Optimal Taxation and R&D Policies

Ufuk Akcigit — Chicago

Douglas Hanley — Pittsburgh

Stefanie Stantcheva — Harvard

November 2016

Motivation: Widespread and Diverse R&D Policies

"The need to foster greater innovation and productivity growth is one of the most important economic challenges we face, and tax policy is one of several important levers that policymakers can use", Jason Furman, chairman of CEA

Businesses spend a lot of resources on R&D... and the government already intervenes heavily.

Large **variety** of policies target **innovation and R&D**

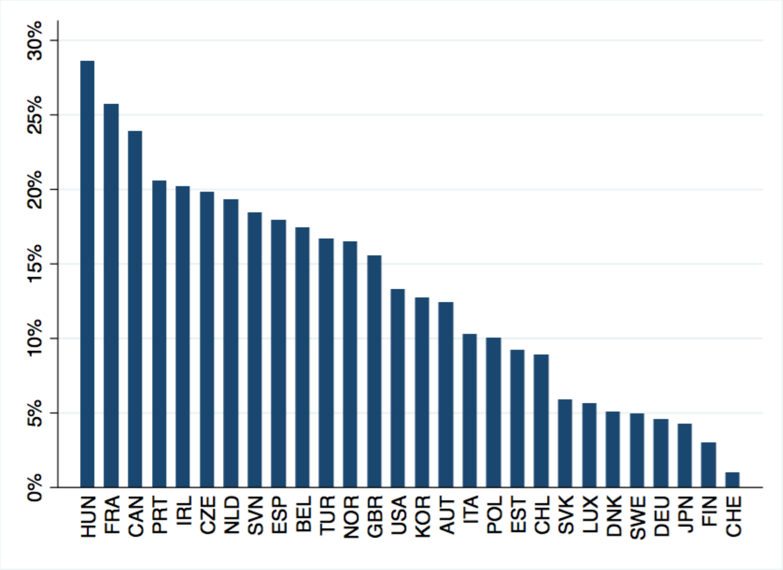
Tax credits, deductions, grants, contracts, direct funding in FFRDCs, Universities, Firms, small business, start-ups..

Large variety of policies across countries as well.

R&D policies are widespread, not fully understood, & very costly:

- "Intramural" R&D cost $35 billion (2014).
- "Extramural" R&D: tax credit $11 bil in 2012, contracting with non FFRDCs 50,6 billion, NSF-NIH $40 billion (econ grant: 0.0025%)

Share of Government Funding in Business R&D



Amount spent by government correlated with better productivity?

This Paper: Optimal Design of R&D Policy and Firm Taxation

- Firms have **heterogeneous**, stochastic productivities.

  Productivity: efficiency of converting R&D inputs into innovation output.

- **Uncertainty** about R&D returns.
- **Spillovers** between firms: one firm's innovations affect other firms.
- Innovation not appropriable unless IPR.
- Firm productivity is **private information** (why important?)
- 1) Mechanism design: no a priori restriction on policy tools.

  Characterize constrained efficient allocations.

  Implementation.

- 2) Quantitative Investigation using patent + Compustat data.
- 3) Losses from "simpler" policies (e.g.: linear age-dependent...).

Related Literature

R&D Policies and Growth: Leahy and Neary (1997), Acemoglu, Akcigit, Bloom and Kerr (2013), Akcigit, Hanley and Serrano-Velarde (2013), Atkeson and Burstein (2014).

Optimal Policy Design: Pavan, Segal and Toikka (2013), Doepke and Townsend (2006), Stantcheva (2012, 2014), Golosov, Tsyvinski and Werning (2006), Farhi and Werning (2013, 2014).

Heterogeneity in Management: Bloom et al. (2013), Bloom, Sadun and Van Reenen (2012).

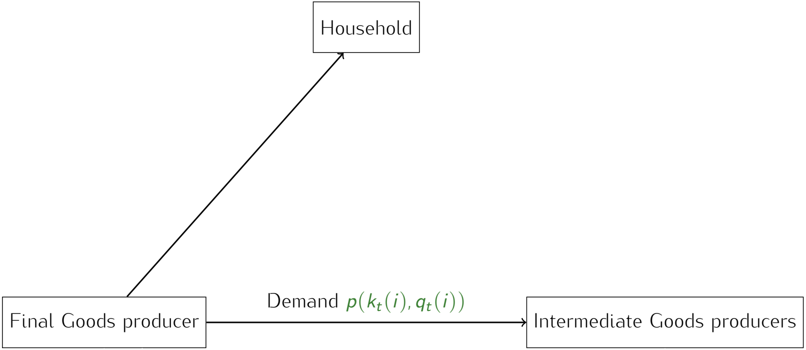
Empirics of R&D Policies: Bloom, Griffith, Van Reenen (2002), Hall and Van Reenen (2000), Bloom, Schankerman and Van Reenen (2013).

IPR Design: Scotchmer (1999), Kremer (1998), Chari, Golosov and Tsyvinski (2012).

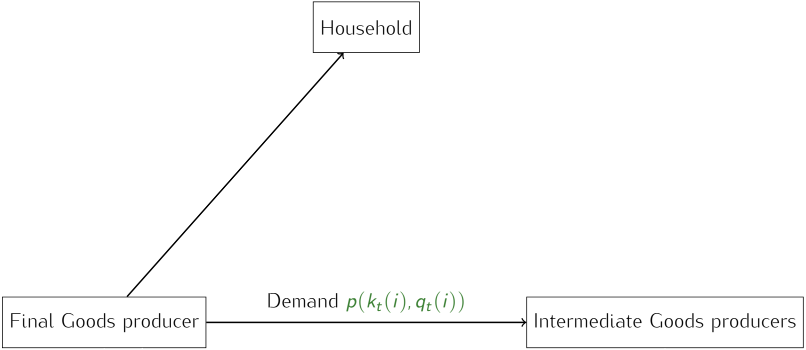
Outline

Model

Intermediate Goods producers



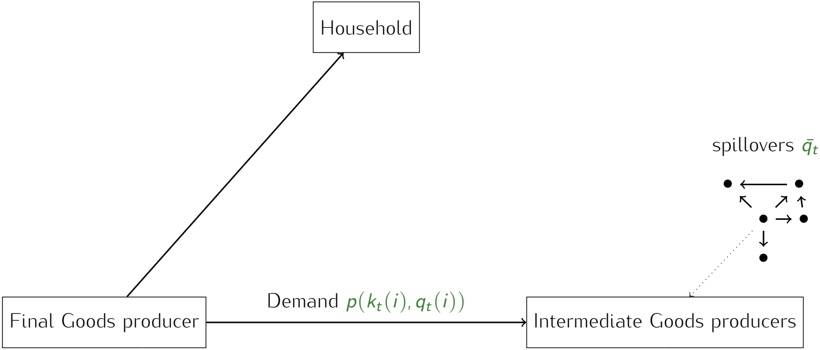
$$Y_t = \int_i Y(q_t(i), k_t(i))di$$



$Y_t = \int_i Y(q_t(i), k_t(i))di$

Production

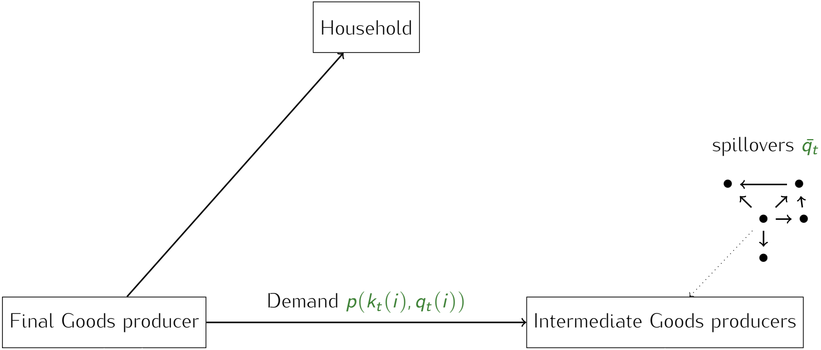
- Quality $q_t(i)$, quantity $k_t(i)$
- Demand: $p(k_t(i), q_t(i))$



