Optimal Tax in Occupational Choice Models: an Application to the Work Decisions of Couples

Guy Laroque & Nicola Pavoni

Institute for Fiscal Studies

Sciences Po & *Bocconi-IGIER*

Bank of Spain Conference on Taxes and Transfers

Introduction

- We lay down a general occupational choice model with multiple dimensions of heterogeneity
- The model subsumes most known models of optimal income taxation (with intensive and extensive margins, ...)
- We apply our framework to study taxation of working couples
- We think that there are more cases that lead to subsidize, not tax, working couples.

Introduction

- Handling a ‘general’ model might be important
- At this stage, we need to address policy questions involving multi-dimensional decisions and multi-dimensional heterogeneity without restricting the framework for tractability
- As in the literature, we give qualitative graphical results, and intuitions in a simplified model
- One outcome of the paper however, is a tool to derive the optimal taxation in a general setting by matrix inversion

Related Literature

1. Occupational Choice: Rothschild and Scheuer (2013 & 2015), Ales et al. (2015), Lockwood et al. (2017), ...
   (general equilibrium effects, externalities, no occupation specific tax)

2. Taxation of Couples: Kleven et al. (2009), Immervoll et al. (2011), Frankel (2014), ...
   (more general model in some aspects, new results)

3. IO literature: McAfee et al. (1989), Salinger (1995), Armstrong (2011), ... (revenue maximization, nonegative taxes)

Model

Fundamentals of the Model

- Finite number of occupations $i = 0, 1, 2, \ldots, I$.
- Continuum of agents of type $\alpha \in \mathcal{A} \subset \mathbb{R}^N$.
- Distribution of types F.
- Production or gross income (observed): ω^i, $i = 0, \ldots I$.
- Net income levels: $c = (c^i, i = 0, \ldots I)$
- c^i cannot depend on α (not observable and pooling)
- Utility of agent α in occupation i when facing c

$$u^i(c^i, \alpha).$$

Assumptions
(i) $u^i(\cdot, \alpha)$ increasing and continuously differentiable $\forall(i, \alpha)$.
(ii) Both $u^i(c^i, \cdot)$ and $u^i_1(c^i, \cdot)$ are continuous $\forall(i, c^i)$.

Occupational Choice Functions

- Set of agents choosing i under c

$$A^i(c) := \{\alpha | u^i(c^i, \alpha) > u^j(c^j, \alpha) \text{ for all } j \neq i\}.$$

- Measure of agents in i

$$\mu(A^i(c)) := \int_{A^i(c)} dF(\alpha).$$

- Since total population is constant (at 1) we have:

$$\frac{\partial \mu(A^0(c))}{\partial c^i} = -\sum_{j=1}^{I} \frac{\partial \mu(A^j(c))}{\partial c^i}.$$

Possible Interpretations: It is all about Borders!

a) For $I = 1$ we are in Choné and Laroque (2005)

b) The case $I = 2$ can be interpreted as Immervoll-Kleven-Kreiner-Verdelin (2011)

c) If cost of work and ω^i increasing in i (e.g., $u^i(c^i, \alpha) = u(c^i) - \alpha^i(\alpha)$) we get the intensive model. The standard Mirrlees model assumes a complete order on $\mathcal{A}$ and increasing differences

$$\alpha^j(\alpha') - \alpha^k(\alpha') \leq \alpha^j(\alpha) - \alpha^k(\alpha) \text{ whenever } \alpha' \succsim \alpha \text{ and } j > k$$

$\Rightarrow$ Occupation i admits only borders with $i - 1$ and $i + 1$.

d) Saez (2002): Occupation i only borders: $i - 1$, $i + 1$ and 0.

