May 28, 2008

Redistributing to the sick: How should health expenditures be integrated into the tax system?*

Abstract

I study the optimal joint taxation of income and health expenditures in a model in which individuals differ in their wage-earning ability and in their health status. First-best redistribution based on potential wage rates and health status is not feasible, and tax payments must be based on endogenous incomes and health expenditures. Within a class of quasi-linear schedules, the conditions for an optimal tax/subsidy system depend on the own and cross price compensated elasticities of demand for leisure and health in a way that generalizes the standard results from the optimal linear income tax literature. Numerical simulations are employed to illustrate the sensitivity of tax and subsidy rates to the correlation between health status and wages. In these simulations, the effective marginal income tax rate optimally increases with health expenditures. However, the welfare gain from optimally incorporating health expenditures into the tax system appears to be very limited, compared with the effect of properly designing the income tax itself.

JEL codes: H2, I1

Keywords: Optimal taxation, health subsidies, redistribution

William Jack
Georgetown University

*Thanks to Robin Boadway, Helmut Cremer, Maurice Marchand, Pierre Pistieau, and Michael Smart for discussion and comments, and to Yamin Ahmad for research assistance.

1 Introduction

There is widespread support in many countries for public intervention in health care delivery markets on redistributive (as well as efficiency) grounds, either in the form of universally available public health care or through subsidized care targeted to certain populations. At the same time, there is a growing literature on the interactions between health and income inequality, recently surveyed by Deaton (2001) and Case (2001). These studies examine links between the level and distribution of income and/or consumption and health outcomes, and investigate, inter alia, the impact of income redistribution on health status. Exactly how a health care subsidy regime should be designed, how large the subsidy should be, and how it should be related to the income tax system, is not necessarily obvious. While the importance of the tax treatment of health care and health insurance, particularly in the US, has long been recognized (e.g., Feldstein, 1973, Pauly, 1986), and policymakers continue to propose tax incentives to address perceived health care and insurance needs (e.g., as reviewed by Pauly and Herring, 2002), a formal integration of health status into a model of income taxation is required to properly design such interventions. This paper aims to provide such a framework.

Atkinson and Stiglitz (1976) showed that as long as individuals differ only in their income earning ability and preferences are separable between leisure and consumption of other goods (so demands for other goods depend only on their prices and the individual's earned income), then commodity taxation cannot increase welfare above the level attained under an optimal non-linear income tax. Under the standard conditions, this theorem implies that health expenditures should not affect tax liability.[1] More generally however, if demand for health care depends on earned income and wage earning ability, then its consumption can be used to relax the incentive constraint constraining the optimal income tax (Nichols and Zeckhauser, 1982). This may occur, for example, if individuals differ by health status which in turn partially determines wage rates

[1] A similar conclusion is drawn by Hylland and Zeckhauser (1979).

and health care demand. If individuals with poor health have lower wages (ability), then as long as health expenditure is income elastic, it can be used as an "indicator" of ability, and thereby increase the efficiency of redistribution from high wage to low wage individuals.

A second reason it might be desirable to integrate health care utilization into a redistributive income tax system is that people with similar wage rates may have very different health needs. Indeed, while income and health status may be correlated, there is no doubt wide variation in the conditional distribution of health needs. Even in the absence of income earning heterogeneity, there is a rationale for redistribution from the sick to the healthy, simply because the budget sets of the former are smaller than those of the latter. In principle this could be effected by a health care subsidy (as suggested by Nichols and Zeckhauser, page 376). However with wage heterogeneity, if health care demand is income elastic, some of the benefits of such a subsidy would accrue to individuals with high wages. The extent to which a health subsidy redistributes from healthy to sick must then be compared to the extent to which it potentially redistributes from poor to rich.

To investigate this trade-off, the model developed in this paper allows individuals to differ on two dimensions - their income earning ability and their health status - and investigates the way tax liability should be related to endogenously determined incomes and health expenditures. Just as the distortionary effects of non-lump-sum income taxes limit the desirable degree of redistribution, distortionary effects of health expenditure subsidies may also limit the extent to which the healthy subsidize the sick. We abstract from all potential market failures in both the health sector and the labor market.[2]

A growing number of contributions to the mechanism design literature (e.g., Mussa and Rosen (1978), Wilson (1993), Armstrong (1996), and Rochet and Choné (1998)) have considered models with multi-dimensional type spaces, in which agents' characteristics differ in more than one dimension. Boskin and Stiglitz (1977), Blomqvist and Horn (1984), and Rochet (1991), made early

[2] This is the normal approach in the one-dimensional optimal income tax literature, where, in particular, employers are assumed to be able to observe an individual's marginal product costlessly. Correspondingly, in the health sector, I assume there is no adverse selection, and individuals with greater needs pay higher prices.

contributions to optimal tax theory when individuals have multi-dimensional types. These have been complemented more recently by papers by Cremer and Pestieau (1996), Cremer, Lozachmeur, and Pestieau (2002), and Boadway, Meite-Monteiro, and Marchand (2002). Cremer and Pestieau (1996) extend Rochet's (1991) analysis of income taxation and social insurance to the case when preferences are not quasi-linear. Cremer et al. (2002) model optimal income taxation in the presence of retirement decisions when individuals differ by labor productivity and health status, where the latter determines the disutility of remaining in the labor force. The model analyzed by Boadway et al. (2002) is the closest in spirit to that of the current paper. While it includes a richer description of the insurance market (allowing for adverse selection), it permits a less general tax instrument.

