Job additionality and deadweight spending in perfectly competitive industries: the case for optimal employment subsidies

Pierre M. Picard

University of Manchester, School of Economic Studies, Oxford Road, Manchester M13 9PL, UK

Received 1 June 1998; received in revised form 1 December 1999; accepted 1 December 1999

Abstract

This paper links the old literature on employment subsidies with the current theories of contract and regulation. One important source of inefficiency of employment subsidies is non-additional employment and deadweight spending which occur when private firms receive a subsidy for jobs that would have been created without the subsidy. We identify the asymmetry of information between the government and the private firm as the source of these problems. When the government proposes optimal incentive contracts to promote employment, we show that all employment creations are additional and that the deadweight spending is equal to the information rent, which may be null when firms’ types are discrete.

Keywords: Employment subsidies; Contract; Regulation

JEL classification: D82; D61; L33; L51

1. Introduction

The debate on the efficiency of employment subsidies is not novel. The generalized subsidy to all workers proposed by Kaldor (1976) has been regularly criticized for its high exchequer cost. This is the reason why Layard and Nickell (1980) introduced the Marginal Employment Subsidy (MES) that is paid only for
the increase in employment. Still, the efficiency of such subsidy schemes is questioned because of the existence of non-additional employment and deadweight spending. Non-additional employment is the number of jobs that would have been created without the subsidy, whereas additional employment corresponds to the part of employment that would not have been created in the absence of the subsidy. On the other hand, deadweight spending — also called windfall gains — refers to the gains that firms benefit from the subsidies; in the MES scheme, deadweight spending is directly related to the non-additional jobs that unduly received a subsidy.

The relative size of non-additional employment and deadweight spending has been highlighted in some empirical studies. Table 1 presents several studies that report the estimated shares of non-additional jobs supported by employment subsidies in Europe. Of course, comparison of those results should be handled with care because of the diversity of evaluation methodologies and the variety of subsidy instruments. Nevertheless, the overall view is that non-additionality and deadweight spending represent a large part of the inefficiencies of employment subsidy policies.¹ (See Foley (1992) and Picard (1998a) for further discussions on these topics.)

A key aspect of this inefficiency is the government’s lack of information over private firms. In order to overcome this problem, some governments propose discretionary subsidies in which the amount of subsidy to each firm is conditioned by its level of additional employment assessed by a civil officer (see for instance, the Regional Selective Assistance in UK). But this assessment process is particularly difficult and costly as it requires identification of additional and non-additional workers in every subsidized firm. Moreover, this assessment process adds uncertainty to the potentially subsidized projects and may lead to collusion between civil officers and representatives of private firms (see Foley

Table 1
Deadweight spending and/or non-additional jobs as percentage of total cost of subsidies and/or total amount of subsidized jobs

<table>
<thead>
<tr>
<th>Study</th>
<th>Non-additional employment or deadweight spending</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breen and Halpin (1989, Ir)</td>
<td>70%</td>
</tr>
<tr>
<td>Disney et al. (1992, UK)</td>
<td>63–80%</td>
</tr>
<tr>
<td>Foley (1992, UK)</td>
<td>40–90%</td>
</tr>
<tr>
<td>de Koning (1993, NL)</td>
<td>40%</td>
</tr>
<tr>
<td>Van der Linden (1995, B)</td>
<td>53%</td>
</tr>
<tr>
<td>Cotta and Mahy (1998, B)</td>
<td>48%, 64%</td>
</tr>
</tbody>
</table>

¹Other inefficiencies like activity shifts that occur between subsidized and non-subsidized firms and substitution effects that take place between workers are usually estimated for 15 to 40% of the total inefficiencies.
In our opinion, the information problem with respect to each private firm must be addressed, which is the focus of the present paper.

In this paper we will concentrate on the design of optimal employment subsidies under asymmetric information. Three comments are appropriate. The first one deals with government power over private firms. In the case of employment subsidies, private firms voluntarily participate in the subsidy scheme. Therefore, the government does not have the power of a central planner who can nationalize firms in order to impose the socially optimal level of employment and output. Firms will participate in the subsidy scheme only if they obtain more than in the laissez-faire situation in which no firm is offered a subsidy. As this setting respects the liberty of enterprise and ownership rights, it can be referred to as a liberal regulation.

The second comment concerns the causes of unemployment. In this paper we depart from the positive investigation of specific imperfections in labor markets. The point of the paper is to show how the problems of deadweight spending and non-additional employment may be analyzed and overcome. We abstract from specific market imperfections by assuming the latter may be summarized by the wage wedge it generates. The wage wedge reflects the difference between the labor price that private firms actually pay and the shadow price that the government would like to impose, given a cost–benefit analysis such as proposed by Drèze and Stern (1987). Such a wage wedge is typically caused by price rigidities in the labor market as result of a minimum wage. Centralized bargaining that takes place between unions and firms in many European economies also leads to this form of wage wedge as the wage agreement is ultimately applied to each firm in some exogenous way.

The last comment concerns the variables that the government can observe and verify. In this paper we will assume that the government observes, compares and verifies the employment level in each firm. Employment is indeed the most appropriate variable to monitor since in most modern economies, tax administrations and labor organizations exert a strong control on employment. Other variables like output are much less easy to observe and verify for the government. There are indeed too many differences in product qualities and too many transactions in the economy for the government to be able to get a precise and comparable measure of each firm’s output.