Optimal Program and FOCs

Typical Optimal Tax Program

$$\max_{c} \quad \sum_{i=0}^{I} \int_{A^i(c)} \beta(\alpha)\psi(u^i(c^i,\alpha))dF(\alpha) \tag{1}$$

$$\sum_{j=0}^{I} \left[\omega^j - c^j\right] \mu(A^j(c)) \geq G. \qquad (\lambda) \tag{2}$$

- Let the average Pareto weight for occupation i be:

$$P(A^i) := \frac{1}{\mu(A^i)} \int_{A^i} \frac{\beta(\alpha)\psi'(u^i(c^i,\alpha))u_1^i(c^i,\alpha)}{\lambda} dF(\alpha).$$

- Using the constant population condition, FOC with respect to c^i reads

$$\mu(A^i)[P(A^i) - 1] = -\sum_{j=1}^{I} \frac{\partial \mu(A^j)}{\partial c^i}[t^j - t^0],$$

where $t^i := \omega^i - c^i$.

Examples

- Choné and Laroque (2005), i.e., $i = 0, 1$

$$t^1 - t^0 = \frac{\mu(A^0)}{\frac{\partial \mu(A^0)}{\partial c^0}}[P(A^0) - 1].$$

- Mirrlees (1973):

$$\mu(A^i)[P(A^i) - 1] = \frac{\partial \mu(A^{i-1})}{\partial c^i}[t^i - t^{i-1}] + \frac{\partial \mu(A^{i+1})}{\partial c^i}[t^i - t^{i+1}].$$

If no income effects:

$$\frac{\partial \mu(A^i)}{\partial c^j} = \frac{\partial \mu(A^j)}{\partial c^i} \qquad \Rightarrow$$

$$t^i - t^{i-1} = \frac{\mu(A^i)}{\frac{\partial \mu(A^i)}{\partial c^i}} \sum_{s=i}^{I} \frac{\mu(A^s)}{\mu(A^i)}[1 - P(A^s)]$$

Structure of FOCs

- With μ^i & P^i given, the FOCs are linear in tax differences
- Ignoring row 0, we have a $I \times I$ system

$$\mu(P-1) = H\Delta t,$$

$\mu(P-1) := (\mu(A^i)(P(A^i)-1),\ i = 1, 2, \ldots I)'$;
$h^{ij} = -\frac{\partial \mu(A^j)}{\partial c^i}$ generic element of H;
Δt vector with generic element $t^i - t^0$.

- If H is invertible:

$$\Delta t = H^{-1}\mu(P-1),$$

- The budget constraint and the FOC_0 determine t^0 and λ.
- Similarly we get differences w.r.t. a generic t^j, $j \neq 0$.

Observations

- Anything goes: Take any two vectors Δt and $\mu(P-1)$, there is at least one H that solves the FOCs (check SOC)
- Computation: Given a matrix H, and a vector $\mu(P-1)$, matrix inversion delivers optimal tax differences
- Question 1: Take a matrix H and a vector μ (data from elasticities), what is the set of Δt compatible with some vector of Pareto weights P?
- Question 2: Fix H, μ, and a property on Δt (e.g., a sign). What is the set of Pareto weights compatible with it?

Optimal Tax of Working Couples

A simple model of couple taxation

- Let $I = 3$ and assume: $\omega^0 = 0$ and $\omega^3 = \omega^1 + \omega^2$;
  ($i = 0$ no one works, $i = 3$ both work)

- Definitions:

$$c^3 + c^0 \geq c^1 + c^2 \qquad \text{(positive reinforcement)},$$
$$c^3 + c^0 \leq c^1 + c^2 \qquad \text{(negative reinforcement)}.$$

- Equivalently, from $t^i = \omega^i - c^i$, positive reinforcement:

$$t^3 - t^2 \leq t^1 - t^0 \quad \text{or} \quad t^3 - t^1 \leq t^2 - t^0.$$

The program gives larger incentives to work to a member of the couple when her/his partner works than when s/he does not work.

