Inventor Location and the Globalization of R&D

Dietmar Harhoff[a] and Grid Thoma[b]
[a]Ludwig-Maximilians-Universität (LMU) Munich and CEPR
[b]University of Camerino

Prepared for the Conference on
Advancing the Study of Innovation and Globalization in Organizations (ASIGO)
Nürnberg, Germany, May 29-30, 2009

Agenda

- The Research Question
- Inventor Counts – The Basic Idea
- Data
- Descriptive Results
- R&D Accounting
- Caveats, Conclusions & Extensions

The Research Question

- There is very little systematic data on the overall extent of R&D globalization.
- R&D data as collected by NSF or statistical agencies like Stifterverband are (mainly) territorial in nature.
- Other data sources have information on total R&D, but not R&D by country.

- But there are long time series on inventor location in patent documents.
- **Can we use these data to get reliable information on globalization processes?**
- Once we can, lots of potential applications to R&D productivity, spillovers, etc.
- Current contribution – a.k.o. feasibility test.

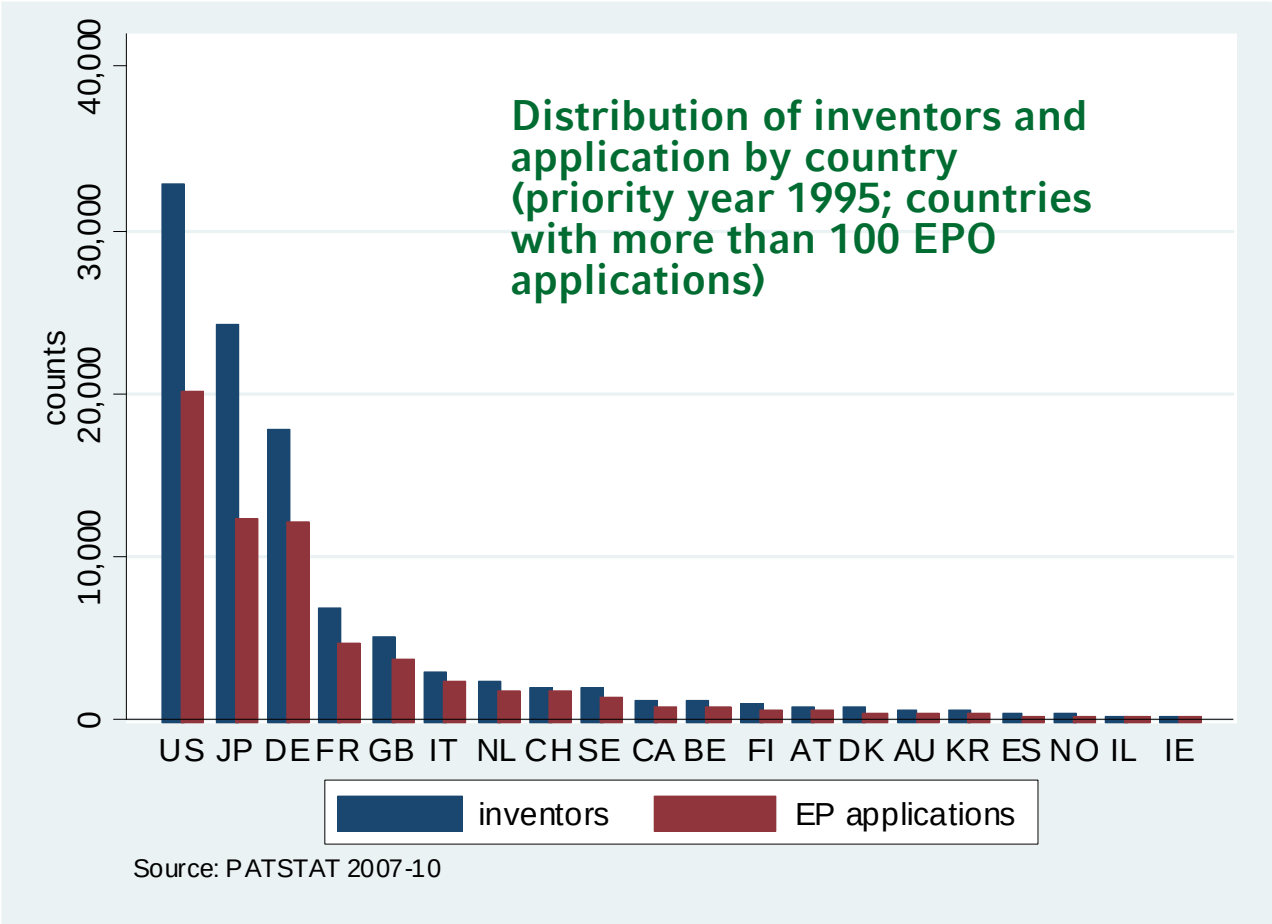
Inventor Counts – The Basic Idea

- identify all patents filed by firm X with headquarter in country Y
- consolidate all inventor names (remove divergent spellings etc.) on these patents
- particular attention to the top ≈2000 R&D performers in Europe – consolidate ownership structures for these firms
- analyze the distribution of inventors by country for different time periods, technical fields, etc.
- compare R&D to inventor counts (in levels and growth rates)
- develop accounting framework and regress R&D expenditures on inventor counts and location information
- derive predictions, e.g. a matrix of R&D linkages between headquarter country and country of execution

