Is distance from innovation a barrier to the adoption of Artifical Intelligence?

Jim Bessen, Iain Cockburn, and Jennifer Hunt

Boston University, Boston University and NBER, Rutgers University and NBER

December 13, 2022

Question

Do firms at a greater distance from US commuting zones with historical strength in AI scientific publications delay adopting AI?

Geographic units

- U.S. commuting zones (CZ)
  - 741
  - Cover all of U.S.

Measure of innovation

- Number of AI publications in CZ through 2006
  - Academic journal papers
  - Conference proceedings
  - Patents (2.7%)
- Created own geocoded dataset
  - Using Microsoft Academic Graph

Measure of adoption

- Share of CZ's online job advertisements requiring AI skill
  - Especially for use of AI–based applications
  - Especially in industries like finance
- Data from Burning Glass Technologies 2007–2019
  - Scrapes and clean data from online job boards
  - 200 million observations
- Use long differences (show here: 7; also 12 years)

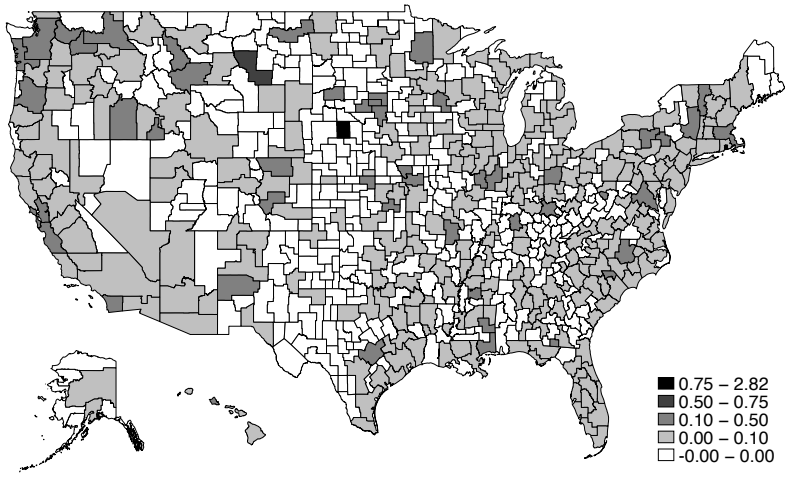
Key distance measure

1. Log km to closest AI publication hotspot CZ
2. Main definition of hotspot
   - Main results: at least 1000 AI publications (32 hotspots)
   - But vary threshold

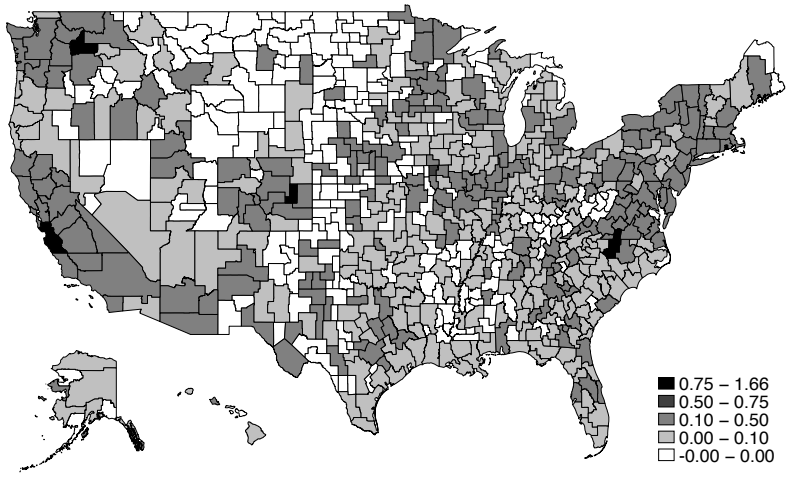
Key other controls

- Help control for unobserved determinants of AI adoption
  - Initial level and change in
    - ★ (Non-AI) IT's share in job ads
    - ★ AI publications
- Distinguish from distance to large city
  - Distance to closest large CZ

