Seminar Paper No. 480

"X"-INEFFICIENCY AND INTERNATIONAL COMPETITION

by

Henrik Horn, Harald Lang and Stefan Lundgren
Seminar Paper No. 480

"X"-INEFFICIENCY AND INTERNATIONAL COMPETITION

by

Henrik Horn, Harald Lang and Stefan Lundgren

Seminar Papers are preliminary material circulated to stimulate discussion and critical comment.

November 1990

Institute for International Economic Studies
S-106 91 Stockholm
Sweden
"X" -INEFFICIENCY AND INTERNATIONAL COMPETITION#

by

Henrik Horn*, Harald Lang*0, and Stefan Lundgren*

*Institute for International Economic Studies
0Department of Economics

Stockholm University

November 1990

# We are grateful for helpful discussions with Jim Anderson, Richard Baldwin, Wilfred Ethier, Chaim Fershtman, Harry Flam, Jan Haaland, Elhanan Helpman, Bengt Holmström, Karl Jungenfelt, James Levinsohn, James Markusen, and Dani Rodrik. We also benefitted from the comments provided by seminar participants at IIES, the Stockholm School of Economics, the Industrial Institute for Economic and Social Research, and at the ERWIT meeting in Aix-en-Provence, 1990.
I. INTRODUCTION

In the policy debate one often encounters arguments in favor of trade liberalization that rely on its alleged positive effect on the internal organization of firms and on the effort supplied by employees. One version of this type of argument is the frequent claim that international economic integration, by increasing competition in product markets, forces firms to reduce organizational (or "X"-) inefficiencies, and that this enhances welfare. This argument played a prominent role in, for instance, the U.K. discussions about whether or not it should join the EEC, with the benefits of a "cold shower" of international competition advanced as a reason for EEC membership. Today the argument reappears in the debate about the effects of the Internal Market. According to the European Commission (1988),

"...the new competitive pressures brought about by the completion of the internal market can be expected to...produce appreciable gains in internal efficiency...[which will] constitute much of what can be called the dynamic effects of the internal market..." (ibid., p. 126).

and furthermore

"[t]he costs of X-inefficiency may often be as great as those resulting from unexploited economies of scale..." (ibid., p. 18).

Similarly, it seems to be widely believed that employees are induced to work harder in firms that are exposed to international competition, than when firms face less fierce competition in their output markets.

In sharp contrast to the importance attached to these kinds of arguments in the policy debate stands the extremely limited attention that they have received in trade theory. It is no exaggeration that trade theory has almost entirely disregarded questions concerning the role of trade for the allocation of resources within firms. The main
exception is a small literature in the 1970s on "X--inefficiency" and trade. In this literature it was assumed that managers *cum* owners choose effort supply in order to maximize their utility, which depends both on profit income and leisure.¹ By foregoing leisure, that is, by expending more effort, profit income could be increased. The literature followed the more informal arguments by identifying reduced internal inefficiency with increased supply of managerial effort, even though it was sometimes emphasized that this usage of the term "inefficiency" does not accord with its conventional usage in economic theory.² Hence, while this literature does consider the above--mentioned informal arguments in a rigorous fashion, its depiction of the internal inefficiency in firms is not very compelling. To the best of our knowledge a formal analysis of these issues within models where firms are internally inefficient in some more reasonable sense, is still lacking. The purpose of this paper is to take a small step in order to remedy this weakness.

A systematic analysis of the welfare consequences of international competition in the presence of internal inefficiencies, requires that the question of why there exists internal inefficiencies at all, is addressed. Most decisions in firms are taken by people who are employed directly, or indirectly, by the owners of the firms. There must hence be some reason why these employees are not induced by their employment contracts to behave efficiently. It seems to us that a natural first approach to such an analysis is to view the firm as a principal--agent relationship where the source of the inefficiency stems from the impossibility for firms to directly observe, and hence contract, employees' supply of effort. Employees dislike effort as such, and therefore the principal -- the owners of the firm -- have to rely on incentive contracts in order to induce employees to supply effort. These contracts are based on some imperfect information the firm receives about the effort


exerted by its employees. However, the extent to which these contracts can exploit the potential gains from trade between firms and employees, is limited by restrictions on the type of contracts that are feasible, such as the necessity to make contracts incentive compatible, and possibly also other restrictions, such as legally imposed minimum wages, etc.

This type of moral hazard problem gives rise to a situation that seems to accord well with popular thinking about the "X-inefficient" firm: First, a small increase in the supply of effort in a firm would, at given contractual wages, have a negligible effect on the utility of the firm's employees, but would have a positive first-order effect on the firm's profit. Second, employees may exert less effort than they would if they worked in a firm that could contract effort — an "informationally efficient" firm — but that otherwise faced identical circumstances. Third, the cost per unit of effort is higher than in the informationally efficient firm that faces identical labor market conditions. Fourth, average costs are higher than in the informationally efficient firm, for any given output volume and labor market situation.³

In the paper we examine a number of views that we believe are commonly held concerning the relationship between international competition, internal inefficiencies in firms, and welfare. To this end, we consider the consequences of international competition for the general equilibrium of an economy in which firms have the above-mentioned type of internal problems. In order to capture varying degrees of competition, we assume that product markets are imperfectly competitive, with free entry and exit, and that economies of scale limit the number of firms. Each firm chooses the number of employees to hire, and the contract to offer each employee, so as to maximize its profit. In order to attract employees each firm has to offer them a utility level which is at least as good as that they

³ In its focus on the individual's supply of effort, the paper also seems to conform to Harvey Leibenstein's notion of X-inefficiency. According to Leibenstein (1976, p.98), "[t]he basic hypothesis upon which the X-efficiency theory rests is that there is always a degree to which effort, in its broadest sense, is a variable".
could obtain in other firms. Hence, the design of the optimal incentive contract is influenced by the general conditions in the labor market. Of course, these conditions are in turn affected by whether or not the country is exposed to international competition, since this affects the elasticity of demand of individual firms, and hence the number of employees they wish to hire. In this fashion, international competition influences the circumstances under which firms and employees agree on the terms of employment, i.e. the reservation utilities of employees. The focus of the analysis is on this interplay between the inefficiencies within firms, and their external environment, as represented by the competition with other firms in factor and product markets.

The paper is to some extent related to a small literature represented by e.g. Hart (1983) and Scharfstein (1988) that analyzes the role of product market competition for the contractual arrangement in a firm in a partial equilibrium setting. Hermelin (1990) studies in a more general context how competitive pressures could spur a firm's managers to work harder. In contrast to these papers, the present paper follows the tradition in trade theory and studies general equilibrium consequences of (international) competition. The reason for this emphasis is, of course, that the opening of trade, as well as international economic integration in the form of the creation of, say, a common market, should be expected to have significant repercussions via factor markets. If one wants to substantiate claims about the relationship between internal inefficiency, trade, and welfare, one thus has to take these effects into consideration. The paper is also related to work by Dixit (1987), who studies the role of trade policy for insurance in the presence of moral hazard. However, as opposed to the analysis by Dixit (1987), the present paper focuses on imperfect competition and increasing returns to scale in product markets.

In our model trade increases welfare, and the mechanism is a familiar one: Trade effectively increases the size of the economy which makes it possible to utilize economies of scale better. The question of interest to the issues considered here is whether the changes in effort supply that results from the exposure to international competition, enhances or
counteracts the exploitations of the increasing returns. As it turns out, depending on the specific properties of the contract problem, effort supply may increase or fall. In the latter case the induced fall in effort supply counteracts the exploitation of the increasing returns to scale. But, when the supply of effort per employee increases as the result of trade, the induced changes in effort supply tend to strengthen the utilization of the increasing returns. However, in contrast to the standard model of imperfect competition and increasing returns, the expansion of output per firm may come about even if firms become smaller in terms of the number of employees.

