The Impact of Academic Patenting on the Rate, Quality, and Direction of (Public) Research Output

Pierre Azoulay[1] **Waverly Ding**[2] **Toby Stuart**[3]

[1]Sloan School of Management
MIT & NBER
pazoulay@mit.edu

[2]Haas School of Business
University of California — Berkeley

[3]Graduate School of Business
Harvard University

January 23, 2007 — Northwestern University

Outline

1. Motivation(s)
2. Methodology
   - Problems with existing approaches
   - Selection on observables with staggered treatment decisions
   - Implementing IPTCW estimation
3. Data & Measurement
   - Data sources
   - Measuring "patentability"
   - Descriptive statistics
4. Results
   - The determinants of selection into patenting
   - The impact of academic patenting on the rate of publications
   - The impact of academic patenting on the quality of publications
   - The impact of academic patenting on the content of publications
5. Caveats, Summary & Future Directions

Outline

1. Motivation(s)
2. Methodology
   - Problems with existing approaches
   - Selection on observables with staggered treatment decisions
   - Implementing IPTCW estimation
3. Data & Measurement
   - Data sources
   - Measuring "patentability"
   - Descriptive statistics
4. Results
   - The determinants of selection into patenting
   - The impact of academic patenting on the rate of publications
   - The impact of academic patenting on the quality of publications
   - The impact of academic patenting on the content of publications
5. Caveats, Summary & Future Directions

Outline

1. Motivation(s)
2. Methodology
   - Problems with existing approaches
   - Selection on observables with staggered treatment decisions
   - Implementing IPTCW estimation
3. Data & Measurement
   - Data sources
   - Measuring "patentability"
   - Descriptive statistics
4. Results
   - The determinants of selection into patenting
   - The impact of academic patenting on the rate of publications
   - The impact of academic patenting on the quality of publications
   - The impact of academic patenting on the content of publications
5. Caveats, Summary & Future Directions

Outline

1. Motivation(s)
2. Methodology
   - Problems with existing approaches
   - Selection on observables with staggered treatment decisions
   - Implementing IPTCW estimation
3. Data & Measurement
   - Data sources
   - Measuring "patentability"
   - Descriptive statistics
4. Results
   - The determinants of selection into patenting
   - The impact of academic patenting on the rate of publications
   - The impact of academic patenting on the quality of publications
   - The impact of academic patenting on the content of publications
5. Caveats, Summary & Future Directions

Outline

1. Motivation(s)
2. Methodology
   - Problems with existing approaches
   - Selection on observables with staggered treatment decisions
   - Implementing IPTCW estimation
3. Data & Measurement
   - Data sources
   - Measuring "patentability"
   - Descriptive statistics
4. Results
   - The determinants of selection into patenting
   - The impact of academic patenting on the rate of publications
   - The impact of academic patenting on the quality of publications
   - The impact of academic patenting on the content of publications
5. Caveats, Summary & Future Directions

Outline

1. Motivation(s)
2. Methodology
   - Problems with existing approaches
   - Selection on observables with staggered treatment decisions
   - Implementing IPTCW estimation
3. Data & Measurement
   - Data sources
   - Measuring "patentability"
   - Descriptive statistics
4. Results
   - The determinants of selection into patenting
   - The impact of academic patenting on the rate of publications
   - The impact of academic patenting on the quality of publications
   - The impact of academic patenting on the content of publications
5. Caveats, Summary & Future Directions

Academic Entrepreneurship in the Life Sciences

A Research Agenda

- Does patenting accelerate or hinder faculty patenters rate of production of public scientific outputs?
- Does patenting directly influence the quality or content of the subsequent-to-the-patent research topics investigated by the scientist?
- Does patenting hinder the flow of information in the scientific community, thus initiating negative spillovers that aggregate to impede scientific progress? (Murray & Stern 2005)
- Does patenting alter the career trajectories of patenters and their associates (e.g., graduate students, post-doctoral fellows, and co-authors)?

Academic Entrepreneurship in the Life Sciences

A Research Agenda

- Does patenting accelerate or hinder faculty patenters rate of production of public scientific outputs?
- Does patenting directly influence the quality or content of the subsequent-to-the-patent research topics investigated by the scientist?
- Does patenting hinder the flow of information in the scientific community, thus initiating negative spillovers that aggregate to impede scientific progress? (Murray & Stern 2005)
- Does patenting alter the career trajectories of patenters and their associates (e.g., graduate students, post-doctoral fellows, and co-authors)?

Academic Entrepreneurship in the Life Sciences

A Research Agenda

- Does patenting accelerate or hinder faculty patenters rate of production of public scientific outputs?
- Does patenting directly influence the quality or content of the subsequent-to-the-patent research topics investigated by the scientist?
- Does patenting hinder the flow of information in the scientific community, thus initiating negative spillovers that aggregate to impede scientific progress? (Murray & Stern 2005)
- Does patenting alter the career trajectories of patenters and their associates (e.g., graduate students, post-doctoral fellows, and co-authors)?

Academic Entrepreneurship in the Life Sciences

A Research Agenda

- Does patenting accelerate or hinder faculty patenters rate of production of public scientific outputs?
- Does patenting directly influence the quality or content of the subsequent-to-the-patent research topics investigated by the scientist?
- Does patenting hinder the flow of information in the scientific community, thus initiating negative spillovers that aggregate to impede scientific progress? (Murray & Stern 2005)
- Does patenting alter the career trajectories of patenters and their associates (e.g., graduate students, post-doctoral fellows, and co-authors)?

Academic Entrepreneurship in the Life Sciences

A Research Agenda

- Does patenting accelerate or hinder faculty patenters rate of production of public scientific outputs?
- Does patenting directly influence the quality or content of the subsequent-to-the-patent research topics investigated by the scientist?
- Does patenting hinder the flow of information in the scientific community, thus initiating negative spillovers that aggregate to impede scientific progress? (Murray & Stern 2005)
- Does patenting alter the career trajectories of patenters and their associates (e.g., graduate students, post-doctoral fellows, and co-authors)?

Academic Entrepreneurship in the Life Sciences

A Research Agenda

- Does patenting accelerate or hinder faculty patenters rate of production of public scientific outputs?
- Does patenting directly influence the quality or content of the subsequent-to-the-patent research topics investigated by the scientist?
- Does patenting hinder the flow of information in the scientific community, thus initiating negative spillovers that aggregate to impede scientific progress? (Murray & Stern 2005)
- Does patenting alter the career trajectories of patenters and their associates (e.g., graduate students, post-doctoral fellows, and co-authors)?

Treatment Effects in Strategy Research

Typical Specification

$$y_{it} = \beta_0 + \beta_1' X_{it} + \beta_2 TREAT_{it} + \varepsilon_{it}$$

- Estimating the effect of "Blah" on "Performance"
  - "Blah" = Firing the CEO — y = Stock Price or Acctng. Profitability
  - "Blah" = Exporting — y = TFP
  - "Blah" = Pro-Pub — y = R&D Productivity among Pharma Firms
- What these settings have in common:
  1. Panel data structure — Variation in treatment both between and within units
  2. *TREAT* is a choice variable, and adoption is staggered over time
  3. Often, we have no good instruments
  4. Lagged dependent variable predicts selection into treatment
- Traditional approach (in the strategy/management literature)
  - Fixed effects estimation — Almost certainly wrong given (4)
  - Dynamic Panels — Problematic for reasons explained below

Treatment Effects in Strategy Research

Typical Specification

$$y_{it} = \beta_0 + \beta_1' X_{it} + \beta_2 TREAT_{it} + \varepsilon_{it}$$

- Estimating the effect of "Blah" on "Performance"
  - "Blah" = Firing the CEO — y = Stock Price or Acctng. Profitability
  - "Blah" = Exporting — y = TFP
  - "Blah" = Pro-Pub — y = R&D Productivity among Pharma Firms
- What these settings have in common:
  1. Panel data structure — Variation in treatment both between and within units
  2. *TREAT* is a choice variable, and adoption is staggered over time
  3. Often, we have no good instruments
  4. Lagged dependent variable predicts selection into treatment
- Traditional approach (in the strategy/management literature)
  - Fixed effects estimation — Almost certainly wrong given (4)
  - Dynamic Panels — Problematic for reasons explained below

Treatment Effects in Strategy Research

Typical Specification

$$y_{it} = \beta_0 + \beta_1' X_{it} + \beta_2 TREAT_{it} + \varepsilon_{it}$$

- Estimating the effect of "Blah" on "Performance"
  - "Blah" = Firing the CEO — y = Stock Price or Acctng. Profitability
  - "Blah" = Exporting — y = TFP
  - "Blah" = Pro-Pub — y = R&D Productivity among Pharma Firms
- What these settings have in common:
  1. Panel data structure — Variation in treatment both between and within units
  2. *TREAT* is a choice variable, and adoption is staggered over time
  3. Often, we have no good instruments
  4. Lagged dependent variable predicts selection into treatment
- Traditional approach (in the strategy/management literature)
  - Fixed effects estimation — Almost certainly wrong given (4)
  - Dynamic Panels — Problematic for reasons explained below

