Optimal Income, Education and Bequest Taxes in an Intergenerational Model

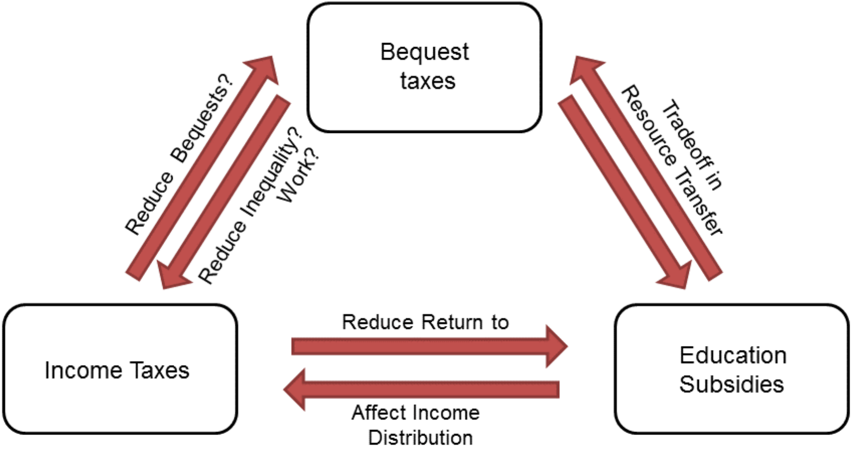
Stefanie Stantcheva
(Harvard Society of Fellows)

May 1, 2015

Introduction

- Parents can transfer resources to children through education or bequests.

Interplay between various tax instruments



An Intergenerational Model of Bequests and Education

- Dynamic intergenerational model *à la* Barro-Becker: altruistic preferences.
- Parents can transfer resources in two ways:
  - **Bequests** yield safe, uniform return.
  - **Education** yields idiosyncratic return: persistent, stochastic "ability."
- Wage of child = f(education, ability)
  - "Ability:" broad, multi-dimensional, exogenous component.
- Government: maximize expected welfare of today's generation.
  - Baseline tools: linear education subsidy, income taxes, bequest taxes.
  - Extend to fully unrestricted mechanism (the "best" we could possibly do).

Goal 1: Derive Simple Operational Optimal Formulas

- For education subsidy, bequest tax, income tax:
- In terms of estimable sufficient stats
  - robust to heterogeneity in preferences and primitives.
- Given all other (not necessarily optimally set) taxes.
- Isolating each tool's redistributive impact.
  - Can use generalized social welfare weights to accommodate any redistributive preferences.
- First, intuition from one-period model. Dynamic formulas look like static ones with appropriately redefined elasticities (of long-term tax base).

Goal 2: How should tax system account for bequests and education investments?

- Should parental human capital expenses be fully tax-deductible?
  - "Siamese Twins" result, Bovenberg and Jacobs (2005).
- Not generally true unless relative efficiency cost = relative distributional effect for bequests and education investments.
- Education subsidies and income taxes need not co-move.
- Bequest and income taxes need not co-move.
- Extend to OLG model to capture credit constraints: will typically $\uparrow$ optimal education subsidy, not change bequest tax.

Goal 3: Introduce and Use Reform Specific Elasticities

- Hard to estimate relevant elasticities in practice: can we target formulas to existing reforms?
- Yes: For any reform: can derive optimal formulas using "reform-specific elasticities."

Goal 4: Solve for Fully Unrestricted Taxes

- Mechanism design approach.
- Optimal to distort parental trade-off between education and bequests.
  - Except in very special case in which Hicksian coefficient of complementarity $\rho_{\theta s} = 1$ for kids.
  - I.e., only if wage $=$ ability $\times$ education.
- If education benefits mostly less able kids – should subsidize it relative to bequests (who benefit everybody equally).

Related Literature

Human Capital: Heckman (1976), Heckman, Lochner and Taber (1997), Heckman, Lochner and Todd (2006), Cunha and Heckman (2007, 2008).

Human Capital and Taxation: Bovenberg and Jacobs (2005), Jacobs (2007), Stantcheva (2014), Findeisen and Sachs (2014).

Bequest taxation: Piketty and Saez (2013), Farhi and Werning (2010, 2013).

Quantitative models with bequests: Krueger and Ludwig (2013, 2014).

Credit constraints for education: Carneiro and Heckman (2002), Jacobs and Yang (2011), Lochner and Monge-Naranjo (2011, 2012).

