Income Risk and Progressive Optimal Taxation*

Patrick A. Pintus† Alain Trannoy‡

Preliminary and VERY incomplete draft (please do NOT circulate). Comments are most welcome.

March 14, 2008

Abstract: We introduce productivity risk in Diamond's [1] formulation of Mirrlees's [5] optimal taxation model with quasi-linear utility. Under multiplicative shocks to productivity, we show that the optimal marginal tax rate can be approximated, with any arbitrarily high level of accuracy, by the solution of a Cauchy-Euler differential equation. This solution takes the simple form of a weighted sum of power functions. We give sufficient conditions such that the approximated marginal tax rate is either an increasing or a U-shaped function of income. In particular, increasing marginal tax rates obtain if the distribution of the productivity shocks has a large enough kurtosis and if one imposes that marginal tax rates should be positive (respectively negative) for large (respectively small) incomes. Finally, we show that with income risk, marginal tax rates generally do not vanish either under a Utilitarian criterion or at the top and bottom of the income distribution.

Keywords: Optimal Income Taxation, Income Risk, Income Tax Progressivity.

Journal of Economic Literature Classification Numbers: H21, H24.

* The first draft of this paper was written while Pintus was visiting the Research Division of the Federal Reserve Bank of St. Louis, whose hospitality and financial support are gratefully acknowledged. First version: september 2007.

† Université de la Méditerranée (GREQAM-IDEP) and Federal Reserve Bank of St. Louis. E-mail: pintus@univ-aix.fr.

‡ EHESS and GREQAM-IDEP. E-mail: Alain.Trannoy@eco.u-cergy.fr.

1 Introduction

We introduce productivity risk in Diamond's [1] formulation of Mirrlees's [5] optimal taxation model with quasi-linear utility. Such a risk makes income uncertain when effort is assumed to be chosen before the realization of the productivity shock. Under multiplicative shocks to productivity, we show that the optimal marginal tax rate can be approximated, with any arbitrarily high level of accuracy, by the solution of a Cauchy-Euler differential equation. This solution takes the simple form of a weighted sum of power functions. We give sufficient conditions such that the approximated marginal tax rate is either an increasing or a U-shaped function of income. In particular, increasing marginal tax rates obtain if the distribution of the productivity shocks has a large enough kurtosis and if one imposes that marginal tax rates should be positive (respectively negative) for large (respectively small) incomes. Finally, we show that with income risk, marginal tax rates generally do not vanish either under a Utilitarian criterion or at the top and bottom of the income distribution.

As in Diamond [1] and Piketty [7], we assume quasi-linear utility (which is assumed away by most papers in the literature on new dynamic optimal taxation that is reviewed by Golosov *et al.* [2] and Kocherlakota [3]). In contrast to the approach of the problem proposed in Mirrlees [6], we do not assume affine taxation and we characterize incentive-compatible tax schedules. We also extend the analysis to incorporate risk aversion.

2 The Model

Continuum of households with private information on their type $\omega \in [\underline{\omega}, \overline{\omega}]$, with $\underline{\omega} > 0$. Each agent has quasi-linear utility $c - v(l)$, where $c > 0$ is consumption and $l > 0$ is effort at work, with $v'(l) > 0$ and $v''(l) > 0$.

Household's effective productivity is $\varepsilon\omega$, that is, the product of two components. Interpretation: ω reflects exogenous "skill" while the random variable ε is an ex-post productivity disturbance that has (small enough) support $[\underline{\varepsilon}, \overline{\varepsilon}]$ with unit mean and $\underline{\varepsilon} \geq 0$, and the distribution of which does not depend on effort ("luck"),. In other words, there is some uncertainty on how markets reward effort. Observed income $\tilde{y} \equiv \varepsilon\omega l > 0$ depends both on skill and on luck. Therefore, agents differ both *ex-ante*, as they have different skills, and *ex-post*, because they may experience different realizations of the productivity shock. Alternatively, ε can be interpreted as a forecasting error.

Densities and cumulative distributions: $f(\omega)$ and $F(\omega)$ for skill, $g(\varepsilon)$ and $G(\varepsilon)$ for productivity shock. Both are continuous. We assume that $\int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon dG(\varepsilon) = 1$.

After-tax income equals consumption $c = \tilde{y} - t[\tilde{y}]$, where $t[\tilde{y}]$ is the tax schedule that is set by the government and taken as given by each household.

