# Incentive Contracts and Total Factor Productivity

BENJAMIN BENTAL* AND DOMINIQUE DEMOUGIN†

February, 2002

**Preliminary version**

ABSTRACT

This paper focuses on the endogenous determination of effort as a source of productivity growth. In the model, workers may be self-employed or hired in an "industrial sector" characterized by double-moral hazard. On one side, workers' effort is not observable, but generates contractible signals. On the other side, detecting signals requires costly monitoring with inputs only observable by firms. Resulting labor contracts are bonus schemes. In addition, firms employ capital. As capital accumulates, labor demand increases, forcing firms to raise bonuses. Economic growth increases the "industrial sector" and employees' effort and is thus associated with increased labor productivity.

[JEL. D82, O40]

KEYWORDS: INCENTIVE CONTRACTS, TOTAL FACTOR PRODUCTIVITY, ECONOMIC GROWTH

*University of Haifa, Haifa Israel. *E-mail* bbental@econ.haifa.ac.il
†Humboldt University at Berlin, Berlin Germany. *E-mail*: ddemoug@wiwi.hu-berlin.de

# 1 Introduction

Understanding the large differences among countries in total factor productivity (hereafter TFP) is a challenge of great importance to the economics profession.[1] The striking fact is that at the aggregate level, TFP is closely correlated with income (see, for example, Jones [1998]). As a matter of fact, the observation that TFP differences are "responsible" for almost the entire differences in income, has motivated Prescott to impress upon the economics profession the "need" for "a theory of total factor productivity" (Prescott [1998]).

Some economists associate the differences in TFP with differences in the access to technology (Romer, [1993]). Others have pointed to differences in factor endowments (in particular of skilled workers) as a source of differences in TFP (Mankiw, Romer and Weil [1992]). Acemoglu and Zilibotti [1999] take the impact of the factor endowment a step further. According to their view, economies which are well endowed with skilled workers are also those that develop new technologies. However, these technologies are suited to skilled workers and not to the less-skilled workers found in LDCs, and therefore the free flow of ideas is insufficient to close the TFP gap.

Prescott [1998] argues that TFP differences are not necessarily due to differences in the stock of knowledge. Prescott cites several studies that demonstrate that TFP differences are associated with differences in work practices and organization. In his conclusion Prescott focuses on "the resistance to the use of better technologies" and "the policy arrangement a society employs" that affects this resistance. Hall and Jones [1999] identify social obstacles that hinder some economies from adopting high-productivity production technologies. They concentrate on "social infrastructure" as an explanatory variable. According to this explanation, countries whose policies are "favorable to productive activities - rather than diversion - produce

[1] Significant TFP differences among OECD countries have also been measured at the industry level (Harrigan [1997]).

much more output per worker". Parente and Prescott [1999] argue that poor economies remain poor because monopolists that control factor supply prevent the adoption of superior technologies. Kocherlakota [2001] concentrates on the technology adoption issue formally. In his paper, it is the ability to enforce a social contract that makes the difference. Economies in which such an enforcement is not possible, do not adopt a superior production technology (which is available at some cost), while economies in which the social contract is enforceable, do.

We take a somewhat different line of explanation. In our paper, total factor productivity reflects neither the knowledge of how to produce, nor the factor endowment or the composition of the labor force. Rather, total factor productivity reflects the effort that workers exert in the production process.[2] Moreover, we show that the amount of effort grows endogenously together with output and therefore total factor productivity also increases as the economy grows.

These results are generated in a framework that combines a standard growth model with models of optimal incentive contracts. Specifically, we consider an economy in which two technologies may be used to produce the same good. The first technology uses only labor as input and workers are "self employed". However, not all workers are identical in their productivity if they choose to be self employed. The second technology uses capital and labor and is operated by "firms".[3] In this technology all workers are ex-ante identical. The productivity of the workers depends on the amount of effort they exert. Workers dislike exerting effort, and their effort level is not directly

[2]Schmitz [2001] provides a very detailed analysis of labor productivity of the U.S. and Canadian iron-ore industries. The study shows very clearly that great productivity gains can be attributed to changes in effort per hour worked and in work rules. Schmitz argues that his case study demonstrates that "productivity differences across other industries may be also due in some substantial part to factors other than production technology, physical capital and human capital."

[3]For the sake of argument, one may think of the two "technologies" as representing "agriculture" and "industry".

observable. This creates a moral hazard problem for the employers. On the other hand, workers are assumed to emit some signals that are positively related to their effort. These signals may be measured by the employers. The precision of the measurement depends on the intensity with which these signals are monitored by the employers. That intensity is generated at a cost, a fact that creates another moral hazard problem between workers and employers.

Demougin and Fluet [1998] have shown that in an environment like this, with risk-neutral workers, the equilibrium is characterized by a bonus contract. This contract stipulates that a bonus is paid if the signal emitted by the worker is favorable. The workers choose the effort level taking as given the monitoring intensity and knowing the impact of effort on the probability that the signal they emit will be detected. The employers choose the monitoring intensity taking into account the effect of their choice on the effort level it induces.

In a dynamic setting, workers are assumed to maximize their infinite-horizon discounted expected utility. The workers are assumed to be risk-neutral with respect to consumption. We show that workers choose every period their employment status (self-employed or employee) and their effort level (if relevant), so as to maximize their periodic expected labor income, given the structure of the employment contract. After the monitoring results are realized (if relevant), workers determine how much to consume and how much to save. Savings turn into next-period capital which is rented to firms in a perfectly competitive capital market.

Firms hire capital and decide on the monitoring intensity every period so as to maximize profits. The particular structure of the bonus contract that is chosen, namely the level of the bonus and the probability with which it is paid, turn out to depend on the amount of capital hired by the firm. This implies that the equilibrium effort level and the number of employees also depend on that amount of capital. In fact, the dynamic path we generate is characterized by increases in capital that induce "better" labor contracts (higher bonus and more monitoring) and higher effort. In addition, more workers become employees, leaving the more efficient workers as self-employed. As a result, productivity in the economy increases without any technical progress or change in human capital.