Outline

Education investments and bequests

- Agents live for 1 period: born, have single child, die.
- Agent from dynasty i at generation t denoted ti.
- Parents in generation t purchase education s_{t+1i} for child.
- Ability θ: stationary, ergodic process with correlation between generations (possibly, multidimensional).
- Wage: $w_{ti}(s) \equiv w(s, \theta_{ti})$
  - How complementary are education and ability ($\frac{\partial^2 w}{\partial\theta\partial s}$)?
  - Early Childhood investments vs. College?
  - Wlog, different types of human capital: $w(s_1, ...s_N, \theta_{ti})$.
- Income: $y_{ti} = w_{ti} l_{ti}$.

Dynastic Setup and Taxes

- Flow utility: $u_{ti}(c, y, s) \equiv u\left(c, \frac{y}{w(s,\theta_{ti})}; \eta_{ti}\right)$
- Expected utility of dynasty i

  $U_{1i} = E\left(\sum_{t=1}^{\infty} \beta^{t-1} u_{ti}(c_{ti}, y_{ti}, s_{ti})\right)$
- Bequests left by generation t, b_{t+1i}, yield R.
- Linear taxes: τ_{Lt}, τ_{St}, τ_{Bt}.
- G_t: lump-sum demogrant.
- Agents' per-generation budget constraint:

  $c_{ti} + b_{t+1i} + (1 - \tau_{St}) s_{t+1i} = Rb_{ti}(1 - \tau_{Bt}) + w_{ti}(s_{ti}) l_{ti}(1 - \tau_{Lt}) + G_t$

Equilibrium and Government Budget

- Aggregate (or per capita): y_t, b_t, and s_t.
- Stochastic processes for θ and η assumed to be ergodic.
  - at constant $(\tau_L, \tau_B, \tau_S, G)$, unique ergodic steady state independent of initial distribution of s_{1i} and b_{1i}.
  - If tax policy $(\tau_{Lt}, \tau_{Bt}, \tau_{St}, G_t)$ converges to long-run constant policy $(\tau_L, \tau_B, \tau_S, G)$ then s_{t+1}, y_t, and b_t also converge to steady state levels and depend on steady tax policies.
- Government budget constraint in equilibrium (per period):

$$G_t = \tau_{Lt} y_t + \tau_B R b_t - \tau_{St} s_{t+1}$$

  - With golden rule followed, such that $\beta = 1/R$, this is wlog.

Outline

Simple One-period Version of the Model

- Utility: $U_i = u_i\left(c_i, y_i, s_i\right)$
- Budget constraint: $c_i + \left(1 - \tau_S\right) s_i = w_i\left(s_i\right) l_i \left(1 - \tau_L\right) + G$
- Social Welfare: $SWF = \int_i \omega_i u_i\left(c_i, y_i, s_i\right) di$
  - For any set of Pareto weights $\{\omega_i\}_i$.
- Government BC: $G = \tau_L y - \tau_S s$

Elasticities and Distributional Characteristics

- Aggregate elasticities of y and s to $1 - \tau_L$:

  $\varepsilon_Y \equiv \frac{dy}{d(1-\tau_L)}\frac{(1-\tau_L)}{y}, \quad \varepsilon_S^Y \equiv \frac{ds}{d(1-\tau_L)}\frac{(1-\tau_L)}{s}$

- Aggregate elasticities of y and s to $\tau_S - 1$:

  $\varepsilon_S \equiv \frac{ds}{d(\tau_S-1)}\frac{(\tau_S-1)}{s}, \quad \varepsilon_Y^S \equiv \frac{dy}{d(\tau_S-1)}\frac{(\tau_S-1)}{y}$

- Distributional characteristic of output and education:

$$\bar{y} \equiv \frac{\int_i \omega_i u_{c,i} y_i di}{y \int_i \omega_i u_{c,i} di} \qquad \bar{s} \equiv \frac{\int_i \omega_i u_{c,i} s_i di}{s \int_i \omega_i u_{c,i} di}$$

- $\bar{s}$ large if s concentrated among high u_c (low c) agents
  - If s and ability not very complementary (Early Childhood Investments)?
  - $\bar{s}$ depends on what type of human capital subsidized (free public education?)
- $\bar{y}$ « 1 typically.