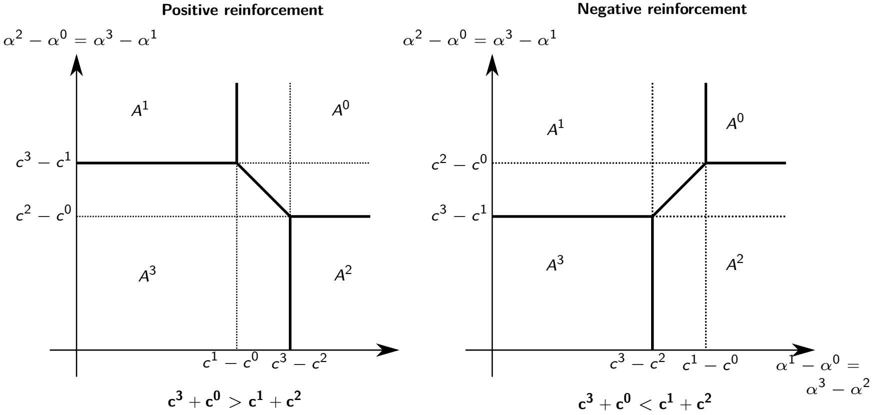


Figure: Look at borders

Subsidising working couples maximizes revenues (if no assortative mating)

- Assume the optimal taxes of the single worker households are non negative, i.e. $t^j - t^0 \geq 0$ for $j = 1, 2$.
- Assume that under negative reinforcement

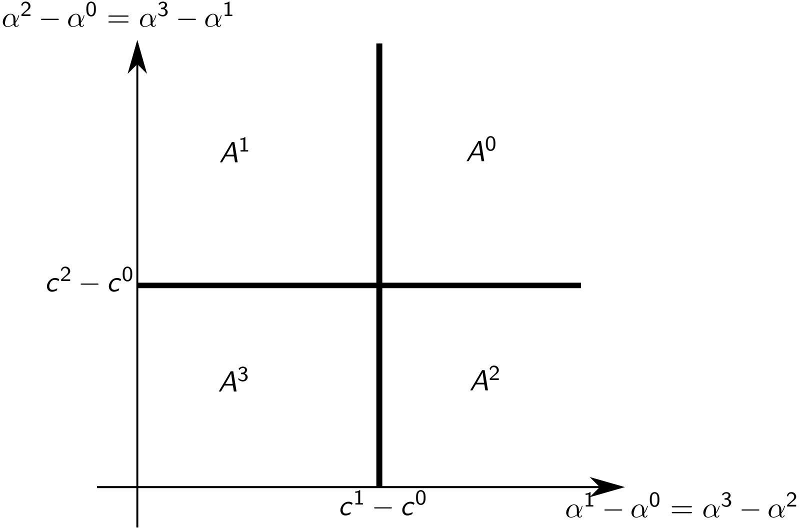
$$\frac{\partial \mu(A^0)(c)}{\partial c^3} = 0.$$

Result 1 If at the optimum $P(A^3) \geq 1$,
the optimum exhibits positive reinforcement.

- Assume no income effects: $u(c^i, \alpha) = c^i - \alpha^i(\alpha)$
- The planner aims at maximizing revenues (Rawls) and $\mu^i > 0$ ($P^i = 0$ for $i \neq 0$)
- Assume, in addition, $\alpha^3(\alpha) + \alpha^0(\alpha) = \alpha^1(\alpha) + \alpha^2(\alpha)\ \forall\alpha$

Result 2 If $F_i(\alpha^i - \alpha^0)$ $i = 1, 2$ is independent and log-concave, then positive reinforcement is optimal.

Start with the optimal independent (no reinforcement)



Non-discriminatory condition

- Now, allow income effects and any distribution on $\mathcal{A}$
- Restrict $t^1 = t^2 := t$ (w.l.o.g. under symmetry)

Result 3 Then the sign of the reinforcement term: $2t - t^3 - t^0$ is the same as that of the following expression (at the optimum)

$$R := \frac{\mu(A^3)}{\frac{\partial \mu(A^3)}{\partial c^3} - \frac{\partial \mu(A^0)}{\partial c^3}} [P(A^3) - 1] + \frac{\mu(A^0)}{\frac{\partial \mu(A^0)}{\partial c^0} - \frac{\partial \mu(A^3)}{\partial c^0}} [P(A^0) - 1].$$

