Optimal Nonlinear Savings Taxation

Charles Brendon
Cambridge & TSE

EEA-ESEM
Milan, Aug 2022

Introduction

- This is an optimal tax paper in the dynamic Mirrlees tradition:
  - *Unobservable types, evolving stochastically*
  - *No ex-ante restriction on policy instruments*
- Dynamic Mirrlees problems are commonly viewed as complex, removed from practical tax design
- Main message here: they can instead **simplify** analyses of tax instruments
- This talk: motivation, overview of main results

Motivation

Sufficient statistics: static to dynamic?

- Huge 'sufficient statistics' tax literature has grown from Diamond (1998), Saez (2001)
  - *Optimal policy expressed by ref. to small no. of measurable statistics*
  - *Simple, intuitive optimality formulae for instrument choice*
- Analytically, a reworking of **static** Mirrlees (1971) mechanism design problem
- Duality between optimising over tax schedules, and optimising over allocations

Motivation

Sufficient statistics: static to dynamic?

- No equivalent 'dual' approach exists for **dynamic** Mirrlees problems
  - *Conceptually difficult: a change to taxes in t might affect behaviour in $t + 58$, so where to start?*
  - *What does the decentralisation even look like?*
- Instead, sufficient statistics papers have taken alternative directions – esp steady state analysis
  - Stantcheva (2020), Piketty & Saez (2013)
- Mech design and sufficient statistics now commonly presented as rival approaches Quote

Overview

What I do

- I analyse a variant of the Atkeson & Lucas (1992) dynamic hidden info problem:
  - *Endowment economy, infinite horizon*
  - *Idiosyncratic,* ***persistent*** *shocks to w/in-period MU of consumption:*

$$U_0 := \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \alpha_t u(c_t)$$

  - α_t *private, noncontractable* $\Rightarrow$ *imperfect insurance*
  - *Simple savings technology (available to policymaker)*
- Basic policy trade-off: **insurance** (equalise $\alpha_t u'(c_t)$) vs **incentives** to misreport

Overview

The mechanism design problem

Objective:

$$\max \int_{\alpha_0} U_0(\alpha_0)\, d\Pi(\alpha_0)$$

Constraints:

1. **Resources**

$$\mathbb{E} \sum_{t=0}^{\infty} R^{-t} c_t(\alpha^t) \leq \bar{Y}$$

2. **First-order IC**

$$\alpha'_t u_t(\alpha'_t) + \beta \omega_{t+1}(\alpha'_t) = \underline{\alpha} u_t(\underline{\alpha}) + \beta \omega_{t+1}(\underline{\alpha}) + \int_{\underline{\alpha}}^{\alpha'_t} \frac{1}{\alpha_t} \left[\alpha_t u_t(\alpha_t) + \beta \omega^{\Delta}_{t+1}(\alpha_t) \right] d\alpha_t$$

3. **Promise keeping**

omegas

Overview

What I show

1. IC allocations have a simple consumption-savings **decentralisation**, s.t. nonlinear savings taxes
2. These tax instruments satisfy an intuitive 'sufficient statistics' optimality condition
   - *Similar to Saez (2001), isomorphic if types iid*
   - *Very **limited dependence** on cross-period effects*
3. This allows for novel qualitative insights:
   - *It is optimal to set **positive marginal savings taxes**, funding a universal lump-sum transfer*

