Colocation of Production and Innovation: Evidence from the United States

Teresa C Fort
Tuck School at Dartmouth & NBER & CEPR

Wolfgang Keller
U Colorado & NBER & CEPR

Peter K Schott
Yale SOM & NBER & CEPR

Stephen Yeaple
Penn State & NBER

Nikolas Zolas
US Census Bureau

Preliminary! April 2021

Disclaimer

This work is unofficial and thus has not undergone the review accorded to official Census Bureau publications. The views expressed in the paper are those of the authors and not necessarily those of the U.S. Census Bureau. The Census Bureau's Disclosure Review Board and Disclosure Avoidance Officers have reviewed this data product for unauthorized disclosure of confidential information and have approved the disclosure avoidance practices applied to this release. DRB Approval Numbers CBDRB-FY20-P1916-R8726 and CBDRB-FY20-P1916-R8756.

Innovation Should Be Made In the USA (WSJ 2019.11.18)

Many U.S. political and economic leaders continue to believe that offshoring is not only profitable but also sound national economic strategy. Instead of manufacturing domestically, the thinking goes, U.S. firms should focus on higher-value work: "innovate here, manufacture there."

Today many Americans are rightly questioning this perspective. **Once manufacturing departs from a country's shores, engineering and production know-how leave as well, and innovation ultimately follows. It's become increasingly clear that "manufacture there" now also means "innovate there."**

Motivation

- In the US, manufacturers traditionally have performed the majority of innovation
- Declining US manufacturing employment has raised concerns that innovation will decline
    - Less resources available for innovation
    - Innovation will follow physical production

Main questions

- How has US innovation evolved over time?
- Does innovation depend on the co-location of R&D and physical production?
- If so, which form of colocation matters most?
  - Within geographic borders?
  - Within firm boundaries?
  - Within both?

Why shocks to manufacturing may affect innovation

- Complementarities between production and R&D?
  - Face-to-face interactions about feasibility, prototypes, etc?
- Gains from reallocation?
  - Lower production costs, *e.g.*, from offshoring, allow reallocation towards R&D?
  - Agglomeration benefits from knowledge spillovers and labor pooling of R&D in cities?

Main findings

- Shift in US patenting to non-manufacturing firms (*NMF*s) over time
- Former manufacturing firms (*FMF*s) continue innovating
- Firms containing manufacturing (*M*) and innovation (*P*) plants patent more
- Within *MP* firms, patenting is (≈12%) greater if they are within 5 miles
- *MP* firms' *M* and *P* plants are spreading out over time
- *Future plans*
  - *Does patenting occur within co-located plants?*
  - *Why is the distance between co-located plants changing?*

Related literature

- Economic geography of innovation and production
  - Jaffe et al. (1993); Audretsch and Feldman (1996); Duranton and Puga (2001); Ellison, Glaeser, Kerr (2010); Pisano and Shih (2012); Buzard and Carlino (2013); Tecu (2013); Alcacer and Delgado (2016); Buzard et al. (2017); Lan (2019); Davis and Dingel (2019); Delgado (2020); Berkes et al. (2020)
- Evolution of manufacturers and innovators
  - Bloom et al. (2015); Bernard and Fort (2015); Bernard et al. (2017); Fort (2017); Kamal (2018); Fort et al. (2018); Ding et al. (2019); Autor et al. (*forthcoming*);
- Innovation and offshoring
  - Naghavi and Ottaviano (2009); Rodriguez-Clare (2010); Fuchs and Kirchain (2010); Fuchs (2014); Bøler et al. (2015); Arkolakis et al. (2018); Bilir and Morales (2020) ; Bernard et al. (2020)

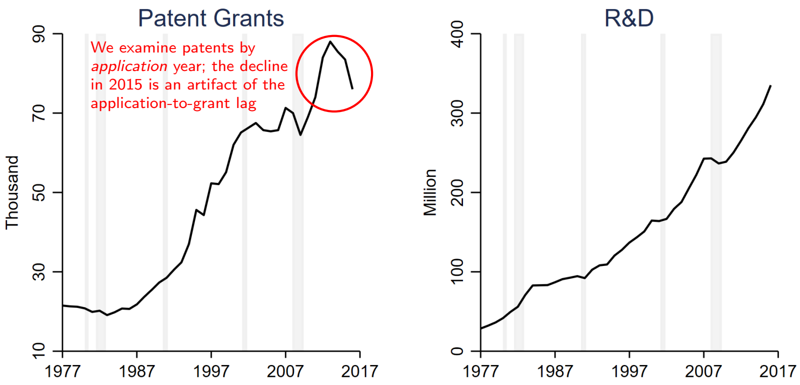
Outline of Talk

- Portrait of US innovation
- Do firms with both manufacturing and innovation plants patent differently?
- Within firms with both types of plants, does spatial proximity matter?
- *Future plans*

