Theta|}-\theta_1}{\theta_1}$ and $\underline{\varrho} = \frac{\beta}{\hat{\beta}_{t+1}} \frac{1+\theta_1-\theta_{|\Theta|}}{\theta_1}$.

Proof. We show this lemma assuming all messages in M_Θ are sent with positive probability, thus, $\sigma(m_1|\theta_1) > 0$, $\sigma\left(m_{2|\Theta|-2}|\theta_{|\Theta|}\right) > 0$, and $|M_{v,\sigma}| = 2|\Theta| - 2$. The other cases are analogous by restricting attention to the subset of M_Θ which is reported with positive probability.

Let $\xi'(\theta, m_\theta, m')$ and $\xi''(\theta, m_\theta, m')$ be the Lagrange multipliers on the constraints (34) and (35), respectively and $\nu^w(m), \nu^u(m)$ be the multipliers on $w(m) \geq \underline{v}$, $u(m) \geq \underline{u}$. We set $\xi'(\theta, m, m') = \xi''(\theta, m, m') = 0$ for all $m \notin M_{v,\sigma}(\theta)$ and $\xi'(\theta, m, m) = \xi''(\theta, m, m) = 0$ for all $m \in M_\Theta$, so that ξ', ξ'' are well defined for all (θ, m, m'). The first order conditions for the optimal choice of $w(m)$ and $u(m)$ in (33) are

$$\begin{aligned}&\sum_{\theta\in\Theta}\left[\frac{\hat{\beta}_{t+1}}{\beta} k'_{t+1}(w_{v,\sigma}(m)) - \gamma_t(v)\right]\sigma(m|\theta)\,\pi(\theta) + \sum_{(\theta,m')\in\Theta\times M_\Theta}\left[\xi'\left(\theta, m, m'\right) + \sigma\left(m'|\theta\right)\xi''\left(\theta, m, m'\right)\right] \\ &- \sum_{(\theta,m')\in\Theta\times M_\Theta}\left[\xi'\left(\theta, m', m\right) + \sigma(m|\theta)\,\xi''\left(\theta, m', m\right)\right] + \nu^w(m) = 0\end{aligned} \tag{77}$$

and

$$\begin{aligned}&\sum_{\theta\in\Theta}\left[(1 - \gamma_t(v))\,\theta - \zeta_t C'(u_{v,\sigma}(m))\right]\sigma(m|\theta)\,\pi(\theta) + \sum_{(\theta,m')\in\Theta\times M_\Theta}\left[\xi'\left(\theta, m, m'\right) + \sigma\left(m'|\theta\right)\xi''\left(\theta, m, m'\right)\right]\theta \\ &- \sum_{(\theta,m')\in\Theta\times M_\Theta}\left[\xi'\left(\theta, m', m\right) + \sigma(m|\theta)\,\xi''\left(\theta, m', m\right)\right]\theta + \nu^u(m) = 0.\end{aligned} \tag{78}$$

Sum (77) and (78) over all m to get

$$\frac{\hat{\beta}_{t+1}}{\beta} \mathbb{E}_\sigma k'_{t+1}(w_{v,\sigma}) + \sum_{m\in M_\Theta} \nu^w(m) = \gamma_t(v) = \mathbb{E}_\sigma\left[1 - \zeta_t C'(u_{v,\sigma})\right] + \sum_{m\in M_\Theta} \nu^u(m). \tag{79}$$

Suppose $(u_{v,\sigma}(m), w_{v,\sigma}(m))$ are the same for all m. Then (79) verifies that $(u_{v,\sigma}(m), w_{v,\sigma}(m))$ satisfies the conditions of the lemma. Hence for the rest of the lemma we assume that not all $(u_{v,\sigma}(m), w_{v,\sigma}(m))$ are the same, in which case it must be true that $\nu^u\left(m_{2|\Theta|-2}\right) = 0$, $\nu^w(m_1) = 0$. Let $G' \subset M_\Theta$ be a set of messages m for which $(u_{v,\sigma}(m), w_{v,\sigma}(m)) = (u_{v,\sigma}(m_1), w_{v,\sigma}(m_1))$ and θ' be the largest θ such that $G' \subset M_{v,\sigma}(\theta)$. Incentive compatibility implies that it is strictly suboptimal for any $\theta < \theta'$ to send any message other than those in G'. Therefore (77) and (78) can be written as

$$\begin{aligned}
\frac{\hat{\beta}_{t+1}}{\beta}k'_{t+1}(w_{v,\sigma}(m_1)) - \gamma_t(v) + \frac{\tilde{\vartheta}(G')}{\Pr(G')} &= 0, &(80)\\
(1-\gamma_t(v))\mathbb{E}_\sigma\left[\theta|G'\right] - \zeta_t C'(u_{v,\sigma}(m_1)) + \frac{\tilde{\vartheta}(G')}{\Pr(G')}\theta' + \frac{\nu^u(G')}{\Pr(G')} &= 0, &(81)
\end{aligned}$$

where $\Pr(G') = \sum_{\theta\in\Theta, m\in G'}\sigma(m|\theta)\pi(\theta)$, $\tilde{\vartheta}(G') = \sum_{m\in G', m'\in M_\Theta}\left[\begin{array}{c}\xi'(\theta', m, m') + \sigma(m'|\theta')\xi''(\theta', m, m')\\ -\xi'(\theta', m', m) - \sigma(m|\theta')\xi''(\theta', m', m)\end{array}\right]$, and $\nu^u(G') = \sum_{m\in G'}\nu^u(m)$. Similarly defining G'' and θ'' for $\left(u_{v,\sigma}\left(m_{2|\Theta|-2}\right), w_{v,\sigma}\left(m_{2|\Theta|-2}\right)\right)$ we get

$$\begin{aligned}
\frac{\hat{\beta}_{t+1}}{\beta}k'_{t+1}\left(w_{v,\sigma}\left(m_{2|\Theta|-2}\right)\right) - \gamma_t(v) + \frac{\tilde{\vartheta}(G'')}{\Pr(G'')} + \frac{\nu^w(G'')}{\Pr(G'')} &= 0, &(82)\\
(1-\gamma_t(v))\mathbb{E}_\sigma\left[\theta|G''\right] - \zeta_t C'\left(u_{v,\sigma}\left(m_{2|\Theta|-2}\right)\right) + \frac{\tilde{\vartheta}(G'')}{\Pr(G'')}\theta'' &= 0. &(83)
\end{aligned}$$

As a preliminary step we establish the signs of $\tilde{\vartheta}(G')$ and $\tilde{\vartheta}(G'')$. Since k_{t+1} is concave and $w_{v,\sigma}(m_1)$ is the largest $w_{v,\sigma}(m)$

$$\frac{\hat{\beta}_{t+1}}{\beta}k'_{t+1}(w_{v,\sigma}(m_1)) \leq \frac{\hat{\beta}_{t+1}}{\beta}\mathbb{E}_\sigma k'_{t+1}(w_{v,\sigma}) \leq \gamma_t(v),$$

where the second inequality follows from (79). Therefore (80) implies that $\tilde{\vartheta}(G') \geq 0$. To establish that $\tilde{\vartheta}(G'') \leq 0$ observe that $w_{v,\sigma}(m) > w_{v,\sigma}\left(m_{2|\Theta|-2}\right)$ for all $m \notin G''$ and therefore $\nu^w(m) = 0$ for all $m \notin G''$. Substitute that into the first equality in (79) to get

$$\begin{aligned}
\gamma_t(v) &= \frac{\hat{\beta}_{t+1}}{\beta}\mathbb{E}_\sigma k'_{t+1}(w_{v,\sigma}) + \nu^w\left(G''\right) \leq \frac{\hat{\beta}_{t+1}}{\beta}k'_{t+1}\left(w_{v,\sigma}\left(m_{2|\Theta|-2}\right)\right) + \nu^w\left(G''\right)\\
&\leq \frac{\hat{\beta}_{t+1}}{\beta}k'_{t+1}\left(w_{v,\sigma}\left(m_{2|\Theta|-2}\right)\right) + \frac{\nu^w(G'')}{\Pr(G'')}.
\end{aligned}$$

Then (82) implies that $\tilde{\vartheta}(G'') \leq 0$.