Overview

Main characterisation

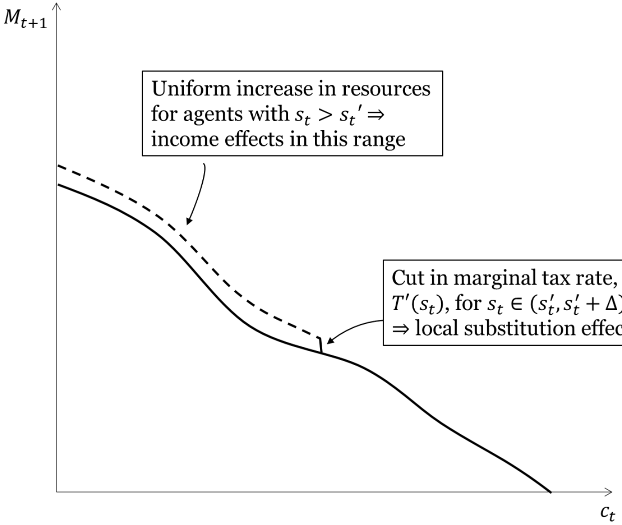
Effects of cutting marginal savings taxes at s_t'?

$$\overbrace{\mathbb{E}_{t-1}\left[1 - T_t'(s_t)\frac{ds_t}{dM_t} - g_t(s_t)\middle| s_t > s_t'\right]}^{\text{Net cost of transfers above } s_t'}$$

$$= \underbrace{\frac{\pi^s(s_t')\,s_t'}{1-\Pi^s(s_t')} T_t'(s_t')\,\varepsilon_t^s + RT_{t-1}'(s_{t-1})\,s_{t-1}\,\epsilon_{t-1,t}^s(s_t')}_{\text{Benefit from behavioural response}}$$

Overview

Perturbing the tax schedule



Taxing savings

Why (and how much)?

- Positive marginal savings taxes redistribute away from states where consumption need is low
- Starting from an allocation with no distortions, higher $T_t'(s_t)$ is desirable at any s_t
- This funds a uniform transfer, raising consumption where α_t is highest
- Sufficient statistics quantification → T_t' around 1 to 2 per cent for highest savers

Recursive multipliers

Mechanism design => welfare weights

- Important feature of paper: mech design and sufficient statistics inform each other
- Good example is 'welfare weight', $g_t(s_t)$
  - MV to policymaker (in resource units) of extra welfare tos_t, as s_t's wealth increases
- I show:

$$g_t(s_t) := \alpha_t u'(c_t) \frac{(1+\lambda_t)}{\eta_t}$$

  -
  - $\lambda_t(\alpha^{t-1})$ is the Marcet-Marimon multiplier (on promise keeping)
  - Policy is cross-sectionally utilitarian, given accumulated wealth

Closing points

- Dynamic mechanism design can yield direct practical insights for tax policy
- Methodology is easily generalisable
- It reveals a complementarity: you need a decentralisation to think practically about taxes...
- ... but structure of information-theoretic problem can make the exercise manageable

Setup

Preferences & shocks

- Time discrete, $t = 0, 1, \dots$
- Committed policymaker, **utilitarian** in period 0
- Continuum of agents, utility (viewed in t):

$$U_t = \mathbb{E}_t \sum_{s=t}^{\infty} \beta^{s-t} \alpha_s u(c_s)$$

- $\alpha_t \in A \subset \mathbb{R}_+$ exog. stochastic, Markov, independent across agents
    - Cdf: $\Pi(\alpha_t | \alpha_{t-1})$, differentiable on A^2
    - Pdf: $\pi(\alpha_t | \alpha_{t-1})$, satisfying MLRP

Setup

Mechanism design problem

Initially, study mechanism design problem:

- Revelation principle $\Rightarrow$ focus on direct allocations
- **First-order approach** to simplify IC