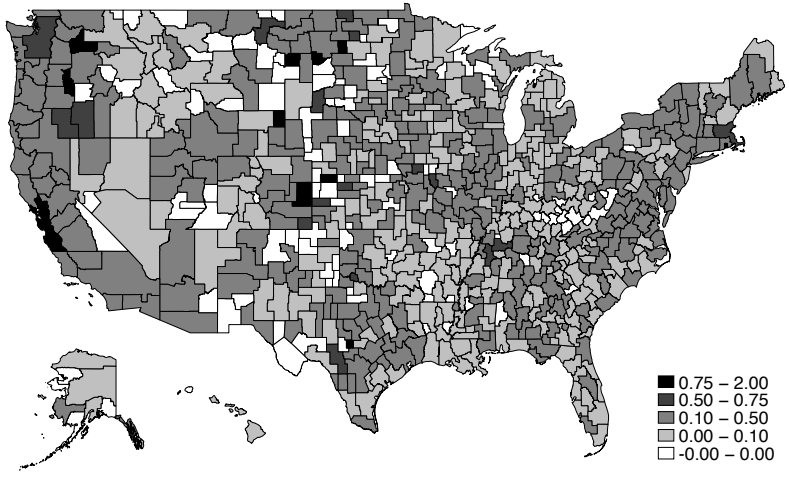
AI job ads as share of all job ads (%)



