Optimal Taxation with Heterogeneous Rates of Return

Nick Hoffman[1] Ali Shourideh [1]

[1]Carnegie Mellon University

April 1, 2022

Summary

- Study optimal *nonlinear* taxation of capital income when agents earn heterogeneous returns on entrepreneurial investment
- Assets: entrepreneurial technology with idiosyncratic return, risk-free bond ("investing" and "saving")
- Assume individual investment outputs imperfect substitutes
  - Eliminates undesirable "solution"
- Static model: differential asset taxation
  - Optimal distortions on both types of assets hump-shaped
  - Tax on investing uniformly higher than saving
  - Distortions driven by *aggregate* DRS, informational frictions
- Dynamic model: broadly progressive taxes in IID model
  - Wedge on investing history dependent, wedge on saving not
  - Additional force governing distortions: self-insurance

Motivation

- Thick tails in wealth distribution driven by capital income risk: Benhabib *et al.* (2011), Benhabib *et al.* (2019)
  - Multiplicative effect, rather than additive
- How should capital income, wealth be taxed?
  - Can tax capital income to redistribute from wealthy to poor
  - Efficiency: excessive capital taxation discourages investment and lowers output
- Mirrlees (1971): framework to consider redistribution/efficiency tradeoff in the presence of informational frictions

Literature Review

- Positive capital income taxes with informational frictions: Golosov *et al.* (2003), Kocherlakota (2005), Albanesi and Sleet (2006)
  - Common rate of return, labor market considerations
- Optimal taxation of entrepreneurs: Albanesi (2006), Scheuer (2014)
  - Entrepreneurial returns dependent on effort
  - Support for differential rates
- Guvenen *et al.* (2019): welfare improvement with *wealth* taxes in a similar model
- Heterogeneous returns: Shourideh (2014), Gerritsen *et al.* (2020)

Static Model: Households and Productivity

Households

- Continuum of households who differ in type θ, which determines both *productivity*, and *variety* of intermediate good
- Assume $\theta \in \Theta = \left[\underline{\theta}, \bar{\theta}\right]$, with CDF $F(\theta)$, privately known by household
- Household can borrow/lend b at common rate R
- Household of type θ can also invest k capital and produce θk of their variety of intermediate good
  - Sells to final good producer at price $p(\theta)$
  - Price-taker, so individual technologies are CRS

Static Model: Aggregate Production and Prices

- Final good producer combines intermediate goods to produce final good using CES aggregation technology:

$$Y = \left(\int_{\Theta} \left[\theta k\left(\theta\right) \right]^{\frac{\varepsilon-1}{\varepsilon}} dF\left(\theta\right) \right)^{\frac{\varepsilon}{\varepsilon-1}}$$

- $\varepsilon > 1$ elasticity of substitution
  - Ensures that at optimal solution, all types will invest
- From FG producer problem, prices are

$$p(\theta) = \left(\frac{Y}{\theta k\left(\theta\right)} \right)^{\frac{1}{\varepsilon}}$$

Static Model: Government

- Government levies taxes on income from investing $\theta kp(\theta)$ and saving Rb according to tax function T
- Taking T and p as given, household solves

$$\max_{k,b} \; u(w-k-b) + \beta u\left[\theta kp(\theta) + Rb - T(\theta kp, Rb)\right]$$

- Benevolent government chooses T to maximize social welfare:

$$\max \int_{\Theta} U(\theta)\, dF(\theta) \tag{1}$$

subject to revenue requirements and household optimality

Static Model: Mechanism Design Problem

- Following from Mirrlees (1971), can recast government's problem in terms of mechanism design
  - Revelation Principle: direct mechanism
  - Household reports θ and receives allocations
- Objective same as in (1)
- Feasibility constraints:

$$w \geq \int_{\Theta} \left[c_0(\theta) + k(\theta) \right] dF(\theta)$$

$$\left(\int_{\Theta} \left[\theta k(\theta) \right]^{\frac{\varepsilon-1}{\varepsilon}} dF(\theta) \right)^{\frac{\varepsilon}{\varepsilon-1}} \geq \int_{\Theta} c_1(\theta)\, dF(\theta)$$