Result 4 If negative enforcement implies $\frac{\partial \mu(A^0)}{\partial c^3} = \frac{\partial \mu(A^3)}{\partial c^0} = 0$ then a sufficient condition for positive reinforcement is

$$\hat{R} := \frac{\mu(A^3)}{\frac{\partial \mu(A^3)}{\partial c^3}} [P(A^3) - 1] + \frac{\mu(A^0)}{\frac{\partial \mu(A^0)}{\partial c^0}} [P(A^0) - 1] > 0.$$

Proof of Result 3.

First order conditions:

$$\begin{aligned}
\mu(A^0)[P(A^0)-1] &= \frac{\partial\mu(A^1)}{\partial c^0}(t^0-t^1) + \frac{\partial\mu(A^2)}{\partial c^0}(t^0-t^2) + \frac{\partial\mu(A^3)}{\partial c^0}(t^0-t^3)\\
\mu(A^1)[P(A^1)-1] &= \frac{\partial\mu(A^0)}{\partial c^1}(t^1-t^0) + \frac{\partial\mu(A^2)}{\partial c^1}(t^1-t^2) + \frac{\partial\mu(A^3)}{\partial c^1}(t^1-t^3)\\
\mu(A^2)[P(A^2)-1] &= \frac{\partial\mu(A^0)}{\partial c^2}(t^2-t^0) + \frac{\partial\mu(A^1)}{\partial c^2}(t^2-t^1) + \frac{\partial\mu(A^3)}{\partial c^2}(t^2-t^3)\\
\mu(A^3)[P(A^3)-1] &= \frac{\partial\mu(A^0)}{\partial c^3}(t^3-t^0) + \frac{\partial\mu(A^1)}{\partial c^3}(t^3-t^1) + \frac{\partial\mu(A^2)}{\partial c^3}(t^3-t^2).
\end{aligned}$$

We rearrange the first and last rows in order to get rid of t^2 (in the first row, we add to both sides $\partial\mu(A^2)/\partial c^0(t^2-t^1)$, in the last row we add $\partial\mu(A^2)/\partial c^3(t^2-t^1)$) we get:

$$\begin{aligned}
\mu(A^0)[P(A^0)-1]+\frac{\partial\mu(A^2)}{\partial c^0}(t^2-t^1) &= \left[\frac{\partial\mu(A^1)}{\partial c^0}+\frac{\partial\mu(A^2)}{\partial c^0}\right](t^0-t^1)+\frac{\partial\mu(A^3)}{\partial c^0}(t^0-t^3)\\
\mu(A^3)[P(A^3)-1]+\frac{\partial\mu(A^2)}{\partial c^3}(t^2-t^1) &= \frac{\partial\mu(A^0)}{\partial c^3}(t^3-t^0)+\left[\frac{\partial\mu(A^1)}{\partial c^3}+\frac{\partial\mu(A^2)}{\partial c^3}\right](t^3-t^1).
\end{aligned}$$

Proof of Result 3, continued

Let $X^0 = \mu(A^0)(P(A^0)-1) + \frac{\partial \mu(A^2)}{\partial c^0}(t^2 - t^1)$ and $X^3 = \mu(A^3)(P(A^3)-1) + \frac{\partial \mu(A^2)}{\partial c^3}(t^2 - t^1)$. The system can be considered as a system of two equations in $t^0 - t^1$ and $t^3 - t^1$.

