ADVERSE SELECTION WITH MULTIPLE INPUTS: MITIGATING INFORMATION RENTS THROUGH INPUT CONTROL

RACHAEL E. GOODHUE AND LEO K. SIMON

FEBRUARY 2, 2004

ABSTRACT. In a principal-agent relationship, inputs that could be chosen by either party are often controlled by the principal. In the presence of adverse selection, the principal's profits are always higher when she controls an input than when she does not. Output is higher when she controls the input, since the second-best input specification reduces information rents. However, this mitigation distorts input use away from the production cost-minimizing level, which is socially costly. The net effect of this tradeoff on social welfare depends primarily on the elasticity of substitution between inputs: the restrictive contract results in higher social surplus than the basic contract if the elasticity of substitution between the inputs is sufficiently small. When the elasticity of substitution is large, the net effect is determined by secondary factors.

JEL Classification: D82 Asymmetric and Private Information

Keywords: Mechanism Design, Asymmetric Information, Input Control

Acknowledgements: We are indebted to Corinne Alexander, Peter Bogetoff, Robert Chambers, Bruno Jullien, Erik Lichtenberg and Jeffrey Williams for their thoughtful comments. We thank Jacques Crémer for suggesting the term "restrictive contract."

Please address correspondence to:
Rachael Goodhue
Department of Agricultural and Resource Economics
One Shields Avenue
University of California, Davis
Davis, CA 95616-8512
Tel: (530) 754 7812
Fax: (530) 752 5614
Email: goodhue@primal.ucdavis.edu

The authors are, respectively, Assistant Professor at the University of California at Davis and and Adjunct Professor at the University of California at Berkeley. Goodhue and Simon are members of the Giannini Foundation of Agricultural Economics.

1. INTRODUCTION

In contractual relationships involving multiple inputs, some of the inputs may be controlled by the principal, though they could just as easily be chosen by the agent. For example, construction contracts may specify building materials. Military procurement contracts may specify component materials. Agricultural contracts between farmers and processors may specify allowable fertilizers, seedstock, and other production inputs. While there are a number of reasons why a principal may seek to control inputs, this paper focuses on an information-driven motivation: the reduction of information rents arising from adverse selection.

We examine the case in which output is a function of the agent's type, as well as the levels of capital and effort that he utilizes. The agent's type, which is his private information, affects the productivity of a non-labor input, which is observable, and his effort, which is not. Our conceptualization of input specification by the principal can be viewed as encompassing two cases: the principal either simply provides the nonlabor input to the agent, or else specifies the required level of this input in the contract, and then verifies it.

In a principal-agent relationship involving multiple inputs, other potential information problems may arise. For example, the principal may be less informed than the agent regarding the precise nature of the production function. When information is asymmetric in this sense, it may be costly for the principal to determine the input mix: she may select a combination that the agent could improve upon. Alternatively, the agent's choice of inputs may provide information regarding his type. In this paper, we assume that the principal and agent are both fully informed about the production function for any given agent type, so that the only information asymmetry regarding the production function is that the principal does not know the agent's type.

We show that by controlling the non-labor input, the principal can reduce information rents. Further, the principal's optimal contract menu will always result in higher profits when she controls this input (the *restrictive* contract), relative to when she does not (the *basic* contract), even though she always chooses an input mix which is sub-optimal from a pure production standpoint. Moreover, output will be less distorted relative to the first-best under the restrictive contract than under the basic contract. The consequences for society as a whole are less clear. In order to evaluate these consequences, we develop a construction which continuously varies the degree of asymmetric information in the principal's maximization problem. We establish that if the elasticity of substitution between effort and the other input is sufficiently small, the restrictive contract will result in higher social surplus than the basic contract. If the elasticity of substitution is large, the welfare comparison depends on the relative importance of the output and the input-mix distortions.

Our findings provide a strong motivation for the principal to integrate; that is, one must look outside our framework to justify a decision by the principal to allow the agent to make input allocation decisions, rather than to control these decisions herself. Conversely, our framework suggest why agents may resist integration attempts by principals. For example, a possible explanation of the widespread opposition by farmers to increased integration between agricultural production and food processing is that high-ability farmers do not want the information rents they can command to be reduced by greater processor control over the production process. Conversely, our model may explain why processors are choosing to source an increasing share of their purchases through integrated production, and contracts specifying non-labor inputs: by doing so, processors can reduce the information rents that they need to pay.

Under strong pressure from farmers, Congress and some state legislatures are considering laws that would limit integration and the use of contracts in the production of products such as beef. Clearly, the reduced information rents that farmers would receive under these new laws provide them with an incentive to lobby. Our framework allows us to assess the social desirability of such changes. Because feeder cattle are extremely complementary with farmer effort in the production of beef, these proposals are almost certainly not socially desirable in the context of our model. To justify them on economic grounds, it would be necessary to identify some other kind of market failure, whose impact would be significantly mitigated by contracted and integrated production.

Other research addresses the principal's choice between monitoring output and monitoring agent effort when both are feasible but costly (Maskin and Riley 1985, Khalil and Lawarree 1995). This work has implicitly assumed that there is no substitutability between effort and other inputs that may be exploited by the principal. In our case, in contrast, we incorporate a non-labor input, substitutable for effort in production, which the principal may observe in addition to observing output. Similarly, other analyses, such as Laffont and Tirole (1986), assume that fixed costs (capital levels, in our framework), and required effort levels are completely independent of each other.

Another approach models agent effort as a substitute for purchased inputs, including capital and labor (Baron and Besanko 1987, Lewis and Sappington 1991). Unlike our approach, these articles incorporate costly monitoring. Dewatripont and Maskin (1995) address the principal's optimal monitoring choice when monitoring is costless and agent effort is a substitute for purchased inputs. They compare a number of monitoring possibilities: monitoring only capital, labor, or total cost savings, monitoring both inputs, or monitoring cost savings and capital. They find that monitoring total cost savings after production has occurred dominates monitoring either purchased input for the principal, and dominates monitoring both total cost savings and capital. This second result is the opposite of ours. Their findings are driven by two assumptions that differ from ours: the principal's observation of non-labor inputs chosen by the agent and the timing of their model, especially the possibility of renegotiation after the agent chooses capital but before production occurs. We assume that capital and effort are chosen simultaneously, and that production is realized without an opportunity for renegotiation.

Our analysis is related to two broader research areas: delegation and control, and property rights. The delegation and authority literature focuses on control issues, as do we. This literature has two common assumptions. First, the agent has superior information regarding a choice, whether pre-specified (Dessein 2002) or due to endogenous agent action (Aghion and Tirole 1997). Second, there is a divergence in preferences between the principal and the agent regarding the choice itself. For example, in Dessein (2002), the agent prefers projects that are larger than the one that is optimal for the principal. In the presence of these two assumptions, there is a tradeoff for the principal between delegating the decisionmaking authority to the agent—who then uses his superior information to choose the project—and communicating with the agent to elicit his information prior to choosing the project. In our model, by contrast, it is never privately optimal for the principal to "delegate" control over capital use decisions to the agent.

We address a question that is critical within the property rights literature: when should a firm integrate, and hire labor to work with its own capital, rather than contract with another firm with independent control over labor and capital (Coase 1937, Williamson 1985)? We find that input control by the principal is socially optimal, provided that effort and capital are sufficiently complementary. This result is consistent with Hart and Moore (1990), who find—for a quite different reason—that when assets are highly complementary with an agent's effort, the asset should be controlled by that agent (or one of his essential trading partners) in order to promote efficiency. More broadly, it is consistent with one of the themes of the property rights literature: the allocation of decision rights affects the ex post efficiency of production (Tirole 1999).

The paper is organized as follows. Section Two introduces the modeling framework, establishes the properties of the second best basic and restrictive contracts, and proves that the principal's profits are higher under the restrictive than the basic contract. Section Three introduces our technique of continuously varying the degree of information asymmetry, then uses this technique to compare quantity provision under the two contracts. Section Four compares social surplus under the two optimal contracts. Section Five concludes.

2. THE MODEL

We begin with a standard principal-agent model. The agent may be one of two types; each type has access to a distinct production function, and one type's function is more productive than the other. Both principal and agent are perfectly informed about the specification of these functions and the probability distribution over types. The agent's realized type, however, is unknown to the principal. The principal's goal is to maximize her profits from production, which depend on the agent's production possibilities. To induce the agent to reveal his true type, she must offer him a menu of contracts that provides him with adequate incentives to do so. We assume, as is the convention in models of this type, that the principal cannot observe the level of effort supplied by the agent. We compare two cases: one in which capital is non-observable and non-verifiable, and one in which it is observable and verifiable. For expository convenience, we refer to non-observable capital as capital supplied by the agent, and observable capital as capital supplied by the principal. We assume that capital is homogeneous, so that only the level of capital and the production set available to the agent who uses it are relevant to production. In this section, we formally develop the components of our analysis, and examine the principal's problem when she can and cannot specify capital.

The Production Function: Production depends on capital (k), effort (e) and the agent's type (θ). There are two types, "low" and "high", denoted by ℓ and h. The agent's true type is $\theta \in \{\theta^\ell, \theta^h\}$. Let $\Theta = \frac{\theta^\ell}{\theta^h} < 1$. For $i = \ell, h$, let ϕ^i denote the probability that an agent's type is i. (Obviously, $\phi^\ell + \phi^h = 1$.)

We impose the following additional assumptions on the production function, $f(e,k;\theta)$.

- A1: There exists a twice continuously differentiable function, $g : \mathbb{R}^2 \to \mathbb{R}$, such that for $\theta \in \{\theta^\ell, \theta^h\}$, $f(\cdot,\cdot,\theta) = \theta g$;
- A2: The function g is homogeneous of degree $\alpha < 1$ in e and k;
- A3: The marginal products of effort and capital, g_e and g_k, are positive and g is strictly concave in (e,k), (i.e., $g_{ee}, g_{kk} < 0$ and $g_{ee}, g_{kk} > (g_{ek})^2$).

For $i = \{\ell, h\}$, we shall usually write f^i rather than $f(\cdot,\cdot,\theta^i)$. For convenience, we will normalize by setting $\theta^\ell = 1$, so that $f^\ell \equiv g$. Assumption A1 implies that θ is "technologically neutral," in the sense that for each e and k, $\frac{f_e^\ell(e,k)}{f_k^\ell(e,k)} = \frac{f_e^h(e,k)}{f_k^h(e,k)}$. Assumption A1 and A2 are much stronger than we need, but the computational convenience that these assumptions provide amply compensates for the loss of generality.

Agent's Utility Function: The agent will receive a lump-sum transfer payment from the principal and in return will deliver a specified level of output, contributing effort and, in one of our two cases, capital. The agent's outside alternative is to provide his effort at the given wage-rate $w > 0$ per unit effort supplied. The wage rate exactly compensates for the agent's constant marginal disutility of effort, so that his reservation utility when he does not supply effort is zero. The price of capital is constant at $r > 0$ per unit. In order to induce the agent to supply effort level e and capital level k, the principal's transfer payment must at least cover the agent's cost, $we + rk$.

Input levels: If an agent of type θ chooses input levels to produce a given q—i.e., if the principal cannot observe and verify capital levels—he will solve the (neo-classical) cost minimization problem $\min_{e,k} \; we + rk$ s.t. $f(e,k,\theta) = q$. Let $\left(\tilde{e}(q,\theta), \tilde{k}(q,\theta)\right)$ denote the solution to this problem. We will refer to this input vector as the *neoclassical input mix for q*. The solution to the agent's problem exhibits the following, well-known properties:

Remark 1. *The neoclassical input mix is uniquely defined by the first order condition:*

$$0 \quad = \quad w f_k\left(\tilde{e}(q,\theta), \tilde{k}(q,\theta), \theta\right) \quad - \quad r f_e\left(\tilde{e}(q,\theta), \tilde{k}(q,\theta), \theta\right) \tag{1}$$

Moreover, there exists a constant $\tilde{\beta}$ *such that for all* q, $\frac{\tilde{k}(\cdot,\theta^\ell)}{\tilde{e}(\cdot,\theta^\ell)} = \tilde{\beta}$. *Finally, for all* q, $\left(\tilde{e}(q,\theta^\ell), \tilde{k}(q,\theta^\ell)\right) = \Theta^{1/\alpha}\left(\tilde{e}(q,\theta^h), \tilde{k}(q,\theta^h)\right)$.

Uniqueness follows from the strict concavity of g (A3). Linearity of agent ℓ's expansion path follows from homogeneity (A2) and the fact that r and w are constants. The proportionality relationship between different types' input vectors follows from A1 and A2.

Let $\tilde{C}^P(q,\theta)$ denote the type θ agent's *production cost* of delivering the output level q with the neoclassical input mix:

$$\tilde{C}^P(q,\theta) \quad = \quad w\tilde{e}(q,\theta) \quad + \quad r\tilde{k}(q,\theta) \tag{2}$$

An immediate consequence of Remark 1 is that when the agent chooses inputs, there is an equivalent, *single-input* characterization of technology, $\ddot{f}(\ddot{e},\theta) = \ddot{e}^\alpha f(1,\tilde{\beta},\theta)$, and of production costs, $\ddot{C}^P(q,\theta) = \ddot{v}\ddot{e}(q,\theta)$, where each unit of the *composite* input $\ddot{e}$ is composed of one unit of e and $\tilde{\beta}$ units of k, and $\ddot{v} = (v + \tilde{\beta}r)$ denotes the unit cost of $\ddot{e}$. Provided that the input mix is always neo-classical, this alternative characterization is equivalent to the original one in the following sense:

Remark 2.[1] *For each q and* θ, $\ddot{f}\left(\ddot{e}(q,\theta),\theta\right) \quad = \quad f\left(\tilde{e}(q,\theta),\tilde{k}(q,\theta),\theta\right)$ *and* $\ddot{C}^P(q,\theta)=\tilde{C}^P(q,\theta)$.

The significance of Remark 2 is that when the agent chooses inputs, the principal's problem in our two-input model is formally equivalent to the corresponding, and routine, textbook problem, in which technology is given by the single-input production function $\ddot{f}$ with constant unit cost $\ddot{v}$.[2]

Now suppose that the principal chooses the level of capital and let $\bar{e}$ denote the level of effort required to produce q given k and θ, and let $\bar{C}^P(q,k,\theta)$ denote the type θ agent's production cost of delivering the output level q with capital level k:

$$\bar{C}^P(q,k,\theta) \quad = \quad w\bar{e}(q,k,\theta) \quad + \quad rk \tag{3}$$

An obvious implication of (A1) and (A2) for $\bar{C}^P$ is that

$$\bar{C}^P(q,k,\theta^\ell) \quad > \quad \bar{C}^P(q,k,\theta^h), \text{for all } q \text{ and all } k. \tag{4}$$

For future reference, note that by definition of $\tilde{k}(q,\theta)$

$$\frac{\partial\bar{C}^P(q,\tilde{k}(q,\theta),\theta)}{\partial k} \quad = \quad 0, \quad \text{for all } q \text{ and all } \theta. \tag{5}$$

<u>Contracts</u>: A *basic contract* assigns to each announced type θ^i, $i\in\{\ell,h\}$, an output level and transfer, $\left(\tilde{q}(\theta^i),\tilde{t}(\theta^i)\right)$. We will sometimes write the contract $\left\{\left(\tilde{q}(\theta^i),\tilde{t}(\theta^i)\right)\right\}_{i=\ell,h}$ as $(\tilde{\mathbf{q}},\tilde{\mathbf{t}}) = \left((\tilde{q}^\ell,\tilde{t}^\ell),(\tilde{q}^h,\tilde{t}^h)\right)$. A *restrictive contract* assigns to each θ^i an output level, capital level and transfer. We will similarly sometimes write $\left\{\left(\bar{q}(\theta^i),\bar{k}(\theta^i),\bar{t}(\theta^i)\right)\right\}_{i=\ell,h}$ as $(\bar{\mathbf{q}},\bar{\mathbf{k}},\bar{\mathbf{t}}) = \left((\bar{q}^\ell,\bar{k}^\ell,\bar{t}^\ell),(\bar{q}^h,\bar{k}^h,\bar{t}^h)\right)$. Invoking the Revelation Principle (Myerson 1979), we restrict our analysis to truth-telling contracts, in which each agent chooses to announce his true type.