Static Model: Mechanism Design Problem

- Additional constraints: incentive compatibility

$$U(\theta) \geq \ln\left(c_0\left(\hat{\theta}\right) + k\left(\hat{\theta}\right) - \frac{\hat{\theta} k\left(\hat{\theta}\right)}{\theta}\right) + \beta \ln c_1\left(\hat{\theta}\right), \quad \forall \theta, \hat{\theta} \in \Theta$$

  where

$$U(\theta) = \ln c_0(\theta) + \beta \ln c_1(\theta)$$

- Simplifying assumption: while the planner (government) cannot observe θ, the market for intermediate goods *can*
- If type θ claims to be of type $\hat{\theta}$, they still receive price $p(\theta)$

Static Model: Optimal Distortions

- From household problem, wedges are

$$\tau_k(\theta) = 1 - \frac{u'(c_0)}{\beta u'(c_1)\,\theta p(\theta)}$$
$$\tau_b(\theta) = 1 - \frac{u'(c_0)}{\beta R u'(c_1)}$$

- First and second partial derivatives of tax function T
- Optimal *distortions*, not necessarily optimal taxes

Static Model: Optimal Distortions

- Assume log utility.
- From optimality conditions in social planner's problem, can state wedges as

$$\tau_k = \left(1 + \frac{k}{c_0}\right)\left(1 - \frac{R}{\theta p(\theta)}\right)$$
$$\tau_b = \frac{k}{c_0}\left(\frac{\theta p(\theta)}{R} - 1\right)$$

- Two key determinants: $\frac{k}{c_0}$, and $\frac{R}{\theta p(\theta)}$
- Both wedges nonnegative, and positive on interior of Θ
- $\frac{k}{c_0}$ pulls wedges upward, effect of $\frac{R}{\theta p(\theta)}$ depends on θ

Static Model: Marginal Product

Lemma

At optimum, $R \leq \theta p(\theta)$*, with equality if and only if* $\theta \in \left\{\underline{\theta}, \overline{\theta}\right\}$.

- $\theta p(\theta)$ is the *societal* marginal product of capital at the optimum for type θ, which takes into account effect of k on p
- Using formula for prices,

$$\theta p(\theta) = \theta^{1-\frac{1}{\varepsilon}} \left(\frac{Y}{k\left(\theta\right)}\right)^{\frac{1}{\varepsilon}}$$

$\implies$ diminishing returns in aggregate

- Full information: planner wishes to equate $\theta p(\theta)$ with R
- Informational frictions prevent this, $\frac{\theta p(\theta)}{R}$ hump-shaped instead
- For high values of θ, this term pulls wedges towards zero

Static Model: Positive Wedge on Investing

- Consequence of Lemma: optimal distortions are positive on the interior of Θ
- Intuition for $\tau_k > 0$: planner distorts investment choice to prevent over-supply of variety θ
- Without uncertainty, these can be interpreted as Pigouvian taxes
  - Externality: when I scale up, I affect not only $p(\theta)$, but also the *entire* pricing schedule
- Counterfactual: if households monopolists, investment wedge is lower and often negative; planner corrects for under-supply of variety

Static Model: Differential Asset Taxation

Proposition

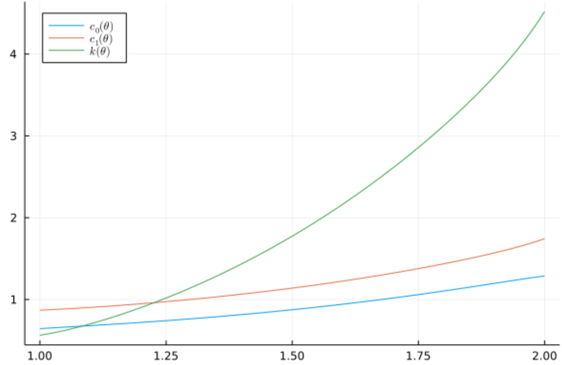
$\tau_k(\theta) \geq \tau_b(\theta)$*, with equality if and only if* $\theta \in \{\underline{\theta}, \overline{\theta}\}$*.*

- Optimality of positive τ_b stems directly from positive τ_k
  - Endowment is perfectly fungible between the two vehicles
- Larger τ_k: investing income is a signal of private information, savings income is not

Static Model: Numerical Example (Dual Problem)