Data

- all EPO filings (publications under an EP publication number)
  - source: PATSTAT 2007-04 (currently in work: 2009-04)
  - develop annual and moving average inventor counts
  - record country of inventor location
  - for details see TTGGHH (2008)
- firm level data (incl. R&D) for 2,157 European firms
  - balance sheet and P&L data
  - "accounting for" a large fraction of European (territorial) R&D
  - ownership consolidated over 150,000 subsidiaries

Descriptive Results (1/3)



Descriptive Results (2/3)

Table 3 - Distribution of Inventors by Applicant and Inventor Country

Top EU R&D Performers - Row Percent

Applicant country	**1986-1990**												
	CA	*CH*	*DE*	*ES*	*FR*	*GB*	*IT*	*JP*	*KR*	*NL*	*ROW*	*SE*	*US*
CH	0.4	31.8	28.8	0.1	4.0	4.5	1.3	1.7	0.0	0.8	3.9	1.4	21.3
DE	0.2	0.8	72.0	0.2	4.4	2.3	1.1	2.6	0.0	0.9	4.0	0.5	11.1
FR	1.2	0.7	19.3	0.4	48.5	5.5	2.1	2.8	0.0	0.8	4.9	0.4	13.4
GB	0.6	0.4	12.2	0.1	6.4	44.1	0.8	4.6	0.0	4.7	2.7	0.8	22.5
IT	0.1	0.6	10.7	0.1	5.7	12.2	58.6	0.1	.	0.2	4.9	0.1	6.7
NL	0.2	0.9	42.2	0.1	6.2	7.3	2.4	3.8	0.0	25.2	2.8	1.0	7.9
other EU	0.2	1.1	9.9	0.4	9.1	5.6	2.5	1.8	.	3.8	51.4	5.0	9.3
SE	0.2	7.7	16.4	0.1	4.9	5.5	3.4	0.6	.	1.7	7.6	43.0	9.1
Applicant country	**2001-2006**												
	CA	*CH*	*DE*	*ES*	*FR*	*GB*	*IT*	*JP*	*KR*	*NL*	*ROW*	*SE*	*US*
CH	0.6	27.3	21.1	0.3	5.6	3.9	2.9	1.7	0.1	0.9	8.2	1.1	26.2
DE	0.6	1.1	69.2	0.9	3.3	1.5	1.6	2.6	0.2	0.9	5.0	0.4	12.8
FR	2.0	0.8	14.0	0.9	48.7	2.6	1.6	4.6	0.6	0.9	4.9	0.5	17.9
GB	1.4	2.7	11.3	0.5	11.3	27.4	1.7	5.1	0.0	3.6	6.8	0.7	27.5
IT	0.3	0.5	9.7	1.1	8.0	3.3	58.3	0.5	.	0.7	7.4	0.9	9.1
NL	0.1	1.2	35.2	0.6	8.9	5.8	5.4	1.6	0.3	24.1	6.0	0.8	9.9
other EU	0.7	0.9	9.1	0.5	5.2	4.5	3.0	1.4	0.2	2.2	55.1	3.1	14.2
SE	1.2	4.0	17.9	0.8	3.3	6.6	3.7	1.5	0.1	2.0	9.3	38.9	10.9

