Optimal Nonlinear Income Tax when Skills and Labor Supply Elasticities are Heterogeneous

Laurence JACQUET
THEMA - University of Cergy-Pontoise

Etienne LEHMANN
CRED - University Panthéon Assas Paris 2

June 15, 2015

Motivation

The Mirrlees (1971) model assumes unobserved heterogeneity to be one-dimensional, which is very restrictive.

- Case with heterogeneous skills and labor supply elasticities,
- Joint taxation of labor and non-labor income (e.g. entrepreneurial income, capital income, income received from renting property,...),
- Nonlinear joint income taxation of households,
- Income taxation with labor supply responses and tax avoidance/evasion responses,
- Optimal nonlinear monopoly pricing when consumers differ in the slope and the intercept of their demand curves.
- Any adverse selection problem with a one-dimensional observed action, many unobserved types and some additive separability in preferences.

⇒ We extend the Mirrlees model to the case with multidimensional unobserved heterogeneity and one observable action (taxable income).

Our contributions

- We characterize the set of individuals of different types who pool at the same income from the equality of their marginal rates of substitution.
- We derive a structural optimal tax formula in terms of the primitives of the model (Proposition 1), using an allocation perturbation method.

⇒ We use this structural formula to find cases where optimal marginal tax rates are positive (Proposition 2).

⇒ We rewrite our structural formula in terms of sufficient statistics (social welfare weights, behavioral elasticity, income density) (Proposition 3).

⇒ We use this structural formula to provide a numerical investigation of how much heterogeneous taxable income elasticities matter for the tax schedule Numerics and for the asymptotic tax rate.

The sufficient statistics approach to optimal taxation

Piketty (RFE 1997), Saez (REStud 2001), Diamond and Saez (JEP 2011), Piketty and Saez (Handbook PE 2013), Golossov Tsyvinsky Werquin (2014), Hendren (2014)

- Derive an optimal marginal tax formula in terms of sufficient statistics using an heuristic proof based on tax perturbation.
- Because sufficient statistics are endogenous, a formula expressed in terms of the policy-invariant primitives is required for the simulations.
- Saez (2001) needs to rely on Mirrlees (1971)'s formula, which is only valid when the unobserved heterogeneity is one-dimensional.
- Saez (2001) conjectures his sufficient statistics formula is also valid in the multidimensional case, but did not prove it.

We rigorously derived a primitive-based formula, which we use in simulations and in verifying the validity of Saez (2001)'s conjecture.

- Papers with a one-dimensional aggregation of multidimensional types:
    - Boadway, Pestieau & Racionero (JPET 2006), Choné & Laroque (AER 2010), Lockwood & Weinzierl (WP 2013) to show cases where optimal marginal tax rates may be negative.
    - Rochet & Stole (REStud 2002), Kleven, Kreiner & Saez (Ectrca 2009), J, L & Van der Linden (JET 2013), Blumkin Sadka & Shem-Tov (2013), L, Simula & Trannoy (QJE 2014) to introduce extensive responses.
    - Rotschild & Scheuer (QJE 2013, NBER 2014a&b), Scheuer (JET 2013, AEJ: EcoPol 2014), Gomes Lozachmeur & Pavan (2014): an aggregator that also depends on the price vector to include general equilibrium effects on the wage distribution.

    $\Rightarrow$ 2 individuals who earn the same income must have the same value for the aggregator, thus face the same income decision program and cannot have distinct labor supply elasticities.
- Papers with n observed actions and $m \leq n$ types: Mirrlees (1976, JPubE), Rochet & Choné (Ectrca 1998), Kleven, Kreiner & Saez (CESifo 2007), Renes & Zoutman (2014)...

Individual preferences

- Individuals of skill $w \in \mathbb{R}_+$ in "group" $\theta \in \Theta$ have preferences over consumption (after-tax income) $c \in \mathbb{R}_+$ and pre-tax income $y \in \mathbb{R}_+$:

$$\mathscr{U}(c, y; w, \theta) = u(c) - v(y; w, \theta) \qquad \text{with} \qquad u', v'_y, v''_{yy} > 0 > v'_w, u''$$

- CDF of θ is $\mu(\theta)$ over the potentially multidimensional set Θ.
- The conditional skill density is $f(\cdot|\theta)$ with support $\mathbb{R}_+$.
- The first-order condition is:

$$1 - T'(Y(w, \theta)) = \frac{v'_y(Y(w, \theta); w, \theta)}{u'(C(w, \theta))} \tag{1}$$

where $\frac{v'_y(y; w, \theta)}{u'(c)}$ is the Marginal Rate of Substitution (MRS).