The paper starts with a formal presentation of the model. In this section we discuss the static problem of the workers and of the firms and derive the optimal bonus contract. We also show that the contract is consistent with the dynamic optimization problem of the workers. Next we parameterize some key functions in our economy and derive the equilibrium conditions for that specific case. We conduct some comparative static experiments on the steady-state of the economy. Finally, we numerically evaluate a dynamic equilibrium path and discuss its properties. In the last section of the paper, we offer some concluding remarks.

# 2 The model

We consider a single good, discrete time economy with a constant population of infinitely lived households indexed on the unit interval by $z$. We describe the households in some detail, before we specify the two technologies with which the good can be produced

## 2.1 Households

Each household $z$ consists of a continuum of identical members over the unit interval. The household owns $k_t(z)$ units of capital at the beginning of period $t$ that are inelastically supplied to the capital market at the rental rate $r_t$. In addition, each member $h$ of the household is endowed with 1 unit of labor per period that is inelastically supplied. Every member of the household may exert effort $a_t^h(z)$. Effort has a potential effect on that member's labor productivity in a way to be specified below. However, household members are not decision-making units. They are agents of the household and carry out its decisions, in particular those concerning effort.[4]

All households in the economy are identical. At every period, they are assumed to care (positively) about their aggregate consumption, $x_t$, and (negatively) about the amount of effort exerted by their members, $a_t$. [5] Households are assumed to be maximizing the discounted stream of momentary utility:

$$\sum_{t=0}^{\infty} \eta^t u(x_t, a_t) \tag{1}$$

For the momentary utility, we specify

$$u(x_t, a_t) = x_t - c(a_t) \tag{2}$$

where the function $c(a)$ measures the disutility of effort. The function $c(a)$ is assumed to be increasing and convex with $c(0) = 0$. In addition, we impose a minimum subsistence level per member denoted by $\underline{x}$.[6]

Labor can be used in either one of two technologies. In one of them, the workers will be referred to as being "self-employed" and in the other as "employees". If self-employed, a member of household $z$ produces output $y(z)$ at no effort cost. Without loss of generality, households are ordered in such a way that $y'(z) > 0$.

Alternatively, a household may decide to send its members to the labor market as employees. Once employed, each member exerts the house-

[4] The goal of this structure is to remove any idiosyncratic uncertainty at the household level (see Shi [1998] for a similarly motivated specification). Specifically, one may think of the household members as "machines" that are "programmed" by the household and have no will of their own. Alternatively, one may think of the household as an institution that can fully and costlessly monitor its members.

[5] Clearly, one may also think of the household preferences as the aggregate over individual member preferences, defined over their consumption and effort, i.e.

$$\int u(x_t^h, a_t^h) dh \ .$$

However (1) follows since $x_t^h = x_t$ and $a_t^h = a_t$ for all members $h \in [0, 1]$ of the household.

[6] The linear specification of the utility function is used for parsimony to keep the subsequent employement contract simple. The subsistence level is introduced to generate more realistic dynamics in the face of that linearity.

hold determined effort level $a$, and obtains a corresponding compensation in terms of output. Despite the fact that members of households have different productivity if self-employed, they are assumed to be equally productive as employees.

The budget constraint of the household depends on the employment status of its members. We discuss the specifics after the introduction of the optimal bonus scheme. To keep the language simple, we refer to the representative household member as "worker", "employee" or "self-employed" according to the corresponding status.

## 2.2 Firms

Firms are employing capital and labor. The effectiveness of labor provided by an employee depends on the effort exerted by that employee. Specifically, a worker who exerts effort $a$, supplies $e = g(a)$ effective units of labor, where the function $g(\cdot)$ is assumed to be strictly increasing. Anticipating that it will be to the firms' advantage to provide equal effort incentive across workers, we write the production function of firms as $F(K, eL)$ where $K$ and $L$ denote the per-firm capital and labor employment.[7] We make standard assumptions on the production function. In particular, it is assumed that the production function is homogeneous of degree 1 in both arguments.

## 2.3 The double moral hazard problem

In this subsection, we solely focus on the contractual game between a worker and a firm. We assume that the game has to be played every period independently of the past. In doing so, we rule out long term contracts and any reputation effect either on the part of firms or employees.

Our key assumption is that workers' effort is not directly contractible. As a result, workers' behavior is affected by problems of moral hazard. On the

[7] Assuming effective labor is additive across workers, production can be written as $F(K, \int_L g(a(h))dh)$ where $L$ is the set of employees of the firm.

other hand, it is assumed that firms can generate contractible information on effort. This introduces the ability to mitigate the moral hazard problem through the use of proper incentives. More concretely, we assume that workers are emitting noisy signals related to their effort level. At some costs, these signals can be measured and made verifiable. Accordingly, these signals become contractible. However, the fact that these measurements are costly to the firm introduces a further moral hazard problem, this time on the part of employers. Our assumption here is that though information is verifiable, the precision of that information is not.[8] This double moral-hazard problem is resolved through a game which determines the precision at which the signals will be measured and the extent to which they will be used in the labor contract.

In order to derive the optimal decision of the firm in the appropriate game, we assume here that in each period workers maximize their income net of effort costs. In the next subsection, we show that under the derived contract, this presumed behavior is consistent with the preference specification as given in (1) and (2). Because the same game is repeated every period, we omit the time index whenever confusion not possible.

The firm faces the problem of how much effort to induce. This decision entails a choice of an employment contract and of the amount of resources it allocates to the process of measuring the emitted signals. The latter determines the precision at which these signals are measured, which is parametrized by $\theta$.

