Optimal Taxation and R&D Policies

Ufuk Akcigit
Chicago

Douglas Hanley
Pittsburgh

Stefanie Stantcheva
Harvard

April 2019
Motivation I: 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”, J. Furman, former 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%)
Is the amount spent by government correlated with better productivity?
Motivation II: Private Information is an Important Constraint

- Take young firms at start of their lifecycle. How much of the variation in subsequent innovation quantity & quality can we explain based on observables?
  - Observables: age, assets, past investments, sales, state FE, year FE, sector FE (+ all interactions), and even past innovations:
    - $R^2$ not above 0.3, improves with age (as info revealed).
    - Conditional on these observables, many “outlier” firms.

- Two ways of possibly addressing asymmetric info problem:
  - **Direct screening**: what the NSF and VCs try to do. Done by the government with public procurement. Hard to do and very costly on a large scale.
  - **Indirect screening**: Design a menu of options (implemented by taxes and subsidies), let firms self-select! “Easy” to decentralize and scalable.
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.

1) Mechanism design: no a priori restriction on policy tools. Characterize constrained efficient allocations. Implementation.

2) Quantitative Investigation using Patent data + Compustat data.

3) Losses from “simpler” policies (e.g.: linear age-dependent...).
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**.

1) Mechanism design: no a priori restriction on policy tools.

  Characterize constrained efficient allocations.

  Implementation.

2) Quantitative Investigation using Patent data + **Longitudinal Business Database (LBD)** data.

3) Losses from “simpler” policies (e.g.: linear age-dependent...).
Related Literature


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


Outline

1. Model

2. Optimal Unrestricted Mechanism

3. Quantitative Investigation

4. Optimal Simpler Policies
Model
Household

Government

Final Goods producer

Intermediate Goods producers

\[ Y_t = \int_i Y_t(q_t(i), k_t(i)) \, di \]

spillovers

\[ \bar{q}_t \]

Intellectual Property Rights

Max consumption

R&D & Tax Policies

Demand

\[ p_t(k_t(i), q_t(i)) \]
\[ Y_t = \int_i Y(q_t(i), k_t(i)) \]
Household

Final Goods producer

Demand $p(k_t(i), q_t(i))$

Intermediate Goods producers

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

Production

- Quality $q_t(i)$, quantity $k_t(i)$
- Demand: $p(k_t(i), q_t(i))$
Final Goods producer \rightarrow \text{Demand } p(k_t(i), q_t(i)) \rightarrow \text{Intermediate Goods producers}

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

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: \( \bar{q}_t = \int_i q_t(i) di \)
- \( \pi(q_t(i), \bar{q}_t) = \max_k \{ p(k, q_t(i))k - C(k, \bar{q}_t) \} \)
Household

Government

Final Goods producer

Intermediate Goods producers

Demand: \( p(k_t(i), q_t(i)) \)

Y_t = \int_Y 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: \( \bar{q}_t = \int \bar{q}_t(i) \, di \)
- \( \pi(q_t(i), \bar{q}_t) = \max_k \{ p(k, q_t(i)) \, k - C(k, \bar{q}_t) \} \)
Household

Government

Final Goods producer

Intermediate Goods producers

Demand $p(k_t(i), q_t(i))$

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

Max consumption

R&D & Tax Policies

spillovers $\tilde{q}_t$

Intellectual Property Rights
Household

Intermediate Goods producers

Final Goods producer

Government

Demand $p(k_t(i), q_t(i))$

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

Max consumption

R&D & Tax Policies

spillovers $\tilde{q}_t$
Intermediate Producers’ static production decisions

- Intermediate good producers: quality $q_t(i)$, quantity $k_t(i)$.

- Final good producer aggregates intermediate goods for consumption:

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

- Return to quality and quantity 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: $C(k_t, \bar{q}_t)$ (↑ or ↓ 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 $\theta_t$ (managerial/firm quality), Markov $f^t(\theta_t|\theta_{t-1})$, history $\theta^t$.

