Wealth Effects and Agency Costs *

Hector Chade[†] Virginia N. Vera de Serio[‡]

First version: June 24, 2010
This version: January 7, 2011

Abstract

In this note, we analyze how the agent's initial wealth affects the principal's expected profits in the standard principal-agent model with moral hazard and additively separable utility function in income and effort.

We show that if the principal prefers a poorer agent for all specifications of action sets, probability distributions, and disutility of effort, then the agent's utility of income *must* exhibit a coefficient of absolute prudence less than three times the coefficient of absolute risk aversion for all levels of income, thus strengthening the sufficiency result of Thiele and Wambach (1999).

We also prove that the principal *always* prefers a poorer agent in applications in which the task involved entails a small (in a precise sense) disutility of effort. The important implication of this result is that there is no condition on the agent's utility of income alone that will make the principal prefer *richer* agents.

Finally, we show that, for an interesting class of principal-agent problems, the principal prefers a relatively poorer agent if the agent's wealth is sufficiently *large*, as it would be the case in applications in which the agent is a CEO.

Keywords: Moral Hazard, Principal-Agent Model, Contracts, Wealth Effects.

*We are grateful to Eddie Schlee for comments and suggestions at several stages of this project.
[†]Arizona State University, Department of Economics, Tempe, AZ 85287-3806.
[‡]Universidad Nacional de Cuyo, Facultad de Ciencias Económicas, Mendoza, Argentina 5500.

1 Introduction

The principal-agent problem with moral hazard is one of the cornerstones of the theory of incentives. In its standard formulation, a risk neutral principal hires a risk averse agent to perform an unobservable task, and their relationship is regulated by a contract that is based on a stochastic observable outcome whose distribution is affected by the agent's actions.[1] In many applications, it is realistic to assume that the principal faces a pool of agents who are heterogeneous in their wealth. A natural question to ask then is whether the principal prefers to hire a poorer or a richer agent. More formally, Are agency costs increasing or decreasing in agent's wealth?

One can grasp the main difficulty in answering this question by looking at the simplest case with two actions ('high' and 'low') and two possible outcomes ('success' and 'failure'). If the principal wants to implement the high action, then the richer the agent is, the more difficult it is to satisfy the incentive constraint (i.e., the larger must be the difference between the wages paid when each outcome is observed), thereby making the incentive problem more severe. The impact of agent's wealth on the participation constraint, however, is ambiguous. For instance, if the agent exhibits increasing absolute risk aversion, then this constraint tightens, and thus agency costs increase as the agent's wealth increases. But if she exhibits decreasing absolute risk aversion, then participation may be easier to satisfy, and if this effect is strong enough it may be possible for the principal to implement the high action in a cheaper way as the agent becomes wealthier.

Thus, it is a priori unclear that anything of interest can be said about wealth effects and agency costs. In a nice paper, Thiele and Wambach (1999) (TW henceforth) proved that if the agent's utility function is additively separable in income and effort, and her utility of income exhibits a coefficient of absolute prudence that is less than three times its coefficient of absolute risk aversion, then the cost of implementing any given effort level increases in agent's wealth. As a result, the principal prefers a poorer agent.

Many common utility functions satisfy TW's sufficient condition, including CARA (constant absolute risk aversion) and CRRA (constant relative risk aversion) with coefficient of relative risk aversion bigger than a half. Hence, their result yields an interesting class of problems in which a clear answer to the aforementioned question obtains.

[1] The obvious references are Holmstrom (1979) and Grossman and Hart (1983). For a recent contribution to the development of the principal-agent framework, see Jewitt, Kadan, and Swinkels (2008).

To be sure, several open questions remain. For instance, Can a weaker condition suffice? Also, Is there an analogous condition under which the reverse result holds? That is, a sufficient condition on the agent's utility function of income such that the principal prefers a richer instead of a poorer agent. Finally, Does the principal prefer a poorer agent in 'extreme' cases, such as when the disutility of effort becomes small or the agent's wealth becomes large? This note provides answers to all these questions.

First, we show that TW's sufficient condition is *tight* in the following meaningful sense: if we want the principal to prefer a poorer agent across all principal-agent problems (i.e., for all action sets, probability distribution of the observable outcomes, disutility of effort, and wealth levels), then the agent's utility of income *must* satisfy TW's condition. We prove that if their condition fails for some level of income, then one can construct a robust principal-agent problem where the principal prefers a richer agent.

Second, we show that given a principal-agent problem, the principal *always* prefers a poorer agent if the task involved entails a small (in a precise sense) disutility of effort. An important implication of this result is that one *cannot* find an analogous condition to TW's that would define a class of problems where the principal prefers a *richer* agent.

Finally, we show that in a class of principal-agent problems that includes those with two outcomes and with HARA (hyperbolic absolute risk aversion) utility functions of income, the principal *always* prefers a poorer agent if her wealth is sufficiently large. As an illustration, consider the shareholders/CEO application of the principal-agent model. Clearly, the pool of available candidates from which shareholders select a CEO consists of fairly wealthy individuals. Our result suggests that shareholders would prefer to hire a relatively poorer CEO from the pool of potential candidates for the job.

As TW illustrate, the effects of agent's wealth on agency costs has important implications for organizational design in a wide variety of economic environments. Moreover, recent contributions show that they are also key for understanding some features of dynamic moral hazard models (e.g., Chiappori, Macho, Rey, and Salanié (1994), Hopenhayn and Nicolini (1997), Park (2004), Spear and Wang (2005), and Chade (2009)). Since contracting settings with moral hazard are usually complex, few comparative static results are known. Thus, it is important for applications to find either a simple condition imposed on preferences or on other primitives such that unambiguous wealth effects ensue.[2] Our results provide an important extension of TW's main result, thereby enlarging

[2]See also Chiu (2010), who provides another instance where the principal prefers a poorer agent by

the class of applications with moral hazard in which wealth effects can be signed.

The next section describes the model and TW's result. Section 3 presents our main results. Section 4 concludes.

2 The Model and Preliminary Results

The Model. The set-up is the standard principal-agent problem with moral hazard (e.g., Grossman and Hart (1983)). A principal hires an agent to perform a certain task, but since her effort is unobservable, the contract is based on a stochastic output generated by her effort. The only difference with the standard model is that we assume that the agent has an observable 'initial wealth,' a scalar denoted by θ.