<u>Timing and Information</u>: Regardless of contract type, the timing of the game is as follows: The principal offers a contract menu to the agent on a take-it-or-leave-it basis. At the time the contract is offered, the agent's type is his private information. The probability of each type occurring is common knowledge. We assume that if the agent is indifferent between accepting and not accepting a contract that he will accept

[1] Because f is homogeneous of degree α (A2), $\ddot{f}\left(\ddot{e}(q,\theta),\theta\right) = \ddot{e}(q,\theta)^\alpha f(1,\tilde{\beta},\theta) = \ddot{e}(q,\theta)^\alpha\left[\tilde{e}(q,\theta)^{-\alpha}f\left(\tilde{e}(q,\theta),\tilde{k}(q,\theta)\right)\right] = f\left(\tilde{e}(q,\theta),\tilde{k}(q,\theta)\right)$.

[2] For an analysis of this problem see Caillaud and Hermalin (2000). A closely related problem is analyzed in chapter 2 of Salanie (1997).

the contract. Similarly, we assume if he is indifferent between the two contracts he will choose the contract intended for his true type. Production and transfers are then implemented as per contract specifications.

Symmetric Information Benchmark: We assume throughout that output is sold on a perfectly competitive market at a price of p. For $\theta \in \{\theta^\ell, \theta^h\}$, let $q^\star(\theta)$ denote the level of output satisfying $\frac{d\tilde{C}^P(q^\star(\theta),\theta)}{dq} = p$. Also, let $(e^\star(\theta), k^\star(\theta)) = (\tilde{e}(q^\star(\theta),\theta), \tilde{k}(q^\star(\theta),\theta))$ denote the neoclassical input mix for $q^\star(\theta)$. If information were symmetric, i.e., if the principal *were* able to observe the agent's type, the (first best) solution to the principal's problem would be to specify the output pair $q^\star(\cdot)$, whether or not she chose capital levels. Regardless of who chose the level of capital, the agent of type θ would then produce $q^\star(\theta)$ with inputs $(e^\star(\theta), k^\star(\theta))$. We shall refer to $(q^\star, e^\star, k^\star)$ as the *symmetric information benchmark solution*. Note that under assumptions A1-A3, the benchmark solution has both types producing a positive quantity.

The principal's problem: basic contracts: Given a basic contract $(\tilde{\mathbf{q}}, \tilde{\mathbf{t}})$, the principal's profit from an agent who declares a type of $\hat{\theta}$ is $p\tilde{q}(\hat{\theta}) - \tilde{t}(\hat{\theta})$. Thus, the principal's problem is to choose the contract $(\tilde{\mathbf{q}}, \tilde{\mathbf{t}})$ that maximizes $\sum_{i\in\{\ell,h\}} \left\{\phi^i \left(p\tilde{q}(\theta^i) - \tilde{t}(\theta^i)\right)\right\}$ subject to incentive and participation constraints. As a consequence of Remark 2, we can characterize the optimal basic contract by drawing on standard results from the mechanism design literature in which production is a function of the agent's effort. The essence of these results is that the constraints that are binding on the principal are type h's incentive compatibility constraint and type ℓ's individual rationality constraint. Consequently, h and ℓ produce, respectively, at and below the symmetric information benchmark solution levels for their types. Moreover, the difference between the transfer offered to ℓ and ℓ's production cost of delivering the designated output level will just equal ℓ's reservation utility, which in our model is zero. On the other hand, the transfer offered to h includes a premium, referred to as his *information rent*, which in the optimal contract will be just sufficient to offset the increment in utility that h would derive by adopting ℓ's contract rather than the one intended for him. In symbols:

Proposition 1. *The optimal basic contract has the following properties:*

(1) agent ℓ produces $\tilde{q}^\ell < q^\star(\theta^\ell)$ and receives a transfer of

$$\tilde{t}^\ell = \tilde{C}^P(\tilde{q}^\ell, \theta^\ell) \tag{6-$\tilde{t}^\ell$}$$

(2) agent h produces $\tilde{q}^h = q^\star(\theta^h)$ and receives a transfer of

$$\tilde{t}^h = \tilde{C}^P(\tilde{q}^h, \theta^h) + \left(\tilde{C}^P(\tilde{q}^\ell, \theta^\ell) - \tilde{C}^P(\tilde{q}^\ell, \theta^h)\right) \tag{6-$\tilde{t}^h$}$$

Proof of Proposition 1: Standard, given Remark 2.[3] □

In what follows, we shall sometimes use the terminology *production costs* and *information costs* to distinguish between costs incurred through production and costs (usually called rents) paid out to ensure truthful revelation. The terms "marginal production" and "marginal information" costs will then have the obvious interpretation.

The principal's problem: restrictive contracts: Given a restrictive contract $(\overline{\mathbf{q}},\overline{\mathbf{k}},\overline{\mathbf{t}})$, the principal's profit from an agent who declares a type of $\hat{\theta}$ is $p\overline{q}(\hat{\theta}) - \overline{t}(\hat{\theta})$. Thus, the principal's problem is to choose the contract $(\overline{\mathbf{q}},\overline{\mathbf{k}},\overline{\mathbf{t}})$ that maximizes $\sum_{i\in\{\ell,h\}}\left\{\phi^i\left(p\overline{q}(\theta^i) - \overline{t}(\theta^i)\right)\right\}$ subject to incentive and participation constraints.

Under a restrictive contract, the input mix is no longer exogenous to the principal's decision. Consequently, the textbook single-input model is no longer of much help to us in characterizing the optimal restrictive contract. Nonetheless, it turns out that this contract, $(\overline{\mathbf{q}},\overline{\mathbf{k}},\overline{\mathbf{t}})$, does share many of the properties of the single-input problem. Once again, the constraints that are binding on the principal are type h's incentive compatibility constraint and type ℓ's individual rationality constraint. Once again, h and ℓ produce, respectively, at and below the benchmark levels for for their types. Once again, agent ℓ receives no information rents while h receives rent equal to the increment in utility that h would derive by adopting ℓ's contract rather than his own. Formally,

Proposition 2. *The optimal restrictive contract has the following properties:*

(1) agent ℓ produces $\overline{q}^\ell < q^\star(\theta^\ell)$ with capital level $\overline{k}^\ell$ and receives a transfer of

$$\overline{t}^\ell = \overline{C}^P(\overline{q}^\ell,\overline{k}^\ell,\theta^\ell) \qquad (7\text{-}\overline{t}^\ell)$$

(2) agent h produces $\overline{q}^h = q^\star(\theta^h)$ using the neo-classical input vector $\left(e^\star(\theta^h),k^\star(\theta^h)\right)$, and receives a transfer of

$$\overline{t}^h = \overline{C}^P(\overline{q}^h,\theta^h) + \left(\overline{C}^P(\overline{q}^\ell,\overline{k}^\ell,\theta^\ell) - \overline{C}^P(\overline{q}^\ell,\overline{k}^\ell,\theta^h)\right) \qquad (7\text{-}\overline{t}^h)$$

(3) agent ℓ's capital-effort ratio exceeds the neo-classical ratio $\tilde{\beta}$.

The proof is relegated to the appendix, as are all subsequent proofs.

[3] See the references cited in footnote 2.

The proof of Proposition 2 makes use of the following lemma, which greatly simplifies the analysis of restrictive contracts. Because we make extensive use of the lemma in the analysis that follows, we state it here in the text. It states that without loss of generality, we can replace the standard formulation of the principal's problem—in which she chooses $(\mathbf{q},\mathbf{k},\mathbf{t})$—with the sparser formulation in which she chooses only $(\mathbf{q},\mathbf{k})$, and the transfers that agent types receive are pre-specified functions of $(\mathbf{q},\mathbf{k})$. Equipped with this lemma, we can henceforth ignore incentive compatibility and individual rationality constraints, because they have been incorporated in the principal's objective function.

Lemma 1. *The problem of choosing the optimal restrictive contract is equivalent to the following problem:*

$$\max_{(\bar{\mathbf{q}},\bar{\mathbf{k}})} \sum_{i\in\{\ell,h\}} \left\{ \phi^i \Big(p\bar{q}(\theta^i) - \bar{t}(\theta^i) \Big) \right\} \tag{8}$$

$$\begin{aligned} \textit{where} \quad \bar{t}^h &= \bar{C}^P(\bar{q}^h,\bar{k}^h,\theta^h) + \Big(\bar{C}^P(\bar{q}^\ell,\bar{k}^\ell,\theta^\ell) - \bar{C}^P(\bar{q}^\ell,\bar{k}^\ell,\theta^h) \Big) \\ \bar{t}^\ell &= \bar{C}^P(\bar{q}^\ell,\bar{k}^\ell,\theta^\ell) \end{aligned}$$

That is, any solution to (8) *is a solution to the principal's restrictive problem and vice versa.*

Let $\bar{C}^I(q,k) = \bar{C}^P(q,k,\theta^\ell) - \bar{C}^P(q,k,\theta^h)$ denote the information cost of contracting with type ℓ to produce q with capital level k under a restrictive contract. It follows from Lemma 1 that the task of choosing the optimal restrictive contract can be reformulated as the following (unconstrained) maximization problem:

$$\max_{(\bar{\mathbf{q}},\bar{\mathbf{k}})} \sum_{i\in\{\ell,h\}} \phi^i \Big(p\bar{q}(\theta^i) - \bar{C}^P(\bar{q}(\theta^i),\bar{k}(\theta^i),\theta^i) \Big) \quad + \quad \phi^h \bar{C}^I(\bar{q}(\theta^\ell),\bar{k}(\theta^\ell)) \tag{9-R}$$

Similarly, from Proposition 1, the task of choosing the optimal basic contract can be reformulated as:

$$\max_{\tilde{\mathbf{q}}} \sum_{i\in\{\ell,h\}} \phi^i \Big(p\tilde{q}(\theta^i) - \tilde{C}^P(\tilde{q}(\theta^i),\theta^i) \Big) \quad + \quad \phi^h \tilde{C}^I(\tilde{q}(\theta^\ell)) \tag{9-B}$$

where $\tilde{C}^I(q) = \tilde{C}^P(q,\theta^\ell) - \tilde{C}^P(q,\theta^h)$ denotes the information cost of contracting with type ℓ to produce q under a basic contract.

Note that since information costs are independent of type h's contractual variables, the presence of information asymmetry has no impact on the choice of these variables. For this reason, in the solution to (9-R), the values of $(\bar{q}(\theta^h),\bar{k}(\theta^h))$ coincide with the corresponding symmetric information benchmark values, $(q^\star(\theta^h),k^\star(\theta^h))$. Similarly, $\tilde{q}(\theta^h)$ coincides with $q^\star(\theta^h)$. For the remainder of the paper, we shall entirely ignore this uninteresting aspect of the principal's problem and focus our attention on the contract targeted for type ℓ. That is, we shall study and compare the following partial problems. To streamline

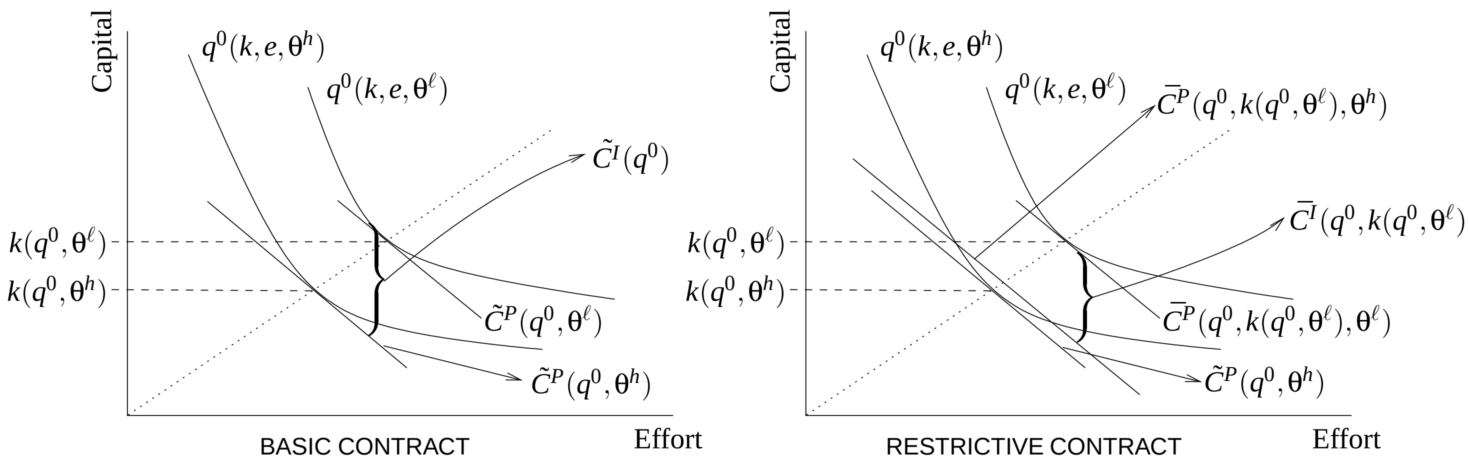

FIGURE 1. Information Cost of Producing q^0 under basic vs restrictive contract

notation,[4] we divide by ϕ^ℓ and write ϕ^h/ϕ^ℓ as Φ.

$$\text{Basic:} \qquad \max_{\tilde{q}} \left(p\tilde{q} - \tilde{C}^P(\tilde{q},\theta^\ell)\right) \quad - \quad \Phi\tilde{C}^I(\tilde{q}) \tag{9-B$'$}$$

$$\text{Restrictive:} \qquad \max_{(\bar{q},\bar{k})} \left(p\bar{q} - \bar{C}^P(\bar{q},\bar{k},\theta^\ell)\right) \quad - \quad \Phi\bar{C}^I(\bar{q}) \tag{9-R$'$}$$

All of the results that follow are consequences of the following inequality

$$\text{for all positive } q, \quad \tilde{C}^I(q) > \bar{C}^I(q). \tag{10}$$

Fig. 1 motivates (10) and, in turn, the essential difference between basic and restrictive contracts. Consider an arbitrary output level q^0. The higher (lower) isoquant in the figure indicates the set of input combinations with which type ℓ (type h) can produce q^0. The parallel lines represent isocost curves. The brace to the left indicates the cost differential if both kinds of agents were to produce q^0 using their respective cost-minimal (i.e., neoclassical) input combinations. The brace to the right indicates the reduced cost differential when agent h is penalized by being forced to use the capital level that is optimal for agent ℓ, i.e., $k(q^0,\theta^\ell)$. This penalty will be positive whenever effort and capital are not perfectly substitutable. The left and right braces also represent *information rents* that the principal would have to pay agent h, under, respectively, a basic and restrictive contract that specified an output level of q^0 and, in the restrictive contract, imposed on agent h the neoclassical input ratio for agent ℓ. We thus demonstrate that the principal can construct a restrictive contract which exactly mimics any basic contract, except for the added restriction on the input mix that h must use if she picks the contract designed for ℓ. Comparing the two contracts, the principal's revenues

4 For reasons that will become clear, type h is in the numerator of Φ but the denominator of $\Theta = \frac{\theta^\ell}{\theta^h}$ (p. 4).

are the same under both, since outputs are the same. Production costs are also the same, since the input mixes are identical. But information rents are lower under the constructed contract, and so profits associated with q^0 are higher. Since q^0 was chosen arbitrarily, this argument applies in particular to $\tilde{q}(\theta^{\ell})$, the output assigned to ℓ in the optimal basic contract. It follows that profits under this contract must be strictly less than profits under the optimal restrictive contract.