Each type-ω household chooses effort l so as to maximize expected utility, defined as:

$$V \equiv \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \{\varepsilon\omega l - t[\varepsilon\omega l] - v(l)\} dG(\varepsilon).$$

FOC of this problem is :

$$v'(l) = \omega[1 - \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon t'[\varepsilon\omega l] dG(\varepsilon)]. \tag{1}$$

Equation (1) defines optimal effort $l(\omega)$ and, therefore, consumption is $c(\omega,\varepsilon) \equiv \varepsilon y(\omega) - t[\varepsilon y(\omega)]$, with $y(\omega) \equiv \omega l(\omega)$. Also $V(\omega) \equiv \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \{\varepsilon\omega l(\omega) - t[\varepsilon\omega l(\omega)] - v(l(\omega))\} dG(\varepsilon)$ for all $\omega \in [\underline{\omega}, \overline{\omega}]$. Note that the second-order conditions are:

$$v''(l) + \omega^2 \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon^2 t''[\varepsilon\omega l] dG(\varepsilon) \geq 0, \tag{2}$$

for all types ω. We prove, in Proposition 3.2 to be presented later on, that the optimal tax schedule satisfies conditions (2).

Government does observe neither effort nor skill. However, agents report their income to the government, whose objective is to design an incentive-compatible tax schedule that maximizes social welfare from an *ex-ante* point of view, that is:

$$\int_{\underline{\omega}}^{\overline{\omega}} W[\int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \{c(\omega,\varepsilon) - v(l(\omega))\} dG(\varepsilon)] dF(\omega), \tag{3}$$

where $W'[x] > 0 > W''[x]$ is a social welfare function that exhibits inequality aversion.

Timing: household learns his type ω; government proposes tax schedule; household chooses labor effort, taking into account taxes; productivity shocks are realized and post-tax income distributed.

2.1 Incentive-Compatible Tax Schedules

Define:

$$\hat{V}(\omega', \omega) \equiv \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \{c(\omega', \varepsilon) - v(\frac{y(\omega')}{\omega})\} dG(\varepsilon)$$

as the expected utility of a type-ω household when he reports a type ω' to the government. Therefore, one has that $V(\omega) = \hat{V}(\omega, \omega)$ by definition. By the Revelation Principle, incentive-compatible tax schedules ensure both truth-telling and optimal (second-best) effort.

Lemma 2.1 (Incentive-Compatibility Constraints)

Suppose that effort disutility is strictly increasing and strictly convex, that is, $v'(l) > 0$ *and* $v''(l) > 0$ *for all* $l > 0$. *Then any incentive-compatible tax schedule satisfies the following first-order conditions:*

$$V'(\omega) = l(\omega)v'(l(\omega))/\omega, \text{ for all types } \omega \in [\underline{\omega}, \overline{\omega}]. \tag{4}$$

If, in addition, $y'(\omega) > 0$ *for all* $\omega \in [\underline{\omega}, \overline{\omega}]$, *then the truth-telling second-order conditions hold.*

Proof: Under the Extended Revelation Principle (see Laffont and Martimort [4, p. 274]), we can restrict the analysis to mechanisms that are truth-telling, that is, such that the following condition is met:

$$\omega = \arg\max_{\omega'} \hat{V}(\omega', \omega) \tag{5}$$

where $\hat{V}(\omega', \omega) \equiv \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \{c(\omega', \varepsilon) - v(y(\omega')/\omega)\} dG(\varepsilon)$.

FOC of problem in equation (5) is $\frac{\partial V}{\partial \omega'}(\omega, \omega) = 0$, that is:

$$d_{\omega'} \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} c(\omega', \varepsilon) dG(\varepsilon) = d_{\omega'} v(\frac{y(\omega')}{\omega})$$

at $\omega' = \omega$, where the notation d_x stands for the total derivative with respect to x. This can be rewritten as:

$$\int_{\underline{\varepsilon}}^{\overline{\varepsilon}} d_{\omega} c(\omega, \varepsilon) dG(\varepsilon) = v'(l(\omega))[l'(\omega) + l(\omega)/\omega]$$

But taking the derivative of $V(\omega)$ with respect to ω yields:

$$V'(\omega) = \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} d_\omega c(\omega, \varepsilon) dG(\varepsilon) - l'(\omega) v'(l(\omega))$$

or, combining with the latter equality, equation (4) in the lemma.

