The Diffusion of Information Technology in the United States and Its Impact on Social Science Research Across Institutions of Higher Education

Anne E. Winkler, University of Missouri-St. Louis
Sharon G. Levin, University of Missouri-St. Louis
Paula E. Stephan, Georgia State Univ. and NBER
Wolfgang Glanzel, Katholieke Universiteit Leuven, Stenupunt O&O

Research Funded by the Mellon Foundation

IASSIST Conference, June 1-4, 2010
Motivation

Information Technology in Higher Education rapidly diffused from the 1980s to present

This prompts a host of research questions and has lead to several papers...
Project Scope

What factors explain the diffusion of early IT in higher education?

Paper under Review

What was the impact of this diffusion on individual publishing productivity of academics?

Paper using cross-section data from SDR, forthcoming in EINT

Paper (with Waverly Ding) using longitudinal data, forthcoming in Management Science

What was the impact of this diffusion on multi-institutional co-authorship patterns, and what are differences by field?

Today’s Focus
This Study

Investigates effect of IT exposure on institutional collaboration and extent of differential effects by field.

• Institutional publication data: Papers indexed by ISI for 1200+ institutions, 1991-2007

• Fields examined are natural sciences (bio, chem, physics) and social sciences (economics)

• Measure of IT: Domain Name System (DNS), e.g. www.umsl.edu
Literature Review: Collaboration Trends

• Increase in co-authors per paper ("team size")

Wuchty, Jones & Uzzi (2007) – ISI data from 1955-2000. Team size doubled from 1.9 to 3.5 authors per paper.

• Increase in collaboration across institutions


Fastest growth occurred in across-university collaborations for all fields.

By 2005, 32.8% of S&E pubs were multi-university
34.4% of Social Science pubs were multi-university
Explanations for Observed Trends

• Rising importance of interdisciplinary research

• With growth of knowledge in each discipline, researchers are becoming more specialized

• Minimize risk by diversifying one’s portfolio via collaboration

• More data available—Genbank database, PubChem, etc.

• Quality found to improve with collaboration

• TECHNOLOGY -- Reduced communication costs
Differences in Research and Collaboration by Field

• **Natural science research**
  Typically involves a physical lab, leading to on-site collaboration. Also, role of grants – they fund multiple scholars in a lab.

• **Social science research (e.g. economics)**
  Rarely involves a lab (except experimental)
  Regarding grants – they fund a PI or co-PI at most.
Role of Technology

• Technology has reduced communication costs in all fields
  => increased formal & informal collab.
  => sharing of data

• Differences in how technology is used by field
  (Walsh & Bayma, 1996; Walsh et al. 2000; Stephan, 2010)
Prior Empirical Studies of IT, Publishing & Collaboration: General Description

Considerable variation in studies depending on:

- **Type of publication data** (individual or institutional-level; cross-section or longitudinal)

- **Measurement of IT** (inferred from period effects, self-reported usage, or institutional adoption of explicit IT measure)

- **Definition of publication productivity** (number articles published or measure of collaboration)

- **Fields examined**
Specific Prior Studies

Natural Sciences

• Hesse et al. (1993)
• Cohen (1996) and Walsh et al. (2006) (and some social science/humanities fields)
• Winkler et al. (forthcoming, *EINT*)
• Ding et al. (forthcoming *Management Science*)
• Agrawal and Goldfarb (2008)

Social Sciences

• Hamermesh & Oster (2002)
• Rosenblat & Mobius (2004)
• Kim, Morse & Zingales (2009)
• Butler et al. (2008)
This Study

- 3 natural science fields (bio, chem, physics) and 1 social science field (economics)
- Institutional-level publication data
- Explicit measure of IT (DNS)
- Focuses on multi-institution collaboration
- Examines US-US and also US-INTL collaborations
IT Measure: DNS

• IT measured using information on institutional adoption of the Domain Name System (DNS). Example: www.umsl.edu