In this paper, the government is modeled as a planner maximizing a social welfare objective that takes into account the consumer and producer surpluses, as well as the social cost of public funds and the welfare valuation of employment through the concept of shadow wage. This is a static partial equilibrium analysis with perfect competition in the product market. The government is ignorant about the true productivity parameter of the firms applying for subsidies, which introduces an asymmetry of information between the private firm and the government à la Baron and Myerson (1982). A main difference between the latter and our model is the introduction of a liberal participation constraint in which the
firms’ reservation profits vary with their types. The combination of type-varying participation constraint and informational constraint yields some technical complexities that have recently been studied by Jullien (1997). When the government designs the employment subsidies as optimal incentive contracts, we will show that the deadweight spending will be minimized and will be equal to the information rent. When the types of firms are discrete, this rent may even be zero. Also, the present model will imply that all subsidized jobs must be additional.

The paper thus focuses on employment policies on the demand side of the labor market. There is also a significant literature on the supply side policies, both empirical and theoretical to which our analysis should be viewed as complementary (see Phelps (1994), Snower (1994), Katz (1996), Lindsay (1987), etc.). Section 6 explores the links between our paper and several aspects of this literature, and more particularly that of optimal income taxation (Mirrlees, 1971; Diamond, 1998). The rest of the paper is structured as follows: Section 2 sets the assumptions of the model, Sections 3 and 4 study the full information case and the asymmetric information case, Section 5 makes an important comment on the dynamics of the optimal employment subsidy and the MES and Section 7 concludes. Most proofs are relegated to Appendix A.

2. The model

2.1. The private firm

In this paper, we assume that production $F(\theta, L)$ depends on the firm’s employment level $L$ and on a discrete productivity parameter $\theta \in \{\theta_i, \theta_j\}$ with $\theta_i < \theta_j$. The production function is such that $F_{\theta L} < 0 < F_{L}^{\prime}$, $F_{\theta}^{\prime}$, where the subscripts denote partial derivatives. The type $\theta$ thus ranks the production schedule positively, the firm with $\theta_j$ being the most productive one.

The current paper focuses on the distortion introduced by the labor market. To focus on this distortion, we assume that the private firm is perfectly competitive and sells its product at a price $P$ fixed by the product market. The revenue from sales is then $PF(\theta, L)$. Without loss of generality, we can normalize the price $P$ to 1.

The firm has a homogeneous work force (same skills, same weekly work hours per employee) and the wage $w$ is exogenously fixed in the labor market. The private firm’s profit has the following form:

$$II = F(\theta, L) - wL + T,$$

where the right hand side terms denote the production $F$, the cost of labor $wL$ and the transfer $T$ granted by the government. By assumption, the profit is concave in $L$.

Finally we make use of the Single Crossing Property for the firm’s profit:
\( \frac{\partial}{\partial \theta} \left( \Pi_L / \Pi_T \right) > 0 \). Since the profit is linear in \( T \) and the wage \( w \) is fixed, this yields:

\[ F_{\theta L}(\theta, L) > 0 \ \forall L, \ \theta. \quad (2) \]

With this hypothesis, the parameter \( \theta \) must be chosen in such a way that it monotonically ranks the values of each firm’s marginal product of labor. The higher the type, the higher the marginal product. Examples of such parameters in industries are productivity parameters and staff performance.

When the firm is free from regulation, it receives no transfer and maximizes its profit. The following first order condition yields the level of employment \( L^F \):

\[ F_L(\theta, L^F) = w. \]

By the strict concavity of the production function, this yields a unique maximum at the level of employment \( L^F(\theta) \) with profit \( \Pi^F(\theta) \) that depends on the firm’s type \( \theta \). For the sake of convenience, we will refer to these values as the free level of employment and the free profit.

By the hypotheses on production functions, the free levels of employment and the free profits are ranked according to the firm’s types: \( L^F(\theta) \) and \( \Pi^F(\theta) \) increase with \( \theta \). So, with two types, \( L^F_1 < L^F_2 \) and \( \Pi^F_1 < \Pi^F_2 \). This property is important as the free profit levels will be used to determine the reservation value of the firm when the latter makes its decision to participate in the subsidy scheme.

2.2. The government

The government operates in an economy with involuntary unemployment caused by distortions in the labor market. To model this we assume that the market wage \( w \) is larger than some shadow wage \( w_s \). This justifies the existence of an excess supply of labor. The intensity of this distortion is denoted by wage wedge \( w - w_s \geq 0 \). When the distortion in the labor market is caused by price rigidities or by central bargaining between firms and unions, the reader may approximate the shadow wage by the value of the wage that would prevail in a perfectly competitive labor market.\(^2\) In a rigorous cost benefit analysis as developed by Drèze and Stern (1987), the shadow wage would incorporate the individual opportunity cost of being unemployed, the financing of government activities and unemployment benefits as well as welfare redistribution concerns.

The government’s objective function is then defined as the total surplus minus the social costs of labor and transfers

\[ \mathcal{W} = \int Pdq - w_s L - \lambda T. \]

\(^2\) Insider–outsider and efficiency wages models might not fit well into our specification as they introduce distortions at the firms’ level rather than at the labor market level.
Since the product market is perfectly competitive and the price is normalized to 1, the government’s objective function reduces to
\[
\mathcal{W} = F(\theta, L) - w_u L - \lambda T
\]  
where \( F(\theta, L) \) is the gross consumer surplus generated by the production in each firm. The group of terms \( F(\theta, L) - w_u L \) is what we will refer to as the social surplus. It consists of the surplus that would be reached if the price of labor was set at the shadow wage, \( w_u \). It is concave in \( L \). The term \( -\lambda T \) is the welfare loss incurred by a transfer \( T \) from the taxpayer to the firm. The social welfare effect of this transfer is negative because of the use of distortionary taxes on income, capital and consumption to finance the transfer. The parameter \( \lambda > 0 \) reflects this distortion and is called the shadow cost of public funds (see Atkinson and Stiglitz, 1980).

Finally under informational asymmetry, the government has only some belief \((p, 1 - p)\) on the possible realizations of type \((\theta_i, \theta_s)\). Its objective function is then the expectation of the social welfare \( \mathcal{W} \).