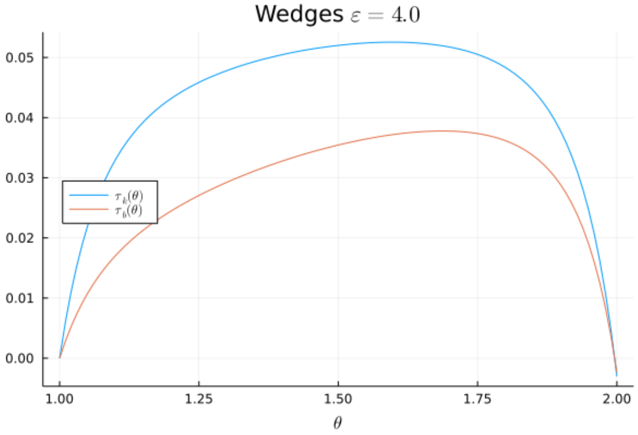
- $\theta \sim N(1.5, 0.2)$ truncated over $\Theta = [1, 2]$
- $\beta = 0.9$, $\varepsilon = 4$, $w = 1$, $R = 1.5$
- Optimal allocations:

Figure: Allocations in the Static Model



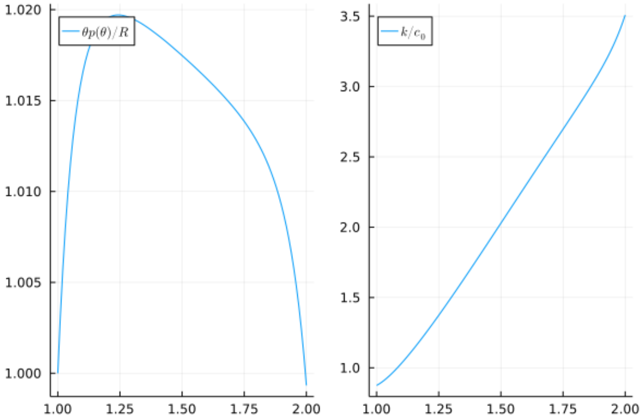
Static Model: Numerical Wedges

Figure: Wedges in the Static Model



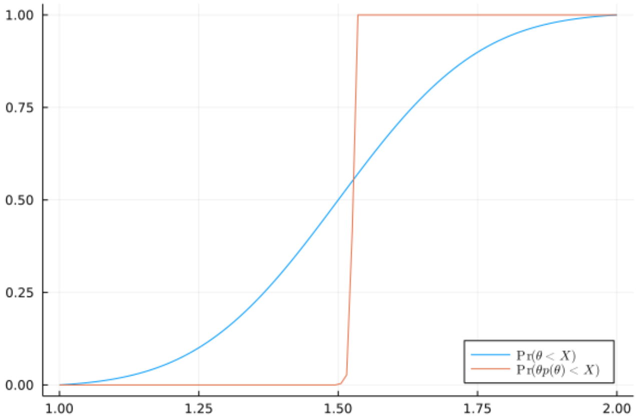
Static Model: Determinants of Wedges

Figure: Determinants of Wedges in the Static Model



Static Wedges: Squeezing

Figure: Squeezing Rates of Return: CDFs



Static Model: Main Takeaways

- Distortions driven by efficiency-equity tradeoff, informational rents, pricing concerns
- Farhi and Werning (2010): distortions "squeeze" post-tax rates of return
  - Entrepreneurs price-takers $\implies$ positive distortions
  - Variance in c_1 much smaller than variance in θ
- Scheuer (2014): intensive rather than extensive margin
  - Rather than determining who selects into entrepreneurship, the tax code ensures that everyone invests the right amount
- Differential taxation: $\tau_k > \tau_b$ on interior of Θ
  - Arbitrage, private information

Dynamic Model: IID Case

Production

- Aggregate producer follows two-step process for Y_t:
  1. Combines intermediate capital goods to create single capital good $K_{t,f}$:

$$K_{t,f} = \left(\int \left[\theta_{t-1} k_t \left(\theta^{t-1} \right) \right]^{\frac{\varepsilon - 1}{\varepsilon}} d\mu_{t-1} \left(\theta^{t-1} \right) \right)^{\frac{\varepsilon}{\varepsilon - 1}}$$