SOC of problem in (5) is $\frac{\partial^2 V}{\partial \omega' \partial \omega}(\omega, \omega) \geq 0$, that is:

$$\frac{y'(\omega)}{\omega^2}[v'(l) + v''(l)\frac{y(\omega)}{\omega}] \geq 0. \tag{6}$$

Under the assumptions that both $v'(l) > 0$ and $v''(l) > 0$ for all $l > 0$, and $y'(\omega) > 0$ for all $\omega \in [\underline{\omega}, \overline{\omega}]$, condition (6) is met. It is well known that under the assumptions of both linear consumption utility and increasing-convex disutility from effort, the Spence-Mirrlees condition is met. To get this result, define expected utility as a function of the type, before-tax income, and disposable income, that is:

$$\tilde{V}(\omega, y, c) = \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \{c(\cdot) - v(\frac{y(\cdot)}{\omega})\} dG(\varepsilon).$$

Then one has that $\partial_c \tilde{V}(\cdot) = 1$ so that $\partial_\omega(\partial_y \tilde{V}(\cdot)/\partial_c \tilde{V}(\cdot)) = \partial^2_{\omega y}\tilde{V}(\cdot)$, where:

$$\partial^2_{\omega y}\tilde{V}(\cdot) = \frac{1}{\omega^2}[v'(l) + v''(l)\frac{y(\omega)}{\omega}] > 0$$

when $v'(l) > 0$ and $v''(l) > 0$, so that the Spence-Mirrlees condition holds. Therefore, our assumptions on $v(l)$ implies both that the Spence-Mirrlees condition is met and that $y'(\omega) > 0$ for all $\omega \in [\underline{\omega}, \overline{\omega}]$ is the SOC (6) of the incentive problem. □

2.2 Optimal Marginal Tax Rates

This section derives, by following closely Diamond [1, pp. 93-94], the optimal marginal tax rates.

Proposition 2.1 (Optimal Marginal Tax Rates)

Under the assumptions of Lemma 2.1, optimal marginal tax rates are characterized by the following conditions:

$$\frac{\int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon t'[\varepsilon\omega l(\omega)]dG(\varepsilon)}{1-\int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon t'[\varepsilon\omega l(\omega)]dG(\varepsilon)} = \{\frac{1-F(\omega)}{\omega f(\omega)}\}\{1-\frac{D(\omega)}{D(\underline{\omega})}\}\{\frac{1}{\epsilon_l}+1\}, \tag{7}$$

for all types $\omega \in [\underline{\omega}, \overline{\omega}]$, where $D(\omega) \equiv [\int_{\omega}^{\overline{\omega}} W'[V(\omega)]f(\omega)d\omega]/[1-F(\omega)]$, and ϵ_l is the elasticity of labor supply with respect to the net wage.

Proof: The problem that the government faces is to maximize social welfare (3) under the incentive constraints (4). We further impose that budget be balanced, that is, $\int_{\underline{\omega}}^{\overline{\omega}} \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} t[\varepsilon\omega l(\omega)]dG(\varepsilon)dF(\omega) = 0$. This budget constraint can be written, using the definition of expected utility $V(\omega) = \int_{\underline{\varepsilon}}^{\overline{\varepsilon}}\{\varepsilon\omega l(\omega) - t[\varepsilon\omega l(\omega)]\}dG(\varepsilon) - v(l(\omega))$, with $\int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon dG(\varepsilon) = 1$, as:

$$\int_{\underline{\omega}}^{\overline{\omega}} \{y(\omega) - v(l(\omega)) - V(\omega)\}dF(\omega) = 0.$$

The problem is then to:

$$\text{maximize} \int_{\underline{\omega}}^{\overline{\omega}} W[V(\omega)]dF(\omega) \tag{8}$$

subject to the incentive constraints:

$$V'(\omega) = l(\omega)v'(l(\omega))/\omega, \text{ for all types } \omega \in [\underline{\omega}, \overline{\omega}], \tag{9}$$

and the balanced-budget requirement:

$$\int_{\underline{\omega}}^{\overline{\omega}} \{y(\omega) - v(l(\omega)) - V(\omega)\}dF(\omega) = 0. \tag{10}$$

Define $\lambda(\omega)$ and μ as the multiplier of, respectively, (9) and (10). Forming the Hamiltonian:

$$H(l(\omega), V(\omega)) \equiv W[V(\omega)]f(\omega) + \lambda(\omega)[l(\omega)v'(l(\omega))/\omega] + \mu[y(\omega) - v(l(\omega)) - V(\omega)]f(\omega), \tag{11}$$

and using Pontryagin's Principle, one gets the following first-order conditions. First, $\partial_{l(\omega)} H(l(\omega), V(\omega)) = 0$ is:

$$\lambda(\omega)[v'(\omega) + l(\omega)v''(\omega)]/\omega + \mu f(\omega)[\omega - v'(l(\omega))] = 0. \tag{12}$$

Second, $-\partial_{V(\omega)} H(l(\omega), V(\omega)) = \lambda'(\omega)$ is:

$$f(\omega)(\mu - W'[V(\omega)]) = \lambda'(\omega). \tag{13}$$

But $\int_{\underline{\omega}}^{\overline{\omega}} \lambda'(\omega) dF(\omega) = \lambda(\overline{\omega}) - \lambda(\underline{\omega}) = 0$. Therefore, (13) can be rewritten as:

$$\mu = \int_{\underline{\omega}}^{\overline{\omega}} W'[V(\omega)]f(\omega)d\omega. \tag{14}$$