3. The full information employment subsidies

Under full information, the government is able to propose a separate employment contract \((L^*(\theta_i), T^*(\theta_i))\) to each type \( \theta \). In this section we characterize this full information contract and we relate it to the issues of deadweight transfers and non-constant reservation profits.

For each type, the government maximizes its objective function given the participation constraint (PC) that imposes a profit above the free profit \( P^F \). Because of full information, the reference to the true types \((\theta_i, \theta_s)\) can be temporarily omitted. This gives:
\[
\max_{(L, T)} F(\theta, L) - w_u L - \lambda T \text{ s.t. } F(\theta, L) - w_u L + T \geq P^F(\theta). \quad \text{(PC)}
\]

Since the transfer \( T \) is costly to the government and appears linearly in the optimization problem, it will be minimized by making the voluntary participation constraint (PC) binding. At the full information level of employment \( L^* \), the transfer is
\[
T^* = P^F - (F(\theta, L^*) - w_u L^*). \quad \text{(4)}
\]
The problem simplifies to
\[
\max_L [F(\theta, L) - w_u L - \lambda(P^F - (F(\theta, L) - w_u L))]
\]
that yields the following first order condition:
The above equality shows that, by concavity of $F(\theta, L)$, the full information level of employment $L^*$ is higher than the free level $L^F$. The government must consider the social wage $(w_s + \lambda w)/(1 + \lambda)$ as the marginal cost of labor. If there was no shadow cost of transfer, the government would correct the market wage $w$ to the level of the shadow wage $w_s$. However, in the presence of a social cost for transfers ($\lambda > 0$), the government considers a larger social wage.

Under full information, the employment contracts $(L^*, T^*)$ are the solution of Eqs. (5) and (4). Employment levels increase with higher distortion in the labor market (lower $w_s$), lower shadow cost of public funds $\lambda$ and more productive firms (higher $\theta$).

Note that the transfer involves no deadweight spending since firms do not get more profit from the subsidy. Also, all created jobs are additional. The government knows each firm’s type and pays exactly what is required to attain the socially optimal level of employment. If a firm considers the group of employees that were supported by the subsidy, it will necessarily report that none of them would have been hired without the subsidy.

Unfortunately this contract cannot be implemented by the government because of informational problems. This is the focus of the next section.

4. Employment subsidies under asymmetric information

Suppose now that the government is no longer informed about the firm’s type $\theta$. Which employment contracts can it give to the firm in order to maximize its objective? How important will be the induced deadweight transfers and non-additional employment?

The government proposes the following one shot game. It offers a menu $(C_i)_{i \in \{1, 2\}}$ of two contracts $C_i = (L_i, T_i)$ without knowing the true productivity parameter $\theta$. Then nature moves by selecting each firm’s parameter in $(\theta_1, \theta_2)$ with the probabilities $(p, 1-p)$. The private firm is informed about its productivity parameter and moves in last place; it either selects one of the two contracts or chooses not to participate, in which case it prefers to hire $L^F(\theta)$ employees and to get the free profit $P^F(\theta)$. Finally the contracted levels of employment and transfers are realized.

The government maximizes the expected welfare subject to the participation constraints (PC) and the incentive compatibility constraints (IC). The latter reflects that each type of firm is enticed not to mimic the other. Consequently it reveals its true type through the choice of the contract. This problem is an adverse selection problem with the peculiarity of non-constant reservation values $II^F(\theta)$ (see Jullien (1997) for continuous type models with non-constant participation constraints).
This peculiarity will modify the binding incentive compatible constraint. The optimization problem can be written in the following way:

\[
P^{IC}: \max_{(t_i, L_i, T_i, J)} p[F(\theta_i, L_i) - w_iL_i - \lambda T_i] + (1 - p)[F(\theta_h, L_h) - w_hL_h - \lambda T_h]
\]

s.t.

\[
F(\theta_h, L_h) - wL_h + T_h \geq \Pi^f_h, \quad (PC_h)
\]

\[
F(\theta_i, L_i) - wL_i + T_i \geq \Pi^f_i, \quad (PC_i)
\]

\[
F(\theta_h, L_h) - wL_h + T_h \geq F(\theta_i, L_i) - wL_i + T_i, \quad (IC_h)
\]

\[
F(\theta_i, L_i) - wL_i + T_i \geq F(\theta_h, L_h) - wL_h + T_h, \quad (IC_i)
\]

where \(\Pi^f_i = \Pi(\theta_i, L^f(\theta_i))\) and \(\Pi^f_h = \Pi(\theta_h, L^f(\theta_h))\).

In the next sections, we show that there are only three relevant classes of contracts according to the states of the incentive and participation constraints of the most productive firm \(\theta_h\). The following section presents the first class in which no constraint binds and thus the full information contract is obtained.

4.1. Incentive compatible full information contract

In this section we derive the condition that allows the government to offer the full information contracts. Intuitively, this will occur when the government requires a not too large increase in employment. Indeed, note first that there is no incentive issue in the laissez-faire situation because there is no gain for firms to mimic each other. Thus, as long as the firms are induced to rather small increases in employment and as long as they have rather different productivities, there is no scope for mimicking. We will refer to those contracts as class (a) contracts.

We define the pooling contract \((L^p, T^p)\) as the single contract that can be adopted by both types of firms while providing them with the profit obtained in the laissez-faire situation:

\[
T^p = \Pi^f_i - (F(\theta_i, L^p) - wL^p) \quad \forall i = l, h.
\]

Thus,

\[
F(\theta_h, L^p) - \Pi^f_h = F(\theta_l, L^p) - \Pi^f_l.
\]  \quad (6)

The pooling level of employment \(L^p\) determines when the full information contracts are incentive compatible.