We first characterize the boundary conditions when $\gamma_t(v) > 1$. In this case (83) together with $\tilde{\vartheta}(G'') \leq 0$ implies that $C'\left(u_{v,\sigma}\left(m_{2|\Theta|-2}\right)\right) < 0$, which is impossible. Therefore $(u_{v,\sigma}(m), w_{v,\sigma}(m))$ must be the same for all m, the case that we already considered above.

Alternatively suppose that $\gamma_t(v) \leq 1$. We establish first that $\nu^u(m) = 0$ for all m. Since our maximization problem is strictly convex, we can guess and verify that all multipliers on the boundary conditions are zero. In this case (81) shows that

$$\zeta_t C'(u_{v,\sigma}(m_1)) = (1 - \gamma_t(v)) \mathbb{E}_\sigma [\theta | G'] + \frac{\tilde{\vartheta}(G')}{\Pr(G')} \theta' \geq 0,$$

where the inequality follows from $\tilde{\vartheta}(G') \geq 0$. This establishes that $u_{v,\sigma}(m_1) \geq \underline{u}$. Monotonicity (73) shows that $u_{v,\sigma}(m) \geq \underline{u}$ for all m, verifying our guess. Since $\mathbb{E}_\sigma[\theta|m] \in [\theta_1, \theta_{|\Theta|}]$, bounds (75) then follow from (81), (83), (73), and $\tilde{\vartheta}(G') \geq 0$, $\tilde{\vartheta}(G'') \leq 0$.

It remains to show the boundary condition (76) when $\gamma_t(v) \leq 1$. To obtain bounds for $k'_{t+1}(w_{v,\sigma})$, substitute for $\tilde{\vartheta}(G'), \tilde{\vartheta}(G'')$ from (81) and (83) into (80) and (82). Then

$$\begin{aligned} \frac{\hat{\beta}_{t+1}}{\beta} k'_{t+1}\left(w_{v,\sigma}\left(m_{2|\Theta|-2}\right)\right) &= \gamma_t(v) + \frac{1-\gamma_t(v)}{\theta''} \mathbb{E}_\sigma[\theta|G''] - \frac{1}{\theta''} + \frac{1}{\theta''}\left(1 - \zeta_t C'\left(u_{v,\sigma}\left(m_{2|\Theta|-2}\right)\right)\right) - \frac{\nu^w(G'')}{\Pr(G'')} \\ &\leq \gamma_t(v) + \frac{1-\gamma_t(v)}{\theta''}\theta_{|\Theta|} - \frac{1-\gamma_t(v)}{\theta''}. \end{aligned}$$

Re-arrange to get

$$\begin{aligned} 1 - k'_{t+1}\left(w_{v,\sigma}\left(m_{2|\Theta|-2}\right)\right) &\geq (1-\gamma_t(v)) \frac{\beta}{\hat{\beta}_{t+1}}\left(1 - \frac{\theta_{|\Theta|}-1}{\theta''}\right) + \left(1 - \frac{\beta}{\hat{\beta}_{t+1}}\right) \\ &\geq (1-\gamma_t(v)) \frac{\beta}{\hat{\beta}_{t+1}}\left(1 - \frac{\theta_{|\Theta|}-1}{\theta_1}\right) + \left(1 - \frac{\beta}{\hat{\beta}_{t+1}}\right). \end{aligned}$$

The other inequality in (76) is shown analogously using the fact that $\nu^w(G') = 0$. ■

The next Corollary states an implication of this lemma that is used throughout in the proofs.

Corollary 3 *There are $\underline{a}_u(v), \bar{a}_u(v), \underline{a}_w(v), \bar{a}_w(v)$ such that C is defined over $[\underline{a}_u(v), \bar{a}_u(v)]$, $u_{v,\sigma}(m) \in [\underline{a}_u(v), \bar{a}_u(v)]$, $w_{v,\sigma}(m) \in [\underline{a}_w(v), \bar{a}_w(v)]$ for all σ and m such that $\sigma(m|\theta) > 0$ for some θ. If utility is either bounded below or $\chi_{t+1} > 0$, then $\underline{a}_u(\cdot), \bar{a}_u(\cdot), \underline{a}_w(\cdot), \bar{a}_w(\cdot)$ can be chosen to be constants for all v sufficiently low.*

Proof. Without loss of generality, suppose all m are sent with positive probability and (73) is satisfied. First, suppose utility is bounded below. If $\gamma_t(v) \leq 1$, then (75) and (76) define compact sets for $u_{v,\sigma}(m)$ and $w_{v,\sigma}(m)$. If $\gamma_t(v) > 1$, then $u_{v,\sigma}(m)$ and $w_{v,\sigma}(m)$ do not depend on m by Lemma 13.

Alternatively, suppose that utility is unbounded below, so in this case $\lim_{v\to-\infty} \gamma_t(v) = 1$ by Lemma 7. Then equation (75) implies that there is $A(v)$ such that $|u_{v,\sigma}(m'') - u_{v,\sigma}(m')| \leq$

$A(v)$ for all σ, m', m'' sent with positive probability. The incentive constraint

$$\theta_{|\Theta|}u_{v,\sigma}\left(m_{2|\Theta|-2}\right) + \beta w_{v,\sigma}\left(m_{2|\Theta|-2}\right) \geq \theta_{|\Theta|}u_{v,\sigma}\left(m_1\right) + \beta w_{v,\sigma}\left(m_1\right)$$

together with monotonicity (73) imply that

$$\frac{\theta_{|\Theta|}}{\beta}\left(u_{v,\sigma}\left(m_{2|\Theta|-2}\right) - u_{v,\sigma}\left(m_1\right)\right) \geq w_{v,\sigma}\left(m_1\right) - w_{v,\sigma}\left(m_{2|\Theta|-2}\right) \geq 0$$

establishing that $|w_{v,\sigma}(m'') - w_{v,\sigma}(m')| \leq \frac{\theta_{|\Theta|}}{\beta}A(v)$. Let $\tilde{w}_v$ be defined by $\frac{\hat{\beta}_{t+1}}{\beta}k'_{t+1}(\tilde{w}_v) = \gamma_t(v)$. Then $w_{v,\sigma}(m_1) \geq \tilde{w}_v \geq w_{v,\sigma}\left(m_{2|\Theta|-2}\right)$ by Lemma 13, which establishes that $w_{v,\sigma}(m) \in \left[\tilde{w}_v - \frac{\theta_{|\Theta|}}{\beta}A(v), \tilde{w}_v + \frac{\theta_{|\Theta|}}{\beta}A(v)\right]$ for all v. We show analogously that $u_{v,\sigma}(m)$ lies in the compact set independent of σ, m.