New dataset on US innovation from 1977 to 2016

- Longitudinal Business Database, 1977-2016
  - Every private, non-farm employer establishment
  - Consistent establishment-level NAICS classification (Fort and Klimek 2018)
- Business Register, 1977-2016
  - Geocodes (addresses, latitude and longitude)
- Economic Censuses, 1977(5)2012
  - Establishment-level sales, inputs, etc. for manuf, wholesale, retail, and services
- Longitudinal foreign trade transactions database, 1992-2016
  - Firm-level import and export transactions
- USPTO PatentView database, 1973-2018
  - Identify manufacturing and processing patents
- SIRD and BRDIS R&D surveys, 1977-2016

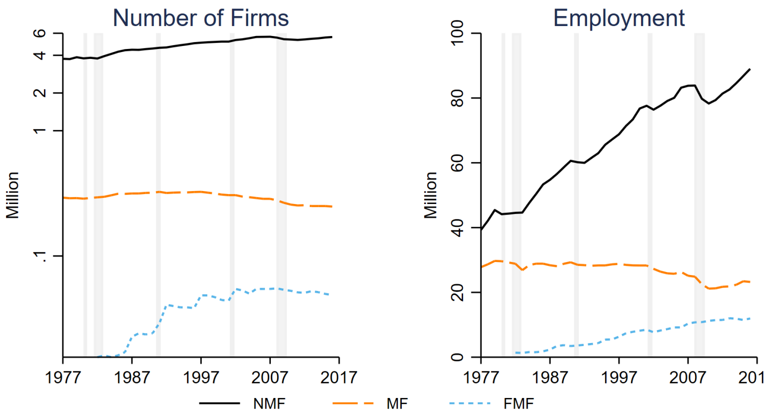
US innovation grows over the last 40 years



We decompose patents by firm *type*

- Classify firms into 3 mutually-exclusive types for each year t
  - MF: manufacturing firm ($\geq$1 manufacturing plant in year t)
  - NMF: non-manufacturing firm (0 manufacturing plants up to t)
  - FMF: former manufacturing firm ($\geq$1 manufacturing plant prior to t; 0 in t)
- Note
  - By definition, NMF can later switch into MF but not back (rare)
  - We later focus on a subset of MFs, MP firms, which have both manufacturing (M) and innovation (P) plants within their boundaries

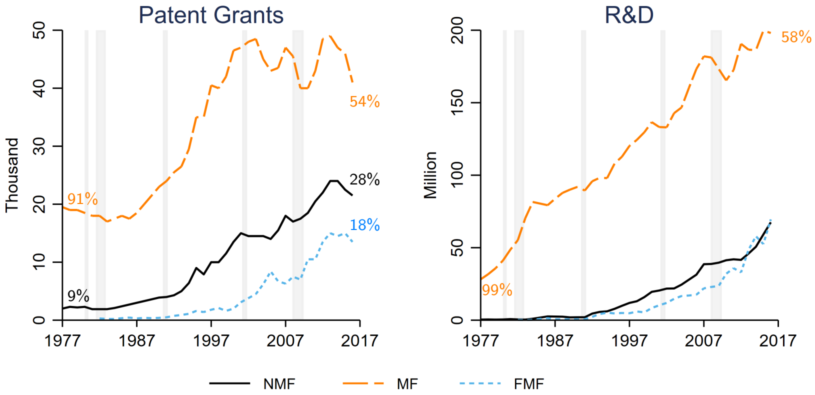
*NMF*s and *FMF*s dominate firms and employment



*MF*s have non-manufacturing employment

*FMF*s have considerable growth

*MF*s dominate US innovation



Summary of new facts

- Manufacturing firms' dominance of innovation has declined substantially
- Some recent cohorts of former manufacturing firms continue patenting intensively
- Imports by patenting firms suggestive of offshoring

Outline of Talk

- Portrait of US innovation
- Do firms with both manufacturing and innovation plants patent differently?
- Within firms with both types of plants, does spatial proximity matter?
- *Future plans*

Example: Bristol Meyers Squibb, North America

Source: Google Maps and author's calculations. Locations of BMS North American facilities are publicly available at `https://www.bms.com/about-us/our-company/worldwide-facilities.html`.