  2. Combines $K_{t,f}$ with labor L_t, to produce Y_t according to Cobb-Douglas technology, $Y_t = K_{t,f}^{\alpha} L_t^{1-\alpha}$
- Assumption: $L_t = L = 1 \; \forall t$, so final good output is

$$Y_t = K_{t,f}^{\alpha}$$

which ensures that a steady state in aggregates (Y, K_f) will exist

Dynamic Model: IID Case

Households

- Time is discrete, each period plays out as follows:
  1. Agent realizes investing income

$$y_t\left(\theta^{t-1}\right) = p_t\left(\theta^{t-1}\right)\theta_{t-1}k_t\left(\theta^{t-1}\right)$$

  2. Agent draws new type θ_t from $F(\theta)$, draws IID across time. Then, agent makes choices c_t, k_t, b_t
- Let $\theta^t = \{\theta_0, \theta_1, ..., \theta_t\}$ denote the history of shocks through time t

Dynamic Model: IID Wedges

Wedges will be given by

$$\tau_{k,t}\left(\theta^t\right) = 1 - \frac{1}{\beta c_t\left(\theta^t\right)\theta_t p_{t+1}\left(\theta^t\right)\mathbb{E}\left[c_{t+1}\left(\theta^{t+1}\right)^{-1}\right]}$$

$$\tau_{b,t}\left(\theta^t\right) = 1 - \frac{1}{\beta c_t\left(\theta^t\right)R_t\mathbb{E}\left[c_{t+1}\left(\theta^{t+1}\right)^{-1}\right]}$$

Dynamic Model: Planning Problem

- Let $\mu_t(\theta^t)$ denote the measure of period-t histories induced by the stochastic process for θ_t.
- The planner chooses allocations $\{c_t(\theta^t), k_{t+1}(\theta^t)\}_{t=0}^{\infty}$ to solve

$$\max \sum_{t=0}^{\infty} \beta^t \int u\left(c_t\left(\theta^t\right)\right) d\mu_t\left(\theta^t\right) \tag{2}$$

- Feasibility assumes entrepreneurs allocated capital share αY:

$$\int \left[c_t\left(\theta^t\right) + k_{t+1}\left(\theta^t\right)\right] d\mu_t\left(\theta^t\right)$$

$$= \alpha \left(\int \left[\theta_{t-1} k_t\left(\theta^{t-1}\right)\right]^{\frac{\varepsilon-1}{\varepsilon}} d\mu_{t-1}\left(\theta^{t-1}\right) \right)^{\frac{\alpha\varepsilon}{\varepsilon-1}}$$

$$= \int \theta_{t-1} p_t\left(\theta^{t-1}\right) k_t\left(\theta^{t-1}\right) d\mu_{t-1}\left(\theta^t\right)$$

Dynamic Model: Incentive Constraints

- Promise utility allocated to an agent of history θ^t is

$$w_{t+1}\left(\theta^t\right) = \sum_{s=t+1}^{\infty} \beta^{s-t-1} \int u\left(c_s\left(\theta^s\right)\right) d\mu_s\left(\theta^s \middle| \theta^t\right)$$

- The *local* incentive constraints require

$$\frac{\partial U_t\left(\theta^t\right)}{\partial \theta_t} = u'\left(c_t\left(\theta^t\right)\right) \frac{k_{t+1}\left(\theta^t\right)}{\theta_t}$$

Dynamic Model: Solving the Planning Problem

- We focus on the dual (cost-minimization) problem of a *component* planner
  - Considers the problem of a single agent of history θ^{t-1}
  - Takes the path of prices $\{p_{s+1}(\theta^s)\}_{s \geq t}$ as given

Dynamic Model: Exploiting Homogeneity

- Although the aggregate value function (cost-minimization) is not homogeneous, the component value function *is*, as we assume that the CP takes prices as given
  - Similar to Angeletos (2007)
- This allows us to derive a recursive formulation
- Implies that the policy function for $k_{t+1}\left(\theta^t\right)$ can be written as

$$
\begin{aligned}
k_{t+1}\left(\theta^t\right) &= \overline{k}_{t+1}\left(\theta_t\right)\exp\left[\left(1-\beta\right)w_t\left(\theta^{t-1}\right)\right] \\
&= \overline{k}_{t+1}\left(\theta_t\right)\exp\left[\left(1-\beta\right)\left(w'\left(\theta_{t-1}\right)+...+w'\left(\theta_0\right)+w_0\right)\right]
\end{aligned}
$$

for some functions $\overline{k}\left(\theta_t\right)$ and $w'\left(\theta_t\right)$, in the case of log utility