The principal is risk neutral and maximizes expected profits defined as the difference between expected output and expected compensation paid to the agent. The agent is risk averse, with utility function for income-action pairs (I, a) given by $V(I+\theta)-\psi(a)$, where $V : (I_\ell, \infty) \to \mathbb{R}$, $I_\ell \geq -\infty$, is three times continuously differentiable, strictly increasing, and strictly concave; i.e., $V'(\cdot) > 0$, and $V''(\cdot) < 0$. Also, $\lim_{I \to I_\ell} V(I) = -\infty$. In turn, $\psi(\cdot)$ is nonnegative for all actions a, and it is strictly increasing in a.

Let $\bar{I}$ be the (certain) income the agent could obtain elsewhere if she does not work for the principal. Then her reservation utility is $V(\bar{I} + \theta)$.

We denote by $R(\cdot) = -V''(\cdot)/V'(\cdot)$ and $P(\cdot) = -V'''(\cdot)/V''(\cdot)$ the coefficients of absolute risk aversion and prudence, respectively, associated with $V(\cdot)$.

Let A be the set of feasible actions (e.g., effort levels) available to the agent. We focus on the two most oft-used cases in applications, namely, A is either a finite set $a_1 < a_2 < \cdots < a_m$, or an interval $[0, \bar{a}]$. Wlog, we assume that the lowest action in each case is costless for the agent (i.e., $\psi(a_1) = \psi(0) = 0$).

The observable output q assumes a finite set of values $q_1 < q_2 < \cdots < q_n$. The probability of observing q_i, $i = 1, ..., n$, when the agent's action is a is denoted by $\pi_i(a)$, and it is positive for all i and a. We denote by $\pi(a)$ the vector $(\pi_1(a), \pi_2(a), ..., \pi_n(a))$.

When $A = [0, \bar{a}]$, we will further assume that $\psi(\cdot)$ and $\pi_i(\cdot)$ are twice continuously differentiable in a, and that $\psi(\cdot)$ is strictly convex in a, i.e., $\psi''(\cdot) > 0$, with $\psi'(0) = 0$.

imposing conditions on other primitives of the model. We comment on this paper below.

Since the agent's action is unobservable, the principal offers a compensation contract $(I_1, I_2, ..., I_n)$ contingent on output and recommends an action a to the agent. Let $B(a) = \sum_{i=1}^{n} \pi_i(a) q_i$ be the expected value of output given action a, and let $C(a, \theta)$ be the minimum cost for the principal of implementing action a if the agent's wealth is θ. As in Grossman and Hart (1983), one can split the analysis of the problem in two steps: first, for each action a, find the contract that minimizes the expected cost to the principal and obtain $C(a, \theta)$; second, find the action that maximizes $B(a) - C(a, \theta)$.

This completes the description of the model. Notice that in terms of primitives, we can succinctly denote a principal-agent problem by $(V(\cdot), \psi(\cdot), \pi(\cdot), A, q, \bar{I}, \theta)$.

Remark 1 *We focus on the canonical case of additively separable utility function in income and effort. Thiele and Wambach (1999) also provide an extension of their result to the nonseparable case. Notice, however, that deterministic contracts need not suffice in the nonseparable case, as the principal might profit from randomization. Hence, that extension is of limited value if stochastic contracts are not allowed. For this reason, we restrict attention to the set up used by most of the moral hazard literature.*

THE COST MINIMIZATION PROBLEM. The function $C(a, \theta)$ solves:

$$C(a, \theta) = \min_{I_1, \ldots, I_n} \sum_{i=1}^{n} \pi_i(a) I_i$$

$$s.t. \quad \sum_{i=1}^{n} \pi_i(a) V(I_i + \theta) - \psi(a) \geq V(\bar{I} + \theta) \tag{1}$$

$$a \in \text{argmax}_{a' \in A} \sum_{i=1}^{n} \pi_i(a') V(I_i + \theta) - \psi(a'), \tag{2}$$

where (1) is the participation constraint and (2) is the incentive constraint.

If the action set is finite, then (2) consists of a finite number of incentive constraints $\sum_{i=1}^{n} \pi_i(a) V(I_i + \theta) - \psi(a) \geq \sum_{i=1}^{n} \pi_i(a') V(I_i + \theta) - \psi(a')$, for all a'. If the action set is an interval, then we replace (2) by the first-order condition of the agent's problem $\sum_{i=1}^{n} \pi_i'(a) V(I_i + \theta) - \psi'(a) = 0$, and we assume that $\pi(\cdot)$ satisfies the monotone likelihood ratio property (henceforth MLRP) and the convexity of the distribution function condition (henceforth CDFC), so that the 'first-order approach' is valid (Rogerson (1985)).[3] Also, to avoid dealing with the uninteresting case in which the constraint set

[3]All we need is the validity of the first-order approach. We use MLRP and CDFC for simplicity.

is empty for some a, we assume that, for each a, there exists an i such that $\pi_i'(a) \neq 0$. (A simple sufficient condition is to assume the strict MLRP for every a.)

An equivalent formulation of the problem with the contract written in utility units is as follows: $\min_{v_1,\ldots,v_n} \sum_{i=1}^n \pi_i(a)h(v_i) - \theta$ subject to $\sum_{i=1}^n \pi_i(a)v_i - \psi(a) \geq V(\bar{I}+\theta)$ and $a \in \text{argmax}_{a' \in A} \sum_{i=1}^n \pi_i(a')v_i - \psi(a')$, where $h(\cdot)$ denotes the inverse function of $V(\cdot)$; i.e., $h(\cdot) = V^{-1}(\cdot)$. We use both formulations interchangeably.

A key issue is to understand the behavior of $C(a,\theta)$ as θ changes. The Envelope Theorem plus the binding constraint (1) yield (see TW Propositions 1 and 2)

$$\begin{aligned} \frac{\partial C(a,\theta)}{\partial \theta} &= -1 + V'(\bar{I}+\theta) \sum_{i=1}^n \pi_i(a) \frac{1}{V'(I_i+\theta)} \\ &= V'(h(\overline{v})) \left(\sum_{i=1}^n \pi_i(a) \frac{1}{V'(h(v_i))} - \frac{1}{V'(h(\overline{v}))} \right), \end{aligned} \tag{3}$$

where we have set $\overline{v} = V(\bar{I}+\theta)$ for notational simplicity.

TW's Result. From (3), the cost of implementing an action is increasing (decreasing) in agent's wealth if $\sum_{i=1}^n (\pi_i(a)/V'(h(v_i)))$ is bigger (smaller) than $1/V'(h(\overline{v}))$.

Thiele and Wambach (1999) provided the following condition on $V(\cdot)$ for the principal to prefer poorer agents for all choices of the other primitives of the model. We include an alternative simple proof that relies on convexity and Jensen's inequality.