It remains to show that the boundaries of this set are independent of v if utility is unbounded, $\chi_{t+1} > 0$ and v is sufficiently low (the bounded utility case it trivial). $\chi_{t+1} > 0$ implies $\beta/\hat{\beta}_{t+1} < 1$ and therefore expression (76) implies that there exists v_t^- such that

$$\frac{1}{2}\left(1 - \frac{\beta}{\hat{\beta}_{t+1}}\right) \leq 1 - k'_{t+1}\left(w_{v,\sigma}(m)\right) \leq \frac{3}{2}\left(1 - \frac{\beta}{\hat{\beta}_{t+1}}\right) \text{ for all } m, \sigma, v \leq v_t^-.$$

This establishes bounds for $w_{v,\sigma}(m)$. The incentive constraint

$$\theta_1 u_{v,\sigma}\left(m_1\right) + \beta w_{v,\sigma}\left(m_1\right) \geq \theta_1 u_{v,\sigma}\left(m_{2|\Theta|-2}\right) + \beta w_{v,\sigma}\left(m_{2|\Theta|-2}\right)$$

together with monotonicity (73) establishes bounds for $u_{v,\sigma}(m)$. ■

Lemma 14 *Suppose Assumption 1 is satisfied. Then C'' is continuous and $\lim_{u\to\bar{u}} \frac{C''(u)}{[C'(u)]^2} = 0$. If, in addition, k_t is twice differentiable then $\lim_{v\to\bar{v}} \frac{k''_t(v)}{[1-k'_t(v)]^2} = 0$.*

Proof. By definition $C(U(c)) = c$. Differentiate twice to obtain $C'U' = 1$ and $C''[U']^2 + C'U'' = 0$. Since U'' continuous, C'' is also continuous from the second expression. The two expressions together imply

$$\frac{C''(U(c))}{[C'(U(c))]^2} = -\frac{U''(c)}{U'(c)}.$$

If assumption Assumption 1 is satisfied then $\lim_{u\to\bar{u}} \frac{C''(u)}{[C'(u)]^2} = 0$.

Suppose k_t is twice differentiable and v satisfies $\gamma_t(v) < 1$. Then by Lemmas 8 and 13 $\mathbf{u}_v(\cdot, z)$ is interior and therefore the proof of (70) applies, establishing that

$$0 \leq -k''_t(v) \leq \zeta_t \mathbb{E}_{\boldsymbol{\sigma}_v} C''(\mathbf{u}_v) = (1-\gamma_t(v))^2 \zeta_t \mathbb{E}_{\boldsymbol{\sigma}_v} \frac{C''(\mathbf{u}_v)}{[C'(\mathbf{u}_v)]^2}\left[\frac{C'(\mathbf{u}_v)}{1-\gamma_t(v)}\right]^2. \tag{84}$$

Since $\mathbf{u}_v(\cdot, z)$ satisfies bounds (75) for each z, we have $\frac{C'(\mathbf{u}_v(m,z))}{1-\gamma_t(v)} \in \left[\theta_1, \theta_{|\Theta|}\right]$, $\frac{C''(\mathbf{u}_v(m,z))}{[C'(\mathbf{u}_v(m,z))]^2} \to 0, \gamma_t(v) \to -\infty$ as $v \to \bar{v}$, uniformly in (m, z). Since $k_t'(v) = \gamma_t(v)$, this establishes $\lim_{v \to \bar{v}} \frac{k_t''(v)}{[1-k_t'(v)]^2} = 0$. ■

Lemma 15 *Suppose* $|\Theta| = 2$.

(a). If either utility is bounded below or $\chi_{t+1} > 0$ *then* $\lim_{v \to \underline{v}} [\kappa_t(v, \sigma) - \kappa_t(v, \sigma^{un})] = 0$ *for all* σ.

(b). If Assumption 1 is satisfied then $\lim_{v \to \bar{v}} \left[\kappa_t\left(v, \sigma^{in}\right) - \kappa_t(v, \sigma)\right] = \infty$ *for all* $\sigma \notin \Sigma^{in}$.

Proof. *(a).* As in the proof of Lemma 13 we assume that all messages are sent with positive probability and (73) holds. Define the allocation $\left(u_{v,\sigma}^*, w_{v,\sigma}^*\right)$ where

$$u_{v,\sigma}^*(m) = \mathbb{E}_\sigma \theta u_{v,\sigma}, w_{v,\sigma}^*(m) = \mathbb{E}_\sigma w_{v,\sigma} \text{ for all } m. \tag{85}$$

Since the profile $\left(u_{v,\sigma}^*, w_{v,\sigma}^*\right)$ is incentive compatible for any σ we must have

$$\kappa_t(v, \sigma^{un}) \geq (1 - \gamma_t(v)) u_{v,\sigma}^* - \zeta_t C\left(u_{v,\sigma}^*\right) + \hat{\beta}_{t+1} k_{t+1}\left(w_{v,\sigma}^*\right) - \gamma_t(v) \beta w_{v,\sigma}^* + \gamma_t(v) v. \tag{86}$$

Therefore,

$$\begin{aligned} 0 &\leq \kappa_t(v, \sigma) - \kappa_t(v, \sigma^{un}) \qquad (87) \\ &\leq \mathbb{E}_\sigma \left\{ \left[(1 - \gamma_t(v)) \theta u_{v,\sigma} - \zeta_t C(u_{v,\sigma}) + \hat{\beta}_{t+1} k_{t+1}(w_{v,\sigma}) - \gamma_t(v) \beta w_{v,\sigma}\right] \right. \\ &\quad \left. - \left[(1 - \gamma_t(v)) \theta u_{v,\sigma}^* - \zeta_t C\left(u_{v,\sigma}^*\right) + \hat{\beta}_{t+1} k_{t+1}\left(w_{v,\sigma}^*\right) - \gamma_t(v) \beta w_{v,\sigma}^*\right] \right\} \\ &= \mathbb{E}_\sigma \left[-\zeta_t \left\{C(u_{v,\sigma}) - C\left(u_{v,\sigma}^*\right)\right\} + \hat{\beta}_{t+1} \left\{k_{t+1}(w_{v,\sigma}) - k_{t+1}\left(w_{v,\sigma}^*\right)\right\} - \gamma_t(v) \beta \left\{w_{v,\sigma} - w_{v,\sigma}^*\right\}\right], \end{aligned}$$

where the second inequality follows from the fact the that right hand side of (86) does not depend on (θ, m) and the equality follows from (85).

First, suppose that utility is bounded below. From Lemma 13, $u_{v,\sigma}(m) \to \underline{u}$ and $w_{v,\sigma}(m) \to \hat{w}$ for all m and σ as $v \to \underline{v}$ and, thus, $u_{v,\sigma}^* \to \underline{u}, w_{v,\sigma}^* \to \hat{w}$. Therefore, $\kappa_t(v, \sigma) - \kappa_t(v, \sigma^{un}) \to 0$ for all σ.