Dynamic Model: Exploiting Homogeneity

- This decomposition implies a similar process for prices:

$$p_{t+1}\left(\theta^{t}\right)=\overline{p}_{t}\left(\theta^{t-1}\right)\hat{p}_{t}\left(\theta_{t}\right)$$

- Can decompose price into $\overline{p}$, which encodes history, and $\hat{p}$, which only depends on θ_t
- "Common price" $\overline{p}$ evolves according to

$$\overline{p}_{t+1}\left(\theta^{t}\right)=\overline{p}_{t}\left(\theta^{t-1}\right)\tilde{p}\left(\theta_{t}\right)$$

where

$$\tilde{p}\left(\theta_{t}\right)=\exp\left[-\frac{(1-\beta)}{\varepsilon}w'\left(\theta_{t}\right)\right]$$

Dynamic Model: Recursive Formulation

Assuming constant R, the recursive problem is:

$$C(w, \overline{p}) = \min_{\substack{c(\theta), k'(\theta), \\ w'(\theta), U(\theta)}} \int \left\{ c(\theta) + k'(\theta) + R^{-1} \left[C\left(w'(\theta), \overline{p} \cdot \tilde{p}(\theta) \right) - \overline{p} \cdot \hat{p}(\theta) \theta k'(\theta) \right] \right\} dF(\theta)$$

subject to

$$w \leq \int U(\theta) dF(\theta)$$

$$U(\theta) = u(c(\theta)) + \beta w'(\theta)$$

$$U'(\theta) = u'(c(\theta)) \frac{k'(\theta)}{\theta}$$

Dynamic Model: Exploiting Homogeneity Once More

Proposition

Suppose $u(c) = \ln c$. Then, the component planner's problem has the following solution:

$$C\left(w,\overline{p}\right) = A\left(\overline{p}\right)e^{(1-\beta)w} \qquad w'\left(\theta,w,\overline{p}\right) = w'\left(\theta,\overline{p}\right) + w$$
$$c\left(\theta,w,\overline{p}\right) = c\left(\theta,\overline{p}\right)e^{(1-\beta)w} \qquad U\left(\theta,w,\overline{p}\right) = U\left(\theta,\overline{p}\right) + w$$
$$k'\left(\theta,w,\overline{p}\right) = k'\left(\theta,\overline{p}\right)e^{(1-\beta)w}$$

for some functions $A\left(\overline{p}\right), c(\theta,\overline{p}), k'(\theta,\overline{p}), w'(\theta,\overline{p}), U(\theta,\overline{p})$.

- Implication: we can solve the above problem for $w = 0$ to obtain "baseline" functions $A\left(\overline{p}\right), c(\theta,\overline{p}), k'(\theta,\overline{p}), w'(\theta,\overline{p}), U(\theta,\overline{p})$.

Dynamic Model: Optimal Wedges

- By definition,

$$\overline{p}_t\left(\theta^{t-1}\right) = \alpha K_f^{\alpha-1} K_f^{\frac{1}{\varepsilon}} \exp\left[-\frac{(1-\beta)}{\varepsilon} w_t\left(\theta^{t-1}\right)\right]$$

- In general equilibrium, there will be only one value of $\bar{p}$ consistent with $w = 0$, call this $\bar{p}_0$.
  - This is the value on which the "baseline" allocations are based.
  - We still need to solve $A(\bar{p})$ for a range of $\bar{p}$ values, but this value will not change from one period to the next

Dynamic Model: Optimal Wedges

- Thus, the optimal wedges will be given by the following:

$$\tau_k(\theta, \bar{p}, w) = 1 - \frac{\exp\left[(1-\beta)\, w'(\theta, \bar{p}_0)\right]}{\beta c(\theta, \bar{p}_0)\, \bar{p}\hat{p}(\theta)\, \theta \mathbb{E}\left[c(\theta', \bar{p}_0)^{-1}\right]}$$

$$\tau_b(\theta, \bar{p}, w) = 1 - \frac{\exp\left[(1-\beta)\, w'(\theta, \bar{p}_0)\right]}{\beta R c(\theta, \bar{p}_0)\, \mathbb{E}\left[c(\theta', \bar{p}_0)^{-1}\right]}$$

- τ_b independent of history, depends only on θ_t
- τ_k depends on θ_t, as well as history through $\bar{p}$

Dynamic Model: Optimal Wedges

Interpretation: Several Forces

- For both τ_b and τ_k holding $\bar{p}$ fixed, wedges will be closely related to their static counterparts
  - In particular: correcting externalities
  - τ_k will be increasing in $\bar{p}$, so decreasing in promise utility w
- There will be an additional force: household's insurance (consumption-smoothing) motive and its effect on incentive constraints.

