# The Pareto-Frontier in a simple Mirrleesian model of income taxation

Felix Bierbrauer
Pierre C. Boyer

MAX PLANCK SOCIETY

# The Pareto-Frontier in a simple Mirrleesian model of income taxation

Felix Bierbrauer / Pierre C. Boyer

May 2010

# The Pareto-Frontier in a simple Mirrleesian model of income taxation[*]

Felix J. Bierbrauer[†]and Pierre C. Boyer[‡]

May 5, 2010

### Abstract

We characterize the Pareto-frontier in a simple Mirrleesian model of income taxation. We show how the second-best frontier which incorporates incentive constraints due to private information on productive abilities relates to the first-best frontier which takes only resource constraints into account. In particular, we argue that the second-best frontier can be interpreted as a Laffer-curve. We also use this second-best frontier for a comparative statics analysis of how optimal income tax rates vary with the degree of inequity aversion, and for a characterization of optimal public-good provision. We show that a more inequity averse policy maker chooses tax schedules that are more redistributive and involve higher marginal tax rates, but chooses a lower public-goods provision level.

*Keywords:* Optimal Income Taxation, Laffer-Curve, Public-Good Provision.

*JEL:* H21, H41, D82.

# 1  Introduction

In this note we characterize the set of Pareto-efficient allocations in a simple version of the Mirrlees (1971) model of optimal nonlinear income taxation. We consider an economy where there are two types of individuals who differ in their productive abilities. Individuals have a linear effort cost, so that an individual's marginal cost of contributing to the economy's output is low if the individual is high-skilled and high if the individual is low-skilled.[1] We obtain the following results.

*First, we show how the second-best Pareto-frontier compares to the first-best Pareto-frontier in a model in which information about productive abilities is publicly observed.* The first-best frontier is linear and there is neither a maximal, nor a minimal utility level for high-skilled or low-skilled individuals. With the second-best Pareto-frontier, by contrast, such maxima and minima exist. Moreover, we show that each point on the second-best Pareto-frontier is associated with a marginal income tax rate for low-skilled individuals, and a marginal income tax rate for high-skilled individuals. Both marginal tax rates are shown to be non-decreasing functions of the utility of the low-skilled individuals.

The fact that there is a minimal of utility for the high-skilled individuals implies that the second-best frontier is similar to a Laffer-curve: if the high-skilled individuals' utility level fell below the minimal level, then an increase of their utility level would also make the low-skilled individuals better off. Moreover, the minimal utility level of the high-skilled is associated with maximal marginal income tax rates. Having a minimal utility level for the high-skilled individuals is hence akin to the possibility, known from models of linear income taxation (see, e.g., Sheshinski (1972), or Hellwig (1986)), that if the income tax rate is very high, then a decrease of the income tax rate will lead to an increase of tax revenue.

*Second, we use the Pareto-frontier to characterize an optimal income tax.* The frontier makes it possible to characterize an optimal income tax in a straightforward way: given a social welfare function, just choose an optimal point on the frontier. This is a optimization problem whose solution can be characterized by a first-order condition, which is shown to be both necessary and sufficient for welfare-maximization.[2] Moreover, we conduct a comparative statics analysis: an increase in inequity aversion – more formally, an increase of the weight on the low-skilled individuals' utility in the social welfare function – is shown to be associated with an increase of the marginal tax rate of the high-skilled individuals, and with an increase of the marginal tax rate of the low-skilled individuals.[3]

*Third, we use the Pareto-frontier to characterize the optimal level of public-good provision.* As an extension, we assume that individuals derive utility from the provision of a public good and solve for an optimal scheme of income taxation and public-good provision. Once more, the Pareto-frontier enables us to characterize the optimal policy via the first order conditions of an unconstrained optimization problem. Again, we show that these first-order conditions are not

---

[1] For an analysis of optimal utilitarian income taxation with linear effort costs, see Weymark (1986, 1987).

[2] The literature on optimal income taxation mostly focusses on necessary conditions. Our simple setup enables us to show sufficiency also.

[3] For a model of linear income taxation, a similar comparative statics analysis can be found in Hellwig (1986).

only necessary but also sufficient.[4] Again, we undertake a comparative statics exercise and find that an increase in inequity aversion is, ceteris paribus, associated with a lower public-goods provision level.

Thus, a main finding of this paper is that an increase in inequity aversion has different implications for redistributive income taxation and public-goods provision. It implies that there are more direct income transfers from high-skilled to low-skilled individuals and also higher marginal tax rates. However, it also implies that there is less public-goods provision. The reason for this last observation is that low-skilled individuals suffer more from the higher taxes that are needed in order to finance increased public-goods provision. A more inequity-averse policy maker attaches more weight to this and, therefore opts for a smaller public-goods provision level.

This paper contributes to the literature in the following way: it complements the work by Weymark (1986, 1987) who provides a comparative statics analysis of optimal income taxes, under the maintained assumption that the policy maker's objective is to redistribute resources from high-skilled to low-skilled individuals. Our work does not rely on such an assumption, but characterizes the whole set of Pareto-efficient allocations. Our comparative static results therefore apply to all efficient income tax schedules, and not just to the subset that is relevant if the policy is chosen by an inequity-averse policy maker.[5]

This present paper's characterization of the Pareto-frontier will prove useful for applications. Having a one-to-one mapping between Pareto-efficient income tax schedules and points on the Pareto-frontier makes it possible to circumvent the solution of constraint optimization problems with incentive and resource constraints. This property has been useful for the characterization of optimal tax and expenditure policies in Bierbrauer (2010). Moreover, a complete characterization of the Pareto-frontier is necessary for any political economy treatment of nonlinear income taxation. If the policy maker's objective function is an endogenous object due to political competition, then one cannot a priori limit attention to tax schedules that would be chosen by an inequity-averse policy maker.

The remainder is organized as follows. Section 2 specifies the environment. The Pareto-frontier is derived in Section 3. Section 4 contains the characterization of an optimal income tax system, and Section 5 extends the model so as to include an optimal decision on public-good provision. The last section concludes. All proofs are relegated to an Appendix.

## 2    The Environment

An individual $i$ has utility function $U_i = u(c_i) - l_i$, where $c_i$ is consumption of private good and $l_i$ denotes hours worked by individual $i$. The function $u(.)$ satisfies $u'(.) > 0$, $u''(.) < 0$, and $\lim_{c \to 0} u'(c) = \infty$, and $\lim_{c \to \infty} u'(c) = 0$. Individuals differ in their productive abilities.

---

[4]Necessary conditions have previously been derived, for instance, by Boadway and Keen (1993), Sandmo (1998), or Gahvari (2006).

[5]The work by Stiglitz (1982, 1987) contains some results about Pareto-efficient income tax schedules. This work, however, neither contains a complete characterization of the Pareto-frontier, nor comparative statics results.