The moral hazard problem in firms may lead to a too low equilibrioum level effort supply. It is therefore perhaps tempting to think that a trade–induced increase in the supply of effort could have a positive welfare effect *per se*. This is indeed possible, but the mere existence of the moral hazard problem is not a sufficient prerequisite. When the only constraints on firms' choice of incentive contract are that they should be incentive compatible, and yield employees an expected utility level equal to their best alternative, contracts will be chosen such that the marginal benefit of a higher effort is equal to the marginal cost of inducing it through the contract. An induced effort change must then have a zero direct impact on firms' costs. It also has a zero effect on the employees' welfare since the induced change in effort must be a movement along the incentive compatibility constraint. Thus, the welfare effect of an induced change in effort vanishes due to an envelope property: for the employees, marginal disutility of effort equals their marginal compensation, and for the firms, the marginal benefit of effort equals its marginal cost. This property is a standard feature of optimal incentive contracts, but its implications for the consequences of trade are not immediately obvious judging from the type of arguments we referred to initially.

But, it is also possible that a change in effort supply that is brought about by exposure to international competition, does have a first–order effect on welfare. This will be the case if there are additional restrictions on the set of feasible contracts, for instance a
minimum wage legislation, or a bound on the maximum permitted divergence between wages. The firm may then be unable to induce an effort level which equalizes the marginal benefit of effort with the marginal cost of effort. In this case, an induced increase in effort supply as a result of trade will have a positive welfare effect per se. But, since effort supply may increase as well as fall, depending on the exact circumstances under which the contract is agreed upon, this may be a source of welfare loss as well as gain.

A related popular belief that is analyzed in the context of our model, is the notion that trade reduces internal inefficiencies in firms. We show that, to the extent that internal inefficiencies are of the type modeled here, they cannot in general be expected to be reduced by international competition. The source of these inefficiencies is the moral hazard problems in firms, and these problems do not vanish because of trade. Instead, trade alters the circumstances under which these internal inefficiencies affect the economy. It is argued that there are several ways in which the inefficiency could be measured, and that even if a particular measure is adopted, the impact of international competition is generally indeterminate.

The paper also discusses the popular claim that when trade reduces internal inefficiencies, this is a source of a welfare gain. It is argued on more general grounds that this type of reasoning is misleading. A change in the amount of internal inefficiency depends on whether the distance to a fictitious economy without moral hazard problems increases or not, as a result of trade. But what is relevant for welfare changes is the extent and direction in which the equilibrium allocation changes in the economy with moral hazard problems, not how its position relative to a fictitious economy changes. Thus, the relevant question is not if exposure to international competition reduces the inefficiencies, but rather if it makes these inefficiencies matter more or less, i.e. does the cost of these inefficiencies increase or decrease as a result of trade. It hence seems as if statements of the type "gains from trade due to reduced X—inefficiency" do not make sense.

The structure of the paper is the following. In section II the model of the firm is
presented. In section III the optimal incentive contract is analyzed. The consequences of the internal inefficiency in general equilibrium is investigated in section IV. In the ensuing section the consequences of trade are examined, including a discussion of the relationship between trade and the amount of internal inefficiency. Section VI, finally, provides some concluding remarks.

II. THE MODEL OF THE FIRM

Consider a firm that produces an output volume $x$. The firm has $n-1$ identical competitors, each one producing $y$ units of the same homogenous good. The price of the good, $p$, is a decreasing function of the total output produced and increases in consumers' total income $I$: $p = P(x+(n-1)y,I)$. The firm's output depends on the total amount of effort expended in the firm: $x = F(me)$, where $m$ is the number of identical employees of the firm and $e$ is their effort level. There are increasing returns to scale, implying that the marginal product of total effort exceeds its average product. We assume that these economies are smaller the larger the output volume, in the sense that the ratio between the marginal and the average product is closer to unity the larger the output volume. Letting $G = F^{-1}$, the number of employees $m$ required to produce $x$ units of output when each employee's effort level is $e$, is given by $m = G(x)/e$. The assumptions about the production technology are that $\theta(x) \equiv zG_x(x)/G(x) < 1$, and $G_{xx} > 0$. It follows from the latter assumption that $\theta_x > 0$.

An employee's effort is not observable and therefore not contractable. But the firm is able to monitor its employees, and through its monitoring the firm obtains some imperfect information about an employee's effort. We do not explicitly model the monitoring of the firm, but simply assume that the monitoring technology is such that it is optimal for the firm to monitor all its employees. As a result of the monitoring, the firm

---

4 Subscripts on function symbols denote partial derivatives.
observes an outcome of the employee’s effort which is either "good" or "bad", and which is contractable. It is known that the probability $\phi$ of the "good" outcome increases, but possibly at a decreasing rate, with the effort exerted by the employee: $\phi = \phi(e)$, with $\phi_e > 0$ and $\phi_{ee} \leq 0$. Since the employees can affect the probability of the outcome, the firm’s imperfect information about effort supply provides a basis for an incentive contract: The employee is paid the real wage $w$ if his performance is classified as "good" and the real wage $v$ if his performance is classified as "bad". The incentive contract is thus a state–contingent contract which with probability $\phi$ gives the employee the real wage $w$ and with probability $1-\phi$ gives him the real wage $v$.

This somewhat abstract description of the firm–employee relationship can be given several interpretations. For instance, suppose that $e$ is the fraction of the day the employee actually works, that the firm makes a random daily inspection of the employees and that the individual upon inspection is either found to be working or to be shirking. Then the choice of $e$ affects the probability that the employee is caught shirking and the contract stipulates that the wage $w$ is paid if the employee is not caught and that $v$ is paid otherwise. Another example is when the firm can observe the quality of the employee’s output, the quality may be either "good" or "bad" and $\phi(e)$ is the probability of an employee producing a "good" quality when his effort is $e$. Here, the incentive contract requires $w$ to be paid if the quality is "good" and $v$ to be paid otherwise. Yet another interpretation is in terms of the "canonical" principal–agent model of moral hazard in

---

5 We assume that each firm employs sufficiently many employees so that each individual conjectures that his choice of effort has a negligible impact on the observable average output. This assumption ensures that the type of contract analyzed by Holmström (1982) is not feasible. Applied to the present model, such a contract would yield each employee a high renumeration if the average effort meets a certain target, while any other observed average effort results in a penalty. This type of contract will provide an incentive for employees to exert effort if each employee believes that his choice of effort has some observable effect on the average effort: If he expects all other employees to supply the average effort, he has no incentive to deviate from that effort level himself. We hence assume that there is a genuine free rider problem which prevents this type of contract to be implemented.
which the output of the individual employee takes on one of two possible values \( s_H > s_L \), and where the probability for the higher output is \( \phi(e) \). With a large number of workers aggregate output is approximately deterministic, so that the production function \( F \) could e.g. be of the form \( x = m[\phi(e)s_H + (1-\phi(e))s_L] - a \), where \( a \) is a fixed component. The worker then receives the wage \( w \) when \( s_H \) is realized and \( v \) when \( s_L \) obtains.

The expected utility of an employee depends on his consumption and his supply of effort. Let \( U(w) \) be the (indirect) utility of the income \( w \), with \( U_w > 0 \) and \( U_{ww} \leq 0 \), and let \( Z(e) \) denote the individual’s disutility of effort, with \( Z_e > 0 \) and \( Z_{ee} > 0 \). The employee’s expected utility is

\[
S(e,w,v) = \phi(e)U(w) + (1 - \phi(e))U(v) - Z(e) \tag{1}
\]

The employee decides on the amount of effort to provide given the real wages \( w \) and \( v \), and given some constraints on \( e, \underline{e} \leq e \leq \bar{e} \). Assuming an interior solution to this problem, the first-‐order condition for a maximum is

\[
S_e(e,w,v) = \phi_e(e)[U(w) - U(v)] - Z_e(e) = 0, \tag{2}
\]

and the second-‐order condition is

\[
S_{ee}(e,w,v) = \phi_{ee}(e)[U(w) - U(v)] - Z_{ee}(e) < 0, \tag{3}
\]

which is fulfilled if \( w \geq v \) since \(-Z(e)\) is strictly concave. Equation (2) gives the individual's optimal choice of effort as a function of the real wages in the two states: \( e = E(w,v) \), with the properties
\[ E_w = -\frac{\phi_e U_w}{\beta e e} > 0 \quad \text{and} \quad E_v = \frac{\phi_e U_v}{\beta e e} < 0 \] (4)

To attract employees a firm must offer an expected utility of at least \( s \) — the "individual rationality" (IR) constraint for the firm. The reservation utility \( s \) is determined through the interaction of labor demand and supply, but the individual firm takes \( s \) as given. The set of feasible contracts between an employee and a firm is also restricted by the employee's discretion over the unobservable effort level, i.e. by \( e = E(w,v) \), the "incentive compatibility" (IC) constraint.

