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

  Jones, Wuchty & Uzzi (2008) analyzed publication patterns (sole-authored, multi-authored within same institution, multi-authored across institutions) using ISI data for 1975-2005.

  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

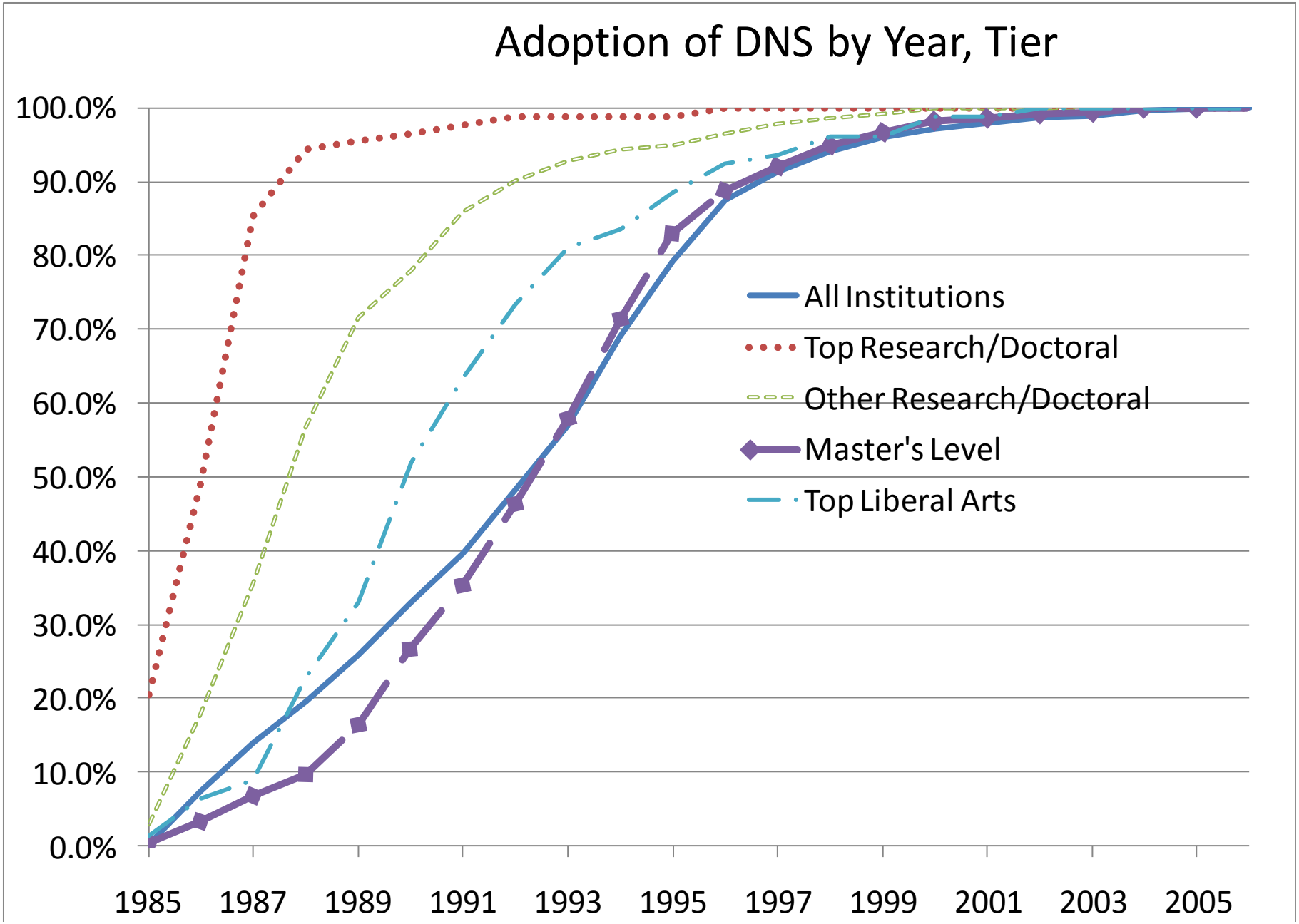
IT Measure: DNS

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

  Invented in 1994; by 2001, virtually fully diffused.