Proposition 1 (Sufficiency, TW) *If $V(\cdot)$ satisfies $P(I+\theta) \leq 3R(I+\theta)$ for all $I+\theta$, then the principal's cost of implementing any action higher than the lowest one is an increasing function of the agent's wealth θ. As a result, the principal's expected profit is a decreasing function of the agent's wealth θ.*

Proof. Since (1) binds and a is not the lowest action (i.e., $\psi(a) > 0$), $\sum_{i=1}^n \pi_i(a)v_i > \overline{v}$. Simple algebra shows that $P(I+\theta) \leq 3R(I+\theta)$ for all $I+\theta$ if and only if $1/V'(h(v))$ is convex in v.[4] Now apply Jensen's inequality to (3) (recall $1/V'(h(v))$ is increasing in v). As $B(a) - C(a,\theta)$ is decreasing in θ for every a, so is $\max_{a \in A} B(a) - C(a,\theta)$. □

As we stated in the Introduction, many common utility functions satisfy this condition, and hence the result holds for an interesting class of principal-agent problems.

[4] See Amir and Czupryna (2004) for more on the behavior of the derivatives of inverse utility and their connections with absolute risk aversion and prudence.

3 Main Results

3.1 Tightness of TW's Condition

How tight is TW's sufficient condition for the principal to prefer a poorer agent? The proof of Proposition 1 reveals that there is some unused *slack* in the participation constraint, namely, $\sum_{i=1}^{n} \pi_i(a)v_i = \overline{v} + \psi(a)$ implies $\sum_{i=1}^{n} \pi_i(a)v_i - \overline{v} > 0$ since $\psi(a) > 0$ for all actions above the lowest one. That is, convexity of $1/V'(h(\cdot))$ and the application of Jensen's inequality to obtain Proposition 1 seem far from being tight. This begs the question of whether a *weaker* condition on $V(\cdot)$ would suffice. The next result shows that the answer is negative: TW's condition is indeed *necessary* in a precise sense.

Proposition 2 (Tightness) *(i) If the principal's cost of implementing an action higher than the lowest one is increasing in the agent's wealth θ for all choices of $(\psi(\cdot), \pi(\cdot), A, q, \bar{I}, \theta)$, then $V(\cdot)$ satisfies TW's condition.*

(ii) If $V(\cdot)$ does not satisfy TW's condition, then the principal prefers a richer agent in some principal-agent problem.

Proof. (i) Suppose that $V(\cdot)$ is such that $P(\hat{I} + \hat{\theta}) > 3R(\hat{I} + \hat{\theta})$ for some $\hat{I}$ and $\hat{\theta}$. We will show that there exists a principal-agent problem with $V(\cdot)$ as the agent's utility function such that, for some action a, $\partial C(a, \theta)/\partial\theta < 0$ in an open neighborhood of $\hat{\theta}$.

To this end, assume a continuum of actions $A = [0, \bar{a}]$; two output levels q_1 and q_2, with $q_2 > q_1$; a probability distribution $\pi(\cdot)$ that is three times continuously differentiable with $\pi_2'(a) > 0$, $\pi_2''(a) < 0$; a disutility of effort $\psi(\cdot)$ that is three times continuously differentiable in a, with $\psi''(0) > 0$; and let $\bar{I}$ and θ be arbitrary.

For any action $a \in (0, \bar{a})$, constraints (1)–(2) are $(1 - \pi_2(a))v_1 + \pi_2(a)v_2 - \psi(a) = \overline{v}$ and $\pi_2'(a)(v_2 - v_1) = \psi'(a)$. Thus, the optimal contract that implements a is given by

$$v_1 = \overline{v} + \psi(a) - \pi_2(a)\frac{\psi'(a)}{\pi_2'(a)} \tag{4}$$

$$v_2 = \overline{v} + \psi(a) + (1 - \pi_2(a))\frac{\psi'(a)}{\pi_2'(a)}. \tag{5}$$

Notice that v_1 and v_2 are twice continuously differentiable in a and $C(a, \theta) = \pi_2(a)h(v_2) + (1 - \pi_2(a))h(v_1)$. It then follows from Lemma 1 in the Appendix that there exists an

$\tilde{a} > 0$ such that $\partial C(a,\theta)/\partial\theta > 0$ for all θ when $a \in (0,\tilde{a})$ if

$$P(\bar{I}+\theta) < 3R(\bar{I}+\theta) + \frac{(\pi_2'(0))^2 V'(\bar{I}+\theta)}{\pi_2(0)(1-\pi_2(0))\psi''(0)}, \tag{6}$$

and, conversely, if there is such an action $\tilde{a}$, then (6) holds with less than or equal to.

To complete the proof, recall that $P(\hat{I}+\hat{\theta}) > 3R(\hat{I}+\hat{\theta})$ for some $\hat{I}$ and $\hat{\theta}$, and set $\bar{I} = \hat{I}$ and $\theta = \hat{\theta}$. Then there exists a threshold $\tilde{k} > 0$ such that if $\psi''(0) > \tilde{k}$,

$$P(\hat{I}+\hat{\theta}) > 3R(\hat{I}+\hat{\theta}) + \frac{(\pi_2'(0))^2 V'(\hat{I}+\hat{\theta})}{\pi_2(0)(1-\pi_2(0))\psi''(0)},$$

and therefore $\partial C(\hat{a},\hat{\theta})/\partial\theta < 0$ for some action $\hat{a} \in (0,\tilde{a})$.[5] But then $\partial C(\hat{a},\theta)/\partial\theta < 0$ for all levels of wealth θ in an open neighborhood of $\hat{\theta}$.

(*ii*) Consider the principal-agent problem constructed in (*i*). Since there are two output levels, $B(a) = q_1 + \pi_2(a)\Delta$, where $\Delta = q_2 - q_1 > 0$. Moreover, $\pi_2''(a) < 0$ implies that any action can be made optimal for the principal (i.e., solve $\max_{a\in A} B(a) - C(a,\theta)$) by a judicious choice of Δ. In particular, this applies to action $\hat{a}$. Thus, for wealth levels in an open neighborhood of $\hat{\theta}$, the principal prefers a richer agent. □

For some intuition, notice that the aforementioned slack is smaller for low levels of effort (i.e., actions such that $\psi(a)$ is small). Thus, if TW's condition does not hold, the natural place to look for a violation of the result is in a principal-agent problem where the principal finds it optimal to implement a relatively low action. This is what the proof of Proposition 2 accomplishes. The task is not straightforward, for one needs to analyze nontrivial second-order effects to pin down the behavior of (3) near the lowest action (see Lemma 1 in the Appendix).