We will also consider the implications of a third restriction on the set of feasible contracts. It restricts the set of real wages which can be entered in the contract and this set may be different for different values of the reservation utility \( s \). For given values of \( s \) and \( w \) this constraint puts a lower bound on \( v \), the real wage in the bad state, and for this reason we shall refer to it as a "limited punishment" (LP) restriction. One example of such a restriction is a minimum wage legislation which puts a lower bound on the wages in the two states. Another example could be a trade union which is able to enforce a limit on the divergence between wages in the two states. A third possibility is a lower bound on the realized income that the employee accepts, because there is some other non-market activity that the employee would turn to if threatened with an income below, say, \( v_0 \), where \( v_0 \) may depend on \( s \). A contract based on \( w \) or \( v \) less than \( v_0 \) would not be feasible.

This type of limitation to the possible punishments occurs in various guises in the literature. For instance, in the efficiency wage models of unemployment the severest punishment is firing; in the credit contracting literature borrowers have limited liability, as do often managers in models of managerial control problems, etc. We do not necessarily believe that such a restriction is an inherent feature of the type of contract problems studied here, even though for instance a lower income bound does have some intuitive appeal. Rather, what we want to show is the difference it makes whether or not there
exists restrictions on the contract in addition to the IR and IC constraints. The exact specification of this restriction is of less importance, and it is therefore represented by the general form $H(w,v,s) \leq 0$, with $H_s \geq 0$. It is taken to be symmetric with respect to $w$ and $v$ in the sense that if $w > v$ then $H_w \geq 0$ and $H_v \leq 0$, while if $w < v$ then $H_w \leq 0$ and $H_v \geq 0$.

The firm's decision problem is to choose its output, the number of employees to hire and the two wages so that its profit is maximized, taking into account the behavior of its employees, and ensuring that they obtain at least the expected utility $s$. Total cost is

$$m[\phi(e)w + (1 - \phi(e))v] = G(z) \frac{\phi(e)w + (1 - \phi(e))v}{e}$$

(5)

which is separable in output and in variables that affect the utility of employees.

Let $C(e,w,v) \equiv [\phi(e)w + (1 - \phi(e))v]/e$, i.e. $C(e,w,v)$ is the expected (or average) cost per efficiency unit. Note that

$$C_e = \frac{1}{e}[\phi_e(w - v) - C],$$

(6)

i.e. an increase in the effort level has two counter-acting effects on the firm's cost per efficiency unit. First, the cost tends to increase, since the higher wage $w$ will be paid with a higher frequency and the lower wage $v$ with a lower frequency. Second, at an unchanged wage bill, the cost per efficiency unit decreases as their number increases.

The firm selects a contract $(w,v)$ and an induced effort level $e$ such that labor cost is minimized, i.e. its optimal contract problem is

$$G(z)[\min_{e, w, v} C(e,w,v) \text{ s.t. } S(e,w,v) \geq s, e = E(w,v), H(w,v,s) \leq 0]$$

(7)

Note that the solution to this problem is independent of the firm's choice of output. Given
the optimal contract determined by (7) the firm's total cost is a function of its output and the reservation utility. The latter is taken as given by the firm, and its only remaining task is to select its output volume so that profits are maximized.

III. THE OPTIMAL INCENTIVE CONTRACT

The optimal incentive contract is determined by the firm's contract problem (7). The Langrangean for this is

\[ C(e,w,v) - \lambda(S(e,w,v) - s) - \mu(e - E(w,v)) + \gamma H(w,v,s) \]  

where the Lagrange multipliers \( \lambda \) and \( \gamma \) are non-negative. Note to begin with that the individual participation constraint need not necessarily bind. The first order conditions with respect to the wages \( w \) and \( v \) are

\[ C_w = \lambda S_w - \mu E_w - \gamma H_w \]  
\[ C_v = \lambda S_v - \mu E_v - \gamma H_v \]  

If the individual rationality constraint does not bind (so that \( \lambda \) equals zero) it can be seen from (9) and (10) that \( \gamma > 0 \), i.e. that the LP-constraint binds. In this case the firm prefers for productivity reasons to pay employees more than is necessary to induce them to accept employment. It can be shown that this case requires unemployment to prevail in general equilibrium. But in order to keep as close as possible to the tradition of trade theory, we will focus on situations of full employment and then the individual participation constraint always binds.

The contract problem may be rewritten in a more convenient form with a binding IR-constraint. To this end, let \( w = W(e,s) \) and \( v = V(e,s) \) be defined by the IR and the IC restrictions. It is straightforward to establish that both wage functions are increasing in
the reservation utility \( s \) and that \( W(e,s) \) increases, while \( V(e,s) \) decreases, in effort \( e \). Also, let \( \bar{C}(e,s) \equiv C(e,W(e,s),V(e,s)) \). \( \bar{C}(e,s) \) is the firm’s cost per efficiency unit, when the level of effort supply is \( e \), and the contract yields utility \( s \) to the employee. Finally, let \( \bar{H}(e,s) \equiv H(W(e,s),V(e,s),s). \) The assumptions about the LP—restriction \( H(w,v,s) \) and the properties of the wage functions \( W(e,s) \) and \( V(e,s) \) implies that \( \bar{H}_e > 0 \). Thus, the LP constraint implies an upper bound on effort \( e \). Denote this upper bound by \( \bar{e}(s) \).

The contract problem can now be restated as

\[
\begin{align*}
\min_{e} & \quad \bar{C}(e,s) \\
\text{s.t.} & \quad e \leq e \leq \min\{\bar{e}(s),\bar{e}\}
\end{align*}
\]

(11)

To begin with, we assume that \( e \) is small enough to rule out a solution at \( e = \underline{e} \) (it is also assumed that \( \bar{e}(s) > \underline{e} \)). Unfortunately, \( \bar{C}(e,s) \) is generally not convex in \( e \) so multiple solutions cannot be ruled out, and \( e \) may furthermore depend discontinuously on the reservation utility \( s \). These features of the contract problem may arise under reasonable circumstances and cannot be removed in any simple manner by more specific assumptions about the functions involved in the contract problem. Nevertheless we consider these features to be of secondary importance to the issue at hand: the relationship between the moral hazard problem in firms and international competition. To avoid that this relationship is obscured by peculiarities in the contract problem, we shall concentrate on situations where the contract problem is well behaved in the sense that there is a monotone relation between effort and the reservation utility, and furthermore that this relation is continuous if necessary.

The first—order condition of the contract problem is

\[
\bar{C}_e(e,s) \leq 0
\]

(12)
with strict inequality if \( e = \min \{ \hat{e}(s), \bar{e} \} \). Using the definition of \( \bar{C}(e,s) \) the first order condition implies that

\[
\bar{C}_e \equiv C_e + C_w W_e + C_y V_e \leq 0.
\]

It can readily be shown that \( C_w W_e + C_y V_e \) sum to zero if workers are risk neutral, while risk aversion implies that this sum is strictly positive. Thus, if employees have risk aversion or if the LP—constraint is binding, \( C_e < 0 \). In such a situation, a higher effort level would reduce the cost per efficiency unit, but would have a zero first—order effect on the employee’s welfare. The impossibility of directly contracting effort results in an inefficient amount of effort, an inefficiency which we shall discuss in more detail below.