  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
100.0%
90.0%
80.0%
70.0%
60.0%
50.0%
40.0%
30.0%
20.0%
10.0%
0.0%
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
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

	All Tiers		Top Research/Doc		Other Research/Doc		Master's		Top Liberal Arts	
Field	Mean Pubs	% Zero Pubs	Mean Pubs	% Zero Pubs	Mean Pubs	% Zero Pubs	Mean Pubs	% Zero Pubs	Mean Pubs	% Zero Pubs
All										
1991-1995	158.61	24.3%	1,730	0.0%	264.96	0.9%	21.93	16.7%	21.63	2.5%
1996-2000	186.55	20.1%	2,029	0.0%	311.47	0.4%	26.60	12.1%	25.94	2.8%
2001-2007	227.69	18.5%	2,459	0.0%	388.07	0.1%	33.12	10.9%	32.76	2.7%
Biology										
1991-1995	17.34	64.2%	204.20	0.0%	23.07	9.9%	1.59	68.4%	1.58	44.8%
1996-2000	22.49	58.3%	263.28	0.0%	30.31	7.5%	2.22	60.8%	2.15	34.2%
2001-2007	26.51	54.1%	307.75	0.0%	37.11	5.1%	2.69	53.8%	2.54	27.3%
Chemistry										
1991-1995	16.26	56.8%	166.45	0.0%	34.94	7.4%	2.00	56.4%	2.21	33.2%
1996-2000	19.26	53.5%	196.22	0.0%	41.22	5.7%	2.51	52.9%	2.64	26.6%
2001-2007	22.91	51.2%	229.09	0.0%	50.28	5.6%	3.36	49.7%	3.31	20.4%
Physics										
1991-1995	19.37	61.2%	215.95	0.0%	32.29	9.4%	2.01	64.3%	2.18	35.4%
1996-2000	22.07	58.3%	245.90	0.0%	36.76	7.5%	2.30	59.7%	2.49	32.9%
2001-2007	27.29	55.1%	292.38	0.0%	49.91	6.6%	3.52	54.7%	3.58	23.9%
Economics										
1991-1995	5.22	62.2%	49.80	1.1%	10.93	10.9%	1.23	62.0%	1.49	40.3%
1996-2000	5.58	58.8%	51.74	0.9%	12.08	8.7%	1.43	56.1%	1.63	38.0%
2001-2007	6.43	58.6%	59.10	1.3%	14.39	8.9%	1.61	55.5%	1.87	36.7%

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

$\Rightarrow$Data are very skewed

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

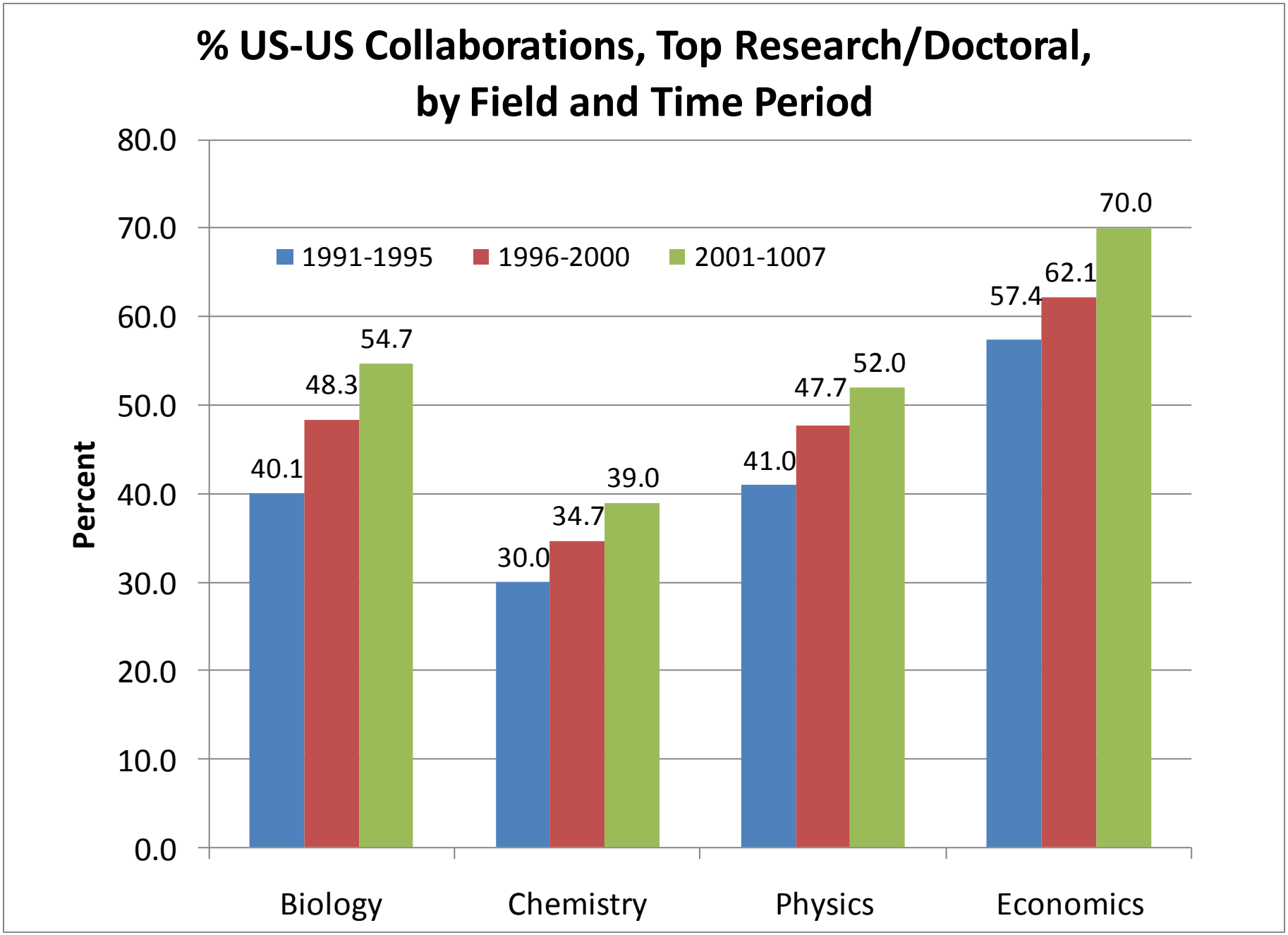
By field:

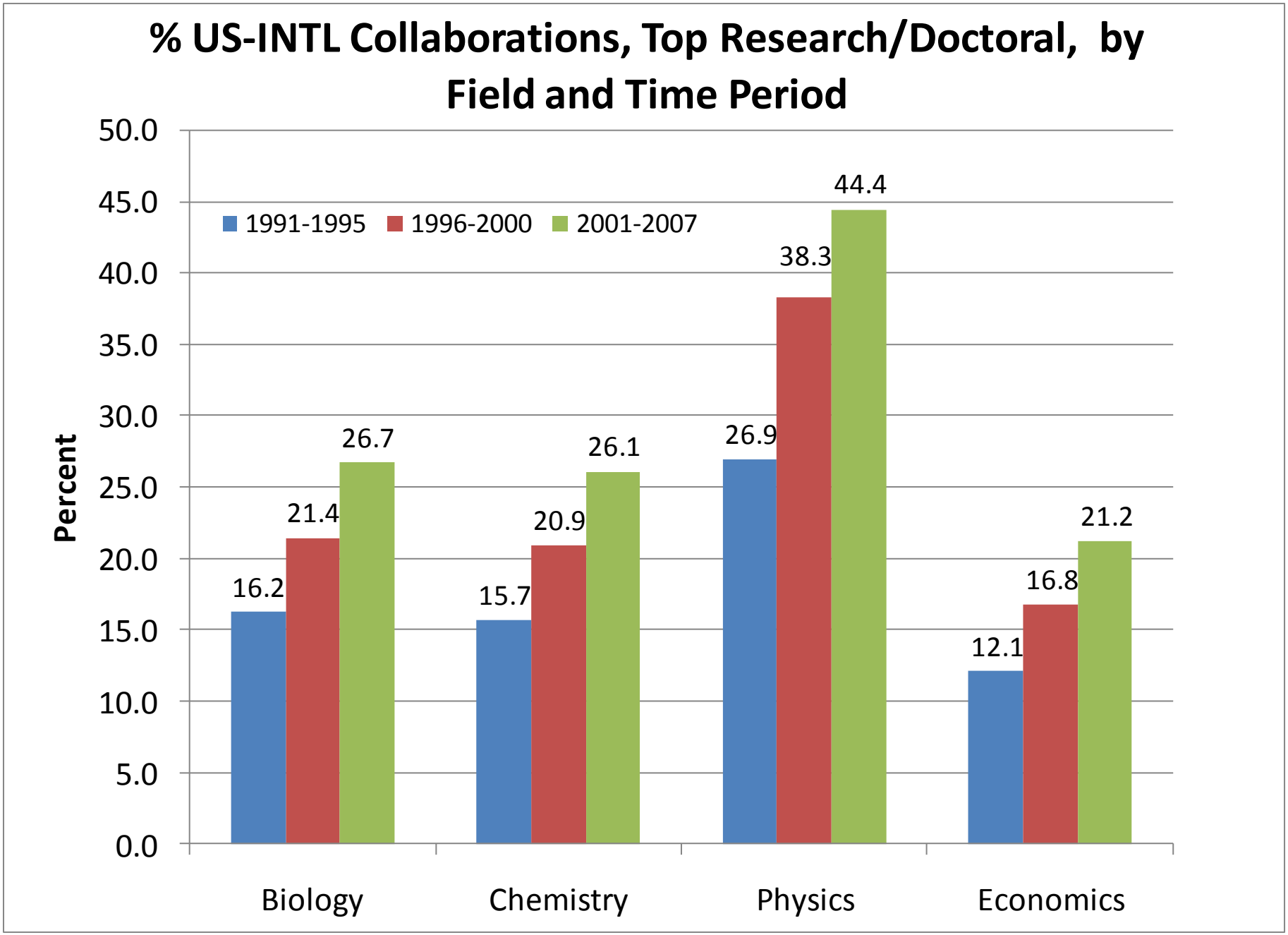
- Mean pubs in Biology increased from 204 to 308
- Mean pubs in Economics increased from 50 to 59

