Volume Title: Taxation in the Global Economy

Volume Author/Editor: Assaf Razin and Joel Slemrod, editors

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-70591-9

Volume URL: http://www.nber.org/books/razi90-1

Conference Date: February 23-25, 1989

Publication Date: January 1990

Chapter Title: The Optimal Taxation of Internationally Mobile Capital in an Efficiency Wage Model

Chapter Author: John D. Wilson

Chapter URL: http://www.nber.org/chapters/c7216

Chapter pages in book: (p. 397 - 432)

11 The Optimal Taxation of Internationally Mobile Capital in an Efficiency Wage Model

John Douglas Wilson

The main purpose of this paper is to investigate the optimal system of taxes and subsidies on capital for an open economy in which similar types of workers are paid different wages. Phrased in popular terminology, the question is, What role do capital taxes and subsidies play in the optimal "industrial strategy" for an economy with "good jobs" and "bad jobs," as distinguished by wage levels? Since the answer is found to depend on the availability of other tax instruments, the paper also investigates the optimal choice of these other instruments. Briefly stated, the case for subsidizing capital investment in "good jobs" appears rather dubious. In fact, a model is presented in which informational asymmetries between the government and private firms justify a *positive* marginal tax on capital investment in the high-wage sector.

The basic reason for wage differentials in this paper is that worker productivity and wages are positively related in some firms but not others. This relation is a special case of the general phenomenon of "dependence of quality on price," which has received substantial attention in recent years, not only in labor markets, but also in credit and product markets. Stiglitz (1987a) provides an extensive review of this literature. For the special case of labor markets, "efficiency wage theories" are reviewed by Stiglitz (1986), Carmichael (1988), and Katz (1988). The main explanations that have been given for the dependence of worker productivity on wage levels include worker supervision problems, labor turnover, morale effects, and

John Douglas Wilson is associate professor of economics at Indiana University.

Some of the research reported here was completed during a visit to Queen's University in Canada, under the auspices of the John Deutsch Institute. The author is grateful to Ben Craig, Patricia Wilson, seminar participants at Purdue University and the University of Toronto, and participants at the NBER Conference on International Aspects of Taxation for helpful comments.

nutritional concerns. In their seminal paper, Shapiro and Stiglitz (1984) use the supervision approach to model long-run unemployment as an efficiency wage problem.

Of special relevance here is the work by Bulow and Summers (1986) on industrial policies. They extend the Shapiro-Stiglitz analysis to include two production sectors, one with a supervision problem and one without. Their analysis shows that high-wage firms should receive a production subsidy. Arvan and Schoumaker (1988) dispute the generality of this result by adding a fixed supply of capital to the model and demonstrating that the optimal commercial policy depends on the relative labor intensities of the two sectors. While neither paper analyzes capital tax policies, Bulow and Summers do conjecture that "keeping capital at home, and in the primary sector, may raise welfare by increasing rents created by primary sector jobs" (1986, 397).

I follow Bulow-Summers and others by assuming that the payment of wages above workers' opportunity costs serves as a worker discipline device: high wages make employment termination a genuine punishment for "shirking" on the job. But my model departs from the previous literature in two significant ways. First, I drop the assumption that utilities are linear in income. In the Bulow-Summers paper, a first-best optimum is obtainable through the use of employment subsidies because total economic welfare depends on national income, not on how it is distributed.[1] In the present paper, however, a first-best optimum is not obtainable, even when the government possesses the same information as private firms, because employment subsidies lead to increased efficiency at the cost of a less equitable income distribution. The cause of income distribution problems here is *not* that workers possess innate differences in preferences or endowments. Rather, distributional issues arise because the only way to deal with worker supervision problems is to provide similar workers in different industries with different incomes. This framework allows me to investigate whether the inherently second-best nature of the problem leads to desirable forms of capital market intervention.

The paper's other distinguishing feature is that consideration is given not only to the traditional case where the government knows the relevant characteristics of each firm but also to a case of "asymmetric information." In particular, the government is assumed not to be certain about the identity of those firms with supervision problems. Rather, it assigns a probability to the possibility that a given firm possesses a supervision problem. This specification makes no presumption about the severity of the information problem; nearly complete information could be obtained as a special case of the model where each firm is assigned a probability close to either zero or one. Note finally that the informational asymmetry does not prevent the government from making the employment and capital subsidies that it provides to a given firm depend on the firm's chosen wage. However, the rule for doing so cannot depend on whether the firm has a supervision problem.

To enrich the economic environment along these lines, I work with a full employment model. This allows me to capture in a reasonably simple manner the distinction between "good jobs" and "bad jobs," which has occupied much of the industrial policy debate. In particular, I make use of Calvo and Wellisz's (1978) insightful way of modeling worker supervision problems within a static framework, the main difference being my assumption that workers caught shirking in the "primary sector" obtain an endogenously determined utility by accepting perfectly supervised work in the "secondary sector" rather than becoming "self-employed."

The paper's organization and main results are summarized as follows. In the next section, I present a two-sector model with both international commodity trade and capital mobility. To eliminate obvious market power reasons for capital market intervention, the economy is assumed to be a price taker on world capital and product markets. Section 11.2 investigates the symmetric information case under the assumption that the government is able to make complete use of its information about private firms without being thwarted by limitations on available tax instruments. Here, the case for capital market intervention disappears: each firm should be allowed to equate the value of the marginal product of capital with the interest rate investors can obtain abroad, as it would in the absence of domestic capital taxes.[2] However, worker supervision problems do create a justification for employment and wage subsidies; and Appendix A demonstrates the desirability of excise taxes. In other words, workers should trade at product prices that differ from world prices, and some firms should face tax incentives to increase both the numbers of workers they employ and the wages they pay them. If these other tax instruments are not available, then positive subsidies on capital investment in the high-wage sector may be warranted. Using a simplified version of the model, Appendix B demonstrates the desirability of such subsidies. But the unavailability of other tax instruments is difficult to justify.

Section 11.3 investigates the asymmetric information case described above. Here, capital taxes and subsidies emerge as a desirable tax instrument, but with a rather surprising property: high-wage firms should face a positive tax on capital at the margin, while low-wage firms should face a positive subsidy. The basic reason for this result is that these capital taxes and subsidies lessen the severity of the adverse selection problem in the model. They allow, for example, the government to further raise its subsidization of high-wage employment relative to low-wage employment without causing those firms without worker supervision problems also to raise their wages so as to obtain the employment subsidies. The result is made more understandable by remembering that the low-wage sector is inherently more efficient than the high-wage sector in the sense that it lacks a supervision problem. Simply stated, the optimal tax policy encourages capital investment in the sector with the relatively efficient production process while discouraging capital investment elsewhere. Taken as a whole, the results of this paper call into question the desirability of encouraging capital investment in high-wage firms. Section 11.4 discusses some possible extensions of the analysis.

11.1 The Basic Model

I consider a simple two-sector model of a small open economy. The two goods produced in the economy are perfectly tradeable internationally at exogenously determined world prices. Each good is produced from labor and capital. The economy's total supply of workers is fixed, but each worker's labor effort is variable, making the economy's "effective labor" supply variable. The economy faces an infinitely elastic supply of capital at an exogenously given world interest rate, r (net of taxes levied abroad). In the following subsections, I describe the individual components of the model.

11.1.1 Production and Trade

The economy contains a "primary sector" (x) and a "secondary sector" (y). Each sector is assumed to behave competitively, the meaning of which is fully specified below. For notational simplicity, the economy is modeled as though each sector contains a single firm, but the analysis clearly applies to a model where there are any fixed number of firms in either sector.[3] Each firm possesses a production technology described by a strictly concave production function. Thus, there are decreasing returns to scale. This function is denoted $f^x(E_x, K_x)$ for the primary sector and $f^y(E_y, K_y)$ for the secondary sector, where K_i and E_i denote the capital and "effective labor" used in sector i (superscripts identify functions, and subscripts of functions denote partial derivatives throughout this paper).[4] My main reason for not assuming constant returns to scale is to allow capital to be perfectly mobile internationally without causing the economy to completely specialize in the production of a single good. Incomplete specialization is assumed throughout the paper, for both the laissez-faire equilibria and the social optima. The interpretation of the decreasing returns assumption is that there is a third factor, say, "entrepreneurial talent," that is omitted from the production function as an explicit argument.

The economy's trade balance constraint may be written

$$(1) \qquad p_x f^x(E_x, K_x) + p_y f^y(E_y, K_y) + r(K^* - K_x - K_y) \geq p_x C_x + p_y C_y,$$

where

C_x = total consumption of the primary good;
C_y = total consumption of the secondary good;

K^* = total ownership of capital by domestic residents;
r = world interest rate;
p_i = world price of sector i output.

This constraint states that the total value of domestic output, plus the value of capital exports, must be at least as great as the total value of domestic consumption, calculated at world prices. It is always satisfied with equality throughout the paper. The assumption that the domestic economy is small means that both r and $p = (p_x, p_y)$ are exogenously fixed from its viewpoint.

11.1.2 Workers and Supervision

I now specify the worker supervision problem, which lies at the heart of the model. To isolate efficiency considerations, all individuals are assumed to be ex ante identical. In particular, they possess identical utility functions, identical labor and capital endowments, and identical ownership shares of domestic profits. The common utility function is denoted $u(c_x, c_y, e)$, where the individual consumption levels, c_x and c_y, contribute positively to utility, and e measures "labor effort," which contributes negatively.[5] If sector i contains N_i workers who each supply labor effort e_i, then its effective labor is $E_i = e_i N_i$.

At the start of the period, the primary-sector firm chooses its desired number of workers. Each chosen worker then makes an irrevocable decision whether to work or to shirk. If he works, he provides the level of labor effort specified by the firm (e_x). In contrast, a shirking worker provides no labor effort and faces a probability $\pi_x < 1$ of being detected. A detected shirker is discharged from the firm and obtains employment in the secondary sector, where supervision is assumed to be perfect. There, he supplies the level of labor effort specified by the firm (e_y) in return for the common wage received by all secondary-sector workers. Undetected shirkers remain in the primary sector and receive the wage given to nonshirkers.

After shirkers have been identified and reassigned jobs, all sector i workers choose their consumption levels to solve the following utility-maximization problem:

$$\max u(c_x, c_y, e)$$

subject to

$$q_x c_x + q_y c_y = I_i , \tag{2}$$

where $q = (q_x, q_y)$ is the vector of consumer prices ($q = p$ in the absence of excise taxes), e equals e_i for nonshirkers and zero for shirkers, and I_i denotes the total income received by a worker who ends up in sector i after all shirkers have been identified. The income variable I_i satisfies

$$I_i = n + w_i , \tag{3}$$

where n is the nonlabor income each worker obtains from capital and profits, and w_i is the wage paid to sector i workers.

This utility-maximization problem yields demand functions and an indirect utility function. Excise taxes are eliminated from the analysis until Appendix A because they do not affect the propositions about capital taxation and because their role as an "antishirking device," while theoretically interesting, seems to be of little practical importance. Consumer prices are then fixed at p and can be omitted as explicit arguments in the demand and utility functions. These functions are denoted $c^j(I, e)$ for good j demand and $v(I, e)$ for utility.