  Source: ALLWHOIS registry site

• We look at IT diffusion and collaboration patterns by tier using 1994 Carnegie codes:
  – Top Research/Doctoral (Carneg 11)
  – Other Research/Doctoral (Carneg 12,13,14)
  – Master’s Level (Carneg 21,22)
  – Top Liberal Arts (per US News & World Report, 1996)
Adoption of DNS by Year, Tier

- All Institutions
- Top Research/Doctoral
- Other Research/Doctoral
- Master's Level
- Top Liberal Arts
Institutional-Level Publication Data

• Data are from Web of Science/ISI for 1,281 four-year colleges and universities located in the U.S. for 1991-2007.

• Fields: All (omits Arts & Humanities), biology, chemistry, physics, economics per Glanzel and Schubert (2003)

  Note: related subfields cannot be aggregated to avoid duplication of publications (some articles are assigned to more than 1 field)

• Data are “whole counts.” An article with authors at two institutions is counted as 1 article at each institution.
Key Publication/Collaboration Measures

- **PUBS**—Number of publications per institution i

- **USUS**—number of publications at institution i where at least one co-author is at another institution within the U.S.

- **USINTL**—number of publications at institution i where at least one co-author is at an institution outside the U.S.

Example: This paper has 2 co-authors at UMSL, 1 at Georgia State, and one at Leuven (outside of US)

UMSL: Pubs = 1; USUS = 1; USINTL = 1
Georgia State: Pubs = 1; USUS = 1; USINTL = 1
## Table 2. Summary Statistics on Institutional Publication Data, by Tier and Field

<table>
<thead>
<tr>
<th>Field</th>
<th>All Tiers</th>
<th>Top Research/Doc</th>
<th>Other Research/Doc</th>
<th>Master's</th>
<th>Top Liberal Arts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean Pubs</td>
<td>% Zero Pubs</td>
<td>Mean Pubs</td>
<td>Mean Pubs</td>
<td>Mean Pubs</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>% Zero Pubs</td>
<td>% Zero Pubs</td>
</tr>
<tr>
<td>Biology</td>
<td>1991-1995</td>
<td>17.34</td>
<td>64.2%</td>
<td>0.0%</td>
<td>23.07</td>
</tr>
<tr>
<td></td>
<td>1996-2000</td>
<td>22.49</td>
<td>58.3%</td>
<td>0.0%</td>
<td>30.31</td>
</tr>
<tr>
<td></td>
<td>2001-2007</td>
<td>26.51</td>
<td>54.1%</td>
<td>0.0%</td>
<td>37.11</td>
</tr>
<tr>
<td>Chemistry</td>
<td>1991-1995</td>
<td>16.26</td>
<td>56.8%</td>
<td>0.0%</td>
<td>34.94</td>
</tr>
<tr>
<td></td>
<td>1996-2000</td>
<td>19.26</td>
<td>53.5%</td>
<td>0.0%</td>
<td>41.22</td>
</tr>
<tr>
<td></td>
<td>2001-2007</td>
<td>22.91</td>
<td>51.2%</td>
<td>0.0%</td>
<td>50.28</td>
</tr>
<tr>
<td>Physics</td>
<td>1991-1995</td>
<td>19.37</td>
<td>61.2%</td>
<td>0.0%</td>
<td>32.29</td>
</tr>
<tr>
<td></td>
<td>1996-2000</td>
<td>22.07</td>
<td>58.3%</td>
<td>0.0%</td>
<td>36.76</td>
</tr>
<tr>
<td></td>
<td>2001-2007</td>
<td>27.29</td>
<td>55.1%</td>
<td>0.0%</td>
<td>49.91</td>
</tr>
<tr>
<td>Economics</td>
<td>1991-1995</td>
<td>5.22</td>
<td>62.2%</td>
<td>1.1%</td>
<td>10.93</td>
</tr>
<tr>
<td></td>
<td>1996-2000</td>
<td>5.58</td>
<td>58.8%</td>
<td>0.9%</td>
<td>12.08</td>
</tr>
<tr>
<td></td>
<td>2001-2007</td>
<td>6.43</td>
<td>58.6%</td>
<td>1.3%</td>
<td>14.39</td>
</tr>
</tbody>
</table>
Summary of Publication Patterns, Full Sample, 1991-2007

