Optimal Taxation and R&D Policies

Ufuk Akcigit Chicago

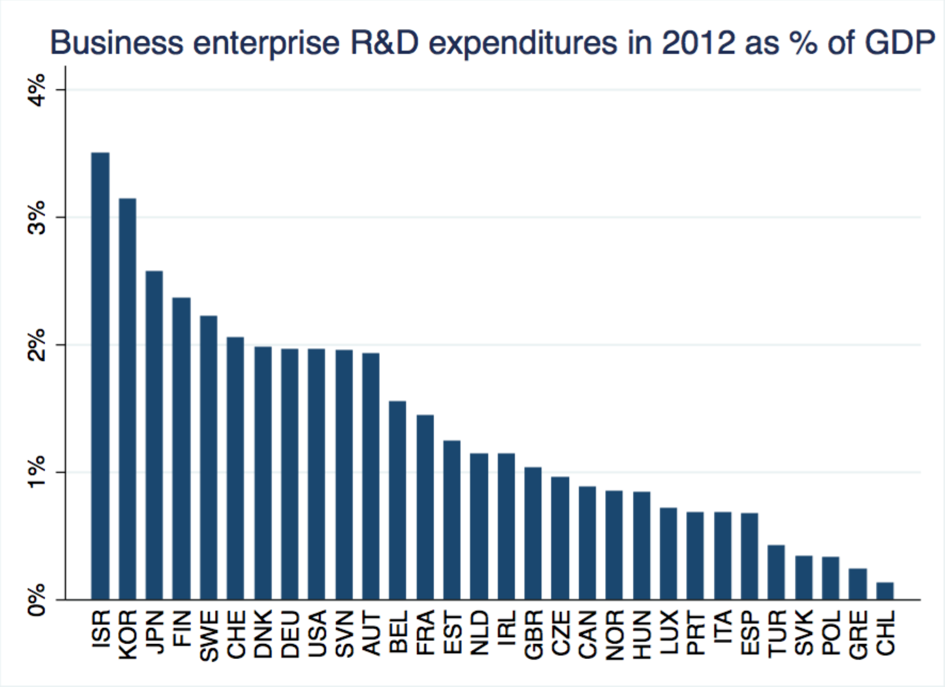
Douglas Hanley Pittsburgh

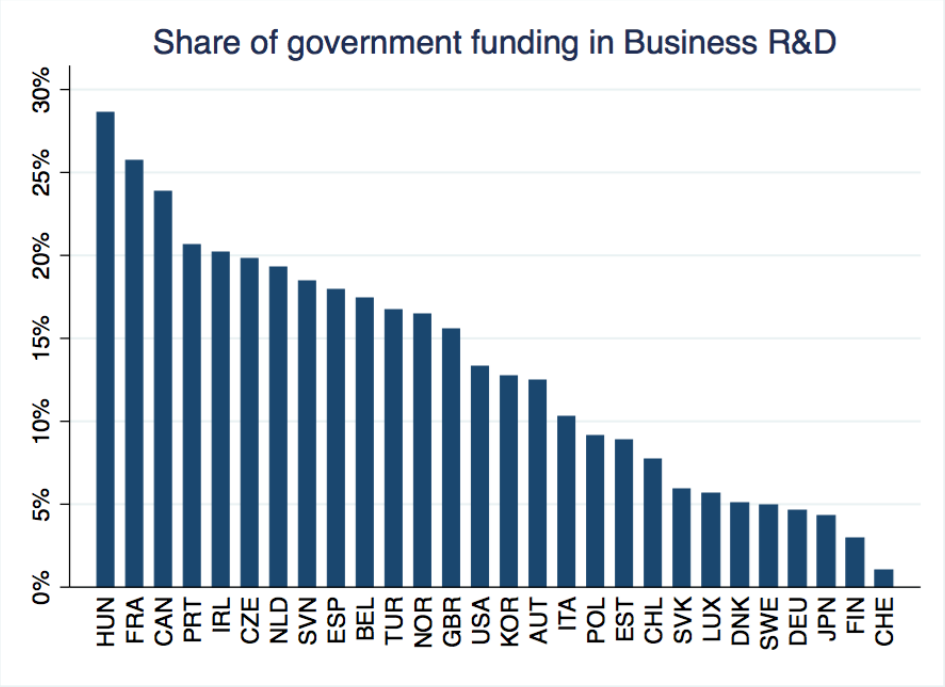
Stefanie Stantcheva Harvard

June 23, 2015

Motivation: Widespread R&D and Industrial Policies

- Industrial policies are widespread, costly and not fully understood.
- EU: 9.6% of EU GDP on industrial policies in 2010.
- Lots of policies target **innovation and R&D**.