FOA

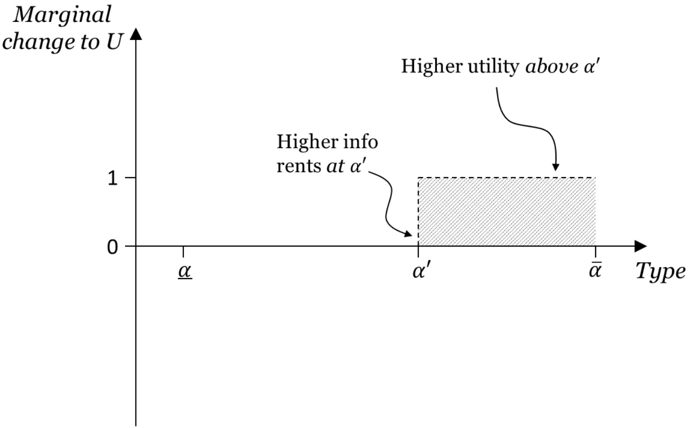
Analysis

Utility perturbations

- It is simplest to characterise this problem by reference to changes in profile of $U_t\left(\alpha_t\right)$, given α^{t-1}
- ‘Suppose I raise info rents at α_t, holding constant elsewhere...’
- Trade off costs of changing **info rents** vs (net) benefits of **utility provision**
- Graphically...

Analysis

Utility perturbations



Analysis

Utility-based characterisation

- Algebraically:

$$\int_{\alpha_t'}^{\bar{\alpha}} \overbrace{\left[\alpha_t\left[1+\lambda_t+\lambda_t^{\Delta}\rho\left(\alpha_t|\alpha_{t-1}\right)\right]-\frac{\eta_t}{u'\left(c_t\left(\alpha_t\right)\right)}\right]}^{\text{Net benefit of utility provision above } \alpha_t'} d\Pi\left(\alpha_t|\alpha_{t-1}\right)$$

$$=\pi\left(\alpha_t'|\alpha_{t-1}\right)\cdot\underbrace{\left(\alpha_t'\right)^2\cdot\left[\lambda_{t+1}^{\Delta}\left(\alpha_t'\right)-\rho\left(\alpha_t'|\alpha_{t-1}\right)\lambda_t^{\Delta}\right]}_{\text{Cost of raising info rents at } \alpha_t'}$$

  - λ_t & λ_t^{Δ}: multipliers on promise constraints
  - η_t: resource multiplier in t
- Useful related results on inverse MU **dynamics**, immiseration

lambdas dynamics

Analysis

The limits of talking about utility

- *These characterisations can be very insightful, but...*
- Dependent on unobservable primitives
  - *Inverse marginal utilities, cost of changing info rents, distn of types...*
  - *[Piketty-Saez quote earlier...]*
- How can I link more closely to market decentralisation?

Analysis

Ways forward?

Two possible approaches:

1. Specify decentralisation, analyse **instrument choice** directly, given responses
   - *Works in static case*
   - *Hard to keep tractable with dynamics*
2. **Reverse-engineer** from what we already have

Reverse engineering

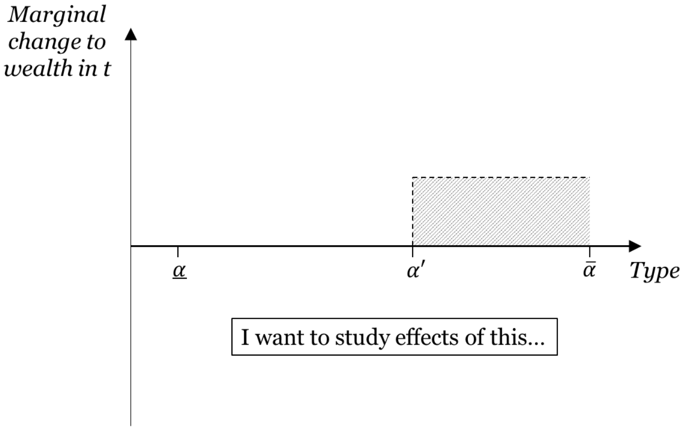
From mechanism design to sufficient statistics

Main methodological innovation:

- I have characterised costs/benefits of x-sectional **utility** perturbations
- I want expressions relating to (still unspecified) **taxes**
- Tax changes $\Rightarrow$ cross-sectional **wealth** perturbations
- But wealth changes *imply* utility changes...
  - *Extra unit of resources in* t *to* $\alpha_t \Rightarrow$ *utility rises by* $\alpha_t u'(c_t)$
- ... which I already have the tools to look at!