$$\begin{aligned} X^0 &= \left[\frac{\partial \mu(A^1)}{\partial c^0} + \frac{\partial \mu(A^2)}{\partial c^0} + \frac{\partial \mu(A^3)}{\partial c^0}\right](t^0 - t^1) - \frac{\partial \mu(A^3)}{\partial c^0}(t^3 - t^1) \\ X^3 &= -\frac{\partial \mu(A^0)}{\partial c^3}(t^0 - t^1) + \left[\frac{\partial \mu(A^0)}{\partial c^3} + \frac{\partial \mu(A^1)}{\partial c^3} + \frac{\partial \mu(A^2)}{\partial c^3}\right](t^3 - t^1), \end{aligned}$$

or

$$\begin{aligned} \frac{\partial \mu(A^0)}{\partial c^0}(t^0 - t^1) + \frac{\partial \mu(A^3)}{\partial c^0}(t^3 - t^1) &= -X^0 \\ \frac{\partial \mu(A^0)}{\partial c^3}(t^0 - t^1) + \frac{\partial \mu(A^3)}{\partial c^3}(t^3 - t^1) &= -X^3. \end{aligned}$$

End of proof of Result 3

The solution of this two by two system is

$$
\begin{aligned}
t^0 - t^1 &= \frac{1}{\Delta}\left[-\frac{\partial\mu(A^3)}{\partial c^3}X^0 + \frac{\partial\mu(A^3)}{\partial c^0}X^3\right] \\
t^3 - t^1 &= \frac{1}{\Delta}\left[\frac{\partial\mu(A^0)}{\partial c^3}X^0 - \frac{\partial\mu(A^0)}{\partial c^0}X^3\right],
\end{aligned}
$$

Δ is the determinant $\partial\mu(A^0)/\partial c^0 \partial\mu(A^3)/\partial c^3 - \partial\mu(A^3)/\partial c^0 \partial\mu(A^0)/\partial c^3 > 0$. Using the non-discriminatory condition $t^2 = t^1 := t$, the sum of the two equations yields

$$
t^3 - 2t + t^0 = \frac{\left[\left(\frac{\partial\mu(A^3)}{\partial c^0} - \frac{\partial\mu(A^0)}{\partial c^0}\right)\mu(A^3)[P(A^3)-1]\right] + \left[\left(\frac{\partial\mu(A^0)}{\partial c^3} - \frac{\partial\mu(A^3)}{\partial c^3}\right)\mu(A^0)[P(A^0)-1]\right]}{\Delta},
$$

The result follows by dividing both sides by the positive number

$$
\left(\partial\mu(A^3)/\partial c^0 - \partial\mu(A^0)/\partial c^0\right) \times \left(\partial\mu(A^0)/\partial c^3 - \partial\mu(A^3)\partial c^3\right)
$$

Tax Reforms

- Look at Result 3
- If we are not at the optimum the sign of R gives a test of optimality of the existing program

Graphical analysis under no income effects

- We now want to explore the meaning of the sufficient condition graphically
- Let

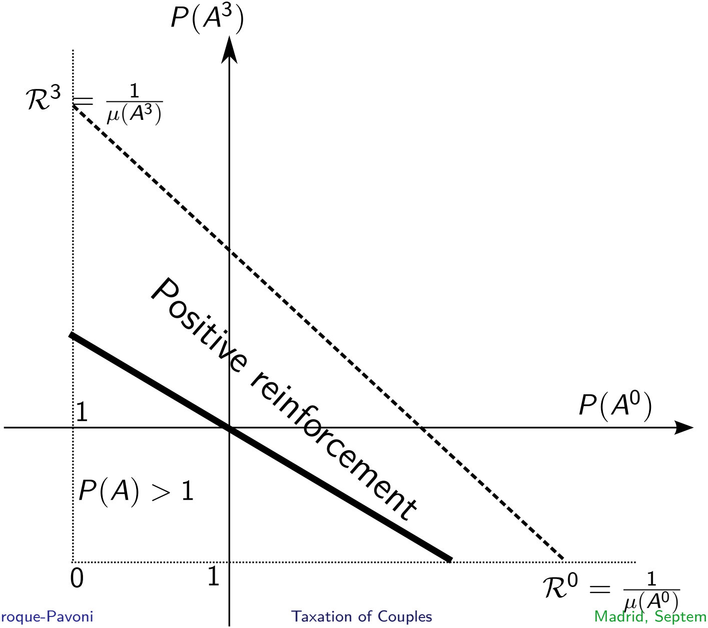
$$r := \frac{\dfrac{\mu(A^0)}{\dfrac{\partial\mu(A^0)}{\partial c^0}}}{\dfrac{\mu(A^3)}{\dfrac{\partial\mu(A^3)}{\partial c^3}}} = \frac{\dfrac{\partial \log \mu(A^3)}{\partial c^3}}{\dfrac{\partial \log \mu(A^0)}{\partial c^0}} \geq 0.$$