Moreover, $\int_{\omega}^{\overline{\omega}} \lambda'(\omega) dF(\omega) = -\lambda(\omega)$ or, using (13) and (14):

$$\lambda(\omega) = \int_{\omega}^{\overline{\omega}} W'[V(\omega)]f(\omega)d\omega - \mu[1 - F(\omega)].$$

The latter equality can be written, by defining a new function:

$$D(\omega) \equiv [\int_{\omega}^{\overline{\omega}} W'[V(\omega)]f(\omega)d\omega]/[1 - F(\omega)],$$

so that $\mu = D(\underline{\omega})$, as:

$$\lambda(\omega) = [1 - F(\omega)][D(\omega) - D(\underline{\omega})] \tag{15}$$

Note then that $\mu = D(\underline{\omega})$, which, if combined with (15), allows to write (12) as:

$$\frac{\omega - v'(l(\omega))}{v'(l(\omega))} = \{\frac{1 - F(\omega)}{f(\omega)}\}\{1 - \frac{D(\omega)}{D(\underline{\omega})}\}\{\frac{1}{\epsilon_l} + 1\}, \tag{16}$$

where ϵ_l denotes the elasticity of labor supply with respect to the net wage. Recalling equation (1), one derives that:

$$\frac{\omega - v'(l(\omega))}{v'(l(\omega))} = \frac{\int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon t'[\varepsilon\omega l(\omega)]dG(\varepsilon)}{1 - \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon t'[\varepsilon\omega l(\omega)]dG(\varepsilon)}. \tag{17}$$

Finally, using the latter relation, equation (16) is written in the following form, which characterizes optimal marginal tax rates:

$$\frac{\int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon t'[\varepsilon\omega l(\omega)]dG(\varepsilon)}{1 - \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon t'[\varepsilon\omega l(\omega)]dG(\varepsilon)} = \{\frac{1 - F(\omega)}{\omega f(\omega)}\}\{1 - \frac{D(\omega)}{D(\underline{\omega})}\}\{\frac{1}{\epsilon_l} + 1\}, \tag{18}$$

as stated in equation (7) of the proposition. □

Equation (7) can be rewritten as:

$$\int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon t'[\varepsilon\omega l(\omega)]dG(\varepsilon) = \frac{\Lambda}{1 + \Lambda}, \tag{19}$$

where $\Lambda \equiv \{\frac{1-F(\omega)}{\omega f(\omega)}\}\{1 - \frac{D(\omega)}{D(\underline{\omega})}\}\{\frac{1}{\epsilon_l} + 1\}$ is the very same term that appears in Diamond's [1] characterization of optimal marginal tax rates when there is no income risk. In particular, if the planner adopts a Utilitarian criterion, then $\Lambda = 0$ and all agents face a zero marginal tax rate in the absence of productivity risk. The next section shows that this implication does not generally hold when risk is introduced.

3 Income Risk and Optimal Marginal Tax Rates

3.1 Approximating the Optimal Marginal Tax Rates

In this section, we provide a way to compute and approximate the integral term that appears in the left-hand side of equation (7). This integral is obviously the sum of the two terms: the average marginal tax rate and the covariance between the random shock and the marginal rate. It turns out to be more convenient to keep the integral and to define $h(\varepsilon) \equiv \varepsilon t'[\varepsilon y]$. If h is assumed to be analytic on $[\underline{\varepsilon}, \overline{\varepsilon}]$ (as we show below to be the case), then a Taylor expansion around the unit mean gives us that:

$$h(\varepsilon) = \sum_{n=0}^{\infty} h^{(n)}(1)\frac{(\varepsilon - 1)^n}{n!} \text{ for all } \varepsilon \in [\underline{\varepsilon}, \overline{\varepsilon}] \tag{20}$$

where $h^{(n)}(x) \equiv d^n h(x)/dx^n$ denotes the n-th order derivative. From the definition of h, we get that $h^{(n)}(1) = ny^{n-1}t^{(n)}[y] + y^n t^{(n+1)}[y]$, so that equation (20) is:

$$h(\varepsilon) = \sum_{n=0}^{\infty} \{ny^{n-1}t^{(n)}[y] + y^n t^{(n+1)}[y]\}\frac{(\varepsilon - 1)^n}{n!}. \tag{21}$$

