Unequal wages for equal utilities

H. Cremer, P. Pestieau and M. Racionero

May, 2008
Introduction

• One of the main rationales for public involvement in education is redistribution

• Design of an optimal educational policy is often combined with that of a redistributive income tax (second stage)

• Problem: first to determine the amount and distribution of public education and then to design structure of income tax

• Standard approach: individuals who differ in their ability to benefit from education
• Typical result: rather regressive distribution of public education
  – resources are concentrated on the most able individuals, to get a “cake” as big as possible
  – sharing among individuals through income taxation.

• Remark: if we restrict exercise to the first stage (Arrow, 1971), solution is different.

• Most of the literature: education together with redistributive income taxation (see, e.g., Bruno (1976), Hare and Ulph (1979) and Ulph (1977), Bovenberg and Jacobs (2005) and Maldonado (2007).

• “Perverse” distribution effect just mentioned is partially mitigated when we introduce decreasing returns of educational spending.
This paper we want to discuss another reason to push for regressive education

- Not linked to heterogeneity in innate ability to benefit from education
- But to pervasive non-convexities in the underlying problem

Related literature:

- Becker’s family economics: specialization in households;
- Stiglitz’s efficiency wage hypothesis.
• To make this point as clear as possible
  – assume away different learning abilities
  – no decreasing returns of education
  – suppose that the amount of public expenditure is given

• Show that (in the two stage problem) the most unequal distribution of wage happens to be desirable from the standpoint of social welfare maximization.
The economy

• In the line of Stiglitz: two types of individuals \( i = 1, 2 \) with relative size \( n_i > 0 \) (so that \( n_1 + n_2 = 1 \)) and identical utilities

\[
  u_i(x_i, \ell_i),
\]

where \( x_i \) is consumption of a numeraire good and \( \ell_i \) labor supply

• Productivities (or wages) \( w_i \) are endogenous and determined by educational technology. We characterize this technology through \( \Gamma \) the set of feasible wages that is defined by

\[
  \Gamma = \{(w_1, w_2) | n_1 w_1 + n_2 w_2 = 2 \text{ and } w_i \geq w\} \text{ with } 1 \geq w \geq 0.
\]
Figure 1:
• Micro foundations: following the literature assume

\[ w_i = \phi_i(\theta_i, e_i), \]

where \( \theta_i \) is (exogenous) ability to benefit from educations and \( e_i \) is education expenditure.

• Our representation is obtained by assuming that

- individuals are identical ex ante (\( \theta_i \) is the same for all)
- \( \phi \) is linear in \( e_i \) (for instance \( w_i = Ae_i \) where \( A \) is some constant)
- the total budget for education expenditures is fixed: \( n_1e_1 + n_2e_2 = E \)
• With $w = 0$ our education technology has two extreme cases
  
  – complete equalization of wages $w_1 = w_2 = 1$
  – maximum wage inequality with say $w_1 = 0$ and $w_2 = 1/n_2$ where all educational resources are devoted to type 2 individuals

• Role of $w$ is to limit the scope of wages inequalities that is feasible

• Symmetric social criterion:
  \[ W = \sum n_i \psi (u_i) \]
  
  where $\psi (\cdot)$ is concave.

• Extreme form of $\psi (\cdot)$ yields the Rawlsian criterion:
  \[ W = \min [u_1, u_2] . \]

• Assume without loss of generality that $n_1 = n_2 = 1/2$. 

9
First-best

- Optimal allocation determined by

\[
\max_{\ell_i, w_i, T_i} W = \Psi[u(w_1\ell_1 - T_1, \ell_1)] + \Psi[u(w_2\ell_2 - T_2, \ell_2)],
\]

s.t.  \( w_1 + w_2 = 2, \)

\( T_1 + T_2 = 0 \)

\( w_i \geq w. \)
• First-order condition with respect to $w_i$. Substituting for $w_2 = 2 - w_1$ and differentiating yields

$$\frac{\partial W}{\partial w_1} = (\ell_1 - \ell_2)\lambda,$$

where $\lambda > 0$ is the Lagrange multiplier of the resource constraint.