**Proposition 1.** The full information contracts \(((L^*_l, T^*_l), (L^*_h, T^*_h))\) are incentive compatible if and only if \(L^*_l \leq L^p\). There is no deadweight spending. All subsidized jobs are additional.

In Fig. 1, the menu of full information contracts \((a_l), (a_h))\) is incentive
compatible. The free profits are represented by the iso-profit curves $\Pi^F_i$ and $\Pi^F_h$.

Larger profits yield higher iso-profit curves. By the Single Crossing Property, the iso-profit curve of the lowest type is drawn at low levels of employment $L$ (to the left), while the highest type’s is drawn at high levels (to the right). By definition, the pooling level of employment $L^p$ is defined by the intersection of the iso-profit curves $\Pi^F_h$ and $\Pi^F_i$. The free profit iso-curves $\Pi^F_i$ and $\Pi^F_h$ determine the participation constraints $PC$ which depict the minimal transfers that make the firms indifferent between participating or not in the contracts. The full information allocation is attained at the tangency point of the iso-profit curves $\Pi^F_i$ and $\Pi^F_h$ and the government’s indifference curve $W$. It is then clear that full information contracts are incentive compatible if $L^*_i \geq L^p$. This implies that, as in the full information case, there is no deadweight spending since no firm can increase its profit by accepting a subsidy. And, all jobs are additional because the government can infer each firm’s true type from the selected contracts. The government will never pay for non-additional jobs.

From this figure, it is clear that the class (a) becomes more likely when the government has weaker employment objectives and when the productivity of the least productive firm is worsened. In this last situation, the gain that the most productive firm may obtain by mimicking the other is reduced. As the incentive issue is weaker, the likelihood of class (a) must increase. More formally, a comparative statics exercise on $(L^*_i - L^p)$ leads to the following corollary.

**Corollary 1.** The range of parameters for class (a) contracts is larger when the distortion in the labor market decreases, when the shadow cost of public funds increases and when the difference between firms productivity increases.

In this class of contracts, the liberal participation constraint relaxes the incentive constraint. Suppose that the least productive firm makes no profit and that the government is allowed to tax firms instead of guaranteeing the laissez-faire profit.
The government would then impose a lump sum tax on the most productive firm. The iso-profit curve $II^*_i$ would move downward and the condition $L^*_i \leq L^P$ may not be satisfied anymore.

The class (a) of incentive contracts allows the socially efficient level of employment in all types of firms. As presented in the next section, this may not be the case if the condition $L^*_i \leq L^P$ is not satisfied.

### 4.2. Information rent as the source of deadweight spending

When the condition for the full information contracts is not satisfied, we show that the least productive is induced to socially inefficient employment levels and gets no rent, whereas the most productive firm is characterized by a socially efficient employment level and gets some information rent. In contrast with the traditional literature on adverse selection models, the information rent may be null for all types of firms. We firstly state the following lemma.

**Lemma 1.** If $L^*_i > L^P$ then (a) levels of employment are monotonically ranked ($L_i \leq L_h$), (b) $PC_i$ is binding, (c) $IC_h$ is binding, (d) $IC_i$ is slack and (e) $PC_h$ may be slack or binding.

The specificity of Lemma 1 derives from the existence of non-constant reservation profits $\Pi^F_i$ and $\Pi^F_h$ in the participation constraints.

It must first be noted that, by part (a) of Lemma 1, each type of firm selects a different contract, so that the optimal employment subsidies are *separating* contracts. The government will always be able to infer the true types of firms from the selected contracts. It will thus be able to avoid paying for any non-additional jobs. *Job additionality is therefore irrelevant when optimal incentive contracts are offered by the government to firms.*

By Lemma 1, the relevant constraints are $(PC_i)$, $(IC_h)$ and $(PC_h)$, which can be transformed to

\[
F(\theta, L_i) - wL_i + T_i = \Pi^F_i, \quad (PC_i)
\]

\[
F(\theta_h, L_h) - wL_h + T_h = \Pi^F_h + \Phi(L_h), \quad (IC_h)
\]

\[
\Phi(L) = 0 \quad [\lambda \mu] \quad (PC_h)
\]

where $\mu$ is the Kuhn–Tucker multiplier of $(PC_h)$ and where $\Phi(L)$ is the information rent that the government must give to the most productive firms to avoid that the latter mimics the least productive types. This rent is

\[
\Phi(L) = [F(\theta, L_i) - wL_i - \Pi^P_i] - [F(\theta_h, L_h) - \Pi^F_i] - [F(\theta, L_i) - \Pi^F_i + \Phi(L_h)]
\]

\[
= [F(\theta, L_i) - F(\theta_h, L_h)] - [\Pi^P_i - \Pi^F_i]. \quad (7)
\]
As the first expression shows, the information rent corresponds to the difference in each firm’s ‘loss’, \( F(\theta, L) - wL \), that would occur if the government had imposed the employment level \( L \) to each private firm. By the Single Crossing Property, we know that the least productive firm has smaller returns to employment increase. Therefore, any rise in employment level \( L \) increases the ‘loss’ more in the least productive firm so that the information rent must increase with \( L \), too.

More formally, by applying (6) and by differentiating (7), we obtain

\[
\Phi(L^P) = 0 \quad \text{and} \quad \Phi'(L^P) > 0. \tag{8}
\]

The information rent is positive and increases as the employment level rises above \( L^P \).

On the other hand, transfers can be written as

\[
T_i = \Pi^F_i - [F(\theta, L)] - wL,
\]

\[
T_h = \Pi^F_h - [F(\theta, L) - wL] + \Phi(L_i).
\]

From the above equations and from expressions (PC), (IC), (PC) and (7), we can observe on the one hand that the lowest productive firm gets the free profit and receives a transfer that exactly compensates the loss in profit due to the contracted level of employment \( L^* \), as was the case with full information. This firm is indifferent between participating or not in the contract: there is thus no deadweight transfer to this type of firm. On the other hand, the highest type may get more profit than in the laissez-faire situation as it may receive an information rent. There may be a deadweight transfer to this type of firm.