Each individual has a skill parameter $w_i$, where $w_i \in \{w_L, w_H\}$ with $0 < w_L < w_H$. There is a continuum of individuals of mass 1. The population share of individuals with a high skill level is commonly known and denoted by $f_H$. Let $f_L = 1 - f_H$. We assume in the following that $1 \geq f_H \frac{w_H}{w_L}$.[6]

Output can produced according to two constant returns to scale technologies. If an individual with productivity $w_t$, $t \in \{L, H\}$, works for one hour, this yields $w_t$ units of output. We denote the output that is provided by individual $i$ in the following by $y_i$, where $y_i = w_i l_i$. We can hence write an individual's utility function also as $U_i = u(c_i) - \frac{y_i}{w_i}$. Individuals have private information on their skills and also on the hours that they work. It is, however, observable how much output an individual provides, $y_i$, and how much the individual consumes, $c_i$. An allocation consists of a consumption-output combination for productive individuals, $(c_H, y_H)$, and a consumption-output combination for less productive individuals, $(c_L, y_L)$.

Our analysis does not incorporate upper and lower bounds on the hours that individuals can work, and the output they can produce. Obviously, we cannot justify this assumption as being realistic. Our reason for imposing it is that it makes the differences between a first-best and a second-best analysis very transparent. As will become clear, the presence of incentive constraints endogenously generates upper and lower bounds on the productive effort of individuals, even if there is no physical constraint at all.

We use a mechanism design approach to study the Mirrleesian income tax problem; that is, instead of assuming that individuals are confronted with an income tax schedule $T$ that relates their pre-tax-income, $y$, to their after-tax income, $c$, and then choose $y$ and $c$ in a utility-maximizing way, we focus immediately on the set of allocations $(c_L, y_L, c_H, y_H)$ that permit a decentralization via some income tax schedule. As is well-known,[7] an allocation can be reached via an income tax schedule if and only if it is *feasible* and *incentive compatible*. Feasibility requires that

$$f_H \, y_H + f_L \, y_L = r + f_H \, c_H + f_L \, c_L \, , \tag{1}$$

where $r$ is an exogenous revenue requirement. Incentive compatibility holds provided that

$$u(c_H) - \frac{y_H}{w_H} \geq u(c_L) - \frac{y_L}{w_H} \, , \tag{2}$$

and

$$u(c_L) - \frac{y_L}{w_L} \geq u(c_H) - \frac{y_H}{w_L} \, . \tag{3}$$

An allocation is said to be *admissible* if it is feasible and incentive compatible.

## 3  The Pareto-Frontier

An allocation $(c_L, y_L, c_H, y_H)$ is said to be Pareto-efficient if it is admissible and there is no admissible allocation $(\tilde{c}_L, \tilde{y}_L, \tilde{c}_H, \tilde{y}_H)$ such that $u(\tilde{c}_H) - \frac{\tilde{y}_H}{w_H} \geq u(c_H) - \frac{y_H}{w_H}$ and $u(\tilde{c}_H) - \frac{\tilde{y}_H}{w_H} \geq u(c_L) - \frac{y_L}{w_L}$, with at least one of these inequalities being strict.

---

[6]This assumptions simplifies the exposition. For the optimization problems studied in the Appendix, it implies that non-negativity constraints on individual consumption levels may be safely ignored.

[7]See, e.g., Guesnerie (1995) for a proof.

To characterize the set of Pareto-efficient allocations, we study a family of optimization problems, which depend on two parameters, namely the revenue requirement $r$, and a given utility level for the low-skilled individuals,

$$u(c_L) - \frac{y_L}{w_L} = v_L \ . \tag{4}$$

Given $r$ and $v_L$, we denote the set of admissible allocations satisfying (4) by $\mathcal{A}(v_L, r)$. We can now define the value function $V_H : (v_L, r) \mapsto V_H(v_L, r)$ of the following optimization problem:

$$V_H(v_L, r) \ := \ \max \ u(c_H) - \frac{y_H}{w_H} \ \text{s.t.} \ (c_L, y_L, c_H, y_H) \in \mathcal{A}(v_L, r) \ . \tag{5}$$

For given $r$, we identify the second best Pareto-frontier with the range of $V_H$ so that $V_{H1} < 0$, where $V_{Hj}$ denotes the derivative of the function $V_H$ with respect to its $j^{th}$ argument.

**Proposition 1** *For given $r$, the function $V_H$ has the following properties:*

I. *There exist numbers $\underline{v}_L(r)$ and $\overline{v}_L(r)$ so that $V_{H1} < 0$ if and only if $v_L \in [\underline{v}_L(r), \overline{v}_L(r)]$.*

II. *There exist numbers $v_L^1(r)$, and $v_L^2(r)$ with $\underline{v}_L(r) < v_L^1(r) < v_L^2(r) < \overline{v}_L(r)$, so that:*

    (a) *For $v_L \in [\underline{v}_L(r), v_L^1(r)[$, the low-skilled individuals' incentive constraint (3) is binding, $V_{H11} < 0$, $\lim_{v_L \to \underline{v}_L(r)} | V_{H1} | = 0$ and $\lim_{v_L \to v_L^1(r)} | V_{H1} | = \frac{f_L \, w_L}{f_H \, w_H}$.*

    (b) *For $v_L \in [v_L^1(r), v_L^2(r)]$, no incentive constraint is binding. Moreover, $V_{H11} = 0$, and $| V_{H1} | = \frac{f_L \, w_L}{f_H \, w_H}$.*

    (c) *For $v_L \in ]v_L^2(r), \overline{v}_L(r)]$, the high-skilled individuals' incentive constraint (2) is binding, $V_{H11} < 0$, $\lim_{v_L \to v_L^2(r)} | V_{H1} | = \frac{f_L \, w_L}{f_H \, w_H}$, and $\lim_{v_L \to \overline{v}_L(r)} | V_{H1} | = \infty$.*

The Proposition shows that along the second-best Pareto-frontier there are maximal utility levels for low-skilled and for high-skilled individuals. For $v_L < \underline{v}_L(r)$, or $v_L > \underline{v}_L(r)$, the function $V_H$ is increasing in $v_L$, which is incompatible with Pareto-optimality. Part $II$ shows that the Pareto-frontier can be divided into three segments: for low values of $v_L$, the low-skilled individuals' incentive constraint is binding, and for high values of $v_L$, the high-skilled individuals' incentive constraint is binding. Moreover, there is a range of intermediate values of $v_L$ so that neither the low-skilled individuals' nor the high-skilled individuals' incentive constraint is binding. Hence, in this range the second-best Pareto-frontier coincides with the first-best Pareto-frontier, which does not take the requirement of incentive compatibility into account. The question which of the incentive constraints is binding is decisive for the curvature of the Pareto-frontier. It is strictly concave only if one of the incentive constraints is binding and linear otherwise. Moreover, the function $V_H$ is shown to satisfy the Inada-conditions

$$\lim_{v_L \to \underline{v}_L(r)} | V_{H1} | = 0 \quad \text{and} \quad \lim_{v_L \to \overline{v}_L(r)} | V_{H1} | = \infty \ . \tag{6}$$