Optimal Static Linear Tax and Subsidy

- Optimal Labor Tax:

$$\tau_L^* = \frac{1 - \bar{y} - \tau_S \frac{s}{y} \varepsilon_S^Y}{1 - \bar{y} + \varepsilon_Y}$$

- Typical trade-off between redistribution ($1 - \bar{y}$) and efficiency (ε_Y).
- Fiscal spillover on education tax base: $\tau_S \frac{s}{y} \varepsilon_S^Y$ (0 if $\tau_S = 0$.)
- Optimal Education Subsidy:

$$\tau_S^* = \frac{1 - \bar{s} + \frac{y}{s} \varepsilon_Y^S \tau_L}{1 - \bar{s} + \varepsilon_S}$$

- Redistributive effect of education $(1 - \bar{s}) \uparrow \tau_S$.
    - $(1 - \bar{s})$ large for Early Childhood Investment.
- Fiscal spillover: $\frac{y}{s} \varepsilon_Y^S \tau_L$ increasing in τ_L.

"Siamese Twins Result" Revisited

- Benchmark: Full deductibility of education expenses.
  $\tau_S = \tau_L \Leftrightarrow$ equivalent to taxable income being $y - s$.

- Full deductibility optimal iff:

$$\frac{\left(\frac{y}{s}\varepsilon_Y^S - \varepsilon_S\right)}{\left(\frac{s}{y}\varepsilon_S^Y - \varepsilon_Y\right)} = \frac{(1-\bar{s})}{(1-\bar{y})}$$

- If $1 - \bar{s} >> 1 - \bar{y}$, then optimal to have: $\tau_S^* > \tau_L^*$.

- Bovenberg and Jacobs (2005) find $\tau_S = \tau_L$, because:
  - $w = \theta s$ and quasilinear utility.
  - Hence: $\bar{y} = \bar{s}$, $\varepsilon_Y^S = \gamma$, $\varepsilon_Y = 1 - \gamma$, $\varepsilon_S^Y = -\gamma$, $\varepsilon_S = \gamma - 1$

Outline

A Variational Approach – One instrument at a time

- Social Welfare:

$$SWF = \max E \sum_{t=1}^{\infty} \beta^{t-1} [u_{ti}((1-\tau_{Lt})y_{ti} - s_{t+1i}(1-\tau_{St})) \\ + R(1-\tau_{Bt})b_{ti} - b_{t+1i} + G_t, y_{ti}, s_{ti})]$$

  subject to

$$G_t = \tau_{Lt} y_t + \tau_{Bt} R b_t - \tau_{St} s_{t+1}$$

- Variation: $d\tau_{St} = d\tau_S$ for $t > T$.
- $dSWF$ = direct welfare (by envelope theorem) + mechanical revenue effect + behavioral effects (anticipatory and post-reform).

Elasticities of the Present Discounted Tax Bases

- Long run elasticities of PDV of tax bases:

$$\varepsilon_S' \equiv (1-\beta)\sum_{t\geq 1}\beta^{t-1-T}\varepsilon_{St+1}$$

$$\varepsilon_Y^{S\prime} \equiv (1-\beta)\sum_{t\geq 1}\beta^{t-1-T}\varepsilon_{Yt}^{S}$$

$$\varepsilon_B^{S\prime} \equiv (1-\beta)\sum_{t\geq 1}\beta^{t-1-T}\varepsilon_{Bt}^{S}$$

  - Mix both children's and parents' responses.
  - Mix income and substitution effects.

- Redistributive factors:

$$\bar{y} \equiv \frac{E\left(u_{c,ti}y_{ti}\right)}{E\left(u_{c,ti}\right)y_t}, \qquad \bar{s} \equiv \frac{E\left(u_{c,ti}s_{t+1i}\right)}{E\left(u_{c,ti}\right)s_{t+1}}, \qquad \bar{b} \equiv \frac{E\left(u_{c,ti}b_{ti}\right)}{E\left(u_{c,ti}\right)b_t}$$

Optimal Linear Taxes and Subsidies

- Optimal education subsidy:

$$\tau_S^* = \frac{1 - \bar{s} + \varepsilon_Y^{S\prime} \tau_L \frac{y}{s} + \varepsilon_B^{S\prime} \tau_B R \frac{b}{e}}{1 - \bar{s} + \varepsilon'_S}$$

  - Decreasing in ε'_S (like in static, but now it's elasticity of full base).
  - Tax deductibility not optimal in general: τ_S and τ_L need not even co-move (unless no income effects).
  - τ_S and τ_B may or may not co-move (substitution vs. income effects).
- Can use formula to evaluate reforms (at any given τ_B and τ_L).
  - Maybe most useful application, only requires knowing ε, $\bar{s}$ at status quo.
- Distributional effects again crucial.
  - Depend on complementarity and institutional setup.
- Can use generalized Social Welfare Weights (Saez and Stantcheva 2014).

