Friedman, Ezra

Working Paper

Risk Sharing and the Dynamics of Inequality

Discussion Paper, No. 1235

Provided in Cooperation with:
Kellogg School of Management - Center for Mathematical Studies in Economics and Management Science, Northwestern University

Suggested Citation: Friedman, Ezra (1998) : Risk Sharing and the Dynamics of Inequality, Discussion Paper, No. 1235, Northwestern University, Kellogg School of Management, Center for Mathematical Studies in Economics and Management Science, Evanston, IL

This Version is available at:
https://hdl.handle.net/10419/221591

Discussion Paper 1235

Risk Sharing And The Dynamics Of Inequality

by

Ezra Friedman

November 13, 1998

http://www.kellogg.nwu.edu/research/math

Risk Sharing and the Dynamics of Inequality

Ezra Friedman[1]

Center for Mathematical Studies of Economics And Management Sciences
Northwestern University

November 13, 1998

Abstract

Risk averse agents who engage in risky production activities frequently participate in some sort of arrangement to share risk. When there are issues of moral hazard, optimal risk sharing typically involves spreading risk over time as well as over space. Agents who suffer bad outcomes can spread risk over time by borrowing against future earnings to supplement present consumption. In this paper I analyze the effect that such intertemporal risk sharing has on the distribution of consumption and utility. I find that in most cases intertemporal risk sharing leads to a spreading of the utility weight distribution. Under the optimal contract agents who suffer negative shocks respond by working harder and bearing more risk, exposing them to the likelihood of more negative shocks.

[1] Math Center, Northwestern University, Room 383, Leverone Hall, 2001 Sheridan Rd, Evanston, IL 60208. email: e-friedman@nwu.edu. I am grateful to Abhijit Banerjee for copious advice and guidance in the writing of this paper . I would also like to thank Bengt Holmstrom and Andrew Atkeson for their comments and advice. Finally, I thank the National Science Foundation for financial support.

Section 1: Introduction

Risk averse agents who have uncertain income streams will seek some way to mitigate the welfare loss caused by this risk. One way to do so is by sharing risk among agents. When individual effort is unobservable, 'pure' risk sharing will generally not be optimal due to problems of moral hazard. Typically, the problem of imperfect risk-sharing due to moral-hazard can be partially mitigated by spreading the risk over time. In other words, rather than punishing an agent with a large one-shot decrease in consumption, it is more efficient to spread the punishment over time. Thus under the optimal dynamic risk-sharing contract, an agent who suffers a bad outcome might receive some immediate insurance at the cost of reduced future consumption.

Spreading risk over time in this way implies that negative shocks will have long term effects; the consumption of an agent in a given period will be affected by all previous shocks the agent has experienced. If the effects of an income shock do not diminish sufficiently over time, and there is not a negative correlation between past and future shocks, the persistence of income shocks might lead to increasing inequality over time.[2] Thus spreading risk over time involves a trade-off between current and future inequality. In fact the optimal contract may decrease the amount of inequality in the present and near future at the cost of increased inequality in the distant future.

Although this paper takes a normative approach to the question of risk sharing and the distribution of income, the results of the paper are of positive interest as well. Work by Udry (1995a,1995b) and Townsend (1994,1995) suggests that real-world risk sharing arrangements which occur in developing countries do share some features of the predictions of the optimal model. Specifically, Udry finds in northern Nigeria a prevalence of informal loan contracts with state dependent repayment schemes. Likewise, Townsend finds in studies of villages in India, Cote D'Ivorie, and Thailand that although he can reject full information perfect risk-sharing, in some villages there is large degree of risk sharing among villagers. Townsend's findings would be consistent with optimal risk sharing under imperfect information.

---

[2] Strictly speaking, this paper makes predictions only about consumption inequality, rather than income inequality. In fact the model used does not even distinguish one agent's income from another's. In practice the changes in the relative positions and indebtedness of agents might be reflected in transfers of land or other durable property which might affect the relative nominal income of the agents. Thus we might expect income inequality, as well as consumption inequality, to result from past income shocks as well.

Deaton and Paxson (1994) provide empirical evidence that inequality within a cohort increases over time. As they note, this is a straightforward prediction of a life cycle model of consumption without borrowing constraints where agents are subject to independent uninsured shocks. This paper relates to Deaton and Paxson in that it shows that even if agents are able to optimally insure against income shocks, as long as they face some moral hazard, inequality will tend to increase over time.

This paper follows more closely in a line of papers concerned with intertemporal risk sharing under asymmetric information. Green (1987) examines an economy in which agents with exponential utility functions insure each other by debt contracts subject to an incentive compatibility constraint. He finds that agents' endowments follow a random walk with negative drift. Thomas and Worrall (1990) generalizes Green's findings, showing that under a range of utility functions an agent's utility tends towards negative infinity. Atkeson and Lucas (1992) examine the evolution of inequality under optimal unemployment insurance with hidden information. In their model, perfect insurance is precluded by the need to prevent agents from falsely claiming to be unemployed. They find that in this model there is no stable distribution of income, in fact, under the utility function they consider, the variance of consumption is strictly increasing over time. Phelan and Townsend (1991) consider a model similar to this one and develop strategies for computationally finding the optimal sharing contracts.

Phelan (1998) considers the role of the assumptions in the findings of much of the previous literature. He considers a multiple agent economy in which aggregate consumption is fixed and agents have exponential utility. He shows that if some agents are perfectly monitored, the utility of all other agents will become arbitrarily low, thus showing that Thomas and Worall's findings do not hinge on decreasing aggregate consumption. On the other hand he shows that there are strictly concave utility functions under which an agent's utility does not become arbitrarily low. Specifically if utility of consumption is bounded away from zero, initial conditions can be found where an agent's consumption is arbitrarily likely to become arbitrarily high with the passage of time.

The papers mentioned above have generally relied on a hidden information model to justify limits on risk sharing. That is they precluded perfect insurance by a need to insure that agents do not falsely report negative shocks. This paper looks at a hidden action model, in which the outcome is observable but is a function of unobservable effort, which is chosen from a continuum of possible values. Allowing agents to choose effort levels allows us to answer the question of how agents' effort levels will respond to negative shocks and whether agents who suffer negative shocks will indeed work harder to catch up.

A lack of convergence in a moral hazard model where agents choose effort from a continuum of values might initially seem surprising. Generally if two agents have the same production possibilities, but one is poorer, the poorer agent would be expected to work harder, so that his income would be greater and he might 'catch up'. There is, however, a simple argument as to why this will not occur under optimal risk sharing: once the poorer agent has caught up he will have the same marginal utility of consumption as the other agents. In the present, since he is poorer, his marginal utility of consumption is greater than that of the other agents. Assuming the absence of liquidity constraints, he will be able to increase his utility by trading away future consumption, which he values as much as other agents, for current consumption, which he values more. As a result, if the poorer agent anticipates catching up he will always wish to trade away future consumption for present consumption. Thus in any optimal contract his expected future position must not be better than his current position.

This is essentially the same intuition that is found in much of the previous literature, and generates their findings that some measure of entitlement or indebtedness will follow a random walk(with the possibility of drift). However, this paper differs from the previous literature in that it considers a model with a small number of agents, so that each agent's effort has a significant effect on aggregate consumption.[3] The effect of an agent's effort on aggregate consumption would provide some incentive for each agent to work even if risk were shared perfectly. If agents are not very risk-averse this effect will be weaker for poor agents(those who have suffered bad shocks) since they receive a smaller share of aggregate consumption. As a result, under the optimal contract, their efforts will be augmented more by the possibilities of future rewards and punishments. For reasons to be explained later, punishments will tend to outweigh rewards over time, so poorer agents will fall farther behind the richer agents.

In contrast to the above literature, several papers present models that result in a stable non-degenerate distribution of wealth or consumption. Banerjee and Newman(1991) obtain this result from a model in which agents choose between engaging in a profitable, but risky project, or investing in a safe diversified portfolio of others' projects. Informed by the results of this paper, one suspects that their finding of a limit to inequality rests on their assumption that rather than caring about the utility of their dynasty, agents care about the bequest they leave, and always wish to leave a positive bequest. Phelan(1994) also

[3] Although Phelan(98) considers a model with a small number of agents, he allows for insurance which negates the relationship between any one agent's realization and aggregate consumption. Similarly, Phelan(94) allows for stochastic variation in aggregate consumption, but assumes this variation is caused by an aggregate shock that is independent of the behavior of the agents.

finds a limiting distribution, but this arises from his use of an overlapping generation model, the inequality within any give cohort is always increasing in his model. In a follow up to their 1992 work, Atkeson and Lucas (1995) show that with a lower bound on continuation utility, there is a stable, non-degenerate distribution of income under optimal insurance. However their result that the lower limit of utility is not absorbing stems from the fact that they place a lower bound on continuation utility, rather than on instantaneous consumption, while their result that the upper bound of continuation utility is not absorbing stems simply from their assumption that agents discount the future more than the principal.

To obtain the result of increasing inequality under optimal risk sharing, I first set up a model of production under moral hazard, and define a risk sharing contract. The first result is that under an optimal risk sharing contract, the state of the economy can be described by a vector which assigns a utility weight to each agent. An optimal contract maximizes the social welfare function given by these utility weights. The contract is recursive in that it maps this old utility weight vector and the realized state of the world into a new utility weight vector which determines both current consumption and expected future utility. Thus, this utility weight vector, which can be interpreted as a measure of wealth or entitlement, forms a sufficient statistic for the effect of history on the optimal contract

In section 3 of this paper I characterize the optimal contract and show how it rewards and punishes agents for outcomes that are suggestive of high or low effort. The punishment or reward takes the form of lowering or raising the agent's utility weight. The condition for the optimal repeated risk sharing contract conforms with the optimal single period risk sharing contract with moral hazard in that punishment and rewards are concentrated in states with high marginal probability ratios.

The general result of increasing inequality is presented Section 4 of this paper. This result takes the form of a proof that under an optimal contract, the expectation of the ratio of the utility weights of any two agents is non-decreasing over time, and it is increasing when both agents are subject to moral hazard. The increasing ratio of utility weights can be interpreted as increasing inequality, although whether or not it corresponds to more conventional measures of increasing inequality depends on the forms of the agents' utility functions. This increasing ratio comes about because each agent's utility weight follows a martingale. This result that each agent's utility weight is a martingale is fundamental to the literature of inequality under risk-sharing and versions of it drive the results of Green (87), Thomas and Worrall (90), Atkeson and Lucas (92) and Phelan(98).

Section 5 of the paper describes the implications of the optimal contract in the special case where one agent is not subject to moral hazard, either because the effort of this agent is observable or because the agent does not take place in production. We think of this

as the principal-agent model with the agent not subject to moral hazard being the principal. The principle finding of this section is that the utility weight of the agent is declining relative to the principal, that is to say the agent expects to receive more punishments in the future than rewards. This is in line with the major result(Proposition 3) of Thomas and Worral(91) which shows that in a hidden information model the utility of the agents tends to negative infinity with probability approaching one.

Section 6 is concerned with describing the conditions under which a model with multiple agents begins to approximate a principal-agent model. This section focuses on the evolution of relative utility weights when one agent faces more moral hazard than another. It is shown that the agent for whom there is a greater moral hazard problem tends to do worse relative to others over time. Furthermore it is shown that for a wide range of parameters and symmetric production functions, under the optimal contract, the moral hazard problem will be more severe for the agents with the lower utility weights, i.e. the poorer agents, than the richer agents. The intuition behind this result is that the richer agent receives a greater share of aggregate production, and thus internalizes more of the effect of his effort on this production. Thus, in expectation, the poorer agents will face more risk and see their utility weight decline relative to the richer agents, and the increasing inequality described in section 4 will be exacerbated.

Section 7 considers the case where utility from consumption is bounded below, so there is a limit to the punishment the social planner can impose. In this case it is possible to have negative utility weights. These negative utility weights correspond to scenarios where an agent is being punished so severely that he is held below his 'efficiency wage' and it is impossible to further punish this agent without reducing his incentive to work in the future. Punishing this agent thus decreases the expected utility of the other agents. In this case a series of short-term contracts would not be able to achieve the constrained optimal social welfare, since the agents would renegotiate whenever the optimal contract called for negative utility weights. Additionally it is shown that in the principal agent model where utility is bounded below, the long run distribution of income is degenerate, and eventually either the agent or the principal will be at the lower bound. This differs from the findings of Atkeson and Lucas(95) which predicts a non-degenerate distribution under bounded utility. The difference in results between these papers derives from the form of the lower bound on utility. Atkeson and Lucas propose a bound on the agent's ability to promise away future utility, will this paper focuses on a lower bound on current utility.

The results of sections 5 and 6, that the agent faces greater moral hazard tends to lose utility weight over time, is derived from the fact that it is more efficient to provide incentive to risk averse agents with material punishments than comparably sized rewards.

The intuition behind this is essentially that it is very socially costly to provide incentives via rewards, because agents are rewarded when their project is successful, and their marginal utility from consumption is low. Hence rewarding them involves transferring large quantities of utility from agents whose marginal utility of consumption is relatively high to agents whose marginal utility is low. Thus, over time, punishments will tend to outweigh rewards for the agents who face the greatest moral hazard. When combined with the result that the poorer agents tend to face more moral hazard this implies that the relative fortune of the poor compared to the rich will tend to decline. This result can be seen as an explanation of the fact that workers everywhere tend to be poorer than capitalists. Even if workers, or agents, are initially better off than capitalists, or principals, this model predicts that over time the punishments imposed on the workers will outweigh the rewards, and wealth will end up in the hands of the capitalists.

Section 2: The Model

I consider an infinite horizon, discrete time, moral hazard model with I individual agents indexed by i. In each period, each agent chooses their own unobservable effort level $e_i \in [0,\infty)$. Assume that agents have exponentially discounted Von-Neumann Morgenstern utility functions, so that $U_{it} = E[V_{it}(c_{it})] - e_{it} + \beta E[U_{it+1}]$, where $V_{it}(c_{it})$ is the instantaneous felicity from consumption. Assume $V_{i_c} > 0$, $V_{i_{cc}} < 0$ as in a standard risk averse utility function, and furthermore that V is thrice differentiable. Their are N possible realizations denoted $\omega \in \Omega$, and each realization is associated with an aggregate resource constraint $x(\omega)$. This is expressed by the condition $\sum_I c_i \leq x(\omega)$. Because this is a model without commitment constraints, and realizations are assumed to be observable, the social planner does not care where the production is coming from, he only cares about the *aggregate* production for each realization and the signal on each agent's effort. Although agents may be producing the resource individually, for the purposes of finding the optimal contract we can treat all production as joint social production. Which agent individually produces more or less is only important insofar as it is a signal about agents' efforts.

The model does not allow for the possibility of savings or aggregate insurance. The production technology is assumed to be stationary and is written as a function P which maps the effort vector **e**, which is not directly observable, into π, a probability distribution over possible states. A particular element in the distribution is referred to as $p(\omega,\mathbf{e})$ which denotes the probability of realization ω given the effort vector **e**. Because **e** is determined in equilibrium, the argument is often suppressed and $p(\omega,\mathbf{e})$ is written $p(\omega)$. I denote the

marginal increase in the probability of state ω with respect to e_i as $p_{e_i}(\omega)$, likewise $p_{e_i e_i}(\omega)$ is the second derivative of the probability of ω with respect to i's effort.

It is assumed that increasing effort always improves the likelihood of a good outcome, specifically

$$\mathbf{e} \geq \mathbf{e}' \Rightarrow \quad \sum p(\omega', \mathbf{e}') \geq \quad \sum p(\omega', \mathbf{e})$$

That is to say that the distribution over resources associated with higher effort by any agent stochastically dominates the distribution associated with lower effort.