- $\frac{\partial \lambda}{\partial \theta} > 0$, $\frac{\partial \lambda}{\partial r} > 0$, $\frac{\partial \lambda}{\partial l} > 0$, $\frac{\partial^2 \lambda}{\partial \theta \partial l} > 0$ (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

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

\[
\phi'_t(l_t(\theta^t)) = \mathbb{E} \left( \sum_{s=t}^{\infty} \left( \frac{1-\delta}{R} \right)^{s-t} \left( \frac{\partial \tilde{Y}^*(q_t(\theta^s), \bar{q}_t)}{\partial q_s} + \frac{\partial \tilde{Y}^*(q_t(\theta^s), \bar{q}_t)}{\partial \bar{q}_s} \right) \frac{\partial \lambda_t(\theta^t)}{\partial l_t(\theta^t)} \right)
\]
Asymmetric Information

- Consider two cases.
  1. Firm productivity $\theta_t$ and R&D effort not observable...
  2. ... and quantity $k_t$ not observable.

- Case (1) $\iff$ can optimize on intellectual property rights policy.
  - Optimal IPR trivial here: prize system or patent system + price subsidy.

- Case (2) $\iff$ 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 revenue requirement (could be $\leq 0$).
Coefficient of Complementarity

- Hicksian coefficient of complementarity between $x$ and $y$:

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

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

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

- When $\rho_{\theta r}$ is large: Higher R&D investments increase informational rent that needs to be forfeited to high quality firms.
Additional observable heterogeneity (sector? product type?) can be conditioned on.

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.
Optimal Unrestricted Mechanism
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 $\theta^t$:
  \[ V_t(\theta^t) = \sum_{t=s}^{T} (\frac{1}{R})^{t-s} \cdot \left\{ \int_{\Omega} \{ T_t(\theta^t) - \phi_t(l_t(\theta^t)) \} P(\theta^t | \theta^s) d\theta^t \right\} \]

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

- Envelope condition (Pavan, Segal and Toikka, 2014):
  \[ \frac{\partial V_t(\theta^t)}{\partial \theta_t} = \mathbb{E}\left\{ \sum_{s=t}^{\infty} I_{t,s} \frac{\partial \tilde{U}(\theta^\infty)}{\partial \theta_s} \right\} \]
  - $I_{t,s}$: impulse response of shock $\theta_t$ on time $s$ shock $\theta_s$. For AR(1) is $p^{s-t}$.
  - Relies on first-order condition (sufficiency?)

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

- Expected PDV of transfers = expected PDV of disutility costs + information ($V_1(\theta_1)$).
Program: Virtual Surplus with Spillovers

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

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

$$P(\bar{q}) = \max W(\bar{q}) = \mathbb{E}\{ \sum_{t=1}^{T} \left( \frac{1}{R} \right)^{t-1} \{ \tilde{Y}^*(q_t(\theta^t), \bar{q}_t) - M_t(r_t(\theta^t)) - \phi_t(l_t(\theta^t)) \\ - V_1(\theta_1) - \frac{1 - F_1(\theta_1)}{f_1(\theta_1)} l_{1,t} \partial \tilde{U}_t/\partial \theta_t \} \}$$

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

and $$q_t(\theta^t) = q_{t-1}(\theta^{t-1})(1 - \delta) + \lambda_t(l_t(\theta^t), r_{t-1}(\theta^{t-1}), \theta_t)$$
Program: Virtual Surplus with Spillovers

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

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

\[
P(\bar{q}) = \max W(\bar{q}) = \mathbb{E}\left\{\sum_{t=1}^{T} \left(\frac{1}{R}\right)^{t-1} \left\{ \tilde{Y}(q_t(\theta^t), \bar{q}_t) - M_t(r_t(\theta^t)) - \phi_t(l_t(\theta^t)) \right. \right.
\]
\[
\left. - V_1(\theta_1) - \frac{1 - F^1(\theta_1)}{f^1(\theta_1)} l^1_{1,t} \frac{\partial \tilde{U}_t}{\partial \theta_t} \right\} \}
\]

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

and \[q_t(\theta^t) = q_{t-1}(\theta^{t-1})(1 - \delta) + \lambda_t(l_t(\theta^t), r_{t-1}(\theta^{t-1}), \theta_t)\]
Wedges: Measures of Distortions in the Allocations