To prevent shirking in the primary sector, the utility of nonshirkers must be set at least as high as the expected utility of shirkers. To write this condition mathematically, note first that the assumption of perfect competition means that each firm treats the *utilities* obtained by workers in other firms as exogenously fixed. Thus, the secondary-sector firm chooses w_y and e_y subject to the constraint

$$v(n + w_y, e_y) \geq \tilde{u}, \tag{4}$$

where $\tilde{u}$ is the utility level at which the firm faces an infinitely elastic supply of workers. Profit maximization obviously requires that (4) hold with equality. The primary-sector firm must then choose w_x and e_x to satisfy the following "no-shirking condition":

$$v(n + w_x, e_x) \geq \pi_x \tilde{u} + (1 - \pi_x) v(n + w_x, 0). \tag{5}$$

This condition also holds with equality under profit maximization (indifference between shirking and not shirking is always resolved in favor of not shirking). Since all primary-sector workers are ex ante identical, either all of them shirk or none of them shirk. In equilibrium, none shirk.

A crucial implication of (5) is that primary-sector workers obtain a higher utility level than secondary-sector workers in equilibrium. While the primary-sector firm can always pay its workers the wage they would get in the secondary sector, it must then require them to provide less labor effort. Mathematically, if the function $e^i(n + w_i, \tilde{u})$ relates firm i's chosen effort level to worker incomes and the secondary-sector utility, then

$$e^x(n + w_x, \tilde{u}) < e^y(n + w_y, \tilde{u}) \tag{6}$$

whenever $w_x = w_y$. This property of the "labor effort functions" is used repeatedly throughout the paper.

11.1.3 Profit Maximization and Taxation

This subsection introduces the tax instruments to be used in the subsequent section and describes the profit-maximizing behavior of firms.

If firm i picks employment N_i, capital K_i, and wage w_i, it obtains revenue $p_i f^i[e^i(n + w_i, \tilde{u})N_i, K_i]$ at a cost equal to $rK_i + w_i N_i + T^i(w_i, N_i, K_i)$, where the function T^i gives the firm's tax liability. Until asymmetric information is introduced into the model, I allow these tax functions to be specified in a way that effectively gives the government complete control over each firm's behavior. For concreteness, I work with tax functions with the following form:

$$T^i(w_i, N_i, K_i) = t_{Ki}K_i + (t_{wi}w_i + \tau_i)N_i. \tag{7}$$

Thus, the firm faces a capital tax, a proportional wage tax, and a per capita employment tax (subsidies are negative taxes). This tax function operates on all the relevant margins: capital is taxed at the marginal rate t_{Ki}, employment is taxed at the marginal rate $t_{wi}w_i + \tau_i$, and the firm's chosen wage is taxed at the marginal rate $t_{wi}N_i$.

To describe the profit-maximizing behavior of firms, it is convenient to write firm i's profits in terms of effective labor, $E_i = e_i N_i$:[6]

$$p_i f^i(E_i, K_i) - (r + t_{Ki})K_i - \{[(1 + t_{wi})w_i + \tau_i]/e^i(n + w_i, \tilde{u})\}E_i. \tag{8}$$

To maximize these profits, the firm chooses the wage to minimize the unit cost of effective labor, given in (8) by the expression in curly brackets. The first-order condition for this minimization problem is[7]

$$\begin{aligned} [w_i + \tau_i/(1 + t_{wi})]/e^i(I_i, \tilde{u}) &= 1/e^i_I(I_i, \tilde{u}) \\ &= [-v_e(I_i, e_i)]/[v_I(I_i, e_i) - (1 - \pi_i)v_I(I_i, 0)], \\ \pi_x < 1, \quad \pi_y &= 1, \end{aligned} \tag{9}$$

where $I_i = n + w_i$, subscripts I and e denote partial derivatives, and the second equality follows from implicit differentiation of (4) and (5). The firm then chooses its capital and effective labor to equate values of marginal products to unit costs:

$$p_i f^i_K(E_i, K_i) = r + t_{Ki} \tag{10}$$

and

$$p_i f^i_E(E_i, K_i) = [(1 + t_{wi})w_i + \tau_i]/e^i(n + w_i, \tilde{u}). \tag{11}$$

First-order conditions (9)–(11) will be used throughout the paper.

The final tax instrument introduced here is a uniform poll tax, collected from each worker. The symbol n then denotes nonlabor income net of this poll tax. Without this tax, the government might not be able to lower the incomes of secondary-sector workers as much as desired. However, its presence adds a fundamental indeterminacy to the model: given any equilibrium, there is an equivalent equilibrium with a higher poll tax. This is easily seen. As the poll tax rises, after-tax incomes can be held constant by raising w_x and w_y by identical amounts. The government budget can then be brought back into balance by using the additional revenue to lower τ_x and τ_y until $(1 + t_{wx})w_x + \tau_x$ and $(1 + t_{wy})w_y + \tau_y$ return to their original values. Thus, neither first-order condition (10) nor (11) is affected by the tax change. Since $w_x + \tau_x/(1 + t_{wx})$ and $w_y + \tau_y/(1 + t_{wy})$ are also clearly unaffected, neither is first-order condition (9) disturbed. Thus, the economy is in a new equilibrium that is identical to the old in all meaningful respects. Essentially, the standard observation about the irrelevance of whether workers or firms pay a tax applies in full force here. In the next section, I shall anchor the tax system without loss of generality by fixing $\tau_y = 0$.

Despite the wide range of tax instruments made available to the government in parts of this paper, I do not allow the government to treat detected shirkers differently than workers who start out in the secondary sector. Without this assumption, the need to use interindustry wage differentials as a worker discipline device disappears, thereby eliminating an essential feature of the economic environment with which this paper is concerned. A possible justification is that significant costs would likely be required to keep track of the past work history of current secondary-sector workers (i.e., did they "shirk"?), especially because these workers would have an incentive to claim those past histories most advantageous to their tax treatment.[8]

11.2 Optimal Government Policy with Symmetric Information

In the absence of taxation, the basic inefficiency in this efficiency wage model may be described as underemployment in the primary sector. To see this, recall that $e^x(n + w_x, \bar{u}) < e^y(n + w_y, \bar{u})$ whenever $w_x = w_y$. This means that the minimum unit cost of effective labor is lower for the secondary-sector firm than for the primary-sector firm: $w_y/e_y < w_x/e_x$. The primary-sector firm will therefore set the value of its marginal product of effective labor above the opportunity cost of effective labor, as measured by forgone secondary-sector output.

This reasoning suggests that the government should design a tax system that effectively subsidizes employment in the primary sector relative to the secondary sector. However, account must be taken of the worsening income inequality that may result from doing so. This section shows that such a tax policy is generally desirable, although the form of the subsidies is generally complex in the sense that it involves the use of both the wage and employment taxes described in the previous section. On the other hand, government intervention in the capital market will be shown *not* to be desirable. This and the other results are demonstrated under the assumption that the government has the same information possessed by private firms and can design a tax system that uses this information in any desired way.

11.2.1 The Basic Setup

Social welfare is defined throughout this paper as the sum of utilities:

$$W = N_x v(w_x, e_x) + N_y v(w_y, e_y). \tag{12}$$

This welfare function can also be thought of as representing a worker's expected utility, assuming that all workers have an equal opportunity of being picked for a primary-sector job. Only nonshirkers appear in W because nobody shirks in equilibrium.

Using its taxing powers, the government effectively exercises complete control over the production and wage policies of all firms. The government's welfare-maximization problem may then be set up with the government treated as though it directly chooses an income-effort vector and input vector for each firm i, denoted (I_i, e_i) and (N_i, K_i). Employment levels N_x and N_y must add up to the total number of workers in the economy, denoted N^*. By using the workers' budget constraints to reexpress the trade balance constraint given by (1), it is then possible to obtain the following formulation of the government's maximization problem:

$$\text{(P1)} \qquad \text{Max } N_x v(I_x, e_x) + N_y v(I_y, e_y)$$

subject to

$$\begin{aligned}(13) \qquad & p_x f^x(e_x N_x, K_x) + p_y f^y(e_y N_y, K_y) + r(K^* - K_x - K_y) \\ & \qquad - N_x I_x - N_y I_y \geq 0\end{aligned}$$

$$(14) \qquad v(I_x, e_x) - \pi_x v(I_y, e_y) - (1 - \pi_x) v(I_x, 0) \geq 0,$$

$$(15) \qquad N_x + N_y = N^*.$$

After the government solves this problem, it can then decentralize the solution using the tax instruments introduced in section 11.1.3. The properties of this tax system are discussed in detail below.

Let λ and β denote Lagrange multipliers for constraints (13) and (14), and substitute $N^* - N_x$ for N_y to get rid of constraint (15). The Lagrangian for problem (P1) may then be written

$$\begin{aligned}(16) \qquad L = {} & N_x v(I_x, e_x) + (N^* - N_x) v(I_y, e_y) \\ & + \lambda\{p_x f^x(e_x N_x, K_x) + p_y f^y[e_y(N^* - N_x), K_y] \\ & \quad + r(K^* - K_x - K_y) - N_x I_x - (N^* - N_x) I_y\} \\ & \quad + \beta\{v(I_x, e_x) - \pi_x v(I_y, e_y) - (1 - \pi_x) v(I_x, 0)\}.\end{aligned}$$

For the subsequent analysis, both constraints are assumed to bind at the margin, implying that

$$(17) \qquad \lambda > 0, \quad \beta > 0.$$

The first inequality must hold since λ represents the marginal value of foreign exchange, which is necessarily positive in this model. The Bulow-Summers paper has the property that $\beta = 0$, but they assume that workers are risk neutral, in which case there is no social cost to income inequality. The government can achieve a first-best allocation simply by making the difference between primary and secondary incomes large enough to eliminate shirking. In the present paper, positive risk aversion is assumed, so increasing the spread between incomes to prevent shirking has a positive social cost, measured at the margin by the multiplier β.

I describe the solution to this problem in the subsequent subsections. First I take up my main concern, capital taxes. Then I discuss the other taxes and argue that, under the optimal tax policy, the primary sector should indeed be the "high-wage" sector, although this is not always true under laissez faire.

11.2.2 The Case against Capital Taxation

Capital taxes play no role in this model. In particular, the following proposition shows that each firm should be allowed to expand its capital stock to the point where the value of the marginal product of capital equals the world interest rate.

Proposition 1: At the optimum,

$$p_x f^x_K(e_x N_x, K_x) = p_y f^y_K(e_y N_y, K_y) = r.$$

Proof: Differentiate the Lagrangian with respect to K_x and K_y, and set the derivatives equal to zero. The result follows immediately. Q.E.D.

This result is related to Diamond and Mirrlees's (1971) finding that aggregate production efficiency is desirable when all commodities can be taxed at any desired rates. Optimal commodity taxation allows consumer prices to be varied independently of producer prices, thereby eliminating any reason to tolerate production inefficiency. In the present case, however, the small country assumption implies that the world product prices, p, are fixed. Thus, deviations from an efficient capital allocation cannot have a desirable effect on the consumer prices, q, even when the government fails to employ an optimal excise tax system. For this reason, proposition 1 holds regardless of whether the government uses excise taxes.