Blomqvist and Horn (1984), Rochet (1991), and Cremer and Pestieau (1996) examined the use of a public, tax-financed, insurance system as a means of redistributing well-being when individuals differ in their risks of suffering a loss.[3] The general outcome of this analysis is that if there is a negative correlation between wages and risk (so that on average the rich have better health), then a public insurance system should complement an optimal redistributive tax system. Bad health (or generically, a realized loss) is simply used as an imperfect, but nonetheless valuable, signal of an individual's *ex ante* risk of falling ill, so that transfers to sick people, say through a public health insurance system or another form of social insurance, are welfare improving. Importantly, there is no deadweight loss associated with these transfers. For example, in Cremer and Pestieau individuals are always perfectly insured, being able to purchase actuarially fair insurance to cover any losses not borne by the social insurance system. Similarly, in both Rochet and Cremer and Pestieau, the size of the loss incurred is not affected by the public (or private) insurance coverage - that is, there is no hidden information moral hazard. Cremer and Pestieau do acknowledge the possibility of hidden action moral hazard (wherein insurance affects the probability of a loss

[3] In Blomqvist and Horn, and Rochet, the loss was associated with an illness. In Cremer and Pestieau is it a generic income loss. In all papers, the size of the loss is fixed and constant across individuals.

occuring), but provide only a cursory discussion.

On the other hand, one aspect of the health care sector that has received a tremendous amount of attention is the impact of insurance on the demand for health care, and the inefficiencies that arise due to overconsumption of health care when it is subsidized.[4] Similarly, subsidization of health insurance, whose demand is elastic, results in distortionary costs of excessive coverage. In order to focus on these efficiency costs most simply, the paper uses a model with no uncertainty and considers the demand for health *care* (as opposed to health *insurance*). Individuals differ not in their risks of falling ill, but in their (certain) health needs.[5] As explained in the presentation of the model below, while the price of health *care* is the same for all individuals, they can be thought of as facing heterogeneous prices for *health improvement*. As long as the price elasticity of demand for health care is between zero and one (an empirically robust observation, _ _ _), for a given income, individuals with greater health needs will spend more on health care services (although they might not buy more health improvement).

A second way in which the model of the current paper differs from some of the related literature is with respect to the information structure. First, individuals' private information is non-trivially two-dimensional. In contrast, in Cremer and Pestieau (1996), although individuals differ on two dimensions (wage and probability of a loss), there is perfect (positive or negative) correlation between the two, so individuals' characteristics are effectively distributed over a one-dimensional subset of $\mathbb{R}^2$. In our model, wages and health status are imperfectly correlated, if at all, and individuals' characteristics are distributed over a two dimensional region. Secondly, the type spaces used in many of the papers referred to above have been limited to a binary support, while we allow for continuous distributions of types (in both dimensions).[6]

[4] See, e.g., Pauly (1986).

[5] In fact, the model can be applied to any situation in which the prices of a particular good, in addition to leisure, vary across individuals. Education may be another example, in which the cost of adding to one's human capital is a function of one's ability. In such a case, the distribution of wages and education prices would be likely to exhibit a high degree of correlation.

[6] An advantage of the binary support assumption is that optimal non-linear tax schedules are relatively easy to characterize.

We do not attempt to solve the two dimensional mechanism design problem that underlies the optimal non-linear taxation of labor income and health expenditures Instead, we restrict ourselves to a subset of admissible tax functions that are characterized by four components: a universal lump-sum grant, a constant marginal tax on income, a constant marginal subsidy to health expenditures, and an interaction term. The interaction term represents a tax payment that is a fixed proportion of income times health expenditure.[7] This functional form allows ask to ask whether effective marginal income tax rates should decrease with health expenditures, or conversely, whether the effective subsidy to health expenditures should fall with income. The functional form also includes, as special cases, income tax deductions and credits for health expenditures.[8]

An advantage of the current model is that it treats health expenditures and income symmetrically. Indeed, one might wonder why, if a public health system were optimal (as in Rochet, 1991), would a fully socialized workforce not also be optimal. The obvious answer is that the distortionary effects on labor supply would be too great. This paper attempts to account for the distortionary effects of national insurance (or full subsidization of medical care) on health care utilization. It allows us to address the issue of how income tax rates should vary (if at all) with health expenditures, or conversely, how health expenditure subsidies should vary (if at all) with income.

The formulae we derive for the four components of the optimal tax function generalize those of the standard linear income tax model with one dimension of heterogeneity. Simulations of the model are used to investigate the quantitative significance of using health expenditures as a redistributive instrument. Initially, tax schedules are restricted to be strictly linear (so the interaction

[7] From a practical standpoint, the marginal income tax and health subsidy can be implemented on a real-time basis (i.e., through income tax withholding and health care price subsidies). However, the payment based on the interaction between income and health spending would necessarily be calculated at the end of the tax year, when individuals complete their income tax returns.

[8] Stiglitz and Boskin (1977) have addressed the issue of whether medical expenses should be creditable or deductible against an income tax, although within a restricted framework. Similarly, Atkinson (1977) examined the issue of housing subsidies within an optimal tax model, but assumed that one dimension of heterogeneity was observable by the government.

term is constrained to be zero). Health expenditures are either subsidized or taxed, depending on the correlation between health status and income earning ability. They are subsidized only if the correlation between wages and poor health status is sufficiently negative. Indeed, even if the correlation between these underlying characteristics is zero, it is optimal to put a small tax on health expenditures, since health *expenditures* are correlated with income, even if health status is not. When the correlation is positive, the case for a tax is strengthened, since then high health spending (associated with greater health needs) is an even more informative signal of relatively high earning ability, and a tax on spending reduces the reliance on the distortionary income tax instrument.

The simulations suggest that when income and health expenditures are subject to a linear tax schedule, the optimal income tax rate is about 30%, across wage-health status correlations. On the other hand, the optimal health subsidy rate ranges from +10% to -10%. These are not insignificant subsidy rates. However, we also present a measure of the welfare impact of optimally incorporating health expenditures in this linear fashion. While introducing a simple linear income tax (with no health subsidy) improves social welfare by an amount equivalent to a budget increase of about 5 percent of GDP, over the range of wage-health status correlations examined the greatest welfare gain (which accrues when the correlation is high) is equivalent to a budget increase of between just 0.02 and 0.04 percent of GDP. For intermediate levels of correlation, the gain is truly infinitessimal, being as small as 0.00004 percent of GDP for modest correlations. This set of simulations thus suggests that the gains to incorporating health expenditures into the tax system could be quite limited. Of course, more extensive variations in the parameterization of the simulations may reveal cases in which the gains are more substantial. Similarly, including health spending in a tax system that is not otherwise optimally designed, could have a larger effect.