In addition the following three assumptions are made to ensure that the optimal contract problem is well-behaved and non-trivial:

(A1) $\forall\, \omega \in \Omega, \forall\, e_i\ p(\omega) > \underline{p}$

(A2) For any effort level e_i, e_{-i}, $\mathbf{p}(e) = \lambda_i(e_i)\mathbf{p}^1(e_{-i}) + (1-\lambda_i(e_i))\,\mathbf{p}^0(e_{-i})$ where λ_i is thrice differentiable with $\lambda_i' > 0$, $\lambda_i'' < 0$

(A3) For any $\varepsilon > 0$, $\exists\ \kappa < \infty$ s.t. if $e_i > \varepsilon$, $\lambda_i'' > -\kappa$

A1 insures that every state is possible at every effort level, and is necessary for existence of an optimal contract. **A2** is taken from Holmstrom(84) and insures that when the first order conditions on an agent's effort choice are met, he is maximizing his utility, so the first order approach is valid[4]. Note that **A2** is a strong condition, implying, among things, that whether a state is a positive or negative signal about an agent's effort does not vary with the agent's expected effort. The last assumption **A3**, simply insures that the returns to effort do not decrease too sharply, so it will not be too easy to hold an agent to the optimal level of effort.

The model is quite general and can apply to a wide range of situations. For example the model could apply to a situation in which there are several agents each of which engage in independent production activities which produce observable output according to a probabilistic production function p_i, where individual output is given by $x_i(\omega_i)$ and the total resource constraint is given by $x(\omega) = \sum_I x_i(\omega_i)$ where the aggregate state is $\omega = \omega^1 \times \omega^2 \ldots \times \omega^I$. It should be pointed out that this is a special case, one in which the signal over agents' outcomes is independent. This refers to the fact that an increase in one agent's

[4] In characterizing the optimal contract, this paper relies on first order conditions. However it has been shown that there are many plausible circumstances under which a first order approach may not lead to a valid characterization of a moral hazard problem when continuous action is possible. The paper uses a condition due to Holmstrom which guarantees that the first order approach is valid. Alternate conditions are possible which are similar to those described by Rogerson[4]. The goal of this paper is not to determine under exactly which conditions a first order approach will be valid in the multiple agent case. Rather we guarantee validity by making a strong assumption about the production technology, and note that our results will hold if these conditions are violated, as long as the first order approach is still valid.

effort will not change the probabilities of getting a good or bad signal regarding another agent's effort. In this case the state ω can be thought of as a vector of signals over individual agents' efforts, leading us to the following definition:

Definition: *The signal over agents' efforts is said to be independent iff* $\omega = \omega^1 \times \omega^2 .. \omega^I$ *and* $p(\omega,e) = \prod_I p(\omega^i)$

A risk sharing contract is a function $F(\omega^t)$ which maps the history of realizations up to time t, $\omega^t = \{ \omega_1 ... \omega_t)$, into a consumption vector $c(\omega^t)$, and an effort vector e_{t+1}. The social planner is permitted to add randomizations to state space, ensuring a convex utility possibility set. The contract is subject to the resource constraint

$$\sum_I c_{it} \leq x(\omega_t) \qquad (1)$$

where $x(\omega_t)$ is the total societal production under realization ω. In addition any contract must satisfy the incentive compatibility constraint that for any history ω^t each agents effort is maximizing his expected utility, which is written:

$$e_{it+1} \in \text{Argmax}_{e_{it+1}} \sum_{\tau=t+1}^{\infty} \beta^\tau (p(\omega^\tau) v(c_i(\omega^\tau)) - e_i(\omega^\tau)) \qquad (2)$$

Our first result is that the optimal contract can be summarized as a recursive map of an expected utility vector U_t and a realization ω_t, into consumption c_t and continuation utility vector U_{t+1}

Lemma 2.1. *The optimal contract can be written as a stationary recursive contract H:* $R^I \times \Omega \rightarrow R^I \times R^I \times R^I$ *where H maps prior utility vector and realized state into an effort vector, consumption vector and continuation utility vector. Furthermore the optimal contract is the solution to the recursive problem:*

$$sup_H \sum_\Omega p(\omega)(\phi \bullet (v(c(\omega)) - e + \beta \phi \bullet z(\omega))),$$

$$s.t.\ (i)\ e_i \in Argmax_{e_i} \sum_\Omega p(\omega)(v(c_i(\omega)) - e_i + \beta z_i(\omega))),\ (ii) \sum_I c_i(\omega) < x(\omega)\ \forall\ \omega,\ (iii)\ z \in Z$$

where Z is the utility possibility set

Lemma 2.1 insures that under the optimal contract, the relevant history can be summarized by an I dimensional vector of continuation utilities which it provides to the participants. However, rather than using the vector of continuation utilities as a state

variable we will use a vector of utility weights.[5] If the utility frontier is strictly convex, for every point on the frontier, there is a unique vector of utility weights such that social welfare is maximized at that point. When this is the case, we can write the contract as G: $R^J \times \Omega \rightarrow R^J \times R^J \times R^J$ so that if $\{c_t, U_{t+1}\}=H(U_t,\omega)$ and $U_t= \text{argmax}_{z\in Z}\ \phi_t \bullet z$, then $\{e_t, c_t, \phi_{t+1}\}=H(\phi,\omega)$, and $U_{t+1} = \text{argmax}_{z\in Z}\ \phi_{t+1} \bullet z$. That is to say ϕ is the inverse slope of the utility frontier at that point, so $G(\phi,\omega)$ is the contract that maximizes $\sum_I \phi_i U_i$. If the continuation utility of the optimal contract is always on the utility frontier then the continuation utility, vector, U_{t+1}, implies a utility weight vector ϕ_{t+1}. Thus the vector ϕ_t is a state variable which completely describes the optimal continuation contract at the beginning of period t. The sum $\sum_i \phi_i U_i$ will be referred to as W, or social welfare, throughout the paper.

It should be noted that the utility possibility frontier need not be strictly convex[6]. In this case the utility weight vector ϕ is not a sufficient statistic for U, the vector of continuation utilities. However, the results of the paper are not predicated on strict convexity. Even if the vector ϕ is not a sufficient statistic for the state of the contract, the results of the paper still hold and we will see, by Lemma 3.1, that the vector ϕ still uniquely describes consumption in any state, although it may not uniquely describe the continuation contract. Using a utility weight vector to describe the evolution of the contract is powerful and convenient because it allows us to characterize the optimal contract without making any assumptions about the actual form of the utility function. Furthermore if we are concerned with inequality, looking at inequality in utility weight, which translates into inequality in the marginal utility from consumption might be more appropriate than simply looking at consumption levels. Inequality in marginal utility is a measure of the marginal efficiency loss to inequality, it tells us how harmful inequality is.

Section 3: Characterization of the Optimal Contract

---

[5] This use of a utility weight vector as a state variable is analogous to the technique of Marcet and Marimon(92) which uses the utility weight of an agent as a state variable in a growth model with commitment constraints.

[6] One situation in which the frontier might not be strictly convex is if the efforts of agents are substitutes. In the case where effort is perfectly substitutable, there might be a linear portion of the utility possibility frontier where the effort of one agent is substituted for that of another at a constant proportion. If this linear portion is long enough, it may be possible to provide the agents with sufficient incentives by just moving along the linear portion, thus holding their utility weights constant. In this case, first best is achieved. On the other hand, it is possible to show that as long as the returns to effort drop off quickly enough the first best utility set is strictly convex. An example that satisfies this is if production consists of individual projects whose returns to efforts decline exponentially. In this case it is possible to show that the first best will not be achieved.

This section begins by showing that the outcome of the optimal contract after the state has been realized can be described by only a utility weight vector ϕ_{t+1} and a societal resource constraint $x(\omega)$. The consumption of individual agents is implicitly determined by ϕ_{t+1} and $x(\omega)$, and the continuation contract is described by ϕ alone.

Lemma 3.1: *If posterior utility is always on the utility frontier in the optimal contract, and* c_i *and* c_j *are above their lower bound, then* $\frac{V_{it_c}(\omega)}{V_{jt_c}(\omega)} = \frac{\phi_{jt+1}}{\phi_{it+1}}$.

Proof: See Appendix

Understanding lemma 3.1 provides the basis for understanding why income or wealth levels will not converge. The central idea of the proof is that if the ratio between two agents' marginal utilities from consumption differs from the marginal rate of transformation of their continuation utilities, Pareto improving trade is possible. Since each agent's utility from consumption is concave, consumption is determined uniquely by the resource constraint and the ratios of agents' marginal utilities of consumption. Thus, when the utility possibility set is strictly convex the optimal risk sharing contract is completely described by a function G: $\mathbf{R}^I \times \Omega \rightarrow \mathbf{R}^I$ where $\phi_{t+1} = G(\phi_t,\omega)$ so G maps prior utility weight and realized state into posterior utility weights which in turn determine the future of the contract.

Before proceeding in describing the optimal contract, the following definitions are useful in making the expression of the contract less cumbersome. We will define W_{e_i} as the marginal effect agent i's effort has on social welfare given weights ϕ. The derivation of W_{e_i} involves considering not only the direct effect of i's effort on production, but also considers the fact that a change in i's effort will also change the effort levels of the other agents, which will in turn affect the welfare of all agents. The algebra and notation necessary for an explicit formulation of W_{e_i} is not terribly enlightening and is relegated to the appendix.

In order to make our expressions of the optimal contract more compact we introduce some more shorthand notation. We define $U_{i_{e_j}}$ as the marginal effect of j's effort on i's utility under the optimal contract. Likewise We define $U_{i_{e_j e_k}}$ as the partial of k's of $U_{i_{e_j}}$ with respect to k's effort.

We will define s_i by:

$$s_i = \frac{-1}{U_{ie_ie_i}}$$

The symbol s_i represents the sensitivity of i's effort to incentives. Since i chooses effort to maximize his utility, at the equilibrium effort choice $U_{e_i} = 0$. Thus the direct change in i's effort with respect to a change in i's incentives is given by s_i which can be thought of as the inverse curvature of the return to effort. When s_i is low so the returns to effort are relatively linear, increasing i's incentives will move him further up along the effort curve than it will when s_i is high and the payoffs to increasing effort are rapidly diminishing. Having set up the notation, it is now possible to express the effect of social welfare of raising any agent's consumption in a particular state.

Proposition 3.2: *The partial derivative of social welfare with respect to $c_i(\omega)$ is given by:*

$$W_{c_i}(\omega) = \phi_i p(\omega) V_{i_c}(\omega) + W_{e_i} s_i p_{e_i}(\omega) V_{i_c}(\omega)$$

Proof: See Appendix

The partial derivative of social welfare, $W_{c_i}(\omega)$ consists of two parts, the first part is the direct effect on welfare through agent i's utility from consumption. As long as utility weights are positive, this component is clearly strictly positive since every agent's utility is increasing in consumption. The second component is the effect of changing i's incentives on everybody else's utility. If $p_{e_i}(\omega) > 0$, that is if increasing i's effort leads to a state becoming more likely, increasing the utility i receives in that state will increase i's effort. Assuming that there is some risk-sharing occurring, increasing i's effort increases everybody else's utility (i.e. $W_{e_i} > 0$) and the second term will be positive for states where p_{e_i} is positive, that is for states which are suggestive of high effort on i's part.

Knowing $W_{c_i}(\omega)$, the welfare consequences of increasing an agent's consumption as a function of the state, enables us to produce the following conditions which must hold in an optimal contract.

Proposition 3.3: *In any state where the resource constraint is binding, the posterior utility weight vector is characterized by the equation:*

$$\frac{\phi_{jt+1}(\omega)}{\phi_{it+1}(\omega)} = \frac{\phi_{jt} + W_{e_j} \sigma_j s_j}{\phi_{it} + W_{e_i} \sigma_i s_i}$$

Proof: See Appendix.

Proposition 3.3 is derived by equalizing the partial derivative of social welfare with regard to each agent's consumption in each state so that for any realization ω, $W_{c_i}(\omega) = W_{c_j}(\omega)$. Thus social welfare is maximized subject to the resource constraint. The result is a general condition which applies to any risk-sharing contract with moral hazard. p_{e_j} and p_{e_i} correspond to the extent to which rewarding i or j will increase their incentives. Thus when p_{e_j} goes up or p_{e_i} goes down, the ratio $\frac{\phi_{it+1}}{\phi_{it+1}}$ increases, so relative to i, j receives more utility in that state. In general when risk sharing is occurring we expect W_{e_i}, the impact of any agent i's effort on social welfare to be positive. Lastly, note that ϕ_{it} and ϕ_{jt}, the prior utility weights, are both multiplied by $p(\omega)$. This implies that when p_{e_i} or p_{e_j} is large in magnitude compared to $p(\omega)$ we expect to see a large difference between the posterior utility weights and the priors, so that one agent is being substantially rewarded or punished. This corresponds to the general mechanism design result that incentives are concentrated in states where the marginal probability is high relative to the absolute probability, and is similar to the findings in Rogerson(1985a) concerning the repeated principal-agent problem.

When proposition 3.3 is satisfied, a local optimum in the social planner's problem is implied. Our assumption **A2** insures the continuity of agent's responses to the optimal contract, and hence insures that the optimal contract will occur at a local optimum. However, we have not shown the converse, that is to say we have not shown that any local maximum is a global maximum, thus the condition in proposition 3.3 must be seen as a necessary but not a sufficient condition for the optimal contract.

Looking back at proposition 3.3, we might ask if there can be a state which sends such a negative signal about an agent's effort that under the optimal contract, the social benefits from punishing an individual in a state outweigh the private cost to him of being punished. In this case $W_{c_i}(\omega)$, the partial of social welfare with respect to his consumption might be negative, implying that the ratio in proposition 3.3 could be negative. However we will now show that if utility is unbounded below, as in a CRRA utility function with $\gamma \geq 1$, $W_{c_i}(\omega)_i$ will never be negative in an optimal contract. The argument is, if utility is unbounded below and $W_{c_i}(\omega)_i$ is negative, social welfare can be improved by taking utility away from agent i. This will increase i's effort, and lessen the moral hazard problem. Since utility is unbounded below, social welfare can always be increased by taking away consumption and utility until $W_{c_i}(\omega) = 0$. If $W_{c_i}(\omega) = 0$, and there is another agent j such that $W_{c_j}(\omega)$ is positive, social welfare can be improved by transferring consumption from i to j. Thus, if utility is unbounded below, either $W_{c_i}(\omega) > 0$ for all i, and the resource constraint is binding, or $W_{c_i}(\omega) = 0$ for all i, and the resource constraint is not binding. Note that $W_{c_i}(\omega)$ will be zero only in states where every agent is being simultaneously

severely punished. As is shown below, under some general conditions about the independence of signal space, it is always possible to punish agents severely enough in states where another agent is not being punished. Thus the ratio in proposition 3.3 will always be strictly positive and the resource constraint will always be binding.

Proposition 3.4 *Suppose* $\exists$ *a partition of* I $\{I_1,...,I_M\}$ *such that* $\Omega = \Omega^1 \times .. \Omega^M$ *so that* $p(\omega) = \prod_M p(\omega^m)$ *and* $p_{e_i}(\omega^m) \equiv 0 \ \forall i \notin I_m$ *If utility is unbounded below, then under an optimal contract* $W_{c_i}(\omega)$ *is always positive and the resource constraint will be binding in every state.*

Proof: See Appendix

According to proposition 3.4, if social production can be broken into projects performed by disjoint teams, the resource constraint will always be binding. Note that this is a slightly weaker condition than the condition for independence of signals presented in section 2. Proposition 3.4 holds because it will always be possible to provide sufficient incentives to a team by punishing them very severely if their project fails while another team's project succeeds. As an aside, it is worth noting that if utility is unbounded below we have a condition for the maximum value of W_{e_i}. By the non-negativity of $W_{c_i}(\omega)$, we know that:

$$W_{e_i} \leq \text{Min}_{(\Omega | p_{e_i} < 0)} \frac{-\phi_{it}\, p(\omega)}{p_{e_i} s_i}$$

This places an upper bound on the marginal moral hazard any agent faces. This upper bound is increasing in his utility weight and decreasing in the marginal informativeness of the state which represents the worst signal about his effort.

Having characterized the optimal contract, and shown how the current utility weights depend on the past utility weight and the signal over the agents' efforts we are now ready to approach the question of the evolution of utility weight shares. We are now turn to the question of what our findings say about the evolution of inequality.

Section 4: Increasing Inequality

This section uses the results of the previous section to show that the expectation of the ratio of utility weights of any two agents is increasing if the resource constraint is always binding and the signals over the agents' efforts satisfy some weak conditions about

independence. In other words as long as the posterior utility weights will always be positive, the utility weights of any two agents are expected to spread. I interpret this as an increase in expected inequality, and show how the effects of this increasing inequality on income and utility from consumption depend on the form of the utility function chosen.