Business enterprise R&D expenditures in 2012 as % of GDP
4%
3%
2%
1%
0%
ISR
KOR
JPN
FIN
SWE
CHE
DNK
DEU
USA
SVN
AUT
BEL
FRA
EST
NLD
IRL
GBR
CZE
CAN
NOR
HUN
LUX
PRT
ITA
ESP
TUR
SVK
POL
GRE
CHL


Share of government funding in Business R&D
30%
25%
20%
15%
10%
5%
0%
HUN
FRA
CAN
PRT
IRL
CZE
NLD
SVN
ESP
BEL
TUR
NOR
GBR
USA
KOR
AUT
ITA
POL
EST
CHL
SVK
LUX
DNK
SWE
DEU
JPN
FIN
CHE

Goal: Optimal Design of R&D and Corporate Taxation

- Little agreement on how effective these policies are and how to design them.
- In this paper: Optimal R&D and Corporate Taxation Policies.
  - Consider heterogeneous firms $\rightarrow$ Firm quality is **private information**.
  - Firm faces uncertainty about its quality over time.
  - R&D returns are stochastic.
  - **Spillovers** between firms.
  - Do not restrict policies ex ante: let optimal policy tools arise endogenously.
  - First: optimal mechanism design, then consider "simpler" policies.

What we do:

- Model of R&D and endogenous growth with heterogeneous firms.
- Mechanism design setup: Firms are agents, government is the Principal.
- Derive analytically fully unrestricted optimal dynamic mechanism $\rightarrow$ characterize allocations using "wedges" or implicit taxes and subsidies.
- Estimate important parameters of the model using patent data matched to Compustat (not yet today)
  - Numerically illustrate life cycle path of optimal policies for firms (by age, quality..).
- How well do simpler policies perform (linear size-independent age-dependent, linear size and age-independent policies, etc..).
  - Formulas in terms of "sufficient statistics."

Related Literature

R&D Policies and Growth: Acemoglu, Akcigit, Bloom and Kerr (2013), Atkeson and Burstein (2014), Lentz and Mortensen (2015).

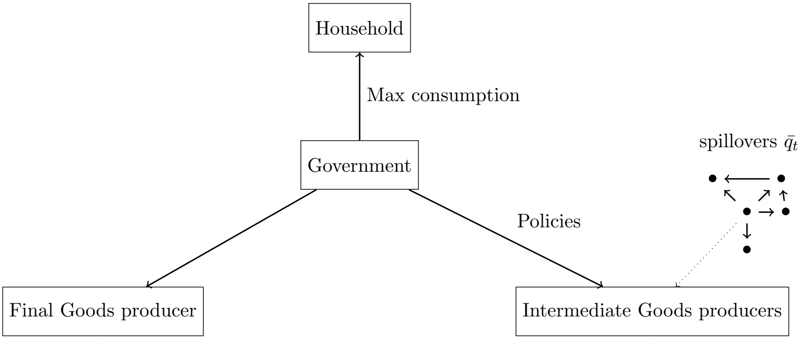
Empirics of R&D Policies: Bloom, Griffith, Van Reenen (2000), Criscuolo *et al.* (2013), Hall (1992), Goolsbee (1998), Romer (2001), Wilson (2009)

Optimal Policy Design: Hopenhayn and Nicolini (1997), Golosov, Tsyvinski and Werning (2006), Shourideh (2012), Farhi and Werning (2013, 2014), Stantcheva (2014), Pavan, Segal and Toikka (2013).

Outline

Outline

Model: Summary



$Y(\bar{q}_t) = \int_{\Theta^t} k(\theta^t)^{1-\beta} q(\theta^t)^{\beta} p(\theta^t)\, d\theta^t$

- Manager: effort l_t
- R&D : $q(\theta^t) = q(\theta^{t-1}) + \lambda_t(r(\theta^{t-1}), l(\theta^t), \theta_t)$
- Produce $k(\theta^t)$
- $\pi(\theta^t, \bar{q}_t) = \max_k \{p(\theta^t)k - C(k, \bar{q}_t)\}$

Model: Intermediate Producers R&D Investments and Efforts

- Each producer can improve his quality q_t through R&D: $q_t = (1-\delta)q_{t-1} + \lambda_t$
  - λ_t: step size, δ: depreciation factor.
- The step size $\lambda_t(r_{t-1}, l_t, \theta_t)$ depends on:
  - R&D resources r_t cost $M_t(r_t)$.
  - Type θ_t (managerial/firm quality).
  - Managerial/firm effort: l_t at cost $\phi_t(l_t)$.
  - $\frac{\partial\lambda}{\partial\theta} > 0$, $\frac{\partial\lambda}{\partial r} > 0$, $\frac{\partial\lambda}{\partial l} > 0$, $\frac{\partial^2\lambda}{\partial\theta\partial l} > 0$ (screening).
  - Returns to R&D are stochastic, depend on stochastic type.