Finally, we ask what sign the interaction term should have in an optimal quasi-linear tax system. Within the class of cases considered in the simulations, welfare is increased by introducing a small positive tax on the product of income and health expenditures.[9] Thus at the optimum, the effective marginal income tax rate increases with health expenditures, and the subsidy to (or tax on) health spending falls (increases) with income.

The next section sets out the government's welfare maximization problem. Section 3 derives the government's first order conditions. Section 4 provides some interpretation to the first order conditions, although the relatively large number of tax parameters (four) makes such interpretation more difficult than in the standard linear income tax model. In light of this, section 5 presents the numerical simulations, and section 6 concludes.

2 Optimal taxation with wage and health status heterogeneity

Individuals consume three goods, health (h), leisure (l), and a general consumption good (c). Each individual has an endowment of one unit of leisure, and labor supply is $L = 1 - l$.[10] Preferences are described by the suitably differentiable utility function $u(h, c, l)$. Individuals face different wage rates, w, and health prices, p, and are distributed on a set Ω in $\mathbb{R}^2_+$, with a continuous and differentiable joint probability density function $f(w, p)$.

The interpretation of leisure and consumption are standard, but a brief discussion of health is warranted. The good h can be thought of as "healthiness", which is produced by purchasing health care services, x. The price of health care services is normalized to unity. The health production function, which transforms health care services into health outcomes, is linear, but the productivity, π, of health care varies across individuals depending on their health status: that is, $h = \pi x$. The cost of acquiring healthiness h is thus $h/\pi = ph$, where we define $p = 1/\pi$.

Taxes are levied on the basis of earned income, $Y = wL$ and health expenditures, $Z = ph$,

[9] Clearly optimally introducing another tax instrument available cannot reduce welfare. The result of interest is that the sign of the interaction term should be positive.

[10] It may be interesting to introduce an endowment of health that differs across individuals.

according to the schedule

$$T = -G + tY - sZ + \theta(Y.Z)$$

where G is a universal lump-sum grant, and t, s, and θ are marginal tax rate parameters. Let $\Sigma = \{G, t, s, \theta\}$ be the set of tax parameters. An individual with wage rate w and health price p takes the tax parameters as given and chooses her consumption bundle according to

$$\max_{h,c,l} u(h, c, l) \text{ s.t. } c = G + (1-t)Y - (1-s)Z - \theta YZ \tag{1}$$

yielding income $Y = w(1-l)$, health expenditures $Z = ph$ and indirect utility $v(w, p; \Sigma)$.

The government chooses G, t, s, and θ to maximize social welfare, defined as

$$W = \int_{\Omega} \Psi(v(w, p; \Sigma)) f(w, p) dw dp$$

where Ψ is a strictly increasing weakly concave function of utilities. This maximization is carried out subject to a revenue constraint, R_0. That is,

$$G + R_0 = \int_{\Omega} (tY(w, p) - sZ(w, p) + \theta Y(w, p)Z(w, p)) f(w, p) dw dp$$

where we have normalized the population to one. Given a tax system defined by parameters Σ, some individuals may choose not to work. This set of individuals is described in Appendix 1.

3 Optimality conditions

Let us follow the standard approach (e.g., Sheshinski (1972), Atkinson and Stiglitz (1980, page 405)) and derive the first order conditions for the government's maximization problem. The Lagrangian for the problem is

$$\mathcal{L} = \int_{\Omega} (\Psi(v) + \lambda[tY - sZ + \theta YZ - G - R_0]) f(w, p) dw dp \tag{2}$$

where λ is the Lagrange multiplier on the revenue constraint.

3.1 Lump-sum grant

Differentiating (2) with respect to G, we obtain

$$\int\limits_{\Omega} \left(\Psi'(v)\frac{\partial v}{\partial G} + \lambda[\phi(w,p) - 1] \right) f(w,p)dwdp = 0 \tag{3}$$

where we define

$$\phi(w,p) = \frac{d}{dm}\left(tY(w,p) - sZ(w,p) + \theta Y(w,p)Z(w,p)\right)$$

as the additional revenue collected due to an increase in an individual's income, m. Following Atkinson and Stiglitz (1980), we define the social marginal valuation of income accruing to a (w,p)-type individual as

$$b(w,p) = \Psi'(v)\frac{\alpha(w,p)}{\lambda} + \phi(w,p),$$

where $\alpha(w,p)$ is the private marginal utility of income of an individual with wage w and health price p. Equation (3) then reduces to

$$\bar{b} = 1. \tag{4}$$

That is, given the other tax parameters, the lump-sum component should be adjusted to the point at which the marginal social valuation of an additional dollar, across all consumers, is equal to the marginal cost (one dollar).

3.2 Income tax rate, t

Differentiating (2) with respect to t yields

$$\int\limits_{\Omega} \left(\Psi'(v)\frac{\partial v}{\partial t} + \lambda Y \left[1 + \frac{tw}{Y}(1 + \frac{\theta Z}{t}) \left.\frac{\partial L}{\partial t}\right|_u - \frac{sp}{Y}(1 - \frac{\theta Y}{s}) \left.\frac{\partial h}{\partial t}\right|_u - \phi \right] \right) f(w,p)dwdp = 0$$

where the terms $\frac{\partial L}{\partial t}\big|_u$ and $\frac{\partial h}{\partial t}\big|_u$ represent the change in labor supply and health demand, respectively, in response to a change in t, keeping the individual on a fixed indifference surface in $\mathbb{R}^3$. Note that $\frac{\partial v}{\partial t} = -\alpha Y$. If we define the effective marginal income tax and health expenditure subsidy rates by, respectively, $\tau = t + \theta Z$ and $\sigma = s - \theta Y$, the first order condition can be written

$$\mathrm{cov}(Y,b) + \int\limits_{\Omega} \left(Y\frac{\tau}{1-\tau}\varepsilon_L - Z\frac{\sigma}{1-\tau}\varepsilon_{h\widehat{w}} \right) f(w,p)dwdp = 0 \tag{5}$$

where ε_L is the compensated elasticity of labor supply with respect to the net wage, $\widehat{w} = w(1 - t - \theta Z)$, i.e.,

$$\varepsilon_L = \frac{\widehat{w}}{L} \left.\frac{\partial L}{\partial \widehat{w}}\right|_u .$$

