The importance (or not) of patents to UK firms

By Bronwyn Hall, Christian Helmers, Mark Rogers, and Vania Sena

Online Appendix: Data description

The dataset consists of four components, which are all linked by a unique enterprise business register number:

Business Structure Database (BSD): the dataset is derived from the Inter Departmental Business Register (IDBR) and provides longitudinal business demography information for the population of businesses in the UK. We use information on a company's industrial classification (SIC 92), employment, turnover, R&D, as well as incorporation and market exit dates from the BSD.[1] The BSD reports data at the level of the reporting unit, which in most cases coincides with the enterprise, which means that we aggregate the data up to the enterprise level.

UK Community Innovation Survey (CIS) 3, 4, and 5: the CIS is a stratified sample of firms with more than 10 employees drawn from the IDBR. The CIS contains detailed information on firms' self-reported innovative activities.[2] We use three surveys: CIS 3 which covers the period 1998-2000, CIS 4 which covers 2002-2004, and CIS 5 which covers 2004-2006. The sample frames differ for the three CIS waves both in terms of size and industry coverage. For CIS 3, the sample frame consists of 19,625 enterprises with responses from 8,172 enterprises (42 percent response rate); CIS 3 covers both production (manufacturing, mining, electricity, gas and water, construction) and services sectors whereas the retail sector was excluded. CIS 4 has the largest sample size out of the three CIS waves with a sample frame of 28,355 enterprises and responses from 16,446 enterprises (58 percent response rate); it also includes the following sectors: sale, maintenance & repair of motor vehicles (SIC 50); Retail Trade (SIC 52); and Hotels & restaurants (SIC 55). CIS 5 was answered by 14,872 firms which corresponds to a response rate of 53 percent (Robson and Haigh, 2008). It covers the same industries as CIS 4 with the addition of SIC 921 (motion picture and video activities) and 922 (radio and television activities).

[1] The definition of market exit is problematic. It is not possible to identify whether a firm has ceased trading or if it has merely undergone a change in structure that leads to its original reference number becoming extinct.

[2] The survey structure follows the Oslo Manual (OECD, 1992). See Mairesse and Mohnen (2010) for a detailed discussion of the CIS data.

Patent data: we use a match of UK patents obtained from Optics and EPO patents (designating the UK and obtained from EPO's Patstat database, version April 2010) with the IDBR. The patents-IDBR match was carried out by the ONS/UKIPO using firms' names as patent documents lack unique firm identifiers.[3] Since the matched data is based on the IDBR, it has population coverage and covers all patents filed at UKIPO, WIPO (possibly designating the UK through PCT route), and EPO (possibly designating the UK through the EPC route) by firms registered in the UK over the sample period. While our analysis relies on the application date of patents, we include all patents that have been published (which occurs at the UKIPO and EPO 18 months after filing).

Trademark data: trademarks were matched to the IDBR by the ONS/ UKIPO using firms' and applicants' names as trademark documents lack unique firm identifiers. The data contain both UK and Community (OHIM) trademarks applied for by firms registered in the UK during the sample period.

The BSD and CIS data were cleaned and modified/adapted in order to combine them into a single integrated dataset. In particular, the structure of CIS 3 differs considerably from CIS 4 and 5, which required a number of changes to make the different datasets compatible and consistent.

References:

Helmers C., P. Schautschick and M. Rogers (2011): Intellectual Property at the Firm-Level in the UK: The Oxford Firm-Level Intellectual Property Database, Oxford University, Department of Economics Discussion Paper No. 546.

OECD (1992): OECD Proposed Guidelines for Collecting and Interpreting Technological Innovation Data – Oslo Manual, OECD, Paris.

Mairesse J. and P. Mohnen (2010): Using Innovation Surveys for Econometric Analysis, Chapter for *Handbook of the Economics of Innovation*, B. H. Hall and N. Rosenberg (eds.), Elsevier.

Robson S. and G. Haigh (2008): First Findings from the UK Innovation Survey 2007, *Economic and Labour Market Review*, Vol. 2, No. 4, pp. 47-53.

[3] For a detailed description of the methodological challenges see Helmers et al. (2011).