We now turn to the marginal income tax rates that are associated with any given point on the Pareto-frontier. To any point on the frontier corresponds the allocation

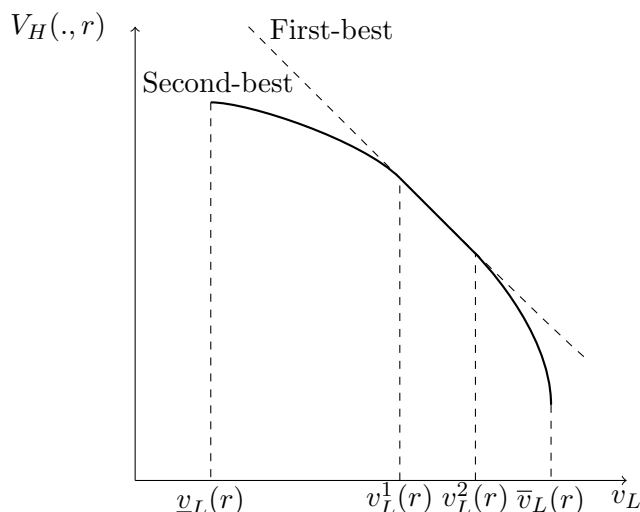$$a(v_L, r) = (c_L(v_L, r), y_L(v_L, r), c_H(v_L, r), y_H(v_L, r)) \ ,$$

Figure 1: Pareto-Frontiers.

which solves Problem (5). Following the literature, we interpret the difference between an individual's marginal rate of transformation between output $y$ and consumption $c$, which equals 1 for each individual, and the individual's marginal rate of substitution, $\frac{1}{wu'(c)}$, as the marginal income tax rate that the individual faces.[8] With reference to the allocation $a(v_L, r)$, we therefore define the marginal tax rates for high-skilled and low-skilled individuals, respectively, as follows:

$$\tau_H(v_L, r) := 1 - \frac{1}{w_H u'(c_H(v_L, r))} \text{ and } \tau_L(v_L, r) := 1 - \frac{1}{w_L u'(c_L(v_L, r))} .$$

**Proposition 2** *Both marginal tax rates are non-decreasing functions of $v_L$. We also have that $\tau_H(v_L, r) \leq 0$ and that $\tau_L(v_L, r) \geq 0$, for all $v_L$ and $r$. More specifically,*

(a) *For $v_L \in [\underline{v}_L(r), v_L^1(r)[$, $\tau_H < 0$, $\tau_{H1} > 0$, and $\tau_L = 0$.*

(b) *For $v_L \in [v_L^1(r), v_L^2(r)]$, $\tau_H = 0$, and $\tau_L = 0$.*

(c) *For $v_L \in ]v_L^2(r), \overline{v}_L(r)]$, $\tau_H = 0$, $\tau_L > 0$, and $\tau_{L1} > 0$.*

According to Proposition 2, both the sign and the comparative statics properties of the marginal income tax rates depend on which incentive constraint is binding. If the low-skilled are very badly off, their incentive constraint is binding, which implies an upward distortion of labor supply for the high-skilled, $\tau_H < 0$, and no distortionary taxation of low-skilled labour, $\tau_L = 0$. Moreover, as the low-skilled are made better off, the upward distortion of high-skilled labour supply becomes smaller and smaller, so that $\tau_{H1} > 0$. In the range where no incentive constraint binds, there are no distortions at all, i.e., both marginal tax rates are equal to 0. Finally, if

---

[8]This interpretation is based on the first-order condition of the utility maximization problem that individuals face when confronted with an income tax schedule $T$: choose $c$ and $y$ in order to maximize $u(c) - \frac{y}{w}$ subject to the constraint $c = y - T(y)$. The first order condition is $T'(y) = 1 - \frac{1}{wu'(c)}$.

the low-skilled individual's utility level is very high, and hence the high-skilled individual's utility level very low, the high-skilled individuals' incentive constraint is binding. This yields a downward distortion of the supply of low-skilled labor, $\tau_L > 0$, and no distortion of high-skilled labor supply, $\tau_H = 0$. Moreover, the downward distortion gets more severe as we make the low-skilled individuals even better off, $\tau_{L1} > 0$.

The second-best Pareto-frontier is akin to a Laffer-curve in that it indicates what marginal tax rates are compatible with efficiency. The typical Laffer-consideration identifies efficiency with an ability to generate tax revenues, so that a tax system is inefficient if higher revenues can be generated with lower tax rates. We, by contrast, work with a notion of Pareto-efficiency under conditions of asymmetric information. The conclusions that emerge are, however, similar: The increased demand for tax revenues in the conventional Laffer-analysis, corresponds in our model to an increased concern for the low-skilled individuals' well being. This goes hand in hand with an increase of marginal tax rates. However, there is a pair of maximal tax rates so that a further increase would fail to channel more resources to the low-skilled.

To complete our characterization of the Pareto-frontier, the following proposition documents the consequences of a change in the exogenous revenue requirement $r$.

**Proposition 3** *The maximal and minimal utility levels of the low-skilled individuals $\underline{v}_L$ and $\overline{v}_L$, and the maximal and minimal utility levels of the high-skilled individuals $\underline{V}_H$ and $\overline{V}_H$, are strictly decreasing functions of $r$, where, for any $r$, we define $\underline{V}_H(r) := V_H(\overline{v}_L(r), r)$ and $\overline{V}_H(r) := V_H(\underline{v}_L(r), r)$.*

An increase in the revenue requirement $r$ yields to a crowding out of utility-possibilities, i.e., if $r$ goes up then the range of possible utility levels for the low-skilled individuals shifts to the left and the range of possible utility levels for the high-skilled individuals shifts downwards; see Figure 2.