Dynamic Model: Numerical Example

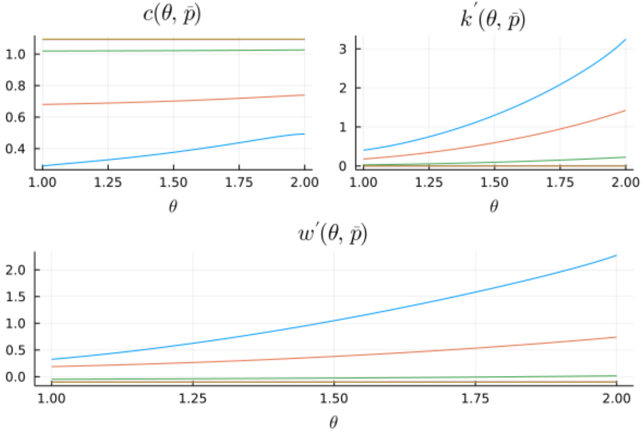
- Can solve the recursive problem for $A(\bar{p})$ using the following parameterization:

$$
\begin{aligned}
\beta &= 0.9 & \varepsilon &= 4 \\
R &= 1.1 & \left\{\underline{\theta}, \overline{\theta}\right\} &= \{1, 2\} \\
\{p_H, p_L\} &= \{0, 0.9\} & \theta &\sim N(1.5, 0.2)
\end{aligned}
$$

- Again assume that the distribution for θ is truncated, only has support on Θ

Dynamic Model: Allocations

Figure: Allocations in Infinite-Horizon Case



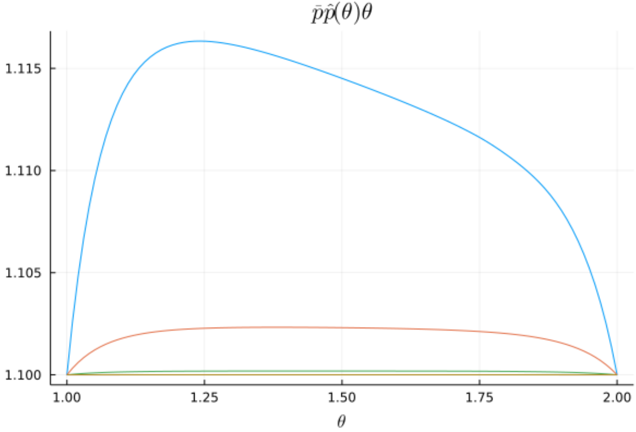
Dynamic Model: Allocations

Main takeaway is the relationship of promised utility to the "baseline" allocations

- Consumption c decreasing in $\bar{p}_0$, promise utility w' increasing
- Recall that $\bar{p}$ and w move in opposite directions: lower $\bar{p}$ $\implies$ higher w
- Returns also increasing in $\bar{p}$
- Upshot: consumption today c delivers on past promise, continuation promise utility w' incentivizes current investment