Identifying innovation (P) establishments

- How to identify P establishments?
- We classify establishments with the following activities as P plants
  - NAICS 5417: Scientific Research and Development Services (i.e., R&D labs)
  - NAICS 551114: Corporate, Subsidiary, and Regional Managing Offices (i.e., HQs)
  - NAICS 5413-5416: Professional Scientific and Technical Services
  - NAICS 5112, 517, 518: Information and Telecommunications
- Descriptive regressions
  - (1) How does firm patenting vary with the presence of *both* M and P plants within the firm?
  - (2) Within MP firms, how does patenting vary with the spatial proximity of M and P plants?
  - (3) Does MP firm patenting occur in their spatially located M-P plants?

▸ NAICS

Firm patenting and the *presence* of both M and P plants within the firm

$$
\begin{aligned}
ln(\widetilde{y}_{ft}) = \quad & \gamma_1 M_{ft} + \gamma_2 P_{ft} + \gamma_3 M_{ft} \times P_{ft} + \\
& \gamma_4 FMF_{ft} + \gamma_5 FMF_{ft} \times P_{ft} + \\
& \beta X_{ft} + \alpha_t + \alpha_c + \varepsilon_{fct}
\end{aligned}
$$

- $ln(\widetilde{y}_{ft})$: log number of patents granted to firm f applied for in years $t{:}t+4$ ($sinh^{-1}$ transform)
- M_{ft}, P_{ft}: indicators that firm f has a M or P plant in year t
- FMF_{ft}: indicator that firm f is a former manufacturing firm in t
- X_{ft}: time-varying firm size and age controls
- α_t, α_c: year and county fixed effects
- Omitted category: firms with no M or P estabs in year t
- Sample: *MP* firms, Census years ending in "2" and "5" from 1977 to 2012

Patenting is highest for firms with both M and P plants

Dependent variable is $ln(Patents_{f,t:t+4})$: firm f's total patent grants applied for in years t:$t+4$

	(1)	(2)	(3)	(4)	(5)	(6)
M_{ft}	0.0374***			0.0149***		
	(0.0003)			(0.0009)		
P_{ft}	0.0172***			0.0047***		
	(0.0004)			(0.0006)		
$M_{ft} \times P_{ft}$	0.665***			0.147***		
	(0.0132)			(0.0082)		
FMF_{ft}						
$FMF_{ft} \times P_{ft}$						
Emp_{ft}, Age_{ft}	Yes			Yes		
County FEs	Yes			Yes		
Year FEs	Yes			Yes		
Firm FEs	No			Yes		
R-squared	0.152			0.742		
N (millions)	27			27		

→ Firms patent 15% more when they have both M and P plants versus when they don't

Notes: Dependent variables is the $sinh^{-1}$ transform of firm's patents of the sum of subsequently granted patents applied for by firm f in years t to $t+4$, with mean and standard deviation of 0.0074 and 0.1360. Standard errors clustered by firm. * p<0.10, ** p<0.05, *** p<0.01.

Patenting is highest for firms with both M and P plants

Dependent variable is $ln(Patents_{f,t:t+4})$: firm f's total patent grants applied for in years $t{:}t+4$