At this point, we can draw from existing results in the literature. In particular, it is known that in the current setting – due to the risk-neutrality of both parties – optimal incentive contracts are of the bonus type where a

[8]To give a concrete example, suppose that university contract promises a positive tenure decision whenever a tenure commission presents two ‘good’ reports from external qualified academics. The precision of such a scheme is obviously manipulable since a tenure commission could always ask for more that two reports and only present those reports that are found advantageous.

worker receives a fixed payment $A$, and depending on the realization of the measured signal, a bonus $B$.[9] These results depend on the aforementioned assumption that the distribution of the signals is affected by the worker's effort. Moreover, consistent with the moral hazard problem of the firm, it is assumed that this distribution also depends on the precision of measurement.

Since the optimal contract is of the bonus type, the measured signals can be aggregated to a binary random variable, $\chi \in \{0, 1\}$, where the worker receives the bonus if $\chi = 1$. We denote

$$p(a, \theta) = \Pr[\chi = 1 \mid a, \theta] \ . \tag{3}$$

We assume that $p_a > 0$ and $p_{aa} < 0$. Heuristically the first requirement means that $\chi = 1$ constitutes a 'favorable' information with respect to the agent's action in the sense of Milgrom (1981). The concavity requirement guarantees that the agent's problem is well behaved. The conditions are necessary and sufficient for any action to be implementable with the binary signal $\chi$. The requirements are implicit conditions on the initial information structure. Demougin and Fluet [1998] have shown that assuming the underlying information system satisfies MLRC and CDFC is sufficient to guarantee $p_a > 0$ and $p_{aa} < 0$.

Finally, we let $\phi(\theta)$ denote the resource cost of precision per worker. Regarding precision, we assume $p_\theta < 0, p_{\theta\theta} < 0$ and $\phi_\theta > 0, \phi_{\theta\theta} < 0$. The conditions on the first derivative are not real restrictions. They would naturally follow if one were to fully model the information acquisition problem of firms. The concavity requirements guarantee that first order conditions are sufficient.

The timing of the game between firms and workers (within a period) is as follows. Firms offer a bonus contract $\{A, B\}$ where $A$ denotes the fixed payment and $B$ the bonus part of the contract. In addition, firms announce a precision level $\theta$ by which they intend to measure the signals. We assume that $\theta$ is not contractible, which creates the double moral hazard problem and

[9] See Park [1995], Kim [1997] and Demougin and Fluet [1998].

requires the firm to make a credible announcement. Workers either decide to work for firms or to remain self-employed. Those who become employees select their level of effort, given the announced precision. The firms select the precision, which, in equilibrium, is the same they have announced. Finally, signals are observed and payments are made.

## 2.4 The bonus contract

For the subgame where firms select precision and workers make a choice of effort, we use the Nash equilibrium concept. Starting with the problem of the employee, suppose a worker has chosen to work for a firm. That worker faces a bonus contract $(A, B)$ and expects the firm to implement a precision level $\theta$. The worker chooses effort to maximize utility derived from the employment contract. Analytically, he solves

$$\max_{a} \quad A + Bp(a, \theta) - c(a) \ . \tag{4}$$

From the foregoing, the first-order condition is sufficient. Rewriting that condition yields the bonus which the firm must pay to induce effort $a$, given the precision level $\theta$:

$$B = \frac{c'(a)}{p_a(a, \theta)} \ . \tag{5}$$

Given that a worker works for a firm according to the bonus scheme $\{A, B\}$ and given that the firm expects the worker to produce the effort $a$, the firm will choose precision, $\theta$, to minimize its expected costs – i.e. the sum of precision costs plus expected bonus and fixed payments to the worker. Hence, a firm will solve

$$\min_{\theta} \quad A + Bp(a, \theta) + \phi(\theta) \ . \tag{6}$$

Again, the first-order condition is sufficient, yielding:

$$Bp_{\theta}(a, \theta) + \phi'(\theta) = 0. \tag{7}$$

Given a contract $\{A, B\}$, the solution to the Nash game is a pair $(a, \theta)$ that solves (5) and (7). [10]

## 2.5 The overall problem of the firm

In this subsection, we embed the contractual game in the larger context of the firm's overall decision problem. At this stage, in addition to determining the bonus contract – thereby inducing the desired effort – the firm must also select precision, capital and employment. The firm takes as given the rental rate of capital. In addition the firm faces a reservation utility for employees. Altogether, the firm's problem can be written as:

$$\max_{a,e,\theta,A,B,K,L} \quad F(K, eL) - [A + Bp(a,\theta) + \phi(\theta)]\, L - rK + (1-\delta)K \tag{8}$$

$$Bp_a(a,\theta) - c'(a) = 0 \tag{9a}$$

$$A + Bp(a,\theta) - c(a) \geq \overline{y} \tag{9b}$$

$$Bp_\theta(a,\theta) + \phi'(\theta) = 0 \tag{9c}$$

$$g(a) = e \tag{9d}$$

where $\delta$ is the depreciation rate, and $\overline{y}$ denotes the reservation utility of workers. The constraints (9a) and (9b) are the agent's incentive and participation conditions and (9c) is the credibility requirement for the firm's announcement of precision. Equation (9d) relates the induced effort level $a$ to the labor productivity factor $e$.

In the remaining, it is advantageous to use the homogeneity of the production function to rewrite the objective function of the firm in per worker terms. Specifically denote

$$Lf(k, e) \equiv F(kL, eL) \ ,$$

[10] Note that the solution may not be unique. This, however, is not a problem since by announcing the contract the firm can select the best equilibrium for itself.

where $k$ measures the capital labor ratio. Substituting this definition in the firm's problem and abstracting from the employment decision, we can rewrite problem (8). Since the firm takes $\overline{y}$ as given, it is easily seen that (9b) will be just binding. Therefore $A$ can be eliminated from the firm's problem. Similarly, $B$ can be eliminated by using the constraints (9a) and (9c). Altogether the optimization problem of the firm can be reduced to a Lagrange problem:

$$\max_{a,\theta,k} \mathcal{L} = f(k, g(a)) - [c(a) + \overline{y} + \phi(\theta)] + (1 - \delta - r)k \\ + \lambda \left( \phi'(\theta) - \frac{p_\theta(a,\theta)}{p_a(a,\theta)} c'(a) \right) \qquad (10)$$

## 2.6 The household's problem revisited

When deriving the optimal bonus contract above, $\mathcal{C} = \{A, B, \theta\}$ we have assumed that households care about the expected income net of effort costs of each of their members (see (4)). Here we show that under the derived bonus contract this assumption is consistent with the dynamic optimization problem of households.