Proposition 1 does require the use of wage and employment subsidies, however. I analyze these instruments below. If they are assumed not to be available, then examples can be constructed in which social welfare is increased by subsidizing capital in the primary sector or taxing capital in the secondary sector. One such example is given in Appendix B. However, the unavailability of all other tax instruments is difficult to justify. A more reasonable approach would be to limit the use of these other instruments by explicitly incorporating informational problems into the model. Section 11.3 follows this approach.

Although I have not explicitly considered foreign tax systems, their existence need not change the results. Suppose that the domestic economy under consideration is a capital importer, and assume, as commonly practiced, that foreign governments allow a tax credit for taxes paid to the domestic government. Then foreign investors are indifferent about where to invest if and only if

$$g[1 - \max(b, b^*)] = g^*(1 - b^{**}),$$

where b and b^* denote the tax rates imposed by the domestic and foreign governments on foreigners' domestically located capital investments, b^{**} is the tax imposed by foreign governments on foreigners' foreign-located capital investments, and g and g^* denote the before-tax returns that these investors receive on domestically and foreign-located capital investments. As argued by Slemrod (1988), the domestic government maximizes social welfare by setting $b = b^*$ since raising b to b^* merely transfers tax revenue from the foreign government to the home government without affecting private investment incentives. But the tax does affect "public tax incentives" by lowering the social cost of capital from $g^*(1 - b^{**})/(1 - b^*)$ to $g^*(1 - b^{**})$ since foreign investors now pay b^* to the domestic government for every unit of their domestic investment. In other words, $g^*(1 - b^{**})$ now serves as the relevant "world interest rate," r, in both the trade balance constraint (eq. [13]) and proposition 1. As a result, the domestic government now finds it advantageous to provide domestic firms with an investment subsidy, s, that is carefully designed to lie outside the tax crediting system used by foreign governments but is set equal to b^* so that domestic firms expand investment to the point where the values of their marginal products equal $g^*(1 - b^{**})$.[9] The net effect of this domestic tax policy is to lower the social opportunity cost of capital without raising any tax revenue. In other words, the statement that capital should not be taxed still holds in the sense that the optimal $b^* - s$ equals zero. Similarly, the subsequent results about capital taxation may be reinterpreted as results about the optimal $b^* - s_i$ for each firm i when foreign tax crediting is practiced.

11.2.3 The Optimal Tax Policy

With the use of capital taxes having been ruled out, it is useful to ask how the government's other tax instruments should be chosen. My first result is that primary-sector wages are higher than secondary-sector wages under the optimal tax system, given reasonable assumptions about the utility function. This result does not follow immediately from the specification of the model since an alternative way of satisfying the no-shirking condition would be to keep the primary-sector effort level (e_x) relatively low. Indeed, Carmichael notes that "it is simply not obvious what (if anything) efficiency wage models predict about wage differentials in the cross section. The results depend on the precise way in which the firm's . . . characteristics combine to affect the position and shape of the entire wage/productivity relationship" (1988, 27–28). His comment concerns the laissez-faire behavior of firms. If the government is able to pursue the optimal tax policy described here, then a rather strong case can be made for providing primary-sector workers with a higher income than secondary-sector workers. In particular, I now prove the following.

Proposition 2: I_x must exceed I_y at the optimum if the following assumptions hold:

i. $v_I(I, e)$ declines with I and is nonincreasing in e;
ii. $v_I(I_x, e_x) - (1 - \pi_x)v_I(I_x, 0) > 0$.

Proof: Differentiate the Lagrangian with respect to income levels to obtain the following first-order conditions for the optimal income levels:

$$(18) \qquad N_x v_I(I_x, e_x) - \lambda N_x + \beta[v_I(I_x, e_x) - (1 - \pi_x)v_I(I_x, 0)] = 0$$

and

$$(19) \qquad N_y v_I(I_y, e_y) - \lambda N_y - \beta\pi_x v_I(I_y, e_y) = 0.$$

Combining (18) and (19) gives,

$$(20) \qquad v_I(I_x, e_x) - v_I(I_y, e_y) = -\beta N_x^{-1}[v_I(I_x, e_x) - (1 - \pi_x)v_I(I_x, 0)] - \beta N_y^{-1}\pi_x v_I(I_y, e_y).$$

Assume now that, contrary to the claim, $I_x \leq I_y$. To satisfy the no-shirking condition, $v(I_x, e_x)$ must exceed $v(I_y, e_y)$. Thus, $e_x < e_y$. By assumption i, it follows that $v_I(I_x, e_x) \geq v_I(I_y, e_y)$. But, under assumption ii, (20) implies that $v_I(I_x, e_x) < v_I(I_y, e_y)$, which is a contradiction. Q.E.D.

Assumption i is quite weak since the marginal utility of income is normally thought of as rising with leisure and an increase in e may be viewed as a reduction in leisure. Assumption ii is also reasonable: although an increase in I_x could increase the incentive to shirk in cases where π_x is near zero and shirking workers possess relatively high marginal utilities of income, such a case is rather extreme. With proposition 2 serving as the justification, I will therefore presume that the primary sector is the high-wage sector throughout this paper.

I now investigate the signs of the optimal per capita employment taxes and proportional wage taxes, τ_i and t_{wi}. Recall that these taxes combine to produce the following *marginal* tax on employment in sector i:

$$(21) \qquad T_i = t_{wi}w_i + \tau_i.$$

The marginal wage tax is $t_{wi}N_i$. The next proposition shows that both employment and the wage should be subsidized at the margin in the primary sector but not in the secondary sector. The subsidies are financed with the poll tax.

Proposition 3: There exists an optimal tax system with the following properties:

$$T_x < 0, \quad t_{wx} < 0, \quad T_y = \tau_y = t_{wy} = 0.$$

Proof: To prove that $t_{wx} < 0$ and $t_{wy} = 0$, first differentiate the Lagrangian with respect to e_x and e_y, giving the first-order conditions

$$N_x v_e(I_x, e_x) + \lambda N_x[p_x f_E^x(E_x, K_x)] + \beta v_e(I_x, e_x) = 0 \tag{22}$$

and

$$N_y v_e(I_y, e_y) + \lambda N_y[p_y f_E^y(E_y, K_y)] - \beta\pi_x v_e(I_y, e_y) = 0. \tag{23}$$

Next use the second equality in (9) to write

$$v_I(I_x, e_x) + v_e(I_x, e_x)(\partial e^x/\partial I_x) = (1 - \pi_x)v_I(I_x, 0) \tag{24}$$

and

$$v_I(I_y, e_y) + v_e(I_y, e_y)(\partial e^y/\partial I_y) = 0. \tag{25}$$

If (22) and (23) are multiplied by $\partial e^x/\partial I_x$ and $\partial e^y/\partial I_y$, respectively, and the results are added to the first-order conditions for I_x and I_y (eqq. [18] and [19]), then (24) and (25) can be used to obtain

$$p_x f_E^x(E_x, K_x)(\partial e^x/\partial I_x)N_x = N_x - \lambda^{-1}(1 - \pi_x)v_I(I_x, 0) \tag{26}$$

and

$$p_y f_E^y(E_y, K_y)(\partial e^y/\partial I_y)N_y = N_y . \tag{27}$$

On the other hand, combining the first-order conditions for profit maximization given by (9) and (11) yields

$$p_x f_E^x(E_x, K_x)(\partial e^x/\partial I_x)N_x = \{1 + t_{wx}\}N_x \tag{28}$$

and

$$p_y f_E^y(E_y, K_y)(\partial e^y/\partial I_y)N_y = \{1 + t_{wy}\}N_y . \tag{29}$$

Equations (26) and (28) then yield $t_{wx} < 0$, while (27) and (29) imply that $t_{wy} = 0$.

As discussed in Section 11.1.3, τ_y may be set equal to zero without loss of generality. With t_{wy} also equal to zero, it follows that $T_y = 0$.

To prove that $T_x < 0$, differentiate the Lagrangian with respect to N_x to obtain the first-order condition

$$\begin{aligned} v(I_x, e_x) - v(I_y, e_y) + \lambda\{[p_x f_E^x(E_x, K_x)e_x - I_x] \\ - [p_y f_E^y(E_y, K_y)e_y - I_y]\} = 0. \end{aligned} \tag{30}$$

Note that

$$I_x - I_y = w_x - w_y . \tag{31}$$

Since $\lambda > 0$ and the no-shirking condition requires that $v(I_x, e_x) > v(I_y, e_y)$, (30) and (31) give

$$p_x f_E^x(E_x, K_x)e_x - w_x < p_y f_E^y(E_y, K_y)e_y - w_y . \tag{32}$$

By the first-order condition for profit maximization given by (11),

$$p_i f_E^i(E_i, K_i)e_i - w_i = t_{wi}w_i + \tau_i = T_i . \tag{33}$$

Substituting (33) into (32) and using $T_y = 0$ gives $T_x < 0$. Q.E.D.

Both Calvo (1985) and Bulow-Summers also demonstrate the desirability of employment subsidies on primary-sector jobs, financed by taxes that impose a burden on self-employed workers (Calvo) or secondary-sector workers (Bulow-Summers). Bulow-Summers find that these subsidies should be used to equate the value of a worker's marginal product across sectors (see their fig. 2). Such a use is not generally desirable in the present model because equity considerations eliminate the desirability of satisfying the standard efficiency conditions (my proposition 1 being a major exception). In fact, the relation between the values of the marginal products of labor in the two sectors cannot be signed in general.

An intuitive explanation may be provided for $t_{wx} < 0$ in proposition 3. By (9), the marginal rate of substitution between labor effort and income in the primary sector is less than the additional incomes that workers must receive to induce them to provide another unit of labor effort:

$$-v_e(I_x, e_x)/v_I(I_x, e_x) < (\partial e^x/\partial I_x)^{-1}. \tag{34}$$

In the absence of employment and wage taxation, however, (28) implies that

$$p_x f_E^x(E_x, K_x) = (\partial e^x/\partial I_x)^{-1}. \tag{35}$$

Thus, we have a situation in which the marginal rate of transformation between effective labor and income exceeds the corresponding marginal rate of substitution; that is, the marginal benefit of additional effort is greater than the marginal cost. For this reason, subsidies should be used to induce the primary-sector firm to raise its wages and thereby induce workers to supply more labor effort without shirking.

In contrast to this result, there is no role for wage subsidies in the Bulow-Summers model because all nonshirking workers are assumed to provide one unit of labor effort, regardless of price incentives. The result also differs from Johnson and Layard's (1986, 963) conclusion that a *positive* proportional tax on a firm's total wage bill is a desirable means of financing a per capita subsidy on employment, the argument being that the combined effect of the two taxes is to lower unemployment in their model. They consider a one-sector efficiency wage model based on labor turnover behavior. The proportional wage tax is completely passed back onto the wage in this model, leaving effective before-tax wages unchanged but lowering all after-tax wages by the same amounts. Such a tax plays the same role as my poll tax: it raises revenue without affecting the marginal behavioral incentives faced by workers and firms.