Now suppose that utility is unbounded. Apply the mean value theorem to (87):

$$0 \leq \kappa_t(v, \sigma) - \kappa_t(v, \sigma^{un}) \leq \mathbb{E}_\sigma \left[-\zeta_t C'(\grave{u}_{v,\sigma}) \left(u_{v,\sigma} - u_{v,\sigma}^*\right) + \hat{\beta}_{t+1} \left\{k_{t+1}'(\grave{w}_{v,\sigma}) - \frac{\beta}{\hat{\beta}_{t+1}}\right\} \left(w_{v,\sigma} - w_{v,\sigma}^*\right)\right] \tag{88}$$

for some $\grave{u}_{v,\sigma}(m) \in \left[u_{v,\sigma}(m_1), u_{v,\sigma}\left(m_{2|\Theta|-2}\right)\right]$, $\grave{w}_{v,\sigma}(m) \in \left[w_{v,\sigma}\left(m_{2|\Theta|-2}\right), w_{v,\sigma}(m_1)\right]$. Therefore $\lim_{v \to -\infty} C'(\grave{u}_{v,\sigma}) = 0$, $\lim_{v \to -\infty} \left\{k_{t+1}'(\grave{w}_{v,\sigma}) - \frac{\beta}{\hat{\beta}_{t+1}}\right\} = 0$ by Lemma 13. If $\chi_{t+1} > 0$ then

$u_{v,\sigma}, u^*_{v,\sigma} \in [\underline{a}_u, \bar{a}_u], w_{v,\sigma}, w^*_{v,\sigma} \in [\underline{a}_w, \bar{a}_w]$ for some reals $\underline{a}_u, \bar{a}_u, \underline{a}_w, \bar{a}_w$ for sufficiently low v by Corollary 3, and the right hand side of equation (88) converges to 0 as $v \to -\infty$.

(b). We first show that $\lim_{v\to\bar{v}} \left[\kappa_t\left(v, \sigma^{in}\right) - \kappa_t\left(v, \sigma^{un}\right)\right] = \infty$ for any uninformative σ^{un}. Since all uninformative strategies give the same payoff, it is sufficient to show this for $\hat{\sigma}$ such that $\hat{\sigma}(m_2|\theta) = 1$ for all θ. We consider v to be sufficiently high so that $\gamma_t(v) < 1$, $(u_{v,\hat{\sigma}}, w_{v,\hat{\sigma}})$ is interior and by Lemma 13 satisfies

$$\frac{\hat{\beta}_{t+1}}{\beta} k'_{t+1}\left(w_{v,\hat{\sigma}}(m_2)\right) = 1 - \zeta_t C'\left(u_{v,\hat{\sigma}}(m_2)\right) = \gamma_t(v). \tag{89}$$

We consider an informative strategy σ^{in} in which type θ_2 reports m_2 with probability 1 and receives $\left(u_{v,\hat{\sigma}}(m_2), w_{v,\hat{\sigma}}(m_2)\right)$, while type θ_1 reports m_1 with probability 1 and receives $\left(u_{v,\hat{\sigma}}(m_2) - \hat{x}_v, w_{v,\hat{\sigma}}(m_2) + \frac{\theta_1}{\beta}\hat{x}_v\right)$ for some $\hat{x}_v > 0$ that we define below. Observe that this allocation is incentive compatible for any $\hat{x}_v \geq 0$. Let $F(v, \hat{x}_v)$ be the value of such strategy. Obviously $\kappa_t\left(v, \sigma^{in}\right) \geq F(v, \hat{x}_v)$.

To make our expressions concise, define

$$h_v(\theta, u, w) = (1 - \gamma_t(v))\theta u - \zeta_t C(u) + \hat{\beta}_{t+1} k_{t+1}(w) - \gamma_t(v)\beta w + \gamma_t(v) v$$

and consider a function $f(x) \equiv h_v\left(\theta_1, u_{v,\hat{\sigma}}(m_2) - x, w_{v,\hat{\sigma}}(m_2) + \frac{\theta_1}{\beta}x\right)$. This function is strictly concave with $f'(0) = (1 - \gamma_t(v))(1 - \theta_1) > 0$ from (89). Let $\hat{x}_v$ be a solution to $f'(\hat{x}_v) = \frac{1}{2}(1 - \gamma_t(v))(1 - \theta_1)$. By strict concavity $\hat{x}_v > 0$. Moreover, it is easy to verify that for any $x \in [0, x^*_v]$, where x^*_v solves $f'(x^*_v) = 0$, the allocation $\left(u_{v,\hat{\sigma}}(m_2) - x, w_{v,\hat{\sigma}}(m_2) + \frac{\theta_1}{\beta}x\right)$ satisfies bounds (75) and (76). Therefore $\left(u_{v,\hat{\sigma}}(m_2) - \hat{x}_v, w_{v,\hat{\sigma}}(m_2) + \frac{\theta_1}{\beta}\hat{x}_v\right)$ satisfies these bounds.

We have

$$F(v, \hat{x}_v) - \kappa_t(v, \hat{\sigma}) = \pi(\theta_1)\left[f(\hat{x}_v) - f(0)\right] = \pi(\theta_1)\frac{f'(\tilde{x}_v)}{1 - \gamma_t(v)}(1 - \gamma_t(v))\hat{x}_v$$

for some $\tilde{x}_v \in (0, \hat{x}_v)$ from the mean value theorem. Convexity of f implies that $\frac{f'(\tilde{x}_v)}{1-\gamma_t(v)} \in \left[\frac{1}{2}(1 - \theta_1), (1 - \theta_1)\right]$. We next show that $\lim_{v\to\bar{v}}(1 - \gamma_t(v))\hat{x}_v = \infty$ if Assumption 1 is satisfied. Since $\kappa_t\left(v, \sigma^{in}\right) - \kappa_t\left(v, \sigma^{un}\right) \geq F(v, \hat{x}_v) - \kappa_t(v, \hat{\sigma})$ it establishes that $\lim_{v\to\bar{v}}\left[\kappa_t\left(v, \sigma^{in}\right) - \kappa_t\left(v, \sigma^{un}\right)\right] = \infty$. To simplify the exposition, we assume that k_{t+1} is twice differentiable. In Supplementary material we extend these arguments to the cases when k_{t+1} does not satisfy this assumption.

If k_{t+1} is twice differentiable, so is f, and applying the mean value theorem we have

$$\frac{1 - \theta_1}{2} = \frac{f'(0) - f'(\hat{x}_v)}{1 - \gamma_t(v)} = \frac{-f''(\check{x}_v)}{\left[1 - \gamma_t(v)\right]^2}(1 - \gamma_t(v))\hat{x}_v \tag{90}$$

for some $\grave{x}_v \in [0, \hat{x}_v]$. Using direct calculations and taking limit as $v \to \bar{v}$

$$\frac{-f''(\grave{x}_v)}{[1-\gamma_t(v)]^2} = \zeta_t \frac{C''(u_{v,\hat{\sigma}}(m_2) - \grave{x}_v)}{[1-\gamma_t(v)]^2} + \hat{\beta}_{t+1}\left(\frac{\theta_1}{\beta}\right)^2 \frac{k''_{t+1}\left(w_{v,\hat{\sigma}}(m_2) + \frac{\theta_1}{\beta}\grave{x}_v\right)}{[1-\gamma_t(v)]^2}$$

$$= \zeta_t \underbrace{\frac{C''(u_{v,\hat{\sigma}}(m_2) - \grave{x}_v)}{[C'(u_{v,\hat{\sigma}}(m_2) - \grave{x}_v)]^2}}_{\to 0 \text{ by Lemma 14}} \left(\underbrace{\frac{C'(u_{v,\hat{\sigma}}(m_2) - \grave{x}_v)}{C'(u_{v,\hat{\sigma}}(m_2))}}_{\leq 1}\right)^2 \left(\underbrace{\frac{C'(u_{v,\hat{\sigma}}(m_2))}{1-\gamma_t(v)}}_{\text{bounded by Lemma 13}}\right)^2$$

$$-\underbrace{\frac{\hat{\beta}_{t+1}\left(\frac{\theta_1}{\beta}\right)^2 k''_{t+1}\left(w_{v,\hat{\sigma}}(m_2) + \frac{\theta_1}{\beta}\grave{x}_v\right)}{\left[1-k'_{t+1}\left(w_{v,\hat{\sigma}}(m_2) + \frac{\theta_1}{\beta}\grave{x}_v\right)\right]^2}}_{\to 0 \text{ by Lemma 14}} \left(\underbrace{\frac{1-k'_{t+1}\left(w_{v,\hat{\sigma}}(m_2) + \frac{\theta_1}{\beta}\grave{x}_v\right)}{1-k'_{t+1}\left(w_{v,\hat{\sigma}}(m_2) + \frac{\theta_1}{\beta}\hat{x}_v\right)}}_{\leq 1}\right)^2 \left(\underbrace{\frac{1-k'_{t+1}\left(w_{v,\hat{\sigma}}(m_2) + \frac{\theta_1}{\beta}\hat{x}_v\right)}{1-\gamma_t(v)}}_{\text{bounded by Lemma 13}}\right)^2.$$