Descriptive Results (3/3)

Correlation between Inventor Counts and R&D in Levels and Growth Rates

		log of annual counts			growth rates		
Year	N	annual	3 yrs mav	5 yrs mav	annual	3 yrs mav	5 yrs mav
1991	159	0.614	0.652	0.660	.	.	.
1992	174	0.692	0.702	0.699	0.607	0.691	0.693
1993	188	0.679	0.699	0.705	0.583	0.743	0.737
1994	217	0.722	0.745	0.757	0.551	0.632	0.684
1995	250	0.709	0.730	0.733	0.492	0.632	0.663
1996	303	0.716	0.733	0.728	0.343	0.526	0.513
1997	377	0.680	0.703	0.716	0.372	0.453	0.480
1998	426	0.690	0.718	0.727	0.319	0.432	0.453
1999	470	0.705	0.711	0.726	0.405	0.527	0.532
2000	485	0.677	0.708	0.718	0.280	0.448	0.489
2001	489	0.691	0.717	0.722	0.269	0.403	0.408
2002	540	0.703	0.715	0.723	0.396	0.542	0.584
2003	570	0.699	0.735	0.741	0.318	0.484	0.477
2004	589	0.674	0.686	0.701	0.311	0.488	0.518
2005	277	0.570	0.679	0.685	0.191	.	.
Total	**5,514**	**0.681**	**0.710**	**0.719**	**0.383**	**0.533**	**0.555**

R&D Accounting

R&D is taken to be the wage bill plus materials and capital expenditures, the latter proportional to R&D labor.

$$R_{it} = \left(\sum_{k=1}^{K} W_{ikt} I_{ikt} \right) \cdot m_{it} \cdot c_{it}$$

Additional assumptions lead to a NLLS estimation framework:.

$$\log R_{it} = \alpha + \log\left(\sum_{k=1}^{K} \beta_k I_{ikt} \right) + \sum_{j=1}^{J} \delta_j D_j + \sum_{t=1}^{T} \gamma_t E_t + \gamma \log P_{jt} + \lambda \log A_{it} + \varepsilon_{it}$$

Note that we need to fix one of the β coefficients to have identification. We allow for industry and time effects and include aggregate patents and age of firm.

A linear approximation will also be used:

$$\log R_{it} = \alpha + \beta_0 \log I_{it} + \sum_{k=1}^{K} \beta_k \left(\frac{I_{ikt}}{I_{it}} \right) + \sum_{j=1}^{J} \delta_j D_j + \sum_{t=1}^{T} \gamma_t E_t + \gamma \log P_{jt} + \lambda \log A_{it} + \varepsilon_{it}$$

Multivariate Results (1/2)

Table 7 - Log linear Regressions - Dependent Variable: R&D
(deflated with a price index of year 2000)
(5514 useful observations, 957 firms, 18 countries, 1991-2005)

Variables	*Model 1*		*Model 2*		*Model 3*	
	yearly counts		*3 yrs moving avg*		*5 yrs moving avg*	
	coeff	*std*	*coeff*	*std*	*coeff*	*std*
Total number of inventors (log)	**0.728**	**0.012**	**0.740**	**0.011**	**0.748**	**0.011**
Share of inventors in EU 15 and Switzerland, Norway and Island	*0.521*	*0.309*	**0.868**	**0.311**	**1.021**	**0.309**
Share of inventors in Other Europe	-	-	-	-	-	-
Share of inventors in USA and Canada	**0.641**	**0.313**	**0.988**	**0.316**	**1.121**	**0.315**
Share of inventors in Far East (AU,HK,ID,JP,KR,MY,NZ,SG,TW)	**0.750**	**0.371**	**1.081**	**0.380**	**1.278**	**0.380**
Share of inventors in Other countries	0.464	0.444	**0.901**	**0.467**	**1.057**	**0.477**
Patents at 3 digit sectoral level	**0.074**	**0.016**	**0.067**	**0.016**	**0.066**	**0.016**
Adjusted R squared	**0.543**		**0.575**		**0.588**	