The main implication of the preceding remarks is summarized in Proposition 3.

Proposition 3. *The principal's profits under the optimal restrictive contract strictly exceed her profits under the optimal basic contract.*

To prove Proposition 3, it is sufficient to construct a feasible restrictive contract that delivers the same output levels as the optimal basic contract but at a strictly lower cost to the principal.

3. Marginal Analysis of the Basic and Restrictive Contracts

In this section we use marginal analysis to study the relationship between the two kinds of contracts. In order to use calculus techniques, we vary continuously the degree of information asymmetry, starting with symmetric information and ending with the incomplete information model described in section 2. Specifically, for each $\chi \in [0,1]$, we solve the principal's two problems (9-B′) and (9-R′), replacing Φ in each of them with $\gamma = \chi\Phi$. We can then compare, under the two types of contracts, the effect on the principal's choice variables of a small increase in information asymmetry. Since our symmetric information benchmark solution (p. 7) is independent of γ, the rates at which output, the principal's profits, etc. decline as γ increases are pure measures of the *marginal* impacts of information asymmetry. We can then integrate these marginal impacts over the interval $[0,\Phi]$ to recover and compare the *total* impacts of information asymmetry on the solutions to the principal's original problems (9-B′) and (9-R′).

§3.1 below presents our marginal analysis of the difficult case, which is the restrictive contract. The analysis in §3.2 of the basic contract is completely routine, but is required so that we can compare expressions. In each case we first derive the principal's marginal cost function for fixed γ, then determine output by equating this function to the price level.

3.1. **Restrictive Contract.** We begin by determining the minimum cost—i.e., production plus information cost—to the principal of having type ℓ produce at least q under a restrictive contract, for given γ. We call the resulting mapping the *restrictive cost function,* $\bar{C}(q,\gamma)$. We decompose $\bar{C}(q,\gamma)$ into $\bar{C}^P(q,\gamma)+\bar{C}^I(q,\gamma)$, where $\bar{C}^P(q,\gamma)=(v\bar{e}^\ell(q,\gamma)+r\bar{k}(q,\gamma))$ is the *production cost* and $\bar{C}^I(q,\gamma)=\gamma v(\bar{e}^\ell(q,\gamma)-\bar{e}^h(q,\gamma))$ is the *information cost* of producing q under the restrictive contract. We then select the profit maximal level of q, given γ, by setting the *restrictive marginal cost function,* $\overline{MC}(q,\gamma)=\frac{d\bar{C}(q,\gamma)}{dq}$, equal to p.

It is convenient to set up our cost minimization problem subject not to the usual equality constraints but to *inequality* constraints.[5] Specifically, we minimize the cost to the principal of having type ℓ produce *at least* q, while requiring that if type h imitates, he produces *at most* q.[6] From (9-R′), the cost minimization problem under the restrictive contract is to pick the nonnegative vector $(\mathbf{e},k)$, $\mathbf{e}=(e^\ell,e^h)$, which minimizes $\left\{\bar{C}^P(\cdot,\cdot,\theta^\ell)+\gamma\bar{C}^I(\cdot,\cdot)\right\}$ subject to these constraints. Specifically, the problem is:

$$\min_{(\mathbf{e},k)}\left\{v\left((1+\gamma)e^\ell-\gamma e^h\right)+rk\right\}\qquad \text{s.t.}\quad f^\ell(e^\ell,k)\geq q,\ f^h(e^h,k)\leq q \text{ and } (\mathbf{e},k)\geq 0 \tag{11}$$

A consequence of a restriction we shall later impose (see (17) below) is that in the solution to (11), $(\mathbf{e},k)$ will necessarily be positive. Because of this, we will omit the nonnegativity constraints from our specification of the Lagrangian:

$$\bar{L}(\mathbf{e},k,\bar{\boldsymbol{\lambda}};q,\gamma) \quad = \quad v\left((1+\gamma)e^\ell-\gamma e^h\right)+rk \quad + \quad \bar{\lambda}^\ell(q-f^\ell(e^\ell,k)) \quad + \quad \bar{\lambda}^h(f^h(e^h,k)-q)) \tag{12}$$

where $\bar{\boldsymbol{\lambda}}=(\bar{\lambda}^\ell,\bar{\lambda}^h)$ is the vector of Lagrangian multipliers for the restricted problem. The first order condition for $\bar{L}$ has five equations in five unknowns:

$$\nabla\bar{L} \quad = \quad \begin{bmatrix} \bar{L}_{e^\ell} \\ \bar{L}_{e^h} \\ \bar{L}_k \\ \bar{L}_{\bar{\lambda}^\ell} \\ \bar{L}_{\bar{\lambda}^h} \end{bmatrix} \quad = \quad \begin{bmatrix} (1+\gamma)v-\bar{\lambda}^\ell f_e^\ell \\ -\gamma v+\bar{\lambda}^h f_e^h \\ r-\bar{\lambda}^\ell f_k^\ell+\bar{\lambda}^h f_k^h \\ q-f^\ell(e^\ell,k) \\ f^h(e^h,k)-q \end{bmatrix} \quad = \quad 0. \tag{13}$$

At the solution $\left(\bar{\mathbf{e}}(q,\gamma),\bar{k}(q,\gamma),\bar{\boldsymbol{\lambda}}(q,\gamma)\right)$ to (13), the constraints are identically zero so that the *restrictive cost function* $\bar{C}(q,\gamma)$—defined as the minimum attainable value of total cost under the restrictive contract for each (q,γ) pair—is identically equal to the minimized value of $\bar{L}$, henceforth denoted by $\bar{L}(\cdot\,;q,\gamma)$.

---

5 This setup ensures that our Lagrangians are non-negative. With equality constraints, the Lagrangians cannot be signed.

6 The point here is that if the principal were able to, she could reduce information rents by requiring that an imitating agent produce *more than* q. Obviously she cannot impose this requirement, hence the constraint. While this setup is nonstandard, it clearly produces the right result, which is that the principal chooses to have both the low-ability and the imitating high-ability agent produce q.

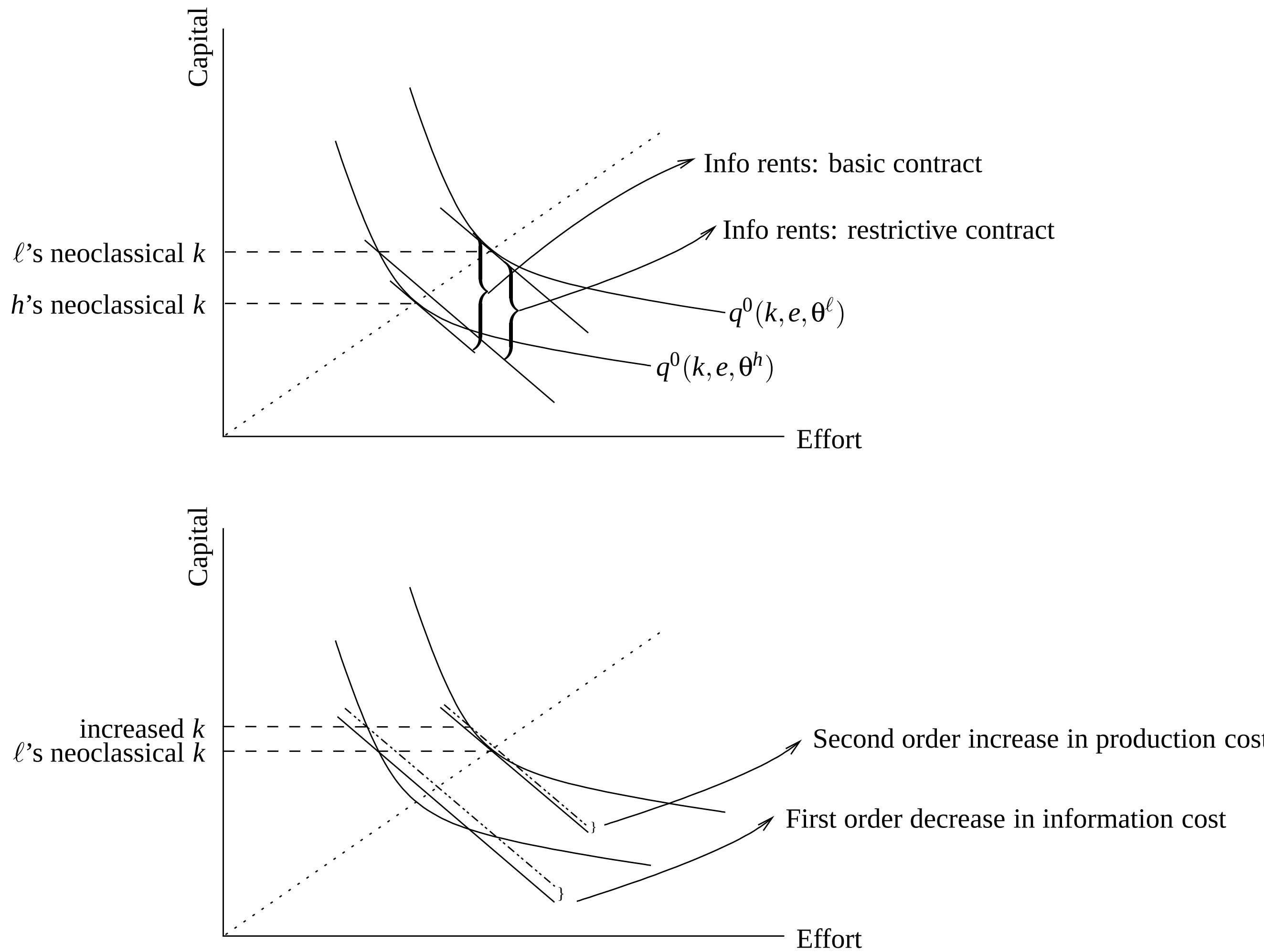

FIGURE 2. Effect on Production and Information Cost of Increasing k

Note that because $\overline{L}_{e^\ell}$, $\overline{L}_{e^h}$ and $\overline{L}_k$ are all zero, we have

$$\overline{\lambda}^\ell \quad = \quad \frac{(1+\gamma)\nu}{f_e^\ell} \quad > \quad \overline{\lambda}^h \quad = \quad \frac{\gamma\nu}{f_e^h} \tag{14}$$

To see that the strict inequality holds, note that the numerator of $\overline{\lambda}^\ell$ is larger than the numerator of $\overline{\lambda}^h$. On the other hand, since h is more efficient than ℓ, and both h and ℓ are using the same level of capital while the f_e^i's are evaluated at the same level of output, the denominator of $\overline{\lambda}^\ell$ is smaller than the denominator of $\overline{\lambda}^h$. An immediate implication of (14) is:

Proposition 4. *In a restrictive contract for a given $(q,\gamma) \gg 0$, the prescribed capital-effort ratio for the low ability agent is greater than the neoclassical ratio $\tilde{\beta}$.*

(For a vector $\mathbf{x} \in \mathbb{R}^n$, we write $\mathbf{x} \gg 0$ if $x_i > 0$, for $i = 1, \ldots n$.) Note that Proposition 4 is more general than, and hence implies property 3 of Proposition 2. Moreover Proposition 4 holds under much more general conditions that A1-A3. Fig. 2 provides some intuition for the result and suggests a weaker sufficient condition. Its top panel reproduces Fig. 1 above. Consider the effect on the principal's problem of increasing γ from zero, for the moment holding the output level constant at an arbitrary output level $\mathring{q}$. By the envelope theorem, a small increase in capital intensity above the neoclassical level has only a second-order impact on the production costs of agent ℓ (see the bottom panel of Fig. 2). On the other hand, since the initial capital level is already super-optimal for agent h, the given increase would result in a *first*-order increase in agent h's production cost if he accepted the contract designed for ℓ. Thus, a small increase in capital intensity beyond the neoclassical level for ℓ results in a first order reduction in information costs, and a second-order increase in production costs. It follows that whenever $\gamma > 0$, the prescribed level of capital for agent ℓ will exceed the neoclassical level for her prescribed level of output. Fig. 2 makes clear that neither A1 or A2 are required for this result to hold. A sufficient, but still far from necessary condition is that the difference between types be technologically neutral, in the sense that for any q, the isoquants associated with that q for the two types be parallel to each other.

Profit maximization under the restrictive contract: The *restrictive marginal cost function*, denoted by $\overline{MC}$, is identically equal to $\frac{d\bar{L}(\cdot\,;q,\gamma)}{dq}$ which, by the envelope theorem, equals $\frac{\partial\bar{L}(\cdot\,;q,\gamma)}{\partial q}$. This partial derivative in turn equals the difference between the two Lagrangians, $\bar{\lambda}^{\ell}$ and $\bar{\lambda}^{h}$, so that $\overline{MC}(q,\gamma) = \bar{\lambda}^{\ell}(q,\gamma) - \bar{\lambda}^{h}(q,\gamma)$. Moreover, at the principal's optimum, $\overline{MC}(\bar{q}(\gamma),\gamma) = p$, where $\bar{q}(\gamma)$ is the profit maximizing level of output produced by the agent of type ℓ at price p under the restrictive contract.

