Global Innovation Spillovers and Productivity: Evidence from 100 Years of World Patent Data

Enrico Berkes [1] Kristina Manysheva [2] Martí Mestieri [3]

[1]University of Maryland – Baltimore County

[2]Columbia GSB

[3]IAE-CSIC, BSE, UPF and CREI

September 4, 2024

Study the Link between Innovation, Knowledge Spillovers, & Productivity

- Productivity is the key driver of modern economic growth
- Modern theories of economic growth: importance of innovation for productivity
- Empirical evidence on relationship between innovation, knowledge spillovers, and productivity is scarce

Study the Link between Innovation, Knowledge Spillovers, & Productivity

- Productivity is the key driver of modern economic growth
- Modern theories of economic growth: importance of innovation for productivity
- Empirical evidence on relationship between innovation, knowledge spillovers, and productivity is scarce
- Use the most comprehensive worldwide patent database (PATSTAT) to:
    1. Study the evolution of innovation over time and across countries
    2. Estimate importance of international knowledge spillovers in innovation activity
    3. Leverage knowledge spillovers to estimate contribution of innovation to productivity

Document Stylized Facts and Effect of Innovation on Productivity

- Document stylized facts on evolution of innovation and knowledge spillovers
  - Technological waves: Mechanical Engineering early 1900s to Computing in 1990's
  - Concentration of innovation across space: Heterogeneity in terms of innovation leaders
  - (International) spillovers: mid 90's increased role of US and Japan as innovation hubs
- Exploit this variation to:
  - Shift-share IV: shock propagation in network of knowledge across sectors-locations
  - Estimate elasticity of innovation to international knowledge spillovers of ≈ 0.5
  - Estimate causal effect of innovation on productivity
    - One st. dev. $\Uparrow$ in log patents $\Rightarrow$ an $\Uparrow$ in sectoral VA per worker growth of 1.1 p.p.
    - Similar results when we look at long term income growth instead

Literature

- **Innovation and Growth (emphasizing knowledge spillovers):**
  - *Coe, Helpman ('95); Coe et al. ('97); Keller ('04), ('09); Cai, Li ('18); Moretti et al. ('20); Huang, Zenou ('20)*
  - This paper: Proxy spillovers with citation network and use network to construct IV
- **Shift-share Instrument**
  - *Acemoglu et al. ('16); Hornbeck, Moretti ('20); Carvalho et al. ('20); Berkes, Gaetani ('22)*
  - This paper: Leverage rich (historical/worldwide) network structure to construct novel instrument
- **Use of Historical Patent Data**
  - *Bottazzi, Peri('03); Nicholas ('10); Petralia et al.('16); Akcigit et al.('17); Kelly et al.('20); Andrews, Whalley('21)*
  - This paper: Document global patterns of innovation, construct fields of knowledge

Outline

1. Data
2. Stylized Facts
3. Theoretical Framework
4. Innovation, Productivity and Output: Empirical Analysis
5. Innovation and Long-term Development

PATSTAT Data

- Use the European Patent Office Global PATSTAT database:
  - Over 110 mln patent records, comes from own patent offices
  - Detailed data for the US, UK, USSR, France, Germany and Switz. start in 1920
  - Majority of countries comprehensive data start from 1970
- Information by patent application (that we use):
  - Patent IPC classification (areas of technology)
  - Patent citations
  - Inventors information (name and residence)
  - Patent family
- Propose a novel approach to map patent classes into fields of knowledge
  - Clustering algorithm based on the name of the inventors ($\sim$ 50mln people) More

    E.g., algorithm groups together B60T: Vehicle brakes *and* F16D: Clutches; Brakes.

World Input Output Database

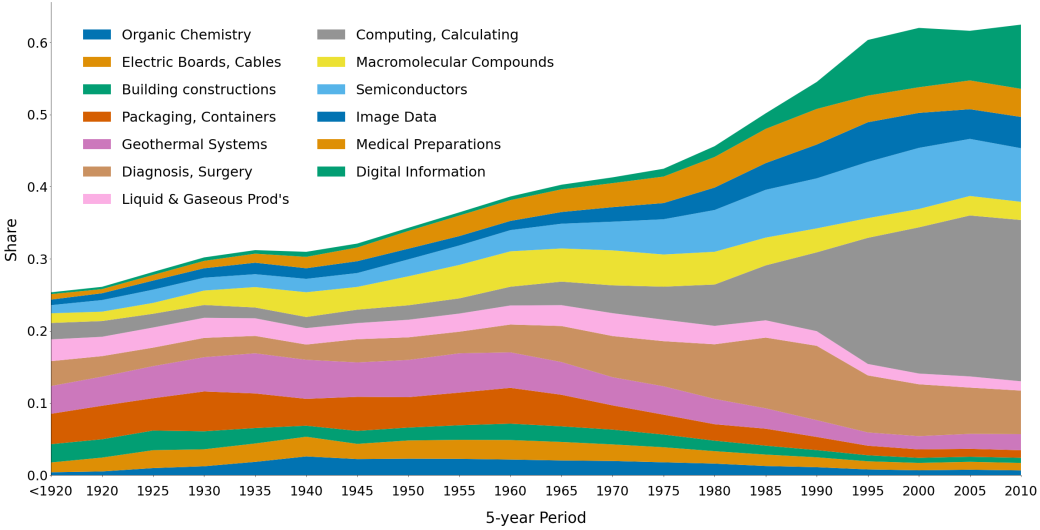
From the World Input Output Database, we collect data on:

