Competing for Inventors

(joint with Michele Fornino)

Andrea Manera

May 3, 2021

Research Questions and Main Idea

We ask two main questions:

1. What are the boundaries of labor markets for *inventors?*
2. Does *competition for scarce inventors* between *different* product markets affect R&D allocation and productivity?

Across product markets, more market power gives:

- Higher *private* returns to R&D, demand for inventors
- Lower *social* returns to R&D, less growth per inventor

Complementary explanation for observed decline in R&D productivity

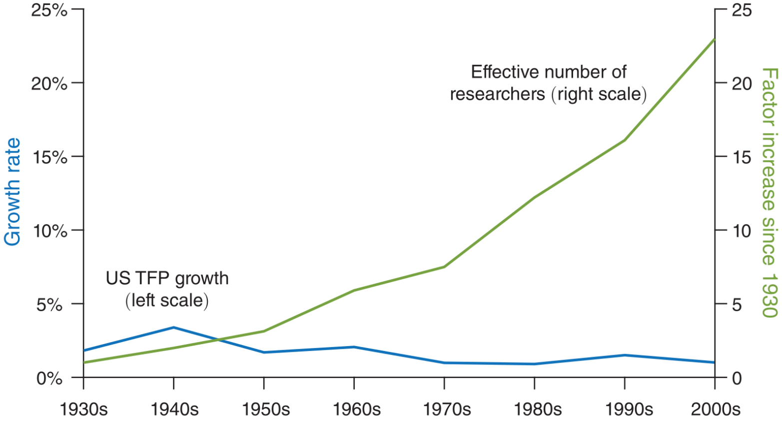
What we Do

1. Build data on inventors' flows across product markets
2. Preliminary facts:
    - Inventors have become more concentrated across product markets (all levels of NAICS)
    - Positive correlation between product market concentration and share of relevant inventors
3. Preliminary model:
    - Acemoglu and Akcigit (2012) + monopolistic firms (CES) in each product market
    - Markets with higher markup hire more inventors
    - Future: work on welfare consequences and policy

Plan of the Talk

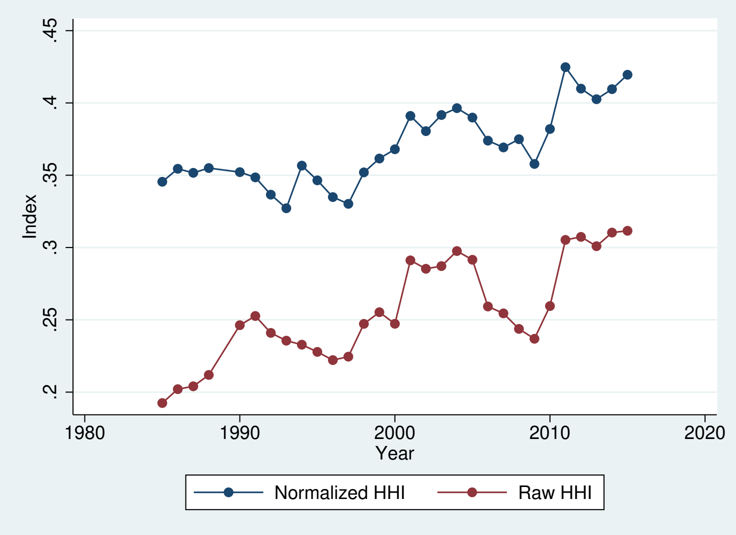
1. Motivation (most from our work) and Literature Review
2. Data construction
3. Model

Are Ideas Harder to Find?