Assumption (1 - Within Group Single-Crossing condition)

$$v''_{yw}(y; w, \theta) < 0 \qquad \Leftrightarrow \qquad \textit{MRS decreasing in } w$$

and limiting conditions:

$$\lim_{w \mapsto 0} v'_y(y; w, \theta) = +\infty \qquad \textit{and} \qquad \lim_{w \mapsto +\infty} v'_y(y; w, \theta) = 0$$

Leading example:

$$\mathscr{U}(c, y; w, \theta) = u(c) - \left(\frac{y}{w}\right)^{1+\frac{1}{\theta}}$$

where $\theta \in \Theta \subset \mathbb{R}_+$ is the (Frish) labor supply elasticity Back

The governement

The gvt maximizes a type-specific $\Phi(\cdot; w, \theta)$ social welfare function:

$$\int_{\theta\in\Theta} \left\{ \int_0^{+\infty} \Phi\left(U(w,\theta); w, \theta\right) \; f(w|\theta) \; dw \right\} d\mu(\theta)$$

subject to the budget constraint (multiplier $\lambda > 0$):

$$\int_{\theta\in\Theta} \left\{ \int_0^{+\infty} \left[Y(w,\theta) - C(w,\theta)\right] f(w|\theta) \; dw \right\} d\mu(\theta) \geq 0$$

and to incentive constraints (IC): $\forall (w, \hat{w}, \theta, \hat{\theta}) \in \mathbb{R}_+^2 \times \Theta^2$

$$u\left(C(w,\theta)\right) - v\left(Y(w,\theta); w, \theta\right) \geq u\left(C(\hat{w},\hat{\theta})\right) - v\left(Y(\hat{w},\hat{\theta}); w, \theta\right)$$

Within-group IC allocations

Incentive Constraints (IC) implies within-Group IC: $\forall (w, \hat{w}, \theta) \in \mathbb{R}_+^2 \times \Theta$:

$$u\left(C(w,\theta)\right) - v\left(Y(w,\theta); w, \theta\right) \geq u\left(C(\hat{w},\theta)\right) - v\left(Y(\hat{w},\theta); w, \theta\right)$$

Within each group θ:

$\Rightarrow$ $w \mapsto Y(w,\theta)$ is nondecreasing (Graphical proof)

$\Rightarrow$ the envelope first-order incentive constraint holds:

$$\dot{U}(w,\theta) = -v'_w(Y(w,\theta); w, \theta) \qquad \text{(IC1)}$$

Pooling versus bunching

Assumption (2 - Smooth allocations)

For each θ, $Y(\cdot,\theta)$ is differentiable, with $\dot{Y}(w,\theta) > 0$, $Y(0,\theta) = 0$ and $\lim_{w \mapsto \infty} Y(w,\theta) = \infty$.

- Bunching: Two individuals in the same group θ but different skill levels earn the same income.
- Pooling: Two individuals in the different groups earn the same income.
- Assumption 2 rules out bunching and makes pooling unavoidable.
- Bunching never occurs in our numerical experiments.
- Assumption 2 is verified with isoelastic preferences. Illustration
- In Rochet & Chone (1998), bunching occurs because there is also a participation constraint.

The pooling function

- The pooling problem: How should be set each within-group allocation $w \mapsto (Y(w,\theta), C(w,\theta))$ to be mutually incentive compatible?
- From Assumption 2, $w \xrightarrow{Y(\cdot,\theta)} Y(w,\theta)$ is increasing from 0 to ∞.

$\Rightarrow$ For each $y \in \mathbb{R}^+$, there exists a single w such that $Y(w,\theta) = y$.