- Value added per worker by sector (our main measure of productivity)
- Total Factor Productivity (TFP) growth by sector
- Capital, labor, and intermediate imports by sector (that we use as controls)

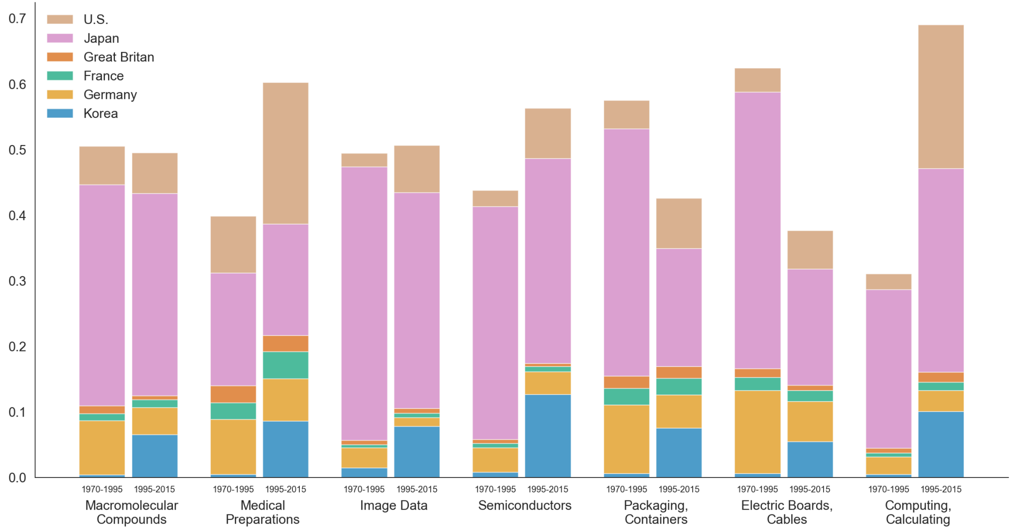
Outline

1. Data
2. Stylized Facts
3. Theoretical Framework
4. Innovation, Productivity and Output: Empirical Analysis
5. Innovation and Long-term Development

Evolution of Top Fields of Knowledge: Rise and Fall, $\Uparrow$ Concentration

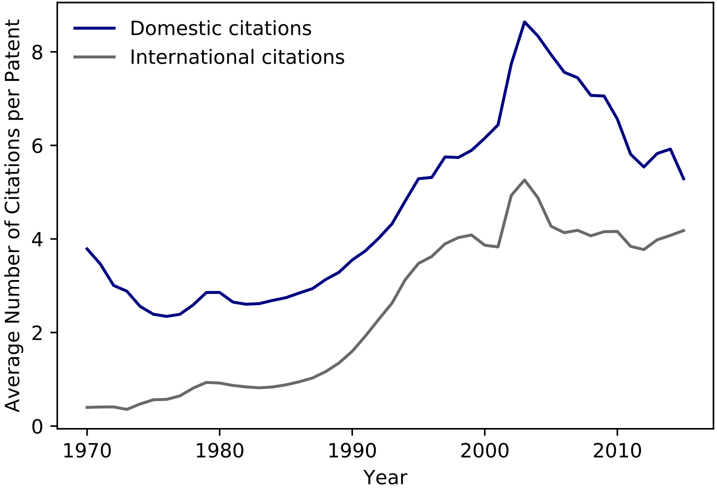


Country Shares in Top Fields, 1970-2015

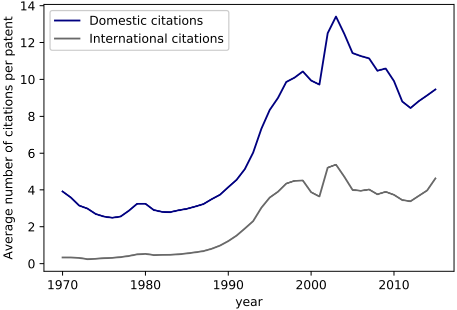


Intĺ Only 1920-70

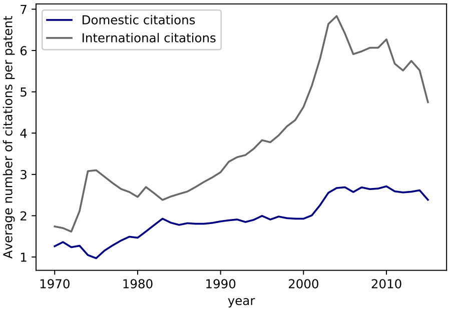
Knowledge Spillovers as Measured by Citations Increased Over Time



Reliance on Knowledge Produced in the US and Japan has Increased



(a) US and Japan



(b) Non-US and Non-Japan

By FoK

Outline

1. Data
2. Stylized Facts
3. Theoretical Framework
4. Innovation, Productivity and Output: Empirical Analysis
5. Innovation and Long-term Development