Treatment Effects in Strategy Research

Typical Specification

$$y_{it} = \beta_0 + \beta_1' X_{it} + \beta_2 TREAT_{it} + \varepsilon_{it}$$

- Estimating the effect of "Blah" on "Performance"
  - "Blah" = Firing the CEO   y = Stock Price or Acctng. Profitability
  - "Blah" = Exporting   y = TFP
  - "Blah" = Pro-Pub   y = R&D Productivity among Pharma Firms
- What these settings have in common:
  1. Panel data structure — Variation in treatment both between and within units
  2. *TREAT* is a choice variable, and adoption is staggered over time
  3. Often, we have no good instruments
  4. Lagged dependent variable predicts selection into treatment
- Traditional approach (in the strategy/management literature)
  - Fixed effects estimation — Almost certainly wrong given (4)
  - Dynamic Panels — Problematic for reasons explained below

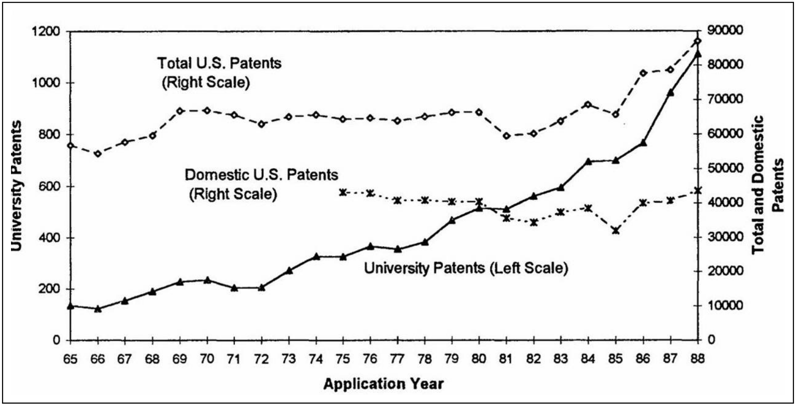
Top 15 Acad. Institutions, Stock of Patents, 1976-2004

1	Massachusetts Institute of Technology	2,650
2	University of California – Berkeley	2,155
3	National Institutes of Health	1,988
4	Stanford University	1,435
5	California Institute of Technology	1,421
6	Wisconsin Alumni Research Foundation	1,177
7	Johns Hopkins University	1,053
8	University of Florida	865
9	University of California – San Francisco	832
10	University of Michigan	771
11	University of Minnesota	764
12	Massachusetts General Hospital	757
13	Cornell University	711
14	Iowa State University Research Foundation	709
15	University of Pennsylvania	671

Source: Authors' Tabulations

Mapping the Rise of Academic Entrepreneurship

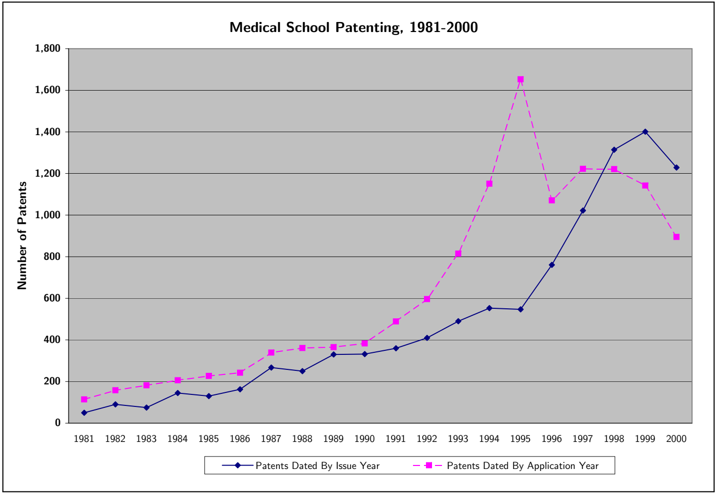
Academic vs. Industry Patents



Source: Henderson, Jaffe and Trajtenberg (1998)

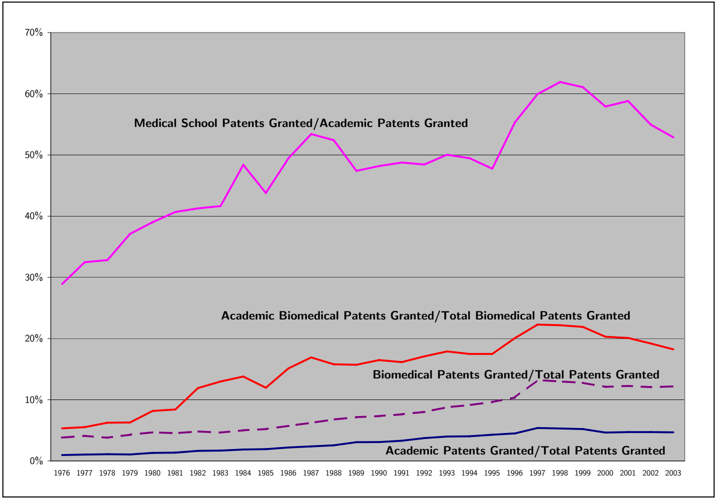
Mapping the Rise of Academic Entrepreneurship (Cont'd)

Concentration in the Life Sciences



Mapping the Rise of Academic Entrepreneurship (Cont'd)

Concentration in the Life Sciences



Two Questions

- Which scientists patent, and when do they patent?
- What is the impact of patenting on [public] research output?
  - Rate of publications
  - Quality of publications
  - Commercial content of publications

Two Questions

- Which scientists patent, and when do they patent?
- What is the impact of patenting on [public] research output?
  - Rate of publications
  - Quality of publications
  - Commercial content of publications

Two Questions

- Which scientists patent, and when do they patent?
- What is the impact of patenting on [public] research output?
  - Rate of publications
  - Quality of publications
  - Commercial content of publications

Literature Review

- Many studies of academic patenting
  - Single university studies (Agrawal & Henderson 2002)
  - Ethnographies (Owen Smith & Powell 2001)
  - Analyses of X-sectional surveys (Stephan et al. 2006)
- What seems to matter:
  - Academic patenters are more likely to be "elite"
  - Important differences across fields in the propensity to patent and in underlying motivations for patenting (life sciences vs. engineering)
  - Peers, institutional environment (TLO,. . .)
- Additional insights to be gained from complete career histories for a random sample of scientists

Literature Review

- Many studies of academic patenting
  - Single university studies (Agrawal & Henderson 2002)
  - Ethnographies (Owen Smith & Powell 2001)
  - Analyses of X-sectional surveys (Stephan et al. 2006)
- What seems to matter:
  - Academic patenters are more likely to be "elite"
  - Important differences across fields in the propensity to patent and in underlying motivations for patenting (life sciences vs. engineering)
  - Peers, institutional environment (TLO,...)
- Additional insights to be gained from complete career histories for a random sample of scientists

Literature Review

- Many studies of academic patenting
  - Single university studies (Agrawal & Henderson 2002)
  - Ethnographies (Owen Smith & Powell 2001)
  - Analyses of X-sectional surveys (Stephan et al. 2006)

- What seems to matter:
  - Academic patenters are more likely to be "elite"
  - Important differences across fields in the propensity to patent and in underlying motivations for patenting (life sciences vs. engineering)
  - Peers, institutional environment (TLO,. . .)

- Additional insights to be gained from complete career histories for a random sample of scientists

Literature Review

- Many studies of academic patenting
  - Single university studies (Agrawal & Henderson 2002)
  - Ethnographies (Owen Smith & Powell 2001)
  - Analyses of X-sectional surveys (Stephan et al. 2006)
- What seems to matter:
  - Academic patenters are more likely to be "elite"
  - Important differences across fields in the propensity to patent and in underlying motivations for patenting (life sciences vs. engineering)
  - Peers, institutional environment (TLO,...)
- Additional insights to be gained from complete career histories for a random sample of scientists

Patenting, Publishing, and Academic Incentives

Old debate on the impact of commercial activities on the rate and direction of scientific progress

- Patents violate the "norm of commonality" in science (Merton 1942)
- Vannevar Bush: *"the perverse law governing research. . . that applied research invariably drives out pure."*

But:

- Scientific reputation is critical to ability to capitalize on intellectual property
  - *ex post* search, screening, and contracting problems in the market for ideas
- Patent application often incidental to the research — co-occurring outputs or *"paper/patent pairs"* (Murray 2002)
- Within-scientist economies of scope
  - Knowledge benefits: access to new social networks; exposure to new ideas
  - Pecuniary benefits: additional sources of financing; access to capital equipment

Patenting, Publishing, and Academic Incentives

Old debate on the impact of commercial activities on the rate and direction of scientific progress

- Patents violate the "norm of commonality" in science (Merton 1942)
- Vannevar Bush: *"the perverse law governing research... that applied research invariably drives out pure."*

But:

- Scientific reputation is critical to ability to capitalize on intellectual property
  - *ex post* search, screening, and contracting problems in the market for ideas
- Patent application often incidental to the research — co-occurring outputs or *"paper/patent pairs"* (Murray 2002)
- Within-scientist economies of scope
  - Knowledge benefits: access to new social networks; exposure to new ideas
  - Pecuniary benefits: additional sources of financing; access to capital equipment

Patenting, Publishing, and Academic Incentives

Old debate on the impact of commercial activities on the rate and direction of scientific progress

- Patents violate the "norm of commonality" in science (Merton 1942)
- Vannevar Bush: *"the perverse law governing research... that applied research invariably drives out pure."*

But:

- Scientific reputation is critical to ability to capitalize on intellectual property
    - *ex post* search, screening, and contracting problems in the market for ideas
- Patent application often incidental to the research — co-occurring outputs or *"paper/patent pairs"* (Murray 2002)
- Within-scientist economies of scope
    - Knowledge benefits: access to new social networks; exposure to new ideas
    - Pecuniary benefits: additional sources of financing; access to capital equipment

Patenting, Publishing, and Academic Incentives

Old debate on the impact of commercial activities on the rate and direction of scientific progress

- Patents violate the "norm of commonality" in science (Merton 1942)
- Vannevar Bush: *"the perverse law governing research... that applied research invariably drives out pure."*

But:

- Scientific reputation is critical to ability to capitalize on intellectual property
  - *ex post* search, screening, and contracting problems in the market for ideas
- Patent application often incidental to the research — co-occurring outputs or *"paper/patent pairs"* (Murray 2002)
- Within-scientist economies of scope
  - Knowledge benefits: access to new social networks; exposure to new ideas
  - Pecuniary benefits: additional sources of financing; access to capital equipment

Contributions

- Data
  - Stratified random sample of 3,862 scientists
  - Matched employee/employer dataset with individual-level measures of output
  - Rich set of covariates
  - Measuring the effect of commercial activities on the direction of scientific progress (rather than just the rate)
- Methodology
  - Hazard models to examine propensity of patenting
  - Novel approach to the selection problem:
    Inverse Probability of Treatment and Censoring Weighted (IPTCW) estimation

Contributions

- Data
  - Stratified random sample of 3,862 scientists
  - Matched employee/employer dataset with individual-level measures of output
  - Rich set of covariates
  - Measuring the effect of commercial activities on the direction of scientific progress (rather than just the rate)
- Methodology
  - Hazard models to examine propensity of patenting
  - Novel approach to the selection problem: Inverse Probability of Treatment and Censoring Weighted (IPTCW) estimation

Contributions

- Data
  - Stratified random sample of 3,862 scientists
  - Matched employee/employer dataset with individual-level measures of output
  - Rich set of covariates
  - Measuring the effect of commercial activities on the direction of scientific progress (rather than just the rate)
- Methodology
  - Hazard models to examine propensity of patenting
  - Novel approach to the selection problem:
    Inverse Probability of Treatment and Censoring Weighted (IPTCW) estimation

Preview of Results

- Self-selection: patenters are more successful scientists, and even more recently successful scientists
- Patenting appears to complement, not substitute, publication output
  - The elasticity of publication count with respect to applying for a patent lies between .195 (fixed effect estimate) and .235 (IPTCW estimate)
- No apparent effect on the *quality* of publication output
  - Order of Authorship
  - Average *Journal Impact Factor* (JIF)
- But genuine impact on the content of publications
  - Patenting entails more coauthored pubs with scientists in industry
  - Patenting increases subsequent patentability
  - Patenting increases the share of articles in journals in which relatively more industry scientists publish

Preview of Results

- Self-selection: patenters are more successful scientists, and even more recently successful scientists
- Patenting appears to complement, not substitute, publication output
  - The elasticity of publication count with respect to applying for a patent lies between .195 (fixed effect estimate) and .235 (IPTCW estimate)
- No apparent effect on the *quality* of publication output
  - Order of Authorship
  - Average *Journal Impact Factor* (JIF)
- But genuine impact on the content of publications
  - Patenting entails more coauthored pubs with scientists in industry
  - Patenting increases subsequent patentability
  - Patenting increases the share of articles in journals in which relatively more industry scientists publish

Preview of Results

- Self-selection: patenters are more successful scientists, and even more recently successful scientists
- Patenting appears to complement, not substitute, publication output
  - The elasticity of publication count with respect to applying for a patent lies between .195 (fixed effect estimate) and .235 (IPTCW estimate)
- No apparent effect on the *quality* of publication output
  - Order of Authorship
  - Average *Journal Impact Factor* (JIF)
- But genuine impact on the content of publications
  - Patenting entails more coauthored pubs with scientists in industry
  - Patenting increases subsequent patentability
  - Patenting increases the share of articles in journals in which relatively more industry scientists publish

Preview of Results

- Self-selection: patenters are more successful scientists, and even more recently successful scientists
- Patenting appears to complement, not substitute, publication output
  - The elasticity of publication count with respect to applying for a patent lies between .195 (fixed effect estimate) and .235 (IPTCW estimate)
- No apparent effect on the *quality* of publication output
  - Order of Authorship
  - Average *Journal Impact Factor* (JIF)
- But genuine impact on the content of publications
  - Patenting entails more coauthored pubs with scientists in industry
  - Patenting increases subsequent patentability
  - Patenting increases the share of articles in journals in which relatively more industry scientists publish

Preview of Results

- Self-selection: patenters are more successful scientists, and even more recently successful scientists
- Patenting appears to complement, not substitute, publication output
  - The elasticity of publication count with respect to applying for a patent lies between .195 (fixed effect estimate) and .235 (IPTCW estimate)
- No apparent effect on the *quality* of publication output
  - Order of Authorship
  - Average *Journal Impact Factor* (JIF)
- But genuine impact on the content of publications
  - Patenting entails more coauthored pubs with scientists in industry
  - Patenting increases subsequent patentability
  - Patenting increases the share of articles in journals in which relatively more industry scientists publish

Methodological Challenge

Classic approach: "Diff-in-Diffs" estimation

$$y_{it} = \beta_0 + \beta_1' X_{it} + \beta_2 TREAT_{it} + \alpha_i + \gamma_t + \varepsilon_{it}$$

- Recovers causal effect if treatment and controls would have followed the same trend in the absence of treatment.
- Likely to be the wrong approach here.

- Our approach: selection on observables
- Key assumption: conditional on observables, "treatment" is randomly allocated across control and treatment observations
  - Is this credible?
  - How does one implement this in practice?

Methodological Challenge

Classic approach: "Diff-in-Diffs" estimation

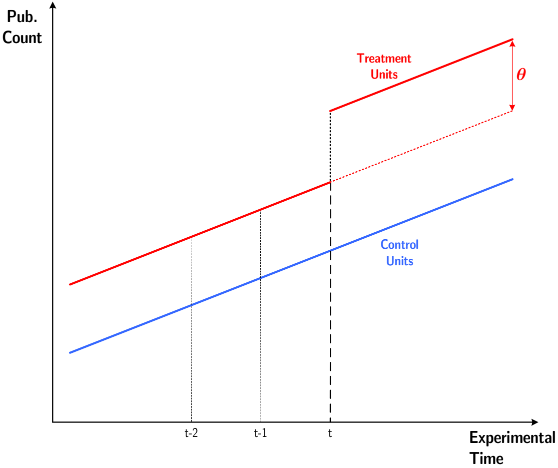
$$y_{it} = \beta_0 + \beta_1' X_{it} + \beta_2 TREAT_{it} + \alpha_i + \gamma_t + \varepsilon_{it}$$

- Recovers causal effect if treatment and controls would have followed the same trend in the absence of treatment.
- Likely to be the wrong approach here.

- Our approach: selection on observables
- Key assumption: conditional on observables, "treatment" is randomly allocated across control and treatment observations
  - Is this credible?
  - How does one implement this in practice?

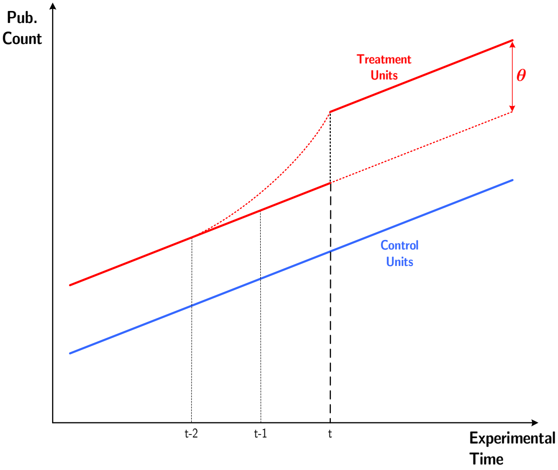
Positive Pre-treatment Trends $\Longrightarrow$ FE Undershoots

"Ashenfelter's Bump"



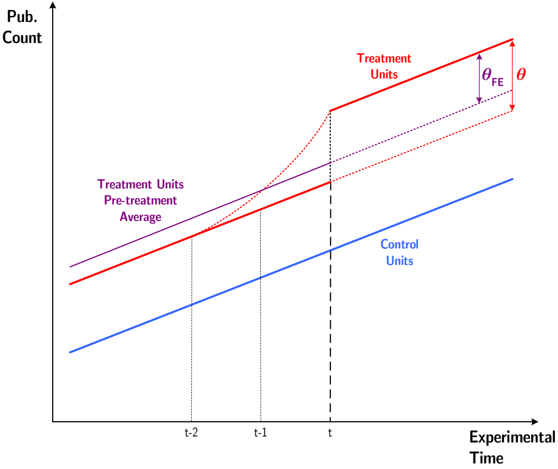
Positive Pre-treatment Trends $\Longrightarrow$ FE Undershoots

"Ashenfelter's Bump"



Positive Pre-treatment Trends $\Longrightarrow$ FE Undershoots

"Ashenfelter's Bump"



Why Not Include a Lagged Dep. Var. on the RHS?