$Y_t = \int_i Y(q_t(i), k_t(i))di$

Production

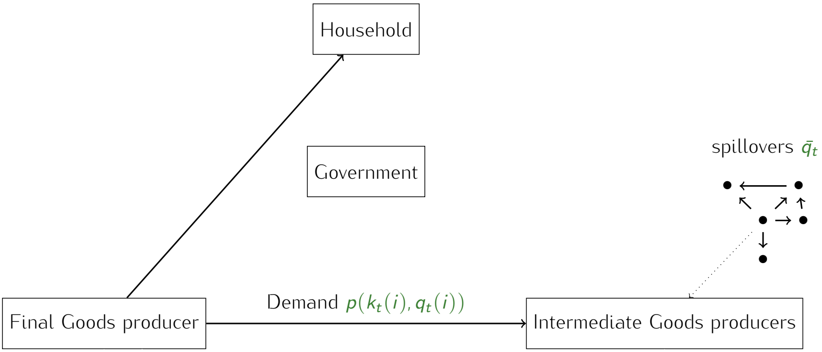
- Quality $q_t(i)$, quantity $k_t(i)$
- Demand: $p(k_t(i), q_t(i))$
- Spillover: aggregate quality: $\bar{q}_t = \int_i q_t(i)di$



$Y_t = \int_i \; Y(q_t(i), k_t(i)) di$

Production

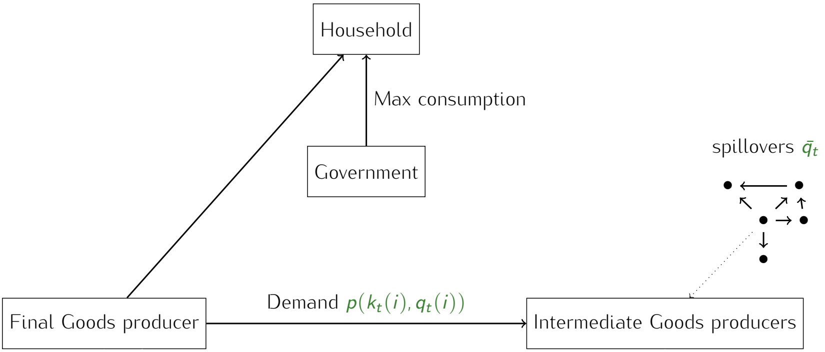
- Quality $q_t(i)$, quantity $k_t(i)$
- Demand: $p(k_t(i), q_t(i))$
- Spillover: aggregate quality: $\bar{q}_t = \int_i q_t(i) di$
- $\pi(q_t(i), \bar{q}_t) = \max_k \{p(k, q_t(i))k - C(k, \bar{q}_t)\}$



$Y_t = \int_i \; Y(q_t(i), k_t(i)) di$

Production

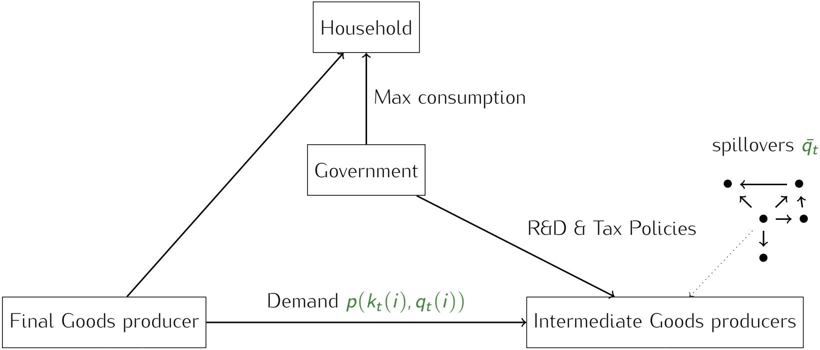
- Quality $q_t(i)$, quantity $k_t(i)$
- Demand: $p(k_t(i), q_t(i))$
- Spillover: aggregate quality: $\bar{q}_t = \int_i q_t(i) di$
- $\pi(q_t(i), \bar{q}_t) = \max_k \{p(k, q_t(i))k - C(k, \bar{q}_t)\}$



$$Y_t = \int_i Y(q_t(i), k_t(i))di$$

Production

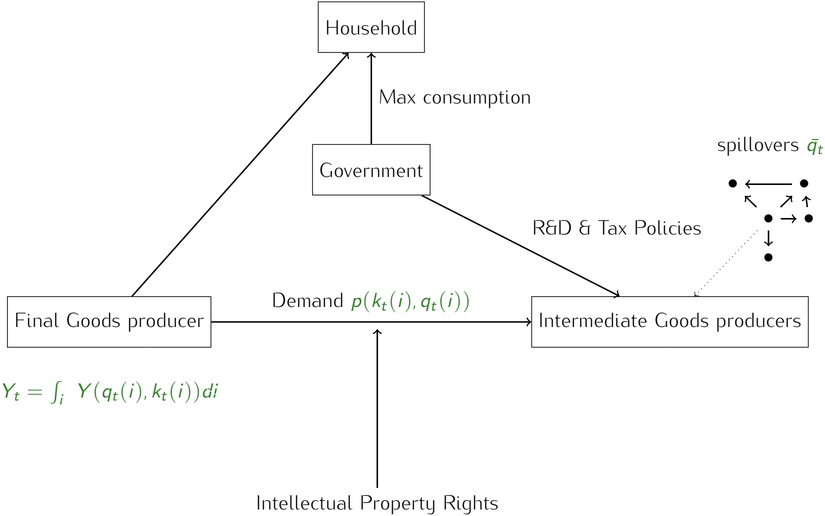
- Quality $q_t(i)$, quantity $k_t(i)$
- Demand: $p(k_t(i), q_t(i))$
- Spillover: aggregate quality: $\bar{q}_t = \int_i q_t(i)di$
- $\pi(q_t(i), \bar{q}_t) = \max_k \{p(k, q_t(i))k - C(k, \bar{q}_t)\}$

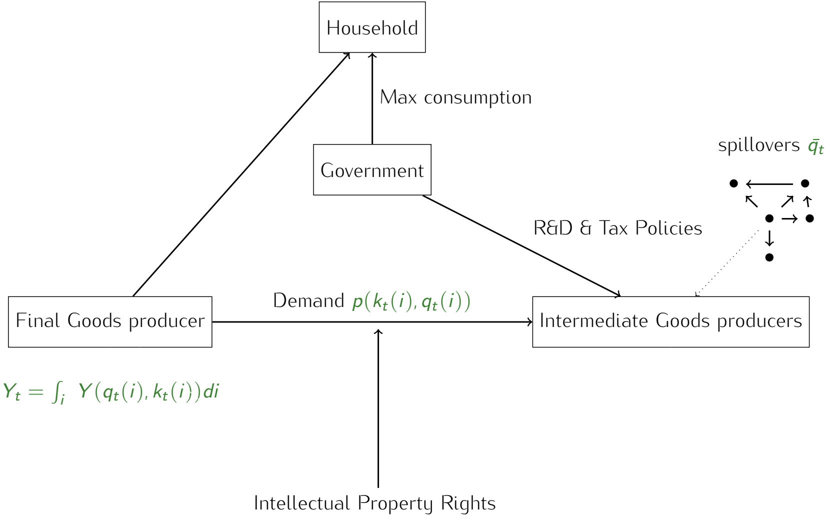


$Y_t = \int_i Y(q_t(i), k_t(i))di$

Production

- Quality $q_t(i)$, quantity $k_t(i)$
- Demand: $p(k_t(i), q_t(i))$
- Spillover: aggregate quality: $\bar{q}_t = \int_i q_t(i)di$
- $\pi(q_t(i), \bar{q}_t) = \max_k \{p(k, q_t(i))k - C(k, \bar{q}_t)\}$


Household
Max consumption
Government
spillovers $\bar{q}_t$
R&D & Tax Policies
Final Goods producer
Demand $p(k_t(i), q_t(i))$
Intermediate Goods producers
$Y_t = \int_i Y(q_t(i), k_t(i))di$
Intellectual Property Rights



- Patent: $p(k_t(i), q_t(i)) = \frac{\partial Y(q_t(i), k_t(i))}{\partial k_t(i)}$
- "Prize": $p(k_t(i), q_t(i)) = \frac{Y(q_t(i), k_t(i))}{k_t(i)}$

Intermediate Producers' static production decisions

- Intermediate good producers: quality $q_t(i)$, quantity $k_t(i)$.
- Final good producer aggregates intermediate goods for consumption:

$$Y_t = \int_i Y(q_t(i), k_t(i))di$$

- Demand function for good i: $p(k_t(i), q_t(i))$, depends on intellectual property rights policy.