	(1)	(2)	(3)	(4)	(5)	(6)
M_{ft}	0.0374***	0.0376***	0.0365***	0.0149***	0.0179***	0.0174***
	(0.0003)	(0.0004)	(0.0003)	(0.0009)	(0.001)	(0.001)
P_{ft}	0.0172***	0.0172***	-0.213***	0.0047***	0.0047***	-0.0199
	(0.0004)	(0.0004)	(0.0206)	(0.0006)	(0.0006)	(0.0126)
$M_{ft} \times P_{ft}$	0.665***	0.665***	0.707***	0.147***	0.147***	0.154***
	(0.0132)	(0.0133)	(0.0149)	(0.0082)	(0.0082)	(0.0090)
FMF_{ft}		0.0046*	-0.016***		0.0081***	0.0061***
		(0.0024)	(0.0015)		(0.0019)	(0.0016)
$FMF_{ft} \times P_{ft}$			0.231***			0.0247*
			(0.0206)			(0.0127)
Emp_{ft}, Age_{ft}	Yes	Yes	Yes	Yes	Yes	Yes
County FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Firm FEs	No	No	No	Yes	Yes	Yes
R-squared	0.152	0.152	0.153	0.742	0.742	0.742
N (millions)	27	27	27	27	27	27

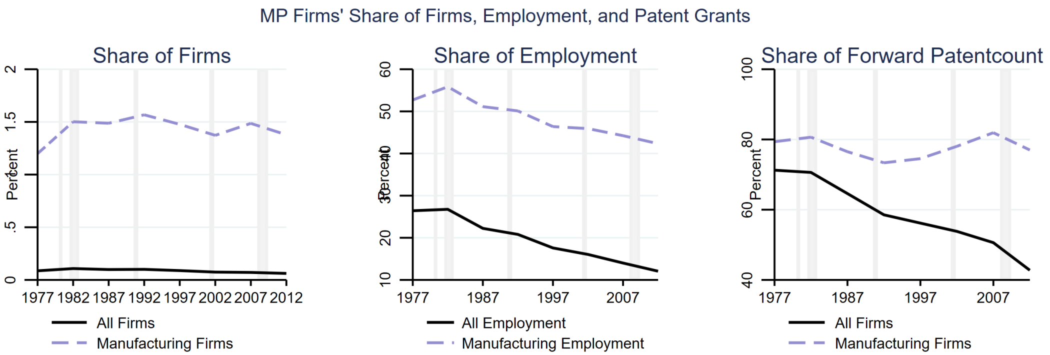
→ *FMF*s patent more than non-*FMF*s and is driven by their having P plants

Notes: Dependent variables is the $sinh^{-1}$ transform of firm's patents of the sum of subsequently granted patents applied for by firm f in years t to $t+4$, with mean and standard deviation of 0.0074 and 0.1360. Standard errors clustered by firm. * p<0.10, ** p<0.05, *** p<0.01.

Outline of Talk

- Portrait of US innovation
- Do firms with both M and P plants patent differently?
- Within MP firms, does spatial proximity matter?
- *Future plans*

MP firms are a minority but dominate patenting



FPatents

Measuring the spatial proximity of *M* and *P* *within* firms

- Use geocodes (latitude and longitude) to measure
  - $dist_{ft}^{avg}$: average distance between plants within firms, in miles
  - $dist_{ft}^{min}$: minimum distance between plants within firms, in miles
  - Examine the median and average of these firm-level measures
- In a future draft we hope to examine
  - Colocation within plants
  - Colocation across firms

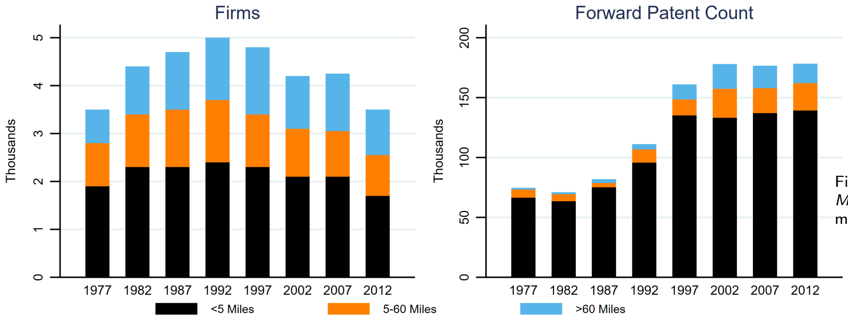
Colocation of *MP* firms' *M* and *P* plants

	$dist_{ft}^{min}$		$dist_{ft}^{avg}$	
	Mean	Median	Mean	Median
1977	95	3	445	301
1982	115	4	457	322
1987	120	5	470	336
1992	141	6	487	359
1997	153	6	502	381
2002	139	5	501	387
2007	142	5	498	383
2012	137	6	517	416

Note: Distances are in miles.