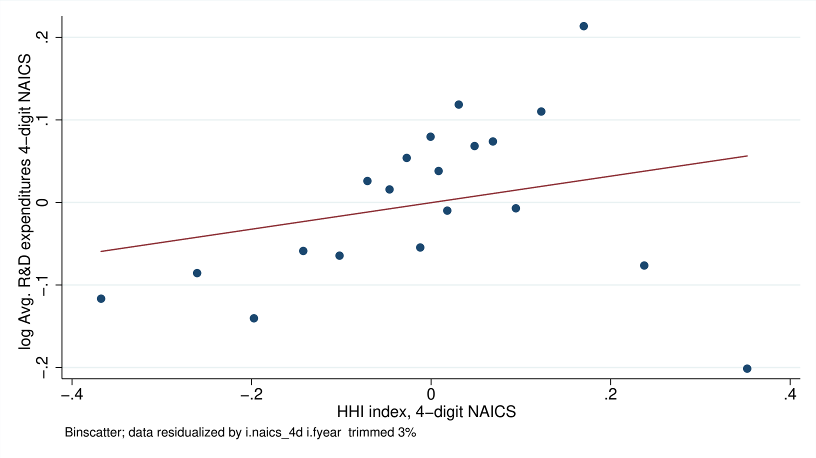
Source: Bloom et al. (2021)

Increased Concentration (4d NAICS)



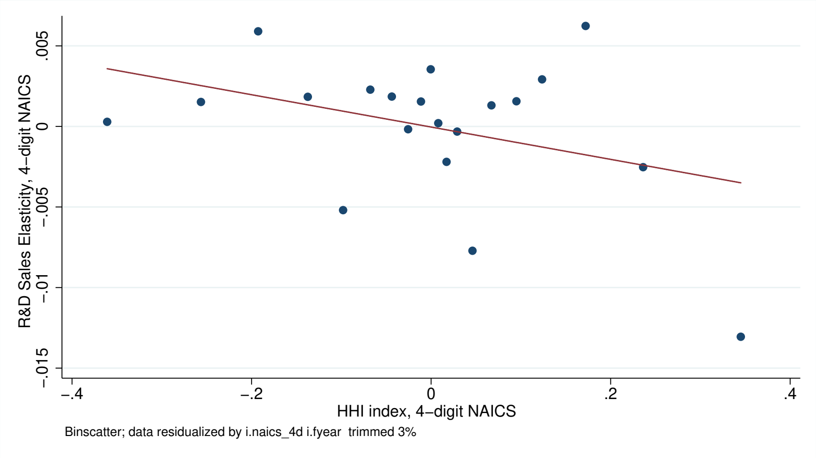
Source: Compustat

R&D spending on Concentration (4d NAICS)