  Monopoly price: $p(k_t(i), q_t(i)) = \frac{\partial Y(q_t(i), k_t(i))}{\partial k_t(i)}$.

  "Prize" mechanism: $p(k_t(i), q_t(i)) = \frac{Y(q_t(i), k_t(i))}{k_t(i)}$

- Technology spillovers: come from aggregate quality: $\bar{q}_t = \int_i q_t(i)di$
- Cost of production: $C(k_t, \bar{q}_t)$ ($\uparrow$ or $\downarrow$ in $\bar{q}_t$).
- Profit maximization: $\pi(q_t, \bar{q}_t) \equiv \max_k \{p(k, q_t)k - C(k, \bar{q}_t)\}$

Intermediate Producers' Innovation Decisions

- Firms can improve their product quality q_t through R&D and effort: $q_t = H(q_{t-1}, \lambda_t)$.
- The step size $\lambda_t(r_{t-1}, l_t, \theta_t)$ depends on:

  R&D investment r_t cost $M_t(r_t)$.

  R&D effort (unobservable R&D input): l_t at cost $\phi_t(l_t)$.

  Productivity θ_t (managerial/firm quality), Markov $f^t(\theta_t|\theta_{t-1})$, history θ^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.

Market Failures and First Best Allocation

1) Lack of appropriability of innovation (need intellectual property rights (IPR)).

2) Technology spillovers.

First best quantity conditional on quality: $k^*(q_t(\theta^t), \bar{q}_t)$.

First best output net of production costs:
$\tilde{Y}^*(q_t(\theta^t), \bar{q}_t) = Y(q_t(\theta^t), k^*(q_t(\theta^t), \bar{q}_t)) - C(k^*(q_t(\theta^t), \bar{q}_t), \bar{q}_t)$.

Optimality: marginal cost = marginal social benefit

$$M_t'(r_t(\theta^t)) = \frac{1}{R}\mathbb{E}\left(\sum_{s=t+1}^{T}\left(\frac{1-\delta}{R}\right)^{s-t-1}\left(\frac{\partial \tilde{Y}^*(q_t(\theta^s), \bar{q}_t)}{\partial q_s} + \frac{\partial \tilde{Y}^*(q_t(\theta^s), \bar{q}_t)}{\partial \bar{q}_s}\right)\frac{\partial \lambda_{t+1}}{\partial r_t}\right)$$

$$\phi_t'(l_t(\theta^t)) = \mathbb{E}\left(\sum_{s=t}^{T}\left(\frac{1-\delta}{R}\right)^{s-t}\left(\frac{\partial \tilde{Y}^*(q_t(\theta^s), \bar{q}_t)}{\partial q_s} + \frac{\partial \tilde{Y}^*(q_t(\theta^s), \bar{q}_t)}{\partial \bar{q}_s}\right)\right)\frac{\partial \lambda_t(\theta^t)}{\partial l_t(\theta^t)}$$

Asymmetric Information

- Consider several cases.

  (1) Firm productivity θ_t and R&D effort not observable...

  (2) ... and quantity k_t not observable.

  (3) ... and innovation quality λ_t (or q_t) unobservable (restricted tools).

- Case (1) $\Leftrightarrow$ can optimize on intellectual property rights policy.

  Optimal IPR trivial here: prize system or patent system + price subsidy.

- Case (2) $\Leftrightarrow$ take IPR as given (partial optimum), e.g.: patents.

- Asymmetric info problem:

  If heterogeneous, but observable types: heterogeneous policies, type-specific lump-sum tax.

  Asymmetric info: cannot extract surplus lump-sum.

  Problem if limited liability and redistribution/revenue requirement.

Coefficient of Complementarity

- Hicksian coefficient of complementarity between x and y:

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

- $\lambda_t(r, I, \theta) = rI\theta \rightarrow \rho_{\theta I} = \rho_{\theta r} = \rho_{Ir} = 1$.
- $\lambda_t(r, I, \theta) = r + I + \theta \rightarrow \rho_{\theta I} = \rho_{\theta r} = \rho_{Ir} = 0$.
- When $\rho_{\theta r}$ is large: Higher R&D investments increase informational rent that needs to be transferred to high quality firms.

Firm Life Cycle, Competition and Patent Policy

- Finite firm lifecycle T (can be large or ∞). "Policy commitment horizon." Terminal value?
- Additional observable heterogeneity (sector? product type?) can be conditioned on.
- Entry and exit: intensive margin here, even if out of market, firm's q contributes to $\bar{q}$, market turnover captured by R_t.
- Competition: exogenous markups.
  - Captured reduced form by (i) cost functions (input market competition) and (ii) substitutability between goods, affects pricing power.
  - Can do comparative statics on R&D policies with respect to competition.
- Different types of innovations: new vs existing product, process vs. product. Common core we focus on: **spillovers**.
- Intellectual Protection Policy: different from R&D policy, but affects it.

Illustration in Toy Model

Illustration in 2-type, 1-period Model

Productivity $\theta_2 > \theta_1$ (fractions f_2 and f_1, with $f_2 = 1 - f_1$).

Quality is $q(\theta_i) = q_0 + \lambda(\theta_i)$, q_0 given.

Step size $\lambda(r, l, \theta_i) = w(r, \theta_i)l$, with increasing and concave w.

Aggregate quality $\bar{q} = f_1 q(\theta_1) + f_2 q(\theta_2)$.

Profits: $\pi(q, \bar{q})$ (depend on market structure, IPR).

Menu of contracts: $(r(\theta_i), l(\theta_i), k(\theta_i), T(\theta_i))$ for $i = 1, 2$.

Maximize household consumption:

$$W = f_1 \left(Y(q(\theta_1), k(\theta_1)) - C(k(\theta_1), \bar{q}) - M(r(\theta_1)) - T(\theta_1) \right) + f_2 \left(Y(q(\theta_2), k(\theta_2)) - C(k(\theta_2), \bar{q}) - M(r(\theta_2)) - T(\theta_2) \right).$$

Incentive Constraints

First best: efficient quantity, marginal cost of quality = marginal social benefit, extract all surplus lump-sum $T(\theta_i) = \phi(I(\theta_i))$.

High productivity firm wants to mimic low productivity firm at FB.

Incentive constraints for $j = 1, 2$:

$$T(\theta_i) - \phi(I(\theta_i)) \geq T(\theta_j) - \phi\left(\frac{w(r(\theta_j), \theta_j) I(\theta_j)}{w(r(\theta_j), \theta_i)}\right)$$

Binding PC for low productivity, binding IC for high productivity.

$$T(\theta_1) = \phi\left(I(\theta_1)\right)$$

$$T(\theta_2) - \phi\left(I(\theta_2)\right) \geq T(\theta_1) - \phi\left(\frac{w(r(\theta_1), \theta_1) I(\theta_1)}{w(r(\theta_1), \theta_2)}\right)$$

Second-Best Program When Quantity Can Be Controlled

Conditional on quality $q(\theta_i)$ and $\bar{q}$, quantity set to max consumption net of production costs: $\tilde{Y}^*(q(\theta_i), \bar{q}) := \max_k \left(Y(k, q(\theta_i)) - C(k, \bar{q})\right)$.

Social objective (virtual surplus):

$$W = f_1 \left(\tilde{Y}^*(q_1(\theta_1), \bar{q}_1) - M(r(\theta_1)) - \phi\left(l(\theta_1)\right)\right)$$
$$+ f_2(\tilde{Y}^*(q(\theta_2), \bar{q}) - M(r(\theta_2)) - \phi\left(l(\theta_2)\right)) - f_2 \left(\phi\left(l(\theta_1)\right) - \phi\left(\frac{w(r(\theta_1), \theta_1) l(\theta_1)}{w(r(\theta_1), \theta_2)}\right)\right)$$

Define profit (R&D effort) wedge τ and R&D investment wedge s:

$$s(\theta_i) = M'(r(\theta_i)) - \frac{\partial \pi(q(\theta_i), \bar{q})}{\partial q(\theta_i)} \frac{\partial \lambda(r(\theta_i), l(\theta_i), \theta_i)}{\partial r(\theta_i)}$$

$$(1 - \tau(\theta_i)) \frac{\partial \pi(q(\theta_i), \bar{q})}{\partial q(\theta_i)} \frac{\partial \lambda(r(\theta_i), l(\theta_i), \theta_i)}{\partial l(\theta_i)} = \phi'(l(\theta_i))$$

Optimal Unrestricted Mechanism

A Direct Revelation Mechanism with Spillovers

Firm reports $\theta'_t(\theta^t)$. History of reports: $\theta'^t = \{\theta'_1(\theta_1), ...\theta'_t(\theta^t)\}$.