An important aspect of the contract for the subsequent analysis is the relation between effort supply \( e \), and the reservation utility \( s \); denote this relation \( \bar{E}(s) \). When the LP restriction does not bind, the relation between effort and the reservation utility is locally given by the first order condition \( \bar{C}_e(e,s) = 0 \). Then \( \bar{E}_s(s) = - \bar{C}_{es}/\bar{C}_{ee} \). Effort is increasing (decreasing) in \( s \) if \( \bar{C}_{es} \) is negative (positive), since \( \bar{C}_{ee} > 0 \) by the second order condition. The cross derivative \( \bar{C}_{es} \) may have either sign depending on the specific properties of the functions \( \phi, U \) and \( Z \). Unfortunately, it is impossible to characterize in any simple way the set of such functions which yields, say, a positive relation between effort and the reservation utility. Therefore, both possibilities regarding the sign of that relation shall be allowed for in the remainder of this paper.

When the LP—constraint binds the solution to the contract problem is simply given by the LP—constraint \( e = \hat{e}(s) \). In this case the relation between effort and the reservation utility is \( \bar{E}_s(s) = \hat{e}_s(s) = - \bar{H}_s/\bar{H}_e \). Effort is increasing (decreasing) in \( s \) if \( \bar{H}_s \) is negative (positive). However, the sign of \( \bar{H}_s \) cannot be determined without more specific assumptions about the LP—constraint.

Finally, using \( \bar{C}(e,s) \) and the Lagrangean (8) we have that

\[
\frac{dC}{ds} = \bar{C}_s + \bar{C}_e \bar{E}_s = \lambda + \gamma \bar{H}_s > 0
\]

(13)
i.e. a higher reservation utility must increase the cost per efficiency unit. This is obvious when \( \bar{C}_e = 0 \), since \( \bar{C}_s > 0 \). But minimization implies that even if \( \bar{C}_e < 0 \) (i.e. the LP constraint binds) and \( \bar{E}_s > 0 \), the cost per efficiency unit must increase in \( s \).

**The informationally unconstrained firm**

In the subsequent analysis we will occasionally use the bench–mark case of an economy in which firms can contract effort directly, i.e. in which firms do not face an IC–restriction. These firms will be referred to as "informationally unconstrained". Such a firm does not need an incentive contract to elicit effort. Disregarding for the moment the LP constraint, it is always optimal for this firm to set \( w = v \), since otherwise the firm must compensate the employee's risk by paying a higher average wage. The assumptions about the function \( H(w,v,s) \) ensures that an informationally unconstrained firm pays a deterministic wage \( w = v \) also if the LP constraint is binding.

The informationally unconstrained firm solves:

\[
\begin{align*}
\min_{e,w} & \quad C(e,w) = \frac{w}{e} \\
\text{s.t.} & \quad S(e,w) = U(w) - Z(e) \geq s \\
& \quad H(w,v,s) \leq 0
\end{align*}
\]

(14)

In this case the IR constraint always binds regardless of whether the LP restriction binds or not. If the IR constraint did not bind it would be possible to increase \( e \) somewhat at a constant \( w \). This would not affect the LP constraint, but it would decrease the firm's cost. So we can use the IR constraint to obtain \( W^0(e,s) \); the wage required to compensate for effort \( e \) when the reservation utility is \( s \). Define \( C^0(e,s) \equiv C(e,W^0(e,s)) \). The LP constraint is \( H(W^0(e,s),W^0(e,s),s) \leq 0 \). Assuming that \( H_w + H_v < 0 \) when \( w = v \) (i.e. the LP constraint imposes a lower bound on a certain wage) and noting that \( W^0 \) is increasing
in $e$, it follows that the LP constraint is equivalent to a lower bound on $e$. Denoting this bound by $\bar{e}(s)$, the informationally unconstrained firm’s contract problem can be expressed as

$$\min C^0(e,s) \text{ s.t. } \max_{\bar{e}} \{ \bar{e}^0(s), \bar{e} \} \leq e \leq \bar{e} \quad (15)$$

Again, the cost function is not globally convex. But it can be shown to be strictly convex at any stationary point. Thus, the solution is unique and the optimal effort depends continuously on $s$.

Denote the solution to the minimization problem by $e^0 = E^0(s)$. The efficient effort supply could in general either be increasing or decreasing in the reservation utility. At an interior solution, the relation between $e$ and $s$ is determined by the first order condition $C^0_e(e,s) = 0$. Thus, $E^0_s = -C^0_{es}/C^0_{ee}$, so there is a positive relation between $e^0$ and $s$ if and only if $C^0_{es} < 0$. It can be shown that this is this case if and only if $-wU_{uw}/U_w < 1$, i.e. as long as the utility function is not too concave. If on the other hand the LP constraint binds, the effect on $e^0$ of a higher reservation utility depends on how the latter affects the lower bound $\bar{e}^0$, i.e. on the specific assumptions about the properties of the LP restriction.

Finally, note that (14) and the definition of $C^0(e,s)$ together imply that

$$\frac{dC^0}{ds} = C^0_s + C^0_e E^0_s = \lambda + \gamma H_s > 0 \quad (16)$$

i.e. an increase in the reservation utility must increase the cost per efficiency unit in the informationally unconstrained firm.

Comparison of informationally constrained and informationally unconstrained firms

We conclude this section with a comparison of the optimal contract in
informationally constrained and unconstrained firms. Note, first, that for given \( e \) and \( s \)

\[
U(W^0(e,s)) = \phi(e)U(W(e,s)) + (1-\phi(e))U(V(e,s)) = s + Z(e) \tag{17}
\]

Since \( W^0(e,s) \) is a certain income which gives the same (expected) utility as the risky contract \{\( W(e,s), V(e,s) \)\} it holds that

\[
0 \leq \bar{W}(e,s) - W^0(e,s) \equiv r(e,s) \tag{18}
\]

where \( \bar{W}(e,s) \equiv \phi(e)W(e,s) + (1-\phi(e))V(e,s) \) is the expected wage, and \( r(e,s) \) is the risk premium associated with the risky contract, which is positive when workers are risk averse.

Using the definition of the risk premium, the relation between the cost functions for informationally constrained and unconstrained firms is

\[
\bar{C}(e,s) = C^0(e,s) + \frac{r(e,s)}{e} \tag{19}
\]

and consequently \( \bar{C}(e,s) \geq C^0(e,s) \). Furthermore, we shall assume that the solution for the informationally constrained firm always yields \( \bar{E}(s) \geq \tilde{e}^0(s) \). Then minimization implies that \( C^0(e^0,s) \leq C^0(\tilde{e},s) \), and hence

\[
C^0(e^0,s) \leq \bar{C}(\tilde{e},s) \tag{20}
\]

i.e. the cost per efficiency unit is lower in the informationally unconstrained firm.

The risk premium per efficiency unit is the cost of implementing a certain effort \( e \) through an incentive contract rather than through a direct contract on effort. If this cost always were increasing in \( e \) then employees in an informationally constrained firm would always exert less effort than their counterparts in an informationally unconstrained firm:
From (19) it can be seen that if the risk premium per efficiency unit is increasing in effort at \( e = \bar{E}(s) \), then \( C^e \) must be decreasing in \( e \) at that effort level, so \( E^0(s) \) must be higher. Unfortunately, this cost can also — for perfectly reasonable utility functions — be decreasing in effort. Under such circumstances the informationally constrained firm will actually induce more effort than a firm without moral hazard problems. But, in the subsequent analysis we concentrate on situations where incentive contracts result in a too low effort supply, i.e. where \( E^0(s) \geq \bar{E}(s) \). The implicit assumption is thus that the cost per efficiency unit of inducing effort through an incentive contract rather than through a direct contract, is higher the higher the effort level is.