Similarly (see the appendix for further discussion), $\varepsilon_{h\widehat{w}}$ is the cross elasticity of health demand with respect to the net wage,[11]

$$\varepsilon_{h\widehat{w}} = \frac{\widehat{w}}{h} \left.\frac{\partial h}{\partial \widehat{w}}\right|_u .$$

Note well however that τ and σ are each functions of endogenous choices by the individuals, and vary according to w and p (in particular, $\tau = t + \theta Z(w,p)$ and $\sigma = s - \theta Y(w,p)$) and so (5) is difficult to simplify further.

3.3 Health expenditure subsidy, s

The tax system allows for a direct credit for a fraction s of incurred health expenditures. By symmetry, the first order condition of s is similar to that for t,

$$\int\limits_{\Omega} \left(\Psi'(v)\frac{\partial v}{\partial s} + \lambda Z \left[1 + \frac{sp}{Z}(1 - \frac{\theta Y}{s}) \left.\frac{\partial h}{\partial s}\right|_u - \frac{tw}{Z}(1 + \frac{\theta Z}{t}) \left.\frac{\partial L}{\partial s}\right|_u - \phi \right] \right) f(w,p)dwdp = 0.$$

With $\frac{\partial v}{\partial s} = \alpha Z$, this becomes

$$\mathrm{cov}(Z,b) + \int\limits_{\Omega} \left[Z\frac{\sigma}{1-\sigma}\varepsilon_h - Y\frac{\tau}{1-\sigma}\varepsilon_{L\widehat{p}} \right] f(w,p)dwdp = 0, \tag{6}$$

where ε_h is the compensated elasticity of demand for health with respect to the net price of health, $\widehat{p} = p(1 - s + \theta Y)$, and $\varepsilon_{L\widehat{p}}$ is the compensated elasticity of labor supply with respect to this price.

3.4 The interaction term, θ

Finally, through θ, the tax system allows the marginal tax rate to vary with health expenditures, or equivalently, it allows the rate of credit for health expenditures to vary with income. The first order condition for θ is

$$\int\limits_{\Omega} \left(\Psi'(v)\frac{\partial v}{\partial \theta} + \lambda YZ \left[1 + \frac{\tau}{YZ} \left.\frac{\partial Y}{\partial \theta}\right|_u - \frac{\sigma}{YZ} \left.\frac{\partial Z}{\partial \theta}\right|_u - \phi \right] \right) f(w,p)dwdp = 0.$$

[11] Own price elasticities are denoted with single subscripts, while cross price elasticities have the corresponding quantity and price subscripts.

Using $\frac{\partial v}{\partial \theta} = -\alpha YZ$, this becomes

$$-\text{cov}(YZ, b) + \int\limits_{\Omega} \left[\tau w \left. \frac{\partial L}{\partial \theta} \right|_u - \sigma p \left. \frac{\partial h}{\partial \theta} \right|_u \right] f(w,p) dw dp = 0,$$

or, using equations (13) and (14) in the appendix

$$\begin{aligned}
\text{cov}(YZ, b) &= \int\limits_{\Omega} \left[\tau w \left. \frac{\partial L}{\partial \theta} \right|_u - \sigma p \left. \frac{\partial h}{\partial \theta} \right|_u \right] f(w,p) dw dp \\
&= -\int\limits_{\Omega} \left[\tau w L \left(\frac{Z}{(1-\tau)} \varepsilon_L - \frac{Y}{(1-\sigma)} \varepsilon_{L\widehat{p}} \right) \right. \\
&\quad \left. - \sigma p h \left(\frac{Z}{(1-\tau)} \varepsilon_{h\widehat{w}} - \frac{Y}{(1-\sigma)} \varepsilon_h \right) \right] f(w,p) dw dp \\
&= -\int\limits_{\Omega} \left[Z \frac{(\tau Y \varepsilon_L - \sigma Z \varepsilon_{h\widehat{w}})}{(1-\tau)} - Y \frac{(\tau Y \varepsilon_{L\widehat{p}} - \sigma Z \varepsilon_h)}{(1-\sigma)} \right] f(w,p) dw dp.
\end{aligned} \tag{7}$$

4 Interpretation of the FOCs

It is imperative to keep in mind that, as in most of the optimal tax literature, the four first order conditions (3), (5), (6), and (7) provide only implicit definitions of the optimal four tax and subsidy rates. Direct interpretation is difficult in this case, but some qualitative insights are possible.

We have already given an interpretation of the FOC for G, equation (3) - the lump-sum component should be adjusted to the point at which the marginal social valuation of an additional dollar, across all consumers, is equal to the marginal cost (one dollar). There is some reason to expect that the optimal lump-sum grant may be smaller than in the standard linear income tax problem, and perhaps negative. Suppose for instance that the wage distribution is degenerate, and that individuals differ *only* in their health prices. In this case, we would expect the optimum to be characterized by a lump-sum tax ($G < 0$) plus a positive marginal health expenditure subsidy ($s > 0$). To the extent that the optimum tax structure when both wages and health prices are heterogeneous is some combination of the two degenerate cases, it is possible (although not obvious) that $|G|$ could be relatively small. This intuition backfires however if the optimal health expenditure subsidy is negative. Now both income and health expenditures are tax sources, and they can feasibly finance a larger lump-sum benefit. (See the simulations below.)