Our research strategy for studying the properties of the restrictive contract will be to apply the implicit function theorem to the first order conditions (13), along the path $\{(\bar{q}(\gamma),\gamma) : \gamma \in [0,\Phi]\}$. This requires, of course, that the determinant of the Hessian of $\bar{L}(\cdot\,;\bar{q}(\gamma),\gamma)$, denoted $\Delta^{\overline{HL}}(\gamma)$, is non-zero along this path. It is easy to verify that $\Delta^{\overline{HL}}(0)$ is positive. By continuity, the previous requirement is then equivalent to requiring that $\Delta^{\overline{HL}}(\cdot)$ is positive on $[0,\Phi]$. To appreciate the implications of this requirement, consider the case in which g is CES. in addition to satisfying A1-A3. In this case, the expression for the determinant reduces to

$$\Delta^{\overline{HL}}(\gamma) \quad = \quad -\tau \times \left\{ \frac{\gamma}{f_e^h(e^h,k)} \, \frac{e^\ell f_e^\ell(e^\ell,k)}{k f_k^\ell(e^\ell,k)} \quad - \quad \frac{1+\gamma}{f_e^\ell(e^\ell,k)} \, \frac{e^h f_e^h(e^h,k)}{k f_k^h(e^h,k)} \right\} \tag{15}$$

where $\xi > 0$ depends only on parameters of the model. Clearly, expression (15) will be positive in a neighborhood of $\gamma = 0$. For large γ, however, positivity is difficult to guarantee when inputs are close substitutes and h is much more efficient than ℓ. Fig. 3 illustrates the problem. When isoquants have minimal curvature, the difference between $f_e^\ell(e^\ell,k)$ and $f_e^h(e^h,k)$ depends on the efficiency gap, but only minimally on the input ratio, while the ratios $\frac{e^\ell f_e^\ell(e^\ell,k)}{k f_k^\ell(e^\ell,k)}$ and $\frac{e^h f_e^h(e^h,k)}{k f_k^h(e^h,k)}$ are very similar. Given all other parameters, therefore, we can construct an example in which e^h is arbitrarily close to zero (as in Fig. 3), ensuring that (15) will be positive except when γ is very small. Lemma 2 below establishes that this problem does not arise when the elasticity of substitution between effort and labor is bounded above by unity.[7]

Lemma 2. *If g is CES in effort and capital, with constant elasticity of substitution parameter* $\bar{\sigma}_{ke} \leq 1$, *then* $\Delta^{\overline{HL}}(\cdot)$ *will be positive on* $[0,\Phi]$

Proof of Lemma 2: The lemma follows immediately from an inspection of (15). Since h is more efficient than ℓ and $e^h < e^\ell$, we have $f_e^h(e^h,k) > f_e^h(e^\ell,k)$ so that $\frac{1+\gamma}{f_e^\ell(e^\ell,k)} > \frac{\gamma}{f_e^h(e^h,k)}$. (15) will therefore be positive if

$$\frac{e^h \;\; f_e^h(e^h,k)}{k \;\; f_k^h(e^h,k)} \quad > \quad \frac{e^\ell \;\; f_e^\ell(e^\ell,k)}{k \;\; f_k^\ell(e^\ell,k)} \quad \text{or, more conveniently,} \quad \frac{k \;\; f_k^\ell(e^\ell,k)}{e^\ell \;\; f_e^\ell(e^\ell,k)} \quad > \quad \frac{k \;\; f_k^h(e^h,k)}{e^h \;\; f_e^h(e^h,k)} \tag{16}$$

Now $e^h < e^\ell$; moreover since $\bar{\sigma}_{ke} \leq 1$, a given change $\left|\frac{d(k/e)}{k/e}\right|$ induces a weakly smaller change $\left|\frac{d(f_k/f_e)}{f_k/f_e}\right|$ (see fn. 7), i.e., we have $\frac{k/e^\ell}{k/e^h} \geq \frac{f_k^h(e^h,k)/f_e^h(e^h,k)}{f_k^h(e^h,k)/f_e^h(e^h,k)} > 1$. Hence (16) is satisfied. □

Since the sufficient condition in Lemma 2 is far from necessary for the property we need, we will hold the condition in reserve for the moment and, in Props. 5 and 7 below, simply *assume* that $\Delta^{\overline{HL}}(\cdot) > 0$. A convenient implication of this assumption—which we invoked when we specified the Lagragian (12)—is:

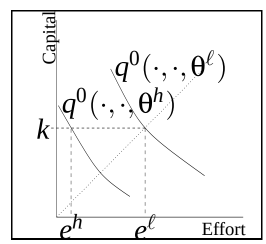

FIGURE 3.

$$\text{For all } \gamma > 0, \text{if } \Delta^{\overline{HL}}(\gamma) \text{ is positive then}$$
$$e^h(\bar{q}(\gamma),\gamma) \text{ and hence } e^\ell(\bar{q}(\gamma),\gamma) \text{ and } \bar{q}(\gamma) \text{ are also positive.} \tag{17}$$

To see this, note that if $e^h(\bar{q}(\gamma),\gamma) = 0$, then the first term in (15) would be positive and the second term zero.

The next result is critical and by no means obvious, although its analog for the basic contract is self-evident.

Proposition 5. *If* $\Delta^{\overline{HL}}(\cdot)$ *is positive on* $[0,\Phi]$, *then for all* $\gamma > 0$, *the restrictive marginal cost function* $\overline{MC}(\cdot,\gamma)$ *is increasing in q. Moreover,* $\bar{q}(\cdot)$ *is a continuously differentiable function of* γ, *with*

$$\bar{q}'(\cdot) \quad = \quad \frac{v(\bar{e}^\ell - \bar{e}^h)}{(\alpha-1)(\lambda^\ell - \lambda^h)} \quad < \quad 0 \tag{18}$$

3.2. **Basic Contract.** The objective of this subsection is to obtain an expression analogous to (18) for the basic contract, i.e., for $\frac{d\tilde{q}(\gamma)}{d\gamma}$. By Remark 2, it is sufficient to analyze the equivalent single-input formulation

---

[7] The elasticity of substitution of f at (e,k) is denoted by (see Silberberg (1990)):

$$\sigma_{ke}(e,k) \quad = \quad -\left.\frac{\partial \ln(k/e)}{\partial \ln(f_k/f_e)}\right|_{f(e,k)=\text{constant}} \quad = \quad -\left.\frac{d(f_k/f_e)}{f_k/f_e} \Big/ \frac{d(k/e)}{k/e}\right|_{f(e,k)=\text{constant}} \quad = \quad -\frac{f_e f_k(f_e e + f_k k)}{ek[(f_k)^2 f_{ee} - 2 f_e f_k f_{ek} + (f_e)^2 f_{kk}]}$$

of the principal's problem under the basic contract (ee Remark 2). Since this analysis is entirely routine, we will be very brief. Let $\ddot{g}$ denote the single-input production function corresponding to $\ddot{f}$. Let $\ddot{e}^i(q)$ denote the level of composite input required for type i to produce q.[8] From Remark 1, $\ddot{e}^h(\cdot) = \Theta^{1/\alpha}\ddot{e}^\ell(\cdot)$ (recall from page 5 that $\Theta = \frac{\theta^\ell}{\theta^h}$), so that the information cost of producing q under a basic contract is $\ddot{C}^I(q) = \ddot{v}\left(1-\Theta^{1/\alpha}\right)\ddot{e}^\ell(q)$. Thus, the *basic cost function* is

$$\tilde{C}(q,\gamma) \quad = \quad \ddot{C}^P(q,\theta^\ell) \;+\; \gamma\ddot{C}^I(q) \quad = \quad \ddot{v}\left\{1+\gamma\left(1-\Theta^{1/\alpha}\right)\right\}\ddot{e}^\ell(q) \tag{19}$$

Since $\frac{d\ddot{e}^\ell(q)}{dq} = \left(\ddot{g}'(\ddot{e}^\ell(q))\right)^{-1}$, the *basic marginal cost function* is

$$\widetilde{MC}(q,\gamma) \quad = \quad \ddot{v}\left\{1+\gamma\left(1-\Theta^{1/\alpha}\right)\right\}\left(\ddot{g}'(\ddot{e}^\ell(q))\right)^{-1}. \tag{20}$$

Proceeding as on page 14, $\widetilde{MC}(\tilde{q}(\gamma),\gamma) = p$ at the principal's optimum, where $\tilde{q}(\gamma)$ is the profit maximizing level of output produced by the agent ℓ at price p under the basic contract. Once again, we apply the implicit function theorem to obtain:

$$\begin{aligned}\frac{d\tilde{q}(\gamma)}{d\gamma} \quad &= \quad -\frac{\partial\widetilde{MC}(\tilde{q}(\gamma),\gamma)}{\partial\gamma}\bigg/\frac{\partial\widetilde{MC}(\tilde{q}(\gamma),\gamma)}{\partial q} \\ &= \quad \ddot{v}\left(1-\Theta^{1/\alpha}\right)/\ddot{g}'\bigg/\frac{\ddot{v}\left\{1+\gamma\left(1-\Theta^{1/\alpha}\right)\right\}}{(\ddot{g}')^2}\frac{\ddot{g}''}{\ddot{g}'}\end{aligned}$$

Since $\ddot{g}'$ is homogeneous of degree $\alpha-1$, Euler's theorem implies that $\ddot{g}'' = (\alpha-1)\ddot{g}'/\ddot{e}^\ell(q)$. By profit maximization, $\left\{1+\gamma\left(1-\Theta^{1/\alpha}\right)\right\} = p\ddot{g}'/\ddot{v}$. Using these two equalities, the expression becomes

$$\frac{d\tilde{q}(\gamma)}{d\gamma} \quad = \quad \frac{\ddot{v}\ddot{e}^\ell(q)\left(1-\Theta^{1/\alpha}\right)}{(\alpha-1)p} \quad = \quad \frac{(v\tilde{e}^\ell(q)+r\tilde{k}^\ell(q))\left(1-\Theta^{1/\alpha}\right)}{(\alpha-1)p} \quad < \quad 0. \tag{21}$$

3.3. **Comparing Restrictive and Basic Contracts.** In this subsection, we establish two factors that contribute to the dominance, from the principal's perspective, of the restrictive over the basic contract. We have already observed (inequality (10)) that information costs are lower under the the optimal restrictive contract than under the optimal basic contract, while (from Proposition 4) *production* costs are higher. Proposition 6 notes that the former inequality dominates.

---

8 Since the principal has only one choice variable under a basic contract, there is no need to set up a Lagrangian corresponding to (12) to determine the cost function.

Proposition 6. *For any given $(q,\gamma) \gg 0$, the principal's total cost of optimally obtaining q under a restrictive contract is less than the corresponding costs under a basic contract. That is,*

$$\text{for all } q \text{ and all } \gamma > 0, \quad \bar{C}(q,\gamma) \quad < \quad \tilde{C}(q,\gamma) \tag{22}$$

The proof of Proposition 6 is immediate: the neoclassical input mix is feasible under the restrictive contract, but, by Proposition 4, violates the first order condition (28). The proposition reflects the fact, illustrated in Fig. 2, that the first order reduction in information costs obtained by moving away from the neoclassical input mix necessarily offsets the resulting, second order increase in production costs.

Our next result is far less immediate. It establishes that output is distorted by less under the restrictive contract than under the basic contract.

Proposition 7. *If $\Delta^{\overline{HL}}(\cdot)$ is positive on $[0,\Phi]$, then output produced by agent ℓ is higher under the second best restrictive contract than under the second best basic contract.*

An interpretation of Proposition 7 is that the marginal cost curve (including both production and information rent costs), is strictly lower under the restrictive contract than under the basic contract. Intuitively, this is because with the restrictive contract the principal can limit the extent to which agent *h* could substitute between effort and capital if she were to accept the contract designed for ℓ.

4. The social cost of information asymmetry

As we have seen, the principal's profits are higher under the optimal restrictive contract than under the optimal basic contract. This does not imply, however, that restrictive contracts are preferable to basic contracts from a *social* perspective. While the principal's objective is to minimize the sum of production and information costs, only production costs matter for social surplus. Information costs are simply a transfer from the principal to agent *h*. Our task in this section is to compare social surplus under the two types of contracts.

In the present model with perfectly elastic demand, social surplus is the sum of the principal's profit and the information rent received by type *h*. Since the information rent is a pure transfer, social surplus is equal to the principal's total revenue minus *production* costs. Although information rents are lower under the optimal restrictive contract, average production costs are higher, because the input mix is sub-optimal from a pure production standpoint. We refer to this distortion as the *input mix effect*. The second factor which affects social surplus is the *level* of production. Proposition 7 establishes that production is always higher under the optimal restrictive contract. We refer to this difference as the *output effect*. The difference between social surplus under the optimal restrictive and basic contracts is the sum of the two effects. Whether the positive output effect or the negative input mix effect dominates depends on a number of considerations, including the elasticity of substition, the relative importance of the two inputs in the production process and the productive efficiency gap between the two types.

The main result of this paper is that if the elasticity of substitution (see footnote 7) is sufficiently small, then social surplus is higher under the restrictive contract than under the basic contract. To obtain this result, we need certain parameters to be bounded away from their natural boundaries. Specifically, we require that the ratio of types' efficiency factors (Θ), the probability of type ℓ (ϕ^ℓ), and the homogeneity factor (α) all belong to compact subsets of $(0,1)$, and that the neoclassical input ratio ($\tilde{\beta}$) and the level of output associated with input vector $(1,\tilde{\beta})$ both belong to compact subsets of $(0,\infty)$. Since any compact sets will do, we define them all in terms of an arbitrarily small scalar, $\breve{\omega} \in (0, 0.5)$. Let

$$G \quad = \quad \left\{ g \text{ satisfying A1-A3} : \Theta, \phi^\ell, \alpha \in [\breve{\omega}, 1-\breve{\omega}], \ \& \ \tilde{\beta}, g(1,\tilde{\beta}) \in [\breve{\omega}, 1/\breve{\omega}] \right\} \tag{23}$$

Given a production function g, let $\sigma(e,k|g)$ denote the elasticity of substitution (see fn. 7) of capital for effort for the function g at (k,e) and let $\bar{\sigma}(g) = \sup\{\sigma(e,k|g) : (e,k) \in \mathbb{R}^2_+\}$. Obviously, for the special case in which g is CES, $\bar{\sigma}(g)$ is just the familiar constant measure. Significantly, the fact that g belongs to G does not impose any restriction on $\bar{\sigma}(g)$.

Proposition 8. *There exists $n \in \mathbb{N}$ such that for all $g \in G$ with $\bar{\sigma}(g) \leq 1/n$, social surplus with technology g is higher under the restrictive contract than under the basic contract.*

The intuition for this result is straightforward. As the elasticity of substitution declines, the principal can obtain a larger and larger "bang for the buck" in terms of information rents. That is, a small distortion in the capital-effort ratio away from the optimal ratio results in a larger and larger decrease in information rents (see Fig. 4). As a consequence, as the elasticity of substitution approaches zero, the quantity produced under the optimal restrictive contract approaches the first-best quantity, while the distortion in the capital-effort ratio goes to zero. It follows that for n sufficiently large, the optimal restrictive contract will socially dominate the optimal basic contract.