## 4 Optimal Income Taxation

For a given revenue requirement $r$, we can use the Pareto-frontier in order to characterize a welfare-maximizing income tax system by the solution to following maximization problem, $\max_{v_L} g_H V_H(v_L, r) + g_L v_L$, where $g_H$ and $g_L = 1 - g_H$ are, respectively, the welfare weights on the utility levels of the high-skilled and the low-skilled individuals. We have shown that $V_H$ is a strictly decreasing and globally concave function of $v_L$, which satisfies the Inada-conditions. Consequently, the following first order condition is both necessary and sufficient for a welfare-maximizing choice of $v_L$,

$$V_{H1} = -\frac{g_L}{g_H} . \tag{7}$$

The solution to this optimization problem is illustrated in Figure 3 as a tangency condition. The downward sloping line which touches the Pareto-frontier at the point $v_L^*$ is a social indifference curve with slope $-\frac{g_L}{g_H}$.
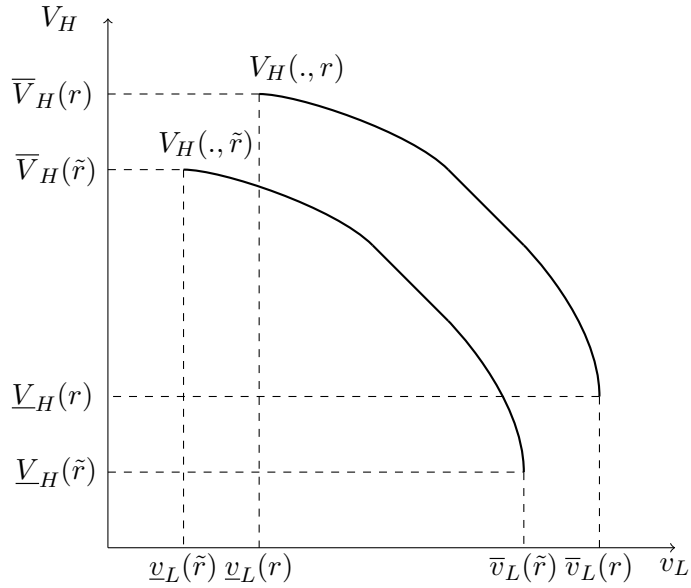
6

Figure 2: Second-Best Pareto-Frontiers with $r < \tilde{r}$.

Given this point on the frontier, we can use Proposition 2 to trace out the implications for marginal income tax rates. In particular, we can do a comparative statics analysis so as to check what an increased concern for the low-skilled – i.e., an increase of $g_L$, so that the social indifference curves become steeper – implies for marginal income tax rates. If we increase $g_L$, then this leads to an increase of the welfare-maximizing utility level of the low-skilled individuals, $v_L^*$. By Proposition 2, the implications for marginal tax rates depend on which incentive constraints are binding.

As is well-known,[9] the optimal utilitarian allocation with $g_L = f_L$ and $g_H = f_H$ is such that the high-skilled individuals' incentive constraint is binding. If, starting from this point on the Pareto-frontier, we increase the utility of the low-skilled, $v_L$, the associated change in consumption levels and output requirements is such that the marginal tax rate of the high-skilled individuals does not change, i.e., we still have "no distortion at the top". At the same time, the "downward distortion" of low-skilled labor supply becomes more severe; that is, the marginal income taxes for the low-skilled go up. If instead, we decrease the utility of the low-skilled, the effect on marginal tax rates depends on how much we reduce $v_L$. A small reduction will again leave the high-skilled individuals' marginal tax rate unchanged, but reduce the downward distortions for the low-skilled individuals. Eventually, the downward distortion completely disappears and we are in the region of the Pareto-frontier in which no incentive constraint binds. In this region, there are no distortions and small changes in $v_L$ have no impact on marginal tax rates. However, if we decrease $v_L$ substantially, we eventually reach the region of the Pareto-frontier so that the low-skilled individuals' incentive compatibility constraint is binding. This is associated with upward distortions in the supply of high-skilled labor (negative marginal tax rates) and no distortions in the supply of low-skilled labor (zero marginal tax

---

[9]See, for instance, Weymark (1987) for a proof.

rates). Moreover, the lower the utility level of the low-skilled, the more severe is the upward distortion for the high-skilled individuals.
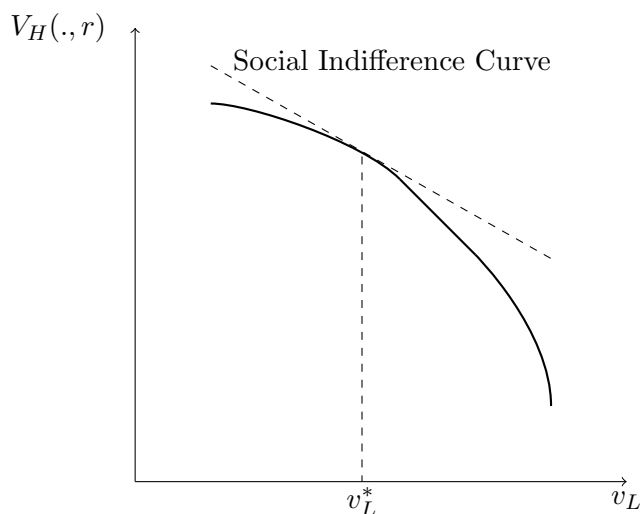


Figure 3: Second-Best Pareto-Frontier and Social Indifference Curve.

## 5 Optimal Public-Good Provision

We now consider an extended version of our model, so as to endogenize the revenue requirement $r$. More specifically, we assume that there is a public good $q$ which can be produced at a cost $r(q)$, where we assume that the cost function $r(\cdot)$ satisfies $r'(.) > 0$, and $r''(.) > 0$. We assume that the preferences of type $t$-individuals, $t \in \{L, H\}$, are now given by $\theta_t\, q + u(c) - \frac{y}{w_t}$, where $\theta_t$ is a parameter which equals the marginal utility from public-good provision for individuals of type $t \in \{L, H\}$. For this extended model, a welfare-maximizing policy solves the following problem

$$\max_{v_L, q}\ g_H(\theta_H\, q + V_H(v_L, r(q))) + g_L(\theta_L\, q + v_L). \tag{8}$$

**Proposition 4** *A pair $(v_L^*, q^*)$ solves the policy problem in (8) if and only if it satisfies condition (7) and*

$$g_H\, \theta_H + g_L\, \theta_L = -g_H\, V_{H2}\, r'. \tag{9}$$

The optimality condition in (9) is a version of the Samuelson rule which characterizes the optimal public-good provision level by the requirement that the marginal welfare gain from increased public-good provision, $g_H\, \theta_H + g_L\, \theta_L$, must be equal to the welfare loss due to the resulting increase of the revenue requirement $-g_H\, V_{H2}\, r'$. According to Proposition 4, this condition, in conjunction with the optimality condition (7) for $v_L$, is both necessary and sufficient for the characterization of a welfare-maximizing policy.