- The median firm has at least one pair of very close *M* and *P* establishments
- Average distances are much larger than minimums
  - Distances grow over time, but the minimum distance stays small j
- We create three $dist_{ft}^{min}$ bins for our regressions:
  - $<$5 miles
  - 5-60 miles
  - $>$60 miles

Distribution of *MP* firms and their patents by $dist_{ft}^{min}$ bins



Firms with closest *M-P* plants patent most

MP firm patenting and M-P plant distance

$$\begin{aligned} ln(\widetilde{y}_{ft}) = \quad & \delta_1 \left[dist_{ft}^{min} \in (0,5)\right] + \delta_2 \left[dist_{ft}^{min} \in (5,60)\right] + \\ & \gamma ln(PatentStock_{f,t-1}^{dep}) + \beta X_{ft} + \alpha_t + \alpha_c + \varepsilon_{fct} \end{aligned}$$

- $ln(\widetilde{y}_{ft})$: log number of patents granted to firm f applied for in years t:$t+4$ ($sinh^{-1}$ transform)
- $ln(\widetilde{y}_{ft})$: $sinh^{-1}$ transform of firm's granted patents applied for in t:$t+4$
- $dist_{ft}^{min}$: indicators for the minimum distance between firm's M and P plants
- $ln(PatentStock_{f,t-1}^{dep})$: firm's depreciated and 1-year lagged patent stock
- X_{ft}: time-varying firm size and age controls
- α_t, α_c: year and county fixed effects
- Omitted category: MP firms with M and P plants over 60 miles apart
- Sample: MP firms, Census years ending in "2" and "5" from 1977 to 2012

MP firm patenting is higher when *M* and *P* estabs are closer

Dependent variable is: $ln(Patents_{f,t:t+4})$	(1)	(2)	(3)	(4)
$dist^{min}_{ft} \in (0,5)$	0.131***	0.149***	0.0201	0.116***
	(0.0284)	(0.0300)	(0.0131)	(0.0279)
$dist^{min}_{ft} \in (5,60)$	-0.0230	0.0984***	0.00690	0.0764***
	(0.0303)	(0.0298)	(0.0148)	(0.0281)
$ln(Patent\ Stock^{dep}_{f,t-1})$			0.833***	0.278***
			(0.00526)	(0.0148)
Emp_{ft}, Age_{ft}	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
FIPS Fixed Effects	Yes	Yes	Yes	Yes
Firm Fixed Effects	No	Yes	No	Yes
R-Squared	0.401	0.875	0.787	0.881
Observations	34,500	34,500	34,500	34,500

Notes: Dependent variables is the $sinh^{-1}$ transform of firm's patents of the sum of subsequently granted patents applied for by firm f in years t to $t+4$, with mean and std deviation of 1.114 and 1.768. Standard errors clustered by firm. * p<0.10, ** p<0.05, *** p<0.01

→ Firms patent 12% more when $dist^{min}_{ft}$ is very small

→ A bit less when it is a bit larger

Estimated impacts are similar for citations as well as manufacturing and process patents

Dependent Variable:	$ln(Citations_{f,t:t+4})$		$ln(ManufPats_{f,t:t+4})$		$ln(ProcessingPats_{f,t:t+4})$	
	(1)	(2)	(3)	(4)	(5)	(6)
$dist_{ft}^{min} \in (0, 5)$	0.015	0.243***	0.038***	0.115***	0.044***	0.068***
	(0.026)	(0.051)	(0.013)	(0.026)	(0.012)	(0.020)
$dist_{ft}^{min} \in (5, 60)$	-0.005	0.133**	0.012	0.072***	0.004	0.042**
	(0.030)	(0.052)	(0.015)	(0.027)	(0.013)	(0.021)
$ln(Patent\ Stock_{f,t-1}^{dep})$	1.201***	0.126***	0.795***	0.264***	0.563***	0.278***
	(0.008)	(0.023)	(0.006)	(0.015)	(0.008)	(0.014)
Emp_{ft}, Age_{ft}	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
FIPS FEs	Yes	Yes	Yes	Yes	Yes	Yes
Firm FEs	No	Yes	No	Yes	No	Yes
R-squared	0.695	0.835	0.780	0.883	0.708	0.872
N (rounded)	34,500	34,500	34,500	34,500	34,500	34,500