- Take a reference group θ_0 and a skill level w, taking $y = Y(w,\theta_0)$, there thus exists a single skill denoted $W(w,\theta)$ such that:

$$Y(W(w,\theta),\theta) \overset{\text{AS 2}}{\equiv} Y(w,\theta_0) \text{ and } C(W(w,\theta),\theta) \overset{\text{IC}}{\equiv} C(w,\theta_0) \quad (2)$$

- The pooling function verifies : $w \xrightarrow{Y(\cdot,\theta_0)} Y(w,\theta_0) \xrightarrow{Y^{-1}(\cdot,\theta)} W(w,\theta)$ (and thus verifies Assumption 2)

Characterization of the pooling function

- The pooling function $W(w,\theta)$ provides the allocation for any group θ from the allocation $w \mapsto (C(w,\theta_0), Y(w,\theta_0))$ specific to group θ_0:

$$\omega \overset{W^{-1}(\cdot,\theta)}{\longrightarrow} W^{-1}(\omega,\theta) \overset{Y(\cdot,\theta_0)}{\longrightarrow} Y(\omega,\theta)$$

Lemma

Under Assumptions 1 and 2, the bundle $(C(w,\theta_0), Y(w,\theta_0))$ designed for types (w,θ_0) is also designed for types $(W(w,\theta),\theta)$ where $W(w,\theta)$ solves:

$$\frac{v'_y\left(Y(w,\theta_0); w,\theta_0\right)}{u'(C(w,\theta_0))} = \frac{v'_y\left(Y(w,\theta_0); W(w,\theta),\theta\right)}{u'(C(w,\theta_0))} \tag{3}$$

- From AS 1, a single skill level $W(w,\theta)$ solves (3) for each θ.

- If two individuals belonging to two different groups earn the same income $Y(w, \theta_0)$
  - $\Rightarrow$ They must face the same marginal tax rate $T'(Y(w, \theta_0))$
  - $\Rightarrow$ They must thus have the same MRS.
  - Formal proof: deriving both sides of the definition (2) and using the individual's foc (1).
- As $\mathscr{U}(c, y; w, \theta) = u(c) - v(y; w, \theta)$, the equality of MRS in Equation (3) simplifies to:

$$v'_y(Y(w, \theta_0); w, \theta_0) = v'_y(Y(w, \theta_0); W(w, \theta), \theta)$$

- We allow for endogenous pooling as it depends on $Y(\cdot, \theta_0)$.

- We consider a set of perturbations of the allocation in the reference group θ_0 such that $Y(\cdot, \theta_0)$ is only modified in the interval $[\omega - \delta\omega, \omega]$ by differentiable amounts $\Delta Y(w, \theta_0, t) = t\Delta Y(w, \theta_0)$, the perturbed allocations remaining increasing. Figure
- Because of the pooling condition, in group $\theta \neq \theta_0$, the allocations are perturbed such that such that $Y(\cdot, \theta)$ is only modified in the interval $[W(\omega - \delta\omega, \theta), W(\omega, \theta)]$ by some $\Delta Y(w, \theta, x)$. Figure
- By IC1, these perturbations induce no change in utility below $W(\omega - \delta\omega, \theta)$ but a uniform change in utility above $W(\omega, \theta)$. Figure
- As pooling and income $Y(\cdot, \theta)$ are not modified above $W(\omega, \theta)$, these uniform increases in utilities above $W(\omega, \theta)$ must be equal to the same $\Delta U(t)$ for all groups θ, which provides the average change in $\Delta Y(w, \theta, t)$ within $[W(\omega - \delta\omega, \theta), W(\omega, \theta)]$ as a function of the average change $\Delta Y(w, \theta_0, t)$ within $[\omega - \delta\omega, \omega]$.

- We normalize these perturbations by ΔU instead of t.
- We compute the Gateaux derivative with respect to these perturbations.
- We take the limit of this derivative when $\delta\omega$ tends to 0.
- Equating this limit to zero provides the necessary conditions:

The optimal tax formula

Proposition (1 - Structural optimal tax formula)

$$\frac{T'(Y(\omega,\theta_0))}{1-T'(Y(\omega,\theta_0))}\int_{\theta\in\Theta}\frac{v'_y\left\langle W(\omega,\theta),\theta\right\rangle}{-v''_{yw}\left\langle W(\omega,\theta),\theta\right\rangle}f(W(\omega,\theta)|\theta)\;d\mu(\theta)$$

$$=u'(C(\omega,\theta_0))\iint_{x\geq W(\omega,\theta),\theta\in\Theta}\left(\frac{1}{u'\left(C(x,\theta)\right)}-\frac{\Phi'_U\left\langle x,\theta\right\rangle}{\lambda}\right)f(x|\theta)dx\;d\mu(\theta)$$

$$0=\iint_{x\in\mathbb{R}_+,\theta\in\Theta}\left(\frac{1}{u'\left(C(x,\theta)\right)}-\frac{\Phi'_U\left\langle x,\theta\right\rangle}{\lambda}\right)f(x|\theta)dx\;d\mu(\theta)$$

These structural formulas being expressed in terms of policy-invariant primitives, it is numerically implementable to real data.