Not all consistent estimates correspond to causal effects

Definition

A time-varying confounder (TVC) is a variable that

1. Predicts selection into treatment
2. Predicts future values of the outcome
3. Is itself predicted by past treatment history

Examples

- *CD4* cell count (HIV example)
- Lagged publication count, latent "patentability" (patenting example)

Why Not Include a Lagged Dep. Var. on the RHS?

Not all consistent estimates correspond to causal effects

Definition

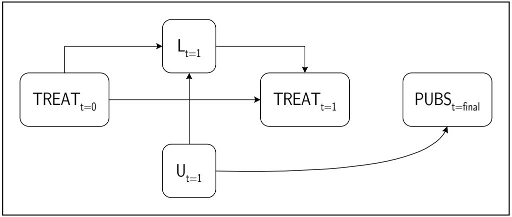
A time-varying confounder (TVC) is a variable that

1. Predicts selection into treatment
2. Predicts future values of the outcome
3. Is itself predicted by past treatment history

Examples

- *CD4* cell count (HIV example)
- Lagged publication count, latent "patentability" (patenting example)

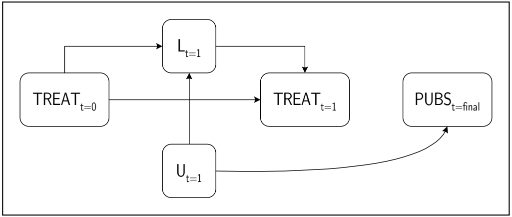
Bias Induced by Controlling for a Variable Affected by Previous Treatment



Legend

- *U* denotes the true, unobserved scientific value of pubs.
- *L* denotes a TVC, e.g. "patentability" or lagged pubs.
- [CIA]: No direct arrow from *U* to *TREAT*

Bias Induced by Controlling for a Variable Affected by Previous Treatment



Legend

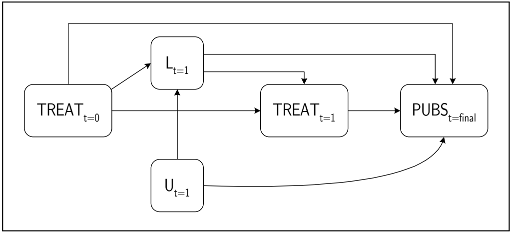
- U denotes the true, unobserved scientific value of pubs.
- L denotes a TVC, e.g. "patentability" or lagged pubs.
- [CIA]: No direct arrow from U to $TREAT$

Bias under the null

Controlling for $L_{t=1}$ creates an effect of $\sum_{t=0}^{1} TREAT_{it}$ where none exists.

- *e.g.*, among those with low patentability at $t = 1$, having patented at $t = 0$ makes it more likely that the true scientific value of the scientist's ideas is low.
- *e.g.*, among those with high patentability at $t = 1$, not having patented at $t = 0$ makes it more likely that the true scientific value of the scientist's ideas is high.

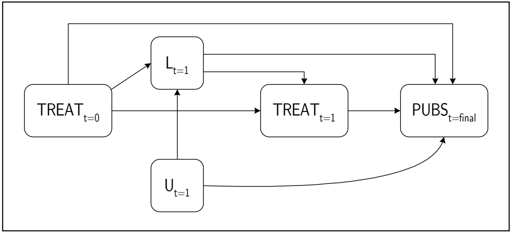
Bias Induced by Controlling for a Variable Affected by Previous Treatment



Legend

- *U* denotes the true, unobserved scientific value of pubs.
- *L* denotes a TVC, e.g. "patentability" or lagged pubs.
- [CIA]: No direct arrow from *U* to *TREAT*

Bias Induced by Controlling for a Variable Affected by Previous Treatment



Legend

- U denotes the true, unobserved scientific value of pubs.
- L denotes a TVC, e.g. "patentability" or lagged pubs.
- [CIA]: No direct arrow from U to $TREAT$

Bias under the alternative

Controlling for $L_{t=1}$ "blocks" the effect of $TREAT_{i,t=0}$ on the outcome of interest.

- $L_{t=1}$ is both a predictor of the final publication count AND is affected by previous treatment $TREAT_{i,t=0}$.
- The corresponding estimates are consistent, but do not correspond to a causal parameter of interest.

Selection on observables

Lessons from the program evaluation literature

- Non-experimental matching estimators "work well" when:
  - Treatment and controls are drawn from similar labor markets
  - There is a long list of covariates to match units on (including lagged outcomes)
  - Outcomes are measured in the same way for control and treatment observations
- Propensity score not appropriate when calendar and experimental time do not coincide, *i.e.*, when treatments are staggered over time

Selection on observables

Lessons from the program evaluation literature

- Non-experimental matching estimators "work well" when:
  - Treatment and controls are drawn from similar labor markets
  - There is a long list of covariates to match units on (including lagged outcomes)
  - Outcomes are measured in the same way for control and treatment observations
- Propensity score not appropriate when calendar and experimental time do not coincide, *i.e.*, when treatments are staggered over time

Selection on observables

Lessons from the program evaluation literature

- Non-experimental matching estimators "work well" when:
  - Treatment and controls are drawn from similar labor markets
  - There is a long list of covariates to match units on (including lagged outcomes)
  - Outcomes are measured in the same way for control and treatment observations
- Propensity score not appropriate when calendar and experimental time do not coincide, *i.e.*, when treatments are staggered over time

Selection on Observables

Counterfactual Outcomes

Notation

1. $i = 1 \cdots n$ scientists; $t = 0 \cdots T$ periods
2. y_{it} is the outcome of interest
3. For any variable W, denote $\widetilde{W}_{it}$ its history up to time t
4. At each time t,
   - scientist i chooses discrete treatment $TREAT_{it}$
   - "prognostic factors" W_{it} are measured
5. We distinguish between exogenous covariates X_{it} and *time-varying confounders* Z_{it}:
   - $W_{it} = (X_{it}; Z_{it})$

Selection on Observables

Counterfactual Outcomes

Definition

- Let $y_{it}^{\tilde{a}}$ be the value of y that would have been observed at time t had i chosen treatment sequence $\tilde{a}_{it} = (a_{i0}, a_{i1}, \ldots, a_{it})$ rather than his observed treatment history $\widetilde{TREAT}_{it}$.
- The average treatment effect of treatment history $\tilde{a}$ on the outcome y is the difference $E[y^{\tilde{a}}] - E[y^{\tilde{0}}]$, the average difference between outcomes when following $\tilde{a}$ and outcomes when never treated.

Selection on Observables
Complications with longitudinal data

Definition

A time-varying confounder (TVC) is a variable that

1. Predicts selection into treatment
2. Predicts future values of the outcome
3. Is itself predicted by past treatment history

Examples

- *CD4* cell count (HIV example)
- Lagged publication count, latent "patentability" (patenting example)

Selection on Observables

Complications with longitudinal data

Definition

A time-varying confounder (TVC) is a variable that

1. Predicts selection into treatment
2. Predicts future values of the outcome
3. Is itself predicted by past treatment history

Examples

- *CD4* cell count (HIV example)
- Lagged publication count, latent "patentability" (patenting example)

Selection on Observables

Key Econometric Result, due to Robins (Multiple Refs.)

Sequential Conditional Independence Assumption [SCIA]

For all i, t, and treatment regime $\widetilde{a}$:

$$y_{it}^{\widetilde{a}} \amalg TREAT_{it} | TREAT_{i,t-1}, Z_{i,t-1}, X_{it}$$

Model for the Counterfactual Mean

We model the mean of $y^{\widetilde{a}}$ conditional on treatment and exogenous covariates X as:

$$E\left[y_{it}^{\widetilde{a}} | TREAT_{it}, X_{it}\right] = \beta_0 + \beta_1' X_{it} + \beta_2 TREAT_{it}$$

Theorem

Under [SCIA], the average treatment effect β_2 is identified and can be recovered by estimating

$$y_{it} = \beta_0 + \beta_1' X_{it} + \beta_2 TREAT_{it} + \varepsilon_{it}$$

by weighted least squares, where the weights correspond to the inverse probability of following actual treatment history $\widetilde{TREAT}_{it}$ up to time t for scientist i.

Selection on Observables

Key Econometric Result, due to Robins (Multiple Refs.)