Equations (5) and (6) for the optimal effective income tax rate and health subsidy rate can be compared to the standard condition for the optimal linear income tax rate in a two-good model. For example, in the absence of a health good, and when individuals only differ according to their wage rates, Dixit and Sandmo (1977) find that the optimal marginal income tax rate t satisfies the condition

$$\frac{t}{1-t} = \frac{-\mathrm{cov}[b, Y]}{\int Y \varepsilon_L dF} \tag{8}$$

where the variables have corresponding meanings to those in this paper (and where the integral is over the one-dimensional domain of wages). Assuming b and Y are negatively correlated, the tax rate should be higher the greater this negative correlation, and the less elastic is the supply of labor by those with high incomes.

If in our model we restrict attention to the optimal choice of the income tax and lump-sum grant by setting $s = \theta = 0$, then (5) becomes

$$\int_{\Omega} \left(\frac{\tau}{1-\tau}\right) Y \varepsilon_L dF = -\mathrm{cov}[b, Y]. \tag{9}$$

But $\tau = t + \theta Z = t$, so this condition is identical to (8), except for the range of integration. We can infer that if the pattern of health status increases the marginal social value of income of the poor relative to the rich (for example when health status and wages are negatively correlated), income tax rates should be correspondingly higher.

The cross elasticity terms in (5) and (6) mean that the simple (though implicit) expressions for the optimal effective tax and subsidy rates similar to (9) are not forthcoming. However, (5) can be interpreted as showing how the effective marginal income tax rates (which vary with health spending) are chosen to balance the positive welfare effects of reductions in the dispersion of income against the negative distortionary effects of the taxes, including their effects on labor supply and on health demand. Equation (6) has a corresponding interpretation. To explore this intuition a little further, let us set $\theta = 0$ and write (5) and (6) as

$$\frac{t}{1-t} = \frac{-\text{cov}[b, Y]}{\int [Y\varepsilon_L - \frac{s}{t} Z\varepsilon_{h\widehat{w}}]dF} \tag{10}$$

and

$$\frac{s}{1-s} = \frac{-\text{cov}[b, Z]}{\int [Z\varepsilon_h - \frac{t}{s} Y\varepsilon_{L\widehat{p}}]dF}. \tag{11}$$

In the standard linear income tax model, the thing that constrains the optimal tax rate to be less than one is the distortionary impact this has on labor supply. This effect is captured by the denominator in (8). If $s > 0$ and health and leisure are Hicksian substitutes ($\varepsilon_{h\widehat{w}} > 0$), condition (10) indicates that the net distortionary impact is smaller, so the tax rate t can be higher. When health expenditures are taxed at the optimum ($s < 0$), the income tax rate should be lower. On the other hand, if health expenditures are taxed, $s < 0$, and health and leisure are Hicksian complements ($\varepsilon_{h\widehat{w}} < 0$), the distortionary impact of the income tax is exacerbated, and the tax rate t should be lower.

Intuition for whether the health expenditure subsidy should be positive or negative can be obtained from equation (11). Indeed, if we constrain the income tax rate to be zero, (11) shows that the subsidy should have the same sign as the covariance between expenditures and the social marginal valuation of income, cov$[b, Z]$.[12] When the tax rate is positive, the optimal subsidy is (algebraically) increased as long as health and leisure are Hicksian complements (i.e., as long as health and labor supply are Hicksian substitutes, $\varepsilon_{L\widehat{p}} < 0$). That is, when health and leisure are Hicksian complements, if s is optimally positive without an income tax, it tends to increase when an income tax is available, but if s is optimally negative without an income tax, it tends to become smaller in absolute value when the income tax instrument is introduced. The reverse is true if health and leisure are Hicksian substitutes.[13]

[12] Recall $\varepsilon_h < 0$.

[13] This discussion is related to Gahvari's (1994) analysis of cash versus in-kind (e.g., health) transfers in the presence of a linear income tax. In that paper it was found that the labor supply effects of the two transfer mechanisms depend on whether leisure and the transfered good (health) are Hicksian substitutes or complements. In the current model, health is not transfered directly in-kind, but its optimal subsidization (or taxation) is affected by the same features of preferences.

Finally, equation (7) describes a similar trade-off between redistribution and distortionary costs of non-zero marginal tax rates. The first order condition is of the form $-\mathrm{cov}(YZ, b) - I = 0$, where I is the integral term on the right hand side of (7). To gain some intuition for the sign of θ at the optimum, one might start by evaluating this expression at $\theta = 0$. In the presence of income effects on the demand for health care, income and health spending are themselves likely to be more closely correlated than wages and health status, suggesting a relatively large negative covariance between their product and b. If this covariance is larger than the value of the integral term, then $\theta > 0$ is optimal. It is clearly difficult to be definitive about the relative values of these terms, but the fact that I is the difference between two components, might induce one to expect θ to be greater than zero.[14] If so, individuals with greater health expenditures would face higher effective marginal tax rates ($\tau = t + \theta Z$), and individuals with greater incomes would receive smaller effective marginal health expenditure subsidies ($\sigma = s - \theta Y$). In the simulations to follow, we find this to be the case.

5 Simulations

In this section we present a selection of simulation results that lend support to the intuition of the preceding paragraphs. The utility function is assumed to be of the simple form

$$u(c, h, l) = \ln(c) - \frac{1}{h} + \ln(l),$$

ensuring that health is a normal good, and that the share of income devoted to health care falls with wealth. In the simulations that follow, we choose a range of health prices ($p \in (0, 1)$) and wage rates ($w \in (10, 100)$), so that the income share of health care is about 12%. This is roughly the share of GDP spent on health care in the US, but somewhat higher than the corresponding share in other OECD countries. We assume that health prices and wages are jointly log-normally

[14] Note that

$$I = \left(\frac{\tau}{1-\tau}\varepsilon_L + \frac{\sigma}{1-\sigma}\varepsilon_{h\widehat{w}}\right) YZ - \left(\frac{\sigma}{1-\tau} Z^2 \varepsilon_{L\widehat{p}} + \frac{\tau}{1-\sigma} Y^2 \varepsilon_h\right).$$

distributed, conditional on falling in the ranges identified above.[15] When they are negatively correlated ($\rho < 0$ - see previous footnote), poor individuals tend to have greater health care needs. With these parameterizations, individual health spending as a share of earned income varies between about 5 percent (for those with high wages and low health needs) and 30 percent (for those with low wages and high health needs) across the population. The elasticity of demand for health care varies across individuals, but is approximately 0.5.