• This FOC is always satisfied for $w_1 = w_2 = 1$ (along with $\ell_1 = \ell_2$) $\Rightarrow$ one might be tempted to think that the solution implies equalization of wages

• However, we cannot simply assume that the SOC to be satisfied here
Quasi-linear illustration

- Assume quasi-linear preferences specified by \( u_i(x_i, \ell_i) = x_i - v(\ell_i) \)

- Pareto efficiency requires the maximization of “surplus” given by \( S = w_1 \ell_1 + w_2 \ell_2 - v(\ell_1) - v(\ell_2) \).

- Individual labor supplies \( \ell_i \) can be expressed solely as a function of \( w_i \) (there is no income effect). Substituting \( w_2 = 2 - w_1 \) we then reduce the problem to a single dimension and write \( \tilde{S}(w_1) \)

\[
\frac{d\tilde{S}}{dw_1} = \ell_1(w_1) - \ell_2(2 - w_1),
\]

\[
\frac{d^2\tilde{S}}{dw_1^2} = \frac{d\ell_1}{dw_1} + \frac{d\ell_2}{dw_2} > 0,
\]

which shows that \( \tilde{S}(w_1) \) is a convex function.
General problem: solution

• Directly compare the relevant levels of welfare, namely $W^I$ (achieved when $w_1 = w_2 = 1$) and $W^C$ (achieved when, say, say $w_1 = (2 - w)$ and $w_2 = w$), and show that $W^C \geq W^I$

• Define the level of labor supply under wage (and consumption) equalization

$$\ell^I = \arg \max_\ell \psi[u(\ell, \ell)] + \psi[u(\ell, \ell)]$$

• Maximum welfare is then given by

$$W^I = \psi[u(\ell^I, \ell^I)] + \psi[u(\ell^I, \ell^I)].$$
• At a corner solution with $w_1 = (2 - w)$ and $w_2 = w$, maximum welfare is given by

$$W^C = \max_{\ell_1, \ell_2} \psi [u (w_1 \ell_1 - T, \ell_1)] + \psi [u (w_2 \ell_2 + T, \ell_2)]$$

where $T$ is lump-sum transfer.

• Set $T$ to equalize consumption levels

$$W^C \geq \max_{\ell_1, \ell_2} \psi \left[ u \left( \frac{w_1 \ell_1 + w_2 \ell_2}{2}, \ell_1 \right) \right] + \psi \left[ u \left( \frac{w_1 \ell_1 + w_2 \ell_2}{2}, \ell_2 \right) \right]$$

• By setting $\ell_1 = \ell_2 = \ell^I$ and using $w_1 = (2 - w)$ and $w_2 = w$ we then obtain

$$W^C \geq \psi \left[ u \left( \ell^I, \ell^I \right) \right] + \psi \left[ u \left( \ell^I, \ell^I \right) \right] = W^1.$$
Proposition 1 Assume that preferences are represented by $u_i(x_i, \ell_i)$ and that wages can be chosen according to $w_1 + w_2 = 2$, with $w_i \geq w$. When individual types are observable (so that personalized lump-sum transfers are available), the level of welfare achieved under maximum wage differentiation (with $w_i = w$ and $w_j = 2 - w$) is always at least as high as (and generally higher than) the level of welfare achieved under wage equalization (with $w_1 = w_2 = 1$).
Intuition

- Consider the simplest case with $w = 0$
- If we go from equal wages to way $w_2 = 2$, individual 2’s wage is effectively doubled
- We can produce the same output as under equal wages by having only one individual work!
The second-best

- Assume that only income \( y_i = w_i \ell_i \) is publicly observable, while wages and labor supplies are private information.

- Case for equalization of wages appears to be considerably strengthened. No need for distortionary redistribution of incomes!

- Quite surprisingly, when \( w = 0 \), maximal wage differentiation continues to dominate even in a second-best setting.

- Show this through two propositions:
  - First one: limited scope but interesting intuition.
  - Second one: general.
• Second-best problem: first-best problem plus two incentive constraints

\[ u(x_i, \ell_i) \geq u(x_j, \tilde{\ell}_j) \quad i, j = 1, 2; i \neq j, \]

where \( \tilde{\ell}_j = w_j \ell_j / w_i \) is the labor supply of individual \( i \) mimicking individual \( j \).