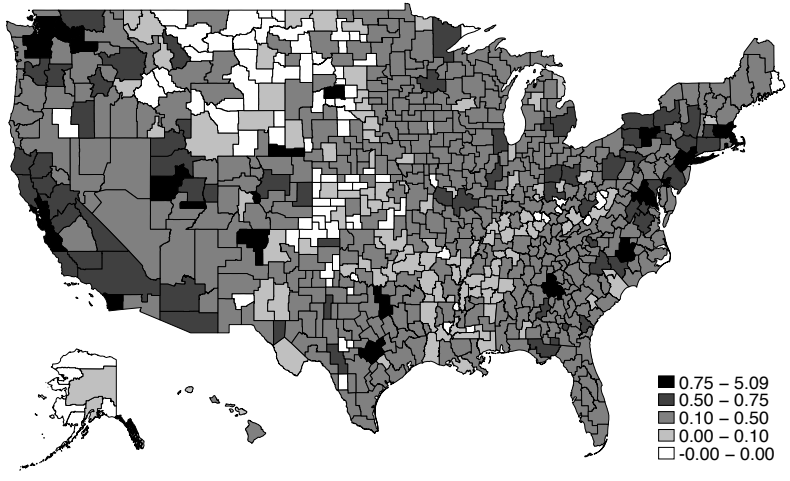
(a) 2007



(b) 2010

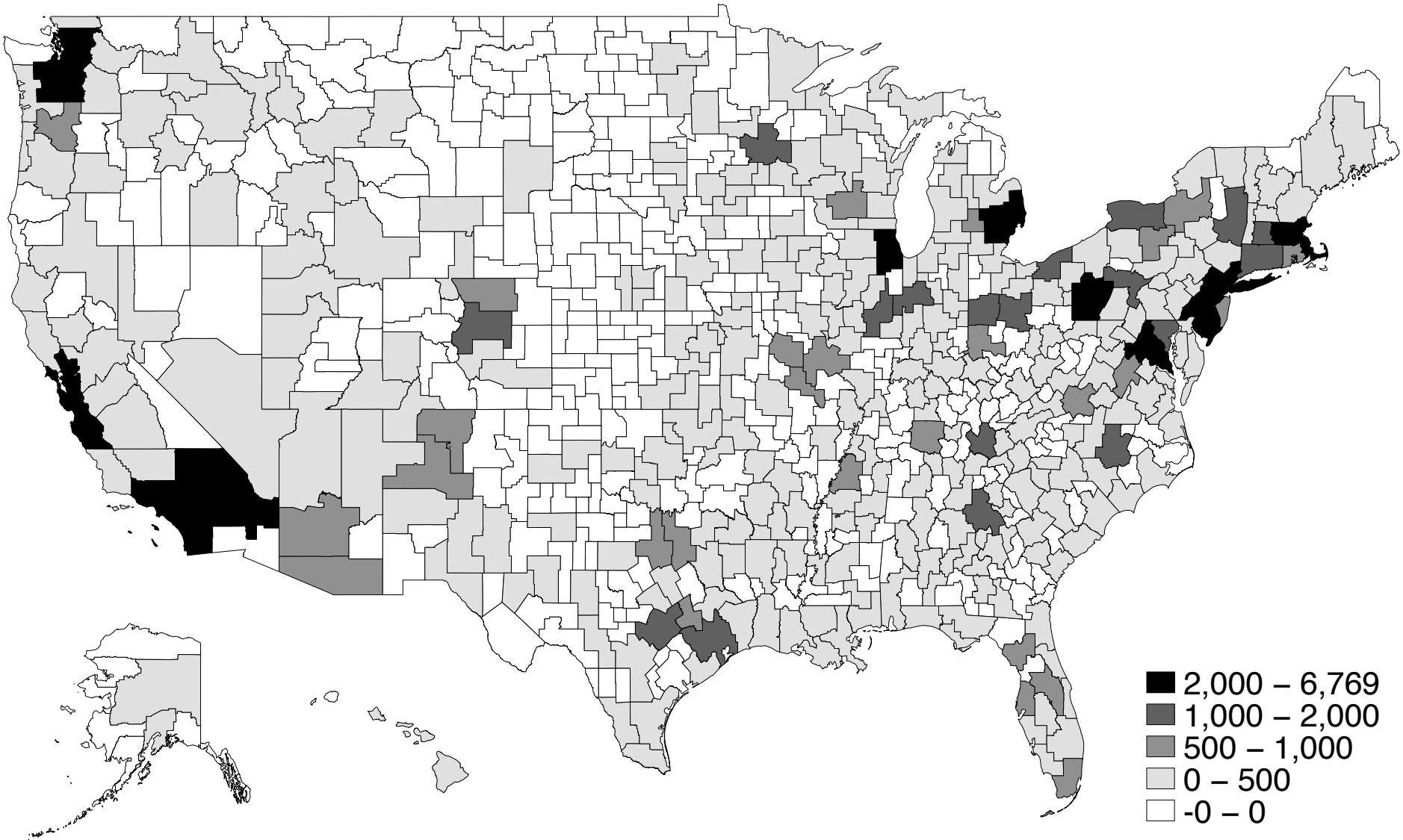


(c) 2014



(d) 2018

Geographic distribution of AI publications through 2006



Main equation

$$
\begin{aligned}
\Delta^k AI^s_{ct} = & \alpha + \sigma log(D^{Hot}_c) \\
& + \beta_1 AI\ Pub > 0_{c,t^*} + \beta_2 AI\ Pubs_{c,t^*} + \beta_3 (AI\ Pubs_{c,t^*})^2 \\
& + \gamma_1 log(All\ job\ ads_{c,t^*}) + \gamma_2 log(Pop_{c,t^*}) \\
& + \nu IT^s_{c,t^*} \\
& + \phi_1 log(\bar{D}_c) + \phi_2 log(D^{Pop}_c) + \phi_3 log(D^{min}_c) \\
& + \rho_1 \Delta^k AI\ Pubs_{c,t} + \rho_2 \Delta^k log(All\ job\ ads_{c,t}) + \rho_3 \Delta^k IT^s_{c,t} \\
& + \eta_t + \Delta^k \epsilon_{ct},
\end{aligned}
$$

Main equation

$$
\begin{aligned}
\Delta^k AI^s_{ct} = & \alpha + \sigma log(D^{Hot}_c) \\
& + \beta_1 AI\ Pub > 0_{c,t^*} + \beta_2 AI\ Pubs_{c,t^*} + \beta_3 (AI\ Pubs_{c,t^*})^2 \\
& + \gamma_1 log(All\ job\ ads_{c,t^*}) + \gamma_2 log(Pop_{c,t^*}) \\
& + \nu IT^s_{c,t^*} \\
& + \phi_1 log(\bar{D}_c) + \phi_2 log(D^{Pop}_c) + \phi_3 log(D^{min}_c) \\
& + \rho_1 \Delta^k AI\ Pubs_{c,t} + \rho_2 \Delta^k log(All\ job\ ads_{c,t}) + \rho_3 \Delta^k IT^s_{c,t} \\
& + \eta_t + \Delta^k \epsilon_{ct},
\end{aligned}
$$

Main equation

$$
\begin{aligned}
\Delta^k AI^s_{ct} = & \alpha + \sigma log(D^{Hot}_c) \\
& + \beta_1 AI\ Pub > 0_{c,t^*} + \beta_2 AI\ Pubs_{c,t^*} + \beta_3 (AI\ Pubs_{c,t^*})^2 \\
& + \gamma_1 log(All\ job\ ads_{c,t^*}) + \gamma_2 log(Pop_{c,t^*}) \\
& + \nu IT^s_{c,t^*} \\
& + \phi_1 log(\bar{D}_c) + \phi_2 log(D^{Pop}_c) + \phi_3 log(D^{min}_c) \\
& + \rho_1 \Delta^k AI\ Pubs_{c,t} + \rho_2 \Delta^k log(All\ job\ ads_{c,t}) + \rho_3 \Delta^k IT^s_{c,t} \\
& + \eta_t + \Delta^k \epsilon_{ct},
\end{aligned}
$$