The government's net revenue requirement (R_0) is zero, and the welfare function is utilitarian ($\Psi(v) = v$). As a benchmark, with s and θ constrained to be zero, when $\rho = 0$ the optimal income tax rate is 30.1 percent. This tax system yields a welfare improvement over the no-tax outcome equivalent to a uniform lump-sum payment to all individuals equal to 4.86 percent of per capita GDP. An optimal linear income tax thus has a non-trivial impact on welfare.

Figure 1 presents simulation results for optimal income tax and health expenditure subsidy rates for the case in which θ is constrained to be zero. (That is, income is taxed at a constant marginal rate t, and health expenditures are subsidized at a constant marginal rate s.) As predicted earlier, both rates are decreasing in ρ, the correlation between the health price and the wage (recall that $\rho < 0$ corresponds to the more likely case, in which lower wage individuals have greater health needs - that is, face a higher price p of health). The optimal income tax rate remains in a narrow range between 29 and 31 percent, except when $\rho > 0.625$ (approximately). Health expenditures are subsidized at the optimum only when $\rho < -0.3$ (approximately): the subsidy reaches nearly 11 percent when $\rho = -0.925$, but it becomes a tax for values of ρ greater than -0.3. When $\rho = +0.925$, this tax is nearly 10 percent.

Two calculations however suggest, at least for the parameterization chosen above, that incor-

[15] In particular, the joint pdf is

$$f(\ln p, \ln w) = C \exp\left(\frac{x_p^2 - 2\rho x_p x_w + x_w^2}{(1-\rho)^2}\right)$$

where $x = (\ln p - \mu_p)/\sigma_p$, and $y = (\ln w - \mu_w)/\sigma_w$, and C is a constant of integration. The average health price is $0.37 (= p_{\min} + 0.3(p_{\max} - p_{\min})$, and μ_p is its log. Similarly, μ_w is the log of the average wage, which is assumed to be $37 (= w_{\min} + 0.3(w_{\max} - w_{\min})$. Both standard deviations are one: $\sigma_p = \sigma_w = 1$. The correlation coefficient ρ falls in the range $(-1, 1)$.

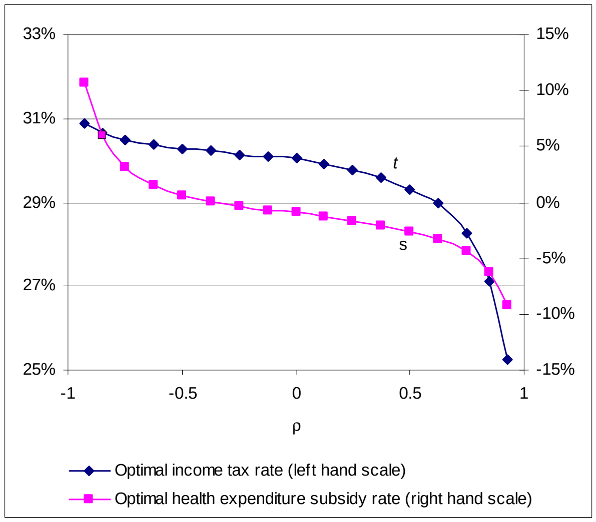

Figure 1: Optimal tax and subsidy rates as functions of the correlation between wages and health needs, in the case of a linear tax system, in which $\theta = 0$.

porating health expenditures into the tax system could have a limited impact. First, we compare the optimal income tax rates calculated above with those that would characterize an optimal linear tax system with s constrained to be zero. The results are shown in Figure 2, which confirms that the tax rates are virtually identical in the two cases, and differ discernibly only at extreme correlations, and even then by less than one percentage point.

Second, it is possible to calculate the increase in the government's budget (reduction in the revenue requirement, R) that would allow an optimal linear income tax without a health expenditure subsidy to generate the same level of welfare as the optimal linear tax system with health expenditures included. Figure 3 presents the results of these calculations, and shows that the welfare gains are very low indeed.

Finally, the change in the Lagrangian associated with a small increase in θ above zero is calculated, evaluated at the optimal linear income tax and health subsidy rates. The numerical values of these changes are not presented here, but it is reported that over the full range of correlations considered, the result was positive. Thus, introducing a small tax on the product of income and health expenditures is robustly welfare improving. Individuals who spend more on

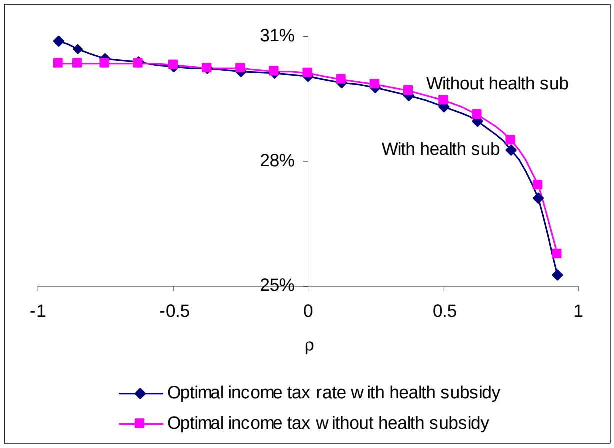

Figure 2: The effect on optimal income tax rates of incorporating health expenditures into a linear income tax system

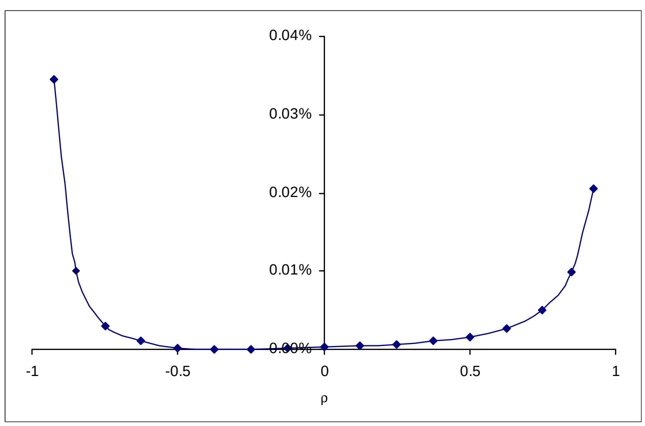

Figure 3: The welfare gain, measured as the equivalent budget increase as a percentage of GDP, from optimally incorporating health expenditures into a linear tax system

health care should thus face a higher marginal income tax rate, and conversely, those with higher incomes should receive a smaller marginal health expenditure subsidy.