Model: Intermediate Producers' Production

- Firm born with θ_1, evolves with Markov $f^t(\theta_t|\theta_{t-1})$.
- History: $\theta^t = \{\theta_1, ..., \theta_t\}$ with probability $P(\theta^t) = f^t(\theta_t|\theta_{t-1})...f^1(\theta_1)$.
- Firm lifecycle is T periods (max time after which exogenous death).
- $k(\theta^t)$: quantity and $q(\theta^t)$: quality.
- Linear production cost: $C(k, \bar{q}_t) = \frac{k}{\bar{q}_t}$
- Aggregate quality (Spillovers): $\bar{q}_t = \int_{\Theta^t} q(\theta^t)P(\theta^t)d\theta^t$
- Hence profits are:
  $\pi(\theta^t, \bar{q}_t) \equiv \max_k\{p(\theta^t)k - C(k, \bar{q}_t)\} = q_t(\theta^t)(1\text{-}\beta)^{\frac{1-\beta}{\beta}} \cdot \beta \cdot \bar{q}_t^{\frac{1-\beta}{\beta}}$
- Final good: $Y_t = \int_{\Theta^t} q(\theta^t)^\beta k(\theta^t)^{1-\beta}P(\theta^t)d(\theta^t)$

Government Policies Studied

- Fully unrestricted mechanism.
- Linear (size-independent) age-dependent corporate tax and R&D subsidies.
- Linear (size- and age-independent) corporate tax and R&D subsidies.
- Nonlinear (i.e., size dependent) corporate tax and R&D subsidies.

What do these policies look like at optimum?

What is the welfare gain from each?

Outline

Direct Revelation Mechanism with Spillovers

- In this section: $\theta_t = p\theta_{t-1} + \psi_t$, with ψ_t distributed $N(0, \sigma_\psi^2)$.
- History θ^t and effort l_t are private info.
- Government sees: step size λ, realized quality q, R&D spending r, and production k.
- Direct revelation: Firm reports $\theta'_t(\theta^t)$.
- History of reports: $\theta'^t = \{\theta'_1(\theta_1), ...\theta'_t(\theta^t)\}$.
- Allocations as function of history of reports: $\{\lambda(\theta'^t)$, $r(\theta'^t), T_t(\theta'^t)\}$.
- Maximize household consumption (wlog, zero weight on manager): $\mathbb{E}\left\{\sum_{t=1}^{T}\left(\frac{1}{R}\right)^{t-1}\{\pi(\theta^t, \bar{q}_t) - M_t(r(\theta^t)) - T_t(\theta^t)\}\right\}$
- First: let's assume full depreciation $\delta = 1$.

Incentive compatibility and a First-order Approach

- Lifetime expected utility from assigned allocations under truthful reporting: $V_1(\theta_1) \equiv V_1(\{\lambda(\theta^s), r(\theta^s), T_s(\theta^s)\}_{s=1}^{T}, \theta_1) = \sum_{t=1}^{T}(\frac{1}{R})^{t-1} \cdot \{\int_{\Theta^t} \{T_t(\theta^t) - \phi_t(\lambda(\theta^t), r(\theta^{t-1}), \theta_t\} P(\theta^t|\theta_1) d\theta_t\}$

- Utility for a given sequence of realizations θ^T:
$\tilde{U}(\theta^T) = \sum_{t=1}^{T} \left(\frac{1}{R}\right)^{t-1} \{T_t(\theta^t) - \phi_t(\lambda(\theta^t), r(\theta^{t-1}), \theta_t)\}$

- Envelope condition (Pavan, Segal and Toikka, 2014):
$\frac{\partial V_t(\theta^t)}{\partial \theta_t} = \mathbb{E}\{\sum_{s=t}^{T} I_{t,s} \frac{\partial \tilde{U}(\theta^T)}{\partial \theta_s}\}$
  - $I_{t,s}$: impulse response of shock θ_t on time s shock θ_s. For AR(1) is p^{s-t}.