Akin to “implicit” taxes and subsidies.

\[ \tau(\theta^t) \equiv \mathbb{E} \left( \sum_{s=t}^{\infty} \left( \frac{1 - \delta}{R} \right)^{s-t} \frac{\partial \pi_s(q_s(\theta^s), \bar{q}_s)}{\partial q_s(\theta^s)} \frac{\partial \lambda_t(\theta^t)}{\partial l_t(\theta^t)} \right) - \phi'(l_t(\theta^t)) \]

\[ s(\theta^t) \equiv M'_t(r(\theta^t)) - \frac{1}{R} \mathbb{E} \left( \sum_{s=t+1}^{\infty} \left( \frac{1 - \delta}{R} \right)^{s-t-1} \frac{\partial \pi_s(q_s(\theta^s), \bar{q}_s)}{\partial q_s(\theta^s)} \frac{\partial \lambda_{t+1}(\theta^{t+1})}{\partial r_t(\theta^t)} \right) \]

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

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

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

\[ Q_t(\theta^t) \equiv \frac{1}{R} \left( \sum_{s=t}^{\infty} \left( \frac{1 - \delta}{R} \right)^{s-t} \frac{\partial \tilde{Y}(q(\theta^s), \bar{q}_s)}{\partial q_s} \right) \] (impact on social surplus, if quantity not controlled).
Optimal Profit wedge and R&D subsidy

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

\[ + \frac{1 - F_1^1(\theta_1)}{f_1^1(\theta_1)} l_{1,t}(\theta^t) \frac{\phi'_t \lambda t}{\lambda_t} \left[ \frac{1}{\varepsilon_{l,1-t}} \frac{1}{\varepsilon \lambda l,t} + \rho \theta l,t \right] \]

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

\[ + \frac{1}{R} \mathbb{E} \left( \frac{1 - F_1^1(\theta_1)}{f_1^1(\theta_1)} l_{1,t+1}(\theta^{t+1}) \phi'_{t+1} (l(\theta^{t+1})) \frac{\lambda_t \lambda_r}{\lambda^l} (\rho_{lr} - \rho_{\theta r}) \right) \]
Optimal Profit wedge and R&D subsidy

\[ \tau(\theta^t) = -\mathbb{E} \left( \sum_{s=t}^{\infty} (1 - \delta)^{s-t} \eta_s \right) \frac{\partial \lambda_t(\theta^t)}{\partial l_t} - \mathbb{E} (Q^*_t - \Pi_t) \frac{\partial \lambda_t(\theta^t)}{\partial l_t} \]

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

\[ \sigma(\theta^t) = \mathbb{E} \left( \sum_{s=t+1}^{\infty} (1 - \delta)^{s-t-1} \eta_s \frac{\partial \lambda(\theta^{t+1})}{\partial r_t} \right) + \mathbb{E} \left( (Q^*_{t+1} - \Pi_{t+1}) \frac{\partial \lambda(\theta^{t+1})}{\partial r_t} \right) \]

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

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

\[
\tau(\theta^t) = -\mathbb{E} \left( \sum_{s=t}^{\infty} (1 - \delta)^{s-t} \eta_s \right) \frac{\partial \lambda_t(\theta^t)}{\partial l_t} - \mathbb{E} (Q^*_t - \Pi_t) \frac{\partial \lambda_t(\theta^t)}{\partial l_t}
\]

\[
+ \frac{1 - F^1(\theta_1)}{f^1(\theta_1)} l_{1,t}(\theta^t) \frac{\phi'_t \lambda_{\theta_t}}{\lambda_t} \left[ \frac{1}{\varepsilon_{l,1-t}} \frac{1}{\varepsilon_{\lambda_{l,t}}} + \rho_{\theta_{l,t}} \right]
\]

\[
s(\theta^t) = \mathbb{E} \left( \sum_{s=t+1}^{\infty} (1 - \delta)^{s-t-1} \eta_s \frac{\partial \lambda(\theta^{t+1})}{\partial r_t} \right) + \mathbb{E} \left( (Q^*_{t+1} - \Pi_{t+1}) \frac{\partial \lambda(\theta^{t+1})}{\partial r_t} \right)
\]

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

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

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

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

Screening:

Efficiency cost of distorting R&D effort:

Allocative efficiency: inverse elasticity rule.