In general, every household decides each period on the employment status of its members, possibly on effort, and on saving and consumption. Saved income turns into capital next period and will be rented out in the capital market. The resulting optimization problem of households becomes:

$$\max_{\{a_t, x_t, k_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \eta^t [x_t - c(a_t)] \qquad (11)$$

$$x_t + k_{t+1} - r_t k_t = \begin{cases} A_t + B_t p(a_t, \theta_t) & \text{if employed} \\ y(z) & \text{if self-employed} \end{cases} \qquad (12)$$

$$x_t \geq \underline{x} \qquad (13)$$

where households are assumed to have perfect foresight over future rental rates and employment contracts and $k_0$ is given. From the optimization problem it should be evident that the following holds.

**Lemma 1** *Given a contract $\mathcal{C}_t = \{A_t, B_t, \theta_t\}$ in period $t$, the effort level chosen by a household of employees is*

$$a_t = \arg\max_a \; A_t + B_t p(a, \theta_t) - c(a) \; . \tag{14}$$

*Moreover, the fraction of population which is employed is given every period by $z$ where $z$ solves:*

$$y(z) = \max_a \; A_t + B_t p(a, \theta_t) - c(a). \tag{15}$$

Thus despite the dynamic structure, due to the linearity of the utility function, households maximize every period income net of effort costs.

## 3 A specification

In this section we introduce specific functional forms. In particular, we assume

$$p(a, \theta) = a^\theta, \; a \in [0, 1] \tag{16a}$$

$$\phi(\theta) = \phi_0 + \phi_1 \cdot \theta^\alpha, \; \alpha \geqslant 1 \tag{16b}$$

$$c(a) = c \cdot a^\beta, \; \beta \geqslant 1 \tag{16c}$$

$$g(a) = \frac{a^\nu}{1 - a^\nu}, \nu > 0. \tag{16d}$$

The exponential specifications of the cost functions (16b) and (16c) are done purely for convenience. The restriction on $a$ (which is a choice variable) in (16a) is needed because $p(a, \theta)$ is a probability. The conversion of effort into labor effectiveness as specified in (16d) transforms the effort level which is represented by a probability-related variable $a$ into an unbounded labor-effectiveness variable $e = g(a)$. The parameter $\nu$ is a measure of the sensitivity of $e$ to $a$: higher values of $\nu$ convert a given level of effort into lower levels of labor-effectiveness.

The particular specification of $p(a, \theta)$ chosen in (16a) is useful because of the relationship implied by the model between the elasticity of that function and the expected bonus. Specifically, notice that (5) implies that the expected bonus needed to induce a given level of effort, $a$, takes the form:

$$Bp(a,\theta) = \frac{ac'(a)}{\varepsilon_{p_a}}, \quad \text{where } \varepsilon_{p_a} = \frac{ap_a}{p}\ . \tag{17}$$

As a result, one may use $\varepsilon_{p_a}$ as a natural measure of the informational usefulness of signals: the more useful a signal becomes, the lower is the expected bonus required to induce a given level of effort. In general, $\varepsilon_{p_a}$ depends both on $\theta$ and on $a$. For the functional form of $p(a,\theta)$ in (16a) however we get $\varepsilon_{p_a} = \theta$. In this case the informational usefulness of the signal depends on the precision of measurement, $\theta$, but not on the worker's effort, $a$.[11]

Under these assumptions, the first-order conditions of the firm's problem in (10) become:

$$f_e(k,e)g'(a) - \beta c a^{\beta-1} + \frac{\lambda}{\theta}\beta c a^{\beta-1}\left[1+\beta\ln(a)\right] = 0 \tag{18a}$$

$$f_k(k,e) + (1-\delta-r) = 0 \tag{18b}$$

$$-\phi_1\alpha\theta^{\alpha-1} + \frac{\lambda}{\theta}\left[\phi_1\alpha(\alpha-1)\theta^{\alpha-1} - \frac{\beta c a^{\beta}\ln(a)}{\theta}\right] = 0 \tag{18c}$$

$$\alpha\phi_1\theta^{\alpha-1} + \frac{\beta c a^{\beta}\ln(a)}{\theta} = 0 \tag{18d}$$

Furthermore, the worker's incentive and participation constraints yield:

$$B\theta a^{\theta-1} - \beta c a^{\beta-1} = 0 \tag{19a}$$

$$A + Ba^{\theta} - ca^{\beta} = \overline{y} \tag{19b}$$

Finally, we also have three market clearing conditions:

$$f(k,e) - \left[A + Ba^{\theta} + (\phi_0 + \phi_1\theta^{\alpha})\right] + (1-\delta-r)k = 0 \tag{20a}$$

$$r = \frac{1}{\eta} \tag{20b}$$

$$y(z) = \overline{y} \tag{20c}$$

Equation (20a) is the zero profit condition for the marginal entrant in the firm sector. Due to the linear homogeneity, it also implies that the economic

[11] The specific functional form $p(a,\theta) = a^{\theta}$ obtains if the underlying information structure of signals follows a Poisson process. For a derivation of the example, see Demougin and Fluet [2001].

profit is zero for all firms. The condition (20b) requires that in the steady state the market clearing interest rate be given by the worker discount factor. Finally, (20c) determines the opportunity cost of the marginal employee in the firm sector.