Notes: Dependent variables is the $sinh^{-1}$ transform of firm's patents of the sum of subsequently granted patents applied for by firm f in years t to $t+4$, with mean and std deviation of 1.114 and 1.768. Standard errors clustered by firm. * p<0.10, ** p<0.05, *** p<0.01

Outline of Talk

- Portrait of US innovation
- Do firms with both M and P patent differently?
- Within MP firms, does spatial proximity matter?
- *Future plans*
    - Exploit geography of inventors

Future work

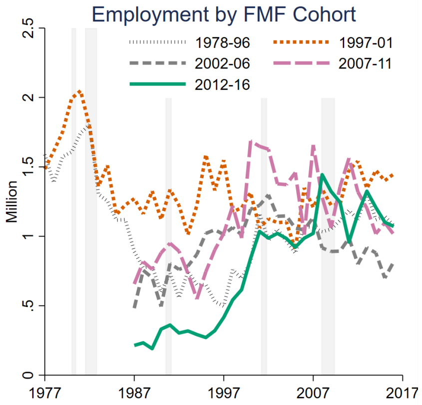
- Where does patenting occur within colocated firms?
  - Exploit inventor locations to validate whether patenting occurs in colocated plants
- Is across-firm within region colocation important?
- What are the margins of colocation adjustment?
- Explore exogenous shocks that would lead to relocation/entry/exit of plants
  - State and City-level R&D tax credits

Conclusion

- Non-manufacturers' share of patents grows from 9% to 46% between 1977 and 2016
- Firms with M and P establishments innovate most throughout period
- MP firms patent more the closer M-P plant pairs
- *Lots more work to be done*

Appendix

Former manufacturing firms' employment by cohort



- Employment dynamics are similar in 2000s
- Cohort that exits in 2002-06 least resilient

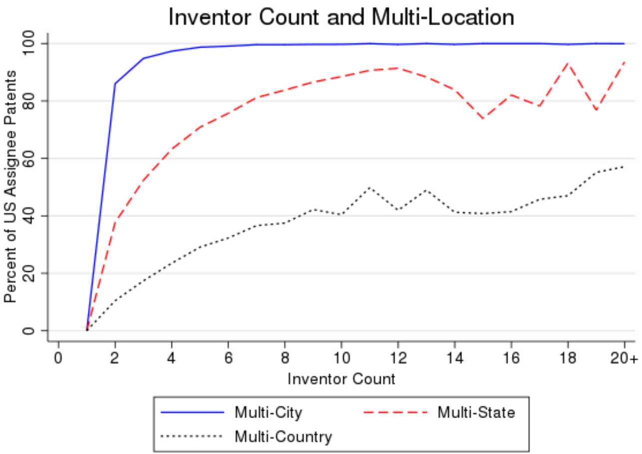
Back

NAICS 5413-5416 and 5112, 517, 518

- Professional, Scientific, and Technical Services
  - 5413: Architectural, Engineering, and Related Services
  - 5414: Specialized Design Services
  - 5415: Computer Systems Design and Related Services
  - 5416: Management, Scientific, and Technical Consulting Services
- Information
  - 5112: Software Publishers
  - 517: Telecommunications
  - 518: Data Processing, Hosting, and Related Services