Organizing Framework: Innovation Technology

- World economy:
  - Multi-country, $c = 1, \ldots, C$
  - Multi-sector, $s = 1, \ldots, S$
  - Fields of Knowledge, $k = 1, \ldots, K$
- World technology: $\mathrm{N}_t \equiv (N_{111t}, \ldots, N_{cskt}, \ldots, N_{CSKt})$
- Production function for Innovation $I(\cdot)$:

$$\Delta N_{cskt} = I\left(S_{csk}(\mathrm{N}_t), R_{cskt}\right),$$

where $S_{csk}(\mathrm{N_t})$: spillover across csk, R_{cskt}: research inputs

Organizing Framework: Innovation Technology

- World economy:
  - Multi-country, $c = 1, \ldots, C$
  - Multi-sector, $s = 1, \ldots, S$
  - Fields of Knowledge, $k = 1, \ldots, K$
- World technology: $\mathrm{N}_t \equiv (N_{111t}, \ldots, N_{cskt}, \ldots, N_{CSKt})$
- Production function for Innovation $I(\cdot)$:

$$\Delta N_{cskt} = I\left(S_{csk}(\mathrm{N}_t), R_{cskt}\right),$$

where $S_{csk}(\mathrm{N_t})$: spillover across csk, R_{cskt}: research inputs

Mappings

- $\Delta N_{cst} \equiv \sum_{k=1}^{K} \Delta N_{cskt} \rightarrow$ Patenting across cst (link FoK to industry codes)
- Linkages across c, s and k embedded in $S \rightarrow$ citations across $cskt$

$$S_{sck}(\mathrm{N_t}) = \sum_{c' \in C} \sum_{s' \in S} \sum_{k' \in K} \alpha_{c's'k't} N_{c's'k't'}$$

Organizing Framework: Innovation, TFP and Output

- Output per worker y_{cst} given by Cobb-Douglas:

$$\log y_{cst} = \phi_{cst} + \log TFP_{cst} + \alpha \log k_{cst}, \qquad \text{where} \quad \phi_{cst} = \tilde{\delta}_{ct} + \tilde{\delta}_{st} + \tilde{\delta}_{cs}.$$

- Isoelastic relationship between TFP and Innovation

$$\log\left(\frac{TFP_{cst+1}}{TFP_{cst}}\right) = \phi_0 + \phi_N \log(1 + \Delta N_{cst}) - \phi_Y \log y_{cst},$$

- Combining both we obtain baseline regression:

$$\log y_{cst+1} \quad = \quad \phi_N \log(1 + \Delta N_{cst}) + \phi_A \log y_{cst} + \delta_{ct} + \delta_{st},$$

where $\phi_A, \phi_N > 0$, and δ's denote Fixed Effects

- Important: no assumption of BGP, free sectoral dynamics

Outline

Empirical Specification Derived from Growth Model

Obtain empirical specification from (semi-)endogenous growth model:

$$\overline{\log y_{cst+n}} = \phi_N \log\left(1 + pat_{cst}\right) + \phi_A \log y_{cst} + \phi_0 X_{cst} + \delta_{ct} + \delta_{st} + \epsilon_{cst},$$

where

- $\overline{\log y_{cst+n}}$ average log value added per worker over next three years
- X_{cst} set of controls: capital, labor, and intermediate inputs trade linkages
- δ_{ct} & δ_{st} country-year and sector-year fixed effects

Innovation and Output per Worker

	Dependent variable is: $\overline{\log(va\_em_{cst,t+3})}$							
	OLS (1)	2SLS (2)	OLS (3)	2SLS (4)	OLS (5)	2SLS (6)	OLS (7)	2SLS (8)
$\log(1 + pat_{cst})$	0.006		0.004		0.005		0.005	
	(0.003)		(0.003)		(0.003)		(0.003)	
$\log(va\_em_{cst})$	0.919		0.942		0.934		0.933	
	(0.012)		(0.016)		(0.016)		(0.016)	
Country-Yr & Sector-Yr FE	✓		✓		✓		✓	
$\ln(capital_t)$			✓		✓		✓	
$\ln(employ_t)$			✓		✓		✓	
$\ln(int\_import_t)$					✓		✓	
$\log(\widehat{go}\_IOlink_{cst})$							✓	
# obs.	8,357		8,357		8,357		8,357	
# countries	36		36		36		36	

Notes: Period of analysis: 2000-14. Two-way st. err. clustered at a country-sector level.

Identification Concerns

- Reverse causality
  - Change in productivity affects innovative activities
- Measurement error
  - Attenuation bias
- Factors that might affect productivity and innovation at the same time, such as:
  - Technological obsolescence
  - Financial shocks
  - Trends in domestic demand

Instrumenting for Innovative Activities

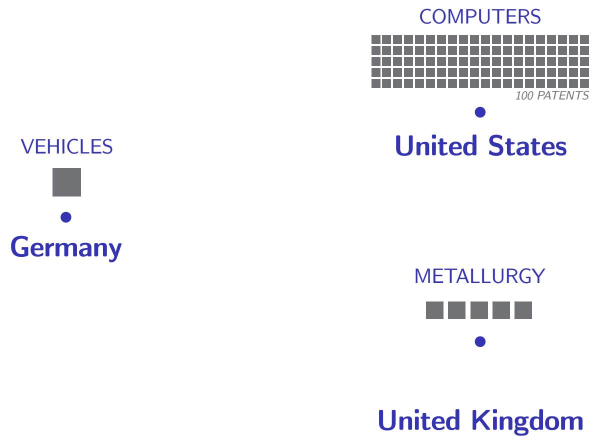
- Use the network of patent citations in the pre-sample period (1970-1990)
- Idea: Each patent is the result of assembling already existing pieces of knowledge (input-output model)
- Observed citations network reveals the knowledge input-output linkages across countries and technological classes