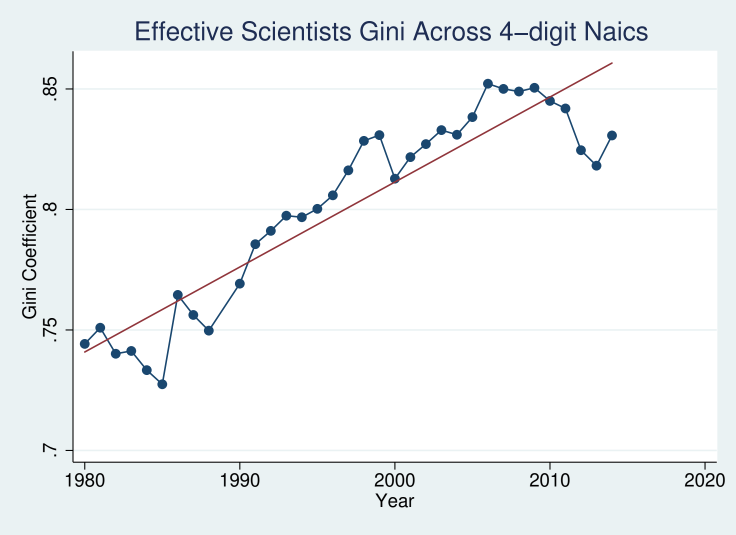
Source: Compustat

"Research Quotient" on Concentration



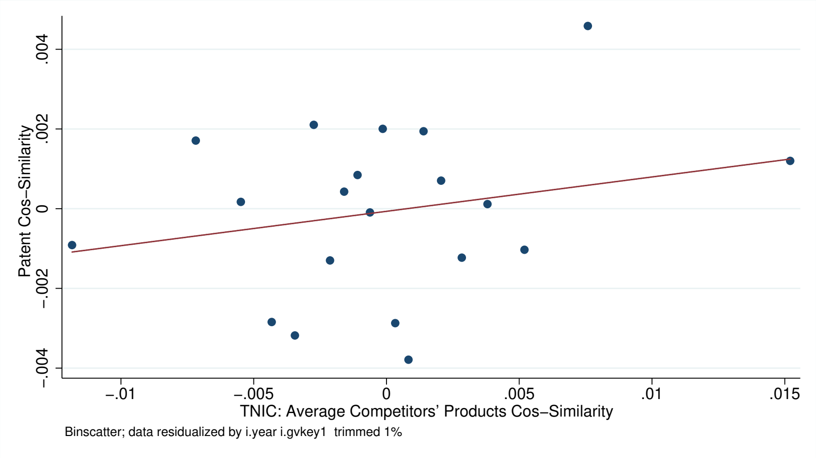
Source: Knott(2012), Compustat

Increased Concentration of Inventors



Source: Compustat and USPTO

Patent vs. Product Similarity among Firm Pairs



Source: Arts et al. (2021) and Hoberg and Phillips (2016)

Relevant Literature

- *Trends in innovation and R&D.*
  Akcigit and Kerr (2018), Akcigit and Ates (2020), **Bloom et al.** (2020, 2021), **Arora et al.** (2021)
- *Decreasing Competition*
  De Loecker et al. (2020), **Gutiérrez and Philippon** (2017, 2018), Grullon et al. (2019)
- *Competition and Innovation*
  Aghion et al. (2005, 2009, 2019), Gutiérrez and Philippon (2017), Autor et al. (2021), Hoberg and Phillips (2016)
- *Models of step-by-step innovation*
  Aghion, Harris, and Vickers (1997), Aghion et al. (2001), **Acemoglu and Akcigit (2012)**

How Do We Differ?

Theory Literature:

- R&D activity is (usually) non-rival
- Focus on homogeneous product market
- Look at competition and innovation *within* product markets

Empirical Literature:

- Focus on patents and citations

This paper:

- R&D is rival (inventors are scarce)
- Competition and innovation across different product markets
- Focus on inventors (although through patents)

Empirics Objectives

- Understand boundary of markets for inventors
  - Identify "knowledge markets" as sets of product markets that hire the same type of inventors
  - Use patent data to build a network of flows of inventors across sectors
  - Identify connected sectors maximizing network's modularity
- Look *within knowledge markets* to see how product markets' share of inventors relate to concentration

Data Sources

- Compustat (firm-year):
  - firm identifiers (*gvkey*), product market (*NAICS* 2007), and sales, build competition measures;
- USPTO (patent-year) :
  - patent citation and disambiguated inventor id's, 1975-present;
- DISCERN, Arora et al. (2021) (patent-inventors-year): crosswalk of patent IDs to ultimate-owner gvkey in Compustat 1980-2015;

Dataset Structure

Patent ID	Inventor ID	UO gvkey (assignee)	NAICS	Year
US00001	00001-1	1010	111	1980
US00001	00001-2	1010	111	1980
US00002	00001-1	1010	111	1980
US00003	00001-1	1044	1121	1982

Dataset Structure

Patent ID	Inventor ID	UO gvkey (assignee)	NAICS	Year
US00001	00001-1	1010	111	1980
US00001	00001-2	1010	111	1980
US00002	00001-1	1010	111	1980
US00003	00001-1	1044	1121	1982

$\Downarrow \quad \Downarrow \quad \Downarrow$

Inventor ID	Gvkey 1	Gvkey 2	NAICS 1	NAICS 2	Year
00001-1	1010	1044	111	112	1982

"Knowledge Markets"

- *Knowledge Market:* set of NAICS (product markets) that employ the same type of inventors
  - To capture similar *required knowledge* to innovate
- From data, undirected network:
  - NAICS (3 or 4 digit) as nodes
  - *Normalized effective inventor flows* as edge weights, W

"Effective inventors"

- *Effective inventors:* "Productivity-adjusted" inventors:

$$\tilde{\text{Inventors}}_{f,t} = \sum_{i \in \{\text{inventors in firm } f\}} \left[\sum_{p \in \{\text{inventor} i\text{'s patents}\}} \frac{1}{N_p} \right],$$

with N_p the number of inventors who worked on patent p

- *Normalized effective inventor flow* between NAICS i and j in period t is:

$$W_{ij,t} = \frac{\tilde{\text{Inventors}}_{i \to j,t} + \tilde{\text{Inventors}}_{j \to i,t}}{\tilde{\text{Inventors}}_{i,t} + \tilde{\text{Inventors}}_{j,t}}$$