Proposition 4.1: *If the resource constraint is always binding, and the signals over agents' action are independent, the expectation of the ratio of the utility weights of any two agents is increasing over time. (i.e.* $E\left(\frac{\phi_{jt+1}}{\phi_{it+1}}\right) \geq \frac{\phi_{jt}}{\phi_{it}}$ *, and if* $W_{e_i} \neq 0$, $E\left(\frac{\phi_{jt+1}}{\phi_{it+1}}\right) > \frac{\phi_{jt}}{\phi_{it}}$ *)*

Proof: See Appendix

Note that the inequality is strict whenever $W_{e_i} \neq 0$, in other words, as long as the social planner has some incentive to punish or reward i by raising or lowering his relative utility weight, utility weights will spread.

Although the proof of proposition 4.1 assumes that the signals about the agents' efforts $\frac{p_{e_i}}{p(\omega)}$ and $\frac{p_{e_j}}{p(\omega)}$ are independent, by second order approximation we find that independence is not necessary. In fact a second order approximation of the necessary condition is that:

$$\frac{\phi_{jt}}{\phi_{it}} \text{Var}\left(s_i W_{e_i} \frac{p_{e_i}}{p(\omega)}\right) > \text{Cov}\left(W_{e_i} \frac{p_{e_i}}{p(\omega)}, W_{e_j} \frac{p_{e_j}}{p(\omega)}\right).$$

This condition is violated only if the covariance of $\ln \phi_{it+1}$ and $\ln \phi_{jt+1}$ is greater than the variance of $\ln \phi_{it+1}$. This is roughly equivalent to the statement that conditional on a good outcome and an increase in ϕ_{it+1}, the expectation of the proportional increase in ϕ_{jt+1} is greater. When i is rewarded he expects that j will be rewarded more, and when he is punished he expects that j will be punished further. Thus increasing i's effort on average actually worsens his position relative to j. This would occur only if the signals of i's effort and j's effort were closely correlated and j faced more moral hazard. Absent such an inter-relationship of signals, the utility weights of i and j will always be spreading.

Proposition 4.1 simply shows that the expectation of the relative utility weights of agent j compared to agent i is increasing, and should not be taken to mean that agent j is doing progressively better in relation to agent i. By symmetry, the expectation of the

relative weight of agent i's utility relative to agent j's is also increasing. Rather, proposition 4.1 implies a spreading of the marginal utilities.

It can be seen that if all else is held equal an increase in W_{e_i}, the marginal social benefit of i's effort, will lead to an increase in the expectation of $\frac{\phi_{jt+1}}{\phi_{it+1}}$. The marginal social benefit of i's effort can be considered a measure of how much risk is shared, thus the more risk is shared, the more quickly the utility weights diverge.

What the result of Lemma 4.1 tells us about the behavior of more conventional measures of inequality, such as inequality in consumption levels depends on the utility function chosen. If utility is given by a CRRA function with $\gamma=1$, i.e. log utility, the ratio of any two agents' consumption will be the same as the ratio of their utility weights. It is thus trivial to show that the expectation of the ratio of the richest agent's consumption to that of the poorest agent is increasing.[7] For any CRRA function with $\gamma<1$ the ratio of consumption is concave in the ratio of utility weights, and is also clearly increasing. If utility is CRRA with $\gamma > 1$ the ratio of consumption is convex in utility weights, and it is no longer possible to state unequivocally that the consumption ratio is increasing. Note that γ, the coefficient of risk aversion, could be considered a measure of how odious inequality is. It is not surprising that in an optimal contract, when γ is high, there will be less inequality in consumption.[8]

The final question this section asks is what will happen to inequality in the long run. As long as W_{e_j} is not zero for every agent with probability 1, there will be some j such that expectation of $\frac{\phi_i}{\phi_j}$ will be rising. Thus if we measure inequality as the sum of the ratios of the utility weights between any two agents($\sum_i \sum_{j \neq i} \frac{\phi_i}{\phi_j}$), inequality is strictly increasing. Furthermore this measure of inequality will not reach an asymptote as long as every agent's utility weight is positive. So as long as utility weights always vary[9], inequality is not bounded in the limit.

---

[7]Suppose i is the richest agent and j is the poorest at time t. At time t+1, $E(\frac{c_{it+1}}{c_{jt+1}}) > \frac{c_{it}}{c_{jt}}$. Since the richest agent as at least as rich as i, and the poorest at least as poor as j, the expectation of the consumption ratio is increasing.

[8] It is questionable whether a decrease in the expectation of the ratio of consumption between the rich and the poor should always be interpreted as decreasing inequality. Differences in need might be more worrisome than differences in consumption. A natural interpretation of marginal utility from consumption is need, thus we can still say that we are always expecting the neediest to become more needy compared to the rich.

[9] This is almost the same as saying the program is first best. A stable distribution would only be optimal when the social planner is randomizing between two effort profiles that are locally first-best when randomization is not permitted. However it is possible that there exists a third globally first-best effort profile, in which case this randomization would not be first best.

This result can be seen as a bare bones in that it implies that if there is any chance of a poorer agent catching up to a richer agent, the poorer agent must also run the risk of falling further behind. The finding in this paper that utility weight follows a martingale, coincides with the findings in Green(1987) and Thomas and Worrall (1990), that the inverse of marginal utility of the agent will follow a martingale. Since the utility weight of the poorer agent and the richer agent are both following martingales, in some sense neither agent is expecting to catch up or fall further behind.

However, measuring inequality by the expectation of utility weight ratios puts a great deal of weight on extreme inequality. As long as there can be some probability that inequality is arbitrarily high, the probability that inequality is low may approach 1. In the following two sections I develop intuition for determining when other, stronger measures of inequality will be increasing. In particular, I show that if the agents are not especially risk averse, under ex-ante symmetrical functions, the "martingale" driven inequality discussed in this section will be exacerbated by the fact that in some sense, the poorer agent expects to do strictly worse relative to the richer agent in the future. In this case , I am able to argue that there will be no stable non-degenerate distribution of utility weight. On the other hand if agents are very risk averse, there is some sense in which poor agents expect their relative standing to improve relative to richer agents. If this is the case there may be a non-degenerate limiting distribution of utility weights, but I argue that due to the extreme risk aversion, every agent's expected utility is negative infinity in the limit, and the detrimental consequences of inequality are in any case unbounded.

Section 5: The Principal- Agent Problem

The implications of the above results when applied to the principal-agent problem are interesting in themselves and also provide guidance as to what to expect when agents are subject to differing degrees of moral hazard. The term principal is meant to describe an agent who faces no moral hazard. This could occur either because his effort is either perfectly observable, or because it does not enter into the production function. Contrary to much of the previous literature, the principal is not assumed to be risk-neutral, rather he is assumed to have a utility function similar to that of the other agents. Because the principal can be seen as the limiting case of an agent that is subject to little moral hazard, the results of this section give us the limit of results when agents bear moral hazard asymmetrically.

---

values of Φ e_i =0 could be optimal, and hence there would not be a moral hazard problem.

The first result is that if utility weights are restricted to lie on the unit simplex, the expectation of the principal's utility weight is always growing. Take note of the result that if i is a principal and bears no moral hazard $E[\frac{\phi_{kt+1}}{\phi_{it+1}}] = \frac{\phi_{kt}}{\phi_{it}}$. Thus the utility weight ratio between an agent and the principal is a martingale. It should be noted that there is a real asymmetry here. Specifically $E(\frac{\phi_{it+1}}{\phi_{kt+1}}) > \frac{\phi_{it}}{\phi_{kt}}$, so it is fair to say that the position of the principal relative to the agent is expected to improve over time.

Proposition 5.1: *If i is a principal and faces no moral hazard, and ∃ some agent j that does face moral hazard, then* $E[\frac{\phi_{it+1}}{\sum_k \phi_{kt+1}}] > \frac{\phi_{it+1}}{\sum_k \phi_{kt+1}}$

Proof: See Appendix

Corollary: *If utility is CRRA with $\gamma=1$, the principal's share of consumption is increasing.*
Proof: if $\gamma = 1$ consumption share is proportional to utility weight share.

The intuition as to why the agent expects to do worse in the future, and the principal better is somewhat subtle. The result stems from the fact that it is more efficient to provide incentives with punishments than rewards. This is because in states where an agent is being rewarded his marginal utility of consumption is low compared to the principal, therefore it is necessary to take a large amount of utility away from the principal in order to provide him with a moderate reward. In states where he is being punished, it is only necessary to take a small amount of consumption away from him to provide incentive to avoid such states. Since the principal only gets a small amount of utility from this, social welfare is still being lost, but the loss is now smaller, because less ex-ante social welfare is being transferred. Since punishment is more socially efficient, the optimal contract will use more punishment and less reward. Since it has already been established that rewards and punishments will take the form of lowering or raising the agent's utility weight in the social welfare function, we'd expect the utility weights of agents who bear moral hazard to decrease over time, since the punishments will tend to outweigh the rewards.

A simple example can clarify the above intuition. The marginal welfare loss from incentivization is the difference between in the weighted marginal utility of the agent and the weighted marginal utility of the principal. Consider a situation where the ex-ante utility weight of the principal and the agent is identical and the agent is being rewarded. Suppose

his marginal utility of consumption is .5, while that of the principal is 1. To transfer 1 util, two units of consumption would have to be transferred causing a welfare loss of one util. If the agent was being punished and his marginal utility was 2 while the principal's was 1, in order to take away 1 util we'd have to transfer one half unit of consumption, and the welfare loss would be only one half of a util. Consequently, if both states were equally likely it would be optimal to reward less and punish more.[10]

The increasing utility weight of the principal does not directly translate into increasing utility for the principal, likewise the decreasing utility weight of the agent does not necessarily imply decreasing expected utility for the agent. However we find that under a wide range of parameters, decreasing utility weight will be accompanied by decreasing utility. In fact, is shown in the appendix, that abstracting from changes in effort and production, a principal's utility from consumption will always be increasing if utility is given by a CRRA function with $\gamma>1$. By the convexity of the utility possibility set the agent's utility must be decreasing under these conditions. In fact it is possible to show that if the utility function is CRRA with $\gamma \geq \frac{1}{2}$ expected utility from consumption is always decreasing for the agent in a principal-agent model. Even when $\gamma < \frac{1}{2}$, the agent's utility will be expected to increase only when his weight is low compared to the principal's.[11]

The central intuition of this section, that the principal's position relative to the agent is improving over time because the agent is being punished more often than rewarded, is interesting in itself, but is also applicable to the question of the dynamics of inequality. If we can identify which agents more closely resemble the principal we can expect that in the long run utility weight will tend to migrate towards them.

Section 6: The evolution of agents' relative welfare.

The findings of the previous section suggest how utility weights and distribution of consumption might evolve when agents face different amounts of moral hazard. Specifically agents who face more moral hazard should on average suffer more punishment than reward. Consequently utility weight should tend to move towards the agents that more

---

[10] Thomas and Worral provide an almost identical intuitive argument in their 1990 paper.

[11] As γ goes to zero j's utility takes the shape of an S-curve when plotted against $\frac{\theta_j}{\theta_i}$, with the steep portion of the curve occurring near where $\frac{\theta_j}{\theta_i}=1$. If $\frac{\theta_j}{\theta_i}<1$, j is on the lower flat part of the S -curve, where it is concave, thus j's utility is increasing with mean-preserving variance in $\frac{\theta_j}{\theta_i}$.

closely resemble a principal and face less moral hazard. With this motivation we turn to the question of what it means for one agent to face more moral hazard than another agent, and describe the conditions under which an agent would expect his utility weight to increase or decrease relative to another agent. We are specifically interested in determining whether agents who have received good realizations and have higher utility (rich agents) will face more or less moral hazard under the optimal contract than agents who have received bad realizations (poor agents). This is of central interest to the paper, because it tells us when there are forces exacerbating or mitigating the dispersion of utility weights and the increase of inequality.

I will now construct a variable I call utility weight risk. I define Z_i as the proportional variance in i's utility weight under the optimal contract. Thus Z_i is defined by:

$$Z_i = \frac{Var(\sigma_i)\ (W_{e_i} s_i)^2}{\phi_i^2} .$$

It should be first noted that this variable is a function of the optimal contract, rather than a direct function of the parameters. It is a measure of how much an agents utility weight is likely to change from the current period to the next. It can be interpreted as moral hazard for the following reason, the more an agent's incentives differ from those of the social planner, the more reason there is for the social planner to modify these incentives by varying the agent's future utility weight as a function of outcome. If an agent's effort were perfectly observable, or if his incentives were perfectly aligned with society's there would be no reason for the planner to vary the agent's weight. The proportional variance in the agents utility weight can be thought of as the marginal cost of providing incentives to this agent under the optimal contract. Since under the optimal contract the marginal social cost of providing incentives is equal to the marginal social benefit, Z_i a measure of how much this agent's interests differ from the planner's, and can thus be seen as a measure of moral hazard.

The measure chosen has the quality that, by looking at a second order approximation, the condition $E[\ln\frac{\phi_{jt+1}}{\phi_{it+1}}] > \ln(\frac{\phi_{jt}}{\phi_{it}})$ is approximately equivalent to the condition $Z_j < Z_i$. Comparing the logs of the prior and posterior utility ratios is an appealing condition for which direction utility weight is being transferred because it is symmetric, in other words $\ln\frac{\phi_{jt}}{\phi_{it}} = -\ln\frac{\phi_{it}}{\phi_{jt}}$, so if $E[\ln\frac{\phi_{jt+1}}{\phi_{it+1}}] > \ln(\frac{\phi_{jt}}{\phi_{it}})$ then $E[\ln\frac{\phi_{it+1}}{\phi_{jt+1}}] < \ln(\frac{\phi_{it}}{\phi_{jt}})$. Thus by comparing the prior and posterior utility weight ratios we can get a definite

answer to the question of which agent expects to improve his relative position in the future. If $Z_j < Z_i$, it is appropriate to say that j expects to improve his utility weight relative to i in the future.

We now turn to the question of whether utility weight risk(Z) is likely to be higher for the poorer or richer agent. We can provide intuition as to why the utility weight risk of a rich agent approaches zero as all utility weight becomes concentrated in the rich agent. As the utility weight of the other agents approaches zero, the social planner cares less about the other agents, and has less incentive to distort the incentives of the one agent he does care about to improve matters for the other agents. The rich agent becomes in effect, a virtual principal, since his utility forms the bulk of social welfare, he has effectively internalized the effects of his effort on social welfare. In fact it is shown that as $\Sigma_{i\neq\varphi}\ \phi_j \rightarrow 0$ for any $j\neq i$, $Z_i \rightarrow 0$.

Our first intuition might be that as an agent's utility weight drops, he internalizes a smaller proportion of social welfare, and should accordingly be rewarded or punished more. A parallel intuition would be that as an agent's utility weight drops, the social planner would like him to work harder, and as a result must increase the incentives on that agent. However there is another countervailing effect: when risk aversion is high, a poorer agent will care more about the aggregate consumption than the rich agent (even though he gets a smaller share). In this case, concern about aggregate consumption alone causes the poor agent to work harder than the richer agent, and since the returns to effort are decreasing, the social planner has less reason to increase the incentives of the poor agent. When risk aversion is very high, the second effect could dominate, and Z might actually decrease as agents get poorer.

In our attempt to determine more precisely when Z will be higher for the poorer agent, we start by taking a closer look at the incentives of any agent. The incentive that an agent faces can be divided into two sources, the incentive from the effects of his effort on the total quantity of resources available, and that from the effect of his effort on his own share utility weight. These will be referred to as size of the pie incentives, and share of the pie incentives respectively. The relationship between an agent's utility weight, and the strength of the incentive he derives from concern over the size of the pie depends on γ, the coefficient of relative risk aversion. If $\gamma=1$, changing the size of the pie, but holding each agent's share constant will have the same effect on each agent's utility. However if $\gamma > 1$, this incentive will be stronger for the poorer agent and if $\gamma < 1$, it will be stronger for the rich agent. The effect of proportionate change in utility weight, or share in the resource is always greater for the poorer agent because the slope of the utility frontier is given by the

inverse ratio of utility weights. Thus the incentive from a change in an agent's share of utility weight will be inversely proportional to the agents utility weight. That is to say that if lowering ϕ_i or increasing ϕ_j by one percent changes U_i by one util, the effect on U_j will be $\frac{\phi_i}{\phi_j}$ utils.