6 Conclusions

A health expenditure subsidy could be a useful redistributive tool if sicker people spend more on health care than the healthy. On the other hand, high income people also spend more on health care than those with low incomes, so such a subsidy would tend to favor the rich ahead of the poor. Which of these two effects dominates in determining the appropriate tax treatment of health expenditures depends on the correlation between health status and income, and the elasticities of demand for health care and supply of labor. This paper has attempted to formalize the trade off within a model that is simple enough to analyze using standard techniques from the optimal (linear) income tax literature.

Optimal linear income tax rates are sensitive to the correlation between wages and health status in a predictable fashion: if low wage individuals have on average worse health status, then higher income tax rates are desirable. This can be thought of as stemming from the fact that the pattern of health status exacerbates the inequality associated with wage differentials, thereby increasing the social value of redistributing income from high wage earners to the lower paid. Formally, the result derives from the fact that such correlation increases the covariance between income and the social marginal value of income across individuals.

When a proportional health expenditure subsidy and an interaction effect are admitted, the first order conditions generalize those of the standard linear income tax model. Perhaps of more interest however are the observations that can be drawn from the numerical simulations presented. First, it is possible that in a system in which taxes paid are based only on income, the optimal tax rate, while responsive to the correlation between health status and wages, varies little over a broad range of such correlations. Second, despite this approximate invariance, the way in which health expenditures are optimally incorporated into the tax system can vary widely with the health status-wage correlation. In particular there is no presumption that health expenditures should be subsidized at the optimum - they may optimally be taxed. Third, as a corollary to the previous observation, health expenditures should be treated non-trivially in the tax system even if health status is uncorrelated with wages. This is simply because when health status and wages are uncorrelated, health expenditures and incomes are likely to be positively correlated. However, the quantitative significance of the impact of optimally treating health expenditures within a linear tax system appears to be modest at best, at least for the simulations reported here. Finally, the simulations suggest that any health expenditure subsidy should be reduced at the margin for higher-income individuals. This conclusion is perhaps not surprising, but it implies that individuals with greater health expenditures should face higher marginal income tax rates. This conclusion may be somewhat less intuitive.

7 Appendix

7.1 Non-working individuals

This appendix describes the set of non-working individuals given a vector of tax parameters $\Sigma = (t, s, \theta, G)$. In general these non-working individuals will constitute a set of positive measure, and will usually have either low wages, high health prices, or both. To see this note that given Σ, the consumer's problem is to

$$\max_{h,c,l} u(h, c, l) \text{ s.t. } \begin{cases} Z + c + T = Y \\ T = -G + tY - sZ + \theta Y Z \\ h \geq 0,\ c \geq 0,\ \text{and } 0 \leq l \leq 1 \end{cases}$$

Assuming infinite marginal utility of consumption of each good at zero, of the four inequality constraints only the last might bind. Substituting T into the first constraint, let ν and μ be the Lagrange multipliers for the budget constraint and the inequality constraint $l \leq 1$. The consumer's

Kuhn-Tucker conditions are then

$$\begin{aligned} u_h &= \nu p[1 - s + \theta w(1-l)] \\ u_c &= \nu \\ u_l &= \nu w[1 - t - \theta p h] + \mu \text{ and } \mu(1-l) = 0 \end{aligned}$$

If $l = 1$, h and c are independent of the wage rate, and are determined by

$$\begin{aligned} u_h(h, c, 1) &= p(1-s)u_c(h, c, 1) \\ \text{and } c &= G - (1-s)ph. \end{aligned}$$

Denote these values of health and general consumption by $\widehat{h}(p; \Sigma)$ and $\widehat{c}(p; \Sigma)$ respectively. The final first order condition is then

$$J(w, p; \Sigma) = w u_c(\widehat{h}, \widehat{c}, 1)[1 - t - \theta p h] - u_l(\widehat{h}, \widehat{c}, 1) = -\mu \leq 0$$

That is, all individuals with wages and health prices such that $J(w, p; \Sigma) \leq 0$ choose not to work. It is straightforward to show that the locus of points $w^*(p; \Sigma)$ defined by $J(w, p; \Sigma) = 0$ is upward sloping and that for all (w, p) such that $w < w^*(p; \Sigma)$ the individual does not work.

7.2 The Slutsky equation with a non-linear budget constraint

This appendix briefly validates the use of the Slutsky equation with a non-linear budget constraint. Around any consumption bundle (c_0, h_0, l_0), the budget constraint in the individual's maximization problem (1) can be linearized as

$$c = G' + \widehat{w}L - \widehat{p}h,$$

where $\widehat{w} = w(1 - t - \theta p h_0)$, $\widehat{p} = p(1 - s + \theta w L_0)$, $G' = c_0 - \widehat{w}L - \widehat{p}h$, and $c_0 = (G + w(1 - t)L_0 - p(1-s)h_0 + \theta w L_0 p h_0)$. The compensated demand for good $x = L, h$ is the optimal choice of $x(\widehat{w}, \widehat{p}; u_0)$ in the program

$$\min_{h,L} c - \widehat{w}L + \widehat{p}h \text{ s.t. } u(c, h, l) = u_0.$$

The compensated elasticity of demand with respect to price $\widehat{q} = \widehat{w}, \widehat{p}$, is just $\frac{\widehat{q}}{x}\frac{\partial x}{\partial \widehat{q}}$.