Allocations for history of reports: $\{\lambda(\theta'^t), r(\theta'^t), T_t(\theta'^t)\}$ (possibly, $k(\theta'^t)$).

Maximize household consumption:

$$\mathbb{E}\left\{\sum_{t=1}^{T}\left(\frac{1}{R}\right)^{t-1}\left\{Y(k_t(\theta^t), q_t(\theta^t)) - C(k_t(\theta^t), \bar{q}_t) - M_t(r(\theta^t)) - T_t(\theta^t)\right\}\right\}$$

Partial problem $P(\bar{q})$:

Fix sequence of $\bar{q} \equiv \{\bar{q}_t\}_t$, solve screening problem subject to consistency of agents' choices with $\bar{q}$.

Full problem P:

P: $max_{\bar{q}} P(\bar{q})$.

Incentive Compatibility and a First-order Approach

- Expected continuation utility of firm after history θ^t:
$V_t(\theta^t) = \sum_{t=s}^{T} (\frac{1}{R})^{t-s} \cdot \left\{ \int_{\Theta^t} \left\{ T_t(\theta^t) - \phi_t(l_t(\theta^t)) \right\} P(\theta^t|\theta^s) d\theta^t \right\}$

- Lifetime utility for a given sequence of realizations θ^T: $\tilde{U}(\theta^T)$

- 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}.
  - Relies on first-order condition (sufficiency?)

- $V_1(\theta_1) = V_1(\underline{\theta}_1) + \int_{\underline{\theta}_1}^{\theta_1} \frac{\partial V_1(\theta)}{\partial \theta} d\theta$.

- Expected PDV of transfers = expected PDV of disutility costs + info rent ($V_1(\theta_1)$).

Program: Virtual Surplus with Spillovers

If quantity can be controlled, set to maximize output net of production costs:

$$\tilde{Y}^*(q_t(\theta^t), \bar{q}_t) = \max_k \{Y(k, q_t(\theta^t)) - C(k, \bar{q}_t)\}$$

$$P(\bar{q}) = \max W(\bar{q}) = \mathbb{E}\{\sum_{t=1}^{T} \left(\frac{1}{R}\right)^{t-1} \{\tilde{Y}^*(q_t(\theta^t), \bar{q}_t) - M_t(r_t(\theta^t)) - \phi_t(I_t(\theta^t))$$

$$- V_1(\underline{\theta_1}) - \frac{1 - F^1(\theta_1)}{f^1(\theta_1)} I_{1,t} \frac{\partial \tilde{U}_t}{\partial \theta_t}\}\}$$

$$\text{s.t.:} \quad \int_{\Theta^t} q_t(\theta^t) P(\theta^t) d\theta^t = \bar{q}_t \qquad [\eta_t]$$

$$\text{and} \quad q_t(\theta^t) = q_{t-1}(\theta^{t-1})(1-\delta) + \lambda_t(I_t(\theta^t), r_{t-1}(\theta^{t-1}), \theta_t)$$

Program: Virtual Surplus with Spillovers

If quantity can not be controlled, set by firm to maximize profits:

$$\tilde{Y}(q_t(\theta^t), \bar{q}_t) = \max_k \{p(k, q_t(\theta^t))k - C(k, \bar{q}_t)\}$$

$$P(\bar{q}) = \max W(\bar{q}) = \mathbb{E}\{\sum_{t=1}^{T} \left(\frac{1}{R}\right)^{t-1} \{\tilde{Y}(q_t(\theta^t), \bar{q}_t) - M_t(r_t(\theta^t)) - \phi_t(I_t(\theta^t)) - V_1(\underline{\theta_1}) - \frac{1 - F^1(\theta_1)}{f^1(\theta_1)} I_{1,t} \frac{\partial \tilde{U}_t}{\partial \theta_t}\}\}$$

$$\text{s.t.:} \quad \int_{\Theta^t} q_t(\theta^t) P(\theta^t) d\theta^t = \bar{q}_t \qquad [\eta_t]$$

$$\text{and} \quad q_t(\theta^t) = q_{t-1}(\theta^{t-1})(1 - \delta) + \lambda_t(I_t(\theta^t), r_{t-1}(\theta^{t-1}), \theta_t)$$

Wedges: Measures of Distortions in the Allocations

Akin to "implicit" taxes and subsidies.

$$\tau(\theta^t) \equiv \underbrace{\mathbb{E}\left(\sum_{s=t}^{T}\left(\frac{1-\delta}{R}\right)^{s-t}\frac{\partial\pi_s(q_s(\theta^s),\bar{q}_s)}{\partial q_s(\theta^s)}\frac{\partial\lambda_t(\theta^t)}{\partial l_t(\theta^t)}\right)}_{\text{Marginal benefit}} - \underbrace{\phi'(l_t(\theta^t))}_{\text{Marginal cost}}$$

$$s(\theta^t) \equiv \underbrace{M_t'(r(\theta^t))}_{\text{Marginal cost}} - \underbrace{\frac{1}{R}\mathbb{E}\left(\sum_{s=t+1}^{T}\left(\frac{1-\delta}{R}\right)^{s-t-1}\frac{\partial\pi_s(q_s(\theta^s),\bar{q}_s)}{\partial q_s(\theta^s)}\frac{\partial\lambda_{t+1}(\theta^{t+1})}{\partial r_t(\theta^t)}\right)}_{\text{Marginal benefit}}$$

Defined relative to laissez-faire with some profits $\pi_s(q_s(\theta^s), \bar{q}_s)$ (e.g.: patent protection).

Introduce Some Notation

$\Pi_t(\theta^t) \equiv \frac{1}{R}\left(\sum_{s=t}^{T}\left(\frac{1-\delta}{R}\right)^{s-t}\frac{\partial\pi(q(\theta^s),\bar{q}_s)}{\partial q_s}\right)$ (impact of q_t on profit stream).

$Q_t^*(\theta^t) \equiv \frac{1}{R}\left(\sum_{s=t}^{T}\left(\frac{1-\delta}{R}\right)^{s-t}\frac{\partial\tilde{Y}^*(q(\theta^s),\bar{q}_s)}{\partial q_s}\right)$ (impact on social surplus, if quantity controlled).

$Q_t(\theta^t) \equiv \frac{1}{R}\left(\sum_{s=t}^{T}\left(\frac{1-\delta}{R}\right)^{s-t}\frac{\partial\tilde{Y}(q(\theta^s),\bar{q}_s)}{\partial q_s}\right)$ (impact on social surplus, if quantity not controlled).

Optimal Profit wedge and R&D subsidy

$$\tau(\theta^t) = -\mathbb{E}\left(\sum_{s=t}^{T}(1-\delta)^{s-t}\eta_s\right)\frac{\partial\lambda_t(\theta^t)}{\partial l_t} - \mathbb{E}\left(Q_t^* - \Pi_t\right)\frac{\partial\lambda_t(\theta^t)}{\partial l_t}$$

$$+ \frac{1-F^1(\theta_1)}{f^1(\theta_1)} l_{1,t}(\theta^t)\frac{\phi_t'\lambda_{\theta t}}{\lambda_t}\left[\frac{1}{\varepsilon_{l,1-\tau}}\frac{1}{\varepsilon_{\lambda l,t}} + \rho_{\theta l,t}\right]$$

$$s(\theta^t) = \mathbb{E}\left(\sum_{s=t+1}^{T}(1-\delta)^{s-t-1}\eta_s\frac{\partial\lambda(\theta^{t+1})}{\partial r_t}\right) + \mathbb{E}\left(\left(Q_{t+1}^* - \Pi_{t+1}\right)\frac{\partial\lambda(\theta^{t+1})}{\partial r_t}\right)$$

$$+ \frac{1}{R}\mathbb{E}\left(\frac{1-F^1(\theta_1)}{f^1(\theta_1)} l_{1,t+1}(\theta^{t+1})\phi_{t+1}'(l(\theta^{t+1}))\frac{\lambda_\theta\lambda_r}{\lambda\lambda_l}(\rho_{lr} - \rho_{\theta r})\right)$$