Main equation

$$
\begin{aligned}
\Delta^k AI^s_{ct} =& \alpha + \sigma log(D^{Hot}_c) \\
&+ \beta_1 AI\ Pub > 0_{c,t^*} + \beta_2 AI\ Pubs_{c,t^*} + \beta_3 (AI\ Pubs_{c,t^*})^2 \\
&+ \gamma_1 log(All\ job\ ads_{c,t^*}) + \gamma_2 log(Pop_{c,t^*}) \\
&+ \nu IT^s_{c,t^*} \\
&+ \phi_1 log(\bar{D}_c) + \phi_2 log(D^{Pop}_c) + \phi_3 log(D^{min}_c) \\
&+ \rho_1 \Delta^k AI\ Pubs_{c,t} + \rho_2 \Delta^k log(All\ job\ ads_{c,t}) + \rho_3 \Delta^k IT^s_{c,t} \\
&+ \eta_t + \Delta^k \epsilon_{ct},
\end{aligned}
$$

Main equation

$$
\begin{aligned}
\Delta^k AI^s_{ct} = & \alpha + \sigma log(D^{Hot}_c) \\
& + \beta_1 AI\ Pub > 0_{c,t^*} + \beta_2 AI\ Pubs_{c,t^*} + \beta_3 (AI\ Pubs_{c,t^*})^2 \\
& + \gamma_1 log(All\ job\ ads_{c,t^*}) + \gamma_2 log(Pop_{c,t^*}) \\
& + \nu IT^s_{c,t^*} \\
& + \phi_1 log(\bar{D}_c) + \phi_2 log(D^{Pop}_c) + \phi_3 log(D^{min}_c) \\
& + \rho_1 \Delta^k AI\ Pubs_{c,t} + \rho_2 \Delta^k log(All\ job\ ads_{c,t}) + \rho_3 \Delta^k IT^s_{c,t} \\
& + \eta_t + \Delta^k \epsilon_{ct},
\end{aligned}
$$