Informational rent: increasing in complementarity effort-type \(\rightarrow\) less costly mimicking of low types.
Optimal Profit wedge and R&D subsidy

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

\[ + \frac{1 - F^1(\theta_1)}{f^1(\theta_1)} l_{1,t}(\theta^t) \phi'_t \lambda_\theta \left[ \frac{1}{\varepsilon_{l,1-t}} \frac{1}{\varepsilon_{l,t}} + \rho_{\theta l,t} \right] \]

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

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

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

$$\tau(\theta^t) = -E \left( \sum_{s=t}^{\infty} (1 - \delta)^{s-t} \eta_s \right) \frac{\partial \lambda_t(\theta^t)}{\partial l_t} - E (Q^*_t - \Pi_t) \frac{\partial \lambda_t(\theta^t)}{\partial l_t}$$
$$+ \frac{1 - F^1(\theta_1)}{f^1(\theta_1)} l_{1,t}(\theta^t) \phi'_t \lambda_{\theta t} \left[ \frac{1}{\varepsilon l_{1-t}} \frac{1}{\varepsilon \lambda_{l,t}} + \rho \theta_{l,t} \right]$$

$$s(\theta^t) = E \left( \sum_{s=t+1}^{\infty} (1 - \delta)^{s-t-1} \eta_s \frac{\partial \lambda(\theta^{t+1})}{\partial r_t} \right) + E \left( (Q^*_{t+1} - \Pi_{t+1}) \frac{\partial \lambda(\theta^{t+1})}{\partial r_t} \right)$$
$$+ \frac{1}{R} E \left( \frac{1 - F^1(\theta_1)}{f^1(\theta_1)} l_{1,t+1}(\theta^{t+1}) \phi'_{t+1} (l(\theta^{t+1})) \frac{\lambda_{\theta l}}{\lambda_{\lambda l}} (\rho_{lr} - \rho_{\theta r}) \right)$$

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.

\[
\frac{\partial \tilde{Y}(q_t(\theta^t), \bar{q}_t)}{\partial q} = \frac{\partial Y(q_t(\theta^t), k_t(q_t(\theta^t), \bar{q}_t))}{\partial q_t(\theta^t)} + \left( p(q_t(\theta^t), k_t(q_t(\theta^t), \bar{q}_t)) - \frac{\partial C}{\partial k} \right) \frac{\partial k_t(q_t(\theta^t), \bar{q}_t)}{\partial q_t(\theta^t)}
\]

Direct benefit + Monopoly distortion

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

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

\[ \lambda_t = \lambda_t(r_{t-1}^1, ..., r_{t-1}^j, ..., r_{t-1}^j, l_t, \theta_t) \]

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

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

(2) **Different externalities:**

\[ C(k, \bar{q}_t^1, ..., \bar{q}_t^J) \text{ with } \bar{q}_t^j = \int_{\Theta_t} q_t^j(\theta^t) d\theta^t \]

and \( q_t^j(\theta^t) = q_t^j(\theta^{t-1})(1 - \delta) + \lambda_t^j(r_{t-1}^j, l_t, \theta_t) \)

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

Allocation when quantity can not be controlled implemented by tax \( T^n_t(q_t, r_t, q_{t-1}, r_{t-1}, q_1) \) (no price subsidy).
Quantitative Investigation
Dataset Information: Compustat/LBD and Patent Data

Patent data from USPTO matched to Compustat or LBD data.