Proposition 4 can be used for a comparative statics of optimal public-goods provision with respect to the policy maker's degree of inequity aversion. In the Appendix it is shown that $V_{H2} = \frac{V_{H1}}{w_L} - \frac{1}{w_H}$, see Lemma 12. If we use this expression together with condition (7) to substitute for $V_{H2}$ in (9), we obtain the following formulation of the Samuleson rule,

$$r' = \frac{g_H \theta_H + g_L \theta_L}{\frac{g_H}{w_H} + \frac{g_L}{w_L}} \ . \tag{10}$$

This equation states that the marginal cost of public-goods provision is equal to the ratio of the marginal welfare gain from increased public-good provision, and the marginal welfare loss from increased output provision (Note that a marginal increase of $y_L$ implies a welfare loss of $\frac{g_L}{w_L}$, and that a marginal increase of $y_H$ implies a welfare loss of $\frac{g_H}{w_H}$).

If we assume that individuals differ only in their disutility of productive effort and set $\theta_L = \theta_H =: \bar{\theta}$, then, upon using that $g_L + g_H = 1$, we may rewrite this expression as

$$r' = \frac{\bar{\theta}}{\frac{1}{w_H} + g_L(\frac{1}{w_L} - \frac{1}{w_H})} \ . \tag{11}$$

Now, it is easily verified that an increased concern for the well-being of the low-skilled implies a lower lower public-goods provision level. Thus, we find that a policy maker who cares more for the well-being of low-skilled individuals provides less public goods. The reason is that such a policy maker attaches more weight to the low-skilled individuals' larger utility loss from having to provide the funds that are needed to pay for the public good.

# 6 Concluding Remarks

This paper has provided a complete analytical characterization of efficient tax and expenditure policies in a simple Mirrleesian model of optimal income taxation. This made it possible to undertake a rigorous comparative statics analysis of optimal policies with respect to the policy maker's degree of inequity aversion. We find that a larger concern for the well-being of low-skilled individuals goes together with larger direct income transfers from high-skilled to low-skilled individuals and with higher marginal income tax rates. At the same time, however, there is a decreased supply of public goods.

# References

Bierbrauer, F. (2010). Optimal income taxation and public-goods provision with preference and productivity shocks. Mimeo, Max Planck Institute for Research on Collective Goods.

Boadway, R. and Keen, M. (1993). Public goods, self-selection and optimal income taxation. *International Economic Review*, 34:463–478.

Gahvari, F. (2006). On the marginal costs of public funds and the optimal provision of public goods. *Journal of Public Economics*, 90:1251–1262.

Guesnerie, R. (1995). *A Contribution to the Pure Theory of Taxation*. Cambridge University Press.

Hellwig, M. (1986). The optimal linear income tax revisited. *Journal of Public Economics*, 31:163–179.

Mirrlees, J. (1971). An exploration in the theory of optimum income taxation. *Review of Economic Studies*, 38:175–208.

Sandmo, A. (1998). Redistribution and the marginal cost of public funds. *Journal of Public Economics*, 70:365–382.

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

Stiglitz, J. (1982). Self-selection and Pareto-efficient taxation. *Journal of Public Economics*, 17:213–240.

Stiglitz, J. (1987). Pareto-efficient and optimal taxation and the new new welfare economics. In Auerbach, A. and Feldstein, M., editors, *Handbook of Public Economics*. North-Holland, Amsterdam.

Weymark, J. (1986). A reduced-form optimal nonlinear income tax problem. *Journal of Public Economics*, 30:199–217.

Weymark, J. (1987). Comparative static properties of optimal nonlinear income taxes. *Econometrica*, 55:1165–1185.

# Appendix

**Proof of Proposition 1**

**Step 1:** We show that the Pareto-frontier can be divided into a part in which only the low-skilled individuals' incentive constraint is binding, a part in which no incentive constraint is binding, and a part in which only the high-skilled individuals' incentive constraint is binding.

**Lemma 1** *If an allocation is Pareto-efficient, then at most one of the constraints (2) and (3) is binding.*

**Proof** Suppose that both incentive constraints are binding. This implies that the given allocation must involve pooling, i.e., $y_L = y_H$ and $c_L = c_H$. To see this, add the two binding incentive constraints in order to obtain $y_H = y_L$. Obviously, incentive compatibility then also requires that $c_H = c_L$. To complete the proof, we show that it is possible to Pareto-improve upon an allocation that involves pooling. Consider an allocation with $y := y_L = y_H$ and $c := c_L = c_H$. The budget constraint implies that $y = c + r$. This pooling allocation is Pareto-dominated by the "laissez-faire" allocation which gives to each type of household the bundle that solves $\max_{c_i, y_i} u(c_i) - \frac{y_i}{w_i}$ s.t. $y_i = c_i + r$. By a standard revealed preferences argument, the laissez-faire allocation is incentive-compatible and makes no individual worse off in comparison to the initial pooling allocation. Moreover, it is easy to verify that the laissez-faire allocation satisfies

$y_L < y_H$ and $c_L < c_H$. Hence the high-skilled or the low-skilled choose a consumption-output combination that differs from the initial bundle $(c, y)$ and are thereby made strictly better off. Finally, note that the laissez-faire allocation is feasible. $\square$

**Lemma 2** *There are Pareto-efficient allocations so that neither the high-skilled individuals' incentive constraint (2) nor the low-skilled individuals' incentive constraint (3) is binding.*

**Proof** It follows from the first fundamental theorem of welfare economics that the laissez-faire allocation, defined formally in the proof of Lemma 1, is, in a first-best sense, Pareto-efficient. Since the laissez-faire allocation is such that the high-skilled individuals' incentive constraint (2) and the low-skilled individuals' incentive constraint (3) are satisfied, it is also second-best Pareto-efficient. The Lemma now follows from the observation that, at the laissez-faire allocation, both incentive constraints hold as strict inequalities. $\square$

**Lemma 3** *There are Pareto-efficient allocations so that the high-skilled individuals' incentive constraint (2) is binding and there are Pareto-efficient allocations so that the low-skilled individuals' incentive constraint (3) is binding.*

**Proof** The arguments in the proofs of Lemmas 1 and 2 imply that the laissez-faire allocation is Pareto-efficient and is such that both incentive constraints are slack. This implies that there is a neighborhood $V_L^{lf}$ of the laissez-faire allocation so that $v_L \in V_L^{lf}$ implies that

$$
V_H(v_L, r) \quad := \quad \max \quad u(c_H) - \frac{y_H}{w_H} \quad \text{s.t.} \quad u(c_L) - \frac{y_L}{w_L} = v_L, \\
f_H \, y_H + f_L \, y_L = r + f_H \, c_H + f_L \, c_L \, . \tag{12}
$$

The solution to this optimization problem can be conveniently characterized using a Lagrangean approach. This yields the following results: the consumption levels $c_L(v_L, r)$ and $c_H(v_L, r)$ satisfy the following first order conditions,

$$
u'(c_H(v_L, r)) = \frac{1}{w_H} \quad \text{and} \quad u'(c_L(v_L, r)) = \frac{1}{w_L} \, . \tag{13}
$$