Optimal Profit wedge and R&D subsidy

$$\tau(\theta^t) = -\mathbb{E}\left(\sum_{s=t}^{T}(1-\delta)^{s-t}\eta_s\right)\frac{\partial\lambda_t(\theta^t)}{\partial l_t} - \mathbb{E}\left(Q_t^* - \Pi_t\right)\frac{\partial\lambda_t(\theta^t)}{\partial l_t}$$

$$+ \frac{1-F^1(\theta_1)}{f^1(\theta_1)} l_{1,t}(\theta^t)\frac{\phi_t'\lambda_{\theta t}}{\lambda_t}\left[\frac{1}{\varepsilon_{l,1-\tau}}\frac{1}{\varepsilon_{\lambda l,t}} + \rho_{\theta l,t}\right]$$

$$s(\theta^t) = \mathbb{E}\left(\sum_{s=t+1}^{T}(1-\delta)^{s-t-1}\eta_s\frac{\partial\lambda(\theta^{t+1})}{\partial r_t}\right) + \mathbb{E}\left(\left(Q_{t+1}^* - \Pi_{t+1}\right)\frac{\partial\lambda(\theta^{t+1})}{\partial r_t}\right)$$

$$+ \frac{1}{R}\mathbb{E}\left(\frac{1-F^1(\theta_1)}{f^1(\theta_1)} l_{1,t+1}(\theta^{t+1})\phi_{t+1}'(l(\theta^{t+1}))\frac{\lambda_\theta\lambda_r}{\lambda\lambda_l}(\rho_{lr} - \rho_{\theta r})\right)$$

Pigouvian Correction:

If positive externality, subsidize profits and R&D.

Larger for high productivity firms as long as $\rho_{\theta l} > 0$ and $\rho_{\theta r} > 0$.

Optimal Profit wedge and R&D subsidy

$$\tau(\theta^t) = -\mathbb{E}\left(\sum_{s=t}^{T}(1-\delta)^{s-t}\eta_s\right)\frac{\partial\lambda_t(\theta^t)}{\partial l_t} - \mathbb{E}\left(Q_t^* - \Pi_t\right)\frac{\partial\lambda_t(\theta^t)}{\partial l_t}$$

$$+ \frac{1-F^1(\theta_1)}{f^1(\theta_1)} l_{1,t}(\theta^t)\frac{\phi_t'\lambda_{\theta t}}{\lambda_t}\left[\frac{1}{\varepsilon_{l,1-\tau}}\frac{1}{\varepsilon_{\lambda l,t}} + \rho_{\theta l,t}\right]$$

$$s(\theta^t) = \mathbb{E}\left(\sum_{s=t+1}^{T}(1-\delta)^{s-t-1}\eta_s\frac{\partial\lambda(\theta^{t+1})}{\partial r_t}\right) + \mathbb{E}\left(\left(Q_{t+1}^* - \Pi_{t+1}\right)\frac{\partial\lambda(\theta^{t+1})}{\partial r_t}\right)$$

$$+ \frac{1}{R}\mathbb{E}\left(\frac{1-F^1(\theta_1)}{f^1(\theta_1)} l_{1,t+1}(\theta^{t+1})\phi_{t+1}'(l(\theta^{t+1}))\frac{\lambda_\theta\lambda_r}{\lambda\lambda_l}(\rho_{lr} - \rho_{\theta r})\right)$$

Screening:

- Stochastic productivity process:
  - Lower persistence: lower wedges over time.
    - Special cases: iid, full persistence, AR(1).
  - Larger inverse hazard ratio: larger wedges (no distortion at the top).

Optimal Profit wedge and R&D subsidy

$$\tau(\theta^t) = -\mathbb{E}\left(\sum_{s=t}^{T}(1-\delta)^{s-t}\eta_s\right)\frac{\partial\lambda_t(\theta^t)}{\partial l_t} - \mathbb{E}\left(Q_t^* - \Pi_t\right)\frac{\partial\lambda_t(\theta^t)}{\partial l_t} + \frac{1-F^1(\theta_1)}{f^1(\theta_1)} l_{1,t}(\theta^t)\frac{\phi_t'\lambda_{\theta t}}{\lambda_t}\left[\frac{1}{\varepsilon_{l,1-\tau}}\frac{1}{\varepsilon_{\lambda l,t}} + \rho_{\theta l,t}\right]$$

$$s(\theta^t) = \mathbb{E}\left(\sum_{s=t+1}^{T}(1-\delta)^{s-t-1}\eta_s\frac{\partial\lambda(\theta^{t+1})}{\partial r_t}\right) + \mathbb{E}\left(\left(Q_{t+1}^* - \Pi_{t+1}\right)\frac{\partial\lambda(\theta^{t+1})}{\partial r_t}\right) + \frac{1}{R}\mathbb{E}\left(\frac{1-F^1(\theta_1)}{f^1(\theta_1)} l_{1,t+1}(\theta^{t+1})\phi_{t+1}'(l(\theta^{t+1}))\frac{\lambda_\theta\lambda_r}{\lambda\lambda_l}(\rho_{lr} - \rho_{\theta r})\right)$$

Screening:

Efficiency cost of distorting R&D effort:

- Allocative efficiency: inverse elasticity rule.
- Informational rent: increasing in complementarity effort-type → less costly mimicking of low types.

Optimal Profit wedge and R&D subsidy

$$\tau(\theta^t) = -\mathbb{E}\left(\sum_{s=t}^{T}(1-\delta)^{s-t}\eta_s\right)\frac{\partial\lambda_t(\theta^t)}{\partial l_t} - \mathbb{E}\left(Q_t^* - \Pi_t\right)\frac{\partial\lambda_t(\theta^t)}{\partial l_t}$$
$$+ \frac{1-F^1(\theta_1)}{f^1(\theta_1)} l_{1,t}(\theta^t)\frac{\phi_t'\lambda_{\theta t}}{\lambda_t}\left[\frac{1}{\varepsilon_{l,1-\tau}}\frac{1}{\varepsilon_{\lambda l,t}} + \rho_{\theta l,t}\right]$$

$$s(\theta^t) = \mathbb{E}\left(\sum_{s=t+1}^{T}(1-\delta)^{s-t-1}\eta_s\frac{\partial\lambda(\theta^{t+1})}{\partial r_t}\right) + \mathbb{E}\left(\left(Q_{t+1}^* - \Pi_{t+1}\right)\frac{\partial\lambda(\theta^{t+1})}{\partial r_t}\right)$$
$$+ \frac{1}{R}\mathbb{E}\left(\frac{1-F^1(\theta_1)}{f^1(\theta_1)} l_{1,t+1}(\theta^{t+1})\phi_{t+1}'(l(\theta^{t+1}))\frac{\lambda_\theta\lambda_r}{\lambda\lambda_l}(\rho_{lr} - \rho_{\theta r})\right)$$

Screening:

Efficiency cost of distorting R&D investments:

Higher $\rho_{lr} \rightarrow$ larger s. Incentivizes unobservable input, relaxes IC.

Higher $\rho_{\theta r} \rightarrow$ smaller s. Increases info rent, tightens IC.

Special case: $\rho_{lr} = \rho_{\theta r}$. Only distort R&D if improves screening and incentives for unobservable input.

Optimal Profit wedge and R&D subsidy

$$\tau(\theta^t) = -\mathbb{E}\left(\sum_{s=t}^{T}(1-\delta)^{s-t}\eta_s\right)\frac{\partial\lambda_t(\theta^t)}{\partial l_t} - \mathbb{E}\left(Q_t^* - \Pi_t\right)\frac{\partial\lambda_t(\theta^t)}{\partial l_t}$$
$$+ \frac{1-F^1(\theta_1)}{f^1(\theta_1)} l_{1,t}(\theta^t)\frac{\phi_t'\lambda_{\theta t}}{\lambda_t}\left[\frac{1}{\varepsilon_{l,1-\tau}}\frac{1}{\varepsilon_{\lambda l,t}} + \rho_{\theta l,t}\right]$$

$$s(\theta^t) = \mathbb{E}\left(\sum_{s=t+1}^{T}(1-\delta)^{s-t-1}\eta_s\frac{\partial\lambda(\theta^{t+1})}{\partial r_t}\right) + \mathbb{E}\left(\left(Q_{t+1}^* - \Pi_{t+1}\right)\frac{\partial\lambda(\theta^{t+1})}{\partial r_t}\right)$$
$$+ \frac{1}{R}\mathbb{E}\left(\frac{1-F^1(\theta_1)}{f^1(\theta_1)} l_{1,t+1}(\theta^{t+1})\phi_{t+1}'(l(\theta^{t+1}))\frac{\lambda_\theta\lambda_r}{\lambda\lambda_l}(\rho_{lr} - \rho_{\theta r})\right)$$

Monopoly Quality Valuation Correction:

- Wedges defined relative to patent system: monopolist does not value quality as much as society, needs extra incentive to invest.
- Disappears if wedges defined relative to prize system: social and private valuations aligned.
- Optimal R&D policy depends on IPR.