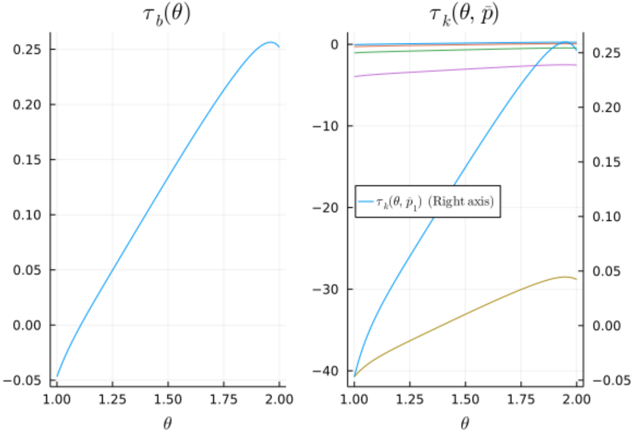
Dynamic Model: Rates of Return

Figure: Rates of Return in Infinite-Horizon Case



Dynamic Model: Numerical Wedges

Figure: Wedges in Infinite-Horizon Case



Dynamic Model: Investing Wedge τ_k

$$\tau_k(\theta, \bar{p}, w) = 1 - \frac{\exp\left[(1-\beta)\, w'(\theta, \bar{p}_0)\right]}{\beta c(\theta, \bar{p}_0)\, \bar{p}\hat{p}(\theta)\, \theta \mathbb{E}\left[c(\theta', \bar{p}_0)^{-1}\right]}$$

- τ_k will depend on $\bar{p}_0$, $\bar{p}$, and θ
- In general: will be increasing in θ over most of the distribution (*progressive*)
- Slope and magnitude (and thus sign) will depend on $\bar{p}_0$ and $\bar{p}$
- *Generally speaking*: decreasing in $\bar{p}_0$, increasing in $\bar{p}$
  - Higher $\bar{p}_0$: less consumption, more w' across all types
  - Higher $\bar{p}$: lower state (w), higher returns across all θ (insurance motive again)

Optimal Capital Taxation: Dynamic Lessons

What do we make of these results?

- Addition of uncertainty in dynamic context produces broadly progressive wedges
  - Insurance motive: if I have a high type today, additional investing can tighten my incentive constraint tomorrow (I insure myself against a bad shock)
- In both cases, taxation "squeezes" returns, as in Farhi and Werning (2010)
- Differential taxation result remains
- When returns driven by private information, optimal distortions are history-dependent
  - Possible role for taxation based on *wealth*

Conclusion

- Studied optimal nonlinear taxation of capital income with heterogeneous returns
- Static model: positive, hump-shaped wedges
  - Investing income subject to strictly larger distortions than savings income
- Dynamic model: derived recursive form, solved simplified planning problem
  - Investing wedge history-dependent and regressive in promise utility
  - Savings wedge independent of history prior to t

References I

Stefania Albanesi and Christopher Sleet. Dynamic optimal taxation with private information. *The Review of Economic Studies*, 73(1):1–30, 2006.

Stefania Albanesi. Optimal taxation of entrepreneurial capital with private information, 2006.

George-Marios Angeletos. Uninsured idiosyncratic investment risk and aggregate saving. *Review of Economic dynamics*, 10(1):1–30, 2007.

Jess Benhabib, Alberto Bisin, and Shenghao Zhu. The distribution of wealth and fiscal policy in economies with finitely lived agents. *Econometrica*, 79(1):123–157, 2011.

Jess Benhabib, Alberto Bisin, and Mi Luo. Wealth distribution and social mobility in the us: A quantitative approach. *American Economic Review*, 109(5):1623–47, 2019.

Emmanuel Farhi and Iván Werning. Progressive estate taxation. *The Quarterly Journal of Economics*, 125(2):635–673, 2010.

References II

Aart Gerritsen, Bas Jacobs, Alexandra Victoria Rusu, and Kevin Spiritus. Optimal taxation of capital income with heterogeneous rates of return. 2020.

Mikhail Golosov, Narayana Kocherlakota, and Aleh Tsyvinski. Optimal indirect and capital taxation. *The Review of Economic Studies*, 70(3):569–587, 2003.

Fatih Guvenen, Gueorgui Kambourov, Burhanettin Kuruscu, Sergio Ocampo-Diaz, and Daphne Chen. Use it or lose it: Efficiency gains from wealth taxation. Technical report, National Bureau of Economic Research, 2019.

Narayana R Kocherlakota. Zero expected wealth taxes: A mirrlees approach to dynamic optimal taxation. *Econometrica*, 73(5):1587–1621, 2005.

James A Mirrlees. An exploration in the theory of optimum income taxation. *The review of economic studies*, 38(2):175–208, 1971.

References III

Florian Scheuer. Entrepreneurial taxation with endogenous entry. *American Economic Journal: Economic Policy*, 6(2):126–63, 2014.

Ali Shourideh. Optimal taxation of wealthy individuals. Technical report, Working paper, 2014.