From this we provide a simple argument that if $\gamma \geq 1$ the poorer agent should work harder under the optimal contract. Since the poorer agent receives as much incentives from the size of the pie, if her utility weight proportionally varies at least as much as the richer agent's (i.e. $Z_j \geq Z_i$) she will receive more incentive from the share of the pie, and thus will work harder. But if the poor agent is not working harder than the rich agent, the moral hazard the poorer agent should face is greater than the richer agents so the poorer agent should work harder. ($e_j < e_i \Rightarrow Z_j > Z_i \Rightarrow e_j > e_i$).

Now we turn our attention to the more difficult question of for whom Z will be higher. In order to express sufficient conditions for our result we first must develop some notation. Recall that if we assume independence the probability over states is the cross product of the probabilities over each individual's realization, and that by **A2** the probability distribution over an individual's realization is given $\lambda_i(e_i)\, p_i^{\,1} + 1 - \lambda_i(e_i) p_i^{\,0}$. We define V_{xy} as the expected utility of consumption of an agent who consumes the aggregate resource given that the probability distribution over states is $p_i^{\,x} \times p_i^{\,y}$. We define e_i^* as the i's first best level of effort as $\frac{\phi_i}{\phi_j} \to \infty$. With these definitions we are able to state the following result about utility weight risk in the limit where one agent's utility weight goes to zero.

Proposition 6.1: *Suppose the economy consists of two agents with independent signals over effort and the following three conditions are met.*

(a) Utility is CRRA with $\gamma < 2$

(b) Returns to effort are 0 for some finite effort ($\forall i$, $\exists \bar{e}_i < \infty$ *s.t.* $\lambda_i'(\bar{e}_i) = 0$)

(c) $\lambda(e_i^*)(W_{11} - W_{10}) + (1 - \lambda(e_i^*))(W_{01} - W_{00}) > |W_{11} + W_{00} - W_{01} - W_{00}|$

Then Lim $_{\phi_j \to 0} \frac{Z_i}{Z_j} < 1$.

As long as the above conditions are met, the fact that the social planner cares much more about the effect of the poor agent's (j) effort on the rich agent(i) then the reverse is enough to show that the planner will augment the poorer agent's incentives with more utility weight risk. In this case the utility weight of the poorer agent will be declining

relative to the richer agent ($E\left[\ln\frac{\phi_{jt+1}}{\phi_{it+1}}\right] < \ln\left(\frac{\phi_{jt}}{\phi_{it}}\right)$ and the random walk spread of inequality described in section 4.1 will be exacerbated.

Conditions (a) and (b) are self explanatory, and condition (c) requires that the effect of j's success on i's incentives be smaller than the direct effect of j's success on i's utility. Since the first best effort levels are easy to find, condition (c) is easy to check. Conditions (b) and (c) are used in the proof only to rule out forms of the optimal contract that are intuitively improbable, and should not be thought of as necessary conditions. Specifically condition (b) is used to rule out an arbitrarily large increase in j's effort resulting from an arbitrarily small proportionate decrease in j's utility weight, and condition (c) rules out the possibility that $W_{e_j}<0$, that is that j is punished when he succeeds. Because it is difficult to imagine why the optimal contract would ever have these forms, one would expect the result to hold even when b) and c) are violated. In fact the poorer agent is more likely to face more utility weight risk if b) is violated: if returns to effort declined exponentially, which would violate b), we could show that for any $\gamma < 3$, $\text{Lim}_{\phi_j \to 0} \frac{Z_i}{Z_j} > 1$. Thus if the optimal contract is sufficiently smooth and well behaved conditions b) and c) are not needed.

However the first condition is likely to be necessary: as γ increases there are two effects which cause the poorer agent to face *less* utility weight risk than the richer agent. First, as γ increases, the effect of the rich agent i's effort on j increases, causing the social planner to augment i's incentives more. Second, as γ increases, j's effort will increase and the marginal effect of his effort will decrease, causing the planner to augment his incentives less. If (b) is satisfied[12] and $\gamma > 2$ these two effects will dominate, and it will be the richer agent who faces more relative risk than the poorer agent, and the spread of inequality will be mitigated by the fact that the richer agent expects to be punished more than the poorer agent.

Outside of the limiting cases, it is more difficult to determine whether the poorer agent faces more or less utility weight risk than the richer agent. It is possible to construct production functions where the magnitude of s_i is decreasing rapidly in e_i, that is where diminishing returns to effort set in suddenly. Therefore as an agent's effort increases it is easier to hold him close to the optimal effort, and he does not need to bear as much extra risk. Even when we assume a specific production function, so that this is not the case, because we do not have a closed form solution for the program, It is difficult to make unequivocal claims about which agent faces more moral hazard. However if we abstract away from effects which we expect to be small we are able to make a strong argument about conditions under which we expect the poorer agent to fall further behind.

---

[12] If the returns to effort decrease exponentially, so that it is not quite so easy to hold the poor agent near his optimal effort level, than these effects will not dominate until $\gamma>3$.

Recall that by **A2**, the returns to any one agent's effort can be summarized by the derivative of the one dimensional variable $\lambda_i(e_i)$. If we assume that rewards to effort decline exponentially it will be the case that s_i does not depend on effort and is constant across individuals[13]. Thus it is only necessary to compare $\frac{Var(\sigma_i)w_{e_i}^2}{\phi_i^2}$ with $\frac{Var(\sigma_j)w_{e_j}^2}{\phi_j^2}$ We will now make two simplifying assumptions about the optimal contract, namely (1) $Var(\sigma_i) = K(\lambda_i')^2 \; \forall$ i and (2) $a_{ij} = a_{ji} = 0$. The first of these assumptions is that the likelihood of the most informative realizations are the same for both agents. The second is that the effort of one agent does not affect the incentives of the other agent, so we can ignore the cross effects of each agent's effort. Neither of these assumptions are generally true, but we argue in the appendix that our result about when the poorer agent faces more moral hazard is not critically sensitive to the first assumption, and the violations of the second assumption is likely to be violated in a way which strengthens rather than weaken the result. Under the simplifying assumptions, in an ex-ante symmetrical model where utilities are CRRA with coefficient $\gamma < 2$, $Z_i > Z_j$ whenever $\phi_j < \phi_i$. That is to say the poorer agent will always fall further behind the richer agent in expectation, and the random walk growth in inequality will always be exacerbated. The derivation of this result is not terribly enlightening and is relegated to the appendix.

It turns out that the behavior in the limit described in proposition 6.1 is enough to give us a much stronger inequality result then that presented in section 4. Specifically, as long as the poorer agent faces more moral hazard in the limit, it is possible to show that there will be no non-degenerate distribution of utility weight, that is to say with probability approaching one, all the utility weight will be concentrated on one agent. As long as in the limit where $\phi_j \to 0$ $Z_j > Z_i$, there will be some ε such that if $\phi_j < \varepsilon$, $Z_i < Z_j$. We can construct a strictly monotone transformation of our utility measure such that $q \sim \ln\left(\frac{\phi_i}{\phi_j}\right)$ if $\phi_i < \varepsilon$ or $\phi_j < \varepsilon$,

---

[13] Recall $s_i = \frac{-1}{U_{ie_ie_i}}$. By the assumption of exponential decreasing returns to effort $\frac{\lambda_i'}{\lambda_i''} = \frac{p_{e_i}}{p_{e_ie_i}}$ and is constant. By incentive compatibility $\int_\Omega p_{e_i}(\omega)U_i(\omega)\,d\omega = 1$ and $U_{ie_ie_i} = \int_\Omega p_{e_ie_i}(\omega)U_i(\omega)\,d\omega$, Thus if $\frac{p_{e_i}}{p_{e_ie_i}}$ is constant, $U_{ie_ie_i} = \frac{p_{e_i}}{p_{e_ie_i}}$ and is constant.

$q \sim \frac{\phi_i}{\phi_j}$ if $\phi_i > \phi_j > \varepsilon$ and $q \sim -\frac{\phi_j}{\phi_i}$ if $\phi_j > \phi_i > \varepsilon$[14]. Whenever $q_t < 0$, q is a supermartingale, and when $q_t > 0$, q is a submartingale. Since a martingale will not converge to a non-degenerate distribution if $\lim_{q \to \infty} \text{Var}(z) < \infty$, and q diverges more quickly than a martingale, there will be no non-degenerate limiting distribution of q or ϕ. In this case the likelihood that the poorer agent's utility is arbitrarily small will be arbitrarily close to one as t goes to infinity. So as long as agents are not very risk averse, the limiting distribution of utility weight or consumption share will be degenerate, and society will become arbitrarily unequal for any measure of inequality.

If agents are very risk averse, so that the conditions for Proposition 6.1 do not hold, there may be a non-degenerate distribution of utility weight. Since the variance in utility weight approaches zero in the tails of the distribution, it is possible to approximate the evolution of utility weight there by Brownian motion with a drift. This allows us to characterize the tails of the limiting distribution when it exists. When we characterize this distribution under the assumption that the poorer agent faces no moral hazard, which is the best possible case for equality, we show that the likelihood that the agents is in tails is finite, so can be a non-degenerate limiting distribution[15]. However we show that for any CRRA utility function the utility from the tails goes to negative infinity, so that even if inequality is bounded by some measure, the negative consequences of inequality are not bounded. This occurs because the variance of utility weight for both agents in the tails is very low, so once the utility weight distribution becomes very unequal, even though it might tend to become more equal in the future it will do so very slowly. As a result, in the long run the economy is very likely to be in a state with high inequality, where one agent is being punished very severely.

The results of this section give us some hint as to how the speed and character of the spread of utility weights depend on the parameters of the model. If rich agents face less relative moral hazard, as is expected to occur when agents are not extremely risk averse and diminishing returns to effort do not set in suddenly, then the 'random walk' dispersion of utility weights is aggravated by the fact that the poor will bear more moral hazard than the

---

[14] The exact function is $q = \ln\left(\frac{\phi_i}{\phi_j}\right)$ if $\phi_i < \varepsilon$, $q = \ln\left(\frac{\varepsilon}{1-\varepsilon}\right) - \left(\frac{\varepsilon}{1-\varepsilon}\right)\frac{\phi_j}{\phi_i}$ if $\phi_j > \phi_i > \varepsilon$. $q = \ln\left(\frac{\varepsilon}{1-\varepsilon}\right) - \left(\frac{\varepsilon}{1-\varepsilon}\right) + \left(\frac{\varepsilon}{1-\varepsilon}\right)\frac{\phi_i}{\phi_j}$ if $\phi_i > \phi_j > \varepsilon$ and $q = 1 - \left(\frac{\varepsilon}{1-\varepsilon}\right) + \ln\left(\frac{\phi_i}{\phi_j}\right)$ if $\phi_j < \varepsilon$

[15] This is done in the Appendix under "Characterization of the Limiting Distribution"

rich and thus will expect to see their utility weight decrease further relative to that of the rich. We might refer to this situation as positive dispersion of utility weights.

On the other hand if the parameters of the model are such that richer agents face more moral hazard under the optimal contract, the random walk dispersion of utility weights will be mitigated by the fact the richer agents actually expect to be punished more than the poorer agents. This countervailing force will slow the spread of inequality, and if this holds in the limit, the limiting distribution of utility will be none-degenerate. However since this will occur only when the agents are very risk averse, the consequences of inequality will still drive any agent's expected utility down to zero.

Section 7: Lower Bounds on Utility

Up to this point the paper has considered situations in which the ability of the social planner to punish a particular agent was unlimited. That is to say no matter how poorly off an agent was he could always be made worse off. However one might believe that there are often limits to society's ability to punish an agent in a given period. For example there might be social norms against giving an agent too little consumption, or the agent might have the possibility of exit, or the agent might simply place a utility of greater than negative infinity on starving to death.

In this cases there is a limit to the quantity of utility that can be transferred away from an agent. Thus it is possible that under an optimal contract $W_{c_i}(\omega)$ will be negative in states which suggest low effort by agent i. This implies that the agents posterior utility weight is negative; the social planner would like to punish the agent more severely and would do so if it were possible to do without harming the other agents. If $W_{c_i}(\omega)$ is negative the following will be true:

- Agent i's consumption will be at the lowest possible level, and lemma 3.1 will not apply since the social planner is not free to transfer more consumption away from agent i.

- Because it is impossible to further decrease i's consumption, the only incentive i faces to contribute effort is the possibility that his future consumption will be increased if a signal indicative of high effort is realized. It is thus impossible to increase his effort without increasing his future utility.

- Society is on an upward sloping portion of the utility frontier as illustrated by the tangency in figure 7.1. Decreasing i's continuation utility will decrease his incentive to work and decrease the continuation utility of the other agents. Thus i can be said to be below his 'efficiency wage' level of utility. Although increasing i's utility to the efficiency wage level would be an ex-post Pareto improvement, it would diminish i's ex-ante incentive to work.

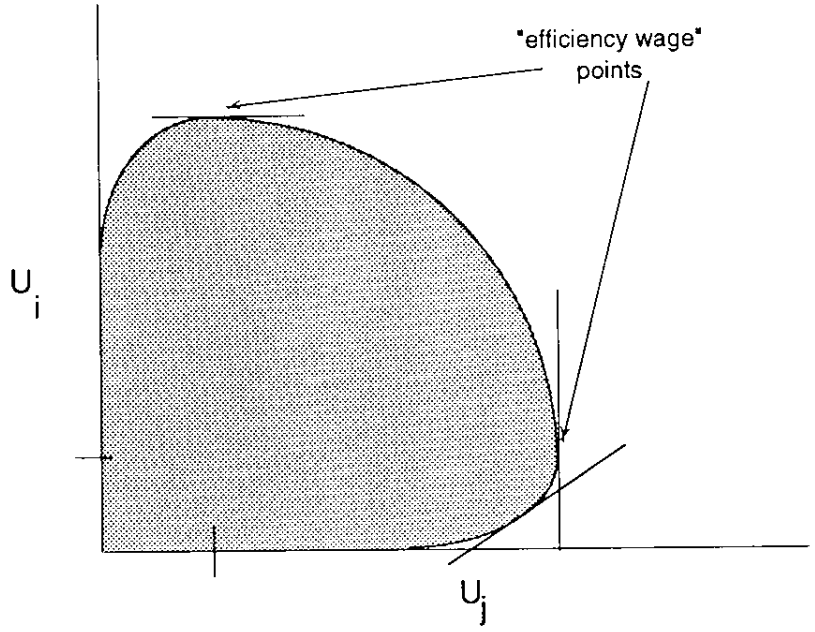

Figure 7.1

The possibility of having negative utility weights has implications about whether the optimal risk-sharing contract can be duplicated by a series of short-term contracts. If utility from consumption is not bounded below, the optimal contract can be achieved in a scenario where agents agree on one period contracts which map states into a consumption vectors and renegotiable continuation contracts. If the continuation contract is on a point on the utility frontier with a slope equal to the ratio's of the agents' marginal utility of consumption they will not renegotiate. However if utility is bounded below and the optimal long-term risk sharing contract calls for a point on the upward sloping portion of the utility frontier, as pointed out in Fudenberg, Holmstrom and Milgrom(1990) it is clear that it cannot be duplicated by a series of short-term contracts. The agents will renegotiate away from any continuation contract that is not Pareto efficient. So if there are lower bounds on the utility from consumption, the optimal social contract is not renegotiation proof, and society may not be able to commit to optimally punishing agents who are poor. In other words society might depart from the optimal contract and redistribute sacrificing some ex-ante efficiency.

Turning away from the question of renegotiation-proofness, we wish to ask what effect placing a lower bound on consumption or utility will have on the distribution of utility weight or consumption. If there are lower bounds on utility it is obvious that inequality cannot continue to increase once it has reached these bounds. However we are

still faced with the question of whether these bounds will be absorbing or reflecting. Do we expect that all agents will be pressed up against these bounds, or will the distribution of utility weight have a non-degenerate distribution, with a substantial proportion of the agents in between the bounds at any given time. In the remainder of the section I show that if there is a floor on consumption, we expect the limiting distribution to be degenerate, with all agents at either the upper or lower bound.