4.3. The condition for deadweight spending

The previous section has shown how the information rent can be the source of deadweight transfer. The current section shows that it can still be null in particular situations. To demonstrate this, we develop the formal solution of the program \( F^{ic} \) when \( L^* > L^P \).

Using (PC), the Lagrangian of the problem can be written in the following form:

\[
\mathcal{L} = p[F(\theta, L) - wL] + (1 - p)[F(\theta, L) - wL] \\
- p\lambda[\Pi^F_i - (F(\theta, L) - wL)] - (1 - p)\lambda[\Pi^F_h - (F(\theta, L) - wL)] \\
+ (\mu - (1 - p))\lambda\Pi(L^*).
\]

The first order conditions are obtained by differentiating \( \mathcal{L} \) by \( L_i \) and \( L_h \). The optimal levels of employment \( L_i^{**} \) and \( L_h^{**} \) are given by

\[
F_i(\theta, L_i^{**}) = \frac{w + \lambda w}{1 + \lambda}; \tag{9}
\]

\[
F_h(\theta, L_h^{**}) = \frac{w + \lambda w}{1 + \lambda};
\]
We have now two classes of contracts depending on whether the Kuhn–Tucker multiplier \( \mu \) is positive or null.

**Class (b):** in this class of contracts, \( \mu > 0 \) and \((PC_n)\) is binding. Thus, by (7),

\[
\Phi(L^{**}_f) = 0 \Leftrightarrow \Pi^{**}_n = \Pi^L_n.
\]

By (8), \( L^{**}_f = L^P \). There is no point for the government to diminish \( L^{**}_f \) below the pooling level of employment \( L^P \). The first order condition (9) holds and is identical to Eq. (5) in the full information setting. The optimal level of employment is then \( L^{**}_n = L^*_n \).

As both participation constraints bind, no firm earns more than its free profit \( \Pi^F \), there is no information rent and no deadweight transfer. The transfer for the most productive firm is the full information transfer \( T^*_n \) and, by (9), the transfer for the least productive firm is \( T^P \). These considerations lead us to the following proposition.

**Proposition 2.** In class (b) incentive contracts, the level of employment in the most productive firm is the full information level \( (L^{**}_n = L^*_n) \). The level of employment in the least productive firm is distorted downward and equal to the pooling level of employment \( (L^{**}_n = L^P) \). The transfers do not include any deadweight spending.

**Class (c):** in this class of contracts, \( \mu = 0 \) so that \((PC_n)\) is slack and \( \Phi(L^{**}_n) > 0 \). Since \( \Phi(L) \) increases in \( L \) and \( \Phi(L^P) = 0 \), then

\[
L^{**}_n > L^P.
\]

The optimal level of employment is always larger than the pooling level of employment in class (c) contracts.

The first order condition (9) for the highest type of firm does not depend on \( \mu \). The optimal level of employment \( L^{**}_n \) is thus \( L^*_n \). On the other hand, the condition (10) determines the lowest type’s optimal level of employment \( L^{**}_f \). It must satisfy

\[
F_L(\theta_i, L^{**}_f) = \frac{w_1 + \lambda w}{1 + \lambda} + \frac{(1 - p)}{p} \frac{\lambda}{1 + \lambda} \Phi'(L^{**}_f).
\]

This equation expresses the government’s trade-off between the social benefit of production (left hand side) and its social cost (right hand side). The social cost includes not only the social wage \((w_1 + \lambda w)/(1 + \lambda)\), as in the full information context, but also the marginal cost of information extraction that is presented by the term in \( \Phi'(L^{**}_f) \). Any employment increase in the least productive firm indeed strengthens the incentive for the most productive firm to mimic the former, which
involves larger information rents. In order to reduce the information rents, the government will reduce its employment requirement in the least productive firms.

**Proposition 3.** In class (c) incentive contracts, the most productive firm achieves the full information level of employment \( L_n^* \), while the least productive one achieves a level of employment lower than the full information level \( L_i^* \) \((L_r^* < L_i^* < L_n^*)\). This employment level increases with lower \( w_i \), with lower \( \lambda \), with higher \( \theta_p \), and with higher \( p \). There exists some deadweight spending.

**Proof.** Since \( \Phi'(L_i^*) > 0 \), \( L_i^* < L_n^* \) by (12). Comparative statics of \( L_i^* \) are obtained by differentiating the above condition with respect to the parameter \( w_i \), \( \lambda \), \( \theta_p \) and \( p \).

Lower distortion in the labor market and larger shadow cost of public funds (larger \( w_i \) and \( \lambda \)) imply smaller interventions from the government. Employment in the least productive firm is therefore less supported. A higher parameter \( \theta_p \) for the most productive firm will make this firm more attractive to the government. The government will desire to reduce further the employment in the least productive firm in order to diminish the social cost of the information rent. Finally, the level of employment \( L_i^* \) decreases when the probability \( p \) of the least productive firm decreases. The smaller the probability of this firm, the smaller weight the government puts on it and the more inclined is the government to lose some employment — and consequently some welfare — in this type of firm.

The three classes of contracts are displayed in Fig. 2. The full information levels \(((a_1), (a_n))\) are reached where the slopes of the iso-profit curves \((II)\) are small. This indicates a small marginal welfare from employment, which characterizes a weak distortion in the labor market or a large shadow cost of public funds. (The marginal social surplus corresponds to the slope of the iso-welfare which is not depicted, but which is tangent at \((a_i)\) and \((a_n)\).)