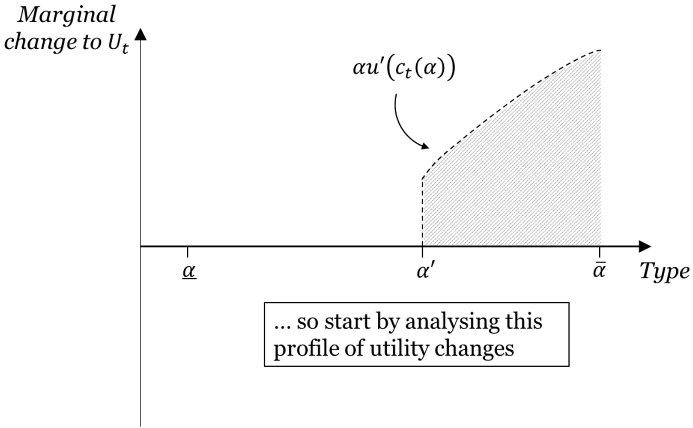
Reverse engineering

From mechanism design to sufficient statistics



Reverse engineering

From mechanism design to sufficient statistics



Reverse engineering

From mechanism design to sufficient statistics

Two-step procedure:

1. ***Characterise*** *effects 'simple' tax changes could engineer*
2. *Link to behavioural statistics from a* ***decentralisation***

Combining previous conditions:

$$
\begin{aligned}
-\int_{\underline{\alpha}}^{\alpha_t'} \left[1 + \alpha_t^2 u''(c_t) \frac{dc_t}{d\alpha_t} \frac{\lambda_{t+1}^{\Delta}(\alpha_t)}{\eta_t} - \frac{\alpha_t u'(c_t)(1+\lambda_t)}{\eta_t} \right] d\Pi(\alpha_t|\alpha_{t-1}) \\
+ \left(\alpha_t'\right)^2 u'(c_t) \frac{\lambda_{t+1}^{\Delta}(\alpha_t')}{\eta_t} \pi\left(\alpha_t'|\alpha_t\right) \\
+ \int_{\underline{\alpha}}^{\alpha_t'} \left\{ \alpha_t^2 u''(c_t) \frac{dc_t}{d\alpha_t} + \alpha_t u'(c_t) \right\} \rho(\alpha_t|\alpha_{t-1}) \beta R \frac{\lambda_t^{\Delta}}{\eta_{t-1}} d\Pi(\alpha_t|\alpha_{t-1}) \\
- \left(\alpha_t'\right)^2 u'(c_t) \rho\left(\alpha_t'|\alpha_{t-1}\right) \beta R \frac{\lambda_t^{\Delta}}{\eta_{t-1}} \pi\left(\alpha_t'|\alpha_t\right) \\
= 0
\end{aligned}
$$

By itself, just a re-working...

2. Tax decentralisation

Consumption-savings choice

Consider a consumption-savings problem with nonlinear taxes...

- Each period individuals have wealth $M_t\left(s^{t-1}\right)$
- Allocate optimally between consumption and savings:

$$M_t = c_t + s_t$$

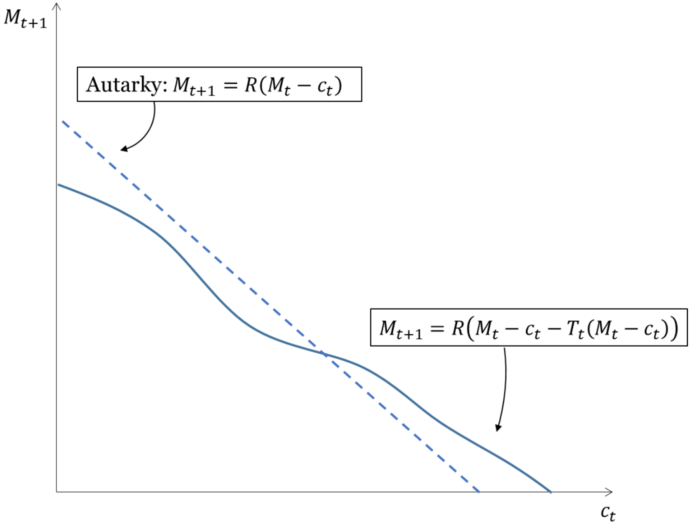
- Savings are taxed, residual earns interest at rate R:

$$M_{t+1} = R\left(s_t - T_t\left(s_t\right)\right)$$

- *[Normalise* $\mathbb{E}_{t-1}\left[T_t\left(s_t\right)\right] = 0$*]*

2. Tax decentralisation

Consumption-savings choice: graphically



2. Tax decentralisation

Consumption-savings choice: observations

- Any allocation w $c_t\left(\alpha^{t-1}, \alpha_t\right)$ **strictly increasing** in α_t can be decentralised this way
  - *[Essentially: set* $M_t = \mathbb{E}_{t-1} \sum_{s=t}^{\infty} R^{t-s} c_s$ *at each node]*
- Complex dependence of future budget constraints on current choice ⇒ no simple Euler eqn
- With enough differentiability, do have:

$$\alpha_t u'\left(c_t\right) = \beta R\left(1 - T_t'\left(s_t\right)\right) \int_{\alpha_{t+1}} V_M\left(M_{t+1}; \alpha_{t+1}\right) d\Pi\left(\alpha_{t+1} | \alpha_t\right)$$

- Use this condition to analyse consumer choice, derive **elasticities**

Optimal taxes

Back to the equation

As previewed, this reduces an unintuitive expression...

$$
\begin{aligned}
-\int_{\underline{\alpha}}^{\alpha_t'} \left[1 + \alpha_t^2 u''(c_t) \frac{dc_t}{d\alpha_t} \frac{\lambda_{t+1}^{\Delta}(\alpha_t)}{\eta_t} - \frac{\alpha_t u'(c_t)(1+\lambda_t)}{\eta_t}\right] d\Pi(\alpha_t|\alpha_{t-1}) \\
+ (\alpha_t')^2 u'(c_t) \frac{\lambda_{t+1}^{\Delta}(\alpha_t')}{\eta_t} \pi(\alpha_t'|\alpha_t) \\
+ \int_{\underline{\alpha}}^{\alpha_t'} \left\{\alpha_t^2 u''(c_t) \frac{dc_t}{d\alpha_t} + \alpha_t u'(c_t)\right\} \rho(\alpha_t|\alpha_{t-1}) \beta R \frac{\lambda_t^{\Delta}}{\eta_{t-1}} d\Pi(\alpha_t|\alpha_{t-1}) \\
- (\alpha_t')^2 u'(c_t) \rho(\alpha_t'|\alpha_{t-1}) \beta R \frac{\lambda_t^{\Delta}}{\eta_{t-1}} \pi(\alpha_t'|\alpha_t) \\
= 0
\end{aligned}
$$

Optimal taxes

Sufficient statistics characterisation

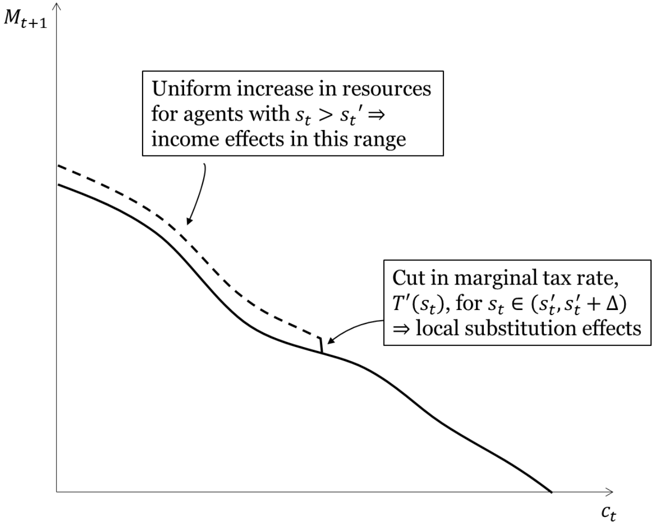
... to something much clearer:

$$\overbrace{\mathbb{E}_{t-1}\left[1 - T_t'(s_t)\frac{ds_t}{dM_t} - g_t(s_t) \middle| s_t > s_t'\right]}^{\text{Net cost of transfers above } s_t'}$$

$$= \underbrace{\frac{\pi^s(s_t')\, s_t'}{1 - \Pi^s(s_t')} T_t'(s_t')\,\varepsilon_t^s + RT_{t-1}'(s_{t-1})\, s_{t-1}\,\epsilon_{t-1,t}^s(s_t')}_{\text{Benefit from behavioural response}}$$

Optimal taxes

Perturbing the tax schedule



Optimal taxes

Observations

- An infinite-horizon problem with continuum of types each period, but...
- At most two elasticities matter!
- The mechanism design problem has simplified sufficient statistics results
  - *C.f. Atkinson-Stiglitz...*
- I show $T_t'(s_t) \geq 0$
  - *Strictly, except possibly at extremes*
  - *Lump-sum transfer, types screened by savings*

How high are top MTRs?

Using sufficient statistics

- Sufficient statistics representations often used to understand top MTRs
- With iid types, characterisation specialises at top to:

$$T_t'(\bar{s}) = \frac{1 - g_t(\bar{s})}{\left.\frac{ds_t}{dM_t}\right|_{\bar{s}} + \varepsilon_t^s a_t(\bar{s})}$$

- $a_t(\bar{s})$: Pareto param for upper tail of savings distn/lower tail of consumption distn
- Unlike static Mirrlees, $g_t(\bar{s}) \nrightarrow 0$

How high are top MTRs

Using sufficient statistics

$$T'_t(\bar{s}) = \frac{1 - g_t(\bar{s})}{\left.\frac{ds_t}{dM_t}\right|_{\bar{s}} + \varepsilon^s_t a_t(\bar{s})}$$

- Difficult objects here are $a_t(\bar{s})$ (conditional!) and $g_t(\bar{s})$
- Latter decomposes via:

$$1 - g_t(\bar{s}) = \chi_t + (1 - \chi_t)\left(\frac{\bar{g}_t - g_t(\bar{s})}{\bar{g}_t}\right)$$

- χ_t: average income effect on taxes; $\bar{g}_t$: average welfare weight...

How high are top MTRs

Implied MTRs

		$a_t = 4$	$a_t = 6$	$a_t = 10$
$\left(\frac{\bar{g}_t - g_t(\bar{s})}{\bar{g}_t}\right) = 0.05$	$\chi_t = 0.01$	0.020	0.015	0.010
	$\chi_t = 0.05$	0.033	0.024	0.016
$\left(\frac{\bar{g}_t - g_t(\bar{s})}{\bar{g}_t}\right) = 0.1$	$\chi_t = 0.01$	0.036	0.027	0.018
	$\chi_t = 0.05$	0.048	0.036	0.024

Two objects in the characterisation less obvious:

1. *Welfare weight* $g_t(s_t)$:

$$g_t(s_t) := \alpha_t u'(c_t) \frac{(1+\lambda_t)}{\eta_t}$$

- 'Value to policymaker of providing extra unit of resources to α_t'
- Policymaker is conditionally utilitarian
- $1+\lambda_t$ an accumulated Pareto weight (martingale)