Finally, inequality (20) demonstrates the inefficiency in the contractual relationship between the firm and its employees when the firm is unable to directly contract effort. The cost per efficiency unit is higher for an informationally constrained firm than for its informationally unconstrained counterpart for two reasons. First, with risk averse employees, the optimal contract involves the usual trade-off between the provision of insurance and the provision of effort incentives. To induce the same amount of effort as a firm which directly can contract effort, the informationally constrained firm must expose the employees to risk and compensate them for this by an amount equal to the risk premium, as shown in (19). As a result its cost per efficiency unit is higher and it is in general beneficial for the firm to adjust the amount of effort induced to counteract some of the higher cost. Given our assumption above that the risk premium per efficiency unit is lower the lower effort is, the optimal adjustment is a reduction of induced effort. Second, a binding LP—restriction exacerbates this inefficiency by imposing an upper bound on the feasible effort supply, which forces the firm to induce a lower effort than it otherwise would have done. In the remaining sections of the paper the general equilibrium ramifications of this inefficiency is explored, in particular its implications for the way in which the economy is affected by international competition.
IV. THE MODEL OF THE ECONOMY

The economy we have in mind consists of a large and fixed number of identical industries. In each industry there are many firms, each facing the incentive problems analyzed above. There is free entry and exit into each industry, and in equilibrium all profits are zero.

For notational convenience the number of industries is set equal to unity. As a consequence, the consumer price index is equal to the price of this good. Each firm will therefore affect the consumer price index, and hence the real wage of its employees, to the same extent that it affects the price of its output. However, since we are interested in situations where each firm is small in relation to the economy as a whole and thus has a negligible impact on the consumer price index, each firm is taken to view the consumer price index as unaffected by its choice of output, while being aware of its influence on the industry's output price.

There is only one factor of production - labor. The supply of labor depends in a simple manner on the utility offered by employment: supply is nil for a utility level below some exogenous value $\bar{s}$ and is equal to $L$ for all higher levels. Equilibrium in the labor market can then be of three different types. First, there can be full employment: $nm = L$. In this case firms face a binding IR constraint $S(E(\frac{wU}{p'}, \frac{uU}{p}, \frac{wU}{p'})) = s > \bar{s}$ in their cost minimization problems. A second possibility is that firms face a binding IR constraint, but there is underemployment, i.e., $S(E(\frac{wU}{p'}, \frac{uU}{p}, \frac{wU}{p'})) = s = \bar{s}$ and $nm < L$. In the third type of equilibrium there is unemployment and a higher expected utility for the employed than the unemployed: $S(E(\frac{wU}{p'}, \frac{uU}{p}, \frac{wU}{p'})) > s = \bar{s}$ and $nm < L$. In this equilibrium firms give workers a higher utility than their reservation utilities for efficiency wage reasons. We will henceforth concentrate on the standard case in trade theory, that of full employment.

Conditions for an autarky (Nash) equilibrium are the following. First, each firm's
choice of output should be a best response against the choices made by other firms.\(^6\) In a symmetric equilibrium the marginal revenue is \(p(1-1/n)\). Normalizing the numeraire \(p\) to unity, the first-order condition for profit maximization is

\[
1 - \frac{1}{n} = G_x(x)\tilde{C}(e,s) \tag{24}
\]

Second, industry equilibrium requires zero profits, which in the the symmetric case is equivalent to the condition

\[
x = G(x)\tilde{C}(e,s) \tag{25}
\]

Third, the induced supply of effort is

\[
e = \tilde{E}(s) \tag{26}
\]

Finally, the labor market equilibrium condition is, taking into account the technological relationship \(G(x) = me\),

\[
nG(x) = eL \tag{27}
\]

Note that since there is full employment, no profits, and \(S(e,w,v) = s\) in equilibrium, welfare is simply given by \(s\).

In section V we employ this simple general equilibrium structure to analyze the relationship between internal inefficiencies of firms and the effects of international competition. The analysis is organized around a number of headlines which we believe

\(^6\)The qualitative results of this paper could also be derived with differentiated products and a Bertrand–Nash equilibrium concept.
reflect "conventional wisdoms" on internal inefficiencies and international competition. The aim is to discuss to what extent they are true in the model under study.

But first, we shall here examine a "conventional wisdom" which is not directly related to questions of international competition, but which serves to highlight how the inefficiency within firms interacts with that due to monopoly pricing and increasing returns to scale.

"Employees exert less effort in economies where firms have internal control problems"

Unobservability of effort leads (by assumption) to a lower level of effort supply for a given level of reservation utility of employees. But it is still an open question whether or not in equilibrium employees exert less effort in an economy with internally inefficient firms compared to an economy with internally efficient ones. Thus, the question to study is whether \( \bar{E}(s) \not\geq E^0(s^0) \).

To this end we introduce a couple of definitions. Let the functions \( x = \bar{\chi}(s) \) and \( x = \chi^0(s) \) be defined by the zero profit condition (25) and the corresponding equation in the case of the efficient economy, respectively. Since \( C \) is increasing in \( s \), the zero profit condition shows that \( \bar{\chi}_s > 0 \) and \( \chi^0_s > 0 \). Note also that the inequality \( \bar{C}(\bar{E}(s),s) > C^0(E^0(s),s) \) in the same way gives \( \bar{\chi}(s) > \chi^0(s) \). Second, let

\[
\Gamma(x) \equiv \frac{G(x)}{1 - \theta(x)}, \quad \text{with} \quad \Gamma'_x = \frac{G_x}{1 - \theta} + \frac{G\theta}{(1 - \theta)^2} > 0
\]

Using (24)–(26) the autarky equilibrium conditions for the two economies can be expressed as:

\[
\Gamma(\bar{\chi}(s)) - \bar{E}(s)L = 0 \tag{28}
\]
\( \Gamma(\chi^0(s)) - E^0(s)L = 0 \)  \hfill (29)

We now impose a "stability" condition on the analysis: Suppose that the number of firms is exogenous, and that the zero profit constraint therefore is irrelevant. An exogenous increase in the number of firms, taking factor market repercussions into account, should then reduce profits. This condition is fulfilled if, for example, effort supply is exogenously given, as in the standard model, or if effort supply falls in the reservation utility of employees: \( \bar{E}_s < 0 \) and \( E^0_s < 0 \). The stability condition ensures that the left hand sides of (28) and (29) are increasing functions of \( s \): \( \Gamma \bar{\chi}_s - \bar{E}_s L > 0 \) and \( \Gamma^0 \chi^0_s - E^0_L > 0 \) (see the appendix for details).

Return now to the question of whether employees work less hard in the economy where firms have problems of internal control. The expressions (28) and (29) implicitly define the autarky equilibrium welfare levels, denoted \( \bar{s}_a \) and \( s^0_a \), respectively. Since \( \bar{\chi} > \chi^0 \) and \( \bar{E} < E^0 \), the autarky welfare level in the economy with internally inefficient firms must be lower than in the one with internally efficient firms.

In order to characterize differences in equilibrium effort levels, one has to distinguish between situations where contracts are such that higher reservation utility leads to higher effort supply and situations where the opposite holds. Consider first the situation in which \( \bar{E}(s) \) increases in \( s \). It is clear that in this case employees exert less effort in the informationally constrained economy: \( \bar{E}(\bar{s}) < \bar{E}(s^0) < E^0(s^0) \). By (28) and (29) the informationally constrained firms have smaller output volumes, thus utilizing economies of scale less. From the profit maximization and zero profit constraints it also follows that their number is smaller. Hence, the price-marginal cost ratio is higher, so there is more of monopoly pricing in the economy with informationally constrained firms.

On the other hand, when effort decreases with the reservation utility level, there are two possibilities. One is that the comparison in the previous case still holds qualitatively. But, it is also possible that employees in equilibrium exert more effort in the economy
with inefficient firms! This is due to the fact that there are two reasons why equilibrium effort levels in the two economies differ. First, effort is (by assumption) lower for any given reservation utility in the inefficient firms; \( \bar{E}(s^0) < E^0(s^0) \) for any \( s^0 \). Second, as a consequence of this inefficiency, the (reservation) utility level is lower in the economy with control problems. This is a general equilibrium ramification of the internal inefficiency that will enforce or dampen the first effect: if \( \bar{E}_s < 0 \), this general equilibrium effect works in the opposite direction, and may even dominate the first effect. In this case the internally inefficient firms are larger in terms of output per firm, i.e. these firms utilize economies of scale better than their internally efficient counterparts, but they are smaller in terms of the number of employees. The number of firms is then larger in the economy with inefficient firms, so there is less of monopoly price mark-up in this economy.