The overall system is block recursive and can be easily simplified. First, we eliminate $r$ and $\overline{y}$ from (20b) and (20c). Second, equations (18c) and (18d) imply that $\lambda/\theta = 1/\alpha$. Using the above results in (18a) and in (18b), the system reduces to:

$$f_e(k,e)g'(a) - \beta c a^{\beta-1} + \frac{\beta c a^{\beta-1}}{\alpha}\left[1 + \beta \ln(a)\right] = 0 \tag{21a}$$

$$f_k(k, g(a)) - \left(\frac{1}{\eta} - (1-\delta)\right) = 0, \tag{21b}$$

where $e = g(a)$, and we obtain $a$ and $k$. Using these results, $\theta$ follows from (18d). Finally, $A, B$ and $\theta$ obtain from (19a)-(20a).

# 4 Comparing steady-states

In this section, we perform some steady state comparisons with respect to changes in the time preference and the effort disutility of workers. We also examine the implications of changes in the production and in the monitoring technology as well as in the outside opportunities for the self-employed sector. Moreover, we discuss dynamic adjustments by examining the implication of a surprise change in the respective parameters at the beginning of a period, i.e. at a point in time where capital is already given. Finally, throughout the section, we only consider the case $f_{ka} > 0$ which means that we restrict the analysis to situations where capital and labor are complements.

## 4.1 Changes in the time preference

Here we compare economies which differ in the subjective discount factor of their population measured by the variable $\eta$. Specifically, in one economy workers are less patient, so that in steady state the rental rate of capital, $r = 1/\eta$, is higher.

In order to simplify notation, we rewrite the firms' problem. We solve for the precision variable from (18d), thus eliminating both $\theta$ and the constraint. The firms' problem then becomes:

$$\max_{a,k} \ \pi = f(k, g(a)) - ca^{\beta} - y(z) - \left(-\frac{\beta}{\alpha} ca^{\beta}\ln(a)\right)^{\frac{1}{\alpha}} + (1 - \delta - r)k \ . \quad (22)$$

Writing the first-order conditions in $a$ and $k$, we obtain from implicit differentiation:

$$a_r = -\frac{\pi_{ak}}{\det H} \leqslant 0 \text{ and } k_r = \frac{\pi_{kk}}{\det H} \leqslant 0 \ , \quad (23)$$

where $H$ is the relevant Hessian. The signs follow from the local concavity at the maximum – which yields $\det H > 0$ and $\pi_{kk} < 0$ – and because capital and labor have been assumed to be complements, i.e. $\pi_{ak} > 0$. We summarize our conclusion in the following result:

**Result 1:** *In an economy in which workers are less patient, in steady state employees apply less effort and the capital-labor ratio is lower.*

The result is very intuitive. Less patient workers want to consume more in the current period. Their saving is lower, which induces a higher interest rate and lower equilibrium capital labor ratio. The reduction in effort follows from the complementarity assumption. Thus, other things being equal, we conclude that a highly capitalized economy (namely, an economy with a high capital-labor ratio) should also be characterized by high effort of employees. The latter implies a high total factor productivity.

Next, we turn to the labor contract, $\mathcal{C}$, required to induce the desired effort, $a$. The ingredients of $\mathcal{C}$ follows from (18d)-(19b) in a block recursive way; (18d) yields $\theta$, then given $\theta$ (19a) yields $B$ and, finally, $A$ results from (19b) given $\theta$ and $B$. Noting that $sign(\theta_a) = -sign\,[a + \beta\ln(a)]$, we introduce the following terminology:

**Definition:** *An economy is a high (low) effort economy when* $[a + \beta\ln(a)] > (<)\ 0$ .

According to this definition, in a high effort economy, $a$ and $\theta$ move in opposite directions. The opposite directions of these changes imply that to induce a higher effort, the contract's power, as measured by $Bp(a,\theta)$, must be increased to provide more incentives. In a low effort economy, firms increase monitoring when they want to induce more effort, but they may or may not raise power.

**Result 2:** *Comparing two high-effort economies, steady-state labor contracts in the economy populated by the less patient population are characterized by more monitoring and less power. In contrast, comparing two low-effort economies, the steady-state labor contract in the economy populated by the less patient population are characterized by less monitoring.*

Finally, in order to obtain the effect of a change in $\eta$ on the fraction of workers employed in the firm sector, we just set firms' profit in the steady state equal to zero. Implicit differentiating the equation yields

$$z_r = -\frac{\pi_r}{\pi_z} = -\frac{k}{y'(z)} < 0 \ . \tag{24}$$

**Result 3:** *In an economy with less patient population the steady-state fraction of the labor force employed in the firm sector is smaller.*

Altogether, differences in the worker discount factor (or, equivalently, in the interest rate), induce co-movements in the capital-labor ratio, in employment and effort. Total factor productivity is changing in the firm sector independently of any technical change in that sector.

## 4.2 Changes in the productivity of self-employment

For this section, we rewrite (20c) as $y(z,\mu) = \overline{y}$ where $\mu$ acts as a shift parameter. Without loss of generality, we assume $y_\mu > 0$ which means that a raise of $\mu$ improves productivity in the self-employed sector. It is easy to see from the equilibrium conditions (18a)-(20c) that in the steady state variations in $\mu$ only affect $z$. Specifically, $y(z,\mu)$ must stay constant. This observation implies that in an economy with a higher $\mu$, the steady state fraction of the labor force who choose to become employees, $z$, is lower. Since in the steady state the capital labor ratio remains unchanged, other things being equal, such an economy has less capital.

The comparison of the steady-state leads to an analysis of the adjustment path which results from such an unexpected change in $\mu$. Initially, total capital is fixed and the outflow of workers from the firm-sector causes the rental rate of capital to fall. From the foregoing section, we know that this induces an increase in the capital labor ratio and in effort (relative to the original steady-state). We summarize our conclusion in the following result:

**Result 4:** *A surprise reduction in the productivity of self-employment causes a temporary reduction in the capital labor ratio and effort and an increase in employment in the firm sector. Along the adjustment path, employment in the firm sector increases towards a new, higher level and the capital-labor ratio and effort increase to their initial levels.*

Assuming a slow adjustment in capital, result 4 suggests sharp jumps in $a, k$ and $z$ at the time of the surprise increase in $\mu$ followed by a period where society is decumulating capital. At some point the interest rate starts raising which will bring the system back to steady state at a higher level of self-employed.