Note finally that my explanation for wage subsidies in the primary sector does not carry over to the secondary section because the absence of a supervision problem there implies that $p_y f_E^y(E_y, K_y) = - v_e(I_y, e_y)/v_I(I_y, e_y)$ in the absence of taxation.

11.3 Asymmetric Information

I now consider informational asymmetries as a possible justification for positive or negative taxes on internationally mobile capital. My basic assumption about information is that the government possesses incomplete information about the identity of those firms with efficiency wage problems. In other words, the government is not certain about whether a given firm is in the "primary sector" (x) or the "secondary sector" (y). To formalize this idea, I assume that the economy contains a fixed number of firms, indexed by $i = 1, 2, \ldots$; and that the government attaches a probability ψ_i to a firm i being a "type x firm," in which shirking workers are caught with probability $\pi_x < 1$, and a probability $1 - \psi_i$ to the firm being a "type y firm," where the detection probability is $\pi_y = 1$.[10] Thus, a given firm's effort function is either $e^x(I, \bar{u})$ or $e^y(I, \bar{u})$, as previously defined.

To isolate this particular source of uncertainty from uncertainty about production technologies, I continue to assume that each firm's production function is known, $f^i(E, K)$ for firm i.[11] Issues concerning unknown characteristics of production functions are discussed at the end of this section. Note, however, that this specification can be made to handle the empirical observation that capital intensive firms tend to pay high wages simply by making the ψ_i's relatively high for firms with relatively capital intensive technologies. In fact, the model may be transformed back into a symmetric information model by assuming that ψ_i equals zero or one for all i. To avoid obvious qualifications on the results, I henceforth assume that $0 < \psi_i < 1$ for every i.

To make a firm's worker detection probability unobservable, additional assumptions must be made about which of the firm's actions the government can or cannot observe. The government could infer the firm's shirker detection probability from observations of the wage that the firm pays workers and the effort level that it demands in return. To eliminate this possibility, the obvious assumption to make is that the government cannot observe effort levels. But the government could still use its knowledge of production functions to infer effort levels from observations on wages, outputs, employment levels, and capital stocks. Of all these variables, a firm's capital stock is by far the most difficult to measure in practice. Thus, I create an asymmetric information problem by making the capital stock unobservable. Baron and Myerson (1982) follow a similar approach in their seminal article on regulation by assuming that some parameters of a firm's cost function are unobservable. Later, I argue that other choices of the unobservable variable do not affect my results. A natural direction for future research would be to construct a model in which the capital stock is imperfectly observed.

Some readers may now ask, How can you study capital taxation using a model in which the government does not observe capital? The answer is that the government's ability to tax those variables it does observe effectively allows it to tax capital. In particular, the total tax paid by a firm i can be made a function of its wage w, output Q, and employment N: $T^i(w, Q, N)$. The marginal tax on another unit of capital is then $(\partial T^i/\partial Q)(\partial Q/\partial K)$, where $\partial Q/\partial K$ is the marginal product of capital. Under profit maximization, this marginal tax equals $p_i f^i_K(E_i, K_i) - r$ for firm i.

Put differently, there is never any loss of generality in arbitrarily picking a single output or input to be untaxed because only relative prices matter. The limitation placed here on the government's taxation powers is not that capital cannot be taxed directly but rather that the government does not possess the information needed to optimally tailor the tax function to differences between primary- and secondary-sector firms. Instead, it must confront any firm i with a tax function that is independent of its type, $T^i(w, Q, N)$.[12] My main concern is whether the optimal tax system effectively taxes or subsidizes a firm's capital at the margin, as measured by $p_i f^i_K(E_i, K_i) - r$.[13]

11.3.1 The Government's Maximization Problem

To pinpoint the role of the informational asymmetry, it is useful first to pose the government's optimization problem for the case where the government knows each firm's type, but with only those variables that are observable in the asymmetric information case treated as control variables. In particular, capital and effort levels may be omitted as control variables by inverting the production relation for each firm i, $Q = f^i[e^{j(i)}(I, \tilde{u})N, K]$ if firm i's type is $j(i)$, to obtain

$$K = K^i[Q, e^{j(i)}(I, \tilde{u})N]. \tag{36}$$

This leaves the equilibrium secondary-sector utility, $\tilde{u}$, and the income-production vector for each i, (I_i, Q_i, N_i), as the control variables. Problem (P1) may then be rephrased as follows:

$$\text{Max} \sum_i N_i v[I_i, e^{j(i)}(I_i, \tilde{u})] \tag{P2}$$

subject to

$$rK^* + \sum_i \{p_i Q_i - rK^i[Q_i, e^{j(i)}(I_i, \tilde{u})N_i] - N_i I_i\} \geq 0, \tag{37}$$

$$\sum_i N_i = N^*. \tag{38}$$

The no-shirking condition from problem (P1) does not appear in (P2) because it has been incorporated into the effort functions. Each secondary-sector firm provides workers with the equilibrium utility, $v[I_i, e^y(I_i, \bar{u})] = \bar{u}$, while primary-sector workers receive higher utilities to prevent shirking: $v[I_i, e^x(I_i, \bar{u})] > \bar{u}$. Note that this utility differential will generally differ across firms with different production functions because incomes will differ. For any pair of primary- and secondary-sector firms, however, propositions 1–3 continue to hold. Of greatest interest here is proposition 1, which says that no firm's capital should be taxed or subsidized at the margin.

Now consider the asymmetric information problem. It is again best to proceed indirectly by setting up the optimization problem with outputs, employment levels, and incomes treated as control variables rather than optimizing directly over the set of permissible tax functions. Since the government does not know whether a given firm i is type x or type y, however, it must choose a wage-production vector for both contingencies: (w_{ix}, Q_{ix}, N_{ix}) and (w_{iy}, Q_{iy}, N_{iy}).[14] Furthermore, these vectors will be feasible only if the firm can construct a tax function, $T^i(w, Q, N)$, such that (w_{ix}, Q_{ix}, N_{ix}) gives the firm at least as high a profit level as (w_{iy}, Q_{iy}, N_{iy}) if the firm's type is x and, conversely, if its type is y. Equivalently, there must exist payments T_{ix} and T_{iy} such that the profit function for the type x firm satisfies

$$(39)\quad p_iQ_{ix} - w_{ix}N_{ix} - rK^i[Q_{ix}, e^x(n + w_{ix}, \bar{u})N_{ix}] - T_{ix} \geq p_iQ_{iy} - w_{iy}N_{iy} - rK^i[Q_{iy}, e^x(n + w_{iy}, \bar{u})N_{iy}] - T_{iy}\,,$$

while the profit function for the type y firm satisfies

$$(40)\quad p_iQ_{iy} - w_{iy}N_{iy} - rK^i[Q_{iy}, e^y(n + w_{iy}, \bar{u})N_{iy}] - T_{iy} \geq p_iQ_{ix} - w_{ix}N_{ix} - rK^i[Q_{ix}, e^y(n + w_{ix}, \bar{u})N_{ix}] - T_{ix}\,.$$

These are completely new types of constraints, known in the principal-agent literature as an "incentive-compatibility constraints." They can be combined into a single constraint by adding (39) and (40) together, canceling common terms on the two sides of the inequality, and rearranging the result to obtain

$$(41)\quad K^i[Q_{iy}, e^x(n + w_{iy}, \bar{u})N_{iy}] - K^i[Q_{ix}, e^x(n + w_{ix}, \bar{u})N_{ix}] \geq K^i[Q_{iy}, e^y(n + w_{iy}, \bar{u})N_{iy}] - K^i[Q_{ix}, e^y(n + w_{ix}, \bar{u})N_{ix}].$$

This constraint can be understood by observing that the optimal tax problem being considered here is equivalent to the design of an optimal "truth-telling mechanism." The government asks firms to name their types, uses the answers to control their production activities, and ensures that these answers are truthful by awarding the firms with positive or negative subsidies based on their answers. Constraint (41) ensures that such subsidies exist by requiring that the incentive to reveal its type as x rather than y, measured in terms of capital cost savings, is at least as great for an actual type x firm as for an actual type y firm. By the famous "Revelation Principle" from the principal-agent literature, no sacrifice in welfare is incurred by considering only truth-telling mechanisms.

There is no need to include a separate constraint requiring that the tax function allow each firm i to earn nonnegative profits: $p_i Q_{ij} - w_{ij}N_{ij} - r \bullet K^i[Q_{ij}, e^j(n + w_{ij}, \tilde{u}) N_{ij}] - T_{ij} \geq 0$ for $j = x, y$, where T_{ij} is again the tax firm i owes if it chooses (w_{ij}, Q_{ij}, N_{ij}). Such constraints would never be binding. To see this, suppose that a given tax function violates one of them. Then the government can lower the total tax owed at every (w, Q, N) by the same amount until profits become nonnegative for both types of firms. This change in the tax function obviously does not affect the profit maximizing (w, Q, N) for either type, and its effect on nonlabor incomes can be offset by a reduction in the poll tax (recall that n is nonlabor income net of this tax).

With (41) representing the only new constraint for the problem, there is no need to include transfers T_{ix} and T_{iy} as explicit variables in the maximization problem. In contrast, the regulator in Baron and Myerson's paper possesses an objective function that contains the subsidies paid to the monopolist. The reason for this difference is that Baron and Myerson treat consumers and the monopolist as separate agents and assume that income in the hands of consumers has a greater social value than income in the hands of the monopolist. This assumption bears a close relation to Laffont and Tirole's (1986) assumption that there is an exogenously determined deadweight loss associated with the transfer of income from consumers to the monopolist. In the present model, however, workers are also owners of the firms, and the transfers provided to the firms can be financed by nondistortionary taxes. Guesnerie and Laffont (1984) also study a class of principal-agent problems in which the income transfers to the agent do not enter the principal's objective function.

The government's maximization problem may now be stated in full. There is no need to include w_{ix} and n as separate control variables in this problem because only their sum matters. Thus, the control variables are the equilibrium utility, $\tilde{u}$, and an income-production vector for each firm i, $(I_{ix}, Q_{ix}, N_{ix}, I_{iy}, Q_{iy}, N_{iy})$. With ψ_i denoting firm i's probability of being type x, these variables are chosen to solve

$$\text{(P3)} \qquad \text{Max} \sum\nolimits_i \{\psi_i[N_{ix}v(I_{ix}, e^x(I_{ix}, \tilde{u}))] + (1 - \psi_i)N_{iy}\tilde{u}\}$$

subject to

$$\text{(42)} \qquad rK^* + \sum\nolimits_i \{\psi_i[p_iQ_{ix} - rK^i(Q_{ix}, e^x(I_{ix}, \tilde{u})N_{ix}) - N_{ix}I_{ix}]\} + \sum\nolimits_i \{(1 - \psi_i)[p_iQ_{iy} - rK^i(Q_{iy}, e^y(_{iy}, \tilde{u})N_{iy}) - N_{iy}I_{iy}]\} \geq 0,$$

$$\text{(43)} \qquad \sum\nolimits_i \{\psi_i N_{ix} + (1 - \psi_i)N_{iy}\} = N^*,$$

$$(44) \quad K^i[Q_{iy}, e^x(I_{iy}, \tilde{u})N_{iy}] - K^i[Q_{ix}, e^x(I_{ix}, \tilde{u})N_{ix}] \geq K^i(Q_{iy}, e^y(I_{iy}, \tilde{u})N_{iy}) - K^i[Q_{ix}, e^y(I_{ix}, \tilde{u})N_{ix}]$$

for all i. As shown, I require only that the trade balance constraint and employment constraint hold in an expected value sense. The assumption underlying this specification is that the number of firms is large enough to eliminate uncertainty about trade and employment in the aggregate.