- Informational rent: $V_1(\theta_1) =$
$\mathbb{E}\left\{\frac{1-F^1(\theta_1)}{f^1(\theta_1)} \sum_{t=1}^{T} \left(I_{1,t} \frac{\partial \tilde{U}(\{\lambda(\theta^s), r(\theta^s), T_s(\theta^s)\}_{s=1}^{T}, \theta^T)}{\partial \theta_t} \right)\right\} + V_1(\underline{\theta}_1)$.

Program: Maximize Virtual Surplus with Spillovers

Notation: $\lambda(\theta^t) = \lambda(r(\theta^{t-1}), l(\theta^t), \theta_t)$.

$$\max \sum_{t=1}^{T} (\frac{1}{R})^{t+1} \{ \int_{\Theta^t} \{ \pi_t(\theta^t, \bar{q}_t) - M_t(r(\theta^t)) - \phi(l(\theta^t)) - \frac{1 - F^1(\theta_1)}{f^1(\theta_1)} p^t [\phi'(l(\theta^t)) \frac{\partial \lambda(\theta^t)/\partial \theta_t}{\partial \lambda(\theta^t)/\partial l_t}] - V_1(\underline{\theta}_1) \} P(\theta^t) d\theta^t \}$$

s.t.:

$$\int_{\Theta_t} \lambda(\theta^t) P(\theta^t) d\theta^t = \bar{q}_t$$

- Solve using this first-order approach, then verify global (IC) ("sufficiency") *ex post* (Farhi and Werning, 2013).
- Cohort-by-cohort problem (pooling across cohorts can be a second step).

Wedges: Measures of Implicit taxes and Subsidies

- Implicit corporate tax:

$$(1-\tau(\theta^t))\frac{\partial \pi_t(\theta^t)}{\partial \lambda_t}\frac{\partial \lambda(\theta^t)}{\partial l_t} = \phi_{lt}(\lambda(\theta^t), r(\theta^{t-1}), \theta_t)$$

- Implicit R&D subsidy:

$$(1-\tilde{s}(\theta^t))M_t'(r(\theta^t)) = \frac{1}{R}\mathbb{E}\left\{\frac{\partial \pi_{t+1}(\theta^{t+1})}{\partial \lambda_{t+1}}\cdot\frac{\partial \lambda(\theta^{t+1})}{\partial r_t}(1-\tau(\theta^{t+1}))\right\}$$

- Net R&D subsidy:

$$s(\theta^t) = \tilde{s}(\theta^t) - \frac{1}{M_t'(r(\theta^t))}\frac{1}{R}\mathbb{E}\left\{\frac{\partial \pi_{t+1}(\theta^{t+1})}{\partial \lambda_{t+1}}\cdot\frac{\partial \lambda(\theta^{t+1})}{\partial r_t}\tau(\theta^{t+1})\right\}$$

Sufficient Statistics: Elasticities and Coefficient of Complementarity

- Elasticity of x_t w.r.t y_t: $\varepsilon_{xy,t} \equiv \frac{\partial x_t}{\partial y_t} \frac{y_t}{x_t}$
- Hicksian coefficient of complementarity between x and y in λ: For $\{x, y\} \in \{\theta_t, r_{t-1}, l_t\}$

$$\rho_{xy} = \frac{\frac{\partial^2 \lambda}{\partial x \partial y} \lambda}{\frac{\partial \lambda}{\partial x} \frac{\partial \lambda}{\partial y}}$$

- $\lambda_t(r, l, \theta) = rl\theta \rightarrow \rho_{\theta l} = \rho_{\theta r} = \rho_{lr} = 1$.
- $\lambda_t(r, l, \theta) = r + l + \theta \rightarrow \rho_{\theta l} = \rho_{\theta r} = \rho_{lr} = 0$.