Sequential Conditional Independence Assumption [SCIA]

For all i, t, and treatment regime $\widetilde{a}$:

$$y_{it}^{\widetilde{a}} \amalg TREAT_{it} | TREAT_{i,t-1}, Z_{i,t-1}, X_{it}$$

Model for the Counterfactual Mean

We model the mean of $y^{\widetilde{a}}$ conditional on treatment and exogenous covariates X as:

$$E\left[y_{it}^{\widetilde{a}} | TREAT_{it}, X_{it}\right] = \beta_0 + \beta_1' X_{it} + \beta_2 TREAT_{it}$$

Theorem

Under [SCIA], the average treatment effect β_2 is identified and can be recovered by estimating

$$y_{it} = \beta_0 + \beta_1' X_{it} + \beta_2 TREAT_{it} + \varepsilon_{it}$$

by weighted least squares, where the weights correspond to the inverse probability of following actual treatment history $\widetilde{TREAT}_{it}$ up to time t for scientist i.

Selection on Observables

Key Econometric Result, due to Robins (Multiple Refs.)

Sequential Conditional Independence Assumption [SCIA]

For all i, t, and treatment regime $\widetilde{a}$:

$$y_{it}^{\widetilde{a}} \amalg TREAT_{it} | TREAT_{i,t-1}, Z_{i,t-1}, X_{it}$$

Model for the Counterfactual Mean

We model the mean of $y^{\widetilde{a}}$ conditional on treatment and exogenous covariates X as:

$$E\left[y_{it}^{\widetilde{a}} | TREAT_{it}, X_{it}\right] = \beta_0 + \beta_1' X_{it} + \beta_2 TREAT_{it}$$

Theorem

Under [SCIA], the average treatment effect β_2 is identified and can be recovered by estimating

$$y_{it} = \beta_0 + \beta_1' X_{it} + \beta_2 TREAT_{it} + \varepsilon_{it}$$

by weighted least squares, where the weights correspond to the inverse probability of following actual treatment history $\widetilde{TREAT}_{it}$ up to time t for scientist i.

Selection on Observables

Inverse Probability of Treatment Weights

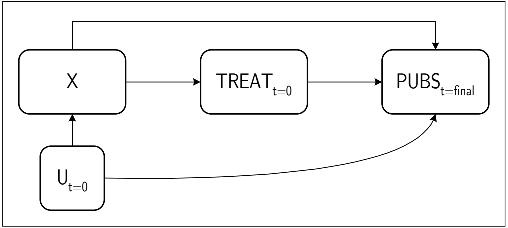
Definition

$$sw_{it} = \prod_{k=0}^{t} \frac{Prob(TREAT_{ik} \mid \widetilde{TREAT}_{i,k-1}, \widetilde{X}_{ik})}{Prob(TREAT_{ik} \mid \widetilde{TREAT}_{i,k-1}, \widetilde{Z}_{i,k-1}, \widetilde{X}_{ik})}$$

- Creates a pseudo population in which the TVCs (the Z variables) do not predict selection, but the relationship between treatment and outcome is identical to that in the original population
- $sw_{it}=1$ to the extent that TVCs (the Z variables) do not matter for selection into treatment

Selection on Observables

Motivating Inverse Probability of Treatment Weighting (in the X-sectional case)



Definitions

- Two potential outcomes, denoted by y_i^0 and y_i^1 for each individual i.
  - y_i^0: outcome that would be realized by i if (possibly contrary to the fact) not treated
  - y_i^1: outcome that would be realized by i if (possibly contrary to the fact) treated
- y_i denotes the realized outcome, $y_i = TREAT_i \cdot Y_i^1 + (1 - TREAT_i) \cdot Y_i^0$
- **Conditional Independence Assumption:** y^1 *and* $y^0 \amalg TREAT|X$
- $p(x) = Prob(TREAT = 1|X = x)$ denotes the **propensity score.**

Selection on Observables

Motivating Inverse Probability of Treatment Weighting

Sketch of proof in the X-sectional case

$$\begin{aligned}
E\left[\frac{TREAT \cdot y}{p(X)}\right] &= E\left[\frac{TREAT \cdot y^1}{p(X)}\right] \\
&= E\left\{E\left[\frac{TREAT \cdot y^1}{p(X)}|X\right]\right\} && \textbf{LIE} \\
&= E\left[\frac{E(TREAT|X) \cdot E(y^1|X)}{p(X)}\right] && \textbf{CIA} \\
&= E\left[\frac{p(X) \cdot E(y^1|X)}{p(X)}\right] = E\left[E(y^1|X)\right] = E\left[y^1\right]
\end{aligned}$$

Similarly,

$$E\left[\frac{(1-TREAT)\cdot y}{1-p(X)}\right] = E\left[y^0\right]$$

And therefore,

$$\tau = E\left[y^1\right] - E\left[y^0\right] = E\left[\frac{TREAT \cdot y}{p(X)} - \frac{(1-TREAT)\cdot y}{1-p(X)}\right]$$

Selection on Observables

Motivating Inverse Probability of Treatment Weighting

	$PATENT_i$=1	$PATENT_i$=0
"should have patented" **p[X]=.75**	Type A 1/p = 1.33	Type B 1/1-p = 4
"shouldn't have patented" **p[X]=.25**	Type C 1/p = 4	Type D 1/1-p = 1.33

Create a "fake" dataset in which X does not predict patenting

1 Copy of Type A scientist, 3 Copies of Type B scientist,
3 Copies of Type C scientist, 1 Copy of Type D scientist

Intuition

Weight relatively more the observations in which the predictions from the selection model and actual treatment choices *disagree*

Selection on Observables

Motivating Inverse Probability of Treatment Weighting

	$PATENT_i=1$	$PATENT_i=0$
"should have patented" p[X]=.75	Type A $1/p = 1.33$	Type B $1/1\text{-}p = 4$
"shouldn't have patented" p[X]=.25	Type C $1/p = 4$	Type D $1/1\text{-}p = 1.33$

Create a "fake" dataset in which X does not predict patenting

1 Copy of Type A scientist, 3 Copies of Type B scientist,
3 Copies of Type C scientist, 1 Copy of Type D scientist

Intuition

Weight relatively more the observations in which the predictions from the selection model and actual treatment choices *disagree*

Selection on Observables

Motivating Inverse Probability of Treatment Weighting

	$PATENT_i$=1	$PATENT_i$=0
"should have patented" **p[X]=.75**	Type A $1/p = 1.33$	Type B $1/1\text{-}p = 4$
"shouldn't have patented" **p[X]=.25**	Type C $1/p = 4$	Type D $1/1\text{-}p = 1.33$

Create a "fake" dataset in which X does not predict patenting

1 Copy of Type A scientist, 3 Copies of Type B scientist,
3 Copies of Type C scientist, 1 Copy of Type D scientist

Intuition

Weight relatively more the observations in which the predictions from the selection model and actual treatment choices *disagree*

Estimation Procedure

1. Estimate 2 logit models for probability of selection
   - Numerator: without including the time-varying confounders
   - Denominator: including the time-varying confounders

2. Multiply fitted values to create the weights:
   1 corresponds to $\frac{1}{\hat{p}}$, 0 corresponds to $\frac{1}{1-\hat{p}}$

3. Deal in a similar way with censoring;
   the product of the selection weight and the censoring weight is the final IPTC weight

4. Estimate the weighted outcome equation

Estimation Procedure

1. Estimate 2 logit models for probability of selection
   - Numerator: without including the time-varying confounders
   - Denominator: including the time-varying confounders
2. Multiply fitted values to create the weights:
   1 corresponds to $\frac{1}{\hat{p}}$, 0 corresponds to $\frac{1}{1-\hat{p}}$
3. Deal in a similar way with censoring;
   the product of the selection weight and the censoring weight is the final IPTC weight
4. Estimate the weighted outcome equation

Estimation Procedure

1. Estimate 2 logit models for probability of selection
   - Numerator: without including the time-varying confounders
   - Denominator: including the time-varying confounders

2. Multiply fitted values to create the weights:
   1 corresponds to $\frac{1}{\hat{p}}$, 0 corresponds to $\frac{1}{1-\hat{p}}$

3. Deal in a similar way with censoring;
   the product of the selection weight and the censoring weight is the final IPTC weight

4. Estimate the weighted outcome equation

Estimation Procedure

1. Estimate 2 logit models for probability of selection
   - Numerator: without including the time-varying confounders
   - Denominator: including the time-varying confounders

2. Multiply fitted values to create the weights:
   1 corresponds to $\frac{1}{\hat{p}}$, 0 corresponds to $\frac{1}{1-\hat{p}}$

3. Deal in a similar way with censoring;
   the product of the selection weight and the censoring weight is the final IPTC weight

4. Estimate the weighted outcome equation

Estimation Procedure

1. Estimate 2 logit models for probability of selection
   - Numerator: without including the time-varying confounders
   - Denominator: including the time-varying confounders

2. Multiply fitted values to create the weights:
   1 corresponds to $\frac{1}{\hat{p}}$, 0 corresponds to $\frac{1}{1-\hat{p}}$