For all fields, all tiers:

• Mean publications per institution increased from 159 to 228
• Median pubs rose from 5 to 8
⇒ Data are very skewed

• % institutions with zero pubs fell from 24% to 19%

By field:

• Mean pubs in Biology increased from 204 to 308
• Mean pubs in Economics increased from 50 to 59
Focus: Multi-Institution Collaborations

• % USUS = USUS/pubs
• % USINTL = USINTL/pubs

These figures are computed for institution-years with at least four publications in the given field
% US-US Collaborations, Top Research/Doctoral, 
by Field and Time Period


Biology
Chemistry
Physics
Economics
% US-INTL Collaborations, Top Research/Doctoral, by Field and Time Period

---|---|---
Biology | 16.2 | 21.4 | 26.7
Chemistry | 15.7 | 20.9 | 26.1
Physics | 26.9 | 38.3 | 44.4
Economics | 12.1 | 16.8 | 21.2
Summary of Key Patterns Regarding Multi-Institution Collaboration

• % USUS and % USINTL collaborations increased for all fields

• % US-US always higher for economics than natural sciences

• % US-INTL always higher for natural sciences than economics
Regression Analysis: Examines Effect of Exposure to IT on Multi-Institution Collaboration

Approach: “Modified Difference Equation”
Nets out changes in institutional factors (and their influence on publications) over time

**Dependent variable:** Year-to-year *change* in number of USUS collaborations (or *change* in number of USINTL collab.)

**Independent variables:**
1) Year-to-year *change* in total pubs
2) *Length* of exposure to DNS (modeled using dummies)
Estimated Model: “Modified First Difference”

\[ USUS_{\text{change}}_{i,t} = B_0 + B_1 \text{ Pub}_{\text{change}}_{i,t} + B_2 \text{ EXP}_{i,t-1} + \varepsilon_{i,t} \]

where

\text{Pub}_{\text{change}} = \text{change in total number of publications at institution } i \text{ in year } j

\text{USUS}_{\text{change}} = \text{year-to-year change in number of publications by institution } i \text{ with at least one co-author from another institution}

\text{EXP} = \text{measure of institutional exposure to DNS (dummy specification)}

Notes:

- USINTL change also used as a dependent variable
- Model estimated for institution-years with > 4 publications, years 1992-2001
- Estimated separately for All Fields, Biology, Chemistry, Physics, and Economics
- Estimated using OLS (with robust standard errors)
Findings, All Fields combined

• Exposure to DNS has a statistically significant positive effect on change in USUS (and change in USINTL) collaborations for All Fields combined

Result holds for all tiers except Top Liberal Arts
Findings, By Subfield

USUS Results:

*Modest* evidence that change in USUS is significantly related to length of exposure to DNS by subfield

- For natural sciences, significant IT effect is generally found for Top Research/Doctoral tier.

  Example: For top tier of chemistry, long exposure to DNS (10+ years) is found to lead to a net addition of 2.1 co-authored articles per year (compared to institutions with 0-4 yrs exposure).

- For economics, significant finding for Master’s level only, and of smaller magnitude than for natural sciences.
Findings, By Subfield, cont’d

USINTL Results:

• Impact of exposure to DNS was greater (in significance and magnitude) than for USUS results.

• Again, significant findings regarding exposure are for top tier in natural sciences only.
Other Models

1) Explicitly compared each natural science field to economics using a fully interactive dummy variable model. Tested for significant differences in IT’s effect on collaboration by field.