Optimal R&D Subsidy and Corporate Tax

- Optimal Corporate Tax:

$$\left(\tau(\theta^t) + \underbrace{\frac{\int \frac{\partial \pi_t(\theta^t, \bar{q}_t)}{\partial \bar{q}_t}}{\frac{\partial \pi_t(\theta^t, \bar{q}_t)}{\partial \lambda_t}}}_{\text{Externality}}\right) \frac{1}{1-\tau(\theta^t)} = \underbrace{\frac{1-F^1(\theta_1)}{f^1(\theta_1)} p^{t-1}}_{\text{Type distribution and persistence}} \underbrace{\frac{\varepsilon_{\lambda\theta,t}}{\theta_t} \left(\frac{1+\rho_{\theta I,t}\varepsilon_{\lambda(1-\tau),t}}{\varepsilon_{\lambda(1-\tau),t}}\right)}_{\text{Efficiency Cost}}$$

- Optimal R&D Subsidy: tension between externality and asymmetric info.

$$s(\theta^t) = \frac{1}{M_t'} \frac{1}{R} \mathbb{E}(\underbrace{\frac{\partial \lambda(\theta^{t+1})}{\partial r_t} \int \frac{\partial \pi_{t+1}(\theta^{t+1}, \bar{q}_{t+1})}{\partial \bar{q}_{t+1}}}_{\text{Externality}}$$

$$+ \underbrace{\frac{1-F^1(\theta_1)}{f^1(\theta_1)} p^t}_{\text{Type distribution and persistence}} (1-\tau(\theta^{t+1})) \frac{\partial \pi_{t+1}(\theta^t, \bar{q}_{t+1})}{\partial \lambda_{t+1}} \frac{\partial \lambda_{t+1}}{\partial \theta_{t+1}} \frac{\partial \lambda_{t+1}}{\partial r_t} \frac{1}{\lambda_{t+1}} \underbrace{(\rho_{Ir} - \rho_{\theta r}))}_{\text{Complementarity coefficients}}$$

Special Cases for the Optimal R&D Subsidy when only Externality Matters

- Multiplicatively separable step size: $\lambda_t(r, l, \theta) = h^1(r)h^2(\theta)h^3(l)$ has $\rho_{lr} = \rho_{\theta l} = \rho_{\theta r} = 1$.

$$s(\theta^t) = \frac{1}{M_t'(r(\theta^t))}\frac{1}{R}\mathbb{E}\left(\frac{\partial h^1(r(\theta^t))}{\partial r_t}h^2(\theta_{t+1})h^3(l(\theta^{t+1}))\int \frac{\partial \pi_{t+1}(\theta^{t+1}, \bar{q}_{t+1})}{\partial \bar{q}_{t+1}}\right)$$

- CES step size: $\lambda(l, r, \theta) = (\alpha_1 l^{(1-\rho)} + \alpha_2 r^{(1-\rho)} + \alpha_3 \theta^{(1-\rho)})^{\frac{1}{1-\rho}}$ has $\rho_{lr} = \rho_{\theta l} = \rho_{\theta r} = \rho$.

$$s(\theta^t) = \frac{1}{M_t'(r(\theta^t))}\frac{1}{R}\mathbb{E}\left(\alpha_2\left(\frac{\lambda(\theta^t)}{r(\theta^t)}\right)^{\rho}\int \frac{\partial \pi_{t+1}(\theta^{t+1}, \bar{q}_{t+1})}{\partial \bar{q}_{t+1}}\right)$$

A Simple Numerical Illustration

- Next step: Estimation from Patent data matched to Compustat.

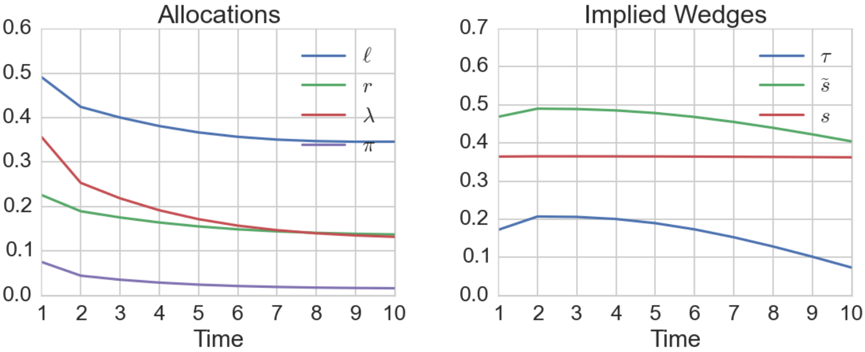
- Functional forms: $\lambda(r, l, \theta) = w(\theta, r) = (\alpha r^{(1-\rho)} + (1-\alpha)\theta^{(1-\rho)})^{\frac{1}{1-\rho}} l$.
  - Two cases: $\rho = 0.8$ and $\rho = 1.2$ (keep average λ constant across two cases).
  - $\alpha = 0.2$.