Obviously, since in steady-state $a$ and $k$ remain unchanged the incentive contract also remain unaffected. However, this is not true along the adjustment path as firms find it optimal to induce higher effort levels. Whether these higher effort levels are induced through more monitoring or more power depends again on whether or not we are in a high or low level economy.

## 4.3 Changes in the monitoring costs

In this section, we consider a decrease in monitoring costs. In principal, we could examine either changes in the fix or the variable monitoring costs. However, it turns out that due to the specification of the model variations in $\phi_1$ are fully compensated by adjustments in $\theta^\alpha$ – see (18d) – such that total monitoring costs remain unchanged. Not surprisingly $a$ and $k$ also remain unchanged as can be verified in (21a) and (21b). Nevertheless, the structure of the incentive contract must adjust to account for the variation in monitoring. For example, a reduction in $\phi_1$ increases $\theta$ thereby reducing the required power of the contract as can be derived from (19a).

Of more interest are variations in $\phi_0$.[12] Suppose, we consider a reduction in the fix costs of monitoring. Again in steady state $a$ and $k$ remain unchanged as monitoring costs do not influence either (21a) or (21b). Moreover, monitoring and power are also unaffected in the long run. However, participation in the firm sector goes up due to the zero profit condition. Just as in the foregoing section, we can think of the adjustment process. Initially, a reduction in $\phi_0$ induces firms to hire more workers. Since capital is temporarily fixed, this generates a reduction in the capital labor ratio. Therefore induces an increase in $r$ and a reduction in $a$. Over time the economy accumulates capital, the firm sector continues to grow while the interest rate falls and effort increases to its steady state level.

## 4.4 Changes in contract enforcement

Until now, we have assumed that the legal environment, in which firms and workers interact, is such that agreed upon contracts will be enforced. For example, in the case underlying the specification, we found that if no errors are detected the principal should pay the agent a bonus. This in turn produced the incentive for firms to undertake monitoring in order to detect mistakes. Credible monitoring then allowed to align incentives between employers and employees. In this section, we briefly explore the possibility of weakening the legal environment. Specifically, we assume that for an exogenous proportion $1 - q$ of employment contracts between a firm and its worker, the agreement is not honored. In that case, we assume that the firm knows it will not pay

[12]Note that the fix cost of monitoring have been defined in per capita terms. As a result a reduction in $\phi_0$ can also be thought of as lowering hiring costs.

the bonus and, thus, decides not monitor at all. Nevertheless, we do assume that offers are all the same across workers, since otherwise offers would work as signals.

Obviously, this is a very crude and add hoc way of introducing variations in the credibility of the legal structures. On the other hand it is very intuitive and, more importantly, does not require large changes in the foregoing model. In particular, the foregoing model is nested for $q = 1$. In contrast, with $q = 0$ the legal environment fully breaks down and no contract is credible.

Introducing these changes affect the constraint of (10) which becomes:

$$\phi'(\theta) + \frac{p_\theta(a,\theta)}{qp_a(a,\theta)}c'(a)$$

Applying the change to the specification of section 3 marginally changes equation (21a):

$$f_a(k, g(a)) - \beta c a^{\beta-1} + \frac{\beta c a^{\beta-1}}{q\alpha}\left[1 + \beta \ln(a)\right] = 0 \tag{25}$$

with the two equation being equivalent at $q = 1$. Together with (21b) the two equations allow us analyze the effect of a change in $q$. We obtain from the implicit function theorem the following result.

**Result 5:** *The effects of variations in the legal system depends on whether the economy is in a low or high effort economy. In the case of a low effort case, $a_q^{**}, k_q^{**} > 0$. Effects are reversed in the high effort environment.*

The result implies that weakening the credibility of the legal environment is always bad. In a low effort economy $a^{**}(q) < a^*$, the result implies that effort is even further reduced. In contrast, in a high effort economy workers are driven to work even harder. Of course, in the latter case, workers are already working too much as compared with the efficient solution.

# 5 Numerical Experiments

## 5.1 Monitoring Cost

We conduct a numerical experiment to assess the impact of the monitoring cost on the steady-state behavior of the economy. In particular, we look at two economies that are identical in all respects, except the monitoring-cost elasticity, $\alpha$ (see (16b)). This elasticity represents the effectiveness of the monitoring device: The higher $\alpha$ becomes, the lower the monitoring cost for a given level of precision $\theta$.

Except for the value of $\alpha$, both economies are characterized by the same parameters. In the monitoring cost function (16c) we set: $\phi_0 = 0$ and $\phi_1 = 3$. For the effort disutility, (16c), we set: $c = 6$ and $\beta = 10$. We specify the production function as:

$$F(K, eL) = T \cdot K^{\gamma}(eL)^{1-\gamma}, \tag{26}$$

with $T = 1.4$ and $\gamma = 0.35$. We set the value of $\nu$ in (16d) to 4. The self-employed produce according to:

$$y(z) = u_0 \cdot z^{\delta}, \tag{27}$$

with $u_0 = 0.5$ and $\delta = 2$.[13] The depreciation rate, *dep*, is set at 0.08. The discount factor, $\eta$, is 0.95. Finally, we set $\underline{x}$ (the subsistence level) to 0.[14]

The following table compares an economy with a highly effective monitoring technology ($\alpha = 4$) to one in which the monitoring technology is much inferior ($\alpha = 1.02$). We also include an intermediate case ($\alpha = 2$).

[13]The parameter $\delta$ seems to be very large in terms of its implication for the average productivity of self-employed workers as a function of their fraction in the labor force. The choice is made, at this stage, so as to make the dynamic effects sufficiently big to be seen easily.