For mathematical simplicity I consider a principal-agent model with one agent. We will refer to the agent as i, and the principal as j. Since the principal will not need to be rewarded or punished his utility weight ϕ_j can be normalized to 1. Even if utility from consumption is bounded below for the principal and the agent, such as when $\gamma < 1$ or when there is a lower bound on consumption, proposition 3.3 will still apply and we will have the result that ϕ_i is a martingale. However ϕ_i will not always correspond to the ratio of marginal utility of consumption between the principal and the agent, since the social planner may run into a lower bound on consumption. Under these circumstances we are able to make some claims about the long run distribution of income. The justifications for these claims all rely on the following proposition.

Proposition 7.1: *If the consumption of the principal is bounded below by* $\underline{c} > 0$, ϕ_i, *the utility weight of the agent is bounded above by some* $\overline{\phi} < \infty$
Proof: See Appendix

The essence of the proof is that as the utility share of the agent grows the consumption and the variation in consumption of the principal shrinks. Since the consumption of the principal is bounded above zero, the principal's marginal utility of consumption is finite. Therefore as the variance in the principal's consumption goes to zero, the variance in his utility goes to zero as well. Thus the amount of moral hazard borne by the agent goes to zero, and the variation in ϕ_i goes to zero.

One can see that if we define $\tilde{\phi}$ as the lowest ϕ_i such that $\frac{v'(\underline{c})}{v'(x(\omega)-\underline{c})} > \phi_i$ for any $\omega \in \Omega$, any time that $\phi_i > \tilde{\phi}$, the principal will always consume $\underline{c}$ and be unaffected by the agent's effort, hence the agent will face no moral hazard. If the agent faces no moral hazard, $W_{e_i} = 0$, and by proposition 3.3, $\phi_{it+1} = \phi_{it}$, so the agents utility weight will be fixed. In this case we might think of the agent as a 'virtual principal'. Because the principal's consumption is fixed the agent bears all the risk and reaps all the gains of a successful outcome. It is worth noting that if the only signal of i's effort is binary, ϕ_i will always be less than $\tilde{\phi}$, so it will never be optimal to reward the agent so much that the

principal is forever held to the lower bound of his utility.[16] However if the signal of i's effort takes on more than 2 possible values, it may be possible to reach an absorbing state.

Having established that there is an upper bound on ϕ_i, we now ask the question of when there will be a lower bound on i's utility weight. We begin by defining the IRIE (infinite returns to infinitesimal effort) condition.

***Definition**: The production technology satisfies IRIE iff*

$\exists\ \omega$ *s.t.* $\lim_{e_i \to 0} p_{e_i}(\omega) = \infty$

Proposition 7.2: *If IRIE is not satisfied (i.e.* $\left|\frac{p_{e_i}(\omega)}{p(\omega)}\right|$ *is bounded above for all* $\omega \in \Omega$*),*

$\exists\ \underset{\sim}{\phi}$ *s.t. if* $\phi_i < \underset{\sim}{\phi}$, $e_i = 0$ *and* ϕ_i *is an absorbing state. Furthermore* ϕ_i *is bounded below by* $\underset{\sim}{\phi}$

Proof: See Appendix

According to proposition 7.2 if IRIE is not satisfied, there is some maximal punishment where agent i is held to the lowest possible utility in the future regardless of outcome. Obviously the agent will exert no effort in these states, and furthermore, the marginal increase in i's effort caused by increasing his incentives is zero. Thus it will not be optimal to provide any incentive to i, so $\phi_{it+1} = \phi_{it}$, and the maximal punishment is an absorbing state. Furthermore there is some $\underset{\sim}{e}_i > 0$ s.t. if $\phi_i > \underset{\sim}{\phi}$ $e_i \geq \underset{\sim}{e}_i$ As ϕ_i decreases below $\underset{\sim}{\phi}$, e_i discretely jumps down to zero.

The intuition for the result is that if the marginal effects of effort are not infinite, it is necessary to increase $U_i(\omega)$ discretely above the minimum in some future state to give i incentive to work. Thus if $e_i > 0$, it is possible to decrease u_i discretely by ceasing to give him incentive to work. If ϕ_i is negative, and of sufficient magnitude the (discretely positive) social benefit from punishing i will outweigh the future cost which must be finite. Social welfare will then be maximized when u_i is at its absolute minimum and e_i is zero. On the other hand, if IRIE is satisfied, it is possible to increase the agent's effort from zero without discretely increasing his utility, nevertheless, ϕ_i may still be bounded from below.

[16] For a proof of this, see Proof of Claim 7.2a in the Appendix.

Proposition 7.3: *If* $\lim_{e_i \to 0} \left| \frac{\lambda_i'^2}{\lambda_i''} \right|$ *is bounded,* $\exists\ \underset{\sim}{\phi}$ *s.t. if* $\phi_i < \underset{\sim}{\phi}$, $e_i = 0$ *and* ϕ_i *is an absorbing state. Furthermore* ϕ_i *is bounded below by* $\underset{\sim}{\phi}$[17]:

Proof: See Appendix

The condition in 7.3 ensures that as $e_i \to 0$, $\frac{de_i}{dU_i}\frac{dU_i}{de_i}$ is bounded. If $-\phi_i$ is greater than the upper bound, the social welfare function will be maximized by minimizing e_i, that is by setting e_i to 0. In any case where $\underset{\sim}{\phi}$ exists, its magnitude is positively related to the maximum of $\frac{de_i}{dU_i}\frac{dU_i}{de_i}$. Thus when $\frac{de_i}{dU_i}$ increases, signifying that i's effort is more cheaply bought, $\underset{\sim}{\phi}$ increases in magnitude, suggesting that it is less likely that U_i be set to the minimum. Intuitively this can be interpreted as showing that the easier it is to motivate the agent when he is very poor, the less likely the agent is to be 'given up on.'

Proposition 7.4: *If Proposition 7.1 and Proposition 7.2 or 7.3 hold then* $\exists\, a < 1$ *such that for any* $\varepsilon > 0$ $\lim_{T\to\infty} Pr(\phi_i > \tilde{\phi} - \varepsilon) = a$ *and* $\lim_{T\to\infty} Pr(\phi_i < \underset{\sim}{\phi} + \varepsilon) = 1-a$

Furthermore if i's initial utility weight is ϕ_{i0}, $\frac{\phi_{i0} - \underset{\sim}{\phi}}{\bar{\phi} - \underset{\sim}{\phi}} < a < \frac{\phi_{i0} - \underset{\sim}{\phi}}{\tilde{\phi} - \underset{\sim}{\phi}}$ *where* $\tilde{\phi}, \underset{\sim}{\phi}, \bar{\phi}$ *and* $\bar{\phi}$ *are given by the definitions above.*

Proof: See Appendix

Proposition 7.4 states that if there is a lower bound on the agent's utility weight, eventually either the agent or the principal is very likely to be very near the lower bound of

---

[17] It is not clear that there are any functions for p(ω) which are bounded which do not satisfy the condition that $\left| \frac{pe_i(\omega)}{pe_ie_i(\omega)} \right|$ is bounded above for all $\omega \in \Omega$ for any $e_i < \varepsilon$. Certainly any function of the form $p(\omega) = x + ae^{1-\gamma}$ satisfies A2, but also satisfies the condition for Lemma 7.3

consumption. The proof of Proposition 7.4 makes use of the fact that any bounded martingale converges to regions where its variance approaches zero. In this case the only regions where variance approaches zero are $(\phi_i > \tilde{\phi} - \varepsilon)$ and $(\phi_i < \underset{\sim}{\phi} + \varepsilon)$. Furthermore since the expectation of ϕ_i must always be ϕ_{i0}, we are able to estimate a, the likelihood that the principal is near the lower bound, by the equation $a\tilde{\phi} + (1-a)\underset{\sim}{\phi} \approx \phi_{i0}$. This result enables us to predict the long run distribution of consumption and income in situations where there is a lower bound on consumption. Note that a is decreasing in $\underset{\sim}{\phi}$, that is to say it is increasing in the magnitude of $\underset{\sim}{\phi}$. From our earlier result the magnitude of $\underset{\sim}{\phi}$ is increasing in the responsiveness of the agent to incentives, thus if an agent is more responsive to incentives, and all else is held equal, he will be more likely to end up at his maximum possible utility.

To summarize, the result is that under very general conditions, if there is a lower bound on the consumption of the principal and of the agent, in the long run, either the principal or the agent will be close to the lower bound with probability approaching one. Furthermore the probability that the agent is at the lower bound is decreasing in how responsive he is to incentives at very low effort levels.

These results differ from those of Atkeson and Lucas (1995) in that here the lower bound of utility is an absorbing state rather than a reflecting state. This difference rests entirely on the differing assumptions about the form the lower bound of utility takes. In Atkeson and Lucas there is a lower bound on agents' expected *future* utility, but it is still possible to punish the agent by decreasing his current consumption. When an agent is punished like this he would like to trade future utility for present consumption but runs into the lower bound on future utility. It is still possible to give this agent incentives by providing him with less than his expected continuation utility in future states which are indicative of low effort, and more utility in states indicative of high effort. In the states where he is rewarded, he will trade current consumption for future utility and lift himself above the lower bound on future utility. In this paper, on the other hand, the lower bound on continuation utility stems directly from a lower bound on instantaneous utility from consumption. Since an agent's utility is lowest when he knows that he will be at the lowest consumption level in every possible future state, the fact that he is at the lower bound of expected utility implies that he must not expect to be rewarded in any state, and is permanently stuck at this lower bound.

It should be pointed out that these results concerning bounded utility are all derived for the one agent case. If there are multiple agents it becomes more difficult to obtain results for the long run distribution of income. However the general result that agents face no moral hazard only when they are being maximally punished, or when everyone else is

being maximally punished, suggests that in the long run at most one agent will be above the lower bound on consumption.

Section 8: Conclusion

This paper shows that the optimal risk sharing contract under moral hazard has a recursive form and maps a prior utility weight vector and a realized state to a posterior utility weight vector. The contract punishes or rewards agents by increasing or decreasing their utility weight in accordance with a positive or negative signal about their effort. If utility is bounded below, following the optimal contract may lead to a situation where the utility weight of an agent is negative, meaning that the social planner is trying to minimize rather than maximize this agent's utility.

Under the optimal contract, utility weights will diverge as long as signals concerning different agents' efforts are sufficiently independent. For some utility functions this implies that inequality, measured as the ratio between the consumption of the richest and poorest agent, is always expected to be increasing. For other utility functions this measure of inequality may not always be increasing, but the ratio of need between the poorest and richest agent, measured as their marginal utility from consumption, is always increasing.

If moral hazard is not borne equally, utility weight will tend to migrate towards those who face the least moral hazard. This occurs because it is more socially efficient to provide incentives by punishment than by reward. Thus the utility of agents who face more moral hazard is decreasing over time for two reasons. Firstly, because they are risk averse, and uncertainty over their utility weight and consumption level is increasing over time, and secondly, because they are being punished more than they are being rewarded so their relative position is actually expected to be worsening over time.

One might not expect to see formal risk sharing contracts exactly like those described here in a real world setting, but as shown by Udry and Townsend it is not unreasonable to expect to see less formal approximations in areas where risk sharing is common. Specifically one might expect contracts where utility weight, although not formally defined, rests in a concept of indebtedness or entitlement, and this entitlement is changed from period to period based on outcomes and consumption. Applying the results of the paper to such situations, one might expect inequality in such a situation to rise over time.

The result of the paper that punishment is more effective in providing incentives than rewards, and that agents tend to lose utility weight relative to principals has widespread implications. It suggests that optimal incentive schemes will take the form of

granting agents entitlements and then, on average, taking them away. We can interpret the results of the model as predicting that lucky agents who experience good realizations will receive a greater share of societies resources, becoming something close to a principal, or capitalist and facing less moral hazard. Agents who suffer bad realizations continue to face more moral hazard, and expect to see their share of consumption continue to decrease.

References

Abreu, D., D. Pearce and E. Stacchetti, 1990 "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring", *Econometrica*, **58**(5), 1041-1063.

Atkeson, A. and R. E. Lucas Jr., 1995. "Efficiency and Equality in a Simple Model of Unemployment Insurance", *Journal of Economic Theory* **66**(1), 64-88

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

Banerjee, A. and A. Newman, 1991. "Risk-Bearing and the Theory of Income Distribution", *Review of Economic Studies* **58**(2), 211-235

Deaton, A. and C. Paxson, 1994. "Intertemporal Choice and Inequality", *Journal of Political Economy* **102**(3), 437-467

Fudenberg, D, B. Holmstrom and P. Milgrom.,1990. "Short-term Contracts and Long-term Agency Relationships", *Journal of Economic Theory* **51**(1), 1-31

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

Hansen, L. P. and J. A. Scheinkman, 1995 "Back to the Future: Generating Moment Implications for Continuous-Time Markov Processes", *Econometrica* **63**(4), 767-804

Holmstrom, B. "A Note on the First-Order Approach in Principal Agent Models," Mimeo, 1984.

Phelan, Christopher, and Robert Townsend, 1991. "Computing Multiperiod Information Constrained Optima," *Review of Economic Studies*, **58**, 853-881

Phelan, Christopher, 1994. "Incentives and Aggregate Shocks" *Review of Economic Studies*, **61**, 681-700

Phelan, Christopher, 1995. "Repeated Moral Hazard and One Sided Commitment " *Journal of Economic Theory*, **66**(2), 488-506

Phelan, Christopher, 1998. "On the Long Run Implications of Repeated Moral Hazard" *Review of Economic Studies*, **69**,174-191

Rogerson, W., 1985a. "Repeated Moral Hazard", *Econometrica*, **53**, 69-76

Rogerson, W., 1985b. "The First-Order Approach to Principal-Agent Problems", *Econometrica*, **53**, 1357-1368

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

Townsend, Robert, 1994. "Risk and Insurance in Village India", *Econometrica,* **62**, 539-591

Townsend, Robert, 1995. "Financial Systems in Northern Thai Villages", *Quarterly Journal of Economics* **110,** 1011-1047

Townsend, Robert, 1995. "Consumption Insurance: An evaluation of Risk Bearing systems in Low Income Economies", *Journal of Economic Perspectives,* **1995** 83-102

Udry, Christopher 1994"Risk and Insurance in a Rural Credit Market: An Empirical Investigation in Northern Nigeria" *Review of Economic Studies,* **61**,495-526

Udry, Christopher, 1995 "Risk and Saving in Northern Nigeria", *American Economic Review,* December 1995, 1287-1300

Appendix

Proof of Lemma 2.1:

Let $C(\Delta^J)$ be the space of continuos bounded functions on Δ^J, the interior of the J dimensional simplex.

Define the operator T on $C(\Delta^J)$ by

$$(TW)(\phi) = \sup_{c,e,z} \sum_{\omega\in\Omega} p(\omega)(\phi\bullet(v(c(\omega))-e+\beta\phi\bullet z(\omega))),$$

$$\text{s.t. } e_i \in \text{Argmax}_{e_i} \sum_{\omega\in\Omega} p(\omega)(v(c_i(\omega))-e_i+\beta z_i(\omega))),\ \theta\bullet z\le W(\theta)\ \forall\ \theta\in\Delta^J$$

Note that this is a well defined problem for any bounded W since it is the maximum of a continuous function and the range of values of $e\times c\times z\times\Omega$ space that must be considered is compact.[18]

We will show that T is a contraction mapping from $C(\Delta^J)$ to $C(\Delta^J)$.