Allocation $\left(u_{v,\hat{\sigma}}(m_2) - \grave{x}_v, w_{v,\hat{\sigma}}(m_2) + \frac{\theta_1}{\beta}\grave{x}_v\right)$ satisfies bounds (75) and (76) since $\grave{x}_v \in [0, x_v^*]$, therefore, it goes to $(\bar{u}, \bar{v})$ as $v \to \bar{v}$. Then Lemmas 13 and 14 imply that $\lim_{v\to\bar{v}} \frac{-f''(\grave{x}_v)}{[1-\gamma_t(v)]^2} = 0$. Equation (90) then implies that $\lim_{v\to\bar{v}} (1-\gamma_t(v))\hat{x}_v = \infty$.

It remains to show our result for any σ that is not uninformative. If $\sigma \notin \Sigma^+$ then no insurance is possible, $\kappa(v,\sigma) = \kappa(v,\sigma^{un})$, and our previous arguments apply. Consider any $\sigma \in \Sigma^+ \backslash \Sigma^{in}$, which in the case of $|\Theta| = 2$ is equivalent to a σ such that there is message m and type θ with $\sigma(m|\theta) \in (0,1)$ and $\sigma(m|\theta') = 0$ for $\theta' \neq \theta$. Without loss of generality let (m_1, θ_1) be such pair. Let σ^{in} be an informative strategy such that $\sigma^{in}(m_1|\theta_1) = 1$ and $\sigma^{in}(m_2|\theta_2) = 1$, and let σ'' be a strategy such that $\sigma''(m_2|\theta) = 1$ for all θ. Since $(u_{v,\sigma}, w_{v,\sigma}) \in X(\sigma^{in})$ and $(u_{v,\sigma}, w_{v,\sigma}) \in X(\sigma'')$,

$$\begin{aligned}
\kappa_t(v,\sigma^{in}) - \kappa_t(v,\sigma) &= \mathbb{E}_{\sigma^{in}}\left[h_v\left(\theta, u_{v,\sigma^{in}}, w_{v,\sigma^{in}}\right)\right] - \mathbb{E}_\sigma\left[h_v(\theta, u_{v,\sigma}, w_{v,\sigma})\right] \\
&\geq \pi(\theta_1)(1-\sigma(m_1|\theta_1))\left[h_v(\theta_1, u_{v,\sigma}(m_1), w_{v,\sigma}(m_1)) - h_v(\theta_1, u_{v,\sigma}(m_2), w_{v,\sigma}(m_2))\right]
\end{aligned}$$

and

$$\begin{aligned}
\kappa_t(v,\sigma) - \kappa_t(v,\sigma'') &= \mathbb{E}_\sigma\left[h_v(\theta, u_{v,\sigma}, w_{v,\sigma})\right] - \mathbb{E}_{\sigma''}\left[h_v\left(\theta, u_{v,\sigma''}, w_{v,\sigma''}\right)\right] \\
&\leq \pi(\theta_1)\sigma(m_1|\theta_1)\left[h_v(\theta_1, u_{v,\sigma}(m_1), w_{v,\sigma}(m_1)) - h_v(\theta_1, u_{v,\sigma}(m_2), w_{v,\sigma}(m_2))\right].
\end{aligned}$$

Combining these inequalities,

$$\kappa_t(v,\sigma^{in}) - \kappa_t(v,\sigma) \geq \frac{1-\sigma(m_1|\theta_1)}{\sigma(m_1|\theta_1)}\left(\kappa_t(v,\sigma) - \kappa_t(v,\sigma'')\right) \geq \frac{1-\sigma(m_1|\theta_1)}{\sigma(m_1|\theta_1)}\left(\kappa_t(v,\sigma) - \kappa_t(v,\sigma^{un})\right).$$

Therefore

$$\kappa_t(v,\sigma^{in}) - \kappa_t(v,\sigma^{un}) = \left\{\kappa_t(v,\sigma^{in}) - \kappa_t(v,\sigma)\right\} + \left\{\kappa_t(v,\sigma) - \kappa_t(v,\sigma^{un})\right\} \leq \frac{\kappa_t(v,\sigma^{in}) - \kappa_t(v,\sigma)}{1-\sigma(m_1|\theta_1)}.$$

Our previous result then implies that $\lim_{v\to\bar{v}}\left\{\kappa_t\left(v,\sigma^{in}\right)-\kappa_t\left(v,\sigma\right)\right\}=\infty$. ■

For any $M_{v,\sigma}(\theta)$, consider the alternative constraint

$$\theta u(m)+\beta w(m)\geq\theta u\left(m'\right)+\beta w\left(m'\right)\text{ for all }\theta,m\in M_{v,\sigma}(\theta),\text{all }m'. \tag{91}$$

Observe that the maximization of (33) subject to (34) and (35) is equivalent to the maximization of (33) over (91).

Remark 1 *Constraint (91) is smaller than constraint (34)-(35) since it imposes restrictions on measure-zero m. However, reporting measure-zero m is not incentive compatible under (34)-(35), so both the value of (33) and the set of maximizers sent with positive probability are the same.*

We now consider some properties of the derivatives of κ_t and W_t. For any $\sigma,\sigma',\alpha\in(0,1)$ let $\sigma_\alpha=(1-\alpha)\sigma+\alpha\sigma'$ and consider the set of messages sent with positive probability under σ_α. This set is independent of α. Let u_α^w be a solution to (26) and (u_α,w_α) be a solution to (33) for σ_α. Since, holding σ_α fixed, these problems are strictly convex, these solutions are unique for any m sent with positive probability. Let $u_0^w(m)=\lim_{\alpha\to0}u_\alpha^w(m)$ and $(u_0(m),w_0(m))=\lim_{\alpha\to0}(u_\alpha(m),w_\alpha(m))$ for such m. u_α^w and (u_α,w_α) can be restricted to lie in a compact set that does not depend on α by (49) and Corollary 3, respectively. Therefore, by the Maximum theorem these limits exists and u_0^w and (u_0,w_0) are, respectively, solutions to (26) and (33) for σ_0, although they may not be unique for the messages sent with zero probability under the reporting strategy σ_0.