3. Deal in a similar way with censoring;
   the product of the selection weight and the censoring weight is the final IPTC weight

4. Estimate the weighted outcome equation

Data

- Random sample of academic scientists, stratified by field to match distribution of academic firm founders
- Outcome variables
  - Pub. count
  - First/last vs. middle author Publication count
  - Average *Journal Impact Factor* (JIF)
  - Proportion of coauthored publications with industry scientists
  - Research "Patentability"
  - Average *Journal Commercial Score* (JCS)
- Observable characteristics
  - Gender, scientific field, characteristics of PhD university, characteristics of current employer, experience
- Patenting measure: flow, "regime" shift, stock

Data

- Random sample of academic scientists, stratified by field to match distribution of academic firm founders
- Outcome variables
  - Pub. count
  - First/last vs. middle author Publication count
  - Average *Journal Impact Factor* (JIF)
  - Proportion of coauthored publications with industry scientists
  - Research "Patentability"
  - Average *Journal Commercial Score* (JCS)
- Observable characteristics
  - Gender, scientific field, characteristics of PhD university, characteristics of current employer, experience
- Patenting measure: flow, "regime" shift, stock

Data

- Random sample of academic scientists, stratified by field to match distribution of academic firm founders
- Outcome variables
  - Pub. count
  - First/last vs. middle author Publication count
  - Average *Journal Impact Factor* (JIF)
  - Proportion of coauthored publications with industry scientists
  - Research "Patentability"
  - Average *Journal Commercial Score* (JCS)
- Observable characteristics
  - Gender, scientific field, characteristics of PhD university, characteristics of current employer, experience
- Patenting measure: flow, "regime" shift, stock

Data

- Random sample of academic scientists, stratified by field to match distribution of academic firm founders
- Outcome variables
  - Pub. count
  - First/last vs. middle author Publication count
  - Average *Journal Impact Factor* (JIF)
  - Proportion of coauthored publications with industry scientists
  - Research "Patentability"
  - Average *Journal Commercial Score* (JCS)
- Observable characteristics
  - Gender, scientific field, characteristics of PhD university, characteristics of current employer, experience
- Patenting measure: flow, "regime" shift, stock

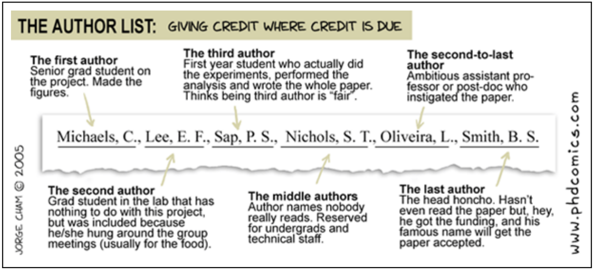
Data

- Random sample of academic scientists, stratified by field to match distribution of academic firm founders
- Outcome variables
  - Pub. count
  - First/last vs. middle author Publication count
  - Average *Journal Impact Factor* (JIF)
  - Proportion of coauthored publications with industry scientists
  - Research "Patentability"
  - Average *Journal Commercial Score* (JCS)
- Observable characteristics
  - Gender, scientific field, characteristics of PhD university, characteristics of current employer, experience
- Patenting measure: flow, "regime" shift, stock

Top 15 Scientific Disciplines Represented

UMI Subject Code	UMI Subject Description	Frequency	
487	Biochemistry	855	22.10%
306	Biology, General	563	14.60%
410	Biology, Microbiology	466	12.10%
419	Health Sciences, Pharmacology	239	6.20%
490	Chemistry, Organic	212	5.50%
786	Biophysics, General	210	5.40%
369	Biology, Genetics	191	4.90%
433	Biology, Animal Physiology	170	4.40%
982	Health Sciences, Immunology	167	4.30%
307	Biology, Molecular	102	2.60%
301	Bacteriology	61	1.60%
287	Biology, Anatomy	54	1.40%
571	Health Sciences, Pathology	52	1.30%
349	Psychology, psychobiology	37	1.00%
572	Health Sciences, Pharmacy	33	0.90%

Inferring Publication "Importance" from Order of Authorship



Measuring Patentability

- Heterogeneity in the commercial value of the research produced by scientists
- Scientific field fixed effects are not going to capture this heterogeneity
- We attempt to compute a direct measure of latent patentability
  - Knowledge of the research foci of academic scientists who have already patented can be used to identify the domains of science in which research is patentable
- With this measure, we ask three questions:
  - Does patentability indeed predicts patenting?
  - Is it the flow or the stock of our measure that most strongly influences patenting behavior?
  - Does the act of patenting influences subsequent research patentability?

Measuring Patentability

- Heterogeneity in the commercial value of the research produced by scientists
- Scientific field fixed effects are not going to capture this heterogeneity
- We attempt to compute a direct measure of latent patentability
  - Knowledge of the research foci of academic scientists who have already patented can be used to identify the domains of science in which research is patentable
- With this measure, we ask three questions:
  - Does patentability indeed predicts patenting?
  - Is it the flow or the stock of our measure that most strongly influences patenting behavior?
  - Does the act of patenting influences subsequent research patentability?

Measuring Patentability

- Heterogeneity in the commercial value of the research produced by scientists
- Scientific field fixed effects are not going to capture this heterogeneity
- We attempt to compute a direct measure of latent patentability
  - Knowledge of the research foci of academic scientists who have already patented can be used to identify the domains of science in which research is patentable
- With this measure, we ask three questions:
  - Does patentability indeed predicts patenting?
  - Is it the flow or the stock of our measure that most strongly influences patenting behavior?
  - Does the act of patenting influences subsequent research patentability?

Measuring Patentability

- Heterogeneity in the commercial value of the research produced by scientists
- Scientific field fixed effects are not going to capture this heterogeneity
- We attempt to compute a direct measure of latent patentability
  - Knowledge of the research foci of academic scientists who have already patented can be used to identify the domains of science in which research is patentable
- With this measure, we ask three questions:
  - Does patentability indeed predicts patenting?
  - Is it the flow or the stock of our measure that most strongly influences patenting behavior?
  - Does the act of patenting influences subsequent research patentability?

Measuring Patentability

- Heterogeneity in the commercial value of the research produced by scientists
- Scientific field fixed effects are not going to capture this heterogeneity
- We attempt to compute a direct measure of latent patentability
  - Knowledge of the research foci of academic scientists who have already patented can be used to identify the domains of science in which research is patentable
- With this measure, we ask three questions:
  - Does patentability indeed predicts patenting?
  - Is it the flow or the stock of our measure that most strongly influences patenting behavior?
  - Does the act of patenting influences subsequent research patentability?

Measuring Patentability (Cont'd)

For all scientists i, keywords j and articles s

Keyword Weight

$$w_{jt}^{i} = \frac{\sum_{s \in I_t^p - \{i\}} \frac{m_{sjt}}{\sum_k m_{skt}}}{\sum_{s \in I_t^{np} - \{i\}} m_{sjt}}$$

We sum over keywords contained in articles published in year t to compute the patentability score for scientist i

Definition

$$PATENTABILITY_{it} = \sum_{j=1}^{J} w_{j,t-1}^{i} \frac{n_{ijt}}{\sum_k n_{ikt}}$$

Measuring Patentability (Cont'd)

For all scientists i, keywords j and articles s

Keyword Weight

$$w_{jt}^{i} = \frac{\sum_{s \in I_t^{p} - \{i\}} \frac{m_{sjt}}{\sum_k m_{skt}}}{\sum_{s \in I_t^{np} - \{i\}} m_{sjt}}$$

We sum over keywords contained in articles published in year t to compute the patentability score for scientist i

Definition

$$PATENTABILITY_{it} = \sum_{j=1}^{J} w_{j,t-1}^{i} \frac{n_{ijt}}{\sum_k n_{ikt}}$$

Measuring Patentability (Cont'd)

For all scientists i, keywords j and articles s

Keyword Weight

$$w_{jt}^{i} = \frac{\sum_{s \in I_t^p - \{i\}} \frac{m_{sjt}}{\sum_k m_{skt}}}{\sum_{s \in I_t^{np} - \{i\}} m_{sjt}}$$

We sum over keywords contained in articles published in year t to compute the patentability score for scientist i

Definition

$$PATENTABILITY_{it} = \sum_{j=1}^{J} w_{j,t-1}^{i} \frac{n_{ijt}}{\sum_k n_{ikt}}$$