Therefore, we conclude that:

$$\int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon t'[\varepsilon y] dG(\varepsilon) = \sum_{n=0}^{\infty} \{ny^{n-1}t^{(n)}[y] + y^n t^{(n+1)}[y]\}\frac{\mu_n}{n!}, \tag{22}$$

where $\mu_n \equiv \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} (\varepsilon - 1)^n dG(\varepsilon)$ is the n-th central moment of the probability distribution associated with G, with $\mu_0 = 1$ and $\mu_1 = 0$. Consequently, the characterization of optimal marginal tax rates in equation (7) can conveniently be written as:

$$\sum_{n=0}^{\infty} \{ny^{n-1}t^{(n)}[y] + y^n t^{(n+1)}[y]\}\frac{\mu_n}{n!} = \frac{\Lambda}{1 + \Lambda}$$

or, equivalently:

$$\sum_{n=0}^{\infty} \{y^n t^{(n+1)}[y]\}\{\frac{\mu_n + \mu_{n+1}}{n!}\} = \frac{\Lambda}{1 + \Lambda}. \tag{23}$$

We have therefore proved the following statement.

Proposition 3.1 (Optimal Marginal Tax Rates II)

Under the assumptions of Proposition 2.1, the optimal marginal tax rates in equation

(7) are equivalently characterized by:

$$\sum_{n=0}^{\infty}\{y^n t^{(n+1)}[y]\}\{\frac{\mu_n+\mu_{n+1}}{n!}\}=\frac{\Lambda}{1+\Lambda}, \tag{24}$$

where $\Lambda \equiv \{\frac{1-F(\omega)}{\omega f(\omega)}\}\{1-\frac{D(\omega)}{D(\underline{\omega})}\}\{\frac{1}{\epsilon_l}+1\}$ *and* $\mu_n \equiv \int_{\underline{\varepsilon}}^{\overline{\varepsilon}}(\varepsilon-1)^n dG(\varepsilon)$ *is the n-th central moment of the income shock distribution G, with* $\mu_0=1$ *and* $\mu_1=0$.

Now we take advantage of the characterization in equation (24) of Proposition 3.1 to compute an approximation of marginal tax rates that allows us to derive analytically the entire tax function. To this end, we state the following:

Definition 3.1 (n-th Order Approximation)

Define $x(y) \equiv t'[y]$. *Then given any integer* $n \geq 1$, *an n-th order approximation of* $\int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon t'[\varepsilon\omega l(\omega)]dG(\varepsilon)$ *is given by* $\sum_{i=0}^{n}\{y^i x^{(i)}(y)\}\{\mu_i+\mu_{i+1}\}/i!$, *with* $\mu_{n+1}=0$.
Then the solution of $\sum_{i=0}^{n}\{y^i x^{(i)}(y)\}\{\mu_i+\mu_{i+1}\}/i! = \Lambda/(1+\Lambda)$, *with* $\mu_{n+1}=0$, *defines an n-approximate optimal marginal tax rate function.*

Finally, we are led to the following main result:

Theorem 3.1 (n-Approximate Optimal Marginal Tax Rates)

Under the assumptions of Proposition 2.1, an n-approximate optimal marginal tax function is the solution of the following Cauchy-Euler differential equation:

$$\sum_{i=0}^{n}\{y^i x^{(i)}(y)\}\{\frac{\mu_i+\mu_{i+1}}{i!}\}=\frac{\Lambda}{1+\Lambda}, \text{ with } \mu_{n+1}=0. \tag{25}$$

If, in addition, Λ *is constant for all* $\omega \in [\underline{\omega},\overline{\omega}]$, *then the n-approximate function is given* *by:*

$$t'[y] = \sum_{i=1}^{n} \alpha_i y^{\gamma_i} + \frac{\Lambda}{1+\Lambda}, \tag{26}$$

where the α_i's are real numbers and the γ_i's are the real roots of the following n-th order polynomial:

$$p(\gamma) \equiv 1 + \sum_{i=1}^{n} \{\frac{\mu_i + \mu_{i+1}}{i!}\}\{\prod_{j=1}^{i} \gamma + 1 - j\}, \text{ with } \mu_{n+1} = 0. \tag{27}$$

Proof: From Definition 3.1, we know that an n-th order approximation of equation (7) is:

$$\sum_{i=0}^{n} \{y^i x^{(i)}(y)\}\{\frac{\mu_i + \mu_{i+1}}{i!}\} = \frac{\Lambda}{1+\Lambda}, \text{ with } \mu_{n+1} = 0, \tag{28}$$

which is an n-th order Cauchy-Euler differential equation (that is, it involves the sum over $i = 0, 1, \cdots, n$ of terms that are of the form $\beta_i y^i x^{(i)}(y)$, where the β_i's are real numbers). It is well known that the solutions of such differential equations are obtained by looking first at the associated, homogeneous equation and by guessing that $x(y) = y^\gamma$, which, once plugged into equation (28), gives:

$$y^\gamma[1 + \sum_{i=1}^{n} \{\frac{\mu_i + \mu_{i+1}}{i!}\}\{\prod_{j=1}^{i} \gamma + 1 - j\}] = 0, \text{ with } \mu_{n+1} = 0. \tag{29}$$