Shift-share Construction Proceeds in **Three** Steps

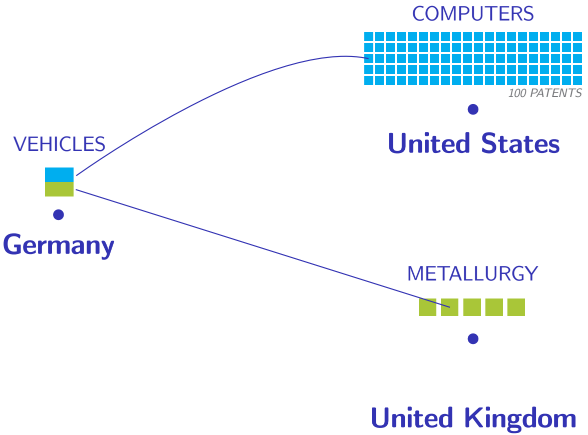
1. **Shares**: Compute citations (c, c', s, s', lag) in "predetermined" network in 1970-90
2. Propagate country-sector innovations for the period 1990-99 (**shifts**) through this network, assuming *probability of innovation is proportional to strength of citations*
3. Iterate forward: using predicted innovations for each country-sector from step 2, propagate again across network to construct **instrument** for 2000-14

Instrumenting for patenting activity: Pre-sample

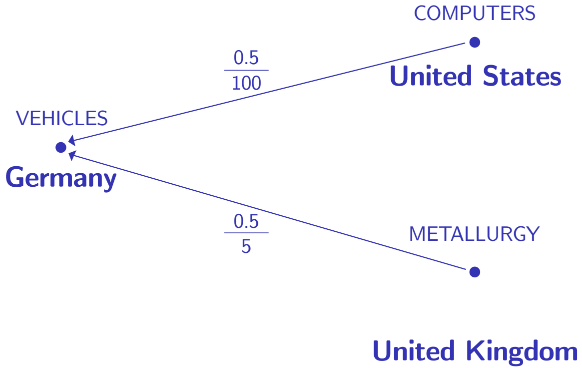
Instrumenting for patenting activity: Pre-sample



Instrumenting for patenting activity: Pre-sample



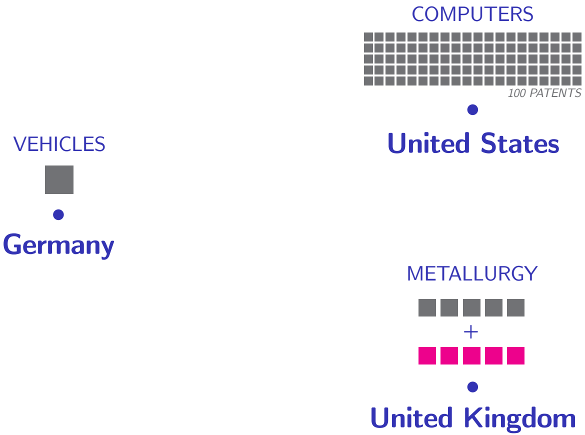
Instrumenting for patenting activity: Spillovers network



Shift-share Construction Proceeds in **Three** Steps

1. **Shares**: Compute citations (c, c', s, s', lag) in "predetermined" network in 1970-90
2. Propagate country-sector innovations for the period 1990-99 (**shifts**) through this network, assuming *probability of innovation is proportional to strength of citations*
3. Iterate forward: using predicted innovations for each country-sector from step 2, propagate again across network to construct **instrument** for 2000-14

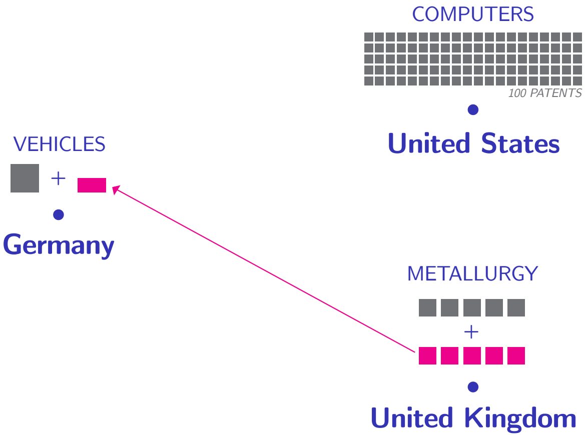
Instrumenting for patenting activity: Sample period



Shift-share Construction Proceeds in **Three** Steps

1. **Shares**: Compute citations (c, c', s, s', lag) in "predetermined" network in 1970-90
2. Propagate country-sector innovations for the period 1990-99 (**shifts**) through this network, assuming *probability of innovation is proportional to strength of citations*
3. Iterate forward: using predicted innovations for each country-sector from step 2, propagate again across network to construct **instrument** for 2000-14