The major difference between problems (P2) and (P3) is the presence of the incentive-compatibility constraint in the latter. But this difference is irrelevant if the incentive-compatibility constraint does not bind. Appendix C demonstrates that there exist cases where the incentive-compatibility constraint does bind and cases where it does not. Since the example given there assumes that consumers are risk neutral, the analysis demonstrates that the informational asymmetry may by itself prevent the attainment of a first-best optimum. I now discuss the implications of a binding incentive-compatibility constraint for capital taxation.

11.3.2 Capital Taxation

This section presents an argument for taxing primary-sector capital at a positive rate and secondary-sector capital at a negative rate. The driving force behind the result is that the primary sector has an inferior supervision technology. Simply stated, tax policy should discourage investment in firms with inferior production processes.

To prove these results, I need the additional assumption that labor and capital are complements in the sense that an increase in either factor raises the marginal product of the other: $f^i_{EK}(E, K) > 0$.[15] Violations of this assumption would be hard to justify at the current level of aggregation, although they are theoretically possible under my assumption of decreasing returns to scale (but not under constant returns).

The specific proposition is stated as follows:

Proposition 4: If capital and labor are complements in all firms, and if the incentive-compatibility constraint binds at the margin for firm i, then the following conditions hold at the optimum:

i. $p_i f^i_K(E_{ix}, K_{ix}) > r$;

ii. $p_i f^i_K(E_{iy}, K_{iy}) < r$.

Proof: Let λ and α_i denote the Lagrange multipliers on constraints (42) and (44). Omitting i as a subscript or superscript to avoid clutter, I may write the first-order conditions for firm i's outputs, Q_x and Q_y, as follows:

$$(45) \quad \lambda\psi_i[p_i - rK_Q(Q_x, e^x(I_x, \tilde{u})N_x)] + \alpha_i[K_Q(Q_x, e^y(I_x, \tilde{u})N_x) - K_Q(Q_x, e^x(I_x, \tilde{u})N_x)] = 0$$

and

$$\lambda(1 - \psi_i)[p_i - rK_Q(Q_y, e^y(I_y, \bar{u})N_y)] + \alpha_i[K_Q(Q_y, e^x(I_y, \bar{u})N_y) - K_Q(Q_y, e^y(I_y, \bar{u})N_y)] = 0. \tag{46}$$

Recall that

$$e^y(I, \bar{u}) > e^x(I, \bar{u}) \tag{47}$$

for any given I and $\bar{u}$. It follows that, if both the type x and the type y firms employ the same numbers of workers and pay them the same wages to produce the same output levels, then the type y firm uses more effective labor and less capital than the type x firm to produce this common output. Under the assumption that labor and capital are complements, the marginal product of capital is then greater in the type y firm than in the type x firm, or, since the derivative $K_Q(\bullet)$ is the inverse of this marginal product,

$$K_Q[Q, e^y(I, \bar{u})N] < K_Q(Q, e^x(I, \bar{u})N] \tag{48}$$

for any Q, N, I, and $\bar{u}$. Conditions (45) and (48) then imply that

$$p_i - rK_Q(Q_x, e^x(I_x, \bar{u})N_x] > 0, \tag{49}$$

while (46) and (48) give

$$p_i - rK_Q[Q_y, e^y(I_y, \bar{u})N_y] < 0. \tag{50}$$

Inequalities (49) and (50) are equivalent to i and ii of the proposition. Q.E.D.

The basic idea behind these results may be simply explained using the equivalence between the tax scheme and truth-telling mechanisms. If the government increases the output it wants a given firm i to choose if its type is x (Q_x), then the profits that this firm receives by choosing the wage-production plan (w_x, Q_x, N_x) change by

$$\Delta_x = p_i - rK_Q[Q_x, e^x(I_x, \bar{u})N_x],$$

where $I_x = n + w_x$. On the other hand, if the given firm were a type y but chose the same (w_x, Q_x, N_x), then its profits would change by

$$\Delta_y = p_i - rK_Q[Q_x, e^y(I_x, \bar{u})N_x].$$

But $\Delta_y > \Delta_x$ because (47) and the complementarity assumption imply that capital is more productive in y than in x. This means that the rise in Q_x gives the type y firm a greater incentive to masquerade as a type x firm, relative to the incentive the x firm faces to reveal its type truthfully. The marginal social cost of these incentive changes is determined by the Lagrange multiplier on the firm i's incentive-compatibility constraint. To offset this cost, a rise in Q_x from the optimum must improve the trade balance, implying that the value of the marginal product of capital must exceed the world interest rate. In other words, primary-sector capital should be positively taxed at the margin. By a similar argument, secondary-sector capital should be subsidized.

11.3.3 Alternative Specifications

Using the type of reasoning just given, I may quickly show that alternative specifications of the informational asymmetry either do not change the results or eliminate any role for capital market intervention. Suppose first that the government can directly tax capital, employment, and wages but finds monitoring a firm's output to be prohibitively costly.[16] Let (w_x, N_x, K_x) be the wage-input vector that it wishes to assign a given firm i if its type is x, and consider an increase in K_x. If the type x firm chooses (w_x, N_x, K_x), then its profits change by

$$\Delta_x = p_i f_K[e^x(I_x, \tilde{u})N_x, K_x] - r,$$

where again $I_x = n + w_x$. On the other hand, if the given firm i were type y but chose to masquerade as a type x firm by also choosing (w_x, N_x, K_x), then it would experience a change in profits given by

$$\Delta_y = p_i f_K[e^y(I_x, \tilde{u})N_x, K_x] - r.$$

Again, $e^y(I_x, \tilde{u})$ exceeds $e^x(I_x, \tilde{u})$ since only the type x firm has the supervision problem. By the assumption of complementary factors, it follows that $\Delta_y > \Delta_x$. Increasing primary-sector capital therefore increases the incentive for the type y firm to masquerade as a type x firm, relative to the incentive of the type x firm to reveal its type truthfully. Again, this cost must be balanced by an improvement in the trade surplus, implying that $p_i \cdot f_K(E_x, K_x)$ exceeds r at the optimum. By similar reasoning, $p_i f_K(E_y, K_y)$ falls short of r.

The story changes if either wages or employment is made the unobservable variable. In either case, if the government observes both a firm's output and capital stock, then it can infer the firm's effective labor from the production function, $Q = f(E, K)$. Thus, a type y firm cannot hide its true identity unless it chooses the same Q, E, and K as a type x firm (although the two types will use different effort and employment levels to obtain the same $E = eN$). But then the marginal product of capital in a type x firm choosing the (Q_x, K_x) assigned to it will be identical to the marginal production of a y firm choosing the same (Q_x, K_x). Raising K_x can therefore have no desirable effect on the relevant incentive-compatibility constraint for the problem, implying that primary-sector capital should not be taxed or subsidized at the margin. By the same reasoning, neither should secondary-sector capital be taxed or subsidized.

This last argument can be used to analyze the case where the detection probabilities depend positively on unobservable managerial effort, m, with π_y exceeding π_x at any given level of m. Suppose that this managerial effort enters the firm's objective function as an unobservable cost.[17] With both output and capital observable, a type x firm will again have to choose the same effective labor as a type y firm in order to hide its true identity. If wages and employment are also observable, this means that the type x firm will be able to masquerade as a type y firm only if it raises its managerial effort enough to equate π_x with π_y. In any case, the type y and x firms will again possess the same marginal products of capital if they choose the same levels of the observable variables. By the previous argument, neither primary- nor secondary-sector capital should then be taxed or subsidized at the margin.

So far, uncertainty about supervision problems has been analyzed in the absence of production function uncertainty by assuming that the government knows each firm's production function but not the supervision technology. A more general specification might also allow the production function to be uncertain. In this case, if the possibility that a firm has a supervision problem is positively related to the possibility that its marginal product of capital is relatively high at any given (E, K), the results may be reversed. This can be explained intuitively. In the previous model, capital taxation effectively steers investment away from firms with inferior supervision technologies and into firms with superior supervision technologies. High wages in the primary sector basically mask inefficiencies in the production process, thereby making the taxation of primary-sector capital desirable. Differences in production functions may offset this inefficiency, however. Loosely stated, the result is still that the sector with the more efficient production process should receive positive capital subsidies, but the identity of this sector will now depend on both differences in production functions and supervision technologies.

11.4 Concluding Remarks

To summarize, this paper finds little support for policies that subsidize capital investment in high-wage industries. Indeed, the opposite conclusion is obtained under reasonable assumptions about informational asymmetries: a positive marginal tax should be placed on primary-sector capital. The driving force behind this result is that the production processes used by high-wage firms are inferior to low-wage production processes in one particular aspect: the supervision technology.

Thus, informational asymmetries create a role for basing tax policy on efficiency differences between sectors. A useful task for future research would be to examine this role in a variety of different contexts. For example, if efficiency wages are paid to reduce labor turnover rather than shirking, then perhaps investment in high-wage jobs should still be taxed at the margin since high wages may again be viewed as masking an inefficiency in the production process, namely, the relatively severe worker turnover problem. This extension would require a dynamic model.

Another role for a dynamic analysis would be to model the intertemporal process by which a government learns the attributes of various firms. One conjecture is that asymmetric information reasons for distortionary capital taxation are unimportant in the long run since the attributes of different firms can eventually be uncovered. But the government's acquisition of information may be severely hampered by the incentives that firms face to hide those activities that may increase their future tax burdens. A recent paper by Laffont and Tirole (1988) suggests that this consideration may be a serious problem for tax policy.

To conclude, three limitations of the asymmetric information analysis deserve emphasis. First, when interpreting all these results, it is important to keep in mind that only marginal taxes have been considered, not average taxes. In fact, the two are likely to depart quite significantly since a highly nonlinear tax schedule would be required to tax a given firm's capital at a positive marginal rate if its type is x and at a negative marginal rate if its type is y, as defined in the text.

Second, the paper has not addressed the issue of how the incentive-compatibility constraint affects the taxation of employment and wages. I have not obtained clear-cut results on this issue, but it deserves further research.

Finally, the degree to which the results on capital taxation are sensitive to the particular informational asymmetry modeled here needs to be further explored. Section 11.3.3 has noted that there exist additional sources of uncertainty that work against proposition 4. As matters now stand, however, I hope to have convinced readers that it is difficult to justify subsidizing capital investment in high-wage industries with any reasonable degree of confidence.