---

![Fig. 2. Three classes of employment contracts.](image-url)
The contracts \((b_1), (b_2)\) and \((c_1), (c_2)\) are reached for larger slopes of iso-profit curves and thus larger marginal welfare from employment. The difference between the two classes of contract lies in the probability that the government puts on the least productive firm. In class (c), the government puts a large probability on this firm. It faces the trade-off between increasing the employment level of this firm and reducing the information rent to the most productive firm. In class (b), the government attaches a smaller probability to this firm. It is then willing to reduce more employment in the least productive firm in order to decrease the cost of information rent accruing to the other highly probable firm. It decreases this employment down to the level \(L^*\); any decrease below this level is inefficient since the information rent would no longer be reduced. In fact, as in Section 4.1, the liberal participation constraint relaxes the incentive constraint and facilitates the screening of firms.

**Example.** The impact of the distortion in the labor market and in the public funding on the employment contracts may be clarified in the following example. Let \(s = (\theta_k - \theta)/\theta\) be the productivity spread. Suppose that the private firms have the production function \(F = \theta \sqrt{L}\) where \(\theta\) is a productivity parameter. Then, the full information class (a) occurs when \(L^p > L^*_p\). After computation this leads to the following inequality

\[
\frac{w}{w_s} \leq \left(\frac{\tilde{w}}{w_s}\right) \quad \text{where} \quad \left(\frac{\tilde{w}}{w_s}\right) = \frac{1 + \frac{1}{2}s}{1 - \frac{1}{2} As} \geq 1.
\]

Class (b) occurs when \(L^p < L^*_p\). This condition is transformed to the following expression:

\[
p < \tilde{p} = \frac{w}{w_s} \cdot \frac{\lambda s}{1 + \frac{1}{2}s}.
\]

The ratio \(\tilde{w}/w_s\) represents the intensity of the distortion in the labor market below which full information contracts are achievable (class (a)). The probability \(\tilde{p}\) is the government’s maximum belief on the least productive firm that allows an incentive contract with no deadweight spending (class (b)). This probability increases with larger \(\lambda\) and \(s\) but with lower \(w/w_s\).

Fig. 3 depicts the three classes for various beliefs about the least productive firms \(p\) and for various intensities of distortion in the labor market \(w/w_s\). A larger cost of public funds or a larger spread will indeed imply a larger difference between the optimal levels of employment in the two types of firm. The government will be more enticed to reduce the employment in the least productive firm down to \(L^*\), which will make class (b) contracts more likely. Finally a lower distortion in the labor market (higher \(w_s\)) means that the government less
intensively desires to promote employment, in particular within the least productive firm. The optimal level of employment in this firm may become sufficiently low so that the most productive firm is not induced to mimic the least productive one. There will be no information rent and thus no deadweight spending. This leads again to class (b) contracts.

In this model, the nature of the information set matters. This issue has been addressed in an earlier version of the present paper (available on request to the author). Some intuition can nevertheless be obtained when $s \to 0$. The current model with a discrete distribution of types tends to a model with a continuous distribution of types. In this particular example, the above formulae show that $(\hat{w}/w_s) \to 1$ and $\hat{p} \to 0$, the classes (a) and (b) may disappear: the full information contracts are no longer incentive compatible and deadweight spending occurs for all types of firms.³

5. Dynamic aspects

In the previous sections, optimal employment subsidies have been designed in a static framework. In a dynamic framework, it is tempting to base the subsidy scheme on past employment levels as in the Marginal Employment Subsidy

³Under regular assumptions, it is shown that, in contrast to the current model, the optimal employment subsidies may not be granted to firms with low productivity. The assumption of discrete or continuous distribution of types matters. Whereas continuous information sets include an infinity of elements, the discrete information set that is used in the current paper includes only a finite and small number of elements. The discrete information set is more informative which eases information extraction. The characteristics of the information set is nevertheless driven by the actual industry characteristics to which employment subsidies should be offered.
(MES) (Layard and Nickell, 1980) that provides a subsidy only to increases in employment. However, Laffont and Tirole (1988) have shown that the re-use of revealed information entices firms to under-report information in order to increase their profit through larger ex-post transfers. This ratchet effect also takes place in the case of employment subsidies, as firms will reduce their ex-ante levels of employment in the first period in order to pretend for larger employment efforts and thus larger subsidies in the next periods. Private firms may even shut down and restart under a new name.

The ratchet effect originates from the absence of government commitment not to re-use information. As soon as it has extracted information from a private firm in the first period of time, the government is able to eliminate any information rents in next periods. This problem is particularly relevant for capital intensive firms that contemplate hires associated to capital investments, because the employment subsidies should be granted over a long period of time.

Baron and Besanko (1984) provide an important argument in favor of the optimal employment subsidies such as developed in the present paper. In their analysis of a repeated adverse selection model, they show that, if information does not vary over time, it is suboptimal for the government to re-use the revealed information over periods. As a result, the information rent and thus the deadweight spending must be borne in each period. Information re-use, such as designed in the MES, is not efficient because of the anticipation problems it generates.

6. Labor supply policies

The present paper focuses on employment policies on the demand side of the labor market. We have considered the firms' heterogeneity and informational advantage. This analysis brings a viewpoint that is complementary to the more documented issue of policies on the supply side of the labor market, in which the implication of worker heterogeneity is analyzed (see for instance Phelps (1994), Snower (1994), Drèze and Sneessens (1997), etc.). In this section, we relate further our paper with this existing literature on employment subsidies to workers, and more specifically to the theory of optimal income taxation.