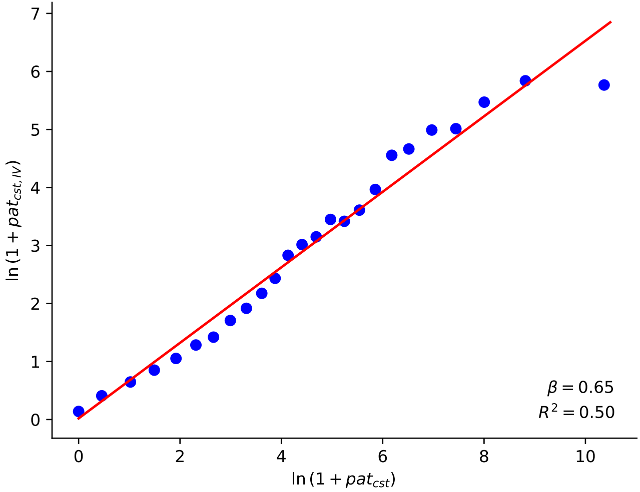
Instrumenting for patenting activity: Instrument



Instrumenting for innovative activities

- Note that the network:
  - is **directed**, $(s, r) \neq (r, s)$
  - takes the **timing of citations** into consideration: Δ_T, with $T = 1, ..., 10$
  - does not consider citations coming from the same field of knowledge and country

Unconditional Correlation between Actual and Predicted Patents (IV)



Elasticity of Innovation to International Knowledge Spillovers

	$\log(1 + pat_{cst})$		
	(1)	(2)	(3)
$\log(1 + \widehat{pat}_{cst})$	0.452	0.470	0.599
	(0.047)	(0.048)	(0.061)
Country-Year FE	✓	✓	✓
Sector-Year FE	✓	✓	✓
# obs.	31,292	31,292	31,292
# countries	198	198	198

Notes: Period of analysis: 2000-14 with citation matrix of 1970-90. Two-way st. err. clustered at country-sector level.

Main Results: Innovation and Value Added per Worker

	$\overline{\log(va\_em_{cst+n})}$ $n \in \{1,2,3\}$							
	OLS (1)	2SLS (2)	OLS (3)	2SLS (4)	OLS (5)	2SLS (6)	OLS (7)	2SLS (8)
$\log(1+pat_{cst})$	0.006	0.019	0.004	0.017	0.005	0.019	0.005	0.018
	(0.003)	(0.008)	(0.003)	(0.008)	(0.003)	(0.007)	(0.003)	(0.007)
$\log(va\_em_{cst})$	0.919	0.917	0.942	0.937	0.934	0.928	0.933	0.927
	(0.012)	(0.012)	(0.016)	(0.016)	(0.016)	(0.015)	(0.016)	(0.015)
Country-Yr & Sector-Yr FE	✓	✓	✓	✓	✓	✓	✓	✓
$\ln(capital_t)$			✓	✓	✓	✓	✓	✓
$\ln(employ_t)$			✓	✓	✓	✓	✓	✓
$\ln(int\_import_t)$					✓	✓	✓	✓
$\log(\widehat{go}\_IOlink_{cst})$							✓	✓
# obs.	8,357	8,357	8,357	8,357	8,357	8,357	8,357	8,357
# countries	36	36	36	36	36	36	36	36
	First-stage estimates							
Predicted $\log(1+pat_{cst})$		0.496		0.461		0.461		0.461
		(0.082)		(0.079)		(0.079)		(0.079)
F-stat		36.7		33.9		34.3		33.9

Notes: Period of analysis: 2000-14 with citation matrix of 1970-90. Two-way st. err. clustered at country-sector level.

Robustness

- Test robustness of specification by:
  1. Alternative log transformation and forward lags Results
  2. Specification in growth rates
  3. Use TFP instead of value added per worker Results: TFP & Growth
- Test robustness of IV procedure by:
  1. Test for pre-trends by showing that the pre-period productivity is uncorrelated with subsequent patent activity predicted by the instrument Results
  2. Control for historical levels of patent activity
  3. Testing for demand-side pull factors
     - Reverse matrix of citations $\Rightarrow$ predicted number of patents in pre-sample period
     - Add predicted by downstream stimulus number of patents in the baseline regression
  4. Add predicted patents using international IO linkages All results

Outline

Innovation and Long-term Development: GDP pc over 1980-2016

	Dep. Var.: $\overline{\log(gdp\_cap)_{ct+n}}$ from 80-16			
	OLS	IV	OLS	IV
$\ln(1 + pat_{ct})$	0.013	0.086	0.005	0.034
	(0.005)	(0.020)	(0.003)	(0.012)
$\ln(gdp\_cap_{ct})$	0.906	0.735	0.852	0.804
	(0.026)	(0.052)	(0.025)	(0.028)
Country FE	Y	Y	Y	Y
Year FE	N	N	Y	Y
# obs.	1,985	1,985	1,985	1,985
# countries	60	60	60	60
F-stat		15.0		7.3

Notes: Pre-determined matrix based on the data prior 1980. St. err. clustered at a country level.

- Magnitude: 1 st. dev. $\Uparrow$ in log pats $\Rightarrow$ $\Uparrow$ in log output (growth) of 2.8 p.p.

Innovation and Long-term Development: GDP pc over 1960/70-2016

	Dependent Variable is: $\overline{\log(gdp\_cap)_{ct+n}}$			
	1960-2016		1970-2016	
	OLS	2SLS	OLS	2SLS
$\ln(1 + pat_{ct})$	0.005	0.015	0.009	0.023
	(0.003)	(0.005)	(0.004)	(0.008)
$\ln(gdp\_cap_{ct})$	0.900	0.879	0.872	0.837
	(0.027)	(0.026)	(0.004)	(0.040)
Country FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
# obs.	2,760	2,760	2,376	2,376
# countries	60	60	60	60
F-stat		27.4		20.7

Std. err. clustered at a country level. Predetermined matrix based on pre-1950 and pre-1960, respectively.