When Quantity Cannot be Controlled

Imagine irremovable patent system $\rightarrow$ monopoly quantity $k_t(q_t(\theta^t), \bar{q}_t)$ chosen for any quality.

Same formulas, but Q_t replaces Q_t^*.

Lesson: Optimal R&D policy depends on IPP.

Improving quality through R&D effort and investment subsidies here generates extra benefit by increasing monopolist's quantity.

$$\frac{\partial \tilde{Y}(q_t(\theta^t), \bar{q}_t)}{\partial q} = \underbrace{\frac{\partial Y(q_t(\theta^t), k_t(q_t(\theta^t), \bar{q}_t))}{\partial q_t(\theta^t)}}_{\text{Direct benefit}} + \underbrace{\left(p(q_t(\theta^t), k_t(q_t(\theta^t), \bar{q}_t)) - \frac{\partial C}{\partial k} \right)}_{\text{Monopoly distortion}} \frac{\partial k_t(q_t(\theta^t), \bar{q}_t)}{\partial q_t(\theta^t)}$$

Larger subsidy, lower tax, but lower investments overall, at higher cost (additional costly constraint).

Age-patterns of optimal policies are ambiguous. Three drivers:

(1) Finite horizon (can make ∞ or specify terminal value).

Both laissez-faire and socially optimal investments $\downarrow$ with time.

All else equal, Pigouvian correction $\downarrow$ with time (for positive spillover).

(2) Policies set at time $t = 1$. Decay as long as $I_{1,t} < 1$.

e.g.: AR(1) has $p^{t-1} < 1$

(3) Age-dependent primitives: $M_t(r)$, $C_t(k, \bar{q})$, $\phi_t(I_t)$, $\lambda_t(r, I, \theta)$, ..

Cross-sectional (cross-productivity-type) patterns:

Depends on primitives, but typically low productivity firms more distorted.

Extensions

(1) **Different types of R&D investments:**

$$\lambda_t = \lambda_t(r^1_{t-1}, ...r^j_{t-1}, ..., r^J_{t-1}, l_t, \theta_t)$$

$s^j(\theta^t)$ depends on i) externality $\frac{\partial \lambda_t}{\partial r^j_{t-1}}$, ii) complementarity: $\rho^j_{\theta l} - \rho^j_{\theta r}$.

$\rightarrow$ subsidize investments with higher externalities, but less so if they are highly complementary with unobservable firm productivity.

(2) **Different externalities:**

$$C(k, \bar{q}^1_t, .., \bar{q}^J_t) \quad \text{with} \quad \bar{q}^j_t = \int_{\Theta^t} q^j_t(\theta^t) d\theta^t$$

$$\text{and} \quad q^j_t(\theta^t) = q^j_t(\theta^{t-1})(1-\delta) + \lambda^j_t(r^j_{t-1}, l_t, \theta_t)$$

Basic vs. Applied research?

Implementation Results

Many possible (theoretically equivalent) implementations.
Administrative/political constraints may matter in practice.

Optimal allocation when quantity can be controlled can be implemented:

1) with price subsidy ($p(k,q)(1+s_p(p,k)) = \frac{Y(k,q)}{k}$) plus age-dependent tax function $T_t(q_t, r_t, q_{t-1}, r_{t-1}, q_1)$.

with constant markup $Y(k,q) = \frac{1}{1-\beta} q^\beta k^{1-\beta}$, constant $s_p = \frac{\beta}{1-\beta}$.

2) with prize $G_t(\lambda_t, r_t, r_{t-1}, q_1)$, government purchases innovation from firms, produces the socially optimal quantity.

Allocation when quantity can not be controlled implemented by tax $T_t^n(q_t, r_t, q_{t-1}, r_{t-1}, q_1)$ (no price subsidy).

Quantitative Investigation

Dataset Information: Compustat and Patent Data

Patent data from USPTO matched to Compustat data.

Select firms as in Bloom, Schankerman and Van Reenen (2013): patent $\geq$ once since 1963, observed $\geq$ 4 times in 1980–2001.

Variable	Mean	Median
Sales (in mil. USD)	3133	494
Citations per patent	7.7	6
Patents per year	18.5	1
R&D spending / sales	0.043	0.014
Number of employees (000's)	18.4	3.8
Number of firms	736	

λ = flow of citations per patent. q = depreciated stock.

Functional Forms for Estimation

Function	Notation	Functional form
Consumer valuation	$Y(q_t, k_t)$	$\frac{1}{1-\beta} q_t^{\beta} k_t^{1-\beta}$
Cost function	$C_t(k, \bar{q}_t)$	$\frac{k}{\bar{q}_t^{\zeta}}$
Quality accumulation	$H(q_{t-1}, \lambda_t)$	$q_t = (1-\delta)q_{t-1} + \lambda_t$
Step size	$\lambda_t(r_{t-1}, l_t, \theta_t)$	$(\alpha r_{t-1}^{1-\rho_{\theta r}} + (1-\alpha)\theta_t^{1-\rho_{\theta r}})^{\frac{1}{1-\rho_{\theta r}}} l_t$
Disutility of effort	$\phi_t(l_t)$	$\kappa_l \frac{l_t^{1+\gamma}}{1+\gamma}$
Cost of R&D	$M_t(r_t)$	$\kappa_r \frac{r_t^{1+\eta}}{1+\eta}$
Stochastic type process	$f^t(\theta_t \mid \theta_{t-1})$	$\log \theta_t = p \log \theta_{t-1} + (1-p)\mu_\theta + \epsilon_t$
Distribution of heterogeneity θ_1	$f^1(\theta_1)$	$f^1(\theta_1) = \frac{I_{\Theta 1}(\theta_1)}{\theta_1[\underline{\theta}_1 - \bar{\theta}_1]}$
Initial quality level	q_0	0

Estimation Targets: Moments

Moment	Simulation	Target
Patent quality-R&D elasticity	0.50	0.57
R&D/Sales median	0.014	0.013
Sales growth (DHS) mean	0.08	0.074
Within-firm patent quality coeff of var	0.67	0.77
Across-firm patent quality coeff of var:		
Young firms	1.17	1.10
Older firms	0.71	0.63
Patent quality young/old	2.00	1.88
Spillover coefficient	0.191	0.188

Parameters to be estimated: $\chi = (\alpha, \rho_{\theta r}, \sigma_{\varepsilon}, p, \kappa_l, \kappa_r, \zeta, \Theta^1)$

Loss function: $L(\chi) = \sum_{k=1}^{8} \left(\frac{\text{moment}_k^{\text{model}}(\chi) - \text{moment}_k^{\text{data}}}{\text{moment}_k^{\text{data}}} \right)^2$

Estimation Targets: Moments

Moment	Simulation	Target
Patent quality-R&D elasticity	0.50	0.57
R&D/Sales median	0.014	0.013
Sales growth (DHS) mean	0.08	0.074
Within-firm patent quality coeff of var	0.67	0.77
Across-firm patent quality coeff of var:		
Young firms	1.17	1.10
Older firms	0.71	0.63
Patent quality young/old	2.00	1.88
Spillover coefficient	0.191	0.188

Replicate Bloom, Schankerman, and Van Reenen (2013) IV estimates.

Randomly draw κ_r in $M(r) = \kappa_r \frac{r_t^{1+\eta}}{1+\eta}$.

Match regression coefficient of q_t on average R&D stock.

Estimated Parameters

Parameter	Symbol	Value
	External Calibration	
Effort cost elasticity	γ	1
Interest rate	R	1.05
Intangibles depreciation	δ	0.1
Knowledge share	β	0.15
R&D cost elasticity	η	1.5
Level of types	μ_θ	0.00
Initial R&D stock	r_0	1.0
	Internal Calibration	
R&D share	α	0.390
R&D-type substitution	$\rho_{\theta r}$	0.861
Type variance	σ_ϵ	0.253
Type persistence	p	0.71
Scale of disutility	κ_l	0.88
Scale of R&D cost	κ_r	0.048
Support width for θ_1	Θ^1	1.98
Production externality	ζ	0.022

Gross and Net Incentives

Gross subsidy vs. net subsidy (on top of making R&D expenses corporate tax-deductible).