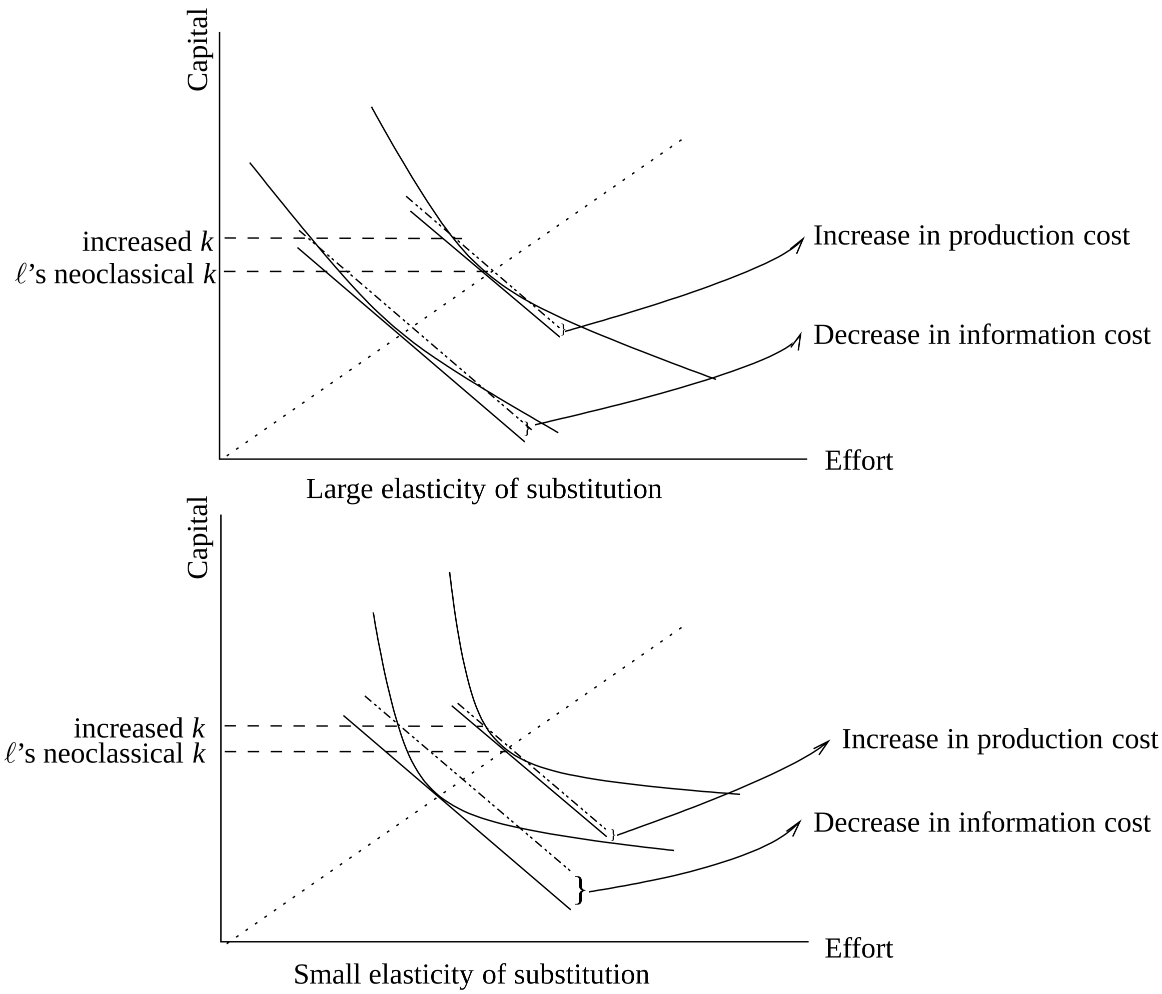

FIGURE 4. Input Mix Distortions and Information Rents

While the proof of Proposition 8 is conceptually very simple, it requires a massive amount of detail. The details arise because we need to find a bound on the elasticity of substitution that implies our result *for all* production functions in the set G. This task would be much easier if the set G were compact. We cannot, however, impose this restriction because we need to include in G functions with arbitrarily small, but positive, elasticities of substitution. The proof is organized into five steps. Step 1 establishes lower and upper bounds on the level of inputs used in any optimal contract, and on the rates at which certain variables change with γ. Steps 2 and 3 are convergence results as the upper bound on the elasticity of substitution goes to zero: step 2 proves that the divergence from the neoclassical input mix of the mix that is optimal for type ℓ in the restrictive contract goes to zero also; step 3 proves that while type h's capital usage in the restrictive contract is bounded away from h's use in the basic contract, the difference between the levels of labor that h uses in the two contracts goes to zero. Step 4 extends Proposition 7 by establishing a lower bound on the gap between optimal outputs under the two kinds of contract that holds uniformly for all technologies in G with sufficiently small elasticities of substitution. Step 5 now completes the proof, by identifying a lower bound on the (positive) output effect, and showing that the (negative) input mix effect can be made arbitrarily small by shrinking the upper bound on the elasticity of substitution.

Matters are less straightforward when the two inputs are close substitutes. First, an interior solution for the restrictive contract may not exist. Specifically, if θ^h were slightly higher than the value represented in Fig. 3 above, then type h would be able to imitate ℓ without utilizing any effort at all. Second (as in Fig. 3), an interior solution may exist but the determinant $\Delta^{\overline{HL}}(\cdot)$ may be negative for sufficiently large γ (see (15)). In this case none of the machinery on which Proposition 8 is based can be applied. Third, suppose, for a sequence (g^n) satisfying A1-A3 with $\bar{\sigma}(g^n)$ increasing without bound, that an interior solution does exist and $\Delta^{\overline{HL}}(\cdot)$ is positive on $[0,\Phi]$. In this case, both the input-mix and output effects will converge to zero and we cannot guarantee that the latter effect will dominate the former. Hence it is possible that when the two inputs are close subsitutes, the optimal basic contract yields a higher level of social surplus than the optimal restrictive contract.

5. Conclusion

We have shown that the principal's profits are higher when she controls capital than when she does not. Further, output is greater in this case, since the principal can allocate capital to mitigate her information costs. In the process, however, she distorts the capital-effort ratio away from its neoclassical level and this distortion is socially costly. Provided that effort and capital are sufficiently complementary, the restrictive contract will result in higher social surplus than the basic contract. When effort and capital are highly substitutable, the relative social desirability of the two contracts depends on the relative severities of the output and the input-mix distortions.

APPENDIX: PROOFS

The proof of Proposition 2 utilizes the proofs of Lemma 1, Proposition 4, and Proposition 5, so it is presented after these proofs.

Proof of Lemma 1: The optimal restrictive contract is the solution to the following program

$$\max_{(\bar{\mathbf{q}},\bar{\mathbf{k}},\bar{\mathbf{t}})} \sum_{i\in\{\ell,h\}} \left\{ \phi^i \left(p\bar{q}(\theta^i) - \bar{t}(\theta^i) \right) \right\} \tag{24}$$

$$\begin{aligned}
\text{subject to} \quad \bar{t}(\theta^\ell) &\geq \bar{C}^P(\bar{q}(\theta^\ell),\bar{k}(\theta^\ell),\theta^\ell) && (\mathrm{IR}_\ell)\\
\bar{t}(\theta^h) &\geq \bar{C}^P(\bar{q}(\theta^h),\bar{k}(\theta^h),\theta^h) && (\mathrm{IR}_h)\\
\bar{t}(\theta^\ell) - \bar{C}^P(\bar{q}(\theta^\ell),\bar{k}(\theta^\ell),\theta^\ell) &\geq \bar{t}(\theta^h) - \bar{C}^P(\bar{q}(\theta^h),\bar{k}(\theta^h),\theta^\ell) && (\mathrm{IC}_\ell)\\
\bar{t}(\theta^h) - \bar{C}^P(\bar{q}(\theta^h),\bar{k}(\theta^h),\theta^h) &\geq \bar{t}(\theta^\ell) - \bar{C}^P(\bar{q}(\theta^\ell),\bar{k}(\theta^\ell),\theta^h) && (\mathrm{IC}_h)
\end{aligned}$$

Now, given $(\bar{\mathbf{q}},\bar{\mathbf{k}},\bar{\mathbf{t}})$ define the *information rent* vector $\overline{\mathbf{dt}}$ by for $\theta \in \{\theta^\ell,\theta^h\}$, $\overline{dt}(\theta) = \bar{t}(\theta) - \bar{C}^P(\bar{q}(\theta),\bar{k}(\theta),\theta)$. The principal's problem can now be rewritten in terms of $(\bar{\mathbf{q}},\bar{\mathbf{k}},\overline{\mathbf{dt}})$ as:

$$\max_{(\bar{\mathbf{q}},\bar{\mathbf{k}},\bar{\mathbf{t}})} \sum_{i\in\{\ell,h\}} \left\{ \phi^i \left(p\bar{q}(\theta^i) - \bar{t}(\theta^i) \right) \right\} \tag{25}$$

$$\begin{aligned}
\text{subject to} \quad \overline{dt}(\theta^\ell) &\geq 0 && (\mathrm{IR}'_\ell)\\
\overline{dt}(\theta^h) &\geq 0 && (\mathrm{IR}'_h)\\
\overline{dt}(\theta^\ell) &\geq \overline{dt}(\theta^h) - \left(\bar{C}^P(\bar{q}(\theta^h),\bar{k}(\theta^h),\theta^\ell) - \bar{C}^P(\bar{q}(\theta^h),\bar{k}(\theta^h),\theta^h)\right) && (\mathrm{IC}'_\ell)\\
dt(\theta^h) &\geq dt(\theta^\ell) + \left(\bar{C}^P(\bar{q}(\theta^\ell),\bar{k}(\theta^\ell),\theta^\ell) - \bar{C}^P(\bar{q}(\theta^\ell),\bar{k}(\theta^\ell),\theta^h)\right) && (\mathrm{IC}'_h)
\end{aligned}$$

Clearly, given any solution $(\bar{\mathbf{q}},\bar{\mathbf{k}},\bar{\mathbf{t}})$ of (24), if $\overline{dt}(\theta^\ell) = 0$ and $\overline{dt}(\theta^h) = \left(\bar{C}^P(\bar{q}(\theta^\ell),\bar{k}(\theta^\ell),\theta^\ell) - \bar{C}^P(\bar{q}(\theta^\ell),\bar{k}(\theta^\ell),\theta^h)\right)$, then $(\bar{\mathbf{q}},\bar{\mathbf{k}})$ must be a solution to (8), since the set of instruments available to the principal in the latter problem is a strict subset of the set of instruments available to her in the former. To prove the lemma, therefore, we need only show that if $(\bar{\mathbf{q}},\bar{\mathbf{k}},\bar{\mathbf{t}})$ is a solution to (24), then

$$\overline{dt}(\theta^\ell) = 0 \quad \text{and} \quad \overline{dt}(\theta^h) = \left(\bar{C}^P(\bar{q}(\theta^\ell),\bar{k}(\theta^\ell),\theta^\ell) - \bar{C}^P(\bar{q}(\theta^\ell),\bar{k}(\theta^\ell),\theta^h)\right) \tag{26}$$

so that $(\bar{\mathbf{q}},\bar{\mathbf{k}})$ is also a solution to (8).

First note that if $\overline{dt}(\theta^\ell) = 0$, then (IC'_h) must hold with equality, so that (26) holds. Otherwise $\overline{dt}(\theta^h)$ could be reduced without violating either (IC'_h) or (IR'_h), resulting in a higher payoff to the principal. In order to complete the proof, therefore, it is sufficient to prove that $\overline{dt}(\theta^\ell) = 0$. Suppose that $\overline{dt}(\theta^\ell) = \varepsilon > 0$. Then by (IC'_h), $\overline{dt}(\theta^h) \geq \varepsilon$, since $\left(\bar{C}^P(\bar{q}(\theta^\ell),\bar{k}(\theta^\ell),\theta^\ell) - \bar{C}^P(\bar{q}(\theta^\ell),\bar{k}(\theta^\ell),\theta^h)\right)$ is positive (see expression (4) above). In this case, we can subtract ε from both $\overline{dt}(\theta^h)$ and $\overline{dt}(\theta^\ell)$, all four inequalities will remain satisfied, and the principal's payoff will be higher. □

Proof of Proposition 3. Let $(\tilde{\mathbf{q}},\tilde{\mathbf{t}}) = \left((\tilde{q}^\ell,\tilde{t}^\ell),(\tilde{q}^h,\tilde{t}^h)\right)$ denote the optimal basic contract. Construct the restrictive contract $(\check{\mathbf{q}},\check{\mathbf{k}},\check{\mathbf{t}}) = \left((\check{q}^\ell,\check{k}^\ell,\check{t}^\ell),(\check{q}^h,\check{k}^h,\check{t}^h)\right)$, where $\check{\mathbf{q}} = \tilde{\mathbf{q}}$ and for $\theta \in \{\theta^\ell,\theta^h\}$, $\check{k}(\theta) = \tilde{k}(\tilde{q},\theta)$. That is, under this constructed restrictive contract, the outputs that were produced under the original basic contract are once again produced using the (neoclassical) input mix that was endogenously selected under the original basic contract. Thus for each θ, the production cost of $\breve{q}(\theta)$ is identical under both contracts. On the other hand, the *information cost* associated with $(\breve{\mathbf{q}},\breve{\mathbf{k}},\breve{\mathbf{t}})$ is lower than the information cost associated with $(\tilde{\mathbf{q}},\tilde{\mathbf{t}})$. To see this, note that since $\theta^\ell < \theta^h$, $\breve{k}^\ell = \tilde{k}(\breve{q}^\ell,\theta^\ell)$ is distinct from the unique solution $\tilde{k}(\breve{q}^\ell,\theta^h)$ to the first order condition (1) for type h. Hence, we have $\overline{C}^P(\breve{q}^\ell,\breve{k}^\ell,\theta^h) > \overline{C}^P(\breve{q}^\ell,\tilde{k}(\breve{q}^\ell,\theta^h),\theta^h) = \tilde{C}^P(\tilde{q}^\ell,\theta^h)$. Therefore,

$$\begin{aligned} \tilde{C}^I(\breve{q}^\ell) &= \tilde{C}^P(\tilde{q}^\ell,\theta^\ell) - \tilde{C}^P(\tilde{q}^\ell,\theta^h) \\ &> \tilde{C}^P(\tilde{q}^\ell,\theta^\ell) - \overline{C}^P(\breve{q}^\ell,\breve{k}^\ell,\theta^h) \\ &= \overline{C}^P(\breve{q}^\ell,\breve{k}^\ell,\theta^\ell) - \overline{C}^P(\breve{q}^\ell,\breve{k}^\ell,\theta^h) = \overline{C}^I(\breve{q}^\ell,\breve{k}^\ell) \end{aligned} \tag{27}$$

The restrictive contract we constructed thus delivers the same output at a strictly lower cost to the principal, and hence achieves strictly higher profits for the principal than the optimal basic contract. Since this constructed contract is feasible, the optimal restrictive contract must achieve strictly higher profits as well. □

Proof of Proposition 4: Observe first that substituting the expressions for the λ's obtained in (14) into the expression for $\overline{L}_k$ obtained in (13) yields:

$$1 - \frac{v}{r}\frac{\overline{f}_k^\ell}{f_e^\ell} = \frac{\gamma v}{r}\left(\frac{\overline{f}_k^\ell}{f_e^\ell} - \frac{\overline{f}_k^h}{f_e^h}\right) \tag{28}$$