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

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales (in mil. USD)</td>
<td>3133</td>
<td>494</td>
</tr>
<tr>
<td>Citations per patent</td>
<td>7.7</td>
<td>6</td>
</tr>
<tr>
<td>Patents per year</td>
<td>18.5</td>
<td>1</td>
</tr>
<tr>
<td>R&amp;D spending / sales</td>
<td>0.043</td>
<td>0.014</td>
</tr>
<tr>
<td>Number of employees (000’s)</td>
<td>18.4</td>
<td>3.8</td>
</tr>
<tr>
<td>Number of firms</td>
<td>736</td>
<td></td>
</tr>
</tbody>
</table>

$\lambda =$ flow of citations per patent. $q =$ depreciated stock.
# Functional Forms for Estimation

<table>
<thead>
<tr>
<th>Function</th>
<th>Notation</th>
<th>Functional form</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consumer valuation</td>
<td>$Y(q_t, k_t)$</td>
<td>$\frac{1}{1-\beta} q_t^\beta k_t^{1-\beta}$</td>
</tr>
<tr>
<td>Cost function</td>
<td>$C_t(k, \bar{q}_t)$</td>
<td>$\frac{k}{\bar{q}_t}$</td>
</tr>
<tr>
<td>Quality accumulation</td>
<td>$H(q_{t-1}, \lambda_t)$</td>
<td>$q_t = (1-\delta)q_{t-1} + \lambda_t$</td>
</tr>
<tr>
<td>Step size</td>
<td>$\lambda_t(r_{t-1}, l_t, \theta_t)$</td>
<td>$(\alpha r_{t-1}^{1-\rho r} + (1-\alpha)\theta_t^{1-\rho r}) \frac{1}{1-\rho r} l_t$</td>
</tr>
<tr>
<td>Disutility of effort</td>
<td>$\phi_t(l_t)$</td>
<td>$\kappa l_t^{1+\gamma}$</td>
</tr>
<tr>
<td>Cost of R&amp;D</td>
<td>$M_t(r_t)$</td>
<td>$\kappa r_t^{1+n}$</td>
</tr>
<tr>
<td>Stochastic type process</td>
<td>$f^t(\theta_t</td>
<td>\theta_{t-1})$</td>
</tr>
<tr>
<td>Distribution of heterogeneity $\theta_1$</td>
<td>$f^1(\theta_1)$</td>
<td>$f^1(\theta_1) = \frac{l_{\theta_1}(\theta_1)}{\theta_1[\theta_1-\bar{\theta}_1]}$</td>
</tr>
<tr>
<td>Initial quality level</td>
<td>$q_0$</td>
<td>0</td>
</tr>
</tbody>
</table>
# Estimation Targets: Moments

<table>
<thead>
<tr>
<th>Moment</th>
<th>Target</th>
<th>Simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1. Patent quality-R&amp;D elasticity</td>
<td>0.879</td>
<td>0.956</td>
</tr>
<tr>
<td>M2. R&amp;D/Sales median</td>
<td>0.041</td>
<td>0.029</td>
</tr>
<tr>
<td>M3. Sales growth (DHS) mean</td>
<td>0.060</td>
<td>0.072</td>
</tr>
<tr>
<td>M4. Within-firm patent quality coeff of var</td>
<td>0.630</td>
<td>0.720</td>
</tr>
<tr>
<td>M5. Across-firm patent quality coeff of var (young)</td>
<td>1.055</td>
<td>1.026</td>
</tr>
<tr>
<td>M6. Across-firm patent quality coeff of var (old)</td>
<td>0.991</td>
<td>0.768</td>
</tr>
<tr>
<td>M7. Patent quality young/old</td>
<td>1.043</td>
<td>1.022</td>
</tr>
<tr>
<td>M8. Spillover regression coefficient</td>
<td>0.191</td>
<td>0.190</td>
</tr>
<tr>
<td>M9. Subsidy regression coefficient</td>
<td>0.350</td>
<td>0.362</td>
</tr>
</tbody>
</table>

Parameters to be estimated: \( \chi = (\alpha, \rho, \sigma, \rho, \kappa, \kappa, \gamma, \zeta, \Theta) \)

Loss function: \( L(\chi) = \sum_{k=1}^{9} \left( \frac{\text{moment}_{k}^\text{model}(\chi) - \text{moment}_{k}^\text{data}}{\text{moment}_{k}^\text{data}} \right)^2 \)
## Estimation Targets: Moments