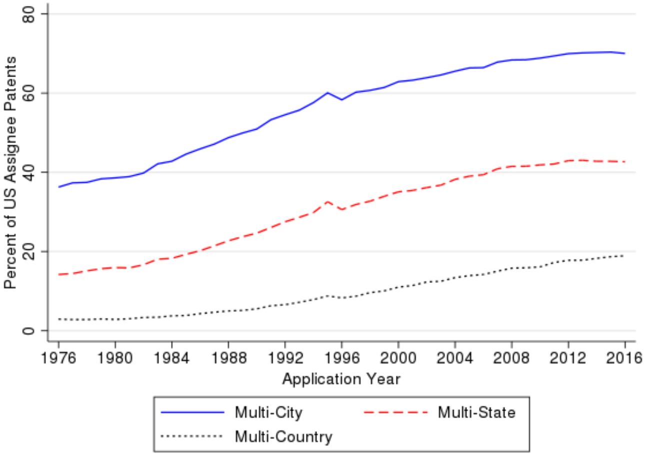
Back

Inventors tend to span cities and states



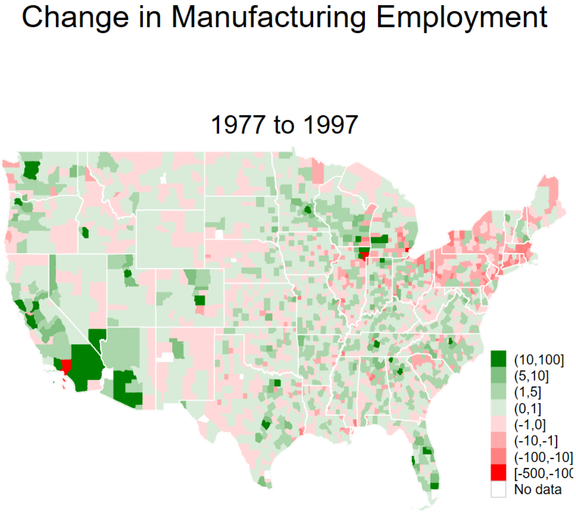
▸ Back

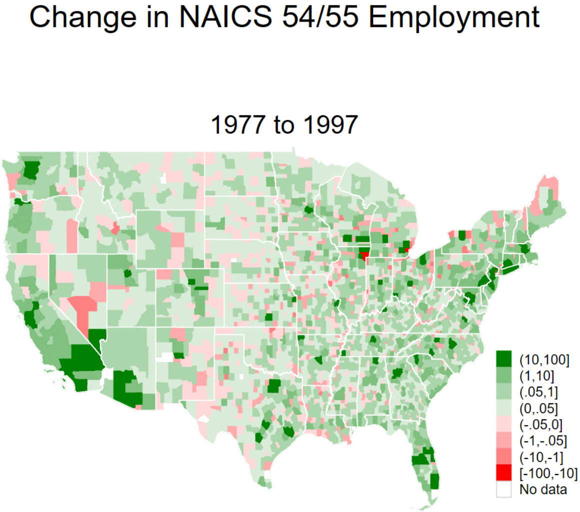
Inventor dispersion has grown over time



▸ Back

Spatial distribution of all US M and P





Differences between colocated and distant man plants

- Document characteristics of colocated man plants
  - Premia regressions on emp, sales, non-prod worker shares, number of products, age
  - Estab design good?
- Document characteristics of colocated P plants (in new project)

Analyze changes in firm colocation patterns

- What drives the changes in colocation?
- Are firms less likely to close the colocated plants?
- Are firms more likely to switch the industry of the colocated plants?

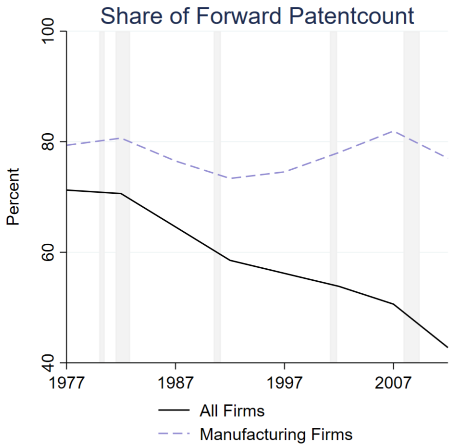
Interpreting firm-region-patent results

- If patenting does not occur in colocated plants, is CL still important?
- If inventors are near manufacturing estabs, is that colocation?
- If inventor teams are more disperse, does that negate colocation? ▸ Dispersion ▸ Over time
- If a growing share of inventors are overseas, is this bad for the US?
- Does the presence of some domestic manufacturing matter?