Generalized Social Welfare Weights

- Instead of standard weights derived from SWF ($\omega_{ti} u_{c,ti}$), use **generalized social welfare weights** g_{ti}
  - g_{ti}: Social marginal value of giving \$1 to person i.

$$\bar{s} = \frac{E(g_{ti} s_{ti})}{E(g_{ti}) s_t}, \qquad \bar{y} = \frac{E(g_{ti} y_{ti})}{E(g_{ti}) y_t}, \qquad \bar{b} = \frac{E(g_{ti} b_{ti})}{E(g_{ti}) b_t}$$

- All redistributive considerations translate into different values for $\bar{s}$, $\bar{y}$, $\bar{b}$.
  - No need to rederive anything.
  - No SWF, only variations/reforms.
- Rawlsian case: $\bar{s} = 0$.
- Pure Efficiency consideration: $\bar{s} = 1$.
- Value altruistic parents most: $\bar{s} >> 1$.
- Worry about kids from poor background: $\bar{s} = \frac{E(s_{ti}|\text{poor background})}{Prob(\text{poor background}) s_t}$.

Optimal Linear Taxes and Subsidies

- Optimal Bequest Tax:

$$\tau_B^* = \frac{1 - \bar{b} + \varepsilon_S^{B\prime} \frac{s}{b} \tau_S - \varepsilon_Y^{B\prime} \tau_L \frac{y}{b}}{1 - \bar{b} + \varepsilon'_B}$$

- Generically not zero – contrast to zero capital taxation result (Chamley, Judd):
- Fiscal spillover/constraint on other tax instruments.
- ε'_B finite (true with uncertainty), breaks down with perfect certainty.
- $\bar{b} \neq 1$: except if utility linear in c, or purely accidental bequests uncorrelated with income.

Reform-Specific Elasticities

- What if we cannot estimate all cross-elasticities needed?
- Target formulas to specific reforms (shifts in several instruments), and care only about total effect. Formulas are "reform-specific."
- E.g.: $d\tau_{St} = d\tau_S$ for $t > T$, with $d\tau_{Lt}$ to maintain budget balance, τ_B unchanged.
- Optimal education subsidy with reform-specific elasticities:

$$\tau_S^* = \frac{1 - \frac{\bar{s}}{\bar{y}}\left(1 - {\varepsilon'}_Y \frac{\tau_L}{1-\tau_L}\right) + R\frac{b}{s}{\varepsilon'}_B \tau_B}{1 - \frac{\bar{s}}{\bar{y}}\left(1 - {\varepsilon'}_Y \frac{\tau_L}{1-\tau_L}\right) + {\varepsilon'}_S}$$

- Long-run elasticities ${\varepsilon'}_B$, ${\varepsilon'}_Y$ and ${\varepsilon'}_S$ estimated from a revenue neutral reform changing τ_S and adjusting τ_L for budget balance.

Reform-Specific Elasticities: Discussion

- Most useful formulation for reforms that have been done so can use "ready" estimates.
- Best to evaluate reforms around status quo where elasticities estimated.
- If we knew primitives (Slutsky matrices), formulas are equivalent.
- Not necessary to assume that τ_L or τ_B optimally set.

Unobservable Education or Human Capital Spending

- Need to provide indirect incentive for human capital *indirectly* through labor and bequest tax only.
- Optimal labor tax with unobserable education:

$$\tau_L^{*,unobs} = \frac{1 - \bar{y} - \frac{b}{y}\varepsilon_B^{Y\prime}\tau_B}{1 - \bar{y} + \varepsilon_Y^{\prime}}$$

- If $\varepsilon_S^{Y\prime} < 0$, then if $\tau_S^* > 0$ was optimal, τ_L lower with unobservable education.
- Optimal bequest tax with unobserable education:

$$\tau_B^{*,unobs} = \frac{1 - \bar{b} - \varepsilon_Y^{B\prime}\tau_L\frac{y}{b}}{1 - \bar{b} + \varepsilon_B^{\prime}}$$

- If education and bequests substitutes overall, $\varepsilon_S^{B\prime} < 0$, and if $\tau_S^* > 0$ had been optimal, τ_B higher to indirectly encourage education.