Appendix A

This appendix discusses the optimal role of excise taxes in the symmetric information case and how this use alters the optimal values of the other tax instruments. To study this role, I must now explicitly include the consumer price vector in the demand and utility functions: $c^i(q, I, e)$ and $v(q, I, e)$. Since only relative prices matter, however, good y may be arbitrarily chosen as the untaxed commodity: $q_y = p_y$.

If $c^x(q, I_x, 0) > (<)\ c^x(q, I_x, e_x)$ in the absence of excise taxes, then placing a small positive (negative) tax on good x raises $v(q, I_x, e_x) - v(q, I_x, 0)$, thereby lessening the shirking problem. This explains the following result.

Proposition A1: At the optimum, $q_x \gtreqless p_x$ as $c^x(q, I_x, 0) \gtreqless c^x(q, I_x, e_x)$.

Proof: The Lagrangian expression given by (16) must be reformulated to take account of excise taxes. To do so, I use the workers' budget constraints to write,

$$\text{(A1)} \quad p_x c^x(q, I_i, e_i) + p_y c^y(q, I_i, e_i) = I_i - (q_x - p_x)c^x(q, I_i, e_i).$$

Using this equality to rewrite the trade balance constraint given by (1), equation (16) may be amended to read

$$\begin{aligned} \text{(A2)} \quad L = {} & N_x v(q, I_x, e_x) + (N^* - N_x)v(q, I_y, e_y) \\ & + \lambda\{p_x f^x(e_x N_x, K_x) + p_y f^y[e_y(N^* - N_x), K_y] \\ & \quad + r(K^* - K_x - K_y) - N_x I_x - (N^* - N_x)I_y \\ & \quad - (q_x - p_x)[N_x c^x(q, I_x, e_x) + (N^* - N_x)c^x(q, I_y, e_y)]\} \\ & + \beta\{v(q, I_x, e_x) - \pi_x v(q, I_y, e_y) - (1 - \pi_x)v(q, I_x, 0)\}. \end{aligned}$$

Starting from the optimum, any combination of changes in consumer prices or incomes must have a zero first-order effect on the Lagrangian. Consider an increase in q_x, accompanied by increases in I_x and I_y that leave unchanged the utilities of all secondary-sector workers and nonshirking primary-sector workers:

$$\text{(A3)} \quad v_q(q, I_y, e_y)dq_x + v_I(q, I_y, e_y)dI_y = 0$$

and

$$\text{(A4)} \quad v_q(q, I_x, e_x)dq_x + v_I(q, I_x, e_x)dI_x = 0,$$

where d denotes a differential change. In other words, I am considering compensated changes in q_x for both types of workers. By Roy's Identity, the income compensations must satisfy

$$\text{(A5)} \quad dI_i = c^x(q, I_i, e_i)dq_x, \quad i = x, y.$$

Using (A5), the first-order change in the Lagrangian from these compensated price changes may then be expressed as follows:

$$\begin{aligned} \text{(A6)} \quad dL = {} & -\lambda(q_x - p_x)[N_x dc^x(q, I_x, e_x) + N_y dc^x(q, I_y, e_y)] \\ & - \beta(1 - \pi_x)dv(q, I_x, 0) = 0. \end{aligned}$$

Another application of Roy's Identity gives the following expression for the first-order change in the utilities of undetected shirkers:

$$\text{(A7)} \quad dv(q, I_x, 0) = v_I(q, I_x, 0)[-c^x(q, I_x, 0)dq_x + dI_x].$$

Substituting (A5) into (A7) for dI_x and then substituting the result into (A6) yields

$$\text{(A8)} \qquad dL = \lambda(q_x - p_x)[N_x dc^x(q, I_x, e_x) + N_y dc^x(q, I_y, e_y)] + \beta(1 - \pi)v_I(q, I_x, 0)[c^x(q, I_x, 0) - c^x(q, I_x, e_x)]dq_x = 0.$$

The demand changes in the first bracketed term are necessarily negative because they result from a compensated price increase. The proposition then follows immediately from (A8). Q.E.D.

Thus, the government should place a positive (negative) tax on consumption of the primary-sector good if undetected shirkers possess higher (lower) demands for this good than nonshirking primary-sector workers. In the case of a separable utility function, $u(c_x, c_y, e) = g(c_x, c_y) - h(e)$, proposition A1 implies that the optimal $q_x - p_x$ equals zero.

Remember that no worker actually shirks in equilibrium. The differential tax burden here acts as an additional incentive not to shirk without penalizing any worker.

Proposition A1 offers an interesting contrast to Dixit's (1989) finding that all marginal rates of substitution should equal world prices, although an adverse selection problem in his model constrains the relative utilities of different workers. The crucial difference between the models is easily pinpointed by Dixit's explanation for this result: "The point is that adverse selection imposes incentive-compatibility constraints, but these apply to the utility *levels*, $U_{ij} = U(x_{ij}, y_{ij})$, not to the means by which they are attained. Therefore the usual efficiency argument for minimizing the resource costs of achieving the desired utility levels remains valid" (238). In my model, not only does the no-shirking condition constrain utility levels, but the government is also constrained to give undetected shirkers and nonshirkers in the primary sector the same incomes and consumer prices. This inability to achieve desired utility levels by any means is responsible for the desirability of excise taxes.

I conclude this appendix by describing how the presence of excise taxes affects propositions 2 and 3 in the text.[18] The case of marginal employment subsidies is straightforward. Proposition 3 shows that work in the primary sector should be subsidized relative to work in the secondary sector in the sense that $T_x < T_y$ (the individual values of these taxes were shown to be indeterminate in sec. 11.1.3). This claim remains valid with excise taxes if the definition of the relative subsidies is modified to include differences in the excise tax payments:

$$\text{(A9)} \qquad T_x + (q_x - p_x)c^x(q, I_x, e_x) < T_y + (q_x - p_x)c^x(q, I_y, e_y).$$

If poll tax payments are added to both sides of (A9), then the government budget constraint can be balanced only if the right side is positive and the left is negative. As before, secondary-sector workers face a positive total tax burden, while primary-sector workers face a negative total tax burden. The only new feature is that this tax burden includes excise taxes (or subsidies).

The required modification of the wage tax results in proposition 3 is described by the following formulae for the marginal wage taxes:

$$(\text{A10}) \qquad t_{wx}N_x = -\lambda^{-1}(1 - \pi_x)v_I(q, I_x, 0) - (q_x - p_x)N_x[c_I^x(q, I_x, e_x) + c_e^x(q, I_x, e_x)\cdot(\partial e^x/\partial I_x)]$$

and

$$(\text{A11}) \qquad t_{wy}N_y = -(q_x - p_x)N_y[c_I^x(q, I_y, e_y) + c_e^x(q, I_y, e_y)\cdot(\partial e^y/\partial I_y)].$$

Account must now be taken of the effects of a rise in income and effort levels on the distorted pattern of consumption. For this reason, it may now be undesirable to subsidize wage increases in the primary sector. For the secondary sector, whether consumption becomes more or less distorted in response to a rise in the wage and effort level completely determines the sign of the wage tax. Since proposition A1 shows that the optimal $q_x - p_x$ can be positive or negative, however, there appears to be no general presumption about the direction in which these new considerations push the signs of the wage taxes at the full optimum.

Turn finally to proposition 2, which shows that primary-sector workers receive higher incomes than secondary-sector workers. If aggregate consumption of the commodity with a positive (negative) tax can be increased (reduced) by transferring income from primary- to secondary-sector workers, then this transfer reduces the distortionary effect of the excise tax on consumption patterns. For this reason, proposition 2 may no longer hold in all cases, although it is hard to imagine that this additional consideration would be strong enough in practice to reverse the desirability of setting I_x above I_y.

Appendix B

This appendix considers the case for subsidizing primary-sector capital when employment and wage subsidies are not available. To keep the government budget balanced, a "neutral" tax instrument is used to balance the budget, namely, a uniform poll tax (positive or negative) imposed on each worker. Since such a tax is "lump sum," it creates no incentive effects by itself. Capital taxes and subsidies are then left as the only means of dealing with the multiple distortions described in section 11.2. Concrete results are therefore hard to come by in the general case. For the following analysis, I alter the model in a way that allows me to concentrate on the employment distortion described in section 11.2. Specifically, I now follow the common practice in the efficiency wage literature of assuming that the effort level takes on only two values: zero and one. All workers provide one unit of labor in equilibrium since nobody shirks. Altered in this way, the model is called the "modified model."

My main result uses the assumption that labor and capital are complements in the sense that an increase in either factor raises the marginal product of the other. In this case, an increase in the price of either must lower the demand for both:

(B1) If $f^i_{NK}(N_i, K_i) > 0$, then $dN_i/dw_i < 0$ and $dK_i/dw_i < 0$,

and similarly for a rise in $r + t_{Ki}$ (see Silberberg 1978, sec. 4.4).

The main result now may be stated as follows.

Proposition B1: Suppose that $t_{Kx} = t_{Ky} = 0$ initially in the modified model. Then a small fall in t_{Kx} raises social welfare if factors are complements in the primary sector, and a small rise in t_{Ky} raises social welfare if factors are complements in the secondary sector.

Proof. A change in either of the two taxes must leave unchanged the trade balance constraint, given by (13) with $e_x = e_y = 1$:

$$\text{(B2)} \quad [p_x f^x_K(N_x, K_x) - r]dK_x + [p_y f^y_K(N_y, K_y) - r]dK_y + [p_x f^x_N(N_x, K_x) - p_y f^y_N(N_y, K_y) - (w_x - w_y)]dN_x - [N_x dI_x + N_y dI_y] = 0,$$

where a d denotes a differential change, and use is made of the equality, $I_x - I_y = w_x - w_y$. Given the absence of any initial taxes, the profit-maximization conditions given by (10) and (11) can be used to reduce (B2) to the following expression:

$$\text{(B3)} \qquad N_x dI_x + N_y dI_y = 0.$$

Equation (B3) implies that any change in primary- and secondary-sector worker incomes must take place in opposite directions. But the no-shirking condition, (5), implies that these changes must take place in identical directions. It follows that I_x and I_y do not change:

$$\text{(B4)} \qquad dI_x = dI_y = 0.$$

Thus, the welfare effects of a tax change are completely determined by the employment change, $dN_x = -dN_y$. Since the no-shirking condition implies that primary-sector workers have higher utilities than secondary-sector workers, welfare rises if and only if the tax change shifts employment from the secondary sector to the primary sector:

$$\text{(B5)} \qquad dW = (u_x - u_y)dN_x > 0$$

if and only if $dN_x > 0$, where $u_i = v(I_i, 1)$.

Since all workers obtain identical nonlabor incomes, (B4) also implies that any tax-induced changes in the primary- and secondary-sector wages must be identical:

(B6) $$dw_x = dw_y .$$

This result will be used to sign dN_x.

Suppose that factors are complements in the secondary sector, and consider a rise in t_{Ky}. If N_x fails to rise, then N_y cannot fall, and the assumption of complements implies that w_y must fall to offset the negative effect of t_{Ky} on N_y. Then w_x falls to satisfy (B6), which implies that N_x rises, contradicting the initial assumption.