Multivariate Results (2/2)

Table 9 - NLLS Regressions with Disaggregated Countries - Dependent Variable: R&D (deflated with a price index of year 2000)

(5514 useful observations, 957 firms, 18 countries, 1991-2005)

Variables	*Annual flow of R&D expenditures*					
	Model 7		*Model 8*		*Model 9*	
	yearly counts		*3 yrs moving avg*		*5 yrs moving avg*	
	coeff	*std*	*coeff*	*std*	*coeff*	*std*
Inventors in Germany	***1.000***		***1.000***		***1.000***	
Inventors in Benelux	**0.658**	**0.120**	**0.882**	**0.168**	**1.071**	**0.214**
Inventors in France	**0.461**	**0.085**	**0.672**	**0.127**	**0.797**	**0.161**
Inventors in Italy	**0.739**	**0.169**	**1.010**	**0.232**	**1.260**	**0.311**
Inventors in Nordic countries (DK,FI,IS,NO,SE)	**0.318**	**0.053**	**0.572**	**0.093**	**0.812**	**0.137**
Inventors in Switzerland and Austria	**0.383**	**0.073**	**0.430**	**0.082**	**0.536**	**0.104**
Inventors in UK and Ireland	**0.561**	**0.076**	**0.999**	**0.135**	**1.495**	**0.210**
Inventors in USA and Canada	**0.678**	**0.077**	**0.986**	**0.122**	**1.302**	**0.172**
Inventors in Other countries	**0.884**	**0.233**	**1.279**	**0.342**	**1.730**	**0.481**
Inventors in Far East (AU,HK,ID,JP,KR,MY,NZ,SG,TW)	**0.547**	**0.106**	**0.701**	**0.145**	**0.918**	**0.198**
Patents at 3 digit sectoral level	**0.050**	**0.018**	0.023	0.018	0.011	0.018
Adjusted R squared	**0.562**		**0.589**		**0.598**	

Caveats, Conclusions & Extensions

- some encouraging results, but still lots to do
- large grain of salt –currently not capturing all inventors due to selection into EPO filings
- wage data could help to improve predictions

- Future Work
  - utilize data on inventors from all national European offices and the US and Canada
  - use wage data
  - calibrate model with more detailed R&D data (composition w.r.t. capital, labor, materials)

Sneak Preview of Productivity Estimates

Table 10 - Productivity Regressions - Dependent Variable: log Sales
(deflated with a price index of year 2000)
(4892 useful observations, 897 firms, 18 countries, 1991-2005)

Variables	*Model 1*		*Model 5*		*Model 8*	
	coeff	*std*	*coeff*	*std*	*coeff*	*std*
Cost of employees	**0.422**	**0.010**	**0.415**	**0.010**	**0.564**	**0.010**
Depreciation of the capital	**0.063**	**0.005**	**0.061**	**0.005**	**0.100**	**0.005**
Annual R&D	**0.506**	**0.012**	**0.493**	**0.012**		
Inventors in the home country - yearly counts						
Inventors in the home country - 3 yrs moving avg						
Inventors in the home country - 5 yrs moving avg					**0.233**	**0.012**
Share of foreign inventors - yearly counts						
Share of foreign inventors - 3 yrs moving avg						
Share of foreign inventors - 5 yrs moving avg			**0.459**	**0.053**	**0.598**	**0.060**
Patents at 3 digit sectoral level	**-0.092**	**0.013**	**-0.088**	**0.013**	**-0.054**	**0.015**
Adjusted R squared	**0.836**		**0.839**		**0.797**	