- 1 st. dev. $\Uparrow$ in log pats $\Rightarrow$ $\Uparrow$ in log output (growth) of 1.6/2.2 p.p.

Conclusions

- Use historical patent data spanning last 100 years to:
  - Map patent classes into fields of knowledge using a clustering algorithm
  - Document stylized facts on evolution of innovation and knowledge spillovers
- Use variation in our data to estimate the effect of innovation on productivity
  - Construct a shift-share instrument leveraging country-sector knowledge spillovers
  - One std. dev. increase in log-patenting increases productivity growth by 1.1 p.p.
- Going forward: Empirical design useful for other applications
  - E.g., trade elasticity in EK-CDK.
- Use these facts and data for a quantitative model

Novel Approach to Map Patent Classes into "Fields of Knowledge"

- Typically, take X first digits of patent class (e.g. Jaffee et al. 93)
  - Justification: Hierarchical system reflecting area of technology
  - Criticized by some due to presence of some arbitrariness
    - B60T: Vehicle brakes or part of thereof
    - F16D: Clutches; Brakes
- Propose an algorithm that groups patent classes into "fields of knowledge"
  - Use information on inventors (almost 50 mln people)
  - Compute conditional probability the same inventor invents in any IPC given that he invents in another IPC for each IPC pair
  - Construct symmetric proximity matrix Details
  - Cluster IPCs according to this information K-medoids Optimality criterion
- Result: 164 fields of knowledge

Back

The inverse measure of similarity between IPC codes i and j equals

$$\phi_{ij} = P(x_{ij} | x_{jj})$$

where for every inventor s

$$x_{ij,s} = \begin{cases} 1, \text{ if inventor has patent with both IPC codes } i \text{ and } j \\ 0, \text{ otherwise} \end{cases}$$

In order to obtain symmetric matrix for the cluster analysis the following dissimilarity measure is obtained:

$$D_{ij} = 1 - (\phi_{ij} + \phi_{ji}) = D_{ji}$$

back

The following steps provide brief description of the *k-medoids* algorithm:

- Select K points as the initial representative objects (i.e., as initial *k-medoids*)
- Repeat
  - Assigning each point to the cluster with the closest medoid
  - Randomly select a non-representative object o_i
  - Compute the total cost S of swapping the medoid m with o_i
    - $S = S_{o_i} - S_m$
    - $S_{o_i} = \sum_{j=1}^{k} s_{j,o_i}$, where $s_{j,o_i} = \sum_{l \in j} D_{l,o_i}$
    - $S_m = \sum_{j=1}^{k} s_{j,m}$, where $s_{j,m} = \sum_{l \in j} D_{l,m}$
    - In words, S_x is the sum of distances to the center within each cluster
  - if $S < 0$, then swap m with o_i to form the new set of medoids
- Until convergence criterion is satisfied

back

Silhouette coefficient:

- For each point x_i, its silhouette coefficient s_i is:

$$s_i = \frac{\mu_{out}^{min}(x_i) - \mu_{in}(x_i)}{\max\{\mu_{out}^{min}(x_i), \mu_{in}(x_i)\}}$$

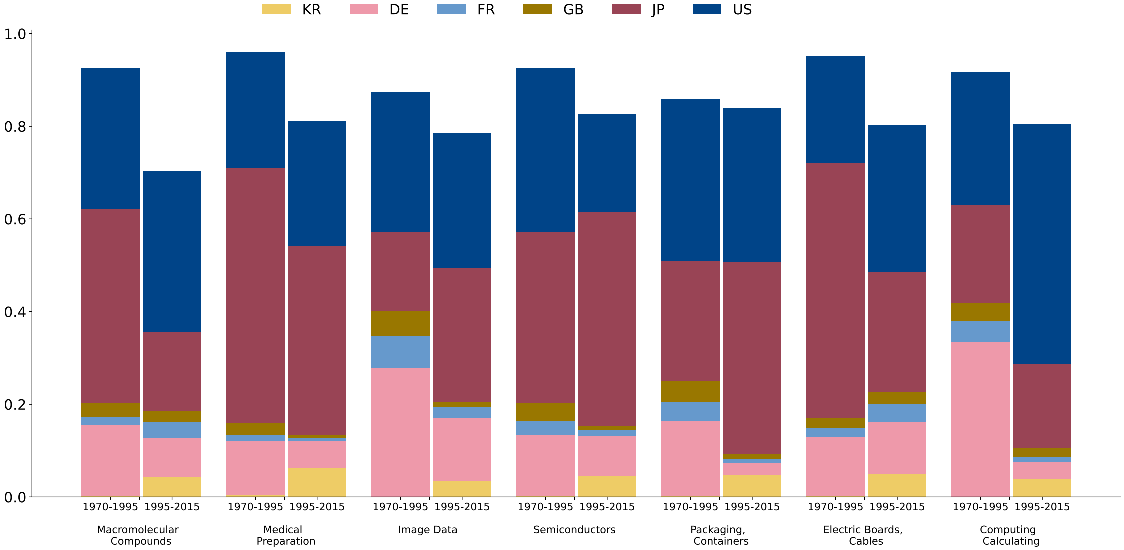
  where
  - $\mu_{in}(x_i)$ is the mean distance from x_i to points in its own cluster
  - $\mu_{out}^{min}(x_i)$ is the mean distance from x_i to points in its closest cluster (defined by finding minimum distance from the center of cluster to the center of other cluster)
- $s_i = 0$ if x_i belongs to singleton cluster
- Silhouette coefficient (SC) is the mean values of s_i across all the points: $SC = \frac{1}{n}\sum_{i=1}^{n} s_i$
- SC close to $+1$ implies good clustering: points are close to their own clusters but far from other clusters