Then if $y > 0$, the solution is known when we know the roots of the n-th order polynomial $p(\gamma) \equiv 1 + \sum_{i=1}^{n} \{\frac{\mu_i + \mu_{i+1}}{i!}\}\{\prod_{j=1}^{i} \gamma + 1 - j\}$, with $\mu_{n+1} = 0$. If the γ_i's are real roots of p, then the solution of the nonhomogeneous equation (28) is obtained by summing the solution to the homogeneous counterpart and a particular solution, which is easily found here to be the constant $\Lambda/(1+\Lambda)$. Then the solution is $x(y) = t'[y] = \sum_{i=1}^{n} \alpha_i y^{\gamma_i} + \Lambda/(1+\Lambda)$, as expressed in equation (26). $\square$

Finally, the next proposition proves that the n-approximate tax schedule satisfies the SOC in (2).

Proposition 3.2

Under the assumptions of Theorem 3.1, the second-order conditions of effort maximization are met.

Proof: Equations (26) imply that $t''[y] = \sum_{i=1}^{n} \alpha_i \gamma_i y^{\gamma_i - 1}$. Therefore, SOC (2) can be written as:

$$v''(l) + \sum_{i=1}^{n} \alpha_i \gamma_i \omega^{\gamma_i + 1} l^{\gamma_i - 1} \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon^{\gamma_i + 1} dG(\varepsilon) \geq 0 \tag{30}$$

for all types ω. To prove that the above inequality is satisfied under our assumptions, we now prove that the second term in the left-hand side of (30) vanishes. To that end, we take a 4-th order expansion of $\varepsilon^{\gamma+1}$ (which is an analytic function of ε) around the unit mean and we get that:

$$\int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon^{\gamma+1} dG(\varepsilon) = 1 + (\gamma + 1) \sum_{i=1}^{n} \{\frac{\mu_{i+1}}{i+1!}\}\{\prod_{j=1}^{i} \gamma + 1 - j\}. \tag{31}$$

Finally, developing and identifying terms shows that $1+(\gamma+1)\sum_{i=1}^{n}\{\frac{\mu_{i+1}}{i+1!}\}\{\prod_{j=1}^{i}\gamma+1-j\}$ indeed equals the polynomial $p(\gamma)$ defined in (27). Because the γ_i's involved in (30) are roots of p, then one has that $\int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon^{\gamma_i+1} dG(\varepsilon) = 0$ for all i. Therefore, this implies that condition (30) is met under our assumption that $v''(l) > 0$ for all $l > 0$. $\square$

3.2 Increasing and U-shaped Patterns of Optimal Marginal Tax Rates

So as to underline the effect of income risk, we now take the example of a Utilitarian planner, which implies that $\Lambda = 0$, as is well known (e.g. Salanié [9, p. 95]). When there is no risk (i.e. $\varepsilon = 1$), equation (7) (or, for that matter, equation (24) with $\mu_n = 0$ for all $n \geq 1$) indicates that all agents uniformly face a zero marginal tax rate. We now show that this conclusion does not generally hold in the presence of risk. In particular, we consider a 4-th order approximation, which is acceptable provided that the support $[\underline{\varepsilon}, \overline{\varepsilon}]$ is small enough. We show that under mild assumptions on the *signs* of the marginal tax rates at both extremes of the income distribution, the *shape* of the marginal tax rate is expected to be either increasing (that is, progressive) or U-shaped.

Two reasons account for choosing a 4-th order approximation. Most importantly, the marginal tax rate then depends on the first four central moments of the shock distribution, so that the discussion is cast in terms of the variance, skewness and kurtosis. For better approximations, the involved higher-order moments are less easily interpreted. In addition, it is well known (from the Abel-Ruffini theorem), that there is no "simple" formula for n-th order polynomial roots when $n \geq 5$. In contrast, restricting the analysis to $n = 4$ gives rise to a quartic equation for which formulas do exist.

Proposition 3.3 (A 4-th Order Approximation with Utilitarian Criterion)

Assume that the planner uses a Utilitarian criterion (so that $\Lambda = 0$) and consider, following Theorem 3.1, a 4-th order approximation. Then:

$$p(\gamma) = \frac{\mu_4}{24}\gamma^4 + (\frac{\mu_3}{6} - \frac{\mu_4}{12})\gamma^3 + (\frac{\mu_2}{2} - \frac{\mu_4}{24})\gamma^2 + (\frac{\mu_2}{2} - \frac{\mu_3}{6} + \frac{\mu_4}{12})\gamma + 1 \qquad (32)$$

If the kurtosis μ_4/μ_2^2 of the income shock distribution is large enough, then $p(\gamma)$ has two negative real roots and two positive real roots.