In short, the analysis reveals that we *cannot* weaken TW's condition if we want a parsimonious condition imposed *solely* on the utility of income that yields a preference for poorer agents for *all* principal-agent problems with primitives $(\psi(\cdot), \pi(\cdot), A, q, \bar{I}, \theta)$.[6]

[5] As an example, let $\psi(a) = ka^2/2$, $k > 0$, so that $\psi''(0) = k$. The result holds for $k > \tilde{k}$.

[6] This does not contradict a recent paper by Chiu (2010) that shows that risk aversion alone is sufficient for the principal to prefer a poorer agent if all wages in the contract are *nonnegative*, since nonnegativity requires restrictions on other primitives besides the agent's utility of income.

3.2 Wealth Effects for Small Disutility of Effort

Suppose we restrict attention to contracting situations in which the task involved entails a small disutility of effort for the agent, as it would be the case in applications where the agent is hired to perform some minor task. Will the principal prefer a poorer agent?

The answer to this question will also shed light on the following important one: Could we find a condition analogous to TW's under which the principal always prefers a *richer* agent? Such a condition would be useful in applications in the same way as TW's condition is. As we shall see below, the answer is negative.

Let us fix a principal-agent problem $(V(\cdot), \psi(\cdot), \pi(\cdot), A, q, \bar{I}, \theta)$ *without* imposing TW's condition. Following Grossman and Hart (1983) (Section 5), we parameterize the agent's disutility of effort by $\eta \geq 0$ as follows: $\tilde{\psi}(a) = \eta\psi(a)$. Also, assume that the solution of the cost minimization problem is continuously differentiable in η (e.g., as when $n = 2$).[7]

Note that if the disutility of effort were zero, then the optimal contract would pay a flat wage equal to the reservation wage; i.e., if $\eta = 0$, then $v_i = \overline{v}$ for all i. Hence, (3) implies that $\partial C(a, \theta)/\partial\theta$ vanishes when evaluated at $\eta = 0$. We will show that this derivative is positive for $\eta > 0$ small enough; i.e., the principal's cost of implementing an action is increasing in agent's wealth when her disutility of effort is small. As a result, the principal prefers a *poorer* agent in this case.

Proposition 3 (Small Disutility of Effort) (*i*) *For any action a above the lowest one, there is an* $\eta_a > 0$ *such that the cost of implementing action a strictly increases in agent's wealth for all* $0 < \eta < \eta_a$*;*

(*ii*) *The principal prefers a poorer agent if her disutility of effort is sufficiently small.*

Proof. (*i*) Since $\partial C(a, \theta)/\partial\theta|_{\eta=0} = 0$, it suffices to show that $\partial^2 C(a, \theta)/\partial\theta\partial\eta|_{\eta=0} > 0$, for then $\partial C(a, \theta)/\partial\theta$ would be positive for values of η in a (right) neighborhood of zero.

Straightforward differentiation yields

$$\partial^2 C(a, \theta)/\partial\theta\partial\eta = V'(h(\overline{v})) \sum_{i=1}^{n} \pi_i(a) \frac{R(h(v_i))}{V'(h(v_i))^2} \frac{\partial v_i}{\partial\eta}.$$

[7] Actually, Proposition 3 only uses continuous differentiability at $\eta = 0$.

Therefore,

$$\begin{aligned}
\partial^2 C(a,\theta)/\partial\theta\partial\eta|_{\eta=0} &= \frac{R(h(\overline{v}))}{V'(h(\overline{v}))}\sum_{i=1}^{n}\pi_i(a)\frac{\partial v_i}{\partial\eta}|_{\eta=0} \\
&= \frac{R(h(\overline{v}))}{V'(h(\overline{v}))}\left(\frac{\partial\left(\sum_{i=1}^{n}\pi_i(a)v_i\right)}{\partial\eta}|_{\eta=0}\right) \\
&= \frac{R(h(\overline{v}))}{V'(h(\overline{v}))}\left(\frac{\partial(\overline{v}+\eta\psi(a))}{\partial\eta}|_{\eta=0}\right) \\
&= \frac{R(h(\overline{v}))}{V'(h(\overline{v}))}\psi(a) \\
&> 0,
\end{aligned}$$

where the third equality follows from the binding participation constraint. By continuity, there exists a threshold η_a such that $C(a,\cdot)$ is strictly increasing in θ if $\eta \in (0,\eta_a)$.

(*ii*) If A is a finite set, then part (*i*) implies that the cost of implementing any action $a \in A$ is increasing in θ if $\eta \in (0,\eta^*)$, where $0 < \eta^* = \min_{a\in A-\{a_1\}}\eta_a$.[8] Thus, $\max_{a\in A} B(a) - C(a,\theta)$ is decreasing in θ for all $\eta \in (0,\eta^*)$.

Suppose $A = [0,\bar{a}]$. Note that the optimal action when $\eta = 0$ is $\bar{a}$, for it solves $\max_{a\in[0,\bar{a}]} B(a) - \bar{I}$ and MLRP implies that $B(\cdot)$ is increasing in a. Consider an action $\tilde{a}$ arbitrarily close to $\bar{a}$. Since $C(\tilde{a},\theta)$ is continuous in η and equal to $\bar{I}$ when $\eta = 0$, there exists a $\tilde{\eta} > 0$ such that $B(\tilde{a}) - C(\tilde{a},\theta) > B(0) - \bar{I}$ for $\eta \in (0,\tilde{\eta}]$. Appealing again to continuity, $B(\tilde{a}) - C(\tilde{a},\theta) > B(a) - C(a,\theta)$ for all $a \in [0,a_\ell)$ and all $\eta \in (0,\tilde{\eta}]$. Hence, if $\eta \in (0,\tilde{\eta})$, the optimal action is in $[a_\ell,\bar{a}]$. Let $\gamma = (R(h(\bar{v}))/V'(h(\bar{v})))\psi(a_\ell) > 0$. For any $a \in [a_\ell,\bar{a}]$, we have that $\partial^2 C(a,\theta)/\partial\theta\partial\eta\ |_{\eta=0} \geq \gamma > 0$. By the continuity of this second derivative, there is an open neighborhood W_a of a and a non-empty neighborhood $U_a = (0,\eta_a]$ of values of η such that $\partial^2 C(a',\theta)/\partial\theta\partial\eta \geq \gamma/2 > 0$, for all $a' \in W_a$ and $\eta \in U_a$. Since $\partial C(a,\theta))/\partial\theta = 0$ at $\eta = 0$, it follows by the Mean Value Theorem that $\partial C(a',\theta)/\partial\theta > 0$ for all $a' \in W_a$ and $\eta \in U_a$. The compactness of $[a_\ell,\bar{a}]$ implies the existence of finitely many $a_1, a_2, \ldots, a_m$ in $[a_\ell,\bar{a}]$ such that $[a_\ell,\bar{a}] \subset \cup_{i=1}^m W_{a_i}$. Setting $\eta^* = \min\{\eta_{a_1},\eta_{a_2},\ldots,\eta_{a_m}\} > 0$ completes the proof. □