back

Center of the Cluster	Members of the Cluster	Respective IPC Name	Suggested Name of Technology Field
A21B	A21B	Bakers' Ovens; Machines or Equipment for Baking	Baking
	A21C	Machines or Equipment for Making or Processing Doughs; Handling Baked Articles Made from Dough	
	A47J	Kitchen Equipment; Coffee Mills; Spice Mills; Apparatus for Making Beverages	
	F24C	Other Domestic Stoves or Ranges; Details of Domestic Stoves or Ranges	

back

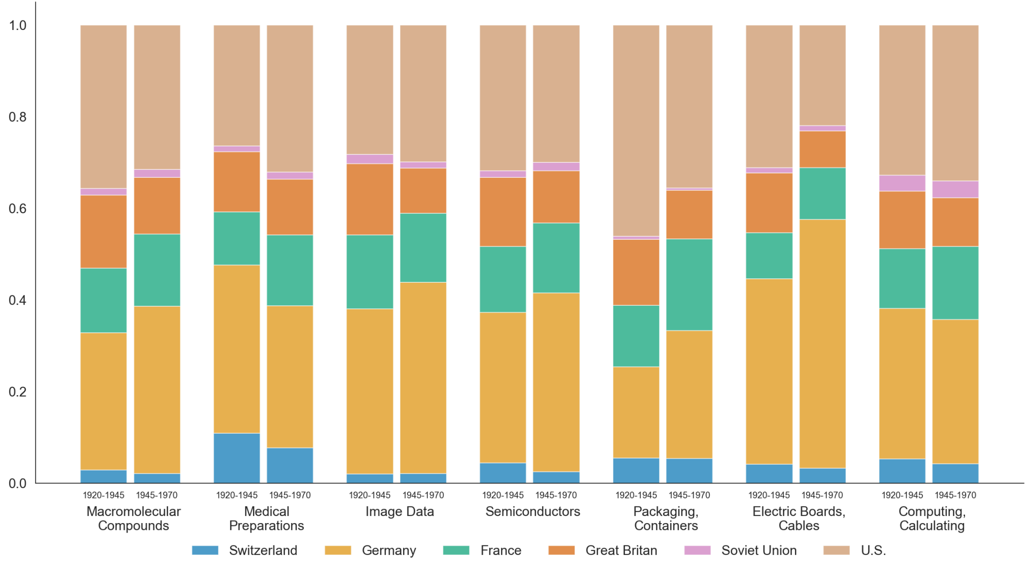
Country Shares in Top Fields, 1970-2015: International Citations



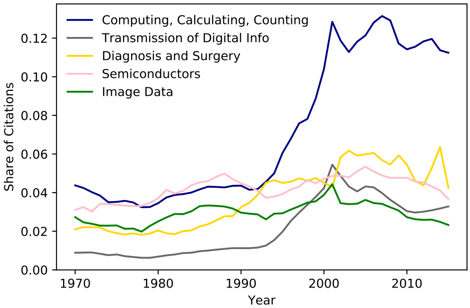
Back

Country Shares in Top Fields, 1920-1970: Advanced Economies

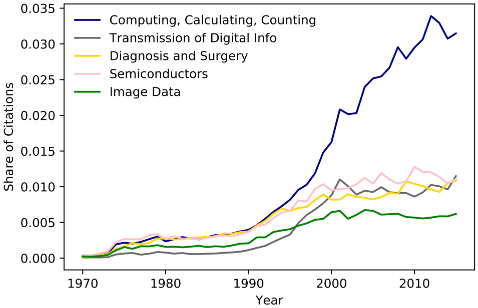
back



Share of Citations to the US and Japan Patents by Field of Knowledge



(c) US and Japan



(d) Non-US and Non-Japan

back

Instrument Construction

The adjency matrix of the knowledge network is (presample):

$$d_{c_o,c_d,s_o,s_d,\Delta} = \begin{cases} 0 & c_o = c_d \\ 0 & s_o = s_d \\ \dfrac{\sum_{t=1970}^{1990} \sum_{p \in \mathcal{P}(c_o,s_o,t)} s_{p \to (c_d,s_d,t-\Delta)}}{\sum_{t=1970}^{1970} \left| \mathcal{P}\left(c_d, s_d, t-\Delta\right) \right|} & o/w \end{cases} \qquad \Delta \in \{1,.,10\}$$

diffuse the observed patents filed in the period 1990-1999 to predict the patenting activity in the sample:

$$\hat{pat}_{c_o,s_o,t} = a_t \sum_{\Delta=1}^{10} \sum_{s_d \in \mathcal{S}} \sum_{c_d \in \mathcal{N}} \left(d_{c_o,c_d,s_o,s_d,\Delta}\right) pat_{c_d,s_d,t-\Delta}$$