For Biology vs. Economics:
- Sig diff. for USUS, Top Research/Doctoral and Top Liberal Arts
- Sig diff. for USINTL, Top Research/Doctoral

For Chemistry vs. Economics,
- Sig diff. for USINTL, Top Research/Doc

For Physics vs. Economics,
- Sig diff. for USINTL, Top Research/Doctoral

2) Quantile regression. Suggests that results from OLS (mean regression) are “driven” by effects for the top quantile.
Conclusion and Next Steps

• Dramatic growth in USUS and USINTL collaboration for all tiers and fields examined

• Preliminary results suggest the impact of IT exposure was more pronounced for top tier natural sciences; larger effects for USINTL vs. USUS

• Future work – The impact of exposure at a point in time also depends on the size of the IT “network”

Comments appreciated! awinkler@umsl.edu
Table 3: Multi-Institution Collaborations, Measured in %

Panel A. % U.S. - U.S. Collaborations (calculated as USUS/Pubs)

<table>
<thead>
<tr>
<th>Year Period</th>
<th>Top Research/Doc %</th>
<th>Other Research/Doc %</th>
<th>Master's %</th>
<th>Top Liberal Arts %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biology</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1991-1995</td>
<td>40.1</td>
<td>40.5</td>
<td>50.8</td>
<td>51.0</td>
</tr>
<tr>
<td>1996-2000</td>
<td>48.3</td>
<td>46.7</td>
<td>54.8</td>
<td>54.2</td>
</tr>
<tr>
<td>2001-2007</td>
<td>54.7</td>
<td>53.7</td>
<td>59.5</td>
<td>57.3</td>
</tr>
<tr>
<td>Chemistry</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1991-1995</td>
<td>30.0</td>
<td>30.2</td>
<td>40.9</td>
<td>37.8</td>
</tr>
<tr>
<td>1996-2000</td>
<td>34.7</td>
<td>35.3</td>
<td>43.4</td>
<td>36.6</td>
</tr>
<tr>
<td>2001-2007</td>
<td>39.0</td>
<td>37.8</td>
<td>47.7</td>
<td>46.2</td>
</tr>
<tr>
<td>Physics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1991-1995</td>
<td>41.0</td>
<td>41.2</td>
<td>49.8</td>
<td>48.6</td>
</tr>
<tr>
<td>1996-2000</td>
<td>47.7</td>
<td>46.4</td>
<td>54.6</td>
<td>60.0</td>
</tr>
<tr>
<td>2001-2007</td>
<td>52.0</td>
<td>50.0</td>
<td>60.1</td>
<td>67.8</td>
</tr>
<tr>
<td>Economics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1991-1995</td>
<td>57.4</td>
<td>54.1</td>
<td>55.4</td>
<td>48.8</td>
</tr>
<tr>
<td>1996-2000</td>
<td>62.1</td>
<td>59.6</td>
<td>58.5</td>
<td>53.5</td>
</tr>
<tr>
<td>2001-2007</td>
<td>70.0</td>
<td>69.1</td>
<td>70.1</td>
<td>57.7</td>
</tr>
</tbody>
</table>

Note: Calculated for institution-year with > 4 pubs.
Table 3: Multi-Institution Collaborations, Measured in %

Panel B. % U.S. - International Collaborations (calculated as USINTL/Pubs)