There is a clean intuition of part (*i*) which follows from the known fact in decision theory that an expected utility maximizer would be willing to take a 'small gamble' with

[8]The cost of implementing the lowest action a_1 is simply $\bar{I}$, which is trivially increasing in θ.

positive mean.[9] To see the relationship, consider (3) and let $u(\cdot) = 1/V'(h(\cdot))$. Recall that $C(a,\theta)$ is increasing in θ if $\sum_{i=1}^{n} \pi_i(a)u(v_i) - u(\bar{v}) \geq 0$. Now, this is *akin* to a decision maker with utility function $u(\cdot)$ considering a choice between $\bar{v}$ for sure, and $\bar{v}$ plus a lottery with mean $\eta\psi(a)$ (recall that the participation constraint binds), and riskiness that also vanishes as η goes to zero. By the aforementioned result, the decision maker would choose the lottery for sufficiently small η, which is tantamount to proving that $C(a,\theta)$ is strictly increasing in θ when η is sufficiently small.

Part (ii) shows that the bound found in (i) for each a can be made uniform.

The most important implication of Proposition 3 is that there cannot exist a class of utility functions of income such that the principal prefers *richer* agents. Hence, *there is no condition on $V(\cdot)$ such that the principal's cost of implementing an action is decreasing in agent's wealth for all $(\psi(\cdot), \pi(\cdot), A, q, \bar{I}, \theta)$.* The proof is simple. If such a condition existed, then the principal would prefer a richer agent when the disutility of effort has the functional form assumed in Proposition 3, i.e., parameterized by η. But the principal always prefers poorer agents when $\eta > 0$ is close to zero, contradiction.

Thus, there is no hope in finding a condition solely on $V(\cdot)$ such that the principal prefers a richer agent. This result would require restrictions on more than one primitive.

3.3 Rich Agents and Agency Costs

Suppose the pool of agents from which the principal draws the one he hires consists of fairly rich individuals. For instance, the principal could be a firm seeking to hire a CEO. In this case, when would the principal prefer a relatively poorer agent from that pool?

In this section we fix a principal-agent problem $(V(\cdot), \psi(\cdot), \pi(\cdot), A, q, \bar{I}, \theta)$ and ask whether the principal prefers a poorer agent when wealth is *sufficiently large*.

Despite the technical complexity of the analysis, we are able to provide a result that holds in a class of principal-agent problems with moral hazard that subsumes cases that are commonly used in applications.[10] We leave it as an open problem the generalization of the result beyond the class of principal-agent problems considered.

[9] For a simple proof, consider $z(\eta) = E[u(\omega + \eta\tilde{x})] - u(\omega)$, where ω is the agent's initial wealth, $u'(\cdot) > 0$, and $\tilde{x}$ is a random variable with positive mean. Then $z(0) = 0$ and $z'(0) = u'(\omega)E[\tilde{x}] > 0$.

[10] To get an idea of the difficulty, notice that θ enters in both $\bar{v}$ and v_i in equation (3). Instead, the disutility of effort, which was the focus of the previous section, only affects v_i.

To simplify the notation, set $u(\cdot) = 1/V'(h(\cdot))$. Since we know from TW that the agent prefers a poor agent when $u(\cdot)$ is convex in v, we focus on the troublesome case with $u(\cdot)$ strictly *concave* in v, or $P(I+\theta) > 3R(I+\theta)$ for all $I+\theta$, and we also assume that $V(\cdot)$ is unbounded above. Notice that $-u''(\cdot)/u'(\cdot) = (P(h(\cdot)) - 3R(h(\cdot)))/V'(h(\cdot))$.

In the next results we will make use of the following three conditions:

(a) $\lim_{v\to\infty} -u''(v)/u'(v) = 0$.

(b) For any $a \in A$ there is an optimal $(v_1, v_2, \ldots, v_n)$ with $|v_i - \bar{v}| \leq K_a$ for all i, where $K_a > 0$ is independent of $\bar{v}$.

(c) $\sup_{a\in A, a\geq\tilde{a}} K_a < \infty$ for $\tilde{a} \in A$.

Proposition 4 (Rich Agents) *(i) Assume (a) and (b). Then, for any action a above the lowest one, there is a threshold $\theta_a < \infty$ such that the cost of implementing action a strictly increases in agent's wealth for all $\theta > \theta_a$. Furthermore, if in addition condition (c) holds for some $\tilde{a} \in A$ then there is some $\theta^* < \infty$ such that the cost of implementing action a strictly increases in agent's wealth for all $a \geq \tilde{a}$ and all $\theta > \theta^*$.*

(ii) Assume (a), (b), and (c). Then the principal prefers a poorer agent when agent's wealth is sufficiently large.

Proof. (i) Let a be any action above the lowest one, and let $v_1, v_2, ...v_n$ be an optimal contract that implements a, such that $|v_i - \bar{v}| \leq K_a$, where $K_a > 0$ and does not depend on $\bar{v}$. We must show that if θ is sufficiently large, then $\sum_i \pi_i(a) u(v_i) > u(\bar{v})$.

Consider a Taylor expansion of $u(\cdot)$ around $v = \bar{v}$. Then

$$u(v_i) = u(\bar{v}) + u'(\bar{v})(v_i - \bar{v}) + \frac{1}{2}u''(\xi_i)(v_i - \bar{v})^2,$$

for some ξ_i between v_i and $\bar{v}$, $i = 1, 2, ..., n$. Multiply by $\pi_i(a)$ and sum over i to obtain

$$\sum_{i=1}^{n} \pi_i(a) u(v_i) = u(\bar{v}) + u'(\bar{v})\psi(a) + \frac{1}{2}\sum_{i=1}^{n} \pi_i(a) u''(\xi_i)(v_i - \bar{v})^2,$$

and thus

$$\sum_{i=1}^{n} \pi_i(a) u(v_i) - u(\bar{v}) = u'(\bar{v}) \left(\psi(a) + \sum_{i=1}^{n} \kappa_i \frac{u''(\xi_i)}{u'(\bar{v})} \right), \tag{7}$$

where $\kappa_i = \frac{1}{2}\pi_i(a)(v_i - \bar{v})^2$ and therefore $0 \leq \kappa_i \leq K_a^2$.