The low-skilled individuals output requirement then follows from

$$
u(c_L(v_L, r)) - \frac{y_L(v_L, r)}{w_L} = v_L \, . \tag{14}
$$

Given that $c_L(v_L, r)$, $c_H(v_L, r)$ and $y_L(v_L, r)$ are determined by (13) and (14), the high-skilled individuals' output requirement follows from

$$
f_H \, y_H(v_L, r) + f_L \, y_L(v_L, r) = r + f_H \, c_H(v_L, r) + f_L \, c_L(v_L, r) \, . \tag{15}
$$

Equations (13)-(15) can be used to undertake a comparative statics exercise of the allocation

$$
a(v_L, r) = (c_L(v_L, r), y_L(v_L, r), c_H(v_l, r), y_H(v_L, r)) \, .
$$

with respect to $v_L$. This yields the following observations:

$$c_{L1} = c_{H1} = 0, \quad y_{L1} = -w_L < 0, \quad y_{H1} = \frac{f_L}{f_H} w_L > 0 . \tag{16}$$

Consequently, if we start from an allocation with $v_L \in V_L^{lf}$, which has slack in both incentive constraints, a marginal change of $v_L$ has the following consequences: (i) For the low-skilled individuals, we have that

$$\frac{d}{dv_L} \left[ \left( u(c_L(v_L, r)) - \frac{y_L(v_L, r)}{w_L} \right) - \left( u(c_H(v_L, r)) - \frac{y_H(v_L, r)}{w_L} \right) \right] = 1 + \frac{f_L}{f_H} > 0 .$$

This implies that, if starting from $v_L \in V_L^{lf}$, we decrease $v_L$, this reduces the slack in the low-skilled individuals' incentive constraint in a linear way. Hence, we eventually reach an allocation where the low-skilled individuals' incentive constraint is binding. This proves the second statement in Lemma 3.

(ii) For the high-skilled individuals, we have that

$$\frac{d}{dv_L} \left[ \left( u(c_H(v_L, r)) - \frac{y_H(v_L, r)}{w_H} \right) - \left( u(c_L(v_L, r)) - \frac{y_L(v_L, r)}{w_H} \right) \right] = -\frac{w_L}{w_H} \left( 1 + \frac{f_L}{f_H} \right) < 0 .$$

Hence, if starting from $v_L \in V_L^{lf}$, we increase $v_L$, this reduces the slack in the high-skilled individuals' incentive constraint in a linear way. Hence, we eventually reach an allocation where the high-skilled individuals' incentive constraint is binding. This proves the first statement in Lemma 3.

$\square$

**Step 2:** We characterize the Pareto-frontier in the region where the high-skilled individuals' incentive constraint (2) is binding.

**Lemma 4** *A solution $a(v_L, r) = (c_L(v_L, r), y_L(v_L, r), c_H(v_L, r), y_H(v_L, r))$ of Problem (5) which is such that the constraint (2) is binding can be characterized by the following system of equations. The consumption-level of the high-skilled individuals satisfies*

$$u'(c_H(v_L, r)) = \frac{1}{w_H} . \tag{17}$$

*Given $c_H(v_L, r)$, $c_L(v_L, r)$ is implicitly characterized by the equation*

$$f_H \frac{w_H}{w_L} u(c_H(v_L, r)) + \left( 1 - f_H \frac{w_H}{w_L} \right) u(c_L(v_L, r)) - \frac{r + f_H c_H(v_L, r) + f_L c_L(v_L, r)}{w_L} = v_L . \tag{18}$$

*Given $c_L(v_L, r)$ and $c_H(v_L, r)$ the output requirements $y_L(v_L, r)$ and $y_H(v_L, r)$ are determined, respectively, as*

$$y_L(v_L, r) = r + f_H c_H(v_L, r) + f_L c_L(v_L, r) - f_H w_H (u(c_H(v_L, r)) - u(c_L(v_L, r))) , \tag{19}$$

*and*

$$y_H(v_L, r) = r + f_H c_H(v_L, r) + f_L c_L(v_L, r) + f_L w_H (u(c_H(v_L, r)) - u(c_L(v_L, r))) . \tag{20}$$

*Finally, the slope of the Pareto-frontier satisfies*

$$V_{H1}(v_L, r) = -\frac{f_L w_L (u'(c_L(v_L, r)) - \frac{1}{w_H})}{f_L - (w_L - f_H w_H) u'(c_L(v_L, r))}. \tag{21}$$

**Proof** We can use the fact that the constraint (2) is binding, and the resource constraint (1) to solve for $y_L(v_L, r)$ and $y_H(v_L, r)$, respectively, as functions of $c_L(v_L, r)$ and $c_H(v_L, r)$. This yields equations (19) and (20). We write the utility level of high-skilled and low-skilled individuals, respectively, as functions $c_L$ and $c_H$. This yields the indirect utility functions

$$U_H(c_L, c_H) := u(c_H) - \frac{y_H}{w_H} = f_H u(c_H) + f_L u(c_L) - \frac{r + f_H c_H + f_L c_L}{w_H} \ ,$$

and

$$U_L(c_L, c_H) := u(c_L) - \frac{y_L}{w_L} = f_H \frac{w_H}{w_L} u(c_H) + \left(1 - f_H \frac{w_H}{w_L}\right) u(c_L) - \frac{r + f_H c_H + f_L c_L}{w_L} \ .$$

The Pareto-problem can now be stated as follows: choose $(c_L, c_H)$ in order to maximize $U_H(c_L, c_H)$ subject to the constraint that $U_L(c_L, c_H) = v_L$. Note that this constraint is equivalent to equation (18) in Lemma 4. Using a Lagrangean approach, the solution $(c_L^*(v_L), c_H^*(v_L))$ to this problem can be characterized by the first-order conditions with respect to $c_H$ and $c_L$. These first order conditions can, respectively, be written as

$$u'(c_H) = \frac{1}{w_H} \ ,$$

which proves equation (17) in Lemma 4, and

$$\gamma = \frac{f_L w_L (u'(c_L) - \frac{1}{w_H})}{f_L - (w_L - f_H w_H) u'(c_L)},$$

where the Lagrange-multiplier $\gamma$ is defined in such a way that

$$\gamma = -\frac{d}{dv_L} U_H(c_L^*(v_L), c_H^*(v_L)) \ ,$$

which proves (21) in Lemma 4. $\qquad\square$

**Lemma 5** *Let $V_L^H$ be defined by the property that $v_L \in V_L^H$ if and only if the solution $a(v_L, r) = (c_L(v_L, r), y_L(v_L, r), c_H(v_L, r), y_H(v_L, r))$ of Problem (5) is such that the high-skilled individuals' incentive constraint (2) is binding. For $v_L \in V_L^H$, we have that*

$$c_{L1} < 0, \ c_{H1} = 0 \quad and \quad V_{H11} < 0 \ . \tag{22}$$

**Proof** It follows immediately from equation (17) that $c_{H1} = 0$. We now seek to show that $c_{L1} < 0$. Observe first that $c_L(v_L, r) < c_H(v_L, r)$. To see this, note that equations (19) and (20) imply that $y_H(v_L, r) > y_L(v_L, r)$. Since the high-skilled individuals' incentive constraint is binding, $u(c_H) - \frac{y_H}{w_H} = u(c_L) - \frac{y_L}{w_H}$, we must have $c_L(v_L, r) < c_H(v_L, r)$. Since $u'(c_H) = \frac{1}{w_H}$, by