2SLS Estimates: Robustness, Different Lags

	$\log(va\_em_{cst+n})$				
	(1)	(2)	(3)	(4)	(5)
$\log(1+pat_{cst})$	0.017	0.019	0.011	0.018	0.026
	(0.008)	(0.008)	(0.004)	(0.008)	(0.010)
Controls	✓	✓	✓	✓	✓
Country-Year FE	Y	Y	Y	Y	Y
Sector-Year FE	Y	Y	Y	Y	Y
# obs.	8,357	8,357	9,744	9,053	8,358
	First-stage estimates				
Predicted	0.461	0.392	0.457	0.460	0.463
$\log(1+pat_{cst})$	(0.079)	(0.070)	(0.077)	(0.079)	(0.079)
F-stat	33.9	31.1	34.7	34.3	34.0

Notes: Column (1) shows the results of our baseline regression, with average level of (log) value added per employment in the next 3 years as a dependent variable. Column (2) is analogous to Column (1) in terms of dependent variable, but uses inverse hyperbolic sine for the log transformation applied to a number of patents used both as an explanatory variable and as an instrument, i.e. $\ln(\sqrt{1+pat^2}+pat)$. Columns (3), (4), and (5) use one, two, and three periods ahead value added per employment as dependent variable. All regressions include (log) values for value added per employment, capital, and employment as controls. Standard errors (in parentheses) are two-way clustered at the country and sector levels. Kleibergen-Paap Wald F-stat is reported for the first stage. back

Growth specification and TFP Regressions

	$\overline{\Delta \log(y_{cst+n})}$ $n \in \{1,2,3\}$								
	VA/EMP			Primal TFP			Dual TFP		
	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\log(1 + patent_{cst})$	0.011	0.009	0.009	0.007	0.010	0.010	0.004	0.008	0.008
	(0.004)	(0.004)	(0.003)	(0.004)	(0.006)	(0.005)	(0.004)	(0.003)	(0.003)
$\log(y_{cst})$	-0.044	-0.031	-0.033	-0.017	-0.010	-0.011	-0.018	-0.009	-0.009
	(0.009)	(0.007)	(0.007)	(0.010)	(0.009)	(0.008)	(0.009)	(0.005)	(0.005)
$\log(capital_{cst})$		-0.005	-0.005		-0.023	-0.022		-0.031	-0.031
		(0.003)	(0.003)		(0.003)	(0.003)		(0.003)	(0.004)
$\log(employ_{cst})$		0.005	0.002		0.021	0.022		0.026	0.023
		(0.004)	(0.005)		(0.004)	(0.004)		(0.004)	(0.003)
$\log(int\_import_{cst})$			0.003			-0.002			0.001
			(0.004)			(0.005)			(0.003)
Country-Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Sector-Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
# obs.	8,834	8,357	8,357	7,931	7,931	7,931	8,554	8,336	8,336
# countries	36	36	36	36	36	36	36	36	36
	First-stage estimates								
Predicted	0.468	0.461	0.461	0.498	0.470	0.472	0.498	0.472	0.473
$\log(1 + pat_t)$	(0.085)	(0.079)	(0.079)	(0.081)	(0.080)	(0.080)	(0.085)	(0.083)	(0.083)
F-stat	30.5	33.9	34.3	34.5	32.5	32.4	38.1	35.0	34.9

Period of the analysis 2000-14, pre-determined matrix based on the data from 1970-90. St. err. clustered at a country level. back

Testing for Pre-trends

	$\log(\overline{va\_emp_{cs}})$			
	Sample Period		Pre-Sample Period	
	(1)	(2)	(3)	(4)
$\log(\overline{1 + pat_{cs2000-14}})$	0.080	0.102	0.032	0.014
	(0.033)	(0.046)	(0.064)	(0.053)
Controls	✓	✓	✓	✓
Country FE	Y	Y	Y	Y
Sector FE	Y	Y	Y	Y
# obs.	641	433	433	424

Notes: Columns (1) and (2) use average value added per employment in the period 2000-14 as a dependent variable computed with WIOD and UNIDO data, respectively. The latter one is included for better compatibility with results in columns (3) and (4), where dependent variable is average value added per employment computed with UNIDO data for the periods 1981-90 and 1971-90, respectively. All regressions include average (log) values for capital, employment and intermediate imports in period 2000-14. Standard errors are clustered at a country level in parentheses. back

Robustness Exercises

	$\overline{\log(va\_em_{cst+n})}$ $n \in \{1,2,3\}$				
	(1)	(2)	(3)	(4)	(5)
$\log(1+pat_{cst})$	0.017	0.029	0.025	0.030	0.017
	(0.008)	(0.010)	(0.010)	(0.010)	(0.008)
$\log(1+\overline{pat_{cs1970-90}})$		-0.009		-0.009	
		(0.005)		(0.005)	
$\log(1+\widehat{pat}_{cst-30})$			-0.006	-0.001	
			(0.007)	(0.006)	
$\log(1+\widehat{pat}\_IOlink_{cst})$					-0.001
					(0.006)
Controls	✓	✓	✓	✓	
Country-Year FE	Y	Y	Y	Y	Y
Sector-Year FE	Y	Y	Y	Y	Y
# obs.	8,357	8,357	8,357	8,357	8,357
	First-stage estimates				
Predicted	0.461	0.264	0.388	0.305	0.459
$\log(1+pat_{cst})$	(0.079)	(0.058)	(0.065)	(0.056)	(0.078)
F-stat	33.9	20.9	35.5	29.3	34.8

back