On the empirical side, we mentioned in the Introduction a part of the literature on the efficiency of employment subsidies to private firms. The focus was on the direct measure of deadweight spending, job additionality and cost per (additional) job. At the same time, many empirical papers have been devoted to the impact of income subsidies to workers, particularly in connection to the US Earned Income Tax Credit (Eissa and Liebman, 1995; Katz, 1996). The effectiveness of employment policies is also often measured by re-employment probabilities and by labor supply elasticities. Felli and Ichino (1988) use an Italian subsidy program as a quasi-experiment to evaluate the increase of re-employment probabilities due to subsidies to workers. Lindsay (1987), Card and Robins (1996) and Piketty (1998)
use quasi-experiments to assess labor supply elasticities of particular groups of workers and to infer the impact of employment subsidies or income tax reforms.

On the theoretical side, our model must be related to the literature on optimal income taxation (Mirrlees, 1971) as subsidies to workers can be seen as negative income taxes. In this literature, the information asymmetry lies between the government and heterogeneous workers (or households) about their individual (work) productivity. The government is nevertheless able to observe and verify each worker’s earning. Piketty (1997) uses such a framework to analyze the employment impact of income tax reforms in France. Whereas this complementary approach shares strong similarities with our analysis, it differs in several respects. We mention only three points.

Firstly, the redistribution issue is much less relevant in our approach. In the theory of optimal income taxation, the government reflects the society’s redistribution concern with respect to individuals who hold private information. The interaction between redistribution and incentives has an important consequence for the optimal taxes and subsidies (see Mirrlees (1971), see also Diamond (1998) for recent developments). In contrast, there is no redistribution issue across the information holders of our model — namely, the private firms (profits are usually supposed to be uniformly distributed among capitalists).

Secondly, in the income taxation literature, the choice of the control variables on individuals is restricted by horizontal equity. The design of income tax based on age, sex, race is considered not ethical. By contrast, neither an ethical nor economic reason imposes horizontal equity in the choice of the control variables on private firms; as mentioned in the Introduction, the choice of employment as the control variable is dictated by a measurability and enforceability argument.

Finally, as stressed in this paper, the government has weaker power in our model involving private firms. The government faces a liberal constraint by which private firms must voluntarily participate in the subsidy scheme. In the literature of optimal income taxation, the income earners usually have no choice over participation in the scheme; since they cannot escape from taxation of their residence country, they are left with the only choice between work time and leisure (or, also, between effort and no effort). If income earners were perfectly mobile across countries, then their participation levels would rise and income taxation would incorporate a participation constraint similar to our liberal constraint.

7. Conclusion

In this paper we have studied the design of optimal employment subsidies to private firms that operate in perfectly competitive product markets and in labor

---

Picard (1998b) provides an application with non-linear employment subsidies to workers.
markets characterized by unemployment. When the government uses such optimal employment subsidies, the deadweight spending is equivalent to the information rent that private firms are capable of extracting because of private knowledge about their hiring capabilities. In the present liberal regulation framework, the private firms participate in the subsidy scheme if and only if they increase their profit compared to the laissez-faire situation. But, since the binding participation constraint is congruent to the incentive constraint, it fosters the screening of firms and the deadweight spending is reduced and sometimes eliminated. Also, this paper showed that optimal employment subsidies are separating incentive contracts. The government infers the firms’ private information from their selection of employment contracts. As it can therefore determine the additional nature of each job, it avoids paying for non-additional jobs and all subsidized jobs must be additional.

Obviously, the issue of job additionality becomes relevant only when the employment subsidies are restricted to some specific forms such as the MES, or when the combination of firms behavior and incentive compatibility forbids the implementation of separating contracts. Separating contracts may not exist for profit functions that do not satisfy the Single Crossing Property, which is excluded in the present paper for simplicity. It may also take place when the firms’ characteristics are multi-dimensional.\(^5\) The issue of non-additional employment would thus be at stake in models that consider the various dimensions of the information asymmetry between firms and government.

**Acknowledgements**

This research has been supported by a grant ‘Actions de Recherche Concertées’ n° 93/98-162 of the Ministry of Scientific Research of the Belgian French Speaking Community. I am also grateful to B. Cockx, M. Dewatripont, L. Eeckhoudt, B. Jullien, P. Madden, M. Marchand, V. Songwe and a referee for their helpful comments and support.

**Appendix A**

**Proof of Proposition 1.** It must be shown that the full information contract solves the program \(P^{IC}\). At the full information contract, the constraint \(IC_h\) requires that
\[
F(\theta_h, L^*_h) - wL^*_h + T^*_h \geq F(\theta_h, L^*_h) - wL^*_h + T^*_h.
\]
Using the expression (4) for the transfers \(T^*_h\) and \(T^*_h\), we get
\[
\Pi^*_h = \Pi^*_h + F(\theta_h, L^*_h) - F(\theta_h, L^*_h).
\]
By the definition of \(L^*_h\), this can be transformed in
\[
0 \geq (F(\theta_h, L^*_h) - F(\theta_h, L^*_h) - F(\theta_h, L^*_h)) - (F(\theta_h, L^*_h) - F(\theta_h, L^*_h)).
\]

\(^5\)The recent developments in multidimensional screening (Rochet and Choné, 1998) have shown that bunching in one dimension is a natural property of multidimensional incentive contracts.
Proof. By the Single Crossing Property this is satisfied if and only if \( L^*_h \leq L^f \). This is the condition in the proposition.

At the full information allocation, the constraint IC, is \( F(\theta, L^*_h) - wL^*_h + T^*_h \geq F(\theta, L^*_h) - wL^*_h + T^*_h \). Using the same argument, we can reduce it to \( 0 \geq (F(\theta, L^*_h) - F(\theta, L^*_h) - (F(\theta, L^f) - F(\theta, L^f)) \). Again, by the Single Crossing Property this is satisfied if and only if \( L^f \leq L^*_h \). But this is always true because, by construction, \( L^f \) takes a value in the range \([L^1_h, L^2_h]\) which implies that \( L^f \leq L^*_h \).