<table>
<thead>
<tr>
<th>Moment</th>
<th>Target</th>
<th>Simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1. Patent quality-R&amp;D elasticity</td>
<td>0.879</td>
<td>0.956</td>
</tr>
<tr>
<td>M2. R&amp;D/Sales median</td>
<td>0.041</td>
<td>0.029</td>
</tr>
<tr>
<td>M3. Sales growth (DHS) mean</td>
<td>0.060</td>
<td>0.072</td>
</tr>
<tr>
<td>M4. Within-firm patent quality coeff of var</td>
<td>0.630</td>
<td>0.720</td>
</tr>
<tr>
<td>M5. Across-firm patent quality coeff of var (young)</td>
<td>1.055</td>
<td>1.026</td>
</tr>
<tr>
<td>M6. Across-firm patent quality coeff of var (old)</td>
<td>0.991</td>
<td>0.768</td>
</tr>
<tr>
<td>M7. Patent quality young/old</td>
<td>1.043</td>
<td>1.022</td>
</tr>
<tr>
<td>M8. Spillover regression coefficient</td>
<td>0.191</td>
<td>0.190</td>
</tr>
<tr>
<td>M9. Subsidy regression coefficient</td>
<td>0.350</td>
<td>0.362</td>
</tr>
</tbody>
</table>

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

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

Match regression coefficient of $q_t$ on average R&D stock.
### Estimation Targets: Moments

<table>
<thead>
<tr>
<th>Moment</th>
<th>Target</th>
<th>Simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1. Patent quality-R&amp;D elasticity</td>
<td>0.879</td>
<td>0.956</td>
</tr>
<tr>
<td>M2. R&amp;D/Sales median</td>
<td>0.041</td>
<td>0.029</td>
</tr>
<tr>
<td>M3. Sales growth (DHS) mean</td>
<td>0.060</td>
<td>0.072</td>
</tr>
<tr>
<td>M4. Within-firm patent quality coeff of var</td>
<td>0.630</td>
<td>0.720</td>
</tr>
<tr>
<td>M5. Across-firm patent quality coeff of var (young)</td>
<td>1.055</td>
<td>1.026</td>
</tr>
<tr>
<td>M6. Across-firm patent quality coeff of var (old)</td>
<td>0.991</td>
<td>0.768</td>
</tr>
<tr>
<td>M7. Patent quality young/old</td>
<td>1.043</td>
<td>1.022</td>
</tr>
<tr>
<td>M8. Spillover regression coefficient</td>
<td>0.191</td>
<td>0.190</td>
</tr>
<tr>
<td>M9. Subsidy regression coefficient</td>
<td>0.350</td>
<td>0.362</td>
</tr>
</tbody>
</table>

Replicate Bloom et al. effects of R&D tax credits estimates.

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

Match regression coefficient R&D stock on cost of R&D (tax credits).
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interest rate</td>
<td>$R$</td>
<td>1.050</td>
</tr>
<tr>
<td>R&amp;D share</td>
<td>$\alpha$</td>
<td>0.398</td>
</tr>
<tr>
<td>Knowledge share</td>
<td>$\beta$</td>
<td>0.150</td>
</tr>
<tr>
<td>Intangibles depreciation</td>
<td>$\delta$</td>
<td>0.100</td>
</tr>
<tr>
<td>Type variance</td>
<td>$\sigma_\epsilon$</td>
<td>0.337</td>
</tr>
<tr>
<td>R&amp;D cost elasticity</td>
<td>$\eta$</td>
<td>1.500</td>
</tr>
<tr>
<td>Effort cost elasticity</td>
<td>$\gamma$</td>
<td>1.095</td>
</tr>
<tr>
<td>Scale of disutility</td>
<td>$\kappa_1$</td>
<td>0.771</td>
</tr>
<tr>
<td>Scale of R&amp;D cost</td>
<td>$\kappa_r$</td>
<td>0.050</td>
</tr>
<tr>
<td>Support for $\theta_1$</td>
<td>$\Theta^1$</td>
<td>2.025</td>
</tr>
<tr>
<td>Level of types</td>
<td>$\mu_\theta$</td>
<td>0.000</td>
</tr>
<tr>
<td>Type persistence</td>
<td>$p$</td>
<td>0.629</td>
</tr>
<tr>
<td>Initial intangibles</td>
<td>$q_0$</td>
<td>0.000</td>
</tr>
<tr>
<td>Initial R&amp;D stock</td>
<td>$r_0$</td>
<td>1.000</td>
</tr>
<tr>
<td>R&amp;D-type substitution</td>
<td>$\rho_{\theta r}$</td>
<td>1.451</td>
</tr>
<tr>
<td>Production externality</td>
<td>$\zeta$</td>
<td>0.033</td>
</tr>
</tbody>
</table>
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:

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

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

Same idea for wedges.
Young vs Old Firms

(a) Profit wedge

(b) Gross & Net R&D wedges

Policies converge to Pigouvian correction. Screening term for R&D ≤ 0 since $\rho_{\theta r} > \rho_{lr} = 1$ (net wedge ↑)
Marginal tax rate and R&D subsidy lower for higher productivity firms.

“No distortion at the top”
Optimal Simpler Policies
How Close can Simpler Policies Come?

Linear Policies

\[ T(\pi) = \tau_0 \pi \quad S(M) = s_0 M \]

Linear Policies with Interaction Term

\[ T(\pi, M) = (\tau_0 + \tau_1 M) \pi \quad S(M) = s_0 M \]

Heathcote-Storesletten-Violante (HSV) Policies

\[ T(\pi) = \tau_0 \pi - \tau_1 1 + \tau_2 \pi^{1+\tau_2} \quad S(M) = s_0 M - s_1 1 + s_2 M^{1+s_2} \]

HSV Policies with Interaction Term

\[ T(\pi, M) = (\tau_0 + \tau_3 M^{s_2}) \pi - \tau_1 1 + \tau_2 \pi^{1+\tau_2} \]

\[ S(M) = s_0 M - s_1 1 + s_2 M^{1+s_2} \]
# Optimal Simpler Policies

<table>
<thead>
<tr>
<th>Externality</th>
<th>Optimal Policy</th>
<th>Revenue (% of optimum)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Linear Policies</td>
<td>$T(\pi) = \tau_0 \pi$</td>
<td>$S(M) = s_0 M$</td>
</tr>
<tr>
<td>$\zeta = 0.033$</td>
<td>$\tau_0 = 0.40$</td>
<td>$s_0 = 0.35$</td>
</tr>
</tbody>
</table>

| B. Linear Policies with Interaction Term | $T(\pi, M) = (\tau_0 + \tau_1 M) \pi$ | $S(M) = s_0 M$ |
| C. Heathcote-Storesletten-Violante (HSV) Policies | $T(\pi) = \tau_0 \pi - \frac{\tau_1}{1+\tau_2} \pi^{1+\tau_2}$ | $S(M) = s_0 M - \frac{s_1}{1+s_2} M^{1+s_2}$ |
| $\zeta = 0.033$ | $\tau_0 = 0.71$ | $\tau_1 = 0.18$ | $\tau_2 = 0.18$ | $s_0 = 0.58$ | $s_1 = 0.25$ | $s_2 = 0.40$ | 37.2% |

| D. HSV Policies with Interaction Term | $T(\pi, M) = (\tau_0 + \tau_3 M^{s_2}) \pi - \frac{\tau_1}{1+\tau_2} \pi^{1+\tau_2}$ | $S(M) = s_0 M - \frac{s_1}{1+s_2} M^{1+s_2}$ |
| $\zeta = 0.033$ | $\tau_0 = 0.48$ | $\tau_1 = 0.24$ | $\tau_2 = 0.35$ | $\tau_3 = 0.40$ | $s_0 = 0.51$ | $s_1 = 0.17$ | $s_2 = 0.42$ | 49.8% |
Simpler Policies with Patent System as given: Revenue Gains

Externality ($\zeta$)

Relative Welfare

- Linear
- Linear + Cross
- HSV (Bounded)
- HSV (Bounded) + Cross
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 → Optimal to subsidize R&D investments.

- Asymmetric information could go other way in theory if R&D very complementary to firm productivity (↑ informational rents to firms).

- Revenue loss from restricted policies is large, but an HSV policy with interaction term between R&D and profits comes close.