- Constant elasticity disutility: $\phi(l) = \frac{l^{1+\gamma}}{1+\gamma}$.
  - $\gamma = 2$ generates Frisch elasticity of 0.5.

- Constant elasticity R&D cost: $M_t(r) = \frac{r^{1+\kappa}}{1+\kappa}$
  - $\kappa = 1$ (typically quadratic R&D cost estimates).

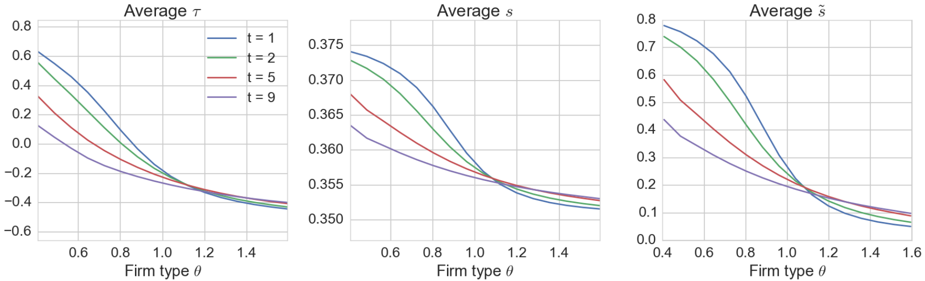
- AR(1) process for type: $\theta_t = p\theta_{t-1} + \psi_t$
  - $p = 0.9$, $\sigma_\psi^2 = 0.01$.

Policies for Old and New Firms, $\rho = 0.8$



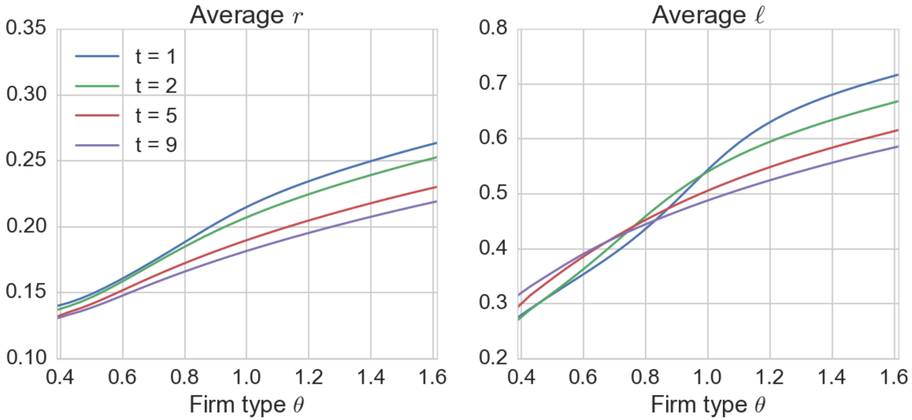
- Can incentivize less and less effort and R&D over time.
- Optimal to make young firms invest most.
- An almost constant net subsidy is optimal (but as $T \to \infty$, $s_t \to 0$.)

Wedges for Higher and Lower Quality Firms, $\rho = 0.8$



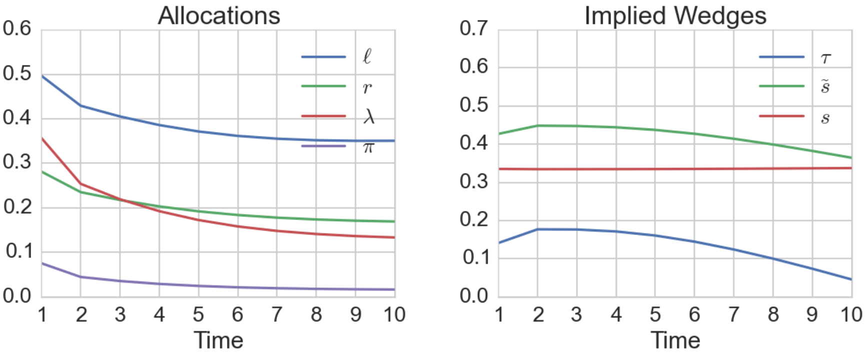
- Gross subsidy tracks corporate tax (else too high disincentive on R&D).
- Higher quality firms are subsidized to produce more (big externality).
- Optimal subsidy highly nonlinear.