Main equation

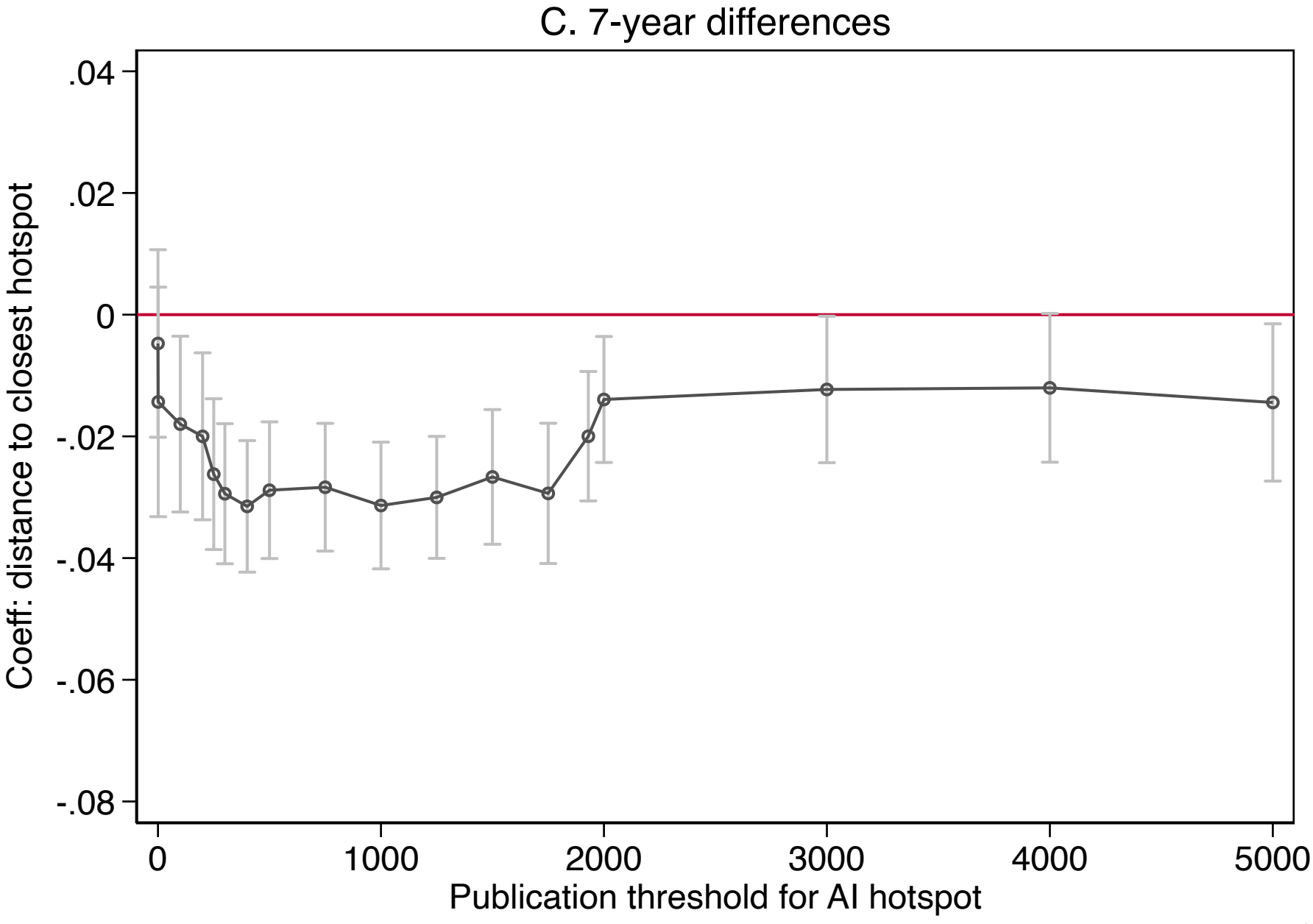
$$
\begin{aligned}
\Delta^k AI^s_{ct} =& \alpha + \sigma log(D^{Hot}_c) \\
&+ \beta_1 AI\ Pub > 0_{c,t^*} + \beta_2 AI\ Pubs_{c,t^*} + \beta_3 (AI\ Pubs_{c,t^*})^2 \\
&+ \gamma_1 log(All\ job\ ads_{c,t^*}) + \gamma_2 log(Pop_{c,t^*}) \\
&+ \nu IT^s_{c,t^*} \\
&+ \phi_1 log(\bar{D}_c) + \phi_2 log(D^{Pop}_c) + \phi_3 log(D^{min}_c) \\
&+ \rho_1 \Delta^k AI\ Pubs_{c,t} + \rho_2 \Delta^k log(All\ job\ ads_{c,t}) + \rho_3 \Delta^k IT^s_{c,t} \\
&+ \eta_t + \Delta^k \epsilon_{ct},
\end{aligned}
$$

Effect of distance to closest hotspot on AI jobs' share in job ads

	Median regression				OLS
Ln distance hotspot	-0.029***	-0.031***	-0.031***	-0.024***	-0.049***
	(0.005)	(0.006)	(0.005)	(0.007)	(0.008)
X's	--	Y	Y	Y	Y
ΔX's	--	--	Y	Y	Y
Ln distance large CZ	--	--	--	Y	--

2964 obs; year dummies, controls for AI publications thru 2006 included

Coefficients on distance to closest AI hotspot, median regressions



Magnitudes of effect on median growth in AI job ad share

- Std dev of distance to closest hotspot is $\approx$10% mean (8.9%)
- 1 standard deviation reduces growth by 2-3% of median

Types of AI required (microdata, ads with valid occ)

	All AI	AI App	Unspecified AI only	AI Tool	AI R&D	Image Processing
% AI type	100%	19.2%	37.1%	9.1%	34.4%	12.5%
% Comp sci	62.6%	52.2%	68.3%	80.2%	66.5%	49.9%

Effect of distance to hotspot on AI jobs' share, by AI type

	Median regression					OLS
	All	App	Unspec only	Tool	R&D	Image Process
Ln distance hotspot	-0.031***	-0.0013	-0.017***	-0.0025***	-0.0069***	0.0030*
	(0.005)	(0.0010)	(0.002)	(0.0004)	(0.0016)	(0.0017)
Median dep var (pp)	0.093	0.004	0.033	0.000	0.017	0.000
Dist ↑ 10% (% of med)	-3.3%	-3.3%	-5.2%	--	-4.1%	--