Outline

An Augmented Dynastic OLG Model

- Generation t born at time t lives for 3 periods:
  1. "Young:" receive s_t from their parents.
  2. "Adult:" have one child each, work to earn y_{t+1}, save k_{t+1}, invest s_{t+1}.
  3. "Old:" Receive bequests b_{t+1} at beginning of period, consume, leave bequests b_{t+2}, die.
- Unit mass of each young, adult, and old at each t.
- Inelastic labor supply for exposition only: $y_{t+1i} = w_{t+1}(s_{ti}, \theta_{t+1i})$.
- Utility (realized in old age at time $t+2$): $u_{t+2}(c_{t+2i}, \eta_{t+2i})$.
- Budget constraint of adult i from generation t:
  $(1-\tau_{Lt+1})w_{t+1}(s_{ti}, \theta_{t+1i}) = k_{t+1i} + s_{t+1i}(1-\tau_{St+1})$
- Budget constraint of old agent i from generation t:
  $k_{t+1i} + Rb_{t+1i}(1-\tau_{Bt+2}) = c_{t+2i} + b_{t+2i}$

Government Transfers, SWF and Credit Constraints

- G_t given at beginning of old age (after bequests received have been taxed). Transfer at time t (to old of generation $t-2$):
  $G_t = \tau_{Lt-1} y_{t-1} + \tau_{Bt} R b_{t-1} - \tau_{St-1} s_{t-1}$
- Social Welfare:
  $SWF_0 = \max E \sum_{t=1}^{\infty} \beta^{t-1} [u_{ti}((1-\tau_{Lt-1}) y_{t-1i} - s_{t-1i}(1-\tau_{St-1})$
  $+R(1-\tau_{Bt}) b_{t-1i} - b_{t+1i} + G_t)]$
- If no credit constraints: all periods collapsed into 1, equivalent to before.
- Credit constraints: $k_t = (1-\tau_{Lt}) w_t(s_{t-1}, \theta_t) - s_t(1-\tau_{St}) \geq 0$, multiplier γ_{ti}.
- Redistributive incidence of credit constraints: $\tilde{s} \equiv \frac{E(\gamma_{ti} s_{t-1i})}{E(u_{c,ti}) s_{t-1}}$
- $\tilde{s}$ higher if credit constraints hit mostly parents who invest a lot in s.

Government Transfers, SWF and Credit Constraints

- Optimal human capital subsidy:

$$\tau_S^{*,cc} = \frac{1-(\bar{s}+\tilde{s})+\varepsilon_Y^{S\prime}\tau_L\frac{y}{s}+\varepsilon_B^{S\prime}\tau_B R\frac{b}{s}}{1-(\bar{s}+\tilde{s})+\varepsilon'_S}$$

- Additional term $\tilde{s}$ acts exactly like $\bar{s}$.
- Credit constraints concentrated among parents who invest a lot in their children $\Leftrightarrow$ high social marginal value on parents investing a lot.
- Tend to increase optimal human capital subsidy, all else equal.
- Bequest tax unchanged: bequests occur too late in life to relieve credit constraints. Could change?

Outline

Optimal Unrestricted Mechanism: Setup

- Simplify: no preference shocks η.
- θ_t follows Markov process $f^t(\theta_t|\theta_{t-1})$.
  - Parents have some advance info, but not full info about kids' abilities.
- Utility separable: $\tilde{u}_t(c_t, y_t, s_t; \theta_t) = u_t(c_t) - \phi_t\left(\frac{y_t}{w_t(\theta_t, s_t)}\right)$
- Key parameter: **Hicksian coefficient of complementarity** between ability and education in the wage function

  $\rho_{\theta s} \equiv \frac{w_{\theta s} w}{w_s w_\theta}$

  - $\rho_{\theta s} < 0$: lower ability kids have a higher marginal benefit from education (Early Childhood Investments, evidence from J. Heckman).
  - $\rho_{\theta s} > 0$: higher ability kids have a higher marginal benefit from education (Heckman and Cunha evidence for College).
  - $\rho_{\theta s} > 1$: higher ability kids have a higher *proportional* benefit from education (Wage elasticity w.r.t ability increasing in education).

Solution Method: First-order Approach + Dynamic Programming

- Farhi and Werning (2013) and Stantcheva (2014).

- Imagine direct revelation mechanism: specify allocations as functions of reported θ^t.