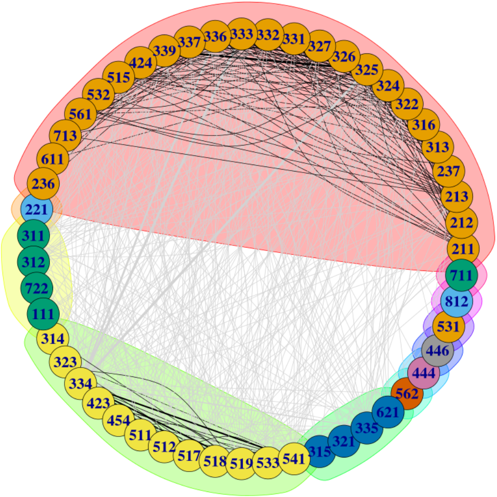
Detecting Knowledge Markets

- Run a weighted community detection algorithm:
  - Maximizes modularity of the network
  - Finds N communities to maximize:

$$Q = \sum_{c=1}^{N} \left[e_{cc} - \left(\sum_{j} e_{cj} \right)^2 \right],$$

    where e_{cj} is the (weighted) fraction of edges that have one end in community c and the other in community j.
  - "How much more the community is connected internally than externally".

Markets at 3-digit NAICS



Features of Flows and Knowledge Markets

- Many connections across product markets even at 3 digits!
- Same inventors are employed by firms in *highly different* product markets
- Broad communities
- Reasonable?
  - Green Cluster collects "Food and Agriculture": Crop Production, Food Manufacturing and Services, Beverage and Tobacco;
  - Orange Cluster is mostly "Mining" and "Heavy Industry": e.g. Petroleum and Coal Products, Chemical, Machinery Manufacturing;
  - Yellow Cluster collects "Communications", "Electronics" and "Publishing":e.g. Computer and Electronic Products, Telecommunications, Data Processing;

Ideal Regression

Ideal causal regression:

$$\text{Share}_{skt} = f_s + f_t + f_k + \beta \text{Conc}_{s,t-1} + \varepsilon_{st},$$

- s is the product market, NAICS 3- or 4-digit ("sector")
- k is a "knowledge" market
- Share_{skt}: share of inventors in market k employed by s
- $\text{Conc}_{s,t-1}$ is a lagged concentration measure: HHI at relevant NAICS, top firms' sale share.

Specification

We run:

$$\hat{\text{Share}}_{skt} = f_s + f_t + f_k + \beta \text{Conc}_{s,t-1} + \varepsilon_{st},$$

with:

$$\hat{\text{Share}}_{skt} \equiv \text{Share}_{skt} - \frac{N_{st}}{N_{kt}},$$

- Deviation from equal distribution (note: f_s, f_k should capture other technological factors)
- N_{st} number of firms in the product market s
- N_{kt} is the number of firms that compete for scientists k.
- Corrects for mechanical downward bias: more concentrated sectors have also less firms.

Results

	(1) Share	(2) $\widehat{\text{Share}}$	(3) $\widehat{\text{Share}}$	(4) $\widehat{\text{Share}}$	(5) $\widehat{\text{Share}}$	(6) $\widehat{\text{Share}}$
L.HHI index in 3-digit NAICS	0.004	0.027***				
	(0.007)	(0.008)				
L.Normalized HHI index in 3-digit NAICS			0.021**			
			(0.007)			
L.Top 3 firms sale share in 3-digit NAICS				0.074***		
				(0.015)		
L.Top 5 firms sale share in 3-digit NAICS					0.102***	
					(0.022)	
L.Top 10 firms sale share in 3-digit NAICS						0.087*
						(0.035)
Constant	0.147***	-0.011**	-0.003	-0.056***	-0.085***	-0.079*
	(0.003)	(0.004)	(0.003)	(0.012)	(0.019)	(0.032)
Observations	1581	1581	1492	1581	1581	1581

Standard errors in parentheses. Fixed effects for sector, year and knowledge market included in all specifications.
+ $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Thoughts

- Note: the correlation is mostly driven by the *contemporaneous concentration*.
  - There is scope for our hypothesis!
- Effects are not insignificant:
  - e.g. a 1pp increase in top-5 product market share gives 0.1pp more scientist market *share*
  - Recall: scientist market spans multiple sectors
- Effects are likely non-linear, more analysis is warranted!