The first step is to note weak monotonicity in T i.e. if $W'(\phi)\ge W(\phi)\ \forall\ \phi\in\Delta^J$, $(TW')(\phi)\ge (TW)(\phi)$

This is true because an increase in $W(\phi)$ simply represents a loosening of constraints on the problem. Now we will show that

if $\sup_{\phi\in\Delta^J}|W(\phi)-W'(\phi)| =\delta$ then $\sup_{\phi\in\Delta^J}|(TW)(\phi)-(TW')(\phi)|\le\beta\delta$

Consider $W_\delta(\phi) =W(\phi)+\delta\ \forall\ \phi\in\Delta^J$. We will show that

$$(TW_\delta)(\phi)-(TW)(\phi)\le\beta\delta\ \forall\ \phi\in\Delta^J$$

If H is a contract which specifies c,e,z as functions of ω, let H_δ specify c,e,z_δ, where the elements of z_δ are defined by $z_{i\delta}(\omega) = z_i(\omega)+\delta\ \forall\ i\in J,\ \omega\in\Omega$. It is clear that if H is feasible and incentive compatible under W, H_δ will be feasible and incentive compatible under W_δ furthermore

$$\sum_{\Omega} p(\omega,e)(\phi\bullet(v(c(\omega))-e+\beta\phi\bullet z(\omega))) - \sum_{\Omega} p(\omega,e)(\phi\bullet(v(c(\omega))-e+\beta\phi\bullet z_\delta(\omega))) = \beta\delta\ \forall\ \phi\in\Delta^J$$ since c and e are the same in both expressions. Thus for any contract H which obtains $TW(\phi)$ under W there is contract H_δ which obtains $TW(\phi)+\beta\delta$ under W_δ so $TW_\delta(\phi)\ge TW(\phi)+\beta\delta\ \forall\ \phi$. Analogously for any contract H which obtains $TW_\delta(\phi)$ under W_δ, there is a contract $H_{-\delta}$ which obtains $TW_\delta(\phi)-\beta\delta$ under W so $TW_\delta(\phi)\le TW(\phi)+\beta\delta\ \forall\ \phi$. Hence $TW_\delta(\phi) - TW(\phi) =\beta\delta$.

---

[18] Even if e is not bounded the spaces of effort levels that must be considered is bounded, it is clear that no effort level greater than $\max_{\phi_i,\Omega^T}\frac{1}{\phi_i p(\Omega^T)}\beta^{-T}(\phi v(\overline{c})-\phi v(\underline{c}))$ will ever be optimal, because a program which prescribed e=0 at all times would be preferable.

WLOG assume $TW'(\phi) \geq TW(\phi)$, then by the definition of W_δ and the fact that $\delta = \sup_{\phi \in \Delta^J} |W(\phi)-W'(\phi)|$. $W_\delta(\phi) \geq W'(\phi)\ \forall\ \phi$ Hence $TW'(\phi) \leq TW_\delta(\phi)$ and $TW'(\phi) \leq \beta\delta$.

Thus T is a contraction mapping from $C(\Delta^J)$ to $C(\Delta^J)$ and has a unique fixed point. This fixed point is analogous to the self generating profile described in Abreu, Pearce and Stacchetti(90)

We now show that if W* is the limit of the above contraction mapping, the optimal contract results in social welfare given by $W^*(\phi)$.
Suppose there were a general (not necessarily recursive) contract F, s.t.

$$\sum_{t=1}^{\infty} \beta^{t-1} p(\omega^t)\ \phi \bullet (v(c(\omega^t)) - e(\omega^{t-1})) > W^*(\phi)$$

which satisfies the incentive constraint that

$$e_{it+1}(\omega^t) \in \text{Argmax}_{e_{it+1}} \sum_{\tau=t+1}^{\infty} \beta^\tau (p(\omega^\tau) v(c_i(\omega^\tau)) - e_i(\omega^\tau)),$$ as well as the feasibility constraints.

Pick $\varepsilon > 0$ s.t. $\sum_{t=1}^{\infty} \beta^{t-1} p(\omega^t)\ \phi \bullet (v(c(\omega^t)) - e(\omega^{t-1})) - W^*(\phi) > \varepsilon$.

By the fact that $TW_\delta(\phi) \leq TW(\phi) + \beta\delta\ \exists\ \omega^1, \phi_2$ s.t.

$$\sum_{t=2}^{\infty} \beta^{t-2} p(\omega^t)\ \phi_2 \bullet (v(c(\omega^t)) - e(\omega^{t-1})) > W^*(\phi_2) + \frac{\varepsilon}{\beta}$$

Likewise it is clear that there is some $\omega^2 \in \Omega \times \Omega$ and ϕ^3 s.t.

$$\sum_{t=3}^{\infty} \beta^{t-3} p(\omega^t)\ \phi_3 \bullet (v(c(\omega^t)) - e(\omega^{t-1})) > W^*(\phi_3) + \frac{\varepsilon}{\beta^2}$$

And there is some $\omega^{n-1} \in \Omega^{n-1}$ $\sum_{t=n}^{\infty} \beta^{t-n} p(\omega^t)\ \phi_n \bullet (v(c(\omega^t)) - e(\omega^{t-1})) > W^*(\phi_3) + \frac{\varepsilon}{\beta^n}$, but this implies a contradiction when $\frac{\varepsilon}{\beta^n} > v(x(\bar{\omega}))$. So no contract can achieve welfare greater than $W^*(\phi)$, and the proposition is proven ♦

Proof of Lemma 3.1:

Without loss of generality, suppose $\frac{V_{i_c}(\omega)}{V_{j_c}(\omega)} = x > \frac{\phi_{jt+1}(\omega)}{\phi_{it+1}(\omega)}$, By Definition G_ϕ

maximizes $\sum_i \phi_{it} EU_{i,}$ so $\frac{\frac{dU_i}{d\phi}}{\frac{dU_j}{d\phi}} = \frac{-\phi_{jt+1}}{\phi_{it+1}}$ (*)

Consider the social welfare effect of perturbing ϕ so that j's utility is increased, i's is decreased, and the utility of all other agents remains constant. By (*) the ratio of the decrease in i's utility to the increase in j's is $\frac{\phi_{jt+1}}{\phi_{it+1}}$. However the ratio of the increase in i's utility to the decrease in j's from moving consumption to i from j, is x, which is greater than $\frac{\phi_{jt+1}}{\phi_{it+1}}$. Therefore total utility can be increased by perturbing ϕ and compensating by shifting consumption, and the contract was not optimal. Hence the lemma must hold in any optimal contract. ♦

Proof of Proposition 3.2:

Define $U_{i_{e_j}}$ as the partial derivative of agent i's utility with respect to agent j's effort, thus:

$$U_{i_{e_j}} = \int_\Omega U_i(\omega)\, p_{e_j}(\omega)\, d\omega$$

Note that if agents choose effort to maximize their own utility, the incentive compatibility constraint implies that $U_{i_{e_i}} = 0$. Let us define $U_{j_{e_j e_k}}$ as the change in j's return to effort directly caused by k's effort, so:

$$U_{j_{e_j e_k}} = \frac{dU_{j_{e_j}}}{de_k}$$

The matrix A is the $I \times I$ matrix whose elements are defined by:

$$a_{ij} = -\frac{U_{ie_ie_j}}{U_{ie_ie_i}}$$

So a_{ij} represents the change in i's effort caused by a change in j's effort.
Since $U_{i_{e_i}} = 0$, by the incentive compatibility constraint a_{ij} represents the change in i's effort caused by a change in j's effort. We define the matrix B as the $I \times I$ matrix whose elements are defined by:

$$b_{ij} = U_{i_{e_j}}$$

Thus b_{ij} represents the effect of e_j on i's utility, note that by incentive compatibility the diagonal elements in B are zero. We define b_k as the k^{th}row of B. Having introduced this

notation it is now possible to continue on with the proof that the marginal effect of an increase in $c_i(\omega)$ is given by:

$$\phi_i p(\omega) V_{i_c} - \sum_k \phi_k b_k (I-A)^{-1} z_i \frac{V_{i_c} p_{e_i}}{U_{ie_i e_i}} = \phi_i p(\omega) V_{i_c} - W_{e_i} \frac{V_{i_c} p_{e_i}}{U_{ie_i e_i}}$$

Increasing i's consumption in state c will directly increase his effort by $\frac{V_{i_c} p_{e_i}}{U_{ie_i e_i}}$. Agent i chooses e_i such that $U_{i_{e_i}} = 0$, thus

$$\sum_{\Omega} U_i(\omega) p_{e_i}(\omega)\, d\omega = 1 \quad (**)$$

Increasing $c_i(\omega)$ increases $U_i(\omega)$ by V_{i_c}, increasing the left side of (**) by $V_{i_c} p_{e_i}(\omega)$. To maintain equality effort must be increased so that $U_{i_{e_i}}$ decreases by $V_{i_c} p_{e_i}$, Hence the change in i's effort generated directly by the increase in c_i is $-\frac{V_{i_c} p_{e_i}}{U_{i_{e_i e_i}}}$

However, this will change the other agents' efforts, which will in turn change i's effort again. Note that **d,** the vector of changes in effort must satisfy
$\mathbf{d} = A\mathbf{d} + \mathbf{g}$ where $\mathbf{g}$ is the direct change in effort, caused by changes in the relative desirability of states and $A\mathbf{d}$ is the indirect change, caused by changes in other agents' effort

Solving for **d** we have $\mathbf{g} = (I-A)\mathbf{d}$, thus $\mathbf{d} = (I-A)^{-1}\mathbf{g}$
Thus the marginal change in effort due to increasing i's consumption is the column vector

$$(\mathbf{I} - A)^{-1} \mathbf{z}_i \frac{V_{i_c} p_{e_i}}{U_{ie_i e_i}}$$

where $\mathbf{z}_i$ is the i^{th} column of **I**, the identity matrix. We define $\mathbf{b}_k$ as the k^{th} row of B so if **d** is a vector of effort changes, $\mathbf{b}_k\mathbf{d}$ is the effect of an effort change of **d** on k's utility. Thus the effect on k's welfare from changed incentives is :

$$\mathbf{b}_k (I-A)^{-1} \mathbf{z}_i \frac{V_{i_c} p_{e_i}}{U_{ie_i e_i}} ,$$

and the total effect on social welfare from changed incentives is:

$$\sum_k \phi_k \mathbf{b}_k (I-A)^{-1} \mathbf{z}_i \frac{V_{i_c} p_{e_i}}{U_{ie_i e_i}}$$

Define W_{e_i} as the change in social welfare with respect to a change in i's incentives , so

$$W_{e_i} = \sum_k \phi_k \mathbf{b}_k (I-A)^{-1} \mathbf{z}_i$$

Hence the partial of social welfare with respect to $c_i(\omega)$ is given by:

$$W_{c_i} = \phi_i p(\omega) V_{i_c}(\omega) - W_{e_i} \frac{V_{i_c} p_{e_i}}{U_{ie_i e_i}}$$ ♦

Proof of Proposition 3.3:
When the resource constraint is binding

$W_{c_i}(\omega) = W_{c_j}(\omega)\ \forall$ i,j,ω. Thus by proposition 3.2

$$\frac{p(\omega)\phi_{it}V_{i_c}(\omega) - W_{e_i}\dfrac{V_{i_c}p_{e_i}}{U_{ie_ie_i}}}{p(\omega)\phi_{jt}V_{j_c}(\omega) - W_{e_j}\dfrac{V_{j_c}p_{e_j}}{U_{je_je_j}}} = 1 \quad (1)$$

From lemma 3.1, we know that $\dfrac{V_{i_c}(\omega)}{V_{j_c}(\omega)} = \dfrac{\phi_{jt+1}}{\phi_{it+1}}$

Dividing (1) Through by $V_{j_c}(\omega)$ we obtain.

$$\frac{\dfrac{\phi_{it}\phi_{jt+1}}{\phi_{it+1}}p(\omega) - W_{e_i}\dfrac{\dfrac{\phi_{jt+1}}{\phi_{it+1}}p_{e_i}}{U_{ie_ie_i}}}{p(\omega)\phi_{jt} - W_{e_j}\dfrac{p_{e_j}}{U_{je_je_j}}} = 1 \quad (2),\ \text{solving for } \phi_{it+1}(\omega) \text{ we obtain}$$

$$p(\omega)\frac{\phi_i\phi_{jt+1}}{\phi_{it+1}} - W_{e_i}\frac{\dfrac{\phi_{jt+1}}{\phi_{it+1}}p_{e_i}}{U_{ie_ie_i}} = p(\omega)\phi_{jt} - W_{e_j}\frac{p_{e_j}}{U_{je_je_j}} \quad (3),\ \text{This reduces to:}$$

$$\frac{\phi_{jt+1}}{\phi_{it+1}} = \frac{p(\omega)\phi_{jt} - W_{e_j}\dfrac{p_{e_j}}{U_{je_je_j}}}{p(\omega)\phi_{it} - W_{e_i}\dfrac{p_{e_i}}{U_{ie_ie_i}}} \quad (4) \qquad \blacklozenge$$

Proof of Proposition 3.4:

From Proposition 3.2 it can be seen that for any agent in I_1, the sign of the partial of the social welfare function depends only on ω^1 and not on ω^2. Likewise the sign of the partial for any agent in I_2 depends only on ω^2. Finally note that for every agent in I there is at least one state where the sign of the partial is positive. This is because $\int_{\Omega} p_{e_i}\, d\omega = 0$ for all i, thus there exists at least one state where p_{e_i} is non negative (or non positive) and where the partial is positive.

Suppose that the partial of an agent $i_1 \in I_1$ was non- positive for some sub-state $\mathring{\omega}^1$.

Consider an agent $i_2 \in I_2$, from above there exists a sub-state $\mathring{\omega}^2$ where his partial is

positive. In the state $\mathring{\omega} = \mathring{\omega}^1 \times \mathring{\omega}^2$ the partial of i_2 is positive, but the partial of i_1 is non-positive so social welfare can be improved by transferring resources from i_1 to i_2 in this state. Thus the original contract was not optimal. It is easy to see that if the partial is positive for all agents in all states, the resource constraint is always binding. ♦

Proof of Proposition 4.1:

We can divide the numerator and denominator of the expression in proposition 3.3 by $p(\omega)$ and we obtain:

$$\frac{\phi_{jt+1}}{\phi_{it+1}} = \frac{\phi_{jt} + W_{e_j} \dfrac{p_{e_j}}{p(\omega) U_{je_je_j}}}{\phi_{it} + W_{e_i} \dfrac{p_{e_i}}{p(\omega) U_{ie_ie_i}}} \qquad (4.1.1)$$

Note that whenever the numerator is positive and held constant this is a strictly convex function of $\dfrac{p_{e_i}}{p(\omega)}$.

Since $\sum_{\Omega} p(\omega) = 1$ by definition, it is a tautology that $\sum_{\Omega} p_{e_i}(\omega) = 0$

Thus the expectation of $\dfrac{p_{e_i}(\omega)}{p(\omega)} = \sum_{\Omega} \dfrac{p_{e_i}}{p(\omega)} p(\omega)\, d\omega$, which is equal to 0.

Since the mean of a convex function is greater than or equal to the function evaluated at its mean, conditional on $\dfrac{p_{e_j}}{p(\omega)}$ the expectation of $\dfrac{\phi_{jt+1}}{\phi_{it+1}}$ is greater than $\dfrac{\phi_{jt} - W_{e_j} \dfrac{p_{e_j}}{p(\omega) U_{je_je_j}}}{\phi_{it}}$. Since the expectation of $\dfrac{p_{e_j}}{p(\omega)}$ is zero, the expectation of $\dfrac{\phi_{jt} - W_{e_j} \dfrac{p_{e_j}}{p(\omega) U_{je_je_j}}}{\phi_{it}}$ is $\dfrac{\phi_{jt}}{\phi_{it}}$ and the expectation of $\dfrac{\phi_{jt+1}}{\phi_{it+1}}$ is greater than $\dfrac{\phi_{jt}}{\phi_{it}}$. ♦

Proof of Proposition 5.1

Since i bears no moral hazard($W_{e_i}=0$), by Lemma 4.1,

$$E\left(\frac{\phi_{kt+1}}{\phi_{it+1}}\right) = \frac{\phi_{kt}}{\phi_{it}} \text{ . Thus } E\frac{\sum_k \phi_{kt+1}}{\phi_{it+1}} = \frac{\sum_k \phi_{kt}}{\phi_{it}}$$

But $\frac{\phi_{it+1}}{\sum_k \phi_{kt+1}}$ is a strictly convex function of $\frac{\sum_k \phi_{kt+1}}{\phi_{it+1}}$

Thus $E\left(\frac{\phi_{it+1}}{\sum_k \phi_{kt+1}}\right) \geq E\left(\frac{\sum_k \phi_{kt+1}}{\phi_{it+1}}\right)^{-1}$, with the inequality being strict if there is any variation in $\frac{\sum_k \phi_{kt+1}}{\phi_{it+1}}$. If anyone faces moral hazard, there will be variation, so the inequality will be strict. ♦

Proof of Proposition 6.1

We begin by normalizing ϕ_i and ϕ_j so $\phi_i+\phi_j=1$.

$$\sqrt{Z_i} = \frac{\sqrt{Var(\sigma_i)}s_i W_{e_i}}{\phi_i} \tag{6.1}$$

$$Var(\sigma_i) = \sum_{\Omega} p(\omega)\left(\frac{p_{e_i}(\omega)}{p(\omega)}\right)^2 \text{ . Now since } p_{e_i}(\omega) = \lambda_i'(\mathbf{p}^0(\omega)-\mathbf{p}^1(\omega)) \text{ and } p(\omega) > \underline{p}$$

$$Var(\sigma_i) < \frac{(\lambda_i')^2}{\underline{p}} \tag{6.2}$$

Secondly because $\lambda'' < 0$, $s_i = \frac{\lambda_i'}{\lambda_i''}$ is bounded as well.