Optimal taxes

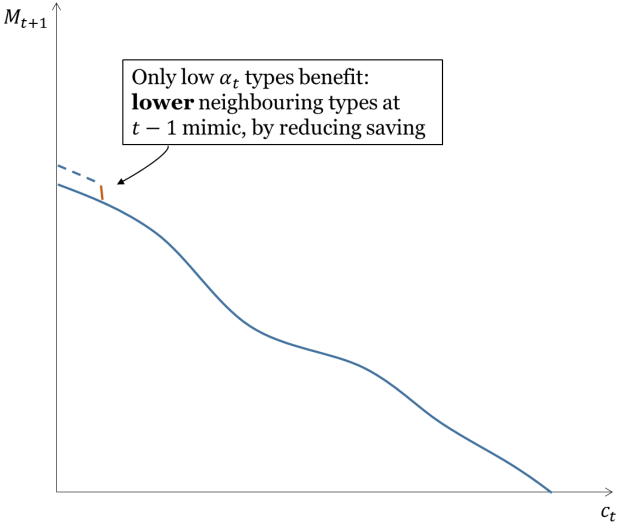
Intertemporal elasticity

2. *Intertemporal elasticity* $\epsilon^{s}_{t-1,t}(s_t)$

   - '**Compensated** responsiveness of savings in $t-1$ to promised tax cut at s_t'
   - Can show $s_{t-1}\epsilon^{s}_{t-1,t}(s_t)$ goes +ve to -ve as s_t increases
   - Tends to make period-t taxes more **progressive**

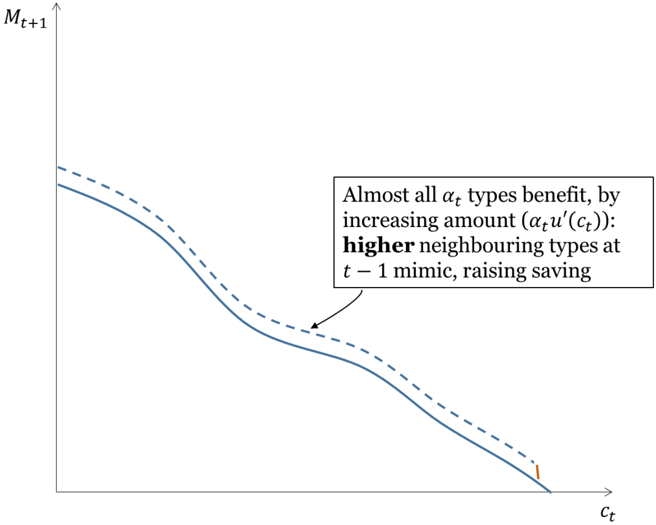
Optimal taxes

Incentives at t-1?



Optimal taxes

Incentives at t-1?



Closing points

- Mechanism design & sufficient statistics approaches are not substitutes!
- They can (and should) inform each other about policy design
- You need a decentralisation to think practically about taxes...
- ... but structure of information-theoretic problem can make this manageable

Value definitions

$$\omega_{t+1}(\alpha_t) = \int_{\alpha_{t+1}} [u_t(\alpha_{t+1}) + \omega_{t+2}(\alpha_{t+1})]\, d\Pi(\alpha_{t+1}|\alpha_t)$$

$$\omega^{\Delta}_{t+1}(\alpha_t) = \int_{\alpha_{t+1}} \left[u_t(\alpha_{t+1}) + \omega^{\Delta}_{t+2}(\alpha_{t+1})\right] \rho(\alpha_{t+1}|\alpha_t)\, d\Pi(\alpha_{t+1}|\alpha_t)$$

back

Multiplier solutions

$$\frac{1+\lambda_{t+1}\left(\alpha^{t}\right)}{\eta_t} = \frac{1}{1-\varepsilon^{\alpha}\left(\alpha_t\right)}\left\{\frac{1}{\mathbb{E}_t\left[\alpha_{t+1}\right]}\mathbb{E}_t\left[\frac{1}{\beta R u'\left(c_{t+1}\right)}\right] - \frac{\varepsilon^{\alpha}\left(\alpha_t\right)}{\alpha_t u'\left(c_t\right)}\right\}$$