Result 4 (restated) Under the condition of the previous slide, if

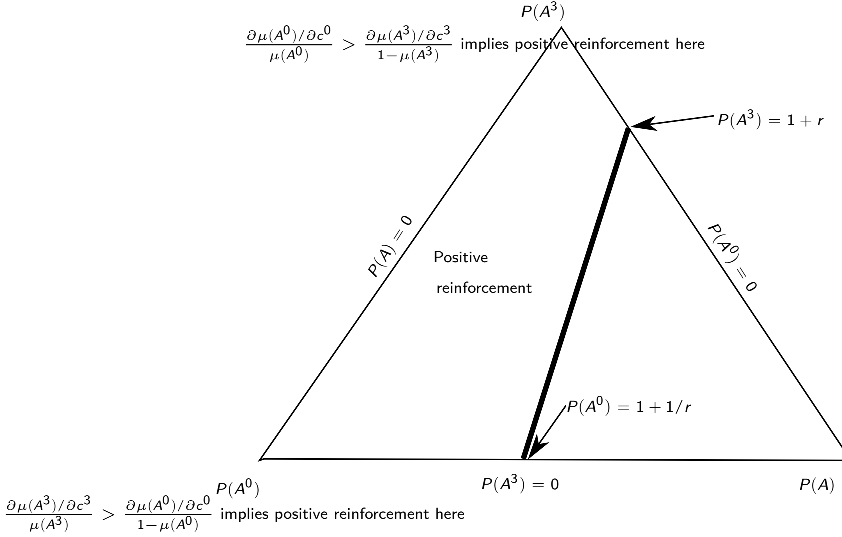
$$P(A^3) - 1 \geq -r[P(A^0) - 1],$$

the optimal program exhibits positive reinforcement.

- Eliminating income effects reduces the dimensionality ⇒


$P(A^3)$
$\mathcal{R}^3 = \frac{1}{\mu(A^3)}$
Positive reinforcement
$P(A^0)$
1
$P(A) > 1$
0
1
$\mathcal{R}^0 = \frac{1}{\mu(A^0)}$

The Marschak-Machina Triangle Representation



Generalizations

Many Productivities with non-discrimination

- Compared to the 'standard intensive' model this model has an extra layer of heterogeneity

- Suppose productivity of both partners observed if one working
- For example, let $\omega^1(\omega) = \omega^2(\omega) = \omega$ and $\omega^3(\omega) = 2\omega$

Result 5 Fix ω and assume $t^1(\omega) = t^2(\omega)$. The sign of the reinforcement term for agents ω is the same as:

$$\sum_{i=1,2} \left[\frac{\partial \mu(A^3(\omega))}{\partial c^i(\omega)} - \frac{\partial \mu(A^0)}{\partial c^i(\omega)} \right] \times RR(\omega),$$

where

$$RR(\omega) := \frac{\mu(A^3(\omega))[P(A^3(\omega)) - 1]}{\frac{\partial \mu(A^3(\omega))}{\partial c^3(\omega)} - \frac{\partial \mu(A^0)}{\partial c^3(\omega)}} + \frac{\mu(A(\omega))[P(A(\omega)) - 1]}{\sum_{i=1,2} \left[\frac{\partial \mu(A^3(\omega))}{\partial c^i(\omega)} - \frac{\partial \mu(A^0)}{\partial c^i(\omega)} \right]}.$$