Model Objectives

- Explain our intuition on decrease in competition driving lower growth through misallocation
- Build a model that generates a negative relation between competition and inventor demand
  - Version of Acemoglu and Akcigit (2012)
  - Two differences:
    - Imperfect substitution of goods within markets $\Rightarrow$No Bertrand competition
    - Look across, rather than within product markets
    - Separate markets for production and research workers
- Issue: concentration not a good measure of competition in this model
- But positive relation between markup and inventor demand!
- Suggestions appreciated

Final Good

- Standard representative household
- Consumes the Cobb Douglas aggregate of intermediate goods $Q(k)$:

$$\log Y = \int_0^1 \log Q(k)\ d\ k$$

- Demand for intermediates:

$$p_k = \frac{Y}{Q_k}.$$

- Y is the numeraire.

Intermediates

- Each intermediate k is a CES aggregate of two differentiated goods:

$$Q_k = \left[q_{i,k}^{\frac{\varepsilon-1}{\varepsilon}} + q_{j,k}^{\frac{\varepsilon-1}{\varepsilon}} \right]^{\frac{\varepsilon}{\varepsilon-1}}$$

- Thus, firm i's problem:

$$\max_{q_{i,k}} p_{i,k} q_{i,k} - \frac{w}{\lambda^{n_{i,k}}} q_i$$

$$\text{s.t.}\, p_{i,k} = p_k \left[\frac{Q_k}{q_{i,k}} \right]^{\frac{1}{\varepsilon}}$$

,

- $\lambda^{-n_{i,k}}$ is firm i's productivity
- $n_{i,k} \in \mathbb{N}$ is the state of technology of firm i
- $\varepsilon > 1$ gives elasticity of substitution between varieties ("mkt power")

Intermediates cont.

- Solution:

$$p_{i,k} = \frac{\varepsilon}{\varepsilon - 1}\frac{w}{\lambda^{n_{i,k}}}, \quad q_{i,k} = \left[\frac{\varepsilon - 1}{\varepsilon}\frac{\lambda^{n_{i,k}}}{w} p_k\right]^{\varepsilon} Q_k$$

- Profits:

$$\Pi_{i,k} = \frac{1}{\varepsilon - 1}\lambda^{-n_{i,k}}\left[1 + \left[\lambda^{n_{j,k} - n_{i,k}}\right]^{\varepsilon - 1}\right]^{-\frac{\varepsilon}{\varepsilon - 1}} w Q_k$$

- Using $Y = p_k Q_k$:

$$\Pi_i = \frac{1}{\varepsilon}\left[1 + \left[\lambda^{n_j - n_i}\right]^{\varepsilon - 1}\right]^{-1} Y$$

Markov Property

- Profits:

$$\Pi_i = \frac{1}{\varepsilon}\left[1 + \left[\lambda^{n_j - n_i}\right]^{\varepsilon - 1}\right]^{-1} Y$$

  depend only on the technology gap wrt rival: $m_{i,k} \equiv n_{j,k} - n_{i,k}$
- Firm's dynamic problem is Markovian with state m
- Use transformed variables $x = X/Y$

R&D

- Technology state n can be improved by one step with success rate x:

$$x = \left(\frac{h}{\alpha}\right)^{\frac{1}{\gamma}}$$

  - h is number of inventors hired
  - $\gamma > 1$ parametrizes how hard research is

- R&D cost:

$$hw^{\text{R\&D}} = \alpha x^{\gamma} w^{\text{R\&D}}$$

R&D Choice

- Firm value:

$$\rho v_m - \dot{v}_m = \max_{x_m} \quad \pi_m - \alpha x_m^{\gamma} w^{\text{R\&D}} + x_m(v_{m+1} - v_m)$$
$$x_{-m}(v_{m-1} - v_m) + \eta_m(v_0 - v_m)$$