Suppose that factors are complements in the primary sector, and consider a fall in t_{Kx}. If N_x fails to rise, then the assumption of complements implies that w_x must rise to offset the positive effect of the decline in t_{Kx} on N_x. Then w_y rises to satisfy (B6), which implies that N_y falls, contradicting the assumption that N_x does not rise. Q.E.D.

The basic explanation for this result is that welfare can be improved by undertaking policies that encourage labor to flow from the secondary sector to the primary sector. Given the complementarity assumption, one such policy is to encourage capital investment in the primary sector, while another is to discourage capital investment in the secondary sector. But dropping this assumption leads to ambiguous results and thereby highlights the roundabout nature of capital taxation as a means of encouraging employment in the primary sector.

Appendix C

This appendix shows by way of an example that the incentive-compatibility constraint in section 11.3 is binding in some cases but not in others. Since the example contains risk neutral consumers, it also demonstrates that the asymmetric information problem described in section 11.3 may render infeasible the first-best allocation, even if there is no social cost attached to the income inequality required to eliminate shirking.

By risk neutrality, I mean that utility is linear in income. It is further assumed that the disutility of labor effort is a perfect substitute with income, in which case the indirect utility function may be written

(C1) $$v(I, e) = I - h(e); \quad h'(e) > 0, \quad h''(e) \geq 0.$$

(The coefficient of I is set equal to one to simplify notation.) Equation (C1) allows the no-shirking condition faced by type x firms (eq. [5]), to be reexpressed

(C2) $$I_x\pi_x - h(e_x) \geq \pi_x\bar{u},$$

while the utility requirement faced by type y firms is

(C3) $$I_y - h(e_y) \geq \bar{u}.$$

For part of my example, I shall work with the following simple form of the effort disutility function:

(C4) $$h(e) = e^{\eta}, \quad \eta \geq 1.$$

On the production side of the model, I simplify matters by assuming that all firms produce the same good using identical production functions and that they possess identical probabilities of being a type x firm:

(C5) $$f^i(E, K) = f(E, K), \quad \psi_i = \psi$$

for all i. Thus, the only difference between firms is their unknown detection probabilities: $\pi_x < \pi_y = 1$.

Following the text, the number of firms is assumed to be large enough for actual aggregate income to be reasonably approximated by expected income. Normalizing the number of firms to equal one, aggregate income may then be written

(C6) $$I^A = \psi N_x I_x + (1 - \psi) N_y I_y .$$

Social welfare may then be written

(C7) $$W = I^A - [\psi N_x h(e_x) + (1 - \psi)(N^* - N_x) h(e_y)].$$

This is also social welfare when each firm's type, x or y, is known to the government since the probability ψ simply becomes the known fraction of type x firms.

I now investigate whether the incentive-compatibility constraint is satisfied under the solution to the symmetric information problem and, therefore, whether it is a binding constraint in the asymmetric information problem.

Since only aggregate income enters the social welfare function, no loss of generality is involved in setting $I_y = 0$ and giving all the income to workers in the primary sector since doing so maximizes the set of effort levels that are consistent with the no-shirking condition (eq. [C2]). (Alternatively, if there is a minimum subsistence income greater than zero, I_y may be set equal to its value.) Except for relatively low values of π_x, the no-shirking condition will then fail to bind and can therefore be omitted from the problem. In other words, the first-best solution will be feasible. This is the case I consider.

The control variables for the government's maximization problem are I^A, e_x, e_y, N_x, K_x, and K_y. They are chosen to maximize social welfare, as defined by (C7), subject to the trade balance constraint,

(C8) $$\psi p f(e_x N_x, K_x) + (1 - \psi) p f[e_y (N^* - N_x), K_y] + r[K^* - \psi K_x - (1 - \psi) K_y] - I^A \geq 0.$$

Given the symmetry of the problem, the solution calls for treating all firms alike along every dimension except income (with $I_x > I_y$ to take care of shirking):

(C9) $$e_x = e_y = e', \quad K_x = K_y = K', \quad N_x = N_y = N'.$$

Let us now consider whether an allocation with properties (C9) can satisfy the incentive-compatibility constraint for the asymmetric information problem. For the two types of firms both to be willing to choose e', I_x and I_y must be set so that

(C10) $$e^x(I_x, \tilde{u}) = e^y(I_y, \tilde{u}) = e', \quad \tilde{u} = I_y - h(e').$$

With incomes so determined, the incentive-compatibility condition ([44] in the text) becomes

(C11) $$K[Q', e^x(I_y, \tilde{u})N'] - K[Q', e^x(I_x, \tilde{u})N'] \geq K[Q', e^y(I_y, \tilde{u})N'] - K[Q', e^y(I_x, \tilde{u})N'],$$

where Q' is the common output level for both types of firms under (C9).

Whether (C11) holds will depend on the properties of both the production function and the effort functions. For the latter functions, I prove the following fundamental result.

Claim: Under assumptions (C1) and (C4), $e^x_I(I, \tilde{u}) < e^y_I(I, \tilde{u})$ for all I and $\tilde{u}$.

Proof: By definition,

(C12) $$h[e^y(I, \tilde{u})] = I - \tilde{u},$$

and, using (C2),

(C13) $$h[e^x(I, \tilde{u})] = \pi_x[I - \tilde{u}].$$

Since $h'(e) = \eta h(e)/e$ under (C4), (C12) and (C13) give

(C14) $$h'(e^y)e^y_I = \eta(I - \tilde{u})(e^y_I/e^y) = 1$$

and

(C15) $$h'(e^x)e^x_I = \pi_x\eta(I - \tilde{u})(e^x_I/e^x) = \pi_x,$$

where the arguments of functions e^x and e^y have been omitted to simplify notation. Since $e^x(I, \tilde{u}) < e^y(I, \tilde{u})$, (C14) and (C15) imply that $e^x_I(I, \tilde{u}) < e^y_I(I, \tilde{u})$. Q.E.D.

Thus, the increase in effort levels that a firm can obtain by increasing worker incomes from I_x to I_y is less for the type x firms than for the type y firms. The crucial implication of this claim is

(C16) $$e^x(I_x, \tilde{u}) - e^x(I_y, \tilde{u}) < e^y(I_x, \tilde{u}) - e^y(I_y, \tilde{u}).$$

By itself, (C16) clearly works against the satisfaction of (C11).

But production considerations work in the opposite direction. In particular, if we assume as before that capital and labor are complements, then it is easily shown that the marginal rate of substitution between effective labor and capital at a given (I, Q, N) is higher for a type x firm than for a type y firm:

$$\text{(C17)} \qquad -K_E[Q, e^x(I, \bar{u})N] > -K_E[Q, e^y(I, \bar{u})N].$$

The basic reasoning comes from the inequality, $e^x(I, \bar{u}) < e^y(I, \bar{u})$. When the type x and y firms pay the same numbers of workers the same incomes to produce the same Q, the type x firm uses more capital and less effective labor. As a result, its marginal product of capital is lower and its marginal product of labor higher than those for the type y firm. Thus, the amount of capital needed to compensate for a unit reduction in effective labor is then greater in the type x firm than in the type y firm; that is, (C17) holds.

This production consideration clearly works in favor of (C11), but the assumption of complements implies nothing about the magnitude of the difference in (C17). Indeed, the form of the production function may be varied to make this difference as small or large as desired, thereby producing examples where (C11) holds and examples where it does not. Thus, the first-best optimum is feasible in some asymmetric problems but not in others.

Finally, I demonstrate that, no matter how severely (C11) limits the solution to the asymmetric information, the assumption that the two factors are complements implies that it is never optimal to implement a "pooling equilibrium," where both types of firms are assigned the same (I, Q, N). Under this solution, the equality $I_x = I_y$ would give $E_x = e_x N < e_y N = E_y$, implying that $K_x > K_y$. By the assumption of complements, it would follow that $f_K(E_x, K_x) < f_K(E_y, K_y)$, which violates proposition 4. While this proposition assumes that the incentive-compatibility constraint is binding at the margin, dropping that assumption would imply that $f_K(E_x, K_x) = f_K(E_y, K_y)$, which is again inconsistent with a pooling equilibrium.

Notes

1. Throughout this paper, the term *first best* refers to the allocation that would be socially optimal in the absence of supervision problems.

2. The interest rate referred to here and throughout the paper is, of course, calculated net of taxes levied abroad. The analysis does not depend on whether the given country is a net capital exporter or importer, but in the latter case my formal model ignores foreign tax credits. I argue in sec. 11.2.2 that my results, if properly interpreted, are not affected by foreign tax credits.

3. An economy with many firms is explicitly considered when I model government uncertainty about the identity of firms with worker supervision problems.

4. The production and utility functions in this paper are assumed to be twice continuously differentiable.

5. Labor effort is measured continuously here, although many studies in the literature assume that e takes only two values, zero for shirkers and one for nonshirkers (Sparks 1986 is an important exception). The present specification is needed for the asymmetric information problem in sec. 11.3, and it is responsible for the role of marginal wage subsidies in sec. 11.2.

6. Since the production function is strictly concave, maximum profits will always be positive if the profit-maximizing output level is positive, as assumed throughout the paper.

7. I assume that the effort functions are strictly concave in income, in which case the profit-maximizing wage varies continuously with the tax parameters. This assumption holds under reasonable assumptions about utility.

8. More generally, no utility differences of any type are allowed between workers in the secondary sector, even when different secondary-sector firms are later explicitly included in the model. Examples could presumably be constructed in which welfare is improved by providing workers in these different firms with different utilities, but such examples would be undesirably sensitive to ad hoc assumptions about the rationing mechanism by which detected shirkers get reallocated across these firms. In a related development, the optimal commodity tax literature has already demonstrated the potential desirability of "random taxation" (e.g., Stiglitz 1982; and Chang and Wildasin 1986), but the principle of "horizontal equity" is often invoked to rule out its use.

9. Given that property taxes paid by domestic firms are not credited by foreign governments, property subsidies might serve the stated purpose. If subsidies outside the crediting system were not available, the domestic government would then face incentives to lower the effective cost of capital through various expenditure programs.

10. The results would not change if detection probabilities were allowed to differ across firms in the primary sector.

11. Different firms may produce identical or different goods.

12. The optimal tax functions will generally vary across firms if production functions differ. My results will concern properties that are common to all these functions, suggesting that the same properties would often hold if the government were forced to confront different firms with identical tax functions.

13. Defining the capital tax in terms of this difference has the desirable feature of not requiring that the optimal tax function $T^{i}(w, Q, N)$ be differentiable, which may not be the case. For a similar procedure with regard to optimal income taxation, see Stiglitz (1987b, 1003).

14. The actual optimization problem contains only total income, I_{ij}, rather than its components, n and w_{ij}, as control variables. In fact, any positive n and w_{ij}'s that sum to the optimal I_{ij}'s are then optimal. (Recall the nonuniqueness of the optimal tax system discussed in sec. 11.1.3.) As a pedagogical device, however, it is useful first to fix the government's choice of n.