Policies for Higher and Lower Quality Firms, $\rho = 0.8$



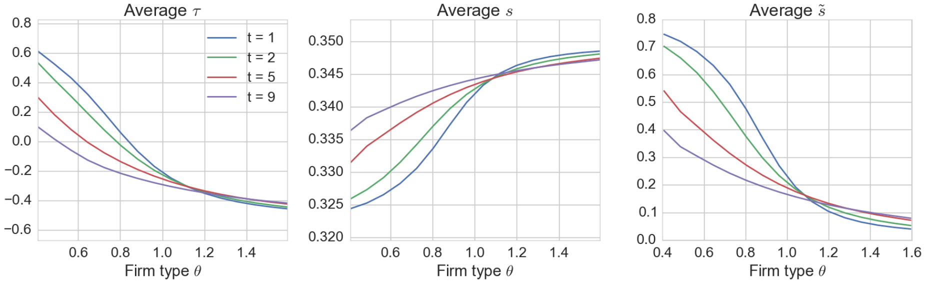
- Monotonic allocations: high quality firms have more inputs.
- Allocations get flatter over time: less able to screen.

Policies for Old and New Firms, $\rho = 1.2$



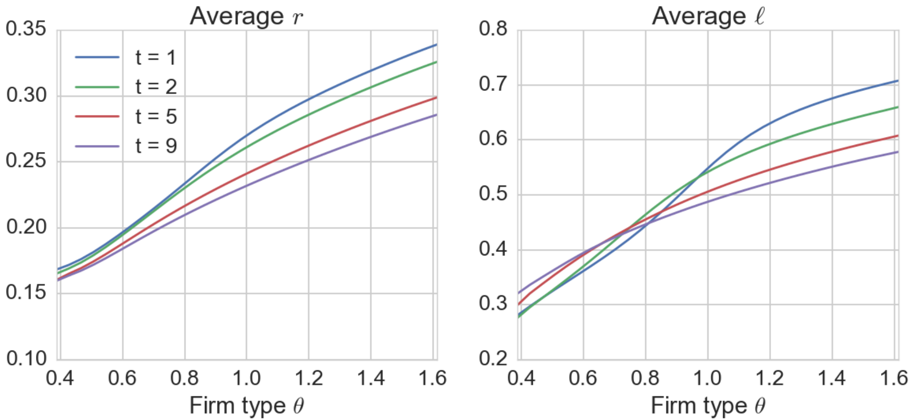
- As ρ is higher, s is lower.

Wedges for Higher and Lower Quality Firms, $\rho = 1.2$



- As ρ is higher, s is lower.

Policies for Higher and Lower Quality Firms, $\rho = 1.2$



Case with no Depreciation: Externality Persists Forever

$$\left(\tau(\theta^t) + \frac{\sum_{s=t}^{T} \int \frac{\partial \pi_s(\theta^s, \bar{q}_s)}{\partial \bar{q}_s}}{E(\sum_{s=t}^{T} (\frac{1}{R})^{s-t} \frac{\partial \pi_s(\theta^s, \bar{q}_s)}{\partial \lambda_t} \frac{\partial \lambda_t}{\partial l_t})}\right) \frac{1}{1-\tau(\theta^t)} =$$

$$\frac{1-F^1(\theta_1)}{f^1(\theta_1)} p^{t-1} \left(\frac{\varepsilon_{\lambda\theta,t}}{\theta_t}\right) \left(\frac{1+\rho_{\theta l,t}\varepsilon_{\lambda(1-\tau),t}}{\varepsilon_{\lambda(1-\tau),t}}\right)$$

$$s(\theta^t) = \frac{1}{M_t'(r(\theta^t))} \frac{1}{R} \mathbb{E}(\frac{\partial \lambda_{t+1}}{\partial r_t} \sum_{s=t+1}^{T} \int \frac{\partial \pi_s(\theta^s, \bar{q}_s)}{\partial \bar{q}_s} +$$

$$\frac{1-F^1(\theta_1)}{f^1(\theta_1)} p^t \frac{\partial \lambda_{t+1}}{\partial r_t} \frac{\partial \lambda_{t+1}}{\partial \theta_{t+1}} \frac{\phi'(l_{t+1})}{\lambda_{t+1}} \frac{1}{\frac{\partial \lambda_{t+1}}{\partial l_{t+1}}} (\rho_{lr} - \rho_{\theta r}))$$

- Same special cases apply when subsidy only set based on (now persistent) externality.

Outline

Optimal Linear (Size-independent) Age-Dependent Policies

- Linear age-dependent corporate tax τ_t and R&D subsidy s_t.