Lemma 16 *(a). For any σ,σ', the derivative $\frac{\partial W_t(\sigma)}{\partial\sigma'}$ exists, is bounded, and*

$$\frac{\partial W_t(\sigma)}{\partial\sigma'}=\mathbb{E}_{\sigma'}\left[\theta u_0^w(m)-\lambda_t^wC\left(u_0^w(m)\right)\right]-\mathbb{E}_\sigma\left[\theta u_0^w(m)-\lambda_t^wC\left(u_0^w(m)\right)\right]\leq W_t\left(\sigma'\right)-W_t(\sigma). \tag{92}$$

For each t, there is $\varepsilon>0$ such that, for any $\sigma^{un}\in\Sigma^{un}$ which is the limit of some sequence $\{\sigma_n\}_n$ with $\sigma_n\in\Sigma^+$, there exists σ^{in} such that $\frac{\partial W_t(\sigma^{un})}{\partial\sigma^{in}}\geq\varepsilon$.

(b) For any v and σ take any $M_{v,\sigma}$. For any strategy σ', with a property that $\sigma'(m|\theta)>0$ only if $m\in M_{v,\sigma}(\theta)$, the derivative $\frac{\partial\kappa_t(v,\sigma)}{\partial\sigma'}$ exists and

$$\frac{\partial\kappa_t(v,\sigma)}{\partial\sigma'}=\mathbb{E}_{\sigma'}\left[(1-\gamma_t(v))\theta u_0-\zeta_tC(u_0)+\hat{\beta}_{t+1}k_{t+1}(w_0)-\gamma_t(v)\beta w_0\right] \tag{93}$$
$$-\mathbb{E}_\sigma\left[(1-\gamma_t(v))\theta u_0-\zeta_tC(u_0)+\hat{\beta}_{t+1}k_{t+1}(w_0)-\gamma_t(v)\beta w_0\right]\leq\kappa_t\left(v,\sigma'\right)-\kappa_t(v,\sigma).$$

Proof. *(a).* For any random variable $x(m) \in X$ for some set X, the family $\{\mathbb{E}_{\sigma_\alpha} x\}_{x \in X}$ is equidifferentiable at any $\alpha \in [0,1)$ since the expectation is linear in α. Therefore the derivative $\frac{\partial W_t(\sigma)}{\partial \sigma'}$ exists and satisfies the equality in (92) by Theorem 3 in Milgrom and Segal (2002). The inequality follows from the fact that $W_t(\sigma') \geq \mathbb{E}_{\sigma'}\left[\theta u_0^w(m) - \lambda_t^w C\left(u_0^w(m)\right)\right]$. $\frac{\partial W_t(\sigma)}{\partial \sigma'}$ is bounded since u_0^w satisfies (49).

Take some $\sigma^{un} \in \Sigma^{un}$, which is the limit of some sequence $\{\sigma_n\}_n$ with $\sigma_n \in \Sigma^+$. Since $\sigma_n \in \Sigma^+$, for all n there is at least one message m which is sent with positive probability by only one type (if all messages were sent by both types, constraints (34)-(35) would imply that $(u_{v,\sigma}(m), w_{v,\sigma}(m))$ are the same for all m sent with positive probability). Without loss of generality, let m_1 and θ_1 be such message and such type. Let σ' be defined as $\sigma'(m_1|\theta_1) = 1$, $\sigma'(m|\theta_2) = \sigma(m|\theta_2)$. Clearly $\sigma' \in \Sigma^{in}$ since $\sigma'(m_1|\theta_2) = 0$. We have $u_0^w(m_1) = \frac{\theta_1}{\lambda_t^w}$ and $u_0^w(m) = \frac{1}{\lambda_t^w}$ for all m sent with positive probability by σ_α for $\alpha > 0$. This implies that there is some $\varepsilon > 0$ such that $\frac{\partial W_t(\sigma^{un})}{\partial \sigma^{in}} \geq \varepsilon$.

(b). Let σ, σ' be as defined in the statement. Then $\sigma_\alpha(m|\theta) > 0$ only if $m \in M_{v,\sigma}(\theta)$. Therefore for all σ_α, $\alpha \in [0,1)$, the constraint set to problem (33) can be written as (91), i.e. independent of α. Therefore we can apply Theorem 3 in Milgrom and Segal (2002) as in part (a). ■

Lemma 17 *If the derivative $\frac{\partial \kappa_t(v,\sigma_v)}{\partial \sigma'}$ exists for some σ' then*

$$\frac{\partial \kappa_t(v,\sigma_v)}{\partial \sigma'} \leq \chi_t \frac{\partial W_t(\sigma_v)}{\partial \sigma'}. \tag{94}$$

Moreover, if $\sigma_v \in \Sigma^+ \backslash \Sigma^{in}$ then there are σ' for which (94) holds with equality. In particular, σ' can be chosen to be in Σ^{in} and in Σ^{un}.

Proof. Since σ_v is optimal,

$$\frac{1}{\alpha}\left[\kappa_t\left(v, \alpha\sigma' + (1-\alpha)\sigma_v\right) - \chi_t W_t\left(\alpha\sigma' + (1-\alpha)\sigma_v\right) - \{\kappa_t(v,\sigma_v) - \chi_t W_t(\sigma_v)\}\right] \leq 0$$

for any $\alpha > 0$. By assumption the limit exists as $\alpha \to 0$, establishing the first part.

Suppose $\sigma_v \in \Sigma^+ \backslash \Sigma^{in}$. Then there must exist some $m', m'', \theta', \theta''$ such that $\sigma_v\left(m'|\theta'\right) > 0$, $\sigma_v\left(m'|\theta''\right) = 0$ and $\sigma_v\left(m''|\theta'\right) > 0$, $\sigma_v\left(m''|\theta''\right) > 0$. Without loss of generality let $m' = m_1, \theta' = \theta_1$. Let σ' be defined as defined in the proof of Lemma 16(a) and let

$$F(\theta, m) = (1 - \gamma_t(v))\,\theta u_0(m) - \zeta_t C\left(u_0(m)\right) + \hat{\beta}_{t+1} k_{t+1}\left(w_0(m)\right) - \gamma_t(v)\,\beta w_0(m).$$

By construction $\sigma'(m|\theta) > 0$ only if $m \in M_{v,\sigma_v}(\theta)$ so the derivative $\frac{\partial \kappa_t(v,\sigma_v)}{\partial \sigma'}$ exists by Lemma 16(b). Also, suppose $\tilde{m}, \hat{m}$ are sent with positive probability by both types under σ_v, then (35) implies that $(u_0(\tilde{m}), w_0(\tilde{m})) = (u_0(\hat{m}), w_0(\hat{m}))$ and, thus, $F(\theta, \tilde{m}) = F(\theta, \hat{m})$ for all θ. Also, from the proof of Lemma 9, since σ_v is optimal, it must be that $\mathbb{E}_{\sigma_v}[\theta|\tilde{m}] = \mathbb{E}_{\sigma_v}[\theta|\hat{m}]$ and, thus, $u_0^w(\tilde{m}) = u_0^w(\hat{m})$. Substitute (92) and (93) into (94) and divide by $\pi(\theta_1)\sigma_v(m''|\theta_1) > 0$ to get

$$\begin{aligned} & \chi_t \left\{ \left[\theta_1 u_0^w(m_1) - \lambda_t^w C(u_0^w(m_1))\right] - \left[\theta_1 u_0^w(m'') - \lambda_t^w C(u_0^w(m''))\right] \right\} \\ \geq & \frac{\partial \kappa_t(v, \sigma_v)/\partial \sigma'}{\pi(\theta_1)\sigma_v(m''|\theta_1)} = F(\theta_1, m_1) - F(\theta_1, m''). \end{aligned} \tag{95}$$

Alternatively, let σ'' be defined as $\sigma''(m''|\theta_1) = 1$, $\sigma''(m|\theta_2) = \sigma_v(m|\theta_2)$ for all m. By construction, $\sigma''(m|\theta) > 0$ only if $m \in M_{v,\sigma_v}(\theta)$, therefore, the same steps as above establish the reverse inequality in (95). Therefore (95) holds with equality. Since $\sigma' \in \Sigma^{in}$, we conclude that (94) holds with equality for some fully informative σ'.