In sum:

*Proposition 1: The economy with informationally constrained firms has lower welfare than the economy with informationally unconstrained firms. It is not possible to unambiguously rank employees' effort supply in the two types of economies.*

V. INTERNAL INEFFICIENCIES AND INTERNATIONAL COMPETITION

In order to capture the consequences of international competition we will consider a world economy consisting of a "Home" economy of the above type, and a "Foreign" country whose firms may or may not be internally inefficient. It is assumed that both countries have the same production technology and the same preferences, i.e. they have functions \( G, U, \) and \( Z \) in common. The international product market is taken to be fully integrated; there is thus a uniform international product price in equilibrium, which is utilized as a numeraire and set to unity. Obviously, since there is only one good, trade need not actually take place in equilibrium. Nevertheless, the mere possibility of trade will have real consequences, since it changes the competitive conditions in the product market.
The revenue of an individual producer \( i \) in, say, the Home country is, in a symmetric equilibrium,

\[
\frac{x^h_i}{x^h_i + (n^h-1)x^h + n^fx^f} (t^h + I^h) \tag{30}
\]

where \( h \) and \( f \) denote "Home" and "Foreign", respectively. Profit maximization requires, taking into account the fact that income equals expenditure in each economy

\[
1 - \frac{x^k}{x^w} = C^k(E^k(s^k), s^k) G(x^k); \quad k = h, f \tag{31}
\]

where \( x^w \equiv n^h x^h + n^f x^f \) is world output, \( C^k \equiv \bar{C} \) and \( E^k \equiv \bar{E} \) if country \( k \) has internally inefficient firms, and \( C^k \equiv \bar{C}^0 \) and \( E^k \equiv \bar{E}^0 \) otherwise. The zero profit constraints are

\[
x^k = C^k(E^k(s^k), s^k) G(x^k); \quad k = h, f \tag{32}
\]

Finally, the conditions for full employment of factors are

\[
n^k G(x^k) = E^k(s^k) L^k; \quad k = h, f \tag{33}
\]

Regardless of whether internal inefficiencies are present in one or both of the two countries, it follows from (31) and (32) that

\[
\frac{x^k}{1 - \theta(x^k)} = x^w; \quad k = h, f \tag{34}
\]

and hence that \( x^h = x^f \). That is, firms in both economies produce the same output volume.
regardless of whether or not they share the problems of internal control.\textsuperscript{7}

Having introduced international competition into the general equilibrium model, we are now ready to address a number of popular views regarding the relations between inefficiencies in firms and international competition.

"International competition yields welfare gains by inducing employees to exert more effort"

Much of the policy discussion of "X-inefficiency" seems to rely on the idea that international competition changes incentives in firms in such a way as to induce employees to increase their supply of effort. Furthermore, it is taken for granted that this increase in the supply of effort is a source of welfare gain. In order to determine the validity of this statement in the context of the present model, we need to establish both the overall effect of international competition for welfare, and to decompose this overall effect into its different sources. For the time being it is immaterial whether or not the Foreign country firms are internally inefficient; Foreign country functions and variables are therefore denoted by superscript $f$.

Taking (31)-(34) into account, we can express those equilibrium conditions that pertain to the Home economy as

$$
\Gamma(\bar{x}(s^h)) - \bar{E}(s^h)L^h = E^f(s^f)L^f
$$

Denote the Home country's equilibrium welfare level under international competition by $s^h_c$. Comparing (35) to (28), it is clear that the only difference is the right hand side of (35), which is now positive instead of zero. It follows immediately from the stability

\textsuperscript{7} This property is essentially due to the combination of a (trivially) homothetic cost function $G(\cdot)\bar{C}(\cdot)$ and the symmetry of the demand, i.e. that firms of both nationalities face identical demand conditions. For instance, it would not any longer hold if markets were segmented, so that firms could price discriminate between national markets, and the countries differed in some respect, such as size.
condition $\bar{r}_x \tilde{x}_s - \bar{E}_s L > 0$ that $s^h_a < s^h_c$. As a consequence, since $\tilde{x}_s > 0$, we have that $x^h_a < x^h_c$.

Intuitively, the only mechanism through which trade directly affects the economy is by increasing the elasticity of product demand for a single firm. This increase in the degree of product market competition will cet. par. induce firms to increase their output volumes. Now suppose for a moment that effort levels are held constant at their autarky levels. The expansion in each firm’s output can then only be brought about by additional hirings. But this is not possible in the aggregate, since there is full employment. In order for the factor market to get back into equilibrium, the number of firms must fall. As a result of the better exploitation of the increasing returns, total consumption and hence welfare, rises.

This is the "story" in the traditional trade model. But here the increased competition in the product market will affect the conditions under which the contract between the firm and its employees is formulated. The latter will now demand a higher utility level to accept employment. Depending on whether $\tilde{E}(s)$ increases or falls as a result of the increase in reservation utility, i.e. depending on properties of the contract problem, this will result in more or less effort. If the supply of effort is reduced, i.e. if $\tilde{E}_s < 0$, the need for additional hirings will increase further, and the equilibrating mechanism in the labor market has to be a reduced labor demand through the exit of firms. But with a sufficiently pronounced positive effect on effort supply, the number of firms per country may actually increase, thus enhancing the pro-competitive effect of trade. The total number of firms in the world economy could then rise as a result of trade.

In order to identify where the welfare gains from international competition stem from, we rewrite the equilibrium conditions as

$$\Gamma(x) - \bar{E}(s^h) L = \ell$$

(36)
\[ \tilde{C}(E(s^h), s^h) = \frac{x}{G(x)} \]  

(37)

where \( \ell = 0 \) in case of autarky and \( \ell = E_f(s^f)L_f \) in case of international competition. Solving for \( x \) and \( s \) in these equations for all values of \( \ell \) in the interval \([0, E_f(s^f)L_f]\) gives us a curve in \((x, s, \ell)\) space that connects the autarky and the international competition allocations. Using the differentials of (36) and (37) we can express the difference in output volumes per firm and the difference in welfare level at the two allocations as line integrals along this curve:

\[ \tilde{z}_c - \tilde{z}_a = \int_0^{\ell_c} \frac{1}{x(z)} \, d\ell + \int_{\tilde{s}_h}^{\tilde{s}_c} \frac{E_s(s^h)}{x(z)} \, ds^h \]  

(38)

\[ \tilde{s}_c - \tilde{s}_a = \int_{\tilde{x}_c}^{\tilde{x}_a} \frac{1 - \theta(z)}{\tilde{C}(E(s^h), s^h) + \tilde{C}(E(s^h), E(s^h))} \, dx \]  

(39)

As we have just seen, international competition need not induce employees to exert more effort, but will do so under certain circumstances. Hence, the popular argument about the beneficial effects of international competition in the form of increased supply of effort is not correct in the present model. But whether changes in effort supply is a source of welfare gain is still an issue.

The first term of (38) gives the standard effect of international competition on output per firm. The second term captures the additional effect — which could be of either sign — due to an induced change in effort level of the employees. Since more output per firm leads to higher welfare through better exploitation of returns to scale, we see that a
change in effort supply does have an effect on welfare. It should be noted, however, that this welfare effect of international competition is not the outcome of any internal or contractual inefficiency in firms, but would exist also if firms could contract effort directly.

Equation (39) describes the effect of better exploitation of returns to scale on welfare: it is clear that if returns to scale were already exhausted — i.e., if $\theta(x)=1$ — then there would not be any welfare gain from competition, whatever happened to the effort level. But we also see from this equation that an induced change in effort supply may have an impact on the transfer mechanism from exploitation of returns to scale to welfare, depending on whether $\tilde{C}_e$ is negative or equal to zero: The more negative $\tilde{C}_e \tilde{E}_s$ is, the larger is the welfare gain from a given increase in output per firm (note that the denominator in (39) is positive according to equation (7)).