Notice that $|v_i - \bar{v}| \leq K_a$ and K_a independent of $\bar{v}$ imply that $v_i \to \infty$ as $\bar{v} \to \infty$, and so does ξ_i. Lemma 2 in the Appendix shows that $\lim_{\bar{v}\to\infty} u''(\xi_i)/u'(\bar{v}) = 0$. It follows that the expression in parenthesis in the right-side of (7) is positive for $\bar{v}$ large, i.e., above a threshold $\bar{v}_a$. Thus $u'(\bar{v})\left(\psi(a) + \sum_{i=1}^{n} \kappa_i u''(\xi_i)/u'(\bar{v})\right) > 0$ for $\bar{v} > \bar{v}_a$. As $\bar{v} = V\left(\bar{I} + \theta\right)$ and $V(\cdot)$ is unbounded, the result follows by taking $\theta_a = h(\bar{v}_a) - \bar{I}$.

Let $A = [0, \bar{a}]$ (the case with A finite is immediate). Assume that (c) holds for some $\tilde{a} \in (0, \bar{a}]$. Take $0 < K < \infty$ with $\sup_{a\in[\tilde{a},\bar{a}]} K_a \leq K$ and choose $\bar{v}^*$ large enough so that

$$\sum_{i=1}^{n} \left| \frac{u''(\xi_i)}{u'(\bar{v})} \right| < \frac{\psi(\tilde{a})}{K^2}$$

for all $\bar{v} > \bar{v}^*$ (by using K instead of K_a in the proof of Lemma 2 in the Appendix one can realize that this threshold $\bar{v}^*$ can be chosen depending only on $\tilde{a}$ and K). From (7) it follows that for any $a \in [\tilde{a}, \bar{a}]$ and $\bar{v} > \bar{v}^*$,

$$\sum_{i=1}^{n} \pi_i(a)\, u(v_i) - u(\bar{v}) \geq u'(\bar{v}) \left(\psi(\tilde{a}) - K^2 \sum_{i=1}^{n} \left| \frac{u''(\xi_i)}{u'(\bar{v})} \right| \right) > 0$$

Finally, letting $\theta^* = h(\bar{v}^*) - \bar{I}$ completes the proof of the result.

(ii) If A is finite, then the result is clear as the cost of implementing any action is increasing in θ if $\theta \in (\theta^*, \infty)$, where $\theta^* = \max_{a\in A-\{a_1\}} \theta_a$. This implies that $\max_{a\in A} B(a) - C(a, \theta)$ is decreasing in θ for all $\theta \in (\theta^*, \infty)$.

Suppose $A = [0, \bar{a}]$. One can show that there exists a constant $\Lambda < \infty$ such that $|v_i'(0)| \leq \Lambda$ for all i and for all values of θ.[11] It follows from Lemma 1 that there is an action $\tilde{a}$ such that $\partial C(a, \theta)/\partial\theta$ for all $a \in (0, \tilde{a})$ if $P(\bar{I} + \theta) < 3R(\bar{I} + \theta) + (\psi''(0)V'(\bar{I} + \theta)/\Lambda^2)$, which can be written as

$$\frac{P(h(\bar{v})) - 3R(h(\bar{v}))}{V'(h(\bar{v}))} < \frac{\psi''(0)}{\Lambda^2}. \tag{8}$$

By condition (a) the left-side of (8) goes to zero as θ goes to infinity. Thus, there exists a threshold θ_1 such that if $\theta > \theta_1$, then $\partial C(a, \theta)/\partial\theta > 0$ for all $a \in (0, \tilde{a})$. Consider now the interval $[\tilde{a}, \bar{a}]$. We know from Proposition 4 that there exists a threshold $\theta_2 <$

[11] The proof is available from the authors upon request. It uses the property that $v_i'(0)$, $i = 1, ..., n$, is the solution of a linear system of equations generated by the derivatives of a Lagrangian function.

∞ such that if $\theta > \theta_2$, then $\partial C(a,\theta)/\partial\theta > 0$, for each $a \in [\tilde{a}, \bar{a}]$. This shows that $\max_{a\in A} B(a) - C(a,\theta)$ is decreasing in θ for all $\theta \in (\theta^*, \infty)$, where $\theta^* = \max\{\theta_1, \theta_2\}$. $\square$

For an intuition, consider once again a decision maker with utility function $u(\cdot) = 1/V'(h(\cdot))$ facing a choice between $\bar{v}$ and $\bar{v}$ plus a lottery with mean $\psi(a)$ and riskiness that is independent of $\bar{v}$. As wealth grows large, the agent becomes *approximately* risk neutral, and the positive mean of the lottery eventually dominates its riskiness.

Proposition 4 provide conditions under which the principal favors a relatively poorer agent when the pool of agents consists of wealthy ones. To be sure, its usefulness hinges on the plausibility of conditions (a)–(c). It is easy to show that (a) is satisfied by *any* HARA utility function, which subsumes *all* the standard utility functions used in applications. But conditions (b)–(c) are more delicate, as they refer to a property of the optimum. Since little is known about the optimal contract's *functional form* in the principal-agent model with moral hazard, it is unclear when they would hold.

This difficulty notwithstanding, we can show that there is at least one important class of principal-agent problems commonly used in applications that satisfy these conditions. To wit, assume condition (a) (e.g., let $V(\cdot)$ be HARA), $n = 2$, and let $\pi(\cdot)$ satisfy MLRP and CDFC. It is easy to verify that, under these assumptions, conditions (b)–(c) hold when the action set is finite or is an interval.

If $A = \{a_1, a_2, ..., a_m\}$, and the principal wants to implement $a_k > a_1$, then, under the general assumptions made in Section 2 plus MLRP and CDFC, only the incentive constraint corresponding to a_{k-1} binds (Grossman and Hart (1983) p. 34). Thus, the optimal contract to implement a_k is

$$\begin{aligned} v_1 &= \bar{v} - \frac{\pi_2(a_k)}{\pi_2(a_k) - \pi_2(a_{k-1})}\psi(a_k), \\ v_2 &= \bar{v} + \frac{1 - \pi_2(a_k)}{\pi_2(a_k) - \pi_2(a_{k-1})}\psi(a_k). \end{aligned}$$

Since $\pi_2(\cdot)$ is strictly increasing, it follows that $\pi_2(a_k) - \pi_2(a_{k-1}) > 0$. Thus, we can set $K_{a_k} = \psi(a_k)\max\{\pi_2(a_k)/(\pi_2(a_k) - \pi_2(a_{k-1})), (1-\pi_2)(a_k)/(\pi_2(a_k) - \pi_2(a_{k-1}))\} > 0$. Clearly, $\sup_{a_k\in A} K_{a_k} < \infty$, thereby showing that (b)–(c) hold.