13

(17) this implies in particular that $u'(c_L) > \frac{1}{w_H}$. Hence, the numerator of the fraction on the left-hand side of (21) is positive. Along the Pareto-frontier it must be that $V_{H1} < 0$, so that the denominator of this fraction is also positive, i.e.,

$$f_L - (w_L - f_H w_H) u'(c_L(v_L, r)) > 0 . \tag{23}$$

Using that $c_{H1} = 0$, we can use equation (18) to obtain

$$c_{L1} = -\frac{w_L}{f_L - (w_L - f_H w_H) u'(c_L)} < 0 . \tag{24}$$

Finally, (21) can be used to obtain

$$V_{H11} = -\frac{f_L w_L (1 + f_H w_H u'(c_L)) u''(c_L) c_{L1}}{(f_L - (w_L - f_H w_H) u'(c_L))^2} < 0 . \tag{25}$$

$\square$

**Lemma 6** *The set $V_L^H$ has a maximal element, denoted by $\overline{v}_L$. Moreover, $\lim_{v_L \to \overline{v}_L} \mid V_{H1} \mid = \infty$.*

**Proof** The maximal element of $V_L^H$ is given as

$$\overline{v}_L = \max_{c_L, c_H} U_L(c_L, c_H) ,$$

where $U_L$ is the indirect utility function defined in the proof of Lemma 4. The first-order conditions of this problem are both necessary and sufficient since the function $U_L$ can be shown to be negative-definite under our assumption that $1 \geq f_H \frac{w_H}{w_L}$. The following first-order condition characterizes $c_L(\overline{v}_L, r)$,

$$f_L - (w_L - f_H w_H) u'(c_L(\overline{v}_L, r)) = 0 .$$

Since $c_{L1} < 0$, we have that $f_L - (w_L - f_H w_H) u'(c_L(v_L, r))$ converges from above to 0, as $v_L$ converges to $\overline{v}_L$. This observation in conjunction with equation (22) in Lemma 4 implies that $\lim_{v_L \to \overline{v}_L} \mid V_{H1} \mid = \infty$. $\square$

**Lemma 7** *The set $V_L^H$ has a minimal element, denoted by $v_L^2$. Moreover, $\lim_{v_L \to v_L^2} \mid V_{H1} \mid = \frac{f_L}{f_H} \frac{w_L}{w_H}$.*

**Proof** We have shown in the proof of Lemma 3, that if we start from an allocation so that no incentive constraint is binding and increase $v_L$ we eventually reach the region where the high-skilled individuals' incentive constraint is binding. The utility level $v_L^2$ marks the boundary of these two regions. The corresponding allocation is characterized by the equations $u'(c_H(v_L^2, r)) = \frac{1}{w_H}$, $u'(c_L(v_L^2, r)) = \frac{1}{w_L}$ as well as by (19) and (20). Plugging $u'(c_L(v_L^2, r)) = \frac{1}{w_L}$ into (21) yields, $\mid V_{H1} \mid = \frac{f_L}{f_H} \frac{w_L}{w_H}$. $\square$

**Step 3**: We characterize the Pareto-frontier in the region where no incentive constraint is binding. The characterization of a Pareto-efficient in this region is given in the proof of Lemma 3. In particular, using this characterization we obtain:

$$V_{H1} = \frac{\partial}{\partial v_L}\left[u(c_H(v_L, r)) - \frac{y_H(v_L, r)}{w_H}\right] = -\frac{f_L\, w_L}{f_H\, w_H}\ .$$

**Step 4:** We characterize the Pareto-frontier in the region where the low-skilled individuals' incentive constraint (3) is binding. This follows from similar arguments as in Step 2. We therefore only state the analog versions of Lemmas 4-7, without giving a formal proof.

**Lemma 8** *A solution $a(v_L, r) = (c_L(v_L, r), y_L(v_L, r), c_H(v_L, r), y_H(v_L, r))$ of Problem (5) which is such that the constraint (3) is binding can be characterized by the following system of equations. The consumption-level of the low-skilled individuals satisfies*

$$u'(c_L(v_L, r)) = \frac{1}{w_L}\ . \tag{26}$$

*Given $c_L(v_L, r)$, $c_H(v_L, r)$ is implicitly characterized by*

$$f_L u(c_L(v_L, r)) + f_H u(c_H(v_L, r)) - \frac{r + f_H c_H(v_L, r) + f_L c_L(v_L, r)}{w_L} = v_L\ . \tag{27}$$

*Given $c_L(v_L, r)$ and $c_H(v_L, r)$ the output requirements $y_L(v_L, r)$ and $y_H(v_L, r)$ are determined, respectively, as*

$$y_L(v_L, r) = r + f_H c_H(v_L, r) + f_L c_L(v_L, r) - f_H w_L(u(c_H(v_L, r)) - u(c_L(v_L, r)))\ , \tag{28}$$

*and*

$$y_H(v_L, r) = r + f_H c_H(v_L, r) + f_L c_L(v_L, r) + f_L w_L(u(c_H(v_L, r)) - u(c_L(v_L, r)))\ . \tag{29}$$

*Finally, the slope of the Pareto-frontier satisfies*

$$V_{H1}(v_L, r) = -\frac{(1 - f_L\frac{w_L}{w_H})u'(c_H(v_L, r)) - \frac{f_H}{w_H}}{f_H(\frac{1}{w_L} - u'(c_H(v_L, r)))}\ . \tag{30}$$

**Lemma 9** *Let $V_H^H$ be defined by the property that $v_L \in V_H^H$ if and only if the solution $a(v_L, r) = (c_L(v_L, r), y_L(v_L, r), c_H(v_L, r), y_H(v_L, r))$ of Problem (5) is such that the low-skilled individuals' incentive constraint (3) is binding. For $v_L \in V_H^H$ we have that,*

$$c_{L1} = 0,\ c_{H1} < 0 \quad and \quad V_{H11} < 0\ . \tag{31}$$

**Lemma 10** *The set $V_H^H$ has a minimal element, denoted by $\underline{v}_L$. Moreover, $\lim_{v_L \to \underline{v}_L} |V_{H1}| = 0$.*

**Lemma 11** *The set $V_H^H$ has a maximal element, denoted by $v_L^1$. Moreover, $\lim_{v_L \to v_L^1} |V_{H1}| = \frac{f_L\, w_L}{f_H\, w_H}$.*

**Proof of Proposition 2**

($a$): It follows from the definition of marginal tax rates, equation (26) in Lemma 8, and the fact that $c_{L1} = 0$ in Lemma 9 that $\tau_L = \tau_{L1} = 0$. The facts that $c_{H1} < 0$ in Lemma 9, and that $u'(c_H(v_L^1, r)) = \frac{1}{w_H}$, imply that $\tau_H < 0$, for all $v_L \in [\underline{v}_L, v_L^1[$. Moreover, $c_{H1} < 0$ also implies that $\tau_{H1} > 0$.