- Continuation utility of the dynasty after history θ^t:

$$\omega\left(\theta^t\right) = u_t\left(c\left(\theta^t\right)\right) - \phi_t\left(\frac{y\left(\theta^t\right)}{w_t\left(\theta_t, s\left(\theta^t\right)\right)}\right) + \beta\int \omega\left(\theta^{t+1}\right) f^{t+1}\left(\theta_{t+1}|\theta_t\right) d\theta_{t+1}$$

- Replace by "envelope condition:"

$$\dot{\omega}\left(\theta^t\right) := \frac{\partial\omega\left(\theta^t\right)}{\partial\theta_t} = \frac{w_{\theta,t}}{w_t} l\left(\theta^t\right)\phi_{l,t}\left(l\left(\theta^t\right)\right) + \beta\int \omega\left(\theta^{t+1}\right)\frac{\partial f^{t+1}\left(\theta_{t+1}|\theta_t\right)}{\partial\theta_t} d\theta_{t+1} \tag{1}$$

Rewrite Problem Recursively

- Rewrite problem recursively using: promised continuation utility v, promised marginal continuation utility Δ.

- The program of the government is:

$$K(v, \Delta, \theta_-, t) = \min \int (c(\theta) + s_{t+1}(\theta) - w_t(\theta, s_t(\theta)) l(\theta) + \frac{1}{R} K(v(\theta), \Delta(\theta), \theta, s_{t+1}(\theta), t+1)) f^t(\theta|\theta_-) d\theta$$

subject to:

$$\omega(\theta) = u_t(c(\theta)) - \phi_t(l(\theta)) + \beta v(\theta)$$

$$\dot{\omega}(\theta) = \frac{w_{\theta,t}}{w_t} l(\theta) \phi_{l,t}(l(\theta)) + \beta \Delta(\theta)$$

$$v = \int \omega(\theta) f^t(\theta|\theta_-) d\theta$$

$$\Delta = \int \omega(\theta) \frac{\partial f^t(\theta|\theta_-)}{\partial \theta_-} d\theta$$

maximization is over functions $(c(\theta), l(\theta), s(\theta), \omega(\theta), v(\theta), \Delta(\theta))$.

Characterize Marginal Distortions Using Wedges

- Distortions relative to *laissez-faire* characterized by "wedges" (pure definitions):

- Intratemporal wedge on labor $\tau_L\left(\theta^t\right)$

$$\tau_L\left(\theta^t\right) \equiv 1 - \frac{\phi_{l,t}(l_t)}{w_t u_t'\left(c_t\right)}$$

- Intertemporal wedge on bequests

$$\tau_B\left(\theta^t\right) \equiv 1 - \frac{1}{R\beta}\frac{u_t'\left(c_t\right)}{E_t\left(u_t'\left(c_{t+1}\right)\right)}$$

Optimal Relation between Bequests and Education

- ε_t^u: uncompensated labor supply elasticity
- ε_t^c: compensated labor supply elasticity (all holding savings fixed).
- At the optimum:

$$R = E\left(w_{s,t+1} l_{t+1} (1 + \tau_{Lt+1} \frac{\varepsilon_{t+1}^c}{1+\varepsilon_{t+1}^u} (1 - \rho_{\theta s,t+1})) \right)$$

- LHS = Return to bequests.
- RHS = ***Social*** return to education = Private return + incentive effect.
- Bequests affect everybody equally, but education does not.

Subsidizing or Taxing Education Relative to Bequests

Education subsidized relative to bequests $\Leftrightarrow \rho_{\theta s,t} \leq 1$

Labor Supply Effect:
Education subsidy increases children's wage
$\rightarrow \uparrow$ labor
$\rightarrow \uparrow$ resources.

$+$

Inequality Effect:
if $\rho_{\theta s} \geq 0$, education benefits more able children more
$\rightarrow \uparrow$ *pre-tax* inequality.

$\rho_{\theta s} \leq 1 \Rightarrow$ subsidy $\downarrow$ *post-tax* inequality
$\Rightarrow$ has **positive redistributive and insurance effects**.

$\rho_{\theta s} = 1 \Rightarrow$ No distortion between bequests and education
Benchmark case in literature $w_t = \theta_t s_t$

Conclusion

- Derive formulas for optimal linear taxes as functions of estimable behavioral elasticities and redistributive factors, robust to heterogeneities and preferences.
  - "Reform elasticities" adapted to existing reforms.
- Not optimal to make education expenses fully tax deductible, as education subsidies have differential distributional impacts.
  - τ_S, τ_B, τ_L can co-move positively or negatively...
- Credit constraints would typically increase optimal education subsidy.
- Fully unrestricted mechanism: if education highly complementary to ability ($\rho_{\theta s} > 1$), tax education relative to bequests.