  - x_m is own research
  - x_{-m} is research by rival
  - η_m is spillover intensity

Distribution of Technology Gaps

- Let x_m^* be the optimal R&D intensity
- LOM of distribution across m-gap markets:

$$\dot{\mu}_m^* = -\left(x_m^* + x_{-(m)}^* + \eta_m\right)\mu_m^* + x_{(m-1)}^*\mu_{m-1}^* + x_{-(m+1)}^*\mu_{m+1}^*, \quad m \geq 1,$$
$$\dot{\mu}_0^* = -2x_0^*\mu_0^* + \sum_{n\geq 1}\eta_n\mu_n^* + x_{-1}^*\mu_1^*.$$

- In matrix form: $\dot{\mu} = A^T\mu$

Labor markets

We assume:

- Supply of production labor perfectly elastic
- Supply of inventors fixed at $L^{R\&D}$
- Inventors' market clearing:

$$L^{\mathrm{R\&D}} = \alpha \sum \mu_n^* \left[\left(x_n^* \left(w^{\mathrm{R\&D}} \right) \right)^\gamma + \left(x_{-n}^* \left(w^{\mathrm{R\&D}} \right) \right)^\gamma \right]$$

- In the future, endogenize supply with human capital investments

BGP

A BGP equilibrium of this model is a set of normalized $\{v_m\}_0^{\infty}$, research intensities, $\{x_n^*\}_{-\infty}^{+\infty}$ wages w^{prod}, $w^{\text{R\&D}}$, and a stationary distribution $\{\mu_m^*\}_0^{\infty}$ such that:

1. Values are determined by the solution to the HJB with $\dot{v}_m = 0$
2. Firm choose x_n^* optimally
3. The stationary distribution solves the KFE $\dot{\mu}^* = 0$
4. Labor markets clear
5. Factor shares are constant

Growth

- Growth is then:

$$g = \frac{\log\left(\frac{\varepsilon-1}{\varepsilon}\lambda\right)\left[\int_0^1 \left(n_k\left(t+\Delta t\right) - n_k\left(t+\Delta t\right)\right)\mathrm{d}k\right]}{\Delta t}$$

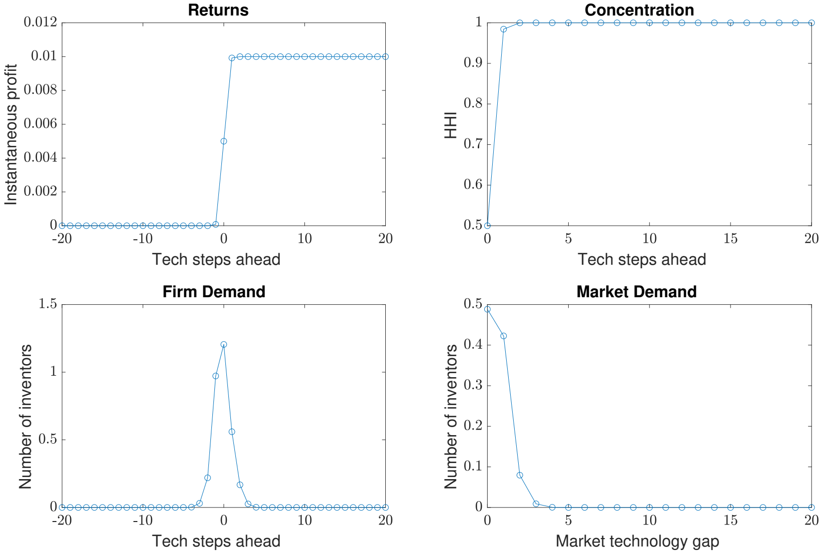
$$= \log\left(\frac{\varepsilon-1}{\varepsilon}\lambda\right)\left[2x_0^*\mu_0^* + \sum_{n>0}\mu_n^* x_n^*\right]$$

- Depends only on $x^*$and μ^*
- Directly increasing in ε for given research intensity and distribution