$$\lambda^{\Delta}_{t+1}\left(\alpha^{t}\right) = \frac{1}{1-\varepsilon^{\alpha}\left(\alpha_t\right)}\left\{\frac{1}{\alpha_t u'\left(c_t\right)} - \frac{1}{\mathbb{E}_t\left[\alpha_{t+1}\right]}\mathbb{E}_t\left[\frac{1}{\beta R u'\left(c_{t+1}\right)}\right]\right\}$$

back

Marginal cost dynamics

$$\frac{1}{\mathbb{E}_t\left[\alpha_s\right]}\mathbb{E}_t\left[\frac{1}{\left(\beta R\right)^{s-t}u'\left(c_s\right)}\right] = \frac{1+\lambda_{t+1}}{\eta_t} + \frac{d\log\mathbb{E}_t\left[\alpha_s\right]}{d\log\alpha_t}\frac{\lambda_{t+1}^{\Delta}}{\eta_t}$$

$$\lambda_{t+1}\left(\alpha_t\right) = \lambda_t + \mu_t\left(\alpha_t\right)$$

$$\lambda_{t+1}^{\Delta}\left(\alpha_t\right) = \rho\left(\alpha_t|\alpha_{t-1}\right)\lambda_t^{\Delta} - \frac{1}{\alpha_t\pi\left(\alpha_t|\alpha_{t-1}\right)}\int_{\alpha_t}^{\bar{\alpha}}\mu_t\left(\tilde{\alpha}_t\right)d\Pi\left(\tilde{\alpha}_t|\alpha_{t-1}\right)$$

back

First-order approach

Integral condition

- **IC** specifies continuum of constraints for *each* $\alpha'_t \in A$ (at each history node):

$$\alpha'_t \in \arg\max_{\alpha_t} \left\{ \alpha'_t u\left(c_t\left(\alpha_t\right)\right) + \beta \int_{\alpha_{t+1}} V\left(\alpha_{t+1}; \alpha_t\right) d\Pi\left(\alpha_{t+1} | \alpha'_t\right) \right\}$$

- Intractably large constraint set ⇒ replace w minimal necessary requirement: **first-order approach**
- Integral representation (c.f. Milgrom & Segal, 2002):

$$\overbrace{\alpha'_t u_t\left(\alpha'_t\right) + \beta \omega_{t+1}\left(\alpha'_t\right)}^{\text{utility at } \alpha'_t} = \overbrace{\underline{\alpha} u_t\left(\underline{\alpha}\right) + \beta \omega_{t+1}\left(\underline{\alpha}\right)}^{\text{utility of lowest type}} + \int_{\underline{\alpha}}^{\alpha'_t} \underbrace{\frac{1}{\alpha_t}\left[\alpha_t u_t\left(\alpha_t\right) + \beta \omega^{\Delta}_{t+1}\left(\alpha_t\right)\right]}_{\text{information rents}} d\alpha_t$$

First-order approach

Validity?

- Integral constraint necessary but not sufficient for global **IC**
- Conventional to check validity ex-post
- Pavan, Segal & Toikka (2014) give conditions in related (QL) environments
- Easiest condition to work with analytically: $c_s\left(\alpha^{t-1}, \alpha_t, ..., \alpha_s\right)$ non-increasing in α_t
- *'Normality' of consumption at all future nodes*

back

Motivation

Dynamic Mirrlees: a view from the literature

- Piketty & Saez (2013b):

> "This [mechanism design] approach ... derives the most general optimum tax compatible with the informational structure. ... [It] tends to generate tax structures that are highly complex and results that are sensitive to the exact primitives of the model."

back