Sign of marginal Tax rates

Proposition (2 - Optimal marginal tax rates are positive)

Under a Benthamite ($\Phi(U; w, \theta) \equiv U$) or a Maximin government, optimal marginal tax rates are positive.

- The Benthamite objective consists in summing individuals' utility level, without any aversion to inequality in utility from the government.
- As $\Phi'_U = 1$, the same argument as in Mirrlees (1971) applies.
- Under Maximin, one has $\Phi'_U = 0$, so the argument of Mirrlees (1971) applies again.
- Benthamite and Maximin objectives imply that all individuals at income y are socially valued identically, unlike in Boadway, Pestieau & Racionero (JPET 2006), Choné and Laroque (AER 2010), Lockwood and Weinzierl (WP 2013).

Proposition (3 - An elasticity-based tax formula)

$$\frac{T'(y)}{1-T'(y)} = \frac{1}{\hat{\varepsilon}(y)} \cdot \frac{\int_y^{\infty} \{1-\hat{g}(z)-\hat{\eta}(z)\cdot T'(Y(z))\}\cdot \hat{h}(z)dz}{y\ \hat{h}(y)}$$

A rewriting of the structural tax formula where: Economic intuition

- $\hat{h}(\cdot)$ is the true income density at the optimum.
- $\hat{g}(y)$ is the mean of welfare weights across individuals who earn y.
- $\hat{\varepsilon}(y)$ is the mean of total compensated elasticity.
- $\hat{\eta}(y)$ is the mean of total income effects.

- $\hat{\varepsilon}(y)$ and $\hat{\eta}(y)$ are the mean of total behavioral responses, taking into account that any income response to a tax reform induces a change in marginal tax rates that triggers a further income response (circularity process due to the nonlinearity of the tax schedule).

$\Rightarrow$ Simulating the elasticity-based optimal tax formula requires in practice to neglect the circularity process, with no way to quantify by how much this approximation matters.

- The elasticity-based formula is useless to show that optimal marginal tax rates are positive under Benthamite or Maximin social objective whenever there is income effects.

Assymptotic marginal tax rates

- The elasticity-based formula is useful to retrieve the Saez (2001) asymptotic formula:
$$T'(\infty) = \frac{1}{1 + a(\infty)\varepsilon(\infty)}$$
where $a(\infty)$ is the asymptotic Pareto parameter of the income distribution.
- If there are two groups with iso-elastic preferences, different elasticity and slightly different asymptotic Pareto parameters, only the group with the fatter tail is present asymptotically and determine $\varepsilon(\infty)$, which thus can be very different from the mean elasticity in the top 1%.

Empirical illustration: By how much heterogeneity matters

- Assume $\mathscr{U}(c, y; w, \theta) = c - \left(\frac{y}{w}\right)^{1+\frac{1}{\theta}}$ and SWF $\Phi(.) = \log(.)$.
- 2 scenarios:
  - The heterogeneous scenario with the elasticity $\theta = 0.6$ for the self-employed and $\theta = 0.2$ for the salary workers.
  - The Mirrlees scenario with an single value of θ with the same sample mean as in the heterogeneous scenario.
- Derive the conditional skill density $f(\cdot|\theta)$ in each scenario using the true income tax schedule and the observed distribution of incomes of singles without dependent children from CPS 2013.
- Because of top-coding of income in CPS, add an exogenous mass at the top of the income grid to approximate a Pareto tail at the top of the skill distributions.