We rewrite the result of proposition 3.2 and obtain

$W_{e_i} = \phi_j b_{ji} + a_{ji} W_{e_j}$

We now show that $\exists$ K s.t. $|a_{ji} W_{e_j}| < K\phi_j$

By our assumption of independence of signals, we can write Ω as $\omega_i \times \omega_j$, where $p(\omega) = p_i(\omega_i)p_j(\omega_j)$ and by **A2**

$$p_i(\omega_i) = \lambda_i(e_i)(p_i^1(\omega_i)) - (1-\lambda_i(e_i))(p_i^0(\omega_i)) \tag{6.3}$$

Let $U_{i_{xy}}$ be agent i's expected utility when $p_i(\omega_i)$ is given by p_i^x and $p_j(\omega_j)$ is given by p_j^y in essence $U_{i_{11}}$ is i's utility if they both succeed, $U_{i_{01}}$ is i's utility if i fails but j succeeds, and so on.

By i's utility maximization:

$$\lambda_i'(\lambda_j(e_j)(U_{i_{11}}-U_{i_{01}})+(1-\lambda_j(e_j))(U_{i_{10}}-U_{i_{00}})) = 1 \qquad (6.4)$$

We are interested in showing that as $\phi_j \to 0$, $a_{ji}W_{e_j} < K\phi_i$
Our preliminary step is to show that $\lim_{\phi_j \to 0} W_{e_j} > 0$

$$W_{e_j} = b_{ij} + a_{ij}W_{e_i} \qquad (6.5)$$

Applying our definition of a_{ji} (The elasticity of i's effort with respect to j's effort)

$$a_{ij} = s_i\lambda_i'\lambda_j'(U_{i_{11}}+U_{i_{00}}-U_{i_{01}}-U_{i_{10}}) \qquad (6.6)$$

It is easy to show that by the optimal static contract $\lim_{\phi_j \to 0} W^{Static}=W_0^{FB}$. That is to say that the optimal static contract approaches the first best contract. Thus the optimal dynamic contract approaches the first best contract as well. Since U_j is bounded above:

$\lim_{\phi_j \to 0} U_i=W^{FB}$. This implies $U_{i_{xy}} = W_{i_{xy}}$

Combining 6.5 and 6.6 and substituting in

$$W_{e_j}= \lambda_j'(\lambda_i(e_i)(W_{11}-W_{01})+(1-\lambda_i(e_i))(W_{10}-W_{00})) + s_i\lambda_i'\lambda_j'(W_{11}+W_{00}-W_{01}-W_{10})W_{e_i}$$

Note that by **A2** $\mathbf{p}(e)$ describes a straight line in probability space, and for any p^0, p^1 on that line we can choose a λ-function. We choose p_i^1 such that $\min_\Omega p_i^1(\omega_i)=0$.
Thus there is a state ω_i^* where $p_i(\omega_i^*) =(1-\lambda_i(e_i))p_i^0(\omega_i^*)$

By Proposition 3.4 we know that $s_iW_{e_i}\underline{\sigma}_i(\omega_i) > -\phi_{e_i}$. Where $\underline{\sigma}_i$ is the lowest possible value of σ_i. However $\sigma_i=\dfrac{p_{e_i}}{p(\omega)}$ So

$$\underline{\sigma}_i=\frac{-\lambda_i'(e_i)p_i^0(\omega_i^*)}{(1-\lambda_i(e_i))p_i^0(\omega_i^*)} < \lambda_i'(e_i) \text{ so}$$

$$s_iW_{e_i}\lambda(\omega_i) < \phi_{e_i} \qquad (6.8)$$

By assumption $(W_{11}+W_{00}-W_{01}-W_{10}) < \lambda_i(e_i)(W_{11}-W_{01})+(1-\lambda_i(e_i))(W_{10}-W_{00})$

Thus we have established that $W_{e_j} > 0$, and we turn our attention to W_{e_i}

$$W_{e_i} = \lambda_i'(\lambda_j(e_j)(U_{j11}-U_{j01})+(1-\lambda_j(e_j))(U_{j10}-U_{j00})) + s_j\lambda_i'\lambda_j'(U_{j11}+U_{j00}-U_{j01}-U_{j10})W_{e_j} \quad (6.9)$$

Since $W_{e_j} > 0$, $(U_{j11}-U_{j01})>0$ and $(U_{j10}-U_{j00}) >0$ and

$$\frac{(U_{j11}+U_{j00}-U_{j01}-U_{j10})}{\lambda_j(e_j)(U_{j11}-U_{j01})+(1-\lambda_j(e_j))(U_{j10}-U_{j00})} < \frac{1}{\lambda(e_j)}$$

Furthermore $s_j\lambda_j'W_{e_j} < \phi_j$ or else there would be a state where $\phi_{jt+1}<0$.

Thus $W_{e_i} < \phi_j(1+\frac{1}{\lambda(e_j)})\,\lambda_i'(\lambda_j(e_j)(U_{j11}-U_{j01})+(1-\lambda_j(e_j))(U_{j10}-U_{j00}))$ (6.9)

Recall that in the limit $U_{ixx}\rightarrow W^{FB}$ and $W_{xx}\rightarrow W^{FB}$, therefore

$\lim_{\phi_j\rightarrow 0}\ \phi_j U_{jxx}\rightarrow 0$, so $\lim_{\phi_j\rightarrow 0}\phi_j U_{jxx}=0$, and

$\lim_{\phi_j\rightarrow 0} W_{e_i}=0$

In fact, we can say a little bit more about how W_{e_i} behaves as $\phi_j\rightarrow 0$

As $\phi_j\rightarrow 0$, $W_{e_i}\rightarrow 0$

As $W_{e_i}\rightarrow 0$, $\frac{\phi_{j11}}{\phi_{j01}}\rightarrow 1$, and $\frac{\phi_{j10}}{\phi_{j00}}\rightarrow 1$. Hence i's effort affects j only through the aggregate resource constraint. so $U_{j1x}-U_{j0x} < \frac{1}{1-\gamma}(\bar{x}^{1-\gamma}-\underline{x}^{1-\gamma})\,\phi_j^{-(1-1/\gamma)}$ (6.10)

Thus $W_{e_i} < \lambda_i'\,(\bar{x}^{1-\gamma}-\underline{x}^{1-\gamma})\,(1+\frac{1}{\lambda(e_j)})\,\phi_j^{1/\gamma}$, since $\lim_{\phi_j\rightarrow 0}\lambda'$ and $\lim_{\phi_j\rightarrow 0} s_i$ finite,

$$\sqrt{K_i} < \lambda_i' s_i \sqrt{Var(\sigma_i)}\,\lambda_i'\,(\bar{x}^{1-\gamma}-\underline{x}^{1-\gamma})\,(1+\frac{1}{\lambda(e_j)})\,\phi_j^{1/\gamma} \quad (6.11)$$

Thus

$$\lim_{\phi_j\rightarrow 0}\sqrt{K_i} < A\phi_j^{1/\gamma} \quad (6.12)$$

Our next step is to prove the following claim:

As $\phi_j\rightarrow 0$, $\exists\ K,>0\ \ \lambda_j' > K\phi_j^{(1-1/\gamma)}$ (6.13)

Let $U^*_{j_{xy}}$ be j's expected utility in state$_{xy}$ given that ϕ_j is held constant, and let $v^*_{j_{xy}}$ be j's expected instantaneous utility from consumption.

The claim will be true if $\lim_{\phi_j\rightarrow 0}\ \frac{\lambda_i(U_{j11}-U_{j10})+(1-\lambda_i)(U_{j01}-U_{j00})}{\lambda_i(U^*_{j11}-U^*_{j10})+(1-\lambda_i)(U^*_{j01}-U^*_{j00})} < \infty$

This implies that the incentive that j receives from the effect of his effort on his utility weight will not arbitrarily dominate the incentive he receives from the effect of his effort on aggregate consumption. By our assumption that $Z_j < Z_i$, we have that $Z_j \leq A\phi_j^{1/\gamma}$,

$$\lim_{\phi_j \to 0} \frac{Var(\phi_j)}{\phi_j} < A\phi_j^{1/\gamma} \tag{6.14}$$

Thus the difference between $v_{jxy} - v^*_{jxy} \sim \frac{1}{1-\gamma}(\bar{x}^{1-\gamma})((1+\phi_j^{1/\gamma})^{-(1-1/\gamma)} - 1)\phi_j^{-(1-1/\gamma)}$.

but $v_{j11} - v^*_{j10} \sim \frac{1}{1-\gamma}(\bar{x}^{1-\gamma} - \underline{x}^{1-\gamma})\,\phi_j^{-(1-1/\gamma)}$,

So for any $\gamma>1$, $\phi_j \to 0$ $\dfrac{\lambda_i(v_{j11}-v_{j10})+(1-\lambda_i)(v_{j01}-v_{j00})}{\lambda_i(v^*_{j11}-v^*_{j10})+(1-\lambda_i)(v^*_{j01}-v^*_{j00})} = 1$

Note that $\lim_{\phi_j \to 0} e_j = \bar{e}_j$. Since $\lim_{\phi_j \to 0} \frac{Var(\phi_j)}{\phi_j} = 0$,

$$e_{jt+1} \approx e_{jt} \approx \bar{e}_j \tag{6.15}$$

Since the limit in the variance in effort is 0,

$$U^*_{j11} - U_{j11} \approx \sum_{\tau=t}^{\infty} \beta^{\tau-t} E(v_{j\tau}|\phi^*_{j11}) - E(v_{j\tau}|\phi_{j11}) < \frac{1}{1-\beta} E(x^{1-\gamma}) \frac{1}{1-\gamma}((1+\phi_j^{1/\gamma})^{-(1-1/\gamma)} - 1)\phi_j^{-(1-1/\gamma)}$$

Since U* holds ϕ constant, $U^*_{j_{11}} - U^*_{j_{10}} = v^*_{j_{11}} - v^*_{j_{10}}$ and $\lim_{\phi_j \to 0} \frac{U_{j11}-U_{j10}}{v^*_{j11}-v^*_{j10}} = 1$

Thus for any $\gamma > 1$

$$\lim_{\phi_j \to 0} \frac{\lambda_i(U_{j11}-U_{j10})+(1-\lambda_i)(U_{j01}-U_{j00})}{\lambda_i(U^*_{j11}-U^*_{j10})+(1-\lambda_i)(U^*_{j01}-U^*_{j00})} = 1 \tag{6.16}$$

And the claim is proven.
Our next step is to show that

$$\frac{W_{e_i}}{\lambda_{j'}} > C \text{ for some C.} \tag{6.17}$$

By (6.5) $\dfrac{W_{e_i}}{\lambda_{j'}} = \dfrac{\phi_j b_{ji}}{\lambda_{j'}} + a_{ji}\dfrac{W_{e_i}}{\lambda_j}$. As $i \to 0$, $a_{ji} \to 0$ and $\dfrac{W_{e_i}}{\lambda_j}$ is bounded , so the second term goes to zero.
The first term, is the difference in i's utility between when j succeeds and when j fails. As $\phi_j \to 0$, $\lambda_{j'} \to 0$ and $\lambda_j \to \lim_{c_j \to \infty} \lambda_j$, also $c_i \to x$. Since j's probability of success will not depend on ϕ_j, and i will consume virtually all consumption in any future state, U_{it+1} will vary negligibly. Thus $\dfrac{b_{ji}}{\lambda_{j'}}$ is the difference in i's expected instantaneous utility of consumption, which is strictly positive and bounded away from zero.

Thus we have

$\sqrt{Z_j} = \mathrm{Var}(\sigma_j) s_j \frac{W_{ej}}{\phi_j}$. By definition $\mathrm{Var}(\sigma_j) > \lambda'^2 \|\mathbf{p}^0 - \mathbf{p}^1\|^2 = \lambda'^2 p^*$. By **A3** $s_j > \frac{\lambda_j'}{\kappa}$.

Combined with (6.17) and (6.13) we have:

$$\lim_{\phi_j \to 0} \sqrt{Z_j} \geq K^2 \phi_j^{(1-1/\gamma)} \frac{\phi_j^{(1-1/\gamma)}}{\kappa} - C \frac{\phi_j^{(1-1/\gamma)}}{\phi_j}$$

$$\lim_{\phi_j \to 0} \frac{\sqrt{Z_j}}{\sqrt{Z_i}} \geq K^2 C/A \frac{\phi_j^{3(1-1/\gamma)}}{\phi_j^{1+1/\gamma}}$$

Thus if $\gamma<2$, $\lim_{\phi_j \to 0} \frac{\sqrt{Z_j}}{\sqrt{Z_i}} = \infty$

Analysis of the Relationship Between Z and Utility Weight:

For simplicity we consider a model where there is a binary signal (success or failure) on each agent's effort. We will refer to the state where both agent's succeed as ω_{11}, and the state where i succeeds and j fails as ω_{10} and so on. We will use δ_i^{size} to refer to the difference between i's expected utility conditional on his success and his utility conditional on failure that is due to increased societal consumption. The symbol δ_i^{size} thus represents i's incentive from the effect of his effort on the 'size of the pie'. We note that if γ is the coefficient of risk aversion, δ_i^{size} is proportional to $\left(\frac{\phi_i^{1/\gamma}}{\phi_j^{1/\gamma} + \phi_i^{1/\gamma}}\right)^{1-\gamma}$. The other part of i's incentive is referred to as δ_i^{share} and is the difference in his utility due to the difference between ϕ_{it+1} conditional on success and ϕ_{it+1} conditional on failure. Note that if $Z_i = Z_j$ so that $\frac{\phi_{it+1}(\omega_{1x})}{\phi_{it+1}(\omega_{0x})} \approx \frac{\phi_{jt+1}(\omega_{x1})}{\phi_{jt+1}(\omega_{x0})}$ and the curvature of the utility function does not change too quickly δ_i^{share} will be proportional to $\frac{1}{\phi_i}$. To see this note that $\frac{dU_j}{dU_i} = -\frac{\phi_i}{\phi_j}$ and since $Z_i = Z_j$ the impact of j's failure on the ratio of utility weights is same as the impact of i's success.

$$\text{so } \frac{\lambda_i'}{\lambda_j'} = \frac{\delta_j^{size} + \delta_j^{share}}{\delta_i^{size} + \delta_i^{share}} = \frac{\left(\frac{\phi_i}{\phi_j}\right)^{1/\gamma - 1} \delta_i^{size} + \frac{\phi_i}{\phi_j} \delta_i^{share}}{\delta_i^{size} + \delta_i^{share}}$$

By simplifying assumption 2, $W_{e_i} \lambda_i' \phi_j (\delta_j^{size} - \delta_j^{share})$

If $Z_i = Z_j$ this implies $W_{e_i} = \lambda_i' \phi_j \left(\left(\frac{\phi_i}{\phi_j}\right)^{1/\gamma - 1} \delta_i^{size} - \frac{\phi_i}{\phi_j} \delta_i^{share}\right)$ and

$$W_{e_j} = \lambda_j'\phi_i(\delta_i^{size}-\delta_i^{share}).$$

$$\text{Thus } \frac{W_{e_i}}{W_{e_j}} = \frac{(\frac{\phi_j}{\phi_i})^{1/\gamma-1}\delta_i^{size}+\frac{\phi_i}{\phi_j}\delta_i^{share}}{\delta_i^{size}+\delta_i^{share}} \frac{\phi_j((\frac{\phi_j}{\phi_i})^{1/\gamma-1}\delta_i^{size}-\frac{\phi_i}{\phi_j}\delta_i^{share})}{\phi_i(\delta_i^{size}-\delta_i^{share})} > \frac{\phi_i}{\phi_j}(\frac{\phi_j}{\phi_i})^{2/\gamma}$$

So if $\gamma < 2$ then if $\phi_i > \phi_j$, $\frac{W_{e_i}}{W_{e_j}} < 1$, by the assumption that λ' decreases exponentially

$s_i=s_j$, thus by simplifying assumption 2) $Z_i < Z_j$

Thus if we assume $Z_i = Z_j$ and $\phi_i < \phi_j$, we obtain a contradiction, Similarly assuming $Z_i > Z_j$ leads to a contradiction leaving only the possibility that $Z_i > Z_j$.