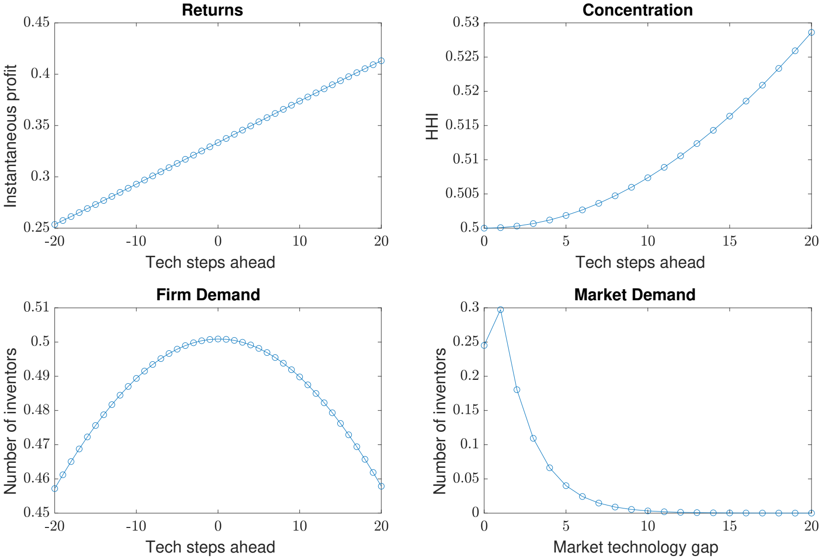
Numerical Exercise

- Parameters from Acemoglu and Akcigit (2012):
  - $\gamma \approx 3.3$ (midrange of estimates, Kortum (1993)), $\alpha \approx 1000$
  - $\lambda = 1.05$
  - $\rho = 5\%$
- Set maximum technology steps $\bar{n} = 100$
- Features of the model with high ($\varepsilon = 100$) and low ($\varepsilon = 1.5$) substitution

Equilibrium with $\varepsilon = 100$



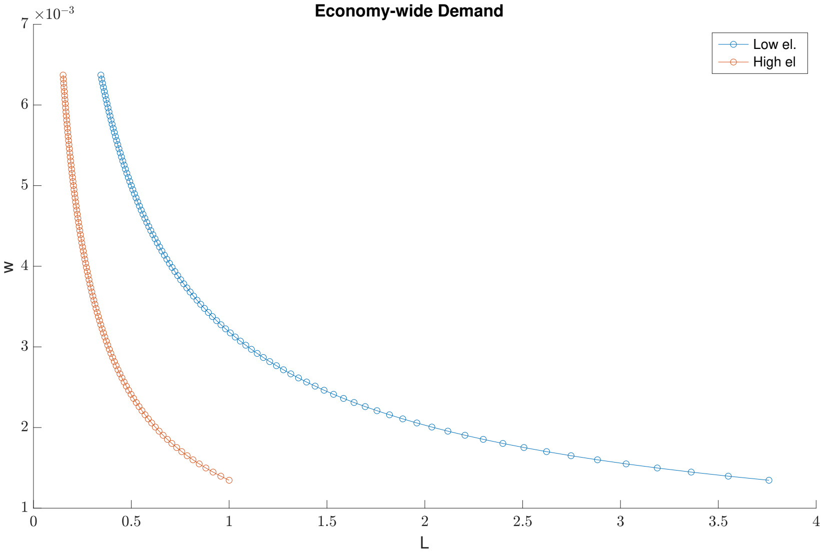
Equilibrium with $\varepsilon = 1.5$



Intuition

- Wih high substitution, even a small tech gap leads to monopoly by most productive firm
- With lower substitution, differentiation can support larger tech gaps
- More substitution gives more concentration cet. par.
- But aggregate demand for inventors (next slide):
    - Higher in market with larger markup
    - Where more of the returns from R&D are appropiated by the firm

Aggregate Labor Demand



Conjectures

- When putting together the two markets, less competitive will attract more.
- This will be regardless of consumer preferences (still C-D).
- Potential for misallocation
- Lower growth and productivity of R&D than overall more competitive economy
- "Issue": competition negatively related to concentration

What Next?

- Find an event for better empirics
- Expand the theory to verify conjectures
    - Adding entry to each good market so that markup does become a function of concentration
- Include human capital and specific inventor types
    - Look at new inventors over time
- Quantitative exploration