• Assume \( w = 0 \) and, say \( w_1 = 0 \) and \( w_2 = 2 \).

• Assume for the time being that the first-best allocation implies

\[ u(x_2, \ell_2) \geq u(x_1, 0) \]

• In that case the full information optimum remains implementable under asymmetric information

• Incentive constraint of type 2 individuals is satisfied because with \( w_1 = 0 \) we have \( \tilde{\ell}_2 = \ell_1 = 0 \)
Proposition 2 Assume $w = 0$. When the first-best solution with $w_i = 0$ and $w_j = 2 (i, j = 1, 2; i \neq j)$ implies $u(x_j, \ell_j) \geq u(x_i, 0)$, then it can also be implemented if individual incomes are observable while wage and labor supplies are not observable.

- Applies for Rawlsian SWF
- But not in utilitarian with separable preferences
- When proposition applies: second-best = first-best
- When it does not apply: second-best $\neq$ first-best, but maximum differentiation remains optimal (when $w = 0$)
General argument (with $w = 0$)

- Best solution with equal wages: $x_i = \ell_i = \ell^I$; individuals are treated identically $\implies$ both a first- and a second-best solution

- Consider again an allocation with maximum wage differentiation $(w_1 = 0, w_2 = 2)$

- Set $\ell_2 = \ell^I$ and define $\hat{T}$ such that
  \[ u\left(\hat{T}, 0\right) = u\left(\ell^I, \ell^I\right) \]

- Observe that $T < \ell^I$. 
• Claim: allocation defined by \( x_1 = \hat{T}, \ell_1 = 0, x_2 = 2\ell^I - \hat{T} \) and \( \ell_2 = \ell^I \), is

- feasible
- Pareto dominates the equal wage solution \( x_i = \ell_i = \ell^I \)
- is incentive compatible

• Pareto dominance: obvious for 1; for 2, \( \hat{T} < \ell^I \) implies

\[
u(x_2,\ell_2) = u(2\ell^I - \hat{T}, \ell^I) > u(\ell^I, \ell^I).
\]

• Combining previous expressions yields

\[
u(x_2,\ell_2) > u(\hat{T}, 0),
\]

\(\implies\) incentive compatibility.
Proposition 3 Assume \( w = 0 \). Under maximum wages differentiation with \( w_i = 0 \) and \( w_j = 2 \) \((i, j = 1, 2; i \neq j)\) there exists a feasible allocation defined by \( x_i = \hat{T}, \ell_i = 0, x_j = 2\ell^I - \hat{T} \) and \( \ell_j = \ell^I \) (where \( \hat{T} \) satisfies \( u(\hat{T}, 0) = u(\ell^I, \ell^I) \)) that Pareto dominates the best equal wage solution \((x_i = \ell_i = \ell^I)\), and is incentive compatible.

- Remark: assumption \( w = 0 \) plays a crucial role in the proofs to both proposition

- With \( w > 0 \) wage equalization cannot be ruled out in second-best, see numerical example
Numerical example

- Quadratic disutility for labor: in the *laissez-faire*:

\[ u_i = w_i \ell_i - \ell_i^2 / 2 = w_i^2 / 2. \]

- Represent (utilitarian) social indifference curves in the plane \((w_1, w_2)\):

\[ S = \frac{w_1^2}{2} + \frac{w_2^2}{2}, \]
Figure 2:
Second best solution

Welfare as a function of $w_2$, with $w_1 = 2 - w_2$. 

25
Conclusion

• To sum up, when (first- or second- best) income redistribution is available maximum differentiation of productivity (concentrating education expenses on a single type) may be best policy, even when individuals are \textit{ex ante} identical.

• Non convexities in the underlying problem

• Assumptions:
  
  – exogenous (total) education expenditures
  – identical individuals \textit{ex ante}
  – no decreasing returns to scale in education technology
• Relaxing assumptions:

– Endogenous education spending (concave production function)
– Different individuals (learning skill)
– Decreasing returns (threshold value)
• Other issues
  
  – equal opportunity
  
  – education may provide utility
  
  – political economy: redistribution through education may be more acceptable than redistribution of incomes