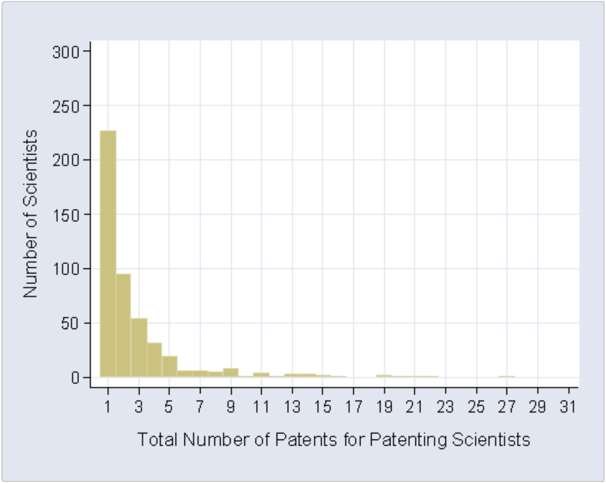
Sample Keywords in 2000

	(1)	(2)	(3)	(4)
	Number of times the keyword was used by patenting scientists	Sum over all patenting scientists of keyword's proportion of total keywords used	Number of times the keyword was used by non-patenting scientists	Keyword weight: Column (2) / Column (3)
Group 1				
HIV-inhibitory	24	0.0110	1	1.100
glaucoma	30	0.0690	25	0.276
ubiquitin	55	0.1450	30	0.483
telomere	37	0.0940	35	0.269
Group 2				
t-cell	424	0.9000	1,242	0.072
antigen	494	1.0940	1,789	0.061
peptide	403	1.0980	1,511	0.073
Group 3				
carnitine	1	0.0004	60	0.001
endothelium-dependent	1	0.0007	51	0.001
aromatase	1	0.0006	70	0.001
aplysia	4	0.0150	102	0.026

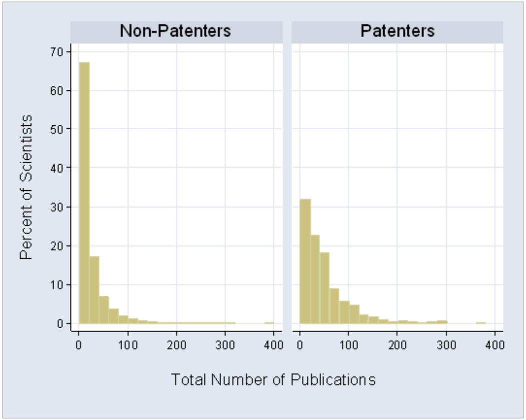
Descriptive Statistics

	Mean	Std. Dev.	Min.	Max.	*N*
Patent Flow (=1 if one or more patent app. in year)	0.030	0.131	0	1	58,562
Patent Regime (=1 after first patent app.)	0.073	0.261	0	1	58,562
Patent Stock	0.184	1.175	0	57	58,562
Research Publication Flow	1.729	2.379	0	35	58,562
Fraction of First or Last Authored Publications (Flow)	0.619	0.397	0	1	38,007
Average JIF of Publications (Flow)	3.956	3.101	0.005	30.334	38,007
Average Journal Commercial Score of Pubs. (Flow)	0.076	0.055	0.001	1	38,007
Fraction of Pubs. with Industry Coauthors (Flow)	0.075	0.223	0	1	38,007
Research Patentability Score (Flow)	0.022	0.049	0	4.173	58,562
Employer Graduate School in Top 20	0.231	0.422	0	1	58,562
Employer has TTO	0.488	0.500	0	1	58,562
Employer Patent Stock ($\times$0.01)	0.718	1.452	0	2.189	58,562
Experience (Career Age)	10.201	7.122	1	32	58,562
Female	0.183	0.387	0	1	3,862
Scientist has one or more patents	0.122	0.328	0	1	3,862
Ph.D. Univ. Grad. School in Top 20	0.308	0.462	0	1	3,862
Ph.D. Univ. 5-year Patent Stock ($\times$0.01)	0.190	0.409	0	5.660	3,862

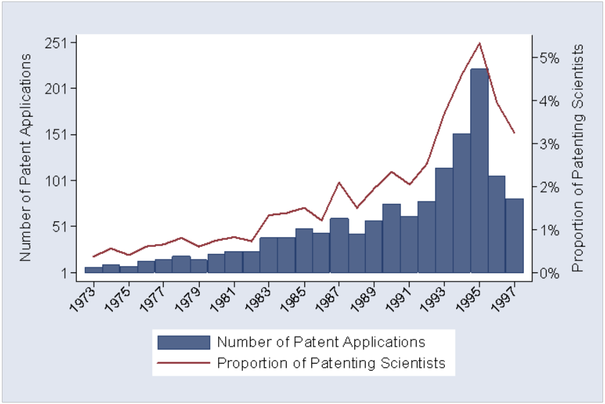
Distribution of Patent Count for Patenting Scientists



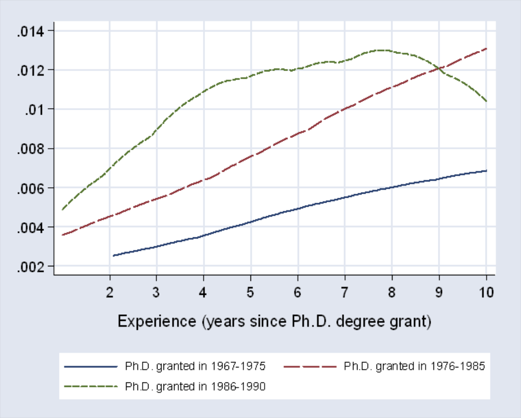
Distribution of Publication Count for Patenting and Non-patenting Scientists



Distribution of Patenting Events over Time



Unconditional Hazard of First Patent Application, by Ph.D. Cohort



The Anatomy of Self-selection into Patenting
Demographics or Opportunities?

- Patenting is concentrated among the group of eminent scientists, but what is the mechanism that generates the relationship between scientific status and patenting behavior?
- Two alternative views
  - Demographics — time-invariant talent
    scientists "cash in" already established reputation when they patent
  - Opportunities — upward deviation from individual trend
    scientists "hit the mother lode" and thereby clear the patenting hurdle
- Empirically, we examine how the flow/stock of publications influences the propensity to patent

The Anatomy of Self-selection into Patenting

Demographics or Opportunities?

- Patenting is concentrated among the group of eminent scientists, but what is the mechanism that generates the relationship between scientific status and patenting behavior?
- Two alternative views
  - Demographics — time-invariant talent
    scientists "cash in" already established reputation when they patent
  - Opportunities — upward deviation from individual trend
    scientists "hit the mother lode" and thereby clear the patenting hurdle
- Empirically, we examine how the flow/stock of publications influences the propensity to patent

The Anatomy of Self-selection into Patenting

Demographics or Opportunities?

- Patenting is concentrated among the group of eminent scientists, but what is the mechanism that generates the relationship between scientific status and patenting behavior?
- Two alternative views
  - Demographics — time-invariant talent
    scientists "cash in" already established reputation when they patent
  - Opportunities — upward deviation from individual trend
    scientists "hit the mother lode" and thereby clear the patenting hurdle
- Empirically, we examine how the flow/stock of publications influences the propensity to patent

The Anatomy of Self-selection into Patenting

Demographics or Opportunities?

- Patenting is concentrated among the group of eminent scientists, but what is the mechanism that generates the relationship between scientific status and patenting behavior?
- Two alternative views
  - Demographics — time-invariant talent
    scientists "cash in" already established reputation when they patent
  - Opportunities — upward deviation from individual trend
    scientists "hit the mother lode" and thereby clear the patenting hurdle
- Empirically, we examine how the flow/stock of publications influences the propensity to patent

The Anatomy of Self-selection into Patenting

Demographics or Opportunities?

- Patenting is concentrated among the group of eminent scientists, but what is the mechanism that generates the relationship between scientific status and patenting behavior?
- Two alternative views
  - Demographics — time-invariant talent
    scientists "cash in" already established reputation when they patent
  - Opportunities — upward deviation from individual trend
    scientists "hit the mother lode" and thereby clear the patenting hurdle
- Empirically, we examine how the flow/stock of publications influences the propensity to patent