- Manager's optimization problem:

$$\max_{\{s_{ti}, r_{ti}\}_{t=1}^{T}} \sum_{t=1}^{T} \left(\frac{1}{R}\right)^{t-1} \left((1-\tau_t)\pi_{ti} - (1-s_t)M_t(r_{ti}) - \phi_t(l_{ti})\right)$$

- Hence $\pi_{ti}((1-\tau_t), s_{t-1}, \bar{q}_t((1-\tau_t), s_{t-1})))$ and $r_{t-1i}((1-\tau_t), s_{t-1}, \bar{q}_t((1-\tau_t), s_{t-1})))$ are solutions as functions of policies.

- Social welfare is now just total revenue:

$$\max_{\{\tau, s\}_{t=1}^{T}} \sum_{t=1}^{T} (\frac{1}{R})^{t-1} (\tau_t \mathbb{E}(\pi_{ti}) - s_t \mathbb{E}(M_t(r_{ti})))$$

- Aggregate (or cross-sectional averages): $r_t = \mathbb{E}(r_{ti})$ and $\pi_t = \mathbb{E}(\pi_{ti})$.

Optimal Formulas as Function of Sufficient Statistics

- Formulas valid even if other instrument not optimally set.
- Useful to evaluate "reforms."
- Could use structural or reduced-form (sufficient stats) approach.
- Fiscal externality from one tax on other tax' base.

$$\tau_t^* = \frac{1 + s_{t-1}\varepsilon_{r,t-1}^{1-\tau}\frac{r_{t-1}}{\pi_t}}{1+\varepsilon_{\pi,t}^{1-\tau}} \qquad s_t^* = \frac{1+\tau_t\varepsilon_{\pi,t}^{s-1}\frac{\pi_t}{r_t}}{1+\varepsilon_{r,t}^{s-1}}$$

with

$$\varepsilon_{r,t-1}^{1-\tau} = \frac{d\mathbb{E}(M(r_{t-1i}))}{d(1-\tau)}\frac{1-\tau}{r_{t-1}} \qquad \varepsilon_{\pi,t}^{1-\tau} = \frac{d\pi_t}{d(1-\tau)}\frac{1-\tau}{\pi_t}$$

$$\varepsilon_{\pi,t}^{s-1} = \frac{d\pi_t}{d(s-1)}\frac{s-1}{\pi_t} \qquad \varepsilon_{r,t}^{s-1} = \frac{d\mathbb{E}(M(r_{ti}))}{d(s-1)}\frac{s-1}{r_t}$$

Optimal Size and Age-Independent Policies

- Same Formula and same sufficient stats – different estimation required in data.

$$\tau^* = \frac{1 + s\varepsilon_r^{1-\tau}\frac{r}{\pi}}{1+\varepsilon_\pi^{1-\tau}} \qquad s^* = \frac{1+\tau\varepsilon_\pi^{s-1}\frac{\pi}{r}}{1+\varepsilon_r^{s-1}}$$

$$\pi = \sum_{t=1}^{T}\left(\frac{1}{R}\right)^{t-1}\pi_t \qquad r = \sum_{t=1}^{T}\left(\frac{1}{R}\right)^{t-1} r_{t-1}$$

$$\varepsilon_r^{1-\tau} = \sum_{t=1}^{T}\left(\frac{1}{R}\right)^{t-1}\varepsilon_{r,t-1}^{1-\tau}\frac{r_{t-1}}{r} \qquad \varepsilon_\pi^{1-\tau} = \sum_{t=1}^{T}\left(\frac{1}{R}\right)^{t-1}\varepsilon_{\pi,t}^{1-\tau}\frac{\pi_t}{\pi}$$

$$\varepsilon_\pi^{s-1} = \sum_{t=1}^{T}\left(\frac{1}{R}\right)^{t-1}\varepsilon_{\pi,t}^{s-1}\frac{\pi_t}{\pi} \qquad \varepsilon_r^{s-1} = \sum_{t=1}^{T}\left(\frac{1}{R}\right)^{t-1}\varepsilon_{r,t}^{s-1}\frac{r_t}{r}$$

Outline

Conclusion

- Model of R&D investments and endogenous growth with heterogeneous firms, private information, and spillovers.
  - Use mechanism design to solve for constrained efficient allocations.
  - Characterize using "wedges."
- Optimal to subsidize R&D investments because of externality.
- Asymmetric info: could tend to reduce R&D subsidy if very complementary to firm quality.
  - But optimal R&D subsidy is nonlinear.
- Simple sufficient stats formula for R&D subsidy and corporate tax in linear, age-dependent or age-independent cases.
- Next step: estimate model using patent + compustat data.