[14]These parameters are chosen so as to obtain a reasoanble-looking charaterization of the steady-state of the intermediate-monitoring-cost economy. In addition, these parameters generate pronounced differences between the low- and high-monitoring cost economies.

Table 1: Steady-States With Different $\alpha$

| Variable | | | |
|---|---|---|---|
| Monitoring-cost elasticity ($\alpha$) | 4.00 | 2.00 | 1.02 |
| Per-capita Total Output ($\overline{y}$) | 0.92 | 0.35 | 0.23 |
| Saving Rate | 0.23 | 0.17 | 0.09 |
| Effort ($a$) | 0.75 | 0.63 | 0.54 |
| $\theta$ | 0.54 | 0.21 | 0.03 |
| $z$ | 0.73 | 0.54 | 0.40 |
| $p(a, \theta)$ | 0.86 | 0.91 | 0.98 |
| $k/\overline{y}$ | 2.89 | 2.11 | 1.18 |
| $wz/\overline{y}^1$ | 0.62 | 0.72 | 0.84 |
| $m/\overline{y}^2$ | 0.20 | 0.20 | 0.13 |
| $y/\overline{y}^3$ | 0.89 | 0.60 | 0.32 |
| $TFP_y^4$ | 0.62 | 0.33 | 0.21 |
| $TFP_z^5$ | 0.38 | 0.31 | 0.26 |

Notes:

[1]Expected wage-bill relative to total output

[2]Expected monitoring cost relative to total output[15]

[3]Share of "industrial" sector in total output

[4]Total factor productivity in "industrial" sector

[5] Output per worker in "self-employed" sector.

Table 1 reveals that the economy that has the most effective monitoring technology has a per-capita output value that is 4 times higher than that of the economy with the worst monitoring technology. The fraction of the labor-force that is employed in the "industrial sector" in the first economy is close to double that of the third. Effort is measured with a much higher precision in the first, and total monitoring costs amount to a much higher fraction of output. The capital/output ratio in the first economy is about as high as that of the US, while that of the third economy is more in the range of a typical LDC. The wage bill in the first economy is lower relative to output because

[15]Notice that the monitoring costs are *not* counted as part of the economy's total output.

of the high importance of capital in that economy. Finally, the productivity measures in the first economy are significantly higher. Specifically, the total factor productivity in the "industrial sector" that amounts to almost 90% of total output in the first economy, is about 3 times as larger than it is in the other economy.[16]

## 5.2 Dynamics

We turn to the growth path of the economy with the intermediate monitoring cost technology ($\alpha = 2$). The economy is started with a quarter of its capital, and with a subsistence 0.22. The economy reaches the steady-state after 22 periods. We describe first the economy's growth performance.

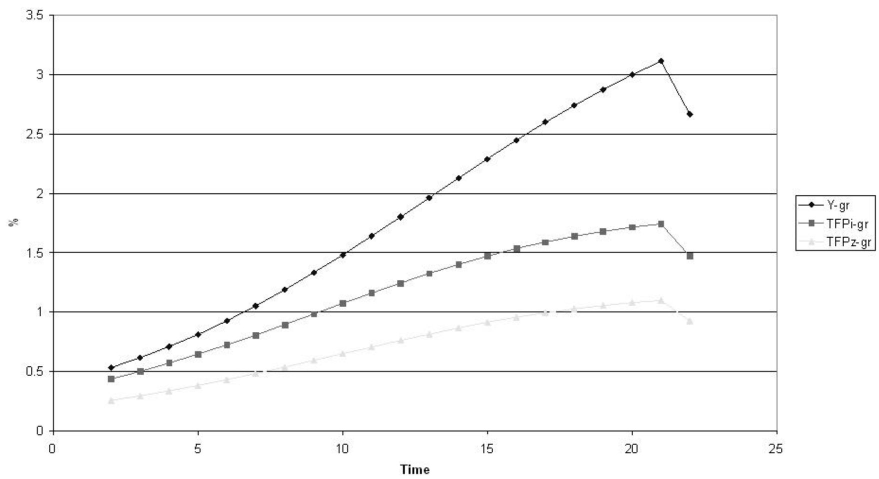


Figure 1: Growth Rates

[16]While this example cannot match the 7.7 factor Hall and Jones [1999] find for the difference between the *U.S.* productivity and that of Niger, our model gets the slope of Figure I in their paper (on page 90) about right. There, a factor of 4 in output per worker is associated with a factor of 2.5 in total factor productivity.

Figure 1 depicts the growth rates of output as well as of the total factor productivity in the "industrial" sector and of the output per worker in the "self-employed" sector. As can be seen from the figure, the growth rates are all accelerating until one period before growth stops (at the steady-state). Total factor productivity growth amounts to approximately a half of the output growth. The result is due to the fact that as the economy is accumulating capital, more workers are needed for the "industrial sector". These workers have to be enticed away from self-employment. As the productivity of the marginal self-employed worker increases, a better contract has to be offered to the industrial workers. This, in turn, entails more monitoring and higher productivity.