We now consider taking into account the effect of one agents effort on the effort supply of the other agents. We can manipulate the intermediate results of Proposition 3.2 and obtain:

$$W_{e_j} = \sum_{i\neq j} b_{ij} + a_{ij} W_{e_i}$$

Let $U_{i_{xy}}$ be i's utility contingent on state ω_{xy}. Evaluating a_{ij} under exponentially decreasing returns to effort we obtain: $\frac{a_{ij}}{a_{ji}} = \frac{U_{i11}+U_{i00}-U_{i10}-U_{i01}}{U_{j11}+U_{j00}-U_{j10}-U_{j01}}$

Intuitively the negative effect of one agent's effort on the incentives of the other agent derives from the fact that if one agent works harder, societal consumption is likely to be higher and the marginal effects of additional consumption will be lower. Under CRRA utility functions an individual's utility from consumption is the product of a societal term and individual term. The negative effect of one agent's effort on the other's incentives will enter through the societal term and be multiplied by the individual term. We have already established that the individual term will be proportional to $\phi_i^{1/\gamma 1}$. So consequently this effect will be greater for the poorer agent whenever $\gamma > 1$. If the coefficient of risk aversion, $\gamma < 1$ then it is greater for the richer agent but $b_{ji}/b_{ij} > a_{ji}/a_{ij}$ so W_{e_i} is still greater for the poorer agent. Thus when we take into account the negative effects of one agent's effort on the other agents effort supply, it should not change the result that the poorer agent faces more moral hazard.

Now let us consider what happens when $\frac{var(\sigma_i)}{var(\sigma_j)} \neq \frac{(\lambda_i')^2}{(\lambda_j')^2}$, that is when the least likely informative state is more likely for one agent than the other. In this case we can still say that as long as long as the ratio of the probabilities of the least likely informative results for both agents is not much greater or less than the ratio of their utility weights the results still hold. By substituting into the earlier demonstration that if $\gamma>2$ $\phi_i>\phi_j \Rightarrow Z_i<Z_j$ We see that if

$$\frac{\phi_i^{1/\gamma}}{\phi_j^{1/\gamma}} \leq \sqrt{\frac{var(\sigma_i)}{var(\sigma_j)}\frac{(\lambda_i')^2}{(\lambda_j')^2}} \leq \frac{\phi_j^{1/\gamma}}{\phi_i^{1/\gamma}}$$ the results will still hold. ♦

Derivation of the Limiting Distribution:

We define $z = \frac{\ln(\phi_i)}{\ln(\phi_j)}$, and look at the limiting distribution of z. We approximate the dynamics of z by assuming it takes the form of Brownian motion with drift, with diffusion speed $\sigma^2 = E[(\frac{\ln(\phi_{it+1})}{\ln(\phi_{jt+1})} - \frac{\ln(\phi_{it+1})}{\ln(\phi_{jt+1})})^2]$ and drift $\mu = E(\frac{\ln(\phi_{it+1})}{\ln(\phi_{jt+1})} - \frac{\ln(\phi_{it+1})}{\ln(\phi_{jt+1})})$. If we assume that only ϕ_i varies when ϕ_j is poorer,

$\sigma^2 = E[\ln(\phi_{i_{t+1}}) - \ln(\phi_{i_t})]^2$, and $\mu = E[\ln(\phi_{i_{t+1}}) - \ln(\phi_{i_t})]$. Recall that $Z_i = \frac{Var(\phi_i)}{\phi_i^2} = Var(\ln \phi_i)$, so $\sigma^2 = Z_i$. It is shown in the proof of Proposition 6.1 that $Z_i \sim \phi_j^{2/\gamma} \sim e^{z(2/\gamma)}$. We find $E[\ln(\phi_{i_{t+1}}) - \ln(\phi_{i_t})]$ using a Taylor approximation, keeping in mind that $E[\phi_{i_{t+1}}] = \phi_{i_t}$, so

$\mu = -\frac{1}{2}\frac{Var(\phi_i)}{\phi_i^2} = -\frac{1}{2}Z_i$.

By equation (2.4) of Hansen and Scheinkman(1995), under Brownian motion, if $\rho(z)$ is the density of the limiting distribution,

$\frac{\rho(z)}{\rho(z')} = \frac{\sigma^2(z')}{\sigma^2(z)} \exp\left[2\int_z^{z'} \frac{\mu(y)}{\sigma^2(y)} dy\right]$. However we have shown that $\frac{\mu(y)}{\sigma^2(y)} = -\frac{1}{2}$, so $\frac{\rho(z)}{\rho(z')} = \frac{\sigma^2(z')}{\sigma^2(z)} e^{-(z'-z)}$. Thus $\frac{\rho(z)}{\rho(z')} \cong \frac{e^{z'(2/\gamma)}}{e^{z(2/\gamma)}} e^{(z'-z)} = e^{-(z'-z)(1-2\gamma)}$

We will now show that the weight in the tail is finite, keeping in mind that we are looking that the left tail of the symmetrical case.

$\text{Prob}(z<z') = \rho(z') \int_{-\infty}^{z} e^{-(z'-z)(1-2/\gamma)} = \rho(z')$. $= \rho(z') \frac{1}{1-2/\gamma} e^{-z'(1-2/\gamma)}$. This will be finite whenever $\gamma > 2$, which is not a relevant condition, because we have already shown that if $\gamma < 2$, there limiting distribution will be degenerate.

Now we show that $E[U_j \to -\infty]$. Under CRRA, $v_j \le \frac{1}{1-\gamma}\bar{x}^{1-\gamma}(\phi_j)^{-(1-1/\gamma)}$, since $\bar{x}$ is the aggregate consumption in the best state. Because the variance in ϕ_j approaches zero in the tail $U_j < \frac{1}{1-\beta}\frac{1}{1-\gamma}\bar{x}^{1-\gamma}(\phi_j)^{-(1-1/\gamma)}$. Normalizing $\phi_i = 1$, $\phi_j = e^{z_t}$ and $U_j < \frac{1}{1-\gamma}e^{-z(1-1/\gamma)}$

Thus j's expected utility from being in the left tail is

$$\rho(z')e^{-z'(1-2/\gamma)} \int_{-\infty}^{z} e^{z(1-2/\gamma)} e^{-z(1-1/\gamma)} = \rho(z')e^{-z'(1-2/\gamma)} \int_{-\infty}^{z} e^{-z/\gamma}$$

This integral does not converge, and since $\rho(z') > 0$, j's expected utility from being in the left tail is negative infinity. Since j's utility is bounded above everywhere, his expected utility in the limiting distribution is negative infinity. ♦

Proof of Proposition 7.1:

Proof:

Step 1) Let $\bar{\omega}$ be the best possible state ($x(\bar{\omega}) \geq x(\omega')\ \forall\ \omega' \in \Omega$), let $\tilde{\phi} = \frac{V_{j_c}(\underline{c})}{V_{i_c}(x(\bar{\omega})-\underline{c})}$. If $\phi_{it} > \tilde{\phi}$, the ratio $\frac{V_{j_c}(\underline{c})}{V_{i_c}(x(\bar{\omega})-\underline{c})} \leq \phi_{it}$, so j's consumption will be $\underline{c}$ regardless of state if $\phi_{it+1}=\phi_{it}$. Suppose $W_{e_i} \neq 0$, this implies $U_{j_{e_i}} \neq 0$, but in any state where $\phi_{it+1} > \phi_{it}$, j's continuation utility will be no higher since it is decreasing in i's utility weight. Furthermore since $\phi_{it+1} > \phi_{it}$, j's consumption will be $\underline{c}$. If $\phi_{it+1} < \phi_{it}$, j's utility will be no lower since continuation utility must be at least the same, and consumption can be no lower. Since increasing effort increases the chance of states where $\phi_{it+1} > \phi_{it}$, increasing effort cannot increase j's utility. Hence If $\phi_{it} > \tilde{\phi}$, $W_{e_i}=0$ and $\phi_{it+1}=\phi_{it}$

Step 2). If A2 does not hold W_{e_i} is bounded and $\frac{p_{e_i}}{p_{e_i}p_{e_i}}$ is bounded as well, so $\phi_{it+1} - \phi_{it}$ is bounded by β. Since if $\phi_{it} > \tilde{\phi}$, $\phi_{it+1}=\phi_{it}$, if $\phi_{it} = \tilde{\phi} + \beta$, $\phi_{it+1} = \tilde{\phi} + \beta$ thus ϕ_{it} is bounded. If A2 holds then if $e_i > \varepsilon > 0$, for any ϕ, then $\frac{p_{e_i}}{p_{e_i}p_{e_i}}$ is bounded and we can use the above proof to show ϕ_{it} is bounded. If $e_i < \varepsilon$, $U_i^+ - U_i^- < \delta$, Where U_i^+ is i's utility in the reward states and U_i^- is his utility in the punishment states. But if $\phi_{it}^+ > \tilde{\phi}$, then i's consumption in the reward states is $x(\omega^+)-\underline{c}$, but then i would have incentive to work since $x(\omega^+) > x(\omega^-)$ (i's effort increases social resources). Hence if $e_i < \varepsilon$, $\phi_i^+ < \tilde{\phi}$. So ϕ_{it} is bounded above. ♦

Proof of Claim 7.1a

Proof: Since the signal over i's effort is binary for any e_i, $\frac{p_{e_i}}{p(\omega)}$ takes on only two possible values. Hence ϕ_{it+1} takes on two possible values $\phi_i^- \leq \phi_{it} \leq \phi_i^+$. Since i is being

rewarded if $\phi_{it+1} > \phi_{it}$, it must be the case that $W_{e_i} > 0$ and thus $U_j(\phi_i^+) > U_j(\phi_i^-)$ but if $\phi_i^+ = \tilde{\phi}$, U_j is at its minimum, so $U_j(\phi_i^+) \leq U_j(\phi_i^-)$, and there is a contradiction. ♦

Proof of Proposition 7.2

By the agent's maximization problem:

If $e_i > 0 \Rightarrow \int_{\Omega} U_i(\omega)\, p_{e_i}(\omega)\, d\omega = 1$. If $\left|\frac{p_{e_i}(\omega)}{p(\omega)}\right|$ is bounded above by β, $\exists\ \omega^-, \omega^+$ s.t.

$U_i(\omega^+) - U_i(\omega^-) \geq 1/2\beta$, But $U_i(\omega^-) \geq \underline{U}_i$ so $U_i(\omega^+) \geq \underline{U}_i + \beta$, since the chance of $\omega^+ > \pi$

$E(U_i) = \underline{U}_i + \pi\beta - e_i$

But $E(U_i)$ is maximized over e_i thus if $e_i > 0 \Rightarrow E(U_i) \geq \underline{U}_i + \pi\beta$

Let $\tilde{U}$ be the lowest possible U_i for which $e_i > 0$. Let $U_j(U_i)$ be j's utility as a function of U_i. Since $\tilde{U} - \underline{U}_i \geq \pi\beta$ and U_j is bounded so $U_j(U_i) - U_j(\underline{U}_i) < \alpha$, $\exists\ \underline{\phi} < 0$ s.t.

$\max_{U_i > \tilde{U}} \underline{\phi} U_i + U_j(U_i) = \underline{\phi}\, \underline{U}_i + U_j(\underline{U}_i)$. Note that since we have assumed an efficiency wage exists, $\exists\ U_i > \tilde{U}$ s.t. $U_j(U_i) > U_j(\underline{U}_i)$, hence $\underline{\phi} < 0$. If $\phi_{it} < \underline{\phi}$, $\phi_{it}\underline{U}_i + U_j(\underline{U}_i) < \max_{U_i > \tilde{U}} \phi_{it} U_i + U_j(U_i)$. So $\underline{U}_i$ maximizes weighted social welfare. However at $\underline{U}_i$ is the absolute minimum utility so the probability that U_i increases in any future state must be zero. Hence setting $\phi_{it+1} = \phi_{it}$ is optimal.

The remainder of the proof that ϕ_i is bounded below is analogous to the proof that it is bounded above, the lower bound is simply the minimum possible ϕ_{it+1} arising from any state where $\phi_{it} > \underline{\phi}$ ♦

Proof of Proposition 7.3:

We will show that $\frac{dU_j(U_i)}{dU_i}$ is bounded above. Hence $\underline{U}_i$ maximizes $\phi_i U_i + U_j(U_i)$ if $\phi_i < \underline{\phi}$ when $\frac{dU_j(U_i)}{dU_i}$ is bounded above by $-\underline{\phi}$.

Since utility is bounded there is a maximum possible value e_i will take on, since $p_{e_i e_i}(\omega)$ is continuous and greater than zero $\left|\frac{p_{e_i}(\omega)^2}{p_{e_i e_i}(\omega)}\right|$ is bounded if $e_i > \varepsilon$. Thus if $\lim_{e_i \to 0} \left|\frac{\lambda_i'^2}{\lambda_i''}\right|$ is bounded, then $\left|\frac{p_{e_i}(\omega)^2}{p_{e_i e_i}(\omega)}\right|$ is bounded. Let $D = U_j(\tilde{U}) - \underline{U}_j$, Let $\bar{p}_{e_i} = \max_{\omega} |p_{e_i}|$ then $\frac{dU_j}{de_i} \leq D\bar{p}_{e_i}$.

$\frac{de_i}{dU_i} < \max_\omega \frac{p_{e_i}(\omega)}{p_{e_i e_i}(\omega)} \frac{1}{\underline{p}}$ Thus $\frac{dU_j}{de_i} \frac{de_i}{dU_i} < D\bar{p}_{e_i} \max_\omega \frac{p_{e_i}(\omega)}{p_{e_i e_i}(\omega)}$

Thus $\frac{dU_j}{dU_i} < \max_\omega \frac{D}{\underline{p}} \frac{p_{e_i} p_{e_i}(\omega)}{p_{e_i e_i}(\omega)}$. Hence if $\left| \frac{p_{e_i}(\omega)^2}{p_{e_i e_i}(\omega)} \right|$ is bounded then $\frac{dU_j}{dU_i}$ is bounded. ♦

Proof of Proposition 7.4:

Part 1) Let $\zeta = \min_{(\phi_i \mid \tilde{\phi} - \varepsilon > \phi_i > \underset{\sim}{\phi} + \varepsilon)} (\text{var}\ \phi_{it+1} | \phi_{it})$. Because ϕ_{it} is bounded above and below, Var ϕ_{it} is bounded. Because ϕ_{it} is a martingale. Var ϕ_{it+1} = Var ϕ_{it} + Var $(\phi_{it+1} | \phi_{it})$.

Suppose in the limiting distribution, $\Pr(\tilde{\phi} - \varepsilon > \phi_i > \underset{\sim}{\phi} + \varepsilon) > \delta$.

Then Var ϕ_{it+1} = Var $\phi_{it} + \zeta$. But then in the limit the variance would be unbounded. So $\text{Lim}_{T \to \infty} \Pr(\tilde{\phi} - \varepsilon > \phi_i > \underset{\sim}{\phi} + \varepsilon) = 0$, and the first part of the proposition is proven.

Part 2) $E(\phi_{it}) = \phi_{i0}$. $E(\phi_{it}) > a(\tilde{\phi} - \varepsilon) + (1-a)\underline{\phi}$, $a = \frac{\phi_{i0} - \underline{\phi}}{\tilde{\phi} - \varepsilon - \underline{\phi}}$

$E(\phi_{it}) < a\bar{\phi} + (1-a)(\underset{\sim}{\phi} + \varepsilon)$, $a = \frac{\phi_{i0} - \underset{\sim}{\phi} - \varepsilon}{\bar{\phi} - \underset{\sim}{\phi} - \varepsilon}$ ♦