It follows that $t'[y] = \sum_{i=1}^{4} \alpha_i y^{\gamma_i}$ is a 4-approximate optimal marginal tax rate function, where the α_i's are real numbers and the γ_i's are the real roots of $p(\gamma)$.

Proof: The polynomial $p(\gamma) = \frac{\mu_4}{24}\gamma^4 + (\frac{\mu_3}{6} - \frac{\mu_4}{12})\gamma^3 + (\frac{\mu_2}{2} - \frac{\mu_4}{24})\gamma^2 + (\frac{\mu_2}{2} - \frac{\mu_3}{6} + \frac{\mu_4}{12})\gamma + 1$ is easily obtained, from equation (27), by developing and arranging terms. The roots of $p(\gamma)$ are shown to tend to $-1, 0, 1$ and 2 when μ_4 tends to ∞ while the other moments μ_2 and μ_3 are held constant. Because the product of the root is given by $24/\mu_4$ (by the Viète's formula), which is positive, by continuity $p(\gamma)$ has four real roots, two positive and two negative, when μ_4 is large enough. Given the variance μ_2, this means that the kurtosis μ_4/μ_2^2 has to be large enough. Then the resulting function $t'[y] = \sum_{i=1}^{n} \alpha_i y^{\gamma_i}$ solves the corresponding 4-th order Cauchy-Euler differential equation and is therefore a 4-approximate optimal marginal tax rate function. □

The foregoing proposition shows that if the distribution of productivity shocks is leptokurtic, with high kurtosis and "fat tails", then p has two pairs of positive and negative real roots. This result applies to distributions that are symmetric ($\mu_3 = 0$) or asymmetric ($\mu_3 \neq 0$) so that it is independent of the sign of the skewness $\mu_3/\mu_2^{3/2}$. We show next that under additional mild assumptions, the resulting pattern of marginal tax rates is either U-shaped or progressive. What remains to be determined in the formula given in Proposition 3.3 (or, for that matter, in (26)) are the α_i's. It seems reasonable to make further assumptions on the *signs* of the marginal tax rates that are acceptable by society at both extremes of the income distribution. First of all, let us suppose that the marginal tax rate should be non-negative at "large" levels of income. If we define and order (without loss of generality) the positive roots of p as $\gamma_1 > \gamma_2 > 0$, then this implies that α_1 and α_2 cannot be both negative. In particular, a sufficient condition for $\lim_{y\to\infty} \sum_{i=1}^{2} \alpha_i y^{\gamma_i} > 0$ is that both $\alpha_1 > 0$, $\alpha_2 > 0$. The latter assumption is stronger than needed and it accordingly reflects the more demanding requirement that marginal tax rates should be positive not only in the limit, but also at large levels of income. Symmetrically, assume that the marginal tax rate should be non-positive at "small" levels of income. Then if $0 > \gamma_3 > \gamma_4$ are the negative roots of p, one cannot have α_3 and α_4 both positive. However, imposing $\alpha_3 < 0$ and $\alpha_4 < 0$ is sufficient for this purpose. Such a stronger condition is interpreted as the fact that marginal tax rates should be non-positive when income is in a right neighborhood of zero, and not only in the limit, when income tends to zero. We can summarize these rather elementary observations in the following statement.

Proposition 3.4 (Increasing Optimal Marginal Tax Rates)

Under the assumptions of Proposition 3.3, let $\gamma_1 > \gamma_2 > 0 > \gamma_3 > \gamma_4$ be the roots of $p(\gamma)$. Moreover, assume that the arginal tax rates are imposed to be positive (respectively negative) for large enough (respectively small enough) incomes, that is, $\alpha_1 > 0$, $\alpha_2 > 0$, $\alpha_3 < 0$, $\alpha_4 < 0$. Then the 4-approximate optimal marginal tax rate $t'[y] = \sum_{i=1}^{4} \alpha_i y^{\gamma_i}$ is strictly increasing for all levels of income $y > 0$.

Proof: It is immediate to show that $t''[y] > 0$ for all $y > 0$ under the assumptions of the proposition. In addition, it is straightforward to show the marginal tax rate is positive (respectively negative) when y tends to ∞ (respectively to 0) only if $\alpha_1 > 0$ (respectively $\alpha_4 < 0$). Finally, the thresholds above (respectively below) which the marginal tax rate is positive (respectively negative) is shown to decrease with α_2 (respectively α_3). $\square$

It is easy to generalize the above result to the case where Λ is positive and constant for all $\omega \in [\underline{\omega}, \overline{\omega}]$. In that case, the theory would predict that, absent risk, agents face a flat pattern of marginal tax rates. However, adding risk dramatically changes the conclusions. In particular, income risk implies that increasing marginal tax rates, as observed in many countries, may well approximate the optimal tax schedule.