General Solution for $I = 3$

$$H = \begin{bmatrix} -\frac{\partial \mu(A^1)}{\partial c^1} & -\frac{\partial \mu(A^2)}{\partial c^1} & -\frac{\partial \mu(A^3)}{\partial c^1} \\ -\frac{\partial \mu(A^1)}{\partial c^2} & -\frac{\partial \mu(A^2)}{\partial c^2} & -\frac{\partial \mu(A^3)}{\partial c^2} \\ -\frac{\partial \mu(A^1)}{\partial c^3} & -\frac{\partial \mu(A^2)}{\partial c^3} & -\frac{\partial \mu(A^3)}{\partial c^3} \end{bmatrix}.$$

- If we *assume away income effects* we gain symmetry

$$H = \begin{bmatrix} a_1 & -b & -e \\ -b & a_2 & -d \\ -e & -d & a_3 \end{bmatrix},$$

$$H^{-1} = \frac{1}{\Delta} \begin{bmatrix} a_2 a_3 - d^2 & ed + a_3 b & a_2 e + bd \\ ed + a_3 b & a_1 a_3 - e^2 & a_1 d + eb \\ a_2 e + bd & a_1 d + eb & a_2 a_1 - b^2 \end{bmatrix}.$$

- We can recover the ‘right’ perturbation from formula
- Similar expression for Result 5 obtainable (arbitrary mating)

Couple Tax with Intensive Margins (Kleven et al. 2013-2014)

- It is all about borders again.
- Let i labor supply for agent 1, ℓ level of labor supply for agent 2, and $-\ell$ the complement.

$$t_\ell^i - t_\ell^{i-1} = \sum_{s=i}^{I} \frac{\mu(A_\ell^s)}{\frac{\partial \mu(A_\ell^i)}{\partial c_\ell^i}} \left[1 - P(A_\ell^s) + \sum_{(k,-\ell)\in K(s,\ell)} \frac{\frac{\partial \mu(A^k_{-\ell})}{\partial c^s}}{\mu(A_\ell^s)} (t_\ell^s - t^k_{-\ell}) \right]$$

where $K(s,\ell)$ is the set of indexes with a border with (s,ℓ)

- Of course, we can get tax alone
- 'Just' inverse of a $2I$ or $I \times I$ matrix (even symbolic, Matlab)
- Basic math and direct interpretation even when intensive margin of both partners is involved
- Positive reinforcement: Δt_ℓ^i decreases with ℓ (for all i?)

A few numbers on the UK

With the help of Barra Roantree from the Institute of Fiscal Studies, we have evaluated the reinforcement term in the UK on a representative sample of couples in the population.

The computations

- The exercise is based on the 2013 Family Expenditure Survey, from which we work on a group of 3077 couples, each couple belonging to a single tax unit.
- Starting from gross income ω, it uses the program TAXBEN to evaluate the disposable family weekly income c, measured in pounds, in the four cases:
  1. Nobody works in the household: ω^0 and c^0,
  2. The man works, the woman does not: ω^1 and c^1,
  3. The man does not work, the woman does: ω^2 and c^2,
  4. Both the man and woman work: ω^3 and c^3.

The computation is easy for the households where both partners work, since the survey then provides their hourly wage.

In the other cases, it requires predicting a counterfactual behaviour for the non workers, and therefore relies on assumptions that might be adapted in future work.

We used the standard assumptions in Mincer wage equations in the labor literature. By construction, the data satisfy our Assumption 3: ω^0 is zero, and productivities of the two members of any couple are additive: $\omega^1 + \omega^2 = \omega^3$.

The sign of reinforcement by household categories

United Kingdom 2013, Source : Family Expenditure Survey

Reinf.	Age	Children	Marg. tax rate	Earnings	Total
Negative	47 (14)	1.03 (1.36)	.14 (.11)	308 (343)	9%
Zero	60 (14)	.34 (.82)	.15 (.12)	199 (427)	36%
Positive	47 (14)	.85 (1.05)	.15 (.13)	372 (471)	55%