2964 observations; year dummies, full controls included

Effect of distance to hotspot on AI jobs' share, by industry

	Share	Mean dep var	OLS
All	100.0%	0.14	-0.049***
			(0.007)
Ag, Utilities, Mining, Construction, Manuf	9.0%	0.08	0.002
			(0.014)
Wholesale trade, Retail trade, Warehousing, Transport	12.3%	0.02	0.004
			(0.008)
Information	3.0%	0.21	-0.079*
			(0.042)
Finance, Insurance	7.6%	0.28	-0.061**
			(0.027)
Real Estate, Prof-scientific services, Administration	17.9%	0.24	-0.090**
			(0.030)
Missing industry	16.0%	0.25	-0.136***
			(0.013)

2964 observations; year dummies, full controls included

Missing industry or firm name

- Most ads with missing industry have missing employer name
- BGT says employer name missing = employment agency

Effect of distance to hotspot on AI jobs' share, by micro sample

Job ads included	All	Valid industry	Missing industry	Valid employer	Missing employer
Log distance hotspot	-0.031***	-0.012***	-0.094***	-0.005	-0.067***
	(0.005)	(0.004)	(0.011)	(0.003)	(0.008)
Median dep var (pp)	0.093	0.060	0.144	0.047	0.103
Distance ↑ 10% (% of med)	-3.3%	-2.0%	-6.5%	-1.1%	-6.5%

2964 observations; year dummies, full controls included

Focusing on adoption (or adaptation)

- Effects driven by employment agency ads (1/3 ads)
  - So for sorts of firms using agencies, distance barrier 3x as large
- Distance barrier for finance industry
- Distance barrier for AI tools and applications

Determinants of missing employer name, micro data

	All	Finance, ins	AI apps
College educated share	-0.727***	-0.666***	-1.596***
of 2000 labor force	(0.067)	(0.082)	(0.233)
Log av distance other CZs	0.064***	0.048*	0.051
	(0.015)	(0.019)	(0.054)
AI publications <2007	0.004**	0.002	0.005
(/1000)	(0.002)	(0.003)	(0.009)
# skills required	-0.0122***	-0.0061***	-0.0019***
	(0.0003)	(0.0003)	(0.0004)
Management	-0.130***	-0.011**	-0.157***
	(0.008)	(0.005)	(0.016)
Business and finance	-0.051***	0.095***	-0092***
	(0.007)	(0.007)	(0.019)
Computer science-math	--	--	--
Architecture-engineering	-0.037***	0.010*	0.002
	(0.008)	(0.006)	(0.017)
IT skill required	0.053***	0.051***	--
	(0.005)	(0.005)	

Conclusions

- Distance from AI innovation is barrier to adoption
  - As long as CZ not itself a hotspot
- For types of firms using employment agencies
  - Firms without suitable local workers
- So distance deters AI innovation workers from moving to locations that do not have workers suitable for AI adoption

Effect of pre-2007 AI publications

Job ads sample:	Employer name is: Valid	Missing
Ln distance hotspot	-0.003	-0.087***
	(0.003)	(0.011)
Pre-2007 AI pubs	p=0.00	p=0.00
Interactions pre-2007, distance	p=0.01	p=0.01
Distance evaluated at		
1 AI pub pre-2007	-0.009	-0.063**
	(0.006)	(0.013)
500 AI pubs pre-2007	-0.001	-0.040***
	(0.013)	(0.010)
1000 AI pubs pre-2007	0.002	-0.021
	(0.022)	(0.016)

2964 observations; year dummies, full controls included

Effect of pre-2007 AI publications

Micro sub-sample	Employer name		Industry	Skills
	Valid	Missing	Finance, ins	AI app
Ln distance hotspot	-0.005	-0.067***	-0.061**	-0.0013
	(0.003)	(0.008)	(0.027)	(0.0010)
Change in AI pubs (x1000)	0.380***	0.187**	0.299***	0.024**
	(0.006)	(0.082)	(0.045)	(0.009)
Pre-2007 AI pub controls	p=0.00	p=0.10	p=0.60	p=0.00
Effect of pre-2007 AI pubs at				
Change 0 to 1 AI pub	0.093***	-0.044**	-0.049	0.007*
	(0.020)	(0.021)	(0.080)	(0.004)
100 AI pubs/1000	0.098***	-0.033	0.016	0.008*
	(0.022)	(0.020)	(0.065)	(0.004)
1000 AI pubs/1000	0.065***	-0.033	0.006	0.005
	(0.015)	(0.016)	(0.047)	(0.004)

2964 observations; year dummies, full controls included