Assume now $A = [0, \bar{a}]$. As in the proof of Proposition 2, the optimal contract that implements an action $a > 0$ is given by equations (4)–(5). Thus, we can set $K_a = \max\{\psi(a) + \pi_2(a)(\psi'(a)/\pi_2'(a)), \psi(a) + (1 - \pi_2(a))(\psi'(a)/\pi_2'(a))\} > 0$. The continuity of the functions involved in the definition of K_a and the fact that $\pi_2'(a) > 0$, yield $\sup_{a \in A} K_a < \infty$. Hence, (b)–(c) hold, and we have thus proved the following result:

Corollary 1 (Two-Outcome Case) *Assume (a). If $n = 2$ and MLRP and CDFC hold, then the principal prefers a poorer agent when wealth is sufficiently large.*

To shed further light on the results of this section, let us contrast them with the case of small disutility of effort. Suppose first that $V(I) = I^\alpha$, $\alpha \in (0.5, 1)$, $A = [0, \bar{a}]$, $\tilde{\psi}(a) = \eta\psi(a)$ and $n = 2$. Easy algebra (using (3) and (4)–(5)) reveals that in this case (3) depends on $\eta/\bar{v}$. Thus, for this utility function (the same is true for any HARA one) and with two outcomes, Proposition 4 can be obtained as a corollary of Proposition 3: 'small' values of η have the same effect as 'large' values of θ (and hence of $\bar{v}$).

To be sure, this special case exploits the power functional form of the utility function and also the property that, in the two-outcome case, the riskiness of the contract $|v_i - \bar{v}|$ is *independent* of $\bar{v}$. Assumptions (a)–(c) ensure that this logic extends beyond this example to principal-agent problems where the riskiness of the contract is *uniformly bounded* by an expression that does not depend on $\bar{v}$. As Corollary 1 reveals, this class contains interesting cases often used in the literature.

4 Concluding Remarks

Using the textbook principal-agent model with moral hazard, we analyze the principal's preferences over agents of differing wealth. First, we show that TW's condition for the principal to prefer a poorer agent is tight. Second, we prove that the principal always prefers a poorer agent if the task involved entails a small disutility. This result implies that there is no analogue of TW's condition such that he would prefer a richer agent. Third, we show that for an important class of problems, if agents are rich enough, then the principal prefers a relatively poorer one.

We focused on the risk neutral principal case, which is the standard assumption in applications. It would be interesting to know if the results generalize to the case where the principal is risk averse, an extension that appears to be challenging.[12]

[12] Let the principal's utility function be $U(q - I)$, with $U'(\cdot) > 0$ and $U''(\cdot) < 0$. He prefers a poorer agent if $\sum_{i=1}^{n} \pi_i(a) U'(q_i - I_i) \left(\frac{V'(\bar{I}+\theta)}{V'(I_i+\theta)} - 1 \right) \geq 0$. Notice that TW's condition does not suffice anymore,

References

AMIR, R., AND M. CZUPRYNA (2004): "On Inverse Utility and Third-Order Effects in the Economics of Uncertainty," CORE Working Paper.

CHADE, H. (2009): "Moral Hazard with Unobservable Consumption," Arizona State University Working Paper.

CHIAPPORI, P.-A., I. MACHO, P. REY, AND B. SALANIÉ (1994): "Repeated Moral Hazard: The Role of Memory, Commitment, and the Access to Credit Markets," *European Economic Review*, 38, 1527–1553.

CHIU, H. (2010): "Non-negativity of Agent's Payoffs and Wealth Effects in the Principal-Agent Model," University of Manchester Working Paper.

GROSSMAN, S., AND O. HART (1983): "An Analysis of the Principal-Agent Problem," *Econometrica*, 51, 7–45.

HOLMSTROM, B. (1979): "Moral Hazard and Observability," *Bell Journal of Economics*, 10, 74–91.

HOPENHAYN, H., AND J. P. NICOLINI (1997): "Optimal Unemployment Insurance," *Journal of Political Economy*, 105, 412–438.

JEWITT, I., O. KADAN, AND J. SWINKELS (2008): "Moral Hazard with Bounded Payments," *Journal of Economic Theory*, 143, 59–82.

PARK, I.-U. (2004): "Moral Hazard Contracting and Private Credit Markets," *Econometrica*, 72, 701–746.

ROGERSON, W. (1985): "The First-Order Approach to Principal-Agents Problems," *Econometrica*, 53, 1357–1367.

SPEAR, S., AND C. WANG (2005): "When to Fire a CEO: Optimal Termination in Dynamic Contracts," *Journal of Economic Theory*, 120, 239–256.

THIELE, H., AND A. WAMBACH (1999): "Wealth Effects in the Principal Agent Model," *Journal of Economic Theory*, 89, 247–260.

---

for the marginal utility of the principal is decreasing, and it is not known in general what the correlation between $q_i - I_i$ and I_i is at the optimum (see Grossman and Hart (1983), Proposition 4).

A Appendix

Let $A = [0, \bar{a}]$. With some abuse of notation, denote by $v_i(a)$, $i = 1, ..., n$ the optimal contract that implements a, and by $v_i'(a)$ and $v_i''(a)$ their derivatives with respect to a.

Lemma 1 *Assume that $v_i(\cdot)$ is twice continuously differentiable in a, $i = 1, 2, ..., n$. Then there exists an action $\tilde{a} > 0$ such that $\partial C(a, \theta)/\partial \theta > 0$ for all θ and $a \in (0, \tilde{a})$ if*

$$P(\bar{I} + \theta) < 3R(\bar{I} + \theta) + \frac{\psi''(0)V'(\bar{I} + \theta)}{\sum_{i=1}^{n} \pi_i(0)(v_i'(0))^2}. \tag{9}$$

Conversely, if there is such an action $\tilde{a}$, then (9) holds with the weak inequality $\leq$.[13]

Proof. Set $z(a) = \sum_{i=1}^{n}(\pi_i(a)/V'(h(v_i(a))))$, and note that $z(0) = 1/V'(h(\bar{v}))$. Thus, $\partial C(a, \theta)/\partial \theta$ is positive for all θ near $a = 0$ if and only if $z(a) > z(0)$ for a close to zero.