($b$): This follows immediately from the definition of marginal tax rates and the observation, in the proof of Lemma 3, that $u'(c_H(v_L, r)) = \frac{1}{w_H}$ and that $u'(c_L(v_L, r)) = \frac{1}{w_L}$, whenever $v_L$ is such that no incentive constraint is binding.

($c$): It follows from the definition of marginal tax rates, equation (17) in Lemma 4, and the fact that $c_{H1} = 0$ in Lemma 5 that $\tau_H = \tau_{H1} = 0$. The facts that $c_{L1} < 0$ in Lemma 5, and that $u'(c_L(v_L^2, r)) = \frac{1}{w_L}$, see the arguments in the proof of Lemma 7, imply that $\tau_L > 0$, for all $v_L \in ]v_L^2, \overline{v}_L]$. Moreover, $c_{L1} < 0$ also implies that $\tau_{L1} > 0$.

**Proof of Proposition 3**

We prove our claim for the maximal utility level of the low-skilled $\overline{v}_L$ and the minimal utility level of the high-skilled $\underline{V}_H$. The proof for $\underline{v}_L$ and $\overline{V}_H$ follows from a symmetric argument.

In the proof of Proposition 1 we have shown that $c_L(\overline{v}_L(r), r)$ and $c_H(\overline{v}_L(r), r)$ do not depend on $r$: Recall that $c_H(\overline{v}_L(r), r)$ satisfies $u'(c_H(\overline{v}_L(r), r) = \frac{1}{w_H}$, as shown in Lemma 4, and that $c_L(\overline{v}_L(r), r)$ satisfies $f_L - (w_L - f_H w_H) u'(c_L(\overline{v}_L(r), r)) = 0$ as shown in the proof of Lemma 6. Hence,

$$\frac{d}{dr} c_L(\overline{v}_L(r), r) = \frac{d}{dr} c_H(\overline{v}_L(r), r) = 0 . \tag{32}$$

To establish that $\overline{v}_L$ and $\underline{V}_H$ are decreasing functions of $r$ it therefore suffices to show that $y_L(\overline{v}_L(r), r)$ and $y_H(\overline{v}_L(r), r)$ are both increasing functions of $r$. Using using equations (19) and (20), see Lemma 4, and equation (32) we obtain:

$$\frac{d}{dr} y_L(\overline{v}_L(r), r) = \frac{d}{dr} y_H(\overline{v}_L(r), r) = 1 . \tag{33}$$

**Proof of Proposition 4**

We seek to show that the first order conditions (7) and (9) are sufficient for an optimal choice of $v_L$ and $q$. To this end, we show that the Hessian of the function

$$W(v_L, q) = g_H(\theta_H \, q + V_H(v_L, r(q))) + g_L(\theta_L \, q + v_L) ,$$

which we henceforth denote by

$$H = \begin{pmatrix} g_H V_{H11} & g_H V_{H21} \, r' \\ g_H V_{H21} \, r' & g_H \left( V_{H22} \, (r')^2 + V_{H2} \, r'' \right) \end{pmatrix}$$

is negative-definite. The following Lemma will prove useful for this purpose.

**Lemma 12** *The function $V_H$ has the following properties*

$$V_{H2} = \frac{V_{H1}}{w_L} - \frac{1}{w_H} \, , \tag{34}$$

*and*

$$V_{H21} = \frac{V_{H11}}{w_L} \quad and \quad V_{H22} = \frac{V_{H11}}{w_L^2} \, . \tag{35}$$

**Proof** To prove this Lemma, we note that a Pareto-efficient allocation $a(v_L, r)$ solves the Pareto-problem in (5). We can setup a Lagrangean for this problem and derive the following set of first order conditions, for the choice of $y_H$ and $y_L$, respectively,

$$-\frac{1}{w_H} + \lambda f_H - \frac{\nu_H}{w_H} + \frac{\nu_L}{w_L} = 0 \, , \tag{36}$$

and

$$\lambda f_L - \frac{\gamma}{w_L} + \frac{\nu_H}{w_H} - \frac{\nu_L}{w_L} = 0 \, , \tag{37}$$

where $\lambda$ is the strictly positive multiplier on the resource constraint, $\gamma$ is the strictly positive multiplier on constraint $u(c_L) - \frac{y_L}{w_L} = v_L$, $\nu_L$ is the non-negative multiplier on the low-skilled individuals' incentive constraint, and $\nu_H$ is the non-negative multiplier on the high-skilled individuals' incentive constraint. Upon adding the first-order conditions (36) and (37), we find that

$$\lambda = \frac{\gamma}{w_L} + \frac{1}{w_H} \, . \tag{38}$$

It follows from the properties of the Lagrangean multipliers that $\lambda = -V_{H2}$ and $\gamma = -V_{H1}$. Using these facts to substitute for $\lambda$ and $\gamma$ in (38) establishes that (34) holds true. Once (34) is established, the statements in (35) follow immediately. $\square$

Let $Q(x_1, x_2) = (x_1 \; x_2) H (x_1 \; x_2)^t$ be the quadratic form associated with the Hessian $H$. Upon using Lemma 12 to substitute $\frac{V_{H11}}{w_L}$ for $V_{H21}$, and $\frac{V_{H11}}{w_L^2}$ for $V_{H22}$, we compute

$$Q(x_1, x_2) = g_H V_{H11} \left( x_1 + \frac{r'}{w_L} x_2 \right)^2 + g_H V_{H2} r'' x_2^2 \, .$$

From Proposition 1, $V_{H11} \leq 0$ and $V_{H1} < 0$; together with (34) this implies that $V_{H2} < 0$. Finally, we have assumed that $r'' > 0$. Consequently, $Q(x_1, x_2) < 0$, whenever $(x_1, x_2) \neq 0$. This implies that $H$ is negative-definite.