15. Paradoxically, the same assumption is also used in App. B to prove that primary-sector capital should be subsidized at the margin while secondary-sector capital should be positively taxed when no other tax instruments are available.

16. For a general analysis of principal-agent problems of this type, with a special application to the optimal regulation of a labor managed firm, see Guesnerie and Laffont (1984).

17. For a detailed analysis of the optimal regulation of a firm when the government can observe both output and costs but where costs can be reduced by means of unobservable managerial effort, see Laffont and Tirole (1986). A particularly exciting feature of their work is that they allow costs to be stochastic at the time the firm chooses output and effort levels.

18. Proofs of the results reported here are available from the author on request.

References

Arvan, Lanny, and Francoise Schoumaker. 1988. Is public opinion right and conventional wisdom wrong about commercial policy? Typescript.

Baron, David P., and Roger B. Myerson. 1982. Regulating a monopolist with unknown costs. *Econometrica* 50 (July): 911–30.

Bulow, Jeremy I., and Lawrence H. Summers. 1986. A theory of dual labor markets with applications to industrial policy. *Journal of Labor Economics* 4 (July): 376–414.

Calvo, Guillermo A. 1985. The inefficiency of unemployment: The supervision perspective. *Quarterly Journal of Economics* 100 (May): 373–87.

Calvo, Guillermo A., and Stanislaw Wellisz. 1978. Hierarchy, ability, and income distribution. *Journal of Political Economy* 87 (October): 991–1012.

Carmichael, H. Lorne. 1988. Efficiency wage models of unemployment: A survey. Typescript.

Chang, Fwu-Ranq, and David E. Wildasin. 1986. Randomization and commodity taxes: An expenditure minimization approach. *Journal of Public Economics* 31 (December): 329–45.

Diamond, Peter A., and James A. Mirrlees. 1971. Optimal taxation and public production I–II. *American Economic Review* 61 (March, June): 8–27, 261–78.

Dixit, Avinash. 1989. Trade and insurance with adverse selection. *Review of Economic Studies* 56 (April): 235–47.

Guesnerie, Roger, and Jean-Jacques Laffont. 1984. A complete solution to a class of principal-agent problems with an application to the control of a self-managed firm. *Journal of Public Economics* 25 (December): 329–69.

Johnson, George E., and Richard Layard. 1986. The natural rate of unemployment: Explanation and policy. In *Handbook of Labor Economics,* vol. 2, ed. Orley C. Ashenfelter and Richard Layard. New York: North-Holland.

Katz, Lawrence F. 1988. Some recent developments in labor economics and their implications for macroeconomics. *Journal of Money, Credit and Banking* 20 (August): 507–22.

Laffont, Jean-Jacques, and Jean Tirole. 1986. Using cost observation to regulate firms. *Journal of Political Economy* 94 (June): 614–41.

______. 1988. The dynamics of incentive contracts. *Econometrica* 56 (September): 1153–75.

Shapiro, Carl, and Joseph E. Stiglitz. 1984. Equilibrium unemployment as a worker discipline device. *American Economic Review* 74 (June): 433–44.

Silberberg, Eugene. 1978. *The structure of economics: A mathematical analysis.* New York: McGraw-Hill.

Slemrod, Joel. 1988. Effect of taxation with international capital mobility. In *Uneasy compromise: Problems of a hybrid income-consumption tax,* ed. H. J. Aaron, H. Galper, and J. A. Pechman. Washington, D.C.: Brookings.

Sparks, Roger. 1986. A model of involuntary unemployment and wage rigidity: Worker incentives and the threat of dismissal. *Journal of Labor Economics* 4 (October): 560–81.

Stiglitz, Joseph E. 1982. Utilitarianism and horizontal equity: The case for random taxation. *Journal of Public Economics* 18 (March): 1–34.

______. 1986. Theories of wage rigidity. In *Keynes' economic legacy: Contemporary economic theories,* ed. J. Butkiewicz, K. Koford, and J. Miller. New York: Praeger.

______. 1987a. The causes and consequences of the dependence of quality on price. *Journal of Economic Literature* 25 (March): 1–48.

______. 1987b. Pareto efficient and optimal taxation and the new welfare economics. In *Handbook of public economics,* vol. 2, ed. A. J. Auerback and M. Feldstein. Amsterdam: North-Holland.

Comment Lawrence F. Katz

The basic question addressed in Wilson's interesting paper is whether noncompetitive interindustry wage differentials provide a rationale for industrial policies in an economy with internationally mobile capital. This is a particularly important question to analyze given recent concern with U.S. competitiveness in strategic sectors and worries about shifts in employment from sectors that provide high-wage, "good jobs" to sectors that provide low-wage, "bad jobs."

Any economic justification for worrying about the sectoral composition of output must rely on the presence of market imperfections that drive a wedge between the marginal productivities of factors in different sectors. Wilson focuses on a labor market distortion within the context of an efficiency wage model in which a worker supervision problem requiring a wage premium is present in one sector (the primary sector) and absent in the other sector (the secondary sector).

In the first part of the paper, Wilson extends the Bulow-Summers dual labor market model by introducing capital in the production function and making capital perfectly internationally mobile. Wilson concludes that there is no role for the use of sectoral capital taxes and subsidies in the optimal government policy in this model when other instruments are available. This is not surprising since the distortion arises in the labor market (the wage is above opportunity costs in the primary sector) and can be directly solved through policies that serve to subsidize employment in the primary sector. The inclusion of capital and open economy considerations adds little to the analysis of this model. On the other hand, this section of the paper is useful in showing how the exact form of labor market policies (wage vs. employment subsidies and taxes) depends on the exact specification of the monitoring technology in a shirking model. The conclusion of no role for sectoral capital tax and subsidy policies also depends on the choice of model. In models where noncompetitive wage differentials arise from the differential ability of workers to extract product market rents and sunk investments in different sectors, the labor market distortion may take the form of a tax on investment that differs across sectors (Grout 1984; Katz and Summers 1989). In this case, even though the distortion arises in the labor market, its interaction with capital investment decisions means that the optimal policy may involve subsidies to investment in sectors where investments are particularly appropriable by labor.

The second part of the paper adds asymmetric information to a shirking model. The government is assumed not to know the identity of firms with supervision problems requiring premium wages. While there are many

Lawrence F. Katz is assistant professor of economics at Harvard University and a research associate of the National Bureau of Economic Research.

reasons to be skeptical of industrial policies and of the ability of the government to determine whether wage differentials represent "labor rents" rather than competitive differentials, the incentive compatibility issues discussed in the second part of the paper do not seem particularly relevant. The government surely has no problem differentiating firms in high-wage sectors (e.g., steel and auto plants) from those in low-wage sectors (e.g., fast food restaurants and retail stores). The more relevant issue is the political economy issue of how to control rent-seeking activities once significant subsidies become potentially available to potentially affected groups that can commission economists and statisticians to document that their industries provide high-rent jobs that should be subsidized. These political problems, emphasized in critiques of sectoral policies by Aaron (1989) and Schultze (1983), are not well illuminated in the mathematics of truth-telling constraints highlighted by Wilson's asymmetric information model.

The Wilson paper takes as given the existence of noncompetitive wage differentials arising from differences in the importance of supervision problems among firms. Many other discussions of industrial policies presume the existence of identifiable "good jobs" and "bad jobs." In fact, much recent empirical research has examined the nature of interindustry wage differentials (e.g., Dickens and Katz 1987; Krueger and Summers 1988; and Katz and Summers 1989). The basic finding is that there are large, systematic, persistent interindustry wage differentials that remain even after controlling for all observable worker and job characteristics available in micro data sets. For example, workers in autos, aircraft, and petroleum consistently earn 20–40 percent more than workers with the same measured characteristics in apparel, retail trade, and repair services. Interindustry wage differentials are remarkably similar across developed economies and highly correlated over time. The persistence over time suggests they are not just transitory differentials. The similarity across countries means that they reflect something fundamental in the nature of advanced industrial economies rather than particular labor market institutions. The differentials appear even larger when one includes employee benefits in measures of labor compensation.

High-wage industries have lower quit rates and face longer queues of job applicants than low-wage industries. These findings indicate that interindustry wage differentials are not largely compensating differentials. The low quit rates and long job queues in high-wage-differential sectors are easy to explain if these jobs provide labor rents. An alternative view is that industry wage differentials reflect the sorting of workers across industries on the basis of unmeasured ability (Murphy and Topel 1987). Longitudinal studies (reviewed in Katz and Summers 1989) find that the wage changes of industry switchers are quite similar to estimated cross-sectional differentials. This evidence casts doubt on the view that these differentials reflect sorting on time-invariant, unobserved productive ability. Furthermore, industry wage differentials are highly correlated across occupations. Industries that pay their managers wage premia also pay wage premia to their secretaries, laborers, and janitors. It is difficult to believe that industries that need high-ability managers also always need high-ability janitors. This strong correlation across occupations combined with the finding that wage differentials are strongly positively correlated with measures of product market rents per worker and appropriable capital per worker suggests that these differentials may reflect rent-sharing considerations.

The overall evidence does appear to be fairly persuasive that there do exist large noncompetitive interindustry wage differentials. Since profits account for a small share of value added relative to labor compensation, even small noncompetitive differences in wages across industries are likely to have more significant allocative consequences than variations in capital rents. In fact, Katz and Summers (1989) find that even conservative estimates of the variation in labor rents across sectors are substantially larger than the variation in capital rents. This suggests that Wilson's emphasis on labor market distortions rather than on profit shifting considerations in his analysis of industrial policies is well placed. More work on the sources of noncompetitive wage differentials is clearly required before strong policy statements can be made given that different models of differentials can lead to quite different predictions. Furthermore, any economic case for activist policy must be tempered by a recognition of the formidable difficulties likely to be encountered in the implementation of policies that pick "good" and "bad" industries.

References

Aaron, Henry J. 1989. Politics and the professors revisited. *American Economic Review* 79 (May): 1–15.

Dickens, William T., and Lawrence F. Katz. 1987. Inter-industry wage differences and industry characteristics. In *Unemployment and the structure of labor markets,* ed. K. Lang and J. Leonard, 48–89. Oxford: Blackwell.

Grout, Paul A. 1984. Investment and wages in the absence of binding contracts: A Nash bargaining approach. *Econometrica* 52 (March): 449–60.

Katz, Lawrence F., and Lawrence H. Summers. 1989. Industry rents: Evidence and implications. *Brookings Papers on Economic Activity: Microeconomics,* 209–75.

Krueger, Alan B., and Lawrence H. Summers. 1988. Efficiency wages and the inter-industry wage structure. *Econometrica* 56 (March): 259–93.

Murphy, Kevin M., and Robert H. Topel. 1987. Unemployment, risk, and earnings: Testing for equalizing differences in the labor market. In *Unemployment and the structure of labor markets,* ed., K. Lang and J. Leonard, 103–40. Oxford: Blackwell.

Schultze, Charles L. 1983. Industrial policy: A dissent. *Brookings Review* (Fall), 3–12.