**Proof of Lemma 1.** Remember that the revenue is \( PF(\theta, L) \) with \( P = 1 \). By the definition of \( L^f \), the inequality \( L^f \geq L^f \) is equivalent to \([F(\theta, L) - \Pi^f_h - [F(\theta, L) - \Pi^f_h] \leq 0 \), which contradicts the Single Crossing Property \( F_{\theta h} \geq 0 \).

**Lemma A.1.** Assume \( \theta_h > \theta \) and \( F_{\theta h} > 0 \). Levels of employment are monotonic:
(a) \( L^*_h \leq L^*_h \). If \( \Phi(L^*_h) > 0 \) then (b) \( PC \) binds, (c) \( IC_h \) binds and (d) \( IC \) does not bind. Only the slackness of constraint \( PC_b \) is undetermined.

**Proof.**

(a) \( L^*_h \leq L^*_h \); we prove that \( L^*_h > L^*_h \) is a contradiction. Suppose \( L^*_h > L^*_h \). If \( IC \) and \( IC_h \) could be satisfied, then the sum of \( IC \) and \( IC_h \) should be satisfied, too: \( (F(\theta, L^*_h) - F(\theta, L^*_h)) - (F(\theta, L^*_h) - F(\theta, L^*_h)) \leq 0 \), which contradicts the Single Crossing Property \( F_{\theta h} \geq 0 \).

(b) If \( \Phi(L^*_h) > 0 \), then \( PC \) binds: we prove here that, if \( \Phi(L^*_h) \geq 0 \) and if \( PC \) is not binding, then \( PC \) is not binding and the transfers \( T^*_h \) and \( T^*_h \) are not optimal for the government. Suppose \( PC \) is not binding at the optimum, i.e. \( F(\theta, L^*_h) - wL^*_h + T^*_h \geq \Pi^f_h \). Add the inequality \( \Phi(L^*_h) > 0 \) from hypothesis and get \( F(\theta, L^*_h) - wL^*_h + T^*_h \geq \Pi^f_h \). Using \( IC_h \), \( F(\theta, L^*_h) - wL^*_h + T^*_h \geq F(\theta, L^*_h) - wL^*_h + T^*_h \geq \Pi^f_h \). The first and last term of this chain of inequalities correspond to the constraint \( PC_b \). So, none of the participation constraints are binding. It is possible to gain some welfare by reducing \( T^*_h \) and \( T^*_h \) without modifying the levels of employment \( L^*_h \) and \( L^*_h \).

(c) If \( \Phi(L^*_h) > 0 \), then \( IC_h \) binds: we prove that, if \( \Phi(L^*_h) > 0 \) then, if \( IC_h \) does not bind then \( PC_b \) does not bind and \( T^*_h \) is not an optimal transfer. If \( IC_h \) is not binding, then \( F(\theta, L^*_h) - wL^*_h + T^*_h > F(\theta, L^*_h) - wL^*_h + T^*_h \). Add the inequality \( \Phi(L^*_h) > 0 \), then \( F(\theta, L^*_h) - wL^*_h + T^*_h + \Phi(L^*_h) > F(\theta, L^*_h) - wL^*_h + T^*_h \). This yields \( F(\theta, L^*_h) - wL^*_h + T^*_h > \Pi^f_h \). Since the parenthesis in the middle term corresponds to the \( PC \) constraint and is positive or zero. So, we have that \( PC_b \) and \( IC_h \) are not binding. Ceteris paribus, the government can therefore reduce \( T^*_h \) and gain some welfare: the transfer \( T^*_h \) is not optimal.

(d) If \( \Phi(L^*_h) > 0 \), then \( IC \) does not bind: we prove here that, if \( IC \) binds, then
by the Single Crossing Property $F_{L_0} > 0$, IC cannot bind, which contradicts the previous paragraph. Suppose IC$_j$ is binding, then $F(\theta_i, L_j^{**}) = wL_j^{**} + T_j^{**}$, By $F_{L_0} > 0$, $F(\theta_j, L_j^{**}) - F(\theta_i, L_i^{**}) < F(\theta_j, L_j^{**}) - F(\theta_i, L_i^{**})$. Add both inequalities and get $F(\theta_i, L_i^{**}) < F(\theta_j, L_j^{**}) - wL_j^{**} + T_j^{**}$, that is IC$_j$ is not satisfied, which is a contradiction. 

**Proof of Corollary 1.** Note first that, by the Single Crossing Property, the pooling level of employment $L^p$ is positively related to the $\theta_i$ and $\theta_j$. It depends also on $w$ and $L^*$.

Then, the occurrence of class (a) depends on whether the condition $L^*(w, w', \lambda; \theta) \leq L^p(w; \theta_i, \theta_j)$ is satisfied. On the one hand, we noted in Section 3 that $L^*(w, w', \lambda; \theta)$ decreased as $\lambda$ and $w'$ increased. Therefore the condition is satisfied for more parameter values. On the other hand, larger decreases in the returns to scale of the lowest type implies lower marginal product and then lower levels of employment $L^*$ at full information. This means that the lowest type has a more convex curve. The likelihood of $L^p < L^*$ increases too.

**References**


The reader may wonder what happens when $\Phi(L_i^{**}) < 0$. In this situation, the constraints PC and PC$_j$ are binding, but none of the constraints IC$_i$ and IC$_j$ are binding. The level of employment $L_i^{**}$ is not optimal since the government can increase $L_i^{**}$ without modifying the states of the constraints. It can increase it up to the level $L^*$ where $\Phi(L^*) = 0$. In fact, there is no point in decreasing $L_i^{**}$ below $L^*$.