Locally, a change in the tax parameters, t, s, and θ affects the linearized budget constraint through changes in $\widehat{w}$ and $\widehat{p}$ in the obvious way. Such a change also affects the position of the budget plane (and not just its slope) through the term G'. However, this shift is irrelevant for the purposes of calculating the compensated demand functions, which only depend on the slope parameters (and u_0). Thus, for example, the compensated derivative of labor supply with respect to the tax instrument t is

$$\begin{aligned} \left.\frac{\partial L}{\partial t}\right|_u &= \left.\frac{\partial L}{\partial \widehat{w}}\right|_u . \frac{\partial \widehat{w}}{\partial t} \\ &= -w \left.\frac{\partial L}{\partial \widehat{w}}\right|_u \\ &= -\frac{Y_0}{\widehat{w}} \varepsilon_L . \end{aligned} \tag{12}$$

Similarly, the compensated derivative of labor supply with respect to the tax instrument θ is

$$\begin{aligned} \left.\frac{\partial L}{\partial \theta}\right|_u &= \left.\frac{\partial L}{\partial \widehat{w}}\right|_u . \frac{\partial \widehat{w}}{\partial \theta} + \left.\frac{\partial L}{\partial \widehat{p}}\right|_u . \frac{\partial \widehat{p}}{\partial \theta} \\ &= -wZ_0 \left.\frac{\partial L}{\partial \widehat{w}}\right|_u + pY_0 \left.\frac{\partial L}{\partial \widehat{p}}\right|_u \\ &= -\frac{Y_0 Z_0}{\widehat{w}} \varepsilon_L + p\frac{Y_0 L_0}{\widehat{p}} \varepsilon_{L\widehat{p}} \\ &= -L_0 \left(\frac{Z_0}{(1-\tau)} \varepsilon_L - \frac{Y_0}{(1-\sigma)} \varepsilon_{L\widehat{p}} \right), \end{aligned} \tag{13}$$

and the compensated derivative of health demand with respect to θ is

$$\begin{aligned} \left.\frac{\partial h}{\partial \theta}\right|_u &= -wZ_0 \left.\frac{\partial h}{\partial \widehat{w}}\right|_u + pY_0 \left.\frac{\partial h}{\partial \widehat{p}}\right|_u \\ &= -h_0 \left(\frac{Z_0}{(1-\tau)} \varepsilon_{h\widehat{w}} - \frac{Y_0}{(1-\sigma)} \varepsilon_h \right) \end{aligned} \tag{14}$$

Bibliography

References

Armstrong, Mark (1996): "Multiproduct Nonlinear Pricing," *Econometrica*, **64(1)**, pp. 51-75.

Atkinson, Anthony B. (1977): "Housing allowances, income maintenance, and income taxation," in *The Economics of Public Services*, M. Feldstein and R. Inman (eds.), Macmillan, London.

Atkinson, Anthony B. and Joseph E. Stiglitz (1976): "The design of tax structure: direct versus indirect taxation," *Journal of Public Economics*, **6**: 55-75.

_________ and ________ (1980): *Lectures on Public Economics*, McGraw-Hill

Blomqvist, Åke and Henrik Horn (1984): "Public Health Insurance and Optimal Income Taxation," *Journal of Public Economics*, **24(3)**, pp. 353-71.

Boadway, Robin, Manuel Leite-Monteiro, and Maurice Marchand (2002): "Social insurance and redistribution with moral hazard and adverse selection," mimeo, CORE, Université Catholique de Louvain.

Boskin, Michael J. and Joseph E. Stiglitz (1977): "Some Lessons from the New Public Finance," *American Economic Review*, **67(1)**: 295-301.

Case, Anne (2001): "Health income, and economic development," Center for Health and Wellbeing working paper, Woodrow Wilson School, Princeton.

Cremer, Helmuth, Jean-Marie Lozachmeur, and Pierre Pestieau (2002): "Social security, retirement age and optimal income taxation," mimeo GREMQ, University of Toulouse.

Cremer, Helmuth and Pierre Pestieau (1996): "Redistributive taxation and social insurance," *International Tax and Public Finance*, **3**, pp. 281-295.

Deaton, Angus (2001): "Health, inequality, and economic development," Center for Health and Wellbeing working paper, Woodrow Wilson School, Princeton.

Dixit, Avinash and Agnar Sandmo (1977): "Some simplified formulae for optimal income taxation," *Scandinavian Journal of Economics*, **79**: 417-423.

Feldstein, Martin (1973): "The welfare loss of excess health insurance," *Journal of Political Economy*, **81(2)**, pp. 251-80.

Gahvari, Firouz (1994): "In-kind transfers, cash grants and labor supply," *Journal of Public Economics*, **55(3)**, pp. 495-504.

Hylland, Aanund and Richard Zeckhauser (1979): "Distributional objectives should affect taxes but not program choice or design," *Scandinavian Journal of Economics*, **81(2)**: 264-284.

Mussa, Michael and Sherwin Rosen (1978): "Monopoly and Product Quality," *Journal of Economic Theory*, **18**, pp. 301-317.

Pauly, Mark V. (1986): "Taxation, Health Insurance, and Market Failure in Medical Care," *Journal of Economic Literature*, June, **24(3)**, pp. 629-75.

Pauly, Mark and Bradley Herring (2002): *Cutting Taxes for Insuring*, American Enterprise Institute Press, Washington DC.

Rochet, Jean-Charles (1991): "Incentives, Redistribution and Social Insurance," *The Geneva Papers on Risk and Insurance Theory*, **16(2)**: 143-165.

Rochet, Jean-Charles and Philippe Choné (1998): "Ironing, Sweeping, and Multi-dimensional Screening," *Econometrica*, **66(4)**: 783-826.

Sheshinski, Etyan (1972): "The optimal linear income tax." *Review of Economic Studies*, **39**: 297-302

Stiglitz, Joseph E. and Michael Boskin (1977): "Some lessons from the new public finance," *American Economic Review*, **67**: 295-301.

Wilson, Robert B. (1993): *Nonlinear Pricing*, Oxford University Press