Determinants of Selection into Patenting

Logit Estimates

	Patent Flow		Patent Regime		Exit Academia	
	Denominator	Numerator	Denominator	Numerator	Denominator	Numerator
Experience = [5, 8]	0.141	0.195	0.166	0.239		
	[0.153]	[0.153]	[0.166]	[0.164]		
Experience = [9, 15]	0.219	0.347	0.305	0.432	0.206	-0.006
	[0.155]	[0.151]*	[0.168]†	[0.162]**	[0.060]**	[0.057]
Experience = [16, 22]	0.022	0.218	0.252	0.401	0.116	-0.264
	[0.174]	[0.162]	[0.196]	[0.180]*	[0.087]	[0.077]**
Experience = [23, 35]	-0.357	-0.097	-0.343	-0.232	0.371	-0.122
	[0.213]†	[0.198]	[0.278]	[0.267]	[0.116]**	[0.101]
Female	-0.649	-0.675	-0.663	-0.7	0.147	0.243
	[0.130]**	[0.133]**	[0.153]**	[0.152]**	[0.054]**	[0.053]**
Patent Flow$_{t-1}$	1.971	2.048			0.299	
	[0.093]**	[0.128]**			[0.174]†	
Patent Stock$_{t-2}$	1.945	2.065			-0.128	
	[0.124]**	[0.093]**			[0.103]	
Publication Flow$_{t-1}$	0.042		0.083		-0.215	
	[0.016]**		[0.022]**		[0.024]**	
Publications Stock$_{t-2}$	0.003		-0.001		-0.013	
	[0.002]		[0.002]		[0.003]**	
High Research Patentability$_{t-1}$	0.309		0.336		-0.097	
	[0.093]**		[0.112]**		[0.068]	
Research Patentability Stock$_{t-2}$	0.129		0.247		0.017	
	[0.309]		[0.300]		[0.203]	
Has Industry Coauthors$_{t-1}$	0.076		0.061		0.055	
	[0.093]		[0.113]		[0.061]	
Employer Grad. School in Top 20	0.143		-0.014		0.054	
	[0.113]		[0.119]		[0.059]	
Employer has TTO	0.137		0.012		-0.05	
	[0.096]		[0.118]		[0.053]	
Employer Patent Stock$_{t-1}$ ($\times$100)	-0.007		0.09		0.031	
	[0.026]		[0.033]**		[0.016]†	
Ph.D. Univ Grad. School in Top 20	0.011	0.053	0.089	0.121	-0.151	-0.181
	[0.092]	[0.089]	[0.104]	[0.104]	[0.053]**	[0.053]**
Ph.D. Univ. 5-year Patent Stock ($\times$100)	0.001	0.001	0.001	0.002	-0.001	-0.001
	[0.001]	[0.001]†	[0.001]	[0.001]*	[0.001]	[0.001]
Constant	-6.043	-5.968	-6.098	-6.039	-4.383	-4.533
	[0.295]**	[0.300]**	[0.304]**	[0.302]**	[0.139]**	[0.139]**
Observations	58,562	58,562	54,746	54,746	58,437	58,437
Number of researchers	3,862	3,862	3,862	3,862	3,862	3,862
Log pseudo-likelihood	-3956.36	-3994.8	-2549.11	-2578.29	-8878.77	-9092.91
Wald χ^2	2263.35	2089.54	348.72	272.91	564.09	308.91

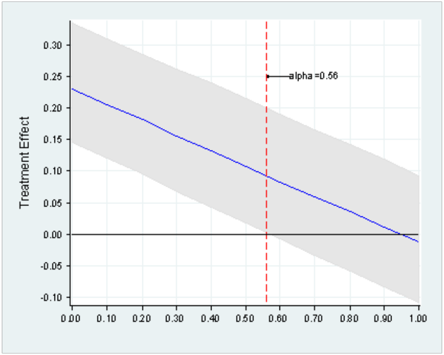
Impact of Acad. Patenting on the *Rate* of Publications
Poisson QML Estimates

	Model 2a	Model 2b	Model 2c
Scientist Fixed Effects	Yes	No	No
IPTC Weights	No	No	Yes
Experience = [5, 8]	0.161 [0.018]**	0.2 [0.018]**	0.206 [0.019]**
Experience = [9, 15]	0.262 [0.029]**	0.43 [0.030]**	0.42 [0.033]**
Experience = [16, 22]	0.228 [0.041]**	0.521 [0.049]**	0.427 [0.047]**
Experience = [23, 32]	0.085 [0.050]†	0.487 [0.073]**	0.335 [0.070]**
Female		-0.203 [0.051]**	-0.224 [0.049]**
PhD Univ. Grad School in Top 20		0.063 [0.042]	0.052 [0.041]
PhD Univ. 5-Year Patent Stock ($\times$100)		0.043 [0.047]	0.048 [0.047]
Patent Regime	**0.195** **[0.031]****	**0.394** **[0.048]****	**0.235** **[0.047]****
Constant		0.034 [0.044]	0.041 [0.045]
Log pseudo-likelihood	-78070	-119953.1	-117057.9
Wald χ^2	2966.37	1301.65	948.59

Sensitivity Analysis

Bias from unmeasured confounding

$$\Delta_{it} = \alpha \cdot pubs_{it} \cdot (2TREAT_{it} - 1)$$



Impact of Acad. Patenting on the *Quality* of Publications

QML Estimates

	Model 1a	Model 1b	Model 2a	Model 2b
	Fractional Logit		Poisson Model	
	Proportion of First or Last-Authored Publications		**Average JIF of Publications**	
	Unweighted	IPTCW	Unweighted	IPTCW
Experience = [5, 8]	-0.096	-0.096	-0.087	-0.088
	[0.029]**	[0.029]**	[0.013]**	[0.013]**
Experience = [9, 15]	0.034	0.029	-0.189	-0.186
	[0.034]	[0.034]	[0.018]**	[0.018]**
Experience = [16, 22]	0.133	0.122	-0.273	-0.275
	[0.046]**	[0.046]**	[0.027]**	[0.027]**
Experience = [23, 32]	0.155	0.137	-0.354	-0.366
	[0.068]*	[0.070]†	[0.039]**	[0.040]**
Female	-0.003	0.0003	0.031	0.033
	[0.038]	[0.038]	[0.022]	[0.022]
PhD Univ. Grad School in Top 20	0.05	0.047	0.135	0.131
	[0.033]	[0.033]	[0.021]**	[0.021]**
PhD Univ. 5-Year Patent Stock ($\times$100)	0.049	0.041	0.086	0.094
	[0.042]	[0.043]	[0.030]**	[0.029]**
Patent Regime	**0.026**	**-0.004**	**0.077**	**0.052**
	[0.048]	**[0.051]**	**[0.029]****	**[0.030]†**
Constant	0.826	0.827	1.37	1.371
	[0.047]**	[0.047]**	[0.023]**	[0.023]**
Log pseudo-likelihood	-22238.9	-21846.2	-91867.7	-90193.4
Wald χ^2	272.6	268.9	642.1	680.8

Impact of Acad. Patenting on the *Content* of Publications
QML Estimates

	Model 1a	Model 1b	Model 2a	Model 2b	Model 3a	Model 3b
	Poisson Models		Fractional Logit		Fractional Logit	
	Research Patentability		**Proportion of Pub. with Industry Coauthors**		**Average Journal Commercial Score**	
	Unweighted	IPTCW	Unweighted	IPTCW	Unweighted	IPTCW
Experience = [5, 8]	0.008	0.005	0.102	0.099	0.016	0.016
	[0.039]	[0.039]	[0.069]	[0.070]	[0.014]	[0.014]
Experience = [9, 15]	-0.025	-0.024	0.13	0.124	0.006	0.006
	[0.038]	[0.037]	[0.086]	[0.086]	[0.019]	[0.019]
Experience = [16, 22]	-0.054	-0.054	0.122	0.128	0.015	0.019
	[0.038]	[0.038]	[0.111]	[0.111]	[0.025]	[0.025]
Experience = [23, 32]	-0.103	-0.104	0.087	0.083	0.057	0.076
	[0.042]**	[0.043]*	[0.154]	[0.155]	[0.035]	[0.035]*
Female	-0.023	-0.023	-0.07	-0.066	-0.007	-0.005
	[0.022]	[0.023]	[0.091]	[0.092]	[0.017]	[0.017]
PhD Univ. Grad School in Top 20	-0.027	-0.025	-0.313	-0.329	-0.069	-0.067
	[0.021]	[0.022]	[0.084]**	[0.086]**	[0.018]**	[0.018]**
PhD Univ. 5-year Patent Stock (×100)	-0.017	-0.018	0.133	0.113	-0.018	-0.018
	[0.020]	[0.020]	[0.098]	[0.091]	[0.025]	[0.026]
Patent Regime	**0.09**	**0.085**	**0.222**	**0.278**	**0.043**	**0.052**
	[0.028]**	**[0.029]****	**[0.088]***	**[0.097]****	**[0.024]†**	**[0.026]***
Constant	-5.7	-5.7	-3.831	-3.827	-2.491	-2.494
	[0.353]**	[0.352]**	[0.153]**	[0.153]**	[0.024]**	[0.024]**
Log pseudo-likelihood	-4887.3	-4750.6	-9099	-8901.8	-7669.4	-7524.1
Wald χ^2	2089.6	1939.8	305.47	295.21	431.53	394.01

Conclusions & Future Directions

- Selection on observables, an econometric free lunch?
  - How much unobserved heterogeneity would lead us to not reject the null?
  - Our sensitivity analysis says: quite a lot!

- Full evaluation of the academic patenting phenomenon would require accounting for externalities:
  - Industrial Firms
  - Trainees (graduate students and postdoctoral fellows)
  - "Invisible College"

Conclusions & Future Directions

- Selection on observables, an econometric free lunch?
  - How much unobserved heterogeneity would lead us to not reject the null?
  - Our sensitivity analysis says: quite a lot!
- Full evaluation of the academic patenting phenomenon would require accounting for externalities:
  - Industrial Firms
  - Trainees (graduate students and postdoctoral fellows)
  - "Invisible College"

Conclusions & Future Directions

- Selection on observables, an econometric free lunch?
  - How much unobserved heterogeneity would lead us to not reject the null?
  - Our sensitivity analysis says: quite a lot!

- Full evaluation of the academic patenting phenomenon would require accounting for externalities:
  - Industrial Firms
  - Trainees (graduate students and postdoctoral fellows)
  - "Invisible College"