Since h is more efficient than ℓ and both are using the same level of capital to produce the same level of output, h's effort level under the restrictive contract must be less than ℓ's. That is, $\frac{\overline{k}}{e^h} > \frac{\overline{k}}{e^\ell}$ which in turn implies $\frac{\overline{f}_k^h}{f_e^h} < \frac{\overline{f}_k^\ell}{f_e^\ell}$. Hence the right hand side of (28) is positive. Hence $\frac{\overline{f}_k^\ell}{f_e^\ell} < \frac{r}{v}$. □

Proof of Proposition 5: The Hessian of $\overline{L}(\cdot\,;\overline{q}(\gamma),\gamma)$ (expression (13)) is:

$$\overline{HL}(\gamma) = \begin{bmatrix} -\overline{\lambda}^\ell f_{ee}^\ell & 0 & -\overline{\lambda}^\ell f_{ek}^\ell & -f_e^\ell & 0 \\ 0 & \overline{\lambda}^h f_{ee}^h & \overline{\lambda}^h f_{ek}^h & 0 & f_e^h \\ -\overline{\lambda}^\ell f_{ek}^\ell & \overline{\lambda}^h f_{ek}^h & (\overline{\lambda}^h f_{kk}^h - \overline{\lambda}^\ell f_{kk}^\ell) & -f_k^\ell & f_k^h \\ -f_e^\ell & 0 & -f_k^\ell & 0 & 0 \\ 0 & f_e^h & f_k^h & 0 & 0 \end{bmatrix}.$$

By assumption, $\Delta^{\overline{HL}}(\gamma)$ is positive for all γ. Hence we can obtain the derivatives of the $\overline{\lambda}^i$'s w.r.t. q and γ, by applying the implicit function theorem to the first order condition (13). Because the inverse of $\overline{HL}(\gamma)$ is complex, we replace terms that we do not need to evaluate in the expression below by $\circledast$'s:

$$\begin{bmatrix} \partial\bar{e}^{\ell}/\partial\gamma & \partial\bar{e}^{\ell}/\partial q \\ \partial\bar{e}^{h}/\partial\gamma & \partial\bar{e}^{h}/\partial q \\ \partial\bar{k}/\partial\gamma & \partial\bar{k}/\partial q \\ \partial\bar{\lambda}^{\ell}/\partial\gamma & \partial\bar{\lambda}^{\ell}/\partial q \\ \partial\bar{\lambda}^{h}/\partial\gamma & \partial\bar{\lambda}^{h}/\partial q \end{bmatrix} = -\overline{HL}(\gamma)^{-1} \begin{bmatrix} \partial\bar{L}_{\bar{e}^{\ell}}/\partial\gamma & \partial\bar{L}_{\bar{e}^{\ell}}/\partial q \\ \partial\bar{L}_{\bar{e}^{h}}/\partial\gamma & \partial\bar{L}_{\bar{e}^{h}}/\partial q \\ \partial\bar{L}_{\bar{k}}/\partial\gamma & \partial\bar{L}_{\bar{k}}/\partial q \\ \partial\bar{L}_{\bar{\lambda}^{\ell}}/\partial\gamma & \partial\bar{L}_{\bar{\lambda}^{\ell}}/\partial q \\ \partial\bar{L}_{\bar{\lambda}^{h}}/\partial\gamma & \partial\bar{L}_{\bar{\lambda}^{h}}/\partial q \end{bmatrix} = \frac{\bar{\lambda}^{\ell}\bar{\lambda}^{h}}{\Delta^{\overline{HL}}(\gamma)} \times \cdots \tag{29}$$

$$\begin{bmatrix} \frac{(f_e^h)^2(f_k^\ell)^2}{\lambda^\ell\lambda^h} & \frac{f_e^h f_k^\ell f_k^h f_e^\ell}{\lambda^\ell\lambda^h} & \circledast & \frac{(f_e^h)^2\beta^\ell\mu(\ell)}{\lambda^h} - \frac{f_e^\ell\xi^h}{\lambda^\ell} & -\frac{f_e^\ell f_k^\ell\mu(h)}{\lambda^\ell} \\ \frac{f_e^h f_k^\ell f_k^h f_e^\ell}{\lambda^\ell\lambda^h} & \frac{(f_e^\ell)^2(f_k^h)^2}{\lambda^\ell\lambda^h} & \circledast & -\frac{f_e^h f_k^h\mu(\ell)}{\lambda^h} & \frac{(f_e^\ell)^2\beta^h\mu(h)}{\lambda^\ell} - \frac{f_e^h\xi^\ell}{\lambda^h} \\ \frac{-(f_e^h)^2 f_k^\ell f_e^\ell}{\lambda^\ell\lambda^h} & \frac{-(f_e^\ell)^2 f_k^h f_e^h}{\lambda^\ell\lambda^h} & \circledast & \frac{(f_e^h)^2\mu(\ell)}{\lambda^h} & \frac{(f_e^\ell)^2\mu(h)}{\lambda^\ell} \\ \frac{(f_e^h)^2\beta^\ell\mu(\ell)}{\lambda^h} - \frac{f_e^\ell\xi^h}{\lambda^\ell} & -\frac{f_e^h f_k^h\mu(\ell)}{\lambda^h} & \circledast & f_{ee}^\ell\xi^h - \frac{\bar{\lambda}^\ell(f_e^h)^2\psi(f,\ell)}{\lambda^h} & \mu(\ell)\mu(h) \\ -\frac{f_e^\ell f_k^\ell\mu(h)}{\lambda^\ell} & \frac{(f_e^\ell)^2\beta^h\mu(h)}{\lambda^\ell} - \frac{f_e^h\xi^\ell}{\lambda^h} & \circledast & \mu(\ell)\mu(h) & f_{ee}^h\xi^\ell - \frac{\bar{\lambda}^h(f_e^\ell)^2\bar{\lambda}^\ell}{\psi(f,h)} \end{bmatrix} \begin{bmatrix} v & 0 \\ -v & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & -1 \end{bmatrix}$$

Where for $i = h, \ell$:

$$\begin{aligned} \mu(i) &= f_{ee}^i f_k^i - f_{ek}^i f_e^i \\ \bar{\beta}^i &= \bar{k}/\bar{e}^i \\ \xi^i &= f_{ee}^i(f_k^i)^2 + f_{kk}^i(f_e^i)^2 - 2f_{ek}^i f_k^i f_e^i \quad = \quad \mu(i)(\bar{e}^i f_e^i + \bar{k} f_k^i)/\bar{k} \end{aligned} \tag{30}$$

$$\begin{aligned} \psi(f,i) &= f_{kk}^i f_{ee}^i - (f_{ek}^i)^2 \\ \Delta^{\overline{HL}}(\gamma) &= \bar{\lambda}^h(f_e^\ell)^2\xi^h \quad - \quad \bar{\lambda}^\ell(f_e^h)^2\xi^\ell \end{aligned} \tag{31}$$

We can now compute the partial derivatives of the marginal cost function at q, γ:

$$\begin{aligned} \frac{\partial\overline{MC}(q,\gamma)}{\partial\gamma} &= \left(\frac{\partial\bar{\lambda}^\ell(q,\gamma)}{\partial\gamma} - \frac{\partial\bar{\lambda}^h(q,\gamma)}{\partial\gamma}\right) && \text{which, after much tedious manipulation} \\ &= \frac{v(\bar{e}^h - \bar{e}^\ell)}{\bar{k}\Delta^{\overline{HL}}(\gamma)}\left(\bar{\lambda}^\ell(f_e^h)^2\mu(\ell) \quad - \quad \bar{\lambda}^h(f_e^\ell)^2\mu(h)\right) && \text{which, since } \mu(i) = \tfrac{\bar{k}\xi^i}{\bar{e}^i f_e^i + k f_k^i} \text{ (expr 30)} \\ &= \frac{v(\bar{e}^h - \bar{e}^\ell)}{\Delta^{\overline{HL}}(\gamma)}\left(\frac{\bar{\lambda}^\ell(f_e^h)^2\xi^\ell}{e^\ell f_e^\ell + k f_k^\ell} - \frac{\bar{\lambda}^h(f_e^\ell)^2\xi^h}{e^h f_e^h + k f_k^h}\right) && \text{which, since } f \text{ is homogenous of degree } \alpha \\ &= \frac{v(\bar{e}^h - \bar{e}^\ell)}{\alpha q(\gamma)\Delta^{\overline{HL}}(\gamma)}\left(\bar{\lambda}^\ell(f_e^h)^2\xi^\ell \quad - \quad \bar{\lambda}^h(f_e^\ell)^2\xi^h\right) && \text{which, from (31)} \\ &= \frac{v(\bar{e}^\ell - \bar{e}^h)}{\alpha q(\gamma)} \quad > \quad 0 \end{aligned}$$

Similarly, we have:

$$\begin{aligned} \frac{\partial\overline{MC}(q,\gamma)}{\partial q} &= \left(\frac{\partial\bar{\lambda}^\ell(q,\gamma)}{\partial q} - \frac{\partial\bar{\lambda}^h(q,\gamma)}{\partial q}\right) && \text{which, after similar manipulation} \\ &= \frac{(\bar{\lambda}^\ell - \bar{\lambda}^h)(1-\alpha)}{k\Delta^{\overline{HL}}(\gamma)}\left(\bar{\lambda}^\ell(f_e^h)^2\mu(\ell) \quad - \quad \bar{\lambda}^h(f_e^\ell)^2\mu(h)\right) \\ &= \frac{(\bar{\lambda}^\ell - \bar{\lambda}^h)(1-\alpha)}{\alpha q} \quad > \quad 0 \end{aligned}$$

proving the first sentence of the proposition. Now for all γ, at the profit maximizing level of q, we have $\overline{MC}(q,\gamma) = p$. Moroever, since $\frac{\partial \overline{MC}(\overline{q}(\gamma),\gamma)}{\partial q}$ is nonzero for all $\gamma \in [0,1]$, the implicit function theorem now implies the existence of neighborhoods U^γ of γ, U^q of $\overline{q}(\gamma)$ and a continuously differentiable function $\overline{q} : U^\gamma \to U^q$ such that $\overline{MC}(\overline{q}(\gamma),\gamma) = p$ on U^γ. It follows that $\overline{q}(\cdot)$ is continuously differentiable on $[0,1]$, with

$$\frac{d\overline{q}(\gamma)}{d\gamma} = -\frac{\partial \overline{MC}(\overline{q}(\gamma),\gamma)}{\partial \gamma} \Big/ \frac{\partial \overline{MC}(\overline{q}(\gamma),\gamma)}{\partial q} = \frac{v(\overline{e}^\ell - \overline{e}^h)}{(\alpha-1)(\overline{\lambda}^\ell - \overline{\lambda}^h)} \tag{18}$$

Expr. (18) is negative because $\overline{e}^\ell > \overline{e}^h$, $\overline{\lambda}^\ell > \overline{\lambda}^h$ & $\alpha < 1$, proving the second sentence of the proposition. □

Proof of Proposition 2: (7-$\overline{t}^\ell$) and (7-$\overline{t}^h$) follow immediately from Lemma 1. $\overline{q}^\ell < q^\star(\theta^\ell)$ is an immediate implication of Proposition 5 (specifically, $\frac{d\overline{q}(\gamma)}{d\gamma} < 0$ for all γ and $\overline{q}^\ell = q^\star(\theta^\ell) + \int_0^1 \frac{d\overline{q}(\gamma')}{d\gamma} d\gamma'$.) Part 3 is a special case of Proposition 4, for $\gamma = 1$. □

Proof of Proposition 7: We begin by comparing (18) for the optimal restrictive contract to the corresponding expression, (21), for the optimal basic contract:

$$\frac{d\overline{q}(\gamma)}{d\gamma} - \frac{d\tilde{q}(\gamma)}{d\gamma} = \frac{1}{1-\alpha}\left(\frac{(v\tilde{e}^\ell + r\tilde{k}^\ell)\left(1-\vartheta^{1/\alpha}\right)}{p} - \frac{v(\overline{e}^\ell - \overline{e}^h)}{\overline{\lambda}^\ell - \overline{\lambda}^h}\right) \tag{32}$$

As we observed on page 14, $(\overline{\lambda}^\ell - \overline{\lambda}^h) = p$ at the optimal restrictive contract. for all Hence for all $\gamma \geq 0$, (38) reduces to

$$\frac{d\overline{q}(\gamma)}{d\gamma} - \frac{d\tilde{q}(\gamma)}{d\gamma} = \frac{1}{p(1-\alpha)}\left(\tilde{C}^I(\tilde{q}(\gamma),\gamma) - \overline{C}^I(\overline{q}(\gamma),\gamma)\right) \tag{32$'$}$$

We use this result to show that $\overline{q}(\Phi) > \tilde{q}(\Phi)$. First note that because $\overline{C}^I(\cdot,\gamma)$ increases with q, and $\overline{q}(\cdot)$ is a continuous function of γ (Proposition 5), inequality (22) implies existence of a continuous, positive function $\varepsilon(\cdot)$ of γ such that

$$\forall \gamma \geq 0, \frac{d\overline{q}(\gamma)}{d\gamma} - \frac{d\tilde{q}(\gamma)}{d\gamma} > 0 \text{ if } \overline{q}(\gamma) \leq \tilde{q}(\gamma) + \varepsilon(\gamma). \tag{33}$$

Now suppose that there exists $\gamma \in [0,\Phi]$ such that $\overline{q}(\gamma) \leq \tilde{q}(\gamma)$ and let $\gamma^\star$ be the infimum of such γ's. Since $\overline{q}(\cdot) - \tilde{q}(\cdot)$ is continuous with respect to γ, $\overline{q}(\gamma^\star) = \tilde{q}(\gamma^\star)$. We will now establish a contradiction.