It remains to show that there is some σ^{un} such that the derivative $\frac{\partial \kappa_t(v,\sigma_v)}{\partial \sigma^{un}}$ exists and satisfies (94) with equality. Define $\sigma^{un}(m''|\theta) = 1$ for all θ. By Lemma 16(b) $\frac{\partial \kappa_t(v,\sigma_v)}{\partial \sigma^{un}}$ exists. Using (92) and (93) and the fact that, if $\tilde{m}, \hat{m}$ are sent with positive probability by both types, then $F(\theta, \tilde{m}) = F(\theta, \hat{m})$, for all θ, and $u_0^w(\tilde{m}) = u_0^w(\hat{m})$, we have

$$\begin{aligned} & \frac{\partial \kappa_t(v, \sigma_v)}{\partial \sigma^{un}} - \chi_t \frac{\partial W_t(\sigma_v)}{\partial \sigma^{un}} \\ = & \sum_{\theta, m} \pi(\theta)\sigma_v(m|\theta)\left(F(\theta, m'') - F(\theta, m)\right) \\ & -\chi_t \sum_{\theta, m} \pi(\theta)\sigma_v(m|\theta)\left\{\left[\theta u_0^w(m'') - \lambda_t^w C(u_0^w(m''))\right] - \left[\theta u_0^w(m) - \lambda_t^w C(u_0^w(m))\right]\right\}, \end{aligned}$$

for all m sent with positive probability only by one type. The last expression is zero by the fact that (95) holds with equality. ■

Proof of Proposition 4. *(a).* Since σ_v is optimal,

$$\left[\kappa_t(v, \sigma_v) - \chi_t W_t(\sigma_v)\right] - \left[\kappa_t(v, \sigma^{un}) - \chi_t W_t(\sigma^{un})\right] \geq 0.$$

This, together with Lemma 15(a) and $W_t(\sigma) \geq W_t(\sigma^{un})$ for all σ by Lemma 7, implies that $\lim_{v \to \underline{v}} W_t(\sigma_v) = W_t(\sigma^{un})$. Suppose a cutoff v_t^- does not exist. Then there is sequence $\{\sigma_{v_n}\}_n$ with $v_n \to \underline{v}$ such that $\sigma_{v_n} \in \Sigma^+$. Since $\{\sigma_{v_n}\}_n$ lie in a compact set, we can choose a convergent subsequence $\{\sigma_{v_{n'}}\}_{n'}$. We must have $\sigma_{v_{n'}} \to \sigma^{un}$ for some σ^{un} since otherwise $\lim_{n' \to \infty} W_t(\sigma_{v_{n'}}) > W_t(\sigma^{un})$ by Lemma 7. Therefore, for n' sufficiently high $\sigma_{v_{n'}} \in \Sigma^+ \backslash \Sigma^{in}$ and by Lemma 17 there exists σ^{in} such that

$$\chi_t \frac{\partial W(\sigma_{v_{n'}})}{\partial \sigma^{in}} = \frac{\partial \kappa_t(v, \sigma_{v_{n'}})}{\partial \sigma^{in}} \leq \kappa_t(v, \sigma^{in}) - \kappa_t(v, \sigma^{un}), \tag{96}$$

where the inequality follows from (93) and $\kappa_t(v,\sigma) \geq \kappa_t(v,\sigma^{un})$ for all σ. Since $\sigma_{v_{n'}} \in \Sigma^+$ there must be a message and a type, say m_1 and θ_1, such that $\sigma(m_1|\theta_1) > 0$ and $\sigma(m_1|\theta_2) = 0$. Then the same arguments in the proof of Lemma 16(a) establish that $\frac{\partial W(\sigma_{v_{n'}})}{\partial \sigma^{in}}$ is bounded away from zero (we define σ^{in} in the same way as in the proof of Lemma 16(a)) and, thus, so is $\frac{\partial \kappa_t(v,\sigma_{v_{n'}})}{\partial \sigma^{in}}$ by (96) for all n' sufficiently high. However, by Lemma 15(a) $\kappa_t\left(v,\sigma^{in}\right) - \kappa_t(v,\sigma^{un})$ converges to 0, which establishes a contradiction. Finally, since by Lemma 8 the optimal strategy $\boldsymbol{\sigma}_v(\cdot|z,\cdot)$ must be a solution to (36) for all z, the arguments above apply for all z, which proves that $\boldsymbol{\sigma}_v$ is uninformative for $v \leq v_t^-$.

(b). Suppose $\sigma_v \in \Sigma^+ \backslash \Sigma^{in}$. By Lemma 17 there exists σ^{un} such

$$-\chi_t \frac{\partial W_t(\sigma_v)}{\partial \sigma^{un}} = -\frac{\partial \kappa(v,\sigma_v)}{\partial \sigma^{un}} \geq \kappa_t(v,\sigma_v) - \kappa_t(v,\sigma^{un}),$$

where the inequality follows from (93). Since the left hand side of the equality is bounded by Lemma 16(a), the right hand side and, therefore, $\kappa_t(v,\sigma_v) - \kappa_t(v,\sigma^{un})$ must also be bounded above. Since $\kappa_t\left(v,\sigma^{in}\right) - \kappa_t(v,\sigma^{un})$ is unbounded for high v by Lemma 15, while $W_t(\sigma)$ is bounded, σ_v cannot be optimal if v is sufficiently high. Finally, since by Lemma 8 the optimal strategy $\boldsymbol{\sigma}_v(\cdot|z,\cdot)$ must be a solution to (36) for all z, the arguments above apply for all z, which proves that $\boldsymbol{\sigma}_v$ is fully informative for v sufficiently high. ■

Proof of Proposition 5. *(a).* The arguments in Lemma 15(a) do not depend on the cardinality of Θ. The key observation is that if a sequence $\{\sigma_n\}_n$ with $\sigma_n \in \Sigma^+$ converges to some σ^{un}, then for sufficiently high n either (i) there is a message m such that $\mathbb{E}_{\sigma_n}[\theta|m] = \theta_1$, or (ii) there is a message m' such that $\mathbb{E}_{\sigma_n}[\theta|m'] = \theta_{|\Theta|}$. To see this, notice that if $\sigma_n \to \sigma^{un}$ then we cannot have some type $\theta \neq \theta_1, \theta_{|\Theta|}$ to be indifferent between two messages m, m' with $u_{v,\sigma_n}(m) < u_{v,\sigma_n}(m')$ for infinitely many n. Otherwise, since the incentive constraints imply that at most one type θ can be indifferent between two distinct allocations, we would necessarily have $M_{v,\sigma_n}(\theta_1) \cap M_{v,\sigma_n}\left(\theta_{|\Theta|}\right) = \emptyset$ for infinitely many n and, thus, violate the assumption $\sigma_n \to \sigma^{un}$. Thus, for high enough n, there can be at most three messages m, m', m'' with $u_{v,\sigma_n}(m) < u_{v,\sigma_n}(m'') < u_{v,\sigma_n}(m')$ such that (i) only type θ_1 is indifferent between m and m'' and (ii) only type $\theta_{|\Theta|}$ is indifferent between m' and m''. Suppose case (i) (case (ii) is analogous), then analogous steps as in the proof of Lemma 16 show how to construct a strategy which reveals full information about type θ_1. This strategy can be used to replace σ^{in} in Lemma 17 and, thus, to replicate the arguments in the proof of part (a) of Proposition 4 for any finite Θ.