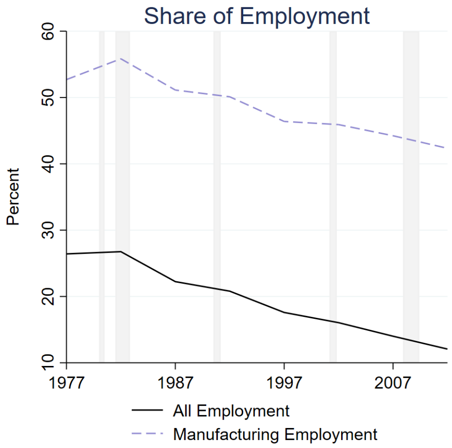
Conclusion: ordering the to do list

1. Spatial analysis of where innovation occurs within firms
2. Decomposition of the margins that drive colocation changes
3. Identification of the colocation changes (e.g., import competition vs offshoring)
4. Characteristics of the colocated and innovating plants
5. Importance of across firm colocation
6. Possibility of colocation within a M (or P) plant
7. Justification of P plants

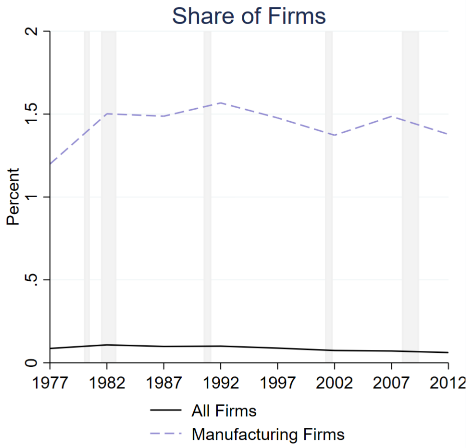
MP firms share of total forward patent counts



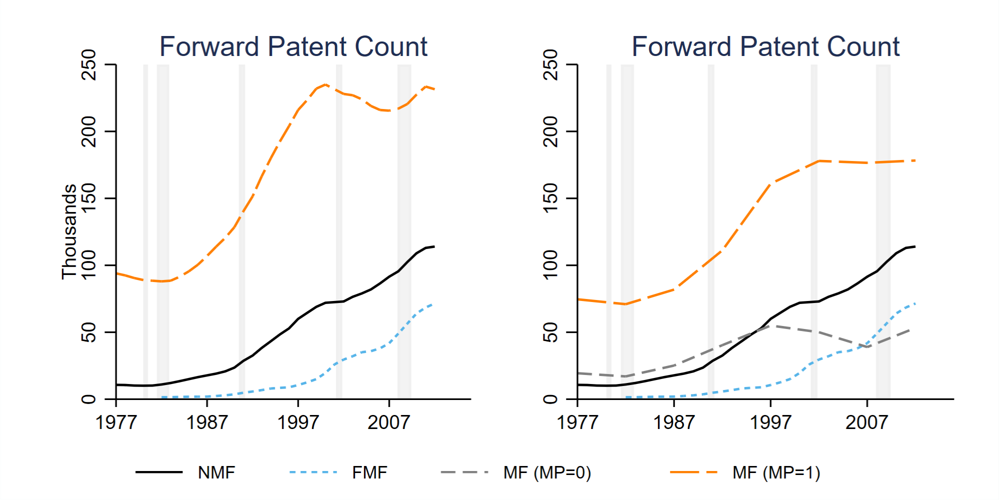
MP firms share of employment



MP firms share of firms

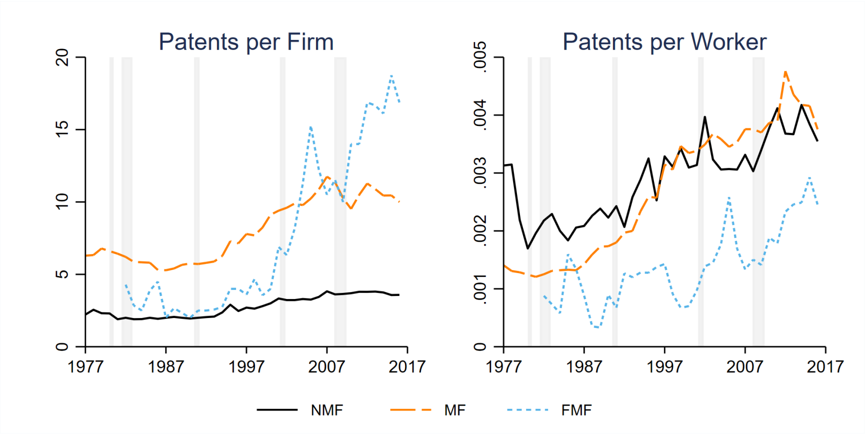


Forward patent count by firm type



Back

Patenting efficiency of patenting firms



- Patenting efficiency does not seem to decline
- Interesting to analyze by worker type