Since $\tilde{q}(0) = \overline{q}(0)$, property (33) plus continuity implies the existence of $\underline{\gamma} > 0$ such that for all $\gamma \in (0,\underline{\gamma}]$, $\tilde{q}(\gamma) = \int_0^\gamma \frac{d\tilde{q}(\gamma')}{d\gamma} d\gamma' < \int_0^\gamma \frac{d\overline{q}(\gamma')}{d\gamma} d\gamma' = \overline{q}(\gamma)$. Therefore, $\gamma^\star > \underline{\gamma} > 0$. Now by assumption $\overline{q}(\cdot) > \tilde{q}(\cdot)$ on $[0,\gamma^\star)$ and $\overline{q}(\cdot) - \tilde{q}(\cdot)$ is continuous with respect to γ, there exists $\overline{\gamma} < \gamma^\star$ such that $\overline{q}(\cdot) - \varepsilon(\cdot) < \tilde{q}(\cdot) < \overline{q}(\cdot)$ on $[\overline{\gamma},\gamma^\star)$. Therefore, from (33),

$$(\overline{q}(\gamma^\star) - \tilde{q}(\gamma^\star)) = (\overline{q}(\overline{\gamma}) - \tilde{q}(\overline{\gamma})) + \int_{\overline{\gamma}}^{\gamma^\star}\left(\frac{d\overline{q}(\gamma')}{d\gamma} - \frac{d\tilde{q}(\gamma')}{d\gamma}\right)d\gamma' > 0$$

contradicting the existence of $\gamma \in [0,\Phi]$ such that $\overline{q}(\gamma^\star) = \tilde{q}(\gamma^\star)$. □

Proof of Proposition 8: As usual, all symbols with bars (tildes) over them are part of the solution to the restrictive (basic) contract. The proof involves five steps:

Step 1: Preliminaries: bounding key variables.
Pick $g \in G$, let $\tilde{\beta}$ denote the neoclassical input mix (see Remark 1) for $f^\ell \equiv g$, let $\ddot{g}$ denote the composite input function corresponding to g and let $\ddot{e}(q)$ denote the level of composite input required to produce q with $\ddot{g}$. Rewriting (20) (p. 16), the *basic marginal cost function* for $\ddot{g}$ is

$$\widetilde{MC}(q,\gamma) \quad = \quad \ddot{v}\left\{1+\gamma\left(1-\Theta^{1/\alpha}\right)\right\}\left(\ddot{g}'(\ddot{e}^\ell(q))\right)^{-1}. \tag{20}$$

Proceeding as on page 14, $\widetilde{MC}(\tilde{q}(\gamma),\gamma) = p$ at the principal's optimum, where $\tilde{q}(\gamma)$ is the profit maximizing level of output produced by the agent ℓ at price p under the basic contract. In this proof, we shall abbreviate $\ddot{e}^i(\tilde{q}(\gamma))$ to $\ddot{e}^i(\gamma)$. Similarly, let $\bar{e}^i(\gamma) = \bar{e}^i(\bar{q}(\gamma))$ and $\bar{k}(\gamma) = \bar{k}(\bar{q}(\gamma))$. Manipulating (20), we obtain:

$$\ddot{g}'(\ddot{e}^\ell(\gamma)) \quad = \quad \frac{\ddot{v}}{p}\left\{1+\gamma\left(1-\Theta^{1/\alpha}\right)\right\}. \qquad \text{Since } \ddot{g}' \text{ is homog of degree } \alpha-1\text{, we have}$$

$$\ddot{e}^\ell(\gamma)^{\alpha-1}\ddot{g}'(1) \quad = \quad \frac{\ddot{v}}{p}\left\{1+\gamma\left(1-\Theta^{1/\alpha}\right)\right\} \qquad \text{Since } \ddot{g} \text{ is h.d. } \alpha,\ \ddot{g}'(1)=\alpha\ddot{g}(1) \text{ and hence}$$

$$\ddot{e}^\ell(\gamma)^{\alpha-1}\alpha\ddot{g}(1) \quad = \quad \frac{\ddot{v}}{p}\left\{1+\gamma\left(1-\Theta^{1/\alpha}\right)\right\} \qquad \text{so that}$$

$$\ddot{e}^\ell(\gamma) \quad = \quad \left[\frac{\ddot{v}}{p\alpha\ddot{g}(1)}\left\{1+\gamma\left(1-\Theta^{1/\alpha}\right)\right\}\right]^{1/(\alpha-1)} \tag{34}$$

Since $g \in G$, the right hand side of (34) is bounded below by $\check{e} = \left(\frac{v\check{\omega}+r}{p\check{\omega}^4}\right)^{-1/\check{\omega}}$ and above by $\hat{e} = \min[1, \left(p/v\check{\omega}\right)^{1/\check{\omega}}]$. To verify the lower bound, note that $\gamma < \frac{1-\check{\omega}}{\check{\omega}}$ and $\left(1-\Theta^{1/\alpha}\right) < 1$. Hence $\left\{1+\gamma\left(1-\Theta^{1/\alpha}\right)\right\} < 1+\frac{1-\check{\omega}}{\check{\omega}} = 1/\check{\omega}$. Moreover, since $\alpha, \ddot{g}(1) \geq \check{\omega}$, and $\ddot{v} \leq v + r/\check{\omega}$, we have an lower bound of $\left(\frac{v\check{\omega}+r}{p\check{\omega}^4}\right)^{-1/\check{\omega}}$. To verify the upper bound, note that since y is negative, the coefficient $\left(\frac{\ddot{v}}{p\alpha\ddot{g}(1)}\right)^y$ is maximized when $\frac{\ddot{v}}{p\alpha\ddot{g}(1)}$ is minimized and y is maximized in absolute value. Moreover, we have $\left(\frac{\ddot{v}}{p\alpha\ddot{g}(1)}\right)^y \geq \left(\frac{v\check{\omega}}{p(1-\check{\omega})}\right)^y \geq \left(\frac{v\check{\omega}}{p}\right)^{-1/\check{\omega}}$. On the other hand $\left\{1+\gamma\left(1-\Theta^{1/\alpha}\right)\right\}^{1/(\alpha-1)} \leq 1$. Hence, we have an upper bound of $\left(\frac{v\check{\omega}}{p}\right)^{-1/\check{\omega}}$. Moreover, $\frac{d\ddot{e}^\ell(\gamma)}{d\gamma}$ is bounded above by $\hat{d}_e = \frac{1}{\check{\omega}}\left(p/v\check{\omega}\right)^{1/\check{\omega}}$. To verify this, note that $\frac{d\ddot{e}^\ell(\gamma)}{d\gamma} = \left(\frac{\ddot{v}}{p\alpha\ddot{g}(1)}\right)^y (1+\gamma x)^y$ where, $x = \left(1-\Theta^{1/\alpha}\right)$ and $y = 1/(\alpha-1) < -1$. The argument we used to construct $\hat{e}$ shows that the upper bound on the coefficient is $\left(\frac{v\check{\omega}}{p}\right)^{-1/\check{\omega}}$. Next note that $\left|\frac{d(1+\gamma x)^y}{d\gamma}\right| = \left|yx(1+\gamma x)^{y-1}\right| \leq |yx| \leq |y| \leq 1/\check{\omega}$. Hence we have $\left|\frac{d\ddot{e}^\ell(\gamma)}{d\gamma}\right| \leq \frac{1}{\check{\omega}}\left(\frac{v\check{\omega}}{p}\right)^{-1/\check{\omega}}$. Finally, from (21), $\left|\frac{d\tilde{q}(\gamma)}{d\gamma}\right|$ is bounded below by $\check{d}_q = \check{e}(v+r\check{\omega})\check{\omega}/p$. To verify this bound, note from (21) that

$$\left|\frac{d\tilde{q}(\gamma)}{d\gamma}\right| \quad = \quad \left|\frac{\ddot{v}\ddot{e}^\ell(q)\left(1-\Theta^{1/\alpha}\right)}{(\alpha-1)p}\right| \quad \geq \quad \left|\frac{\check{e}(v+\check{\omega}r)\left(1-(1-\check{\omega})^{1/\alpha}\right)}{(\alpha-1)p}\right| \quad \geq \quad \frac{\check{e}\check{\omega}(v+\check{\omega}r)}{p}$$

Step 2: Given $\varepsilon \in (0,1]$, $q \in \mathbb{R}$, and $g \in G$, let $(\tilde{e}^\ell, \tilde{k}^\ell)$ denote the cost-minimal vector for producing q with technology g. There exists $n, N \in \mathbb{N}$ such that if $\bar{\sigma}(g) < 1/n$ and agent ℓ uses the input vector $(\bar{e}^\ell, \bar{k}^\ell)$ in the solution to the restrictive FOC (13) for some $\gamma \in (0,\Phi]$, then (a) $(\tilde{e}^\ell - \bar{e}^\ell) < \varepsilon$ and (b) $\bar{k}^\ell < N$.

Proof of Step 2. Fix $\varepsilon \in (0,1]$, $q \in \mathbb{R}$, and $g \in G$. For $i = \ell, h$, let $\mathsf{mrs}^i(e,k)$ denote the marginal rate of substitution $\frac{f_k^i(e,k)}{-f_e^i(e,k)}$. Observe first that substituting the expressions for the λ's obtained in (14) into the expression for L_k obtained in (13) yields:

$$\begin{aligned} 1 - \frac{v}{r}\mathsf{mrs}^\ell(\bar{e}^\ell,\bar{k}^\ell) &= \frac{\gamma v}{r}\left(\mathsf{mrs}^\ell(\bar{e}^\ell,\bar{k}^\ell) - \mathsf{mrs}^h(\bar{e}^\ell,\bar{k}^\ell)\right) && \text{or, rearranging} \\ 1 + \frac{\gamma v}{r}\mathsf{mrs}^h(\bar{e}^\ell,\bar{k}^\ell) &= (1+\gamma)\frac{v}{r}\mathsf{mrs}^\ell(\bar{e}^\ell,\bar{k}^\ell) && (35) \end{aligned}$$

Since the left hand side of (35) is bounded below by unity, and $1+\gamma \le 1+\Phi \le (1+\breve{\omega})/\breve{\omega}$, the expression $\mathsf{mrs}^\ell(\bar{e}^\ell,\bar{k}^\ell)$ is bounded below by $\frac{r\breve{\omega}}{v(1+\breve{\omega})}$. Moroever, $\mathsf{mrs}^\ell(\tilde{e}^\ell,\tilde{k}^\ell) = \frac{r}{v}$. Pick n, N sufficiently large that $\frac{\hat{e}}{n\breve{\omega}} < \varepsilon$ (where $\hat{e}$ was defined on p. 25) while $\frac{\hat{e}}{\breve{\omega}}\left(1-\frac{\varepsilon}{\hat{e}}\right)^{-1} < N$. For all (e,k) such that $g(e,k) = q$, we have $-\frac{d(k/e)}{d\left(\mathsf{mrs}^\ell(e,k)\right)}\frac{\mathsf{mrs}^\ell(e,k)}{k/e}\Big|_{g\text{ const}} \le \frac{1}{n}$ so that $\frac{d(k/e)}{d\left(\mathsf{mrs}^\ell(e,k)\right)}\Big|_{g\text{ const}} \ge -\frac{1}{n}\frac{k/e}{\mathsf{mrs}^\ell(e,k)}$ and hence

$$\begin{aligned} \frac{\tilde{k}^\ell}{\tilde{e}^\ell} - \frac{\bar{k}^\ell}{\bar{e}^\ell} &\ge -\frac{1}{n}\int_{\mathsf{mrs}^\ell(\bar{e}^\ell,\bar{k}^\ell)}^{\mathsf{mrs}^\ell(\tilde{e}^\ell,\tilde{k}^\ell)}\left(\frac{k'/e'}{\mathsf{mrs}^\ell(e',k')}\right) d\left(\mathsf{mrs}^\ell(e',k')\right) \\ &\ge -\frac{1}{n}\frac{\bar{k}^\ell/\bar{e}^\ell}{\mathsf{mrs}^\ell(\bar{e}^\ell,k^\ell)}\left(\mathsf{mrs}^\ell(\tilde{e}^\ell,\tilde{k}^\ell) - \mathsf{mrs}^\ell(\bar{e}^\ell,\bar{k}^\ell)\right) \\ &\ge -\frac{1}{n}\frac{v}{r}\frac{\bar{k}^\ell/\bar{e}^\ell}{\breve{\omega}/(1+\breve{\omega})}\frac{r}{v}\left(1-\frac{\breve{\omega}}{1+\breve{\omega}}\right) = -\frac{1}{n}\frac{\bar{k}^\ell/\bar{e}^\ell}{\breve{\omega}} && \text{or} \\ \frac{\tilde{k}^\ell}{\tilde{e}^\ell} &\ge \left(1-\frac{1}{n\breve{\omega}}\right)\frac{\bar{k}^\ell}{\bar{e}^\ell} && \text{so that} \\ \tilde{k}^\ell\bar{e}^\ell &\ge \left(1-\frac{1}{n\breve{\omega}}\right)\bar{k}^\ell\tilde{e}^\ell && (36) \end{aligned}$$

Since $\bar{k}^\ell \ge \tilde{k}^\ell$, (36) implies that

$$\begin{aligned} \tilde{k}^\ell\bar{e}^\ell &\ge \left(1-\frac{1}{n\breve{\omega}}\right)\tilde{k}^\ell\tilde{e}^\ell \quad \text{so that} \quad \bar{e}^\ell \ge \left(1-\frac{1}{n\breve{\omega}}\right)\tilde{e}^\ell && \text{and hence} \\ \tilde{e}^\ell - \bar{e}^\ell &\le \frac{1}{n\breve{\omega}}\tilde{e}^\ell \le \frac{1}{n\breve{\omega}}\hat{e} \le \varepsilon \end{aligned}$$

Inequality (36), together with the facts that $\bar{e}^\ell < \tilde{e}^\ell$, $\tilde{k}^\ell \le \frac{\hat{e}}{\breve{\omega}}$, and $\frac{1}{n\breve{\omega}} < \frac{\varepsilon}{\hat{e}}$ imply that

$$\bar{k}^\ell \le \tilde{k}^\ell\frac{\bar{e}^\ell}{\tilde{e}^\ell}\left(1-\frac{1}{n\breve{\omega}}\right)^{-1} \le \frac{\hat{e}}{\breve{\omega}}\left(1-\frac{1}{n\breve{\omega}}\right)^{-1} \le \frac{\hat{e}}{\breve{\omega}}\left(1-\frac{\varepsilon}{\hat{e}}\right)^{-1} \le N \qquad (37)$$

■

Step 3: Fix $\varepsilon, \delta, \eta > 0$. There exists $n \in \mathbb{N}$ such that if $g \in G$ with $\bar{\sigma}(g) \le \frac{1}{n}$ and $\mathsf{mrs}^\ell(e^0,k^0) = \eta$, if $k^1 \le k^0 + \delta$ and if $g(e^1,k^1) \ge g(e^0,k^0)$, then $e^1 > e^0 - \varepsilon$.

Intuition for the proof of Step 3 is provided by Fig. 5.

(1) From the starting point (e,k), move northwest along the budget line to $(e - 0.5\varepsilon, k + 0.5\varepsilon/\eta)$, labelled as A in the figure

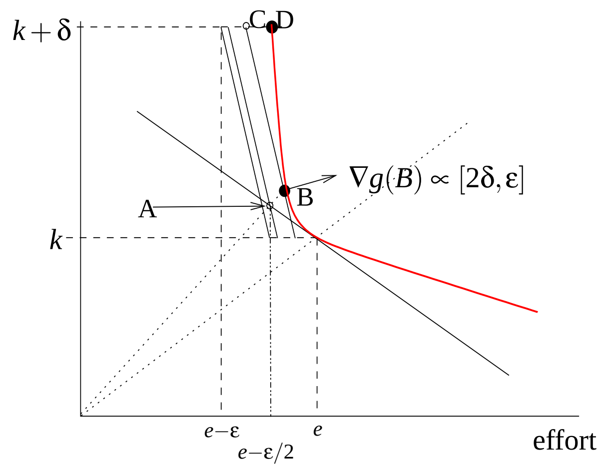

FIGURE 5. Intuition for Step 3

(2) Because A lies on the same iso-cost line as (e,k) and $g \in G$ we know that the level set thru (e,k) intersects the ray thru the origin and A at a point to the north east of A. Call this point B.
(3) Figure out what $\bar{\sigma}$ has to be to ensure that the gradient vector thru B is proportional to $[2\delta,\varepsilon]$, so that as drawn, the tangent plane thru B has slope $\frac{2\delta}{\varepsilon}$.
(4) Now by the now familiar argument, if we go out along the tangent plane to B, it has to intersect the horizontal line starting at $k+\delta$ at a point to the right of $e-\varepsilon$. Call this intersection point C.
(5) By strict quasi-concavity, the the level set thru (e,k) has to intersect the horizontal line starting at $k+\delta$ to the right of C. Call this point D. Hence D is a point $k+\delta, e-\omega$, where $\omega < \varepsilon$, proving the step.