<table>
<thead>
<tr>
<th>Year Range</th>
<th>Top Research/Doc %</th>
<th>Other Research/Doc %</th>
<th>Master's %</th>
<th>Top Liberal Arts %</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Biology</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1991-1995</td>
<td>16.2</td>
<td>14.0</td>
<td>16.6</td>
<td>12.9</td>
</tr>
<tr>
<td>1996-2000</td>
<td>21.4</td>
<td>18.2</td>
<td>19.1</td>
<td>20.4</td>
</tr>
<tr>
<td>2001-2007</td>
<td>26.7</td>
<td>24.4</td>
<td>21.9</td>
<td>17.7</td>
</tr>
<tr>
<td><strong>Chemistry</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1991-1995</td>
<td>15.7</td>
<td>14.8</td>
<td>15.1</td>
<td>16.1</td>
</tr>
<tr>
<td>1996-2000</td>
<td>20.9</td>
<td>19.9</td>
<td>21.4</td>
<td>12.6</td>
</tr>
<tr>
<td>2001-2007</td>
<td>26.1</td>
<td>24.3</td>
<td>25.0</td>
<td>16.3</td>
</tr>
<tr>
<td><strong>Physics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1996-2000</td>
<td>38.3</td>
<td>34.8</td>
<td>37.8</td>
<td>36.8</td>
</tr>
<tr>
<td>2001-2007</td>
<td>44.4</td>
<td>40.0</td>
<td>46.0</td>
<td>37.1</td>
</tr>
<tr>
<td><strong>Economics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1991-1995</td>
<td>12.1</td>
<td>7.6</td>
<td>6.4</td>
<td>6.5</td>
</tr>
<tr>
<td>1996-2000</td>
<td>16.8</td>
<td>10.8</td>
<td>10.9</td>
<td>8.7</td>
</tr>
<tr>
<td>2001-2007</td>
<td>21.2</td>
<td>16.7</td>
<td>15.9</td>
<td>14.7</td>
</tr>
</tbody>
</table>

Note: Calculated for institution-year with > 4 pubs.
<table>
<thead>
<tr>
<th></th>
<th>Top Research/Doc</th>
<th>Other Research/Doc</th>
<th>Master's Level</th>
<th>Top Liberal Arts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Median</td>
<td>Mean</td>
<td>Median</td>
</tr>
<tr>
<td>USUS_change</td>
<td>9.225</td>
<td>6</td>
<td>1.634</td>
<td>1</td>
</tr>
<tr>
<td>USINTL_change</td>
<td>4.670</td>
<td>3</td>
<td>0.719</td>
<td>0</td>
</tr>
<tr>
<td>pub_change</td>
<td>11.619</td>
<td>10</td>
<td>1.957</td>
<td>1</td>
</tr>
<tr>
<td>exp 0-4</td>
<td>0.086</td>
<td>0</td>
<td>0.224</td>
<td>0</td>
</tr>
<tr>
<td>exp 5-9</td>
<td>0.473</td>
<td>0</td>
<td>0.473</td>
<td>0</td>
</tr>
<tr>
<td>exp 10+</td>
<td>0.441</td>
<td>0</td>
<td>0.303</td>
<td>0</td>
</tr>
</tbody>
</table>

**Economics**

<table>
<thead>
<tr>
<th></th>
<th>Top Research/Doc</th>
<th>Other Research/Doc</th>
<th>Master's Level</th>
<th>Top Liberal Arts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Median</td>
<td>Mean</td>
<td>Median</td>
</tr>
<tr>
<td>USUS_change</td>
<td>0.672</td>
<td>0</td>
<td>0.542</td>
<td>0</td>
</tr>
<tr>
<td>USINTL_change</td>
<td>0.438</td>
<td>0</td>
<td>0.198</td>
<td>0</td>
</tr>
<tr>
<td>pub_change</td>
<td>0.144</td>
<td>0</td>
<td>0.597</td>
<td>1</td>
</tr>
<tr>
<td>exp 0-4</td>
<td>0.083</td>
<td>0</td>
<td>0.226</td>
<td>0</td>
</tr>
<tr>
<td>exp 5-9</td>
<td>0.472</td>
<td>0</td>
<td>0.472</td>
<td>0</td>
</tr>
<tr>
<td>exp 10+</td>
<td>0.446</td>
<td>0</td>
<td>0.302</td>
<td>0</td>
</tr>
</tbody>
</table>

Notes: All observations are restricted to >4 observations for each field for each year. Years 1992-2001.