(b). It is easy to see that the arguments in Lemma 15(b) still hold if we replace σ^{in} with a strategy σ such that $\sigma(m|\theta_1) = 1$ and $\mathbb{E}_\sigma[\theta|m] = \theta_1$. Thus, we conclude that there is $v_t^+ < \bar{v}$

such that any σ which does not reveal full information about θ_1 must be suboptimal for all $v \geq v_t^+$.

To prove the statement about type $\theta_{|\Theta|}$, we can repeat the steps in the proof of Lemma 15(b), replacing the function $f(x)$ defined in that proof with the analogous function $\hat{f}(x) \equiv h_v\left(\theta_{|\Theta|}, u_{v,\sigma}\left(m_{2|\Theta|-2}\right) + x, w_{v,\sigma}\left(m_{2|\Theta|-2}\right) - \frac{\theta_{|\Theta|-1}}{\beta}x\right)$. Note that the perturbation we consider does not change the allocation for type $\theta_{|\Theta|-1}$, but gives the different allocation $\left(u_{v,\sigma}\left(m_{2|\Theta|-2}\right) + \hat{x}_v, w_{v,\sigma}\left(m_{2|\Theta|-2}\right) - \frac{\theta_{|\Theta|-1}}{\beta}\hat{x}_v\right)$ to type $\theta_{|\Theta|}$. If inequality $\pi\left(\theta_{|\Theta|-1}\right)\left(\theta_{|\Theta|} - \theta_{|\Theta|-1}\right) > \left(\pi\left(\theta_{|\Theta|-1}\right) + \pi\left(\theta_{|\Theta|}\right)\right)\left(\theta_{|\Theta|-1} - \theta_{|\Theta|-2}\right)$ holds and $1 + \theta_1 - \theta_{|\Theta|} \geq 0$, we can show that $\hat{f}'(0) = \rho\left(1 - \gamma_t(v)\right)$, for some positive constant ρ. Similar arguments as those in Lemma 15(b) establish that $\hat{f}(\hat{x}_v) - \hat{f}(0) \to \infty$.

We sketch the analysis of the last part of the proposition, leaving the details for the Supplementary material. Let $a = \frac{1}{1-\rho} > 1$ and that $C(u) = \frac{1}{a}u^a$. For all $x > 0$, define a function

$$k_t(v,x) = \frac{a}{\bar{\beta}_t}\max_{\mathbf{u},\boldsymbol{\sigma}}\mathbb{E}_{\boldsymbol{\sigma}}\left[\sum_{s=0}^{\infty}\bar{\beta}_{t+s}\left(\theta_s x^{-1}\mathbf{u}_s - \zeta_{t+s}C\left(x^{-1}\mathbf{u}_s\right) - \chi_{t+s}W_{t+s}\right)\right]$$

subject to (23) and (25). The change of variable $\tilde{\mathbf{u}}_s = \mathbf{u}_s/x$ then establishes that $x^{-a}k_t(v,x) = k_t(v/x)$. Thus solution to the maximization problem that define $k_t(1,x)$ is a normalized solution for the maximization problem that defined $k_t(1/x)$. We show that $\lim_{x\to 0}k_t(v,x) = k_t(v,0)$ where

$$k_t(v,0) = \frac{1}{\bar{\beta}_t}\max_{\mathbf{u},\boldsymbol{\sigma}}\mathbb{E}_{\boldsymbol{\sigma}}\left[\sum_{s=0}^{\infty}\bar{\beta}_{t+s}\left(-\zeta_{t+s}C\left(\mathbf{u}_s\right)\right)\right]$$

subject to (23) and (25). Function $k_t(v,0)$ is a version of the standard cost-minimization problem with commitment. Equation (40) provides a sufficient condition to rule out bunching in that problem. This, in turn implies that the normalized solution to $k_t(1/x)$ must converge to a no-bunching allocation in which each agent reports his type truthfully. The arguments similarly to those used in the previous part then establish that it should be true for all x sufficiently low. ■

Proof of Lemma 10. Suppose constraint (22) is slack in an invariant distribution so that $\chi = 0$. Then $\hat{\beta} = \beta$ and the maximization (24) can be written in its dual form $\min_{\mathbf{u}}\mathbb{E}_{\sigma^{in}}\sum_{t=0}^{\infty}\beta^t C(\mathbf{u}_t)$ subject to (23) and (25). Golosov, Tsyvinski, and Werquin (2013) show in Proposition 6 that the only invariant distribution implied by the policy functions to this problem assigns mass 1 to $\underline{v}$. Such distribution violates (22), a contradiction. Similarly, constraint (22) is slack if all agents play an uninformative strategy. Since $\chi > 0$, by Lemma 7 we have $\lim_{v\longrightarrow\underline{v}}k'(v) = \infty$ when utility is bounded below. Therefore $w_{v,\sigma}(m)$ is interior for all $v > \underline{v}$ by Lemma 13, and (74) becomes (41).

To show the existence of $\underline{w}$ observe that the assumption $1 + \theta_1 - \theta_{|\Theta|} \geq 0$ guarantees $\underline{\varrho} \geq 0$ in Lemma 13. If utility is unbounded below, then Lemma 7 and $\gamma(v) = k'(v)$ give $1 - k'(v) \geq 0$. Then (74) and $\hat{\beta} > \beta$ imply that $1 - k'(\mathbf{w}_v(m, z))$ is bounded away from 0 and, thus, $\mathbf{w}_v(m, z) \geq \underline{w}$ for some finite $\underline{w}$, for all m, z and v. If utility is bounded below (wlog by 0) we show that the invariant distribution can have no mass at any point $v > 0$ with $k'(v) \geq 1$. To see this, suppose $v > 0$ is such that $k'(v) \geq 1$ then, by Lemma 13, $\mathbf{u}_v(m, z) = 0$ and $\mathbf{w}_v(m, z) = \hat{w}_v > 0$ for all m and z where $\hat{w}_v$ satisfies $k'(\hat{w}_v) = \beta k'(v) / \hat{\beta} < k'(v)$. If instead v is such that $k'(v) \leq 1$ then (74) implies $k'(\mathbf{w}_v(m, z)) \leq \beta/\hat{\beta} < 1$ for all m and z. This shows that $\mathbf{w}_v(m, z) \geq \underline{w}$ for some finite $\underline{w}$, for all m, z, and $v > 0$.

It remains to show that $\underline{w}$ is not absorbing. An absorbing point $\underline{w} > \underline{v}$ can satisfy (41) only if $k'(\underline{w}) = 0$. If this this the case then equation (41) implies that $k'(v_t)$ is a negative submartingale for any $v_0 \geq \underline{w}$ and the martingale convergence theorem implies that the unique invariant distribution assigns all mass to $\{\underline{v}\} \cup \{\underline{w}\}$. Observe that if $\kappa(v, \sigma) > \kappa(v, \sigma^{un})$ for any v, σ, then $\{w_{v,\sigma}(m)\}_m$ do not take the same values for all m. Therefore both $\sigma_{\underline{v}}$ and $\sigma_{\underline{w}}$ are uninformative, which contradicts the first part of this lemma. ■