When firms are unconstrained by the LP restriction, a small change in the supply of effort which is induced by a change of equilibrium wages $w$ and $v$, and which does not have a first order effect on the welfare of employees, leaves costs per efficiency unit unchanged:

$$\tilde{C}_e(e,s) = 0.$$  Firms are choosing the best possible effort level, constrained only by the inability to contract effort directly. International competition does not remove the source of this inefficiency, but only changes the circumstances under which it affects the economy.

However, when the LP constraint binds, $\tilde{C}_e(e,s) < 0$. In this situation the utility of an employee can be increased, at a constant cost per unit of effort, if he is induced to supply more effort along the IC restriction; the extent to which this is possible is measured by $-\tilde{C}_e(e,s)/\tilde{C}_s(e,s)$. In this case there is an additional degree of inefficiency in the contract, since there exists another contract under which the firm could produce its output volume with fewer employees at a lower cost. The reason for this inefficiency may be, for instance, that the preferred level of effort may require a lower wage in the "bad" outcome state than what is legally permitted, or it may require a larger difference between the low and the high income than a labor union permits. In such a case a change in effort supply induced by international competition will give rise to a welfare effect that stems from the
internal inefficiency in firms. Note, however, that since the direction in which effort changes is indeterminate, we cannot conclude that its effect on welfare is positive when it exists.

A binding LP restriction does not introduce a separate channel through which welfare is affected because of an inherent feature of this particular restriction. Rather, it is because the three restrictions IR, IC, and LP together imply an upper bound on the effort level. As long as the optimal incentive contract does not exceed this bound, the effort level is chosen such that the cost of a marginal unit of effort equals the benefit to the firm thereof (given that the contract fulfills IR and IC). But, if the optimal incentive contract is constrained by the upper bound on effort, there will be a welfare effect if effort is changed. In order to determine the direction in which international competition affects the bound on effort, one has to impose further restrictions on the LP constraint than is done here. For instance, in the special case where there is a minimum wage restriction \( v > v_0 \), the bound on effort is loosened as a result of international competition. In this case the induced increase in effort contributes positively to the welfare gains from international competition.

To what extent is this potential welfare effect the result of an internal inefficiency? It depends on the particular form of the LP restriction. It may bind for both an internally efficient and an inefficient firm. If so, the effect have nothing to do with the moral hazard problem in the firm, even if it leads to a contract that is inefficient in the sense defined above. On the other hand, if it is of a form that only inhibits firms that are informationally constrained, such as a limit on the extent to which wages are permitted to differ in the two states, then we would define this welfare effect of a change in effort as stemming from an internal inefficiency.

*Proposition 2: International competition leads to higher national welfare. It induce employees to exert more (less) effort if incentive contracts yield higher (lower) effort supply*
at higher reservation utilities. This gives rise to a welfare gain (loss) since it affects the degree to which returns to scale are exploited. There may be an additional effect on welfare that can be related to the control problem in firms, if and only if the LP-constraint binds.

"International competition forces firms to reduce X-inefficiencies"

Let us now turn to the question of whether international competition, by increasing competition in product markets, fosters improved internal efficiency, as is often claimed. In this model, the opening of trade does indeed result in an equilibrium with a higher degree of product market competition, in the sense that the discrepancy between marginal revenue and price is smaller. Does this result in a reduction of internal, or "X-", inefficiency? In order to judge whether the inefficiency increases or decreases as a result of international competition, we need to define a measure of the "amount" of inefficiency that exists in a given situation. This is not an entirely trivial task, partly because the concept is commonly used in a rather loose manner.

However, almost all discussion of "X-inefficiency" is in terms of an individual firm, and involves an examination of the extent to which this firm does not manage its business as profitably as it could under some other circumstances. A measure of this type must thus involve an explicit comparison of the actual allocation in the firm with a counter-factual allocation, one that would be chosen in a situation where the source of the inefficiency is absent, but all other circumstances are the same. It is probably also desirable that the measure does not confuse internal inefficiency with inefficiencies due to unexploited economies of scale. These requirements indicates to some extent how X-inefficiency should be measured. But, there are nevertheless several more or less "reasonable" measures that could be (and are) used. We will now discuss how some of them could be defined in this model.

The least satisfactory measure, but one that seems to be in the minds of some debaters, is the difference between the "maximum" effort workers in a firm could supply
and what they actually supply. It is not clear in general what one should mean by "maximum" effort, but it may be well-defined in certain interpretations of the model. For instance, if $e$ is the fraction of the day the worker does not shirk, the maximum $e$ is unity.

A more interesting approach involves a comparison of the allocation within the firm when the problems of internal control are present with what the firm would choose could it directly contract effort, and all other circumstances are identical. A step in this direction would be to consider the difference between the effort level that the efficient firm would choose and the level that the inefficient firm chooses, when they both face the same reservation utility of workers: $E^0(s) - \tilde{E}(s)$. A third possibility is to look at the firm's cost per efficiency unit, given the reservation utilities of workers: $\tilde{C}(\tilde{E}(s),s) - C^0(E^0(s),s)$. A fourth measure answers the question "how much cheaper could the firm produce its chosen output volume, did it not have problems of internal control, but facing the same reservation utility constraint for its employees?". Now the measure of X-inefficiency would be $G(x)[\tilde{C}(\tilde{E}(s),s) - C^0(E^0(s),s)]$ and the effect from trade would show up both through a change in $G(x)$ and a change in $\tilde{C}(\tilde{E}(s),s) - C^0(E^0(s),s)$. A fifth measure could be the difference in average costs. The empirical literature contains various measures of this type. They have the drawback of also capturing scale economy effects, but may also be problematic in other respects. If the measure is obtained by comparing average costs in equilibrium, i.e. by calculating $G(\bar{z})\tilde{C}/\bar{z} - G(x^0)C^0/x^0$, we see that, by the zero profit constraint, trade has no effect at all on X-inefficiency: it is always zero. Yet another measure would involve a comparison of labor productivity, i.e. output per employee: $\tilde{x}/\tilde{m} - x^0/m^0$. This would be subject to the same type of scale problem as the previous measure.

At this stage it shouldn't come as a surprise that international competition could, depending on the exact circumstances, increase or reduce the degree of X-inefficiency, with each of the above mentioned measures. One reason for this indeterminacy is the fact that the response to international competition of the economy with inefficient firms is quite involved, because of the complexity of the incentive contracts. Another reason is the fact
that international competition not only affects the economy under study, but also the benchmark economy with internally efficient firms.

The point of this catalogue of measures is that there seems to exist several "reasonable" candidates for a measure, each with advantages and draw-backs, and each with its own response to international competition. In order to choose a measure one thus has to make clear exactly for what purpose the measure is employed. However, we will argue below that if the purpose is to make inferences about welfare, neither of these measures are relevant.

"Reduced X-inefficiency is a source of welfare gain from international competition"

We have already examined the welfare consequences of international competition in our model. The purpose here is to argue on more general grounds that the notion that a reduced amount of inefficiency is directly related to welfare, is problematic. The reason is the following. A measure of the welfare loss of an inefficiency must involve a comparison between an inefficient allocation and an allocation that is in some sense efficient. The measure hence indicates the gain that would be obtained if the source of the inefficiency were removed. However, international competition cannot in general be expected to remove the source of the inefficiency, only to change the circumstances under which it affects the economy. That is, it results in a move from one inefficient equilibrium to another. The welfare implication of such a move depends on how the equilibrium allocation changes in the inefficient economy; it is in general irrelevant whether some measure of the distance to a fictitious economy is increased or reduced. The latter may well increase when the equilibrium changes, even though the change in the equilibrium allocation is welfare improving.

An example may help to clarify this point. Consider the standard partial equilibrium model of a monopoly with linear demand and constant marginal cost. In this industry there exists an inefficiency due to monopoly pricing, and the degree of inefficiency
can be measured by the dead-weight loss. Now assume that the firm suddenly gets access to another technology with a lower marginal cost. Does this constitute a welfare improvement? If the dead-weight loss of the monopoly is used as a criterion, the answer is no: with linear demand and constant marginal cost, a lower marginal cost implies a higher excess burden of monopoly. But welfare should not be measured by what is foregone, but by what is obtained, and the lower marginal cost implies that the sum of consumer and producer surplus is increased. The increase in the dead-weight loss only says that with a lower marginal cost it becomes costlier to society to have a monopoly.