Proof of Step 3. Given $\varepsilon,\delta,\eta > 0$, fix $(e^0,k^0) \gg 0$. Let $\mu^0 = \frac{k^0}{e^0}$ and $\mu^\dagger = \frac{k^0+\varepsilon/2\eta}{e^0-\varepsilon/2}$. Now pick $n \in \mathbb{N}$ such that $\frac{\mu^\dagger\eta}{\mu^\dagger+n(\mu^\dagger-\mu^0)} < \frac{\varepsilon}{2\delta}$ and a production function $g \in G$ with $\bar{\sigma}(g) \leq \frac{1}{n}$ and $\mathsf{mrs}^\ell(e^0,k^0) = \eta$. By construction, the vector $(k^0+\varepsilon/2\eta, e^0-\varepsilon/2)$—which we used to define $\mu^\dagger$—belongs to the line perpendicular to the gradient of g through (e^0,k^0), so that, since g is strictly quasi-concave, $g(e^0-\varepsilon/2, k^0+\varepsilon/2\eta) < g(e^0,k^0)$. Define $(e^\dagger,k^\dagger)$ by: $\frac{k^\dagger}{e^\dagger} = \mu^\dagger$ and $g(e^\dagger,k^\dagger) = g(e^0,k^0)$, so that $(e^\dagger,k^\dagger) \gg (e^0-\varepsilon/2, k^0+\varepsilon/2\eta)$. Let $\nu = \mathsf{mrs}^\ell(e^0-\varepsilon/2, k^0+\varepsilon/2\delta)$. Since g is homothetic, $\nu = \mathsf{mrs}^\ell(e^\dagger,k^\dagger)$. Now, for all e,k, we have $\frac{-d(k/e)}{d\left(\mathsf{mrs}^\ell(e,k)\right)} \frac{\mathsf{mrs}^\ell(e,k)}{k/e}\Big|_{g\text{ const}} \leq \frac{1}{n}$ so that $\frac{d\left(\mathsf{mrs}^\ell(e,k)\right)}{d(k/e)}\Big|_{g\text{ const}} \leq -n\frac{\mathsf{mrs}^\ell(e,k)}{k/e}$, and hence

$$
\begin{aligned}
\nu - \eta &= \int_{\mu^0}^{\mu^\dagger} \left(\frac{d\left(\mathsf{mrs}^\ell(e',k')\right)}{d\left(k'/e'\right)}\Bigg|_{g\text{ const}} \right) d\left(k'/e'\right) \\
&\leq -n\int_{\mu^0}^{\mu^\dagger} \left(\frac{\mathsf{mrs}^\ell(e',k')}{k'/e'}\Bigg|_{g\text{ const}} \right) d\left(k'/e'\right) \\
&\leq -n(\mu^\dagger-\mu^0)\frac{\min\left\{\mathsf{mrs}^\ell(e',k') : \frac{k'}{e'} \in [\mu^0,\mu^\dagger], g(e',k') = q\right\}}{\max\left\{(k'/e') : \frac{k'}{e'} \in [\mu^0,\mu^\dagger], g(e',k') = q\right\}} = -n\frac{\nu(\mu^\dagger-\mu^0)}{\mu^\dagger} \qquad \text{so that} \\
\nu &\leq \frac{\mu^\dagger\eta}{\mu^\dagger+n(\mu^\dagger-\mu^0)} \quad \text{which is, by assumption} \quad \leq \frac{\varepsilon}{2\delta}.
\end{aligned}
$$

Now pick $dk > 0$ such that $[dk, -\varepsilon/2]'[g_k(e^\dagger, k^\dagger), g_e(e^\dagger, k^\dagger)] = 0$ so that $dk = \varepsilon/(2\nu) \geq \delta$. Since g is strictly quasi-concave, $g(e^\dagger - \varepsilon/2, k^\dagger + dk) < g(e^\dagger, k^\dagger)$. But since $e^\dagger > e^0 - \varepsilon/2$ and $k^\dagger > k^0$, it follows that $(e^0 - \varepsilon, k^0 + \delta) \ll (e^\dagger - \varepsilon/2, k^\dagger + dk)$ and hence $g(e^0 - \varepsilon, k^0 + \delta) < g(e^\dagger, k^\dagger)$. Conclude that for (e^1, k^1) with $k^1 \leq k^0 + \delta$, $g(e^1, k^1) \geq g(e^0, k^0)$ implies that $e^1 > e^0 - \varepsilon$. ■

Step 4: There exists $\delta \in \mathbb{N}$ and $n \in \mathbb{N}$ such that for all $g \in G$ with $\bar{\sigma}(g) \leq 1/n$, the level of output under the restrictive contract, using this technology, exceeds by at least δ the level of output under the basic contract.

Proof of Step 4. Let $\delta = \frac{r\breve{\omega}\breve{e}}{4} \min\left[\frac{\check{d}_q}{v\hat{d}_e}, \frac{\breve{\omega}^2}{p}\right]$, where where $\breve{e}$, $\check{d}_q$ and $\hat{d}_e$ were constructed on p. 25. Invoking Step 2 and Step 3, pick n sufficiently large that for $i = \ell, h$ and all $g \in G$ with $\bar{\sigma}(g) < 1/n$, $(\tilde{e}^i - \bar{e}^i) < r\breve{\omega}^2\breve{e}/4v$. where $\breve{e}$ was constructed on p. 25. Note from Proposition 7 that $\left(\tilde{q}(\cdot) - \bar{q}(\cdot)\right)$ is negative and continuous on $(0, \Phi]$. Hence there exists a continuous function $\upsilon : [0, \Phi] \to [0, \Phi]$, with $\upsilon(\gamma) < \gamma$ on $(0, \Phi]$, such that for all γ, $\tilde{q}(\upsilon(\gamma)) = \bar{q}(\gamma)$. Subtracting (21) from (18), we now obtain:

$$\begin{aligned}
\frac{d\bar{q}(\gamma)}{d\gamma} - \frac{d\tilde{q}(\gamma)}{d\gamma} &= \frac{1}{p(1-\alpha)}\Big((v\tilde{e}^\ell(\gamma) + r\tilde{k}^\ell(\gamma))\big(1 - \Theta^{1/\alpha}\big) - v(\bar{e}^\ell(\gamma) - \bar{e}^h(\gamma))\Big) \\
&= \frac{1}{p(1-\alpha)}\Bigg(\big(1-\Theta^{1/\alpha}\big)\Big[r\tilde{k}^\ell(\gamma) + \underbrace{v(\tilde{e}^\ell(\gamma) - \tilde{e}^\ell(\upsilon(\gamma)))}_{\text{Term 1}}\Big] \\
&\qquad - v\Big[\underbrace{(\bar{e}^\ell(\gamma) - \tilde{e}^\ell(\upsilon(\gamma)))}_{\text{Term 2}} + \underbrace{(\bar{e}^h(\gamma) - \tilde{e}^h(\upsilon(\gamma)))}_{\text{Term 3}}\Big]\Bigg) \qquad (38)
\end{aligned}$$

Since $\tilde{\beta}$ is bounded below by $\breve{\omega}$, $r\tilde{k}^\ell(\gamma)$ is bounded below by $r\breve{\omega}\breve{e}$. There are now two possibilities to consider:

(1) Suppose that $\gamma - \upsilon(\gamma) \geq r\breve{\omega}\breve{e}/(4v\hat{d}_e)$ for some $\gamma^* \in [0, \Phi]$. In this case, since $\upsilon(\cdot)$ is continuous and $\left(\frac{d\bar{q}(\gamma)}{d\gamma} - \frac{d\tilde{q}(\gamma)}{d\gamma}\right)$ is positive whenever $\gamma - \upsilon(\gamma) < r\breve{\omega}\breve{e}/(4v\hat{d}_e)$, it follows that $\gamma - \upsilon(\gamma) \geq r\breve{\omega}\breve{e}/(4v\hat{d}_e)$ for all $\gamma \in [\gamma^*, \Phi]$. But since $|\frac{d\tilde{q}(\gamma)}{d\gamma}| \geq \check{d}_q$, we have

$$\Phi - \upsilon(\Phi) \geq \frac{r\breve{\omega}\breve{e}}{4v\hat{d}_e} \implies \bar{q} - \tilde{q} = \tilde{q}(\upsilon(\Phi)) - \tilde{q}(\Phi) = -\int_{\upsilon(\Phi)}^{\Phi} \frac{d\tilde{q}(\gamma')}{d\gamma} d\gamma' \geq \frac{r\breve{\omega}\breve{e}\check{d}_q}{4v\hat{d}_e}$$

(2) Suppose that $\gamma - \upsilon(\gamma) < r\breve{\omega}\breve{e}/(4v\hat{d}_e)$ on $[0, \Phi]$. Since $\left|\frac{d\ddot{e}^\ell(\gamma)}{d\gamma}\right| \leq \hat{d}_e$, the absolute value of Term 1 is bounded above by $r\breve{\omega}\breve{e}/4v$ whenever $\gamma - \upsilon(\gamma) < r\breve{\omega}\breve{e}/(4v\hat{d}_e)$. Moreover, we have chosen n sufficiently large that Terms 2 and 3 are both bounded above by $(r\breve{\omega}^2\breve{e}/4v)$. It now follows from (38) that whenever $\gamma - \upsilon(\gamma) < r\breve{\omega}\breve{e}/(4v\hat{d}_e)$,

$$\frac{d\bar{q}(\gamma)}{d\gamma} - \frac{d\tilde{q}(\gamma)}{d\gamma} \geq \frac{r}{p(1-\alpha)}\left(\big(1-\Theta^{1/\alpha}\big)\Big[\breve{\omega}\breve{e} - \breve{\omega}\breve{e}/4\Big] - \breve{\omega}\Big[\breve{\omega}\breve{e}/4 + \breve{\omega}\breve{e}/4\Big]\right)$$

but since $\Theta \leq 1 - \breve{\omega}$ and $\alpha < 1$, $\left(1 - \Theta^{1/\alpha}\right) \geq \breve{\omega}$, so that

$$\frac{d\bar{q}(\gamma)}{d\gamma} - \frac{d\tilde{q}(\gamma)}{d\gamma} \geq \frac{r\breve{e}\breve{\omega}^2}{4p(1-\alpha)} \geq \frac{r\breve{e}\breve{\omega}^2}{4p} \qquad (39)$$

Since $\Phi \geq \breve{\omega}$, it now follows from (39) that

$$\bar{q} \quad - \quad \tilde{q} \qquad \geq \qquad \int_0^{\Phi} \Big(\frac{d\bar{q}(\gamma')}{d\gamma} \quad - \quad \frac{d\tilde{q}(\gamma')}{d\gamma} \Big) d\gamma' \qquad \geq \qquad \frac{r\breve{e}\breve{\omega}^3}{4p}$$

We have established, therefore, that for all $g \in G$ with $\bar{\sigma}(g) < 1/n$,

$$\bar{q} - \tilde{q} \quad \geq \quad \delta \quad = \quad \frac{r\breve{\omega}\breve{e}}{4} \min \left[\frac{\breve{d}_q}{v\hat{d}_e}, \frac{\breve{\omega}^2}{p} \right]$$

■

Step 5: Proof of the proposition.

From Step 4, we can pick $N \in \mathbb{N}$ and $\delta > 0$, such that for all $g \in G$ with $\bar{\sigma}(g) < 1/N$. $(\bar{q} - \tilde{q}) \geq \delta$. From Step 2, we can pick $n \geq N$ such that for all $g \in G$ with $\bar{\sigma}(g) < 1/n$, $\bar{e}^{\ell} - \tilde{e}^{\ell} < p\delta\breve{\omega}/v$. We now have

$$\begin{aligned}
\Delta\text{SS} &= (p\bar{q} - \bar{C}^P(\bar{q})) - (p\tilde{q} - \tilde{C}^P(\tilde{q})) \\
&= \underbrace{(p\bar{q} - \tilde{C}^P(\bar{q})) \quad - \quad (p\tilde{q} - \tilde{C}^P(\tilde{q}))}_{\text{output effect}} \quad - \quad \underbrace{(\bar{C}^P(\bar{q}) - \tilde{C}^P(\bar{q}))}_{\text{input mix effect}} \\
&= p(\bar{q} - \tilde{q})(1 - \alpha) \quad - \quad \big(v(\bar{e}^{\ell} - \tilde{e}^{\ell}) + r(\bar{k}^{\ell} - \tilde{k}^{\ell})\big) \\
&\geq p\delta\breve{\omega} \quad - \quad v(\bar{e}^{\ell} - \tilde{e}^{\ell}) \quad > \quad 0
\end{aligned}$$

□

References

Aghion, Phillippe and Jean Tirole, "Formal and Real Authority in Organizations," *J. Political Economy*, 1997, *105* (1).

Baron, David P. and David Besanko, "Monitoring, Moral Hazard, Asymmetric Information, and Risk Sharing in Procurement Contracting," *RAND Journal of Economics*, Winter 1987, *18* (4), 509–532.

Caillaud, Bernard and Ben Hermalin, "Hidden-Information Agency," mimeo, University of California at Berkeley 2000. http://faculty.haas.berkeley.edu/hermalin/mechread.pdf.

Coase, Ronald, "The Nature of the Firm," *Economica*, 1937, *4*, 386–405.

Dessein, Wouter, "Authority and Communication in Organizations," *Review of Economic Studies*, 2002, *69*, 811–838.

Dewatripont, M. and Eric Maskin, "Contractual Contingencies and Renegotiaion," *RAND Journal of Economics*, Winter 1995, *26* (4), 704–719.

Hart, Oliver and John Moore, "Property Rights and the Nature of the Firm," *J. Political Economy*, 1990, *98* (6), 1119–1158.

Khalil, Fahad and Jacques Lawarree, "Input versus Output Monitoring: Who is the Residual Claimant?," *Journal of Economic Theory*, June 1995, *66* (1), 139–157.

Laffont, Jean-Jacques and Jean Tirole, "Using Cost Observation to Regulate Firms," *J. Political Economy*, June 1986, *94* (3), 614–641.

Lewis, Tracy R. and David E. Sappington, "Technological Change and the Boundaries of the Firm," *American Economic Review*, September 1991, *81* (4), 887–900.

Maskin, Eric and John Riley, "Input versus Output Incentive Schemes," *J. of Public Economics*, October 1985, *28* (1), 1–23.

Myerson, Roger, "Incentive Compatibility and the Bargaining Problem," *Econometrica*, January 1979, *47* (1), 61–74.

Salanie, Bernard, *The Economics of Contracts: A Primer*, Cambridge, MA: MIT Press, 1997.

Silberberg, Eugene, *The Structure of Economics: A Mathematical Analysis*, second ed., McGraw-Hill, Inc., 1990.

Tirole, Jean, "Incomplete Contracts: Where Do We Stand?," *Econometrica*, July 1999, *67* (4), 741–781.

Williamson, Oliver E., *The Economic Institutions of Capitalism*, New York: The Free Press, 1985.