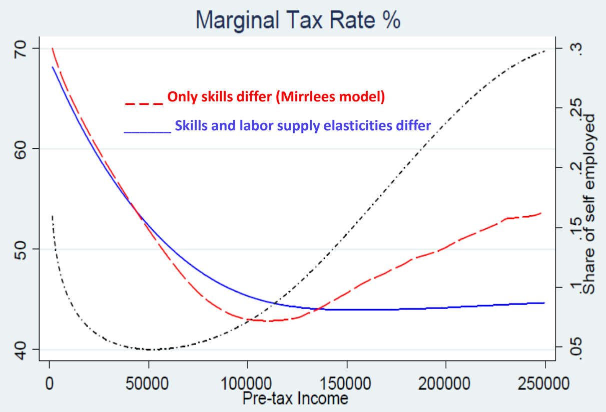


Figure 1: Optimal marginal tax rates (in the U.S.) when self-employed and salary workers have distinct skills and different labor supply elasticities.

Back

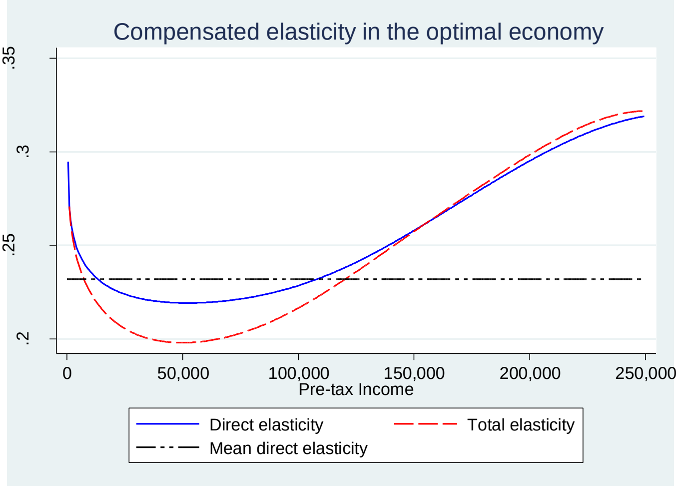
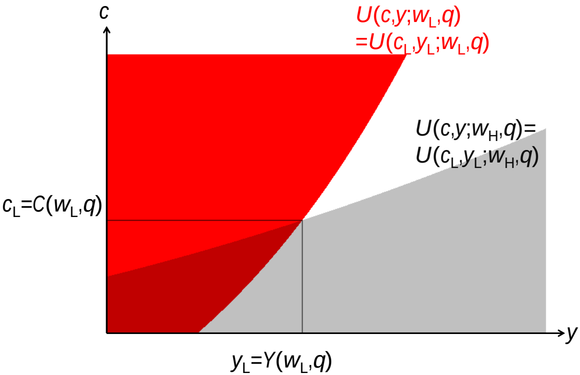
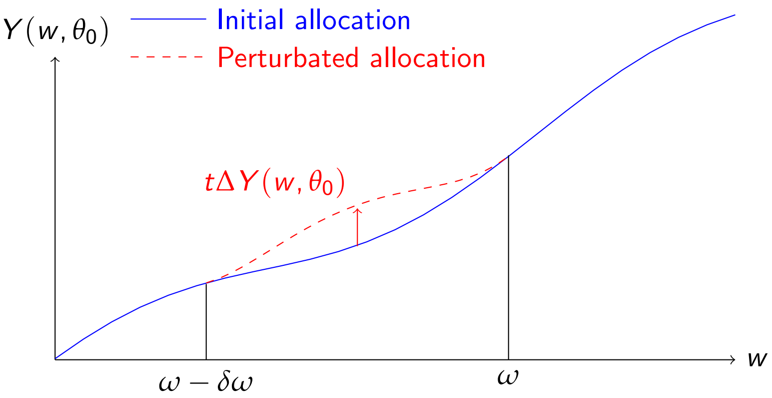


Figure 2: Mean compensated total and direct elasticities.

Thank you !