Gross subsidy $\tilde{s}$:

$$\pi(1-\tau) - (1-\tilde{s})M(r)$$

Net incentive for R&D is s such that:

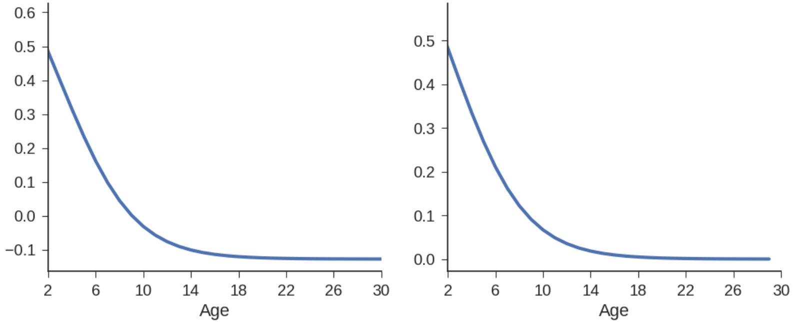
$$\underbrace{(\pi - M(r))}_{\text{Deduct R\&D expenses}} (1-\tau) - \underbrace{(1-s)}_{\text{Net subsidy}} M(r)$$

Relation: $s = \tilde{s} - \tau$.

Same idea for wedges.

Young vs Old Firms

(a) Profit wedge (b) Gross R&D wedge

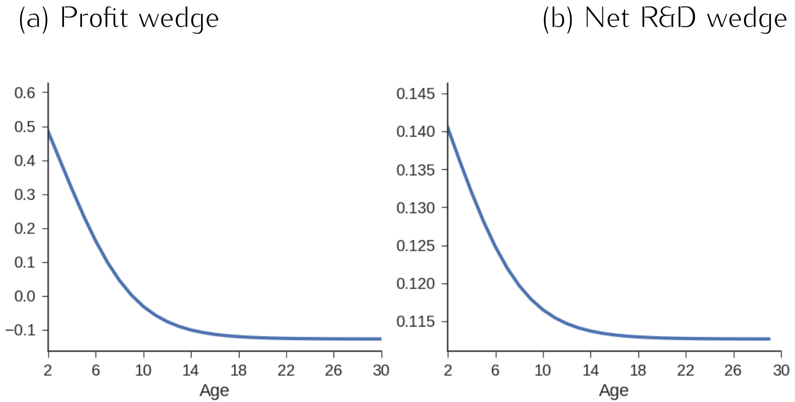


Policies fall with age monotonically: screening term decays.

Converge to Pigouvian correction.

Screening term is positive for R&D investments since $\rho_{\theta r} < \rho_{Ir} = 1$.

Young vs Old Firms



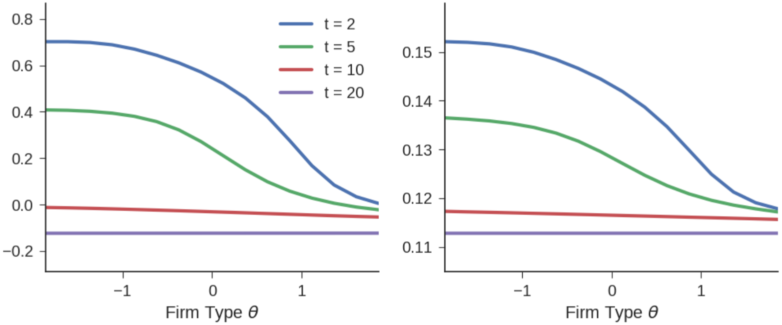
Policies fall with age monotonically: screening term decays.

Converge to Pigouvian correction.

Screening term is positive for R&D investments since $\rho_{\theta r} < \rho_{Ir} = 1$.

High vs Low Productivity Firms

(a) Profit wedge (b) Net R&D wedge



Marginal tax rate and R&D subsidy lower for higher productivity firms.

"No distortion at the top" ▸ Comparative Statics ▸ No IPR ▸ Allocations

Optimal subsidy highly nonlinear (not the case for current policies).

Optimal Simpler Policies

Simpler Policy Results: Optimal Linear Policies

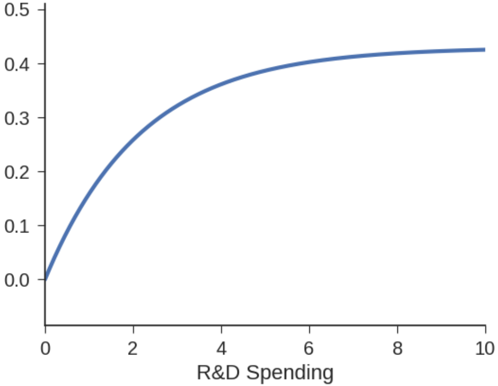
Age	Optimal τ	Optimal $\tilde{s}$	Optimal s	Revenue Loss
		A. Age-Independent Policies		
1-15	31.4%	40.5 %	9.1%	65%
		B. Age-Dependent, 2 age brackets		
1-15	31.6%	40.6%	9%	64%
16-30	31.0%	39.7%	8.7%	
		C. Age-Dependent, 4 age brackets		
1-7	32.7%	40.8%	8.1%	63%
8-15	30.8%	39.8%	9%	
16-22	30.9%	40.1%	9.2%	
23-30	31.5%	40.15%	8.3 %	

Optimal Nonlinear R&D Subsidy

$s(M) = c_0 + (c_1 - c_0) \cdot (1 - e^{-c_2 M})$

Optimum: $c_0 = 0$, $c_1 = 43\%$ and $c_2 = 46\%$. $\tau = 31.56\%$.

Revenue loss: 62%.



Comparative Statics on Key Parameters

Parameter	Optimal τ	Optimal $\tilde{s}$	Optimal s	Revenue Loss
A. Role of Persistence p				
$p = 0.5$	31.4%	40.5 %	9.1%	70.5%
$p = 0.9$	31.3%	40.3 %	9%	45.8%
B. Role of Complementarity $\rho_{\theta r}$				
$\rho_{\theta r} = \rho_{Ir} = 1$	33.6%	41.3%	7.7%	61.7%
$\rho_{\theta r} = 1.2 > \rho_{Ir}$	35.6%	41.3%	5.7%	58.9
C. Role of the Technology Spillover				
$\zeta = 0.01$	36.5%	41.4%	4.9%	73.8%
$\zeta = 0.03$	28%	39.7%	11.7%	73%

Parameters that matter for revenue loss are those that affect screening term.

More persistent types $\rightarrow$ lower value of fine-tuning mechanism.

Comparative Statics on Key Parameters

Parameter	Optimal τ	Optimal $\tilde{s}$	Optimal s	Revenue Loss
A. Role of Persistence p				
$p = 0.5$	31.4%	40.5 %	9.1%	70.5%
$p = 0.9$	31.3%	40.3 %	9%	45.8%
B. Role of Complementarity $\rho_{\theta r}$				
$\rho_{\theta r} = \rho_{lr} = 1$	33.6%	41.3%	7.7%	61.7%
$\rho_{\theta r} = 1.2 > \rho_{lr}$	35.6%	41.3%	5.7%	58.9
C. Role of the Technology Spillover				
$\zeta = 0.01$	36.5%	41.4%	4.9%	73.8%
$\zeta = 0.03$	28%	39.7%	11.7%	73%

Higher complementarity $\rho_{\theta r}$ decreases subsidy.

Revenue loss larger because unrestricted mechanism also does less well.

Comparative Statics on Key Parameters

Parameter	Optimal τ	Optimal $\tilde{s}$	Optimal s	Revenue Loss
A. Role of Persistence p				
$p = 0.5$	31.4%	40.5 %	9.1%	70.5%
$p = 0.9$	31.3%	40.3 %	9%	45.8%
B. Role of Complementarity $\rho_{\theta r}$				
$\rho_{\theta r} = \rho_{lr} = 1$	33.6%	41.3%	7.7%	61.7%
$\rho_{\theta r} = 1.2 > \rho_{lr}$	35.6%	41.3%	5.7%	58.9
C. Role of the Technology Spillover				
$\zeta = 0.01$	36.5%	41.4%	4.9%	73.8%
$\zeta = 0.03$	28%	39.7%	11.7%	73%

Higher externality justifies lower taxes and higher subsidies.