If the conditions in Proposition 3.4 are not satisfied, then U-shaped patterns may arise, as we now show.

[TO BE WRITTEN]

3.3 Top and Bottom n-Approximate Optimal Marginal Tax Rates

It is well known that, in the absence of risk, zero marginal tax rates obtain at both extremes of the skill distribution $[\underline{\omega}, \overline{\omega}]$. To see this, consider (7). One has that $\Lambda \equiv \{\frac{1-F(\omega)}{\omega f(\omega)}\}\{1 - \frac{D(\omega)}{D(\underline{\omega})}\}\{\frac{1}{\epsilon_l} + 1)\}$ is zero when either $\omega = \underline{\omega}$ (because the second bracketed term in the expression of Λ equals zero) or $\omega = \overline{\omega}$ (because the first term in parentheses vanishes). Therefore, one concludes from (7) that the marginal tax rate is then zero when there is no risk. However, with income risk the marginal tax rates generally do not vanish at the top and bottom of the type distribution. This is the case, for instance, in the example of section 3.2, where a large kurtosis implies that the marginal rates at the minimum and maximum incomes are not zero.

Corollary 3.1 (Top and Bottom Optimal Marginal Tax Rates)

Under the assumptions of Proposition 3.3, the 4*-approximate marginal tax rate is non-zero both at the top and at the bottom of the income distribution.*

It is evidently possible to generalize the above result to the case of any n-th order approximation.

[TO BE WRITTEN]

3.4 Income Risk and The Empirical Distribution of Marginal Tax Rates

It is maybe most interesting to examine the implications of our main result when applied to the interior of the skill distribution (that is, when $\Lambda > 0$). It is then useful to compare our results with the sufficient conditions for increasing marginal tax rates that are provided in Diamond [1]. A main implication of our results is that increasing marginal tax rates are expected to obtain under weaker conditions when the "social insurance" role of income taxation is introduced in the Mirrlees [5] model, in agreement with intuition.

For instance, one could interpret our model as best describing the middle and upper parts of the income distribution, rather than the lower part because optimal labor supply may be zero at low γ's. In comparison with Diamond [1], our result suggests that the marginal tax rate curve is predicted, in the model with risk, both to slope up above a *lower* threshold value and to have *higher* values above the threshold. Saez [8] has estimated such a threshold in the model without risk and he has found it to be about $\$75,000$ on a yearly basis. In view of our results, it remains to be seen if the presence of risk makes the threshold closer to what we actually observe. Another important question is whether, in the presence of risk, increasing marginal tax rates would still obtain, under a Pareto skill distribution and a weighted utilitarian social welfare function, when the labor supply elasticity is not constant or, more generally, when Λ is not constant.

[TO BE WRITTEN]

4 Extension: Risk Aversion

In this section, we ask whether our main result is robust with respect to the introduction of risk aversion. We focus on the following utility functions: either $u[c - v(l)]$ or $u[c] - v(l)$, where u is increasing and concave. The former case incorporates no income effects while the latter allows for income effects.

[TO BE WRITTEN]

References

[1] Diamond, P. (1998). "Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates". *American Economic Review* 88: 83-95.

[2] Golosov, M., Tsyvinski, A., and Werning, I. (2006). "New Dynamic Public Finance: A User's Guide". In *NBER/Macroeconomics Annual 2006* (D. Acemoglu, K. Rogoff and M. Woodford, Eds), Volume 21: 317-363. Cambridge: MIT Press.

[3] Kocherlakota, N. (2006). "Advances in Dynamic Optimal Taxation". In *Advances in Economics and Econometrics*, Theory and Applications, Ninth World Congress (R. Blundell, W. K. Newey, T. Persson, Eds), Volume I: 269-299. New-York: Cambridge University Press.

[4] Laffont, J.-J., and Martimort, D. (2002). *The theory of incentives: the principal-agent model*. Princeton: Princeton University Press.

[5] Mirrlees, J. (1971). "An Exploration in the Theory of Optimal Income Taxation". *Review of Economic Studies* 38: 175-208.

[6] Mirrlees, J. (1990). "Taxing Uncertain Incomes". *Oxford Economic Papers* 42: 34-45.

[7] Piketty, T. (1997). "La redistribution fiscale face au chômage". *Revue Française d'Economie* 12: 157-201.

[8] Saez, E. (2001). "Using elasticities to derive optimal tax rates". *Review of Economic Studies* 68: 205-229.

[9] Salanié, B. (2003). *The economics of taxation*. Cambridge: MIT Press.