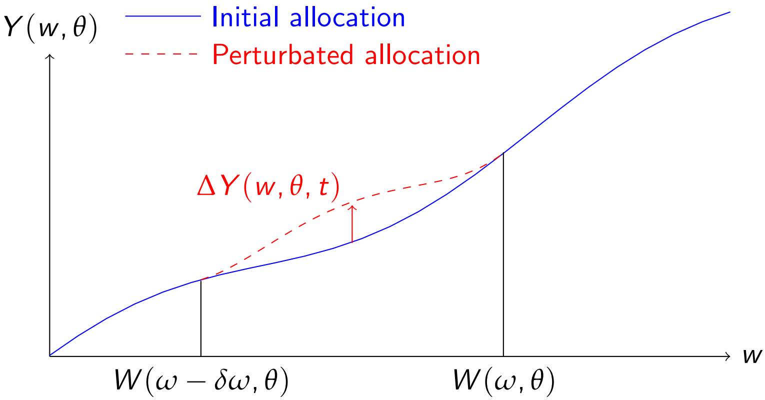


Back



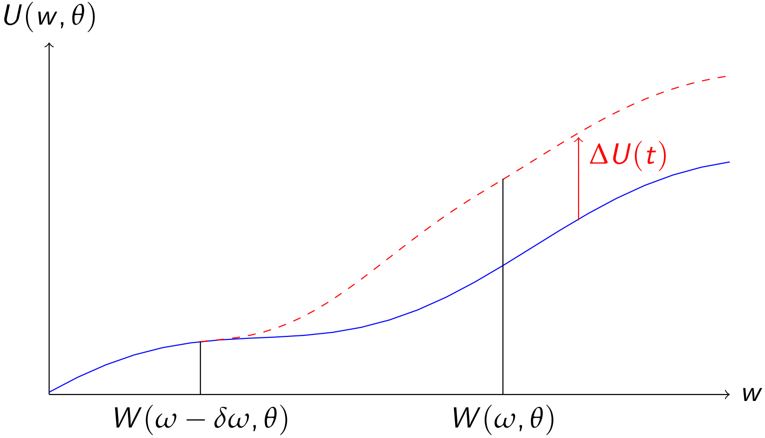
$Y(w, \theta_0)$
Initial allocation
Perturbated allocation
$t\Delta Y(w, \theta_0)$
w
$\omega - \delta\omega$
ω

Back



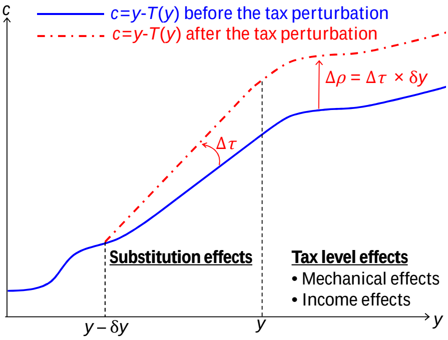
$Y(w,\theta)$
Initial allocation
Perturbated allocation
$\Delta Y(w,\theta,t)$
w
$W(\omega-\delta\omega,\theta)$
$W(\omega,\theta)$

Back



Back

$$\underbrace{\hat{\varepsilon}(y) \cdot y\ \hat{h}(y) \cdot \frac{T'(y)}{1 - T'(y)}}_{\text{Substitution effects}} = \int_y^\infty \left[\underbrace{1 - \hat{g}(z)}_{\text{Mechanical}} - \underbrace{\hat{\eta}(z) \cdot T'(Y(z))}_{\text{Income effects}} \right] \cdot \hat{h}(z) dz$$



Back

Examples of applications

1 Heterogeneous labor supply elasticities θ when $v(y; w, \theta) = \left(\frac{y}{w}\right)^{1+\frac{1}{\theta}}$

2 Joint Taxation of additional income z (capital part of entrepreneurial income, income received from renting property, inherited wealth...):

$$\max_{y,z} \quad u(y - T(y)) - V(y - z, z; w, \theta)$$

where taxable income is y and θ is the ability to earn z. Then:

$$v(y; w, \theta) \stackrel{\text{def}}{\equiv} \min_{z} \quad V(y - z, z; w, \theta)$$

3 Optimal joint taxation of households where $w = w_1$, $\theta = w_2$ and $v(y; w, \theta) = \min_{\ell_1, \ell_2} \quad V(\ell_1, \ell_2) \quad s.t : w_1\ell_1 + w_2\ell_2 = y$

Examples of applications (2)

4 Tax avoidance/evasion where z is evaded income, y is taxable income, so $y+z$ is labor income, w is productivity and θ is the ability to evade income and .

$$\max_{y,z} \quad y - T(y) + z - V\left(y+z, z; w, \theta\right)$$

Then

$$v(y; w, \theta) \stackrel{\text{def}}{\equiv} \min_{z} \quad V\left(y+z, z; w, \theta\right) - z$$

5 Nonlinear pricing in IO (heterogeneity in the factor scale of demand and in price elasticity of demand)