Does not affect losses from linear policies.

Conclusion

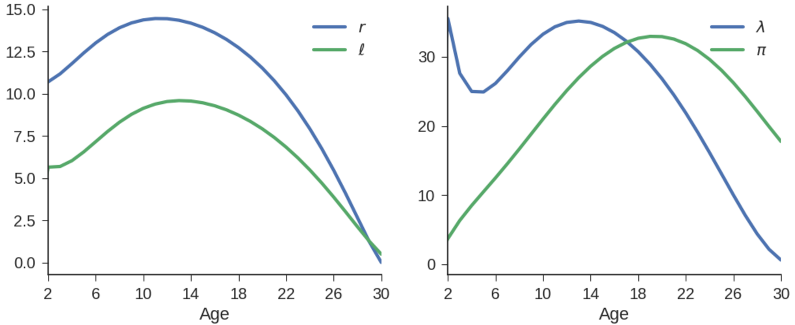
- Model of innovation with heterogeneous firms, private information, and spillovers.
  - Use mechanism design to solve for constrained efficient allocations.
  - Implementation by a tax/subsidy or prize mechanism.
- Externality $\rightarrow$ Optimal to subsidize R&D investments.
- Asymmetric information could go other way in theory if R&D very complementary to firm productivity ($\uparrow$ informational rents to firms).
- Revenue loss from restricted policies is large.

APPENDIX

Young vs. Old firms: Allocations

(a) R&D effort and investment

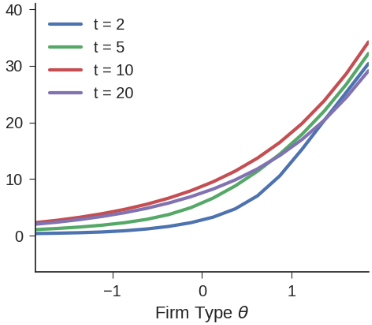
(b) Step size and profits



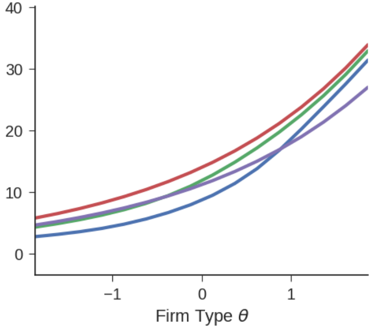
▸ Back

High vs. Low Productivity Firms: Allocations

(a) R&D effort and investment

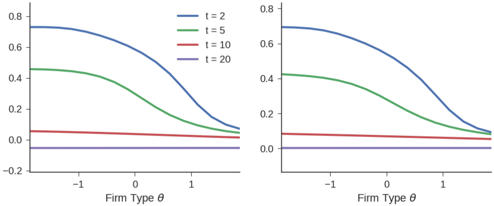


(b) Step size and profits

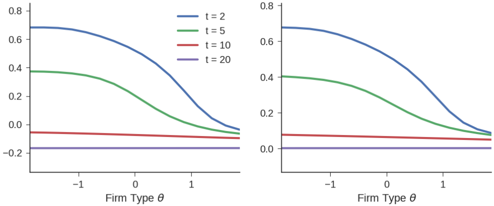


Back

Role of spillovers: profit wedge (left) and R&D wedge (right)

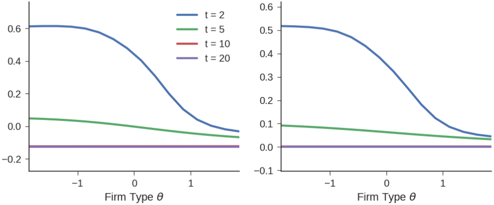


(a) Small spillover $\zeta = 0.01$

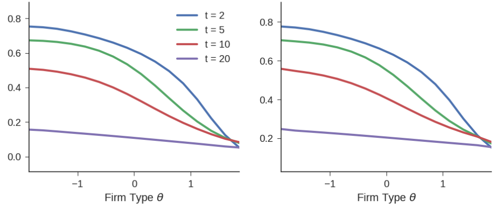


(b) Large spillover $\zeta = 0.03$

Role of persistence: profit wedge (left) and R&D wedge (right)

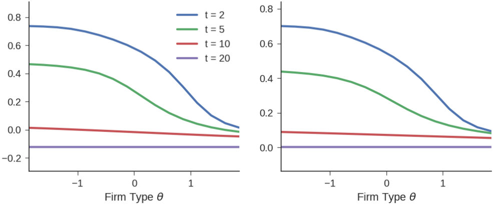


(c) Low persistence, $p = 0.5$

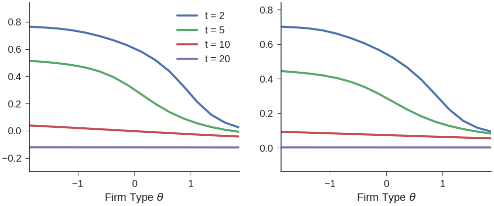


(d) High persistence, $p = 0.9$

Role of complementarity: profit wedge (left) & R&D wedge (right)



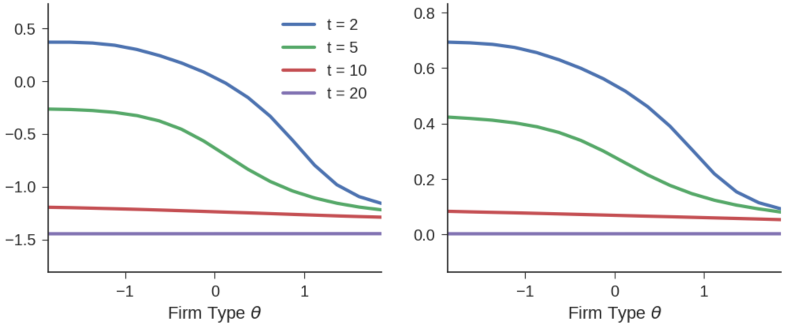
(e) Case with $\rho_{\theta r} = \rho_{rl} = 1$



(f) Case with $\rho_{\theta r} = 1.2 > \rho_{rl} = 1$

Wedges when quantity cannot be controlled

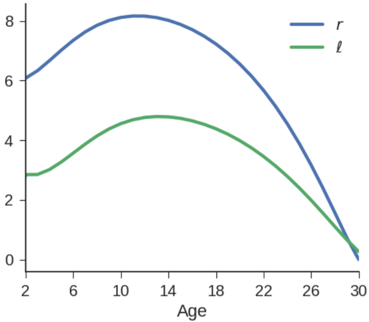
(a) Profit wedge (b) R&D wedge



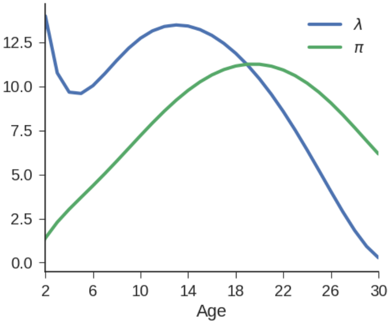
▸ Back

Young vs. Old firms: Allocations when Quantity not Controlled

(a) R&D effort and investment



(b) Step size and profits

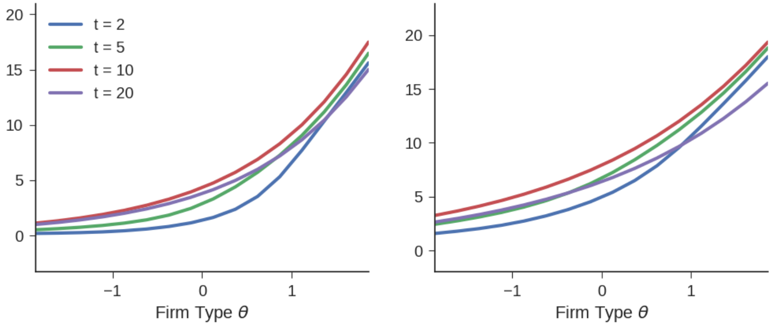


▸ Back

High vs. Low Productivity Firms: Allocations when Quantity not Controlled

(a) R&D effort and investment

(b) Step size and profits



▸ Back