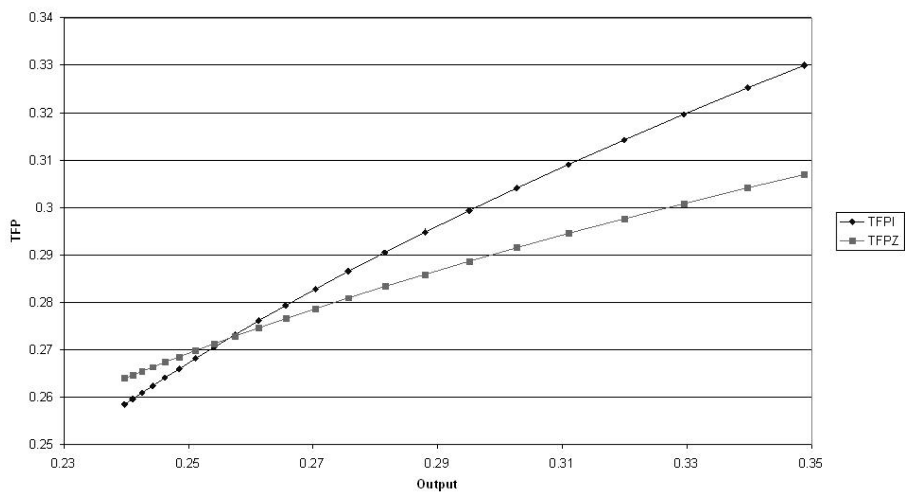


Figure 2: Output and Productivity Levels

This effect can also be seen when one looks at levels, rather than growth rates. Figure 2 shows that as output increases, so do the productivity measures. In fact, the relationship between the endogenously driven productivity in the "industrial" sector and output is even more pronounced than that of the productivity gain in the self-employed sector.

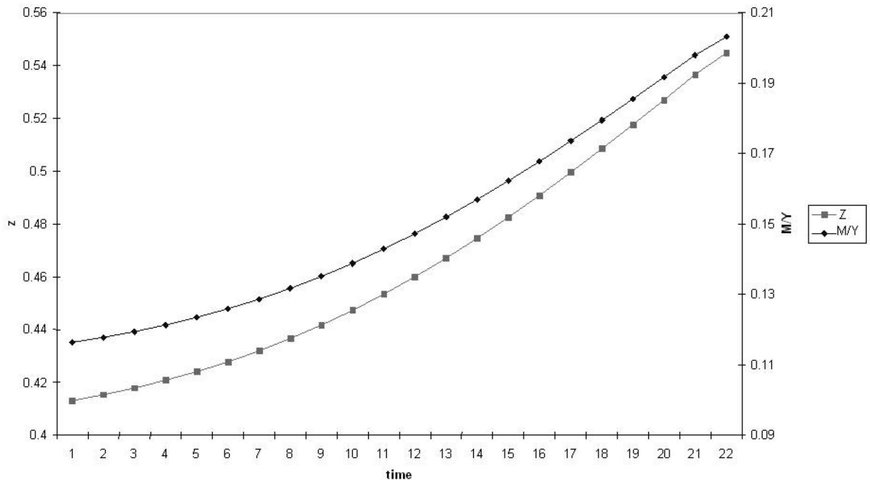


Figure 3: Monitoring and Employment

Finally, in Figure 3 we show the evolution of the fraction of workers employed in the "industrial sector" and the monitoring costs relative to output. The figure clearly shows that the need to entice more workers into the "industrial sector" is indeed associated with higher monitoring costs.

# 6 Conclusion

This paper generates total factor productivity gains that are unrelated to ant technological progress. In fact, production technologies are kept constant throughout the analysis. However, out of steady-state the economy is accumulating capital. The workers who use this capital need to be enticed away by the "industrial sector" from an alternative occupation. The productivity of the marginal worker in that alternative occupation is assumed to increase when the number of workers who are not yet employed in "industry" decreases. This requires a higher wage. To justify that higher wage, workers need to exert more effort. To induce that higher effort, employers must increase their investment in monitoring, and the result is higher productivity.

Thus, the mechanism we have shows how the increased pressure from an alternative sector (in our case - the "self employed" sector) induces increased productivity in the "industrial sector". Clearly, one may think of other sources of pressure that may trigger the same mechanism. Competition from other countries generated by trade-liberalization can clearly be one such factor.

## REFERENCES

Acemoglu, D. and F. Zilibotti, 1999: "Productivity Differences", NBER Working Paper Series # 6879.

Demougin, D. and C. Fluet, 1998: "Mechanism Sufficient Statistic in the Risk-Neutral Agency Problem", *Journal of Institutional and Theoretical Economics*, 154(4), pp. 622-639.

Demougin, D. and C. Fluet, 2001: "Monitoring versus Incentives", *European Economic Review*, 45, pp. 1741-1764.

Hall, R. E. and C. I. Jones, 1999: "Why Do Some Countries Produce So Much More Output Per Worker Than Others?", *The Quarterly Journal of Economics*, pp. 83-116.

Hall, Robert E. and Charles I. Jones, 1999: "Why Do Some Countries Produce So Much More Output per Worker Than Others?", *The Quartery Journal of Economics*, pp.83-116.

Harrigan, J., 1997: "Estimation of Cross-Country Differences in Industry Production Function", NBER WorkingPaper Series # 6121.

Jones, C. I., 1998: *Introduction to Economic Growth*, W. W. Norton and Company.

Kim, S. K., 1995: "Efficiency of an Information System in an Agency Problem", *Econometrica*, 63, pp. 89-102.

Kocherlakota, N. R., 2001: "Building Blocks for Barriers to Riches", Federal Reserve Bank of Minneapolis, Research Department Staff Report 288.

Milgrom, P. 1981: "Good News and Bad News: Representation Theorems and Applications", *Bell Journal of Economics*, 12, pp. 380-391.

Parente, S. L. and E. C. Prescott, 1999: "Monopoly Rights: A Barrier to Riches", *The American Economic Review*, 89(3), pp. 1216-1233.

Park, E. S., 1995: "Incentive Constraints Under Limited Liability", *Journal of Economics and Management Strategy*, 4, pp. 477-490.

Prescott, E. C., 1998: "Needed: A Theory of Total Factor Productivity", *International Economic Review*, 39(3), pp. 525-551.

Schmitz, J. A., Jr., 2001: "What Determines Labor Productivity? Lessons from the Dramatic Recovery of the U.S. and Canadian Iron-ore Industries", Federal Reserve Bank of Minneapolis, Research Department Staff Report 286.

Shi, S. , 1998: "Search for a Monetary Propagation Mechanism", *Journal of Economic Theory*, 81, pp. 314-352.