Differentiating $z(a)$, we obtain (after some algebra):

$$z'(a) = \sum_i \pi_i'(a) \frac{1}{V'(h(v_i(a)))} + \sum_i \pi_i(a) \frac{R(h(v_i(a)))}{(V'(h(v_i(a))))^2} v_i'(a).$$

Notice that $z'(0) = 0$. For

$$z'(0) = \frac{1}{V'(h(\bar{v}))} \sum_i \pi_i'(0) + \frac{R(h(\bar{v}))}{(V'(h(\bar{v})))^2} \sum_i \pi_i(0) v_i'(0),$$

and $\sum_i \pi_i'(0) = 0$ while $\sum_i \pi_i(0)v_i'(0) = (\sum_i \pi_i(a)v_i(a))'|_{a=0} = (\bar{v} + \psi(a))'|_{a=0} = 0$, where the second equality follows from the participation constraint.

Thus, to assess the behavior of $z(a)$ near $a = 0$, we need to look at the second derivative $z''(0)$. If this derivative is positive at $a = 0$, then by continuity it will be positive in a right neighborhood of zero, i.e., for a sufficiently small.

Differentiating $z'(a)$, we obtain after some tedious algebra:

$$\begin{aligned} z''(0) &= 2\frac{R(h(\bar{v}))}{(V'(h(\bar{v})))^2} \sum_i \pi_i'(0)v_i'(0) + \frac{R(h(\bar{v}))}{(V'(h(\bar{v})))^2} \sum_i \pi_i(0)v_i''(0) \\ &\quad + \frac{R(h(\bar{v}))(3R(h(\bar{v})) - P(h(\bar{v})))}{(V'(h(\bar{v})))^3} \sum_i \pi_i(0)(v_i'(0))^2. \end{aligned} \tag{10}$$

[13] If there are only two outcomes, then (4)–(5) yield $\sum_i \pi_i(0)(v_i'(0))^2 = \pi_2(0)(1 - \pi_2(0))(\psi''(0)/\pi'(0))^2$, and (9) reduces to (6), which is the expression used in the proof of Proposition 2.

Differentiation of the participation constraint $\sum_i \pi_i(a)v_i(a) = \bar{v}+\psi(a)$ yields $\sum_i \pi_i'(a)v_i(a)+\sum_i \pi_i(a)v_i'(a) = \psi'(a)$. Since the incentive constraint is $\sum_i \pi_i'(a)v_i(a) = \psi'(a)$, it follows that $\sum_i \pi_i(a)v_i'(a) = 0$, and its derivative yields $\sum_i \pi_i(a)v_i''(a) = -\sum_i \pi_i'(a)v_i'(a)$. Thus,

$$z''(0) = \frac{R(h(\bar{v}))}{(V'(h(\bar{v})))^2}\sum_i \pi_i'(0)v_i'(0) + \frac{R(h(\bar{v}))(3R(h(\bar{v})) - P(h(\bar{v})))}{(V'(h(\bar{v})))^3}\sum_i \pi_i(0)(v_i'(0))^2.$$

The derivative of the incentive constraint is $\sum_i \pi_i''(a)v_i(a)+\sum_i \pi_i'(a)v_i'(a) = \psi''(a)$, which converges to $\sum_i \pi_i'(0)v_i'(0) = \psi''(0)$ as a goes to zero. Then $z''(0)$ becomes

$$z''(0) = \frac{R(h(\bar{v}))}{(V'(h(\bar{v})))^2}\left(\psi''(0) + \frac{(3R(h(\bar{v})) - P(h(\bar{v})))}{V'(h(\bar{v}))}\sum_i \pi_i(0)(v_i'(0))^2\right). \tag{11}$$

Since $h(\bar{v}) = \bar{I} + \theta$, (11) implies that $z''(0) > 0$ if and only if (9) holds. Also, $z''(0) > 0$ implies that $z'(a) > 0$ for a near zero. Thus, there is an $\tilde{a} > 0$ such that $\partial C(a,\theta)/\partial\theta > 0$ for all θ if $a \in (0,\tilde{a})$. To prove the converse, note that if (9) did not hold with sign $\leq$, then $z''(0) < 0$, and thus $\partial C(a,\theta)/\partial\theta$ would be negative for a near zero. □

Lemma 2 *Under the conditions of Proposition 4,* $\lim_{\bar{v}\to\infty} u''(\xi_i)/u'(\bar{v}) = 0$, $i = 1, 2, ..., n$.

Proof. Let $0 < \varepsilon < (2K_a)^{-1}$. Since $\lim_{v\to\infty} u''(v)/u'(v) = 0$, there exists an $M > 0$ such that $|u''(v)/u'(v)| < \varepsilon/2$ for $v > M$. Suppose that $\bar{v} > M + K_a$. If $\bar{v} \leq \xi_i \leq v_i$, then $u'(\xi_i)/u'(\bar{v}) \leq 1$, as $u'(\cdot)$ is positive and decreasing. If $v_i \leq \xi_i \leq \bar{v}$, apply a linear Taylor expansion to $u'(\cdot)$ around $v = \bar{v}$ to obtain $u'(\bar{v} - K_a) = u'(\bar{v}) + u''(\delta)K_a$, for some δ, $\bar{v} - K_a < \delta < \bar{v}$. Using $|u''(v)/u'(v)| < \varepsilon/2$ and $0 < \varepsilon < (2K_a)^{-1}$, we obtain

$$|u'(\bar{v} - K_a) - u'(\bar{v})| = |u''(\delta)| K_a < \varepsilon u'(\delta) K_a < \frac{1}{2}u'(\delta) < \frac{1}{2}u'(\bar{v} - K_a),$$

which gives $|1 - (u'(\bar{v})/u'(\bar{v} - K_a))| < 1/2$ and thus $2/3 < u'(\bar{v} - K_a)/u'(\bar{v}) < 2$. In particular, if $\bar{v} - K_a \leq v_i \leq \xi_i \leq \bar{v}$, then $0 < u'(\xi_i)/u'(\bar{v}) \leq u'(\bar{v} - K_a)/u'(\bar{v}) < 2$.

Therefore, for $\bar{v} > M + K_a$,

$$\left|\frac{u''(\xi_i)}{u'(\bar{v})}\right| = \left|\frac{u''(\xi_i)}{u'(\xi_i)}\right|\frac{u'(\xi_i)}{u'(\bar{v})} \leq 2\left|\frac{u''(\xi_i)}{u'(\xi_i)}\right| < \varepsilon,$$

where the last inequality follows from $\lim_{\bar{v}\to\infty} u''(\xi_i)/u'(\xi_i) = 0$. This proves that $\lim_{\bar{v}\to\infty} u''(\xi_i)/u'(\bar{v}) = 0$, $i = 1, 2, ..., n$. □