The same type of reasoning applies to measures of "X"—inefficiency, since they are measures of what is foregone because of the source of the inefficiency. We would therefore argue that measures of "X"—inefficiency should not in general be expected to be of relevance as indicators of welfare effects of international competition.

"An economy with internally inefficient firms stands more to gain from trading with an economy whose firms lack these problems than with an economy of its own type"

Arguments about the beneficial effects of the "cold shower" of international competition sometimes seem to be based on the idea that the efficiency of foreign firms "spills over" to otherwise inefficient domestic firms. However, it is important to specify the exact mechanism through which internal conditions in one firm affect those in other firms. In our model Home firms are not directly affected by the internal efficiency of Foreign firms, but only through their impact on the price elasticity faced by individual Home firms. As was shown above, it could happen that an economy where firms face internal control problems, in autarky has a larger number of firms than it would have absent these problems. This could be the case when the optimal incentive contract induced lower effort levels, the higher the reservation utility of employees. Hence, a Foreign economy with inefficient firms could yield a stronger competitive pressure than one with efficient firms. In the present model it is therefore not true in general that the Home
country is better off when Foreign firms are internally efficient relative to when they are inefficient.

To see this point analytically, note that since $x^h = x^f$, equation (32) shows that $C^f(s^f) = C^h(s^h)$, i.e., $s^f = (C^f)^{-1}(C^h(s^h))$. Inserting this into equation (35) gives

$$\Gamma (x^h(s^h)) - E^h(s^h)L^h - E^f((C^f)^{-1}(C^h(s^h)))L^f = 0$$

(37)

A stability assumption similar to the one used for the model in autarky ensures that the LHS is increasing in $s^h$. Obviously $C^{-1}(c) < (C^0)^{-1}(c)$, and we have assumed that $E(s) < E^0(s)$; hence the equation above shows that the domestic welfare level $s^h$ is higher if the Foreign country is efficient ($C^f = C^0$, $E^f = E^0$) rather than inefficient ($C^f = \bar{C}$, $E^f = \bar{E}$) as long as $E^f(s)$ is increasing. However, if $E^f(s)$ is decreasing, the converse might be true.

**Proposition 3:** The Home economy gains more from trade when the Foreign economy has efficient firms than when its firms are inefficient, if $E^f(s) > 0$. If on the other hand $E^f(s) < 0$, it is possible that trade with an inefficient economy is more advantageous.

"Gains from international competition are larger for the economy with internally inefficient firms than for the economy with internally efficient firms"

The final notion that will be considered is the argument that international competition is more important to an economy with internally inefficient firms than to one whose firms are efficient. This claim is indeed persuasive, but it gives only half the story. On the one hand, it is true that when firms are internally inefficient, international competition may give rise to particular welfare effects. Furthermore, the economy with internally inefficient firms may exploit returns to scale less than one with internally efficient firms, and may therefore stand more to gain from a given increase in output per firm. However, even if this economy has potentially more to gain from international
competition, it may for the same reason have less ability to take advantage of the opportunities trade offers. The above--mentioned claim can in terms of our model formally be expressed as $\tilde{s}_c - \tilde{s}_a > s_0^c - s_0^a$, which does not seem to be the case in general. Hence, it appears as if an economy with internally inefficient firms in this sense may gain more or less from trade than one with efficient firms. But, it can be shown that one case where the intuitive argument does hold, but probably for a less intuitive reason, is when effort supplies are at maximum before the opening of trade, and so remains with international competition.

VI. CONCLUDING COMMENTS

This paper is inspired by the discrepancy between, on the one hand, the emphasis of the policy debate on reduced inefficiencies within firms as one important source of gains from international economic integration, and, on the other hand, the lack of systematic analysis of this phenomenon. In the paper a number of popular claims are investigated. Despite the fact that the structure of the model economy is very simple, and the internal inefficiencies arise from a straight--forward moral hazard problem, these popular views are in general not correct within the model. One reason for this is no doubt to be found in our simplistic description of the internal inefficiencies. However, it also appears to us as if the relationship between firms' internal organization and their external environment, is more involved than the arguments of the policy debate seem to suggest.

Finally, to avoid misunderstandings, we want to emphasize that we certainly believe that inefficiencies in firms' internal organizations are of considerable practical importance, and that international competition possibly could yield significant gains which somehow are related to these inefficiencies. The purpose of the paper is not to deny their existence, but to show that the mere existence of internal inefficiencies in firms are not sufficient to produce the sort of welfare gains from trade which are recurrently referred to in the policy debate. To develop a better understanding of the roots and implications of these
phenomena we must embark on a systematic analysis of how the internal organization of firms is determined and how this determination interacts with conditions in firms' product and factor markets.
APPENDIX

Here we will establish that the stability condition introduced in section IV implies that
\[ \Gamma x \chi_s - E_s L > 0. \]

The general equilibrium equations (24)-(27) can be written

\[
\begin{align*}
1 - \frac{1}{n} &= G_x(z)C(s) \\
x - G(z)C(s) &= \pi \\
nG(z) &= E(s)L
\end{align*}
\]

(A1) \hspace{2cm} (A2) \hspace{2cm} (A3)

where \( \pi = 0 \) and the three equations determine \( z, s \) and \( n \). Now assume instead that \( n \) is exogenous, and that the three equations determine \( z, s \) and \( \pi \), where \( \pi \) is firms' profits.

Let (A2) define \( \chi(s, \pi) = z \), so that

\[
\chi(s, \pi) - G(\chi(s, \pi))C(s) = \pi
\]

Differentiating this expression w.r.t. \( s \) and \( \pi \) yields

\[
\begin{align*}
(1 - G_z C)\chi_s &= GC_s \\
(1 - G_z C)\chi_\pi &= 1
\end{align*}
\]

These expressions give in combination with (A1) that \( \chi_s > 0 \) and \( \chi_\pi > 0 \).

Substitute \( \chi(s, \pi) \) for \( z \) in (A1) and differentiate w.r.t. \( n \):

\[
\frac{1}{n^2} - G_{xx} \chi_\pi \frac{d\pi}{dn} = (G_{xx} \chi_s + G_z C_s) \frac{ds}{dn}
\]

The parenthesis on the R.H.S. is positive, so the stability condition \( d\pi/dn < 0 \) implies \( ds/dn > 0 \).
(A1) and (A2) give

\[
\frac{1}{n} = 1 - G_x C = 1 - G_x \frac{x - \pi}{G(x)} = 1 - \theta + \pi \frac{G_x}{G}
\]

Substituting this into (A3) gives us

\[
\frac{G(x)}{1 - \theta + \pi G_x/x} - E(s)L = 0
\]

Now define \( \Gamma(x, \pi) = \frac{G(x)}{1 - \theta + \pi G_x/x} \) and substitute \( \chi(s, \pi) \) for \( x \):

\[
\Gamma(\chi(s, \pi), \pi) - E(s)L = 0
\]

Differentiating this w.r.t. \( n \) at \( \pi = 0 \):

\[
[\Gamma_x(x, 0)\chi_\pi + \Gamma_\pi(x, 0)] \frac{d\pi}{dn} + \frac{\partial}{\partial s}[\Gamma(\chi(s, 0), 0) - E(s)L] \frac{ds}{dn} = 0 \quad (A4)
\]

Here

\[
\Gamma_x(x, 0)\chi_\pi + \Gamma_\pi = \frac{G\theta_x}{(1 - \theta)^3} > 0
\]

so (A4) together with \( d\pi/dn < 0 \) and \( ds/dn > 0 \) yield

\[
\frac{\partial}{\partial s}[\Gamma(\chi(s)) - E(s)L] > 0
\]

Q.E.D.
REFERENCES