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
Percent
80.0
70.0
60.0
50.0
40.0
30.0
20.0
10.0
0.0
1991-1995
1996-2000
2001-1007
40.1
48.3
54.7
30.0
34.7
39.0
41.0
47.7
52.0
57.4
62.1
70.0
Biology
Chemistry
Physics
Economics


% US-INTL Collaborations, Top Research/Doctoral, by Field and Time Period
Percent
50.0
45.0
40.0
35.0
30.0
25.0
20.0
15.0
10.0
5.0
0.0
1991-1995
1996-2000
2001-2007
16.2
21.4
26.7
15.7
20.9
26.1
26.9
38.3
44.4
12.1
16.8
21.2
Biology
Chemistry
Physics
Economics

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"

$$USUS\_change_{i,t} = B_o + B_1\ Pub\_change_{i,t} + B_2\ EXP_{i,t-1} + \varepsilon_{i,t}$$

where

Pub_change = change in total number of publications at institution i in year j

USUS_change = year-to-year change in number of publications by institution i with at least one co-author from another institution

EXP = 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)

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

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)

	Top Research/Doc	Other Research/Doc	Master's	Top Liberal Arts
	%	%	%	%
Biology				
1991-1995	16.2	14.0	16.6	12.9
1996-2000	21.4	18.2	19.1	20.4
2001-2007	26.7	24.4	21.9	17.7
Chemistry				
1991-1995	15.7	14.8	15.1	16.1
1996-2000	20.9	19.9	21.4	12.6
2001-2007	26.1	24.3	25.0	16.3
Physics				
1991-1995	26.9	24.2	24.8	23.3
1996-2000	38.3	34.8	37.8	36.8
2001-2007	44.4	40.0	46.0	37.1
Economics				
1991-1995	12.1	7.6	6.4	6.5
1996-2000	16.8	10.8	10.9	8.7
2001-2007	21.2	16.7	15.9	14.7

Note: Calculated for institution-year with > 4 pubs.

Table 4. Summary Statistics for Variables Used in Regressions (Biology and Economics)

Biology

	Top Research/Doc		Other Research/Doc		Master's Level		Top Liberal Arts	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median
USUS_change	9.225	6	1.634	1	1.339	1	1.420	2
USINTL_change	4.670	3	0.719	0	0.481	0	0.352	0
pub_change	11.619	10	1.957	1	2.042	2	2.295	2
exp 0-4	0.086	0	0.224	0	0.375	0	0.239	0
exp 5-9	0.473	0	0.473	0	0.397	0	0.420	0
exp 10+	0.441	0	0.303	0	0.228	0	0.341	0

Economics

	Top Research/Doc		Other Research/Doc		Master's Level		Top Liberal Arts	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median
USUS_change	0.672	0	0.542	0	0.946	1	1.452	2
USINTL_change	0.438	0	0.198	0	0.170	0	0.242	0
pub_change	0.144	0	0.597	1	1.481	2	2.677	3
exp 0-4	0.083	0	0.226	0	0.386	0	0.161	0
exp 5-9	0.472	0	0.472	0	0.451	0	0.419	0
exp 10+	0.446	0	0.302	0	0.163	0	0.419	0

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