Online Appendix Tables

Table A-1: Sector breakdown

	Population										*Innovating sample*		
Sector	*All firms*	*Sector share*	*Pop sector share*	*Inno only*	*R&D only*	*R&D & inno*	*R&D*	*UK or EPO patent*	*Share with patents*		*Innovating firms*	*Sector share*	*Share innov-ating*
Chemicals	1791	5.9%	4.5%	46	486	822	1308	185	10.3%		868	8.4%	48.5%
Food etc	1079	3.5%	2.3%	20	264	478	742	17	1.6%		498	4.8%	46.2%
Hightech	1476	4.8%	3.3%	46	300	790	1090	149	10.1%		836	8.1%	56.6%
Metals & machinery	3154	10.3%	8.0%	73	910	1220	2130	200	6.3%		1293	12.5%	41.0%
Other mfg	1244	4.1%	2.6%	31	355	481	836	53	4.3%		512	5.0%	41.2%
Printing	1116	3.7%	3.6%	41	331	423	754	12	1.1%		464	4.5%	41.6%
Textiles & apparel	637	2.1%	2.0%	21	158	247	405	18	2.8%		268	2.6%	42.1%
Wood & paper	654	2.1%	1.8%	16	205	232	437	28	4.3%		248	2.4%	37.9%
Business services	6658	21.8%	22.2%	175	2153	1666	3819	91	1.4%		1841	17.9%	27.7%
Computer services	757	2.5%	3.2%	24	155	757	912	19	2.5%		631	6.1%	83.4%
R&D services	1031	3.4%	4.4%	28	311	388	699	36	3.5%		180	1.7%	17.5%
Construction	2236	7.3%	10.0%	54	759	317	1076	*	*		371	3.6%	16.6%
FIRE	1781	5.8%	5.6%	50	536	506	1042	*	*		556	5.4%	31.2%
Trade	3919	12.8%	19.5%	162	1211	907	2118	46	1.2%		1069	10.4%	27.3%
Transportation & Utilities	2973	9.7%	7.1%	91	1009	584	1593	17	0.6%		675	6.5%	22.7%
Manufacturing	**11151**	**36.6%**	**28.2%**	**294**	**3009**	**4693**	**7702**	**662**	**5.9%**		**4987**	**48.4%**	**44.7%**
KIBS	**8446**	**27.7%**	**29.8%**	**226**	**2619**	**2811**	**5430**	**146**	**1.7%**		**2592**	**25.1%**	**30.7%**
Other non-mfg	**10909**	**35.8%**	**42.1%**	**357**	**3515**	**2314**	**5829**	**63**	**0.6%**		**2671**	**25.9%**	**24.5%**
Total	**30506**		**100.0%**	**878**	**9143**	**9818**	**18961**	**871**	**2.9%**		**10310**		

* Cells suppressed for disclosure reasons.

Table A-2

Importance of various IP protection methods for the sample

Method of IP protection	*All firms*			*Patenting firms*		
	Number	*Share*	*Pop share*	*Number*	*Share*	*Pop share*
All firms	10093			645		
Design	2623	26.0%	21.2%	387	40.0%	55.3%
Trademark	3353	33.2%	28.9%	416	43.0%	62.4%
Patent	2737	27.1%	21.7%	524	54.2%	83.3%
Registered IP	**2760**	**27.3%**	**22.4%**	**459**	**47.5%**	**70.7%**
Copyright	2854	28.3%	26.5%	323	33.4%	46.6%
Confidentiality	5174	51.3%	46.5%	504	52.1%	77.4%
Formal IP	2951	29.2%	24.7%	448	46.3%	67.9%
Secrecy	4713	46.7%	42.1%	472	48.8%	72.5%
Complexity	3810	37.7%	34.8%	408	42.2%	63.1%
Leadtime	5328	52.8%	49.7%	469	48.5%	73.1%
Informal IP	**4622**	**45.8%**	**42.0%**	**474**	**49.0%**	**74.1%**

The cells show the numbers and shares of firms for whom the indicated form of IP is of medium or high importance.

Correlation of the IP protection mechanisms

	Design	TM	Patents	Copyright	Conf.	Secrecy	Complex.	Leadtime
Design	1.000							
Trademarks	0.605	1.000						
Patents	0.632	0.597	1.000					
Copyright	0.518	0.525	0.484	1.000				
Confidentiality	0.372	0.406	0.403	0.449	1.000			
Secrecy	0.347	0.365	0.385	0.388	0.574	1.000		
Complexity	0.377	0.305	0.375	0.359	0.401	0.495	1.000	
Leadtime	0.310	0.305	0.305	0.302	0.394	0.450	0.497	1.000

Table A-3: Descriptive statistics for estimation sample of innovative firms (10,093 observations)

*Source****	*Name*	*Variable*	*Mean*	*S.D.*	*Median*
UKIPO/ Patstat	d_patent	D (has a UK or EPO patent)	0.064	0.244	0
	qFIP	Firm formal IP rating (0-3)	0.987	0.942	0.8
	qRIP	Firm registered IP rating (0-3)	0.867	1.021	0.33
	qIIP	Firm informal IP rating (0-3)	1.314	0.975	1.33
	sFIP	Average formal IP rating in industry (0-3)*	0.699	0.348	0.65
	sIIP	Average informal IP rating in industry (0-3)*	0.905	0.389	0.90
CIS		Importance of design IP (0-3)	0.789	1.101	0
		Importance of trademarks (0-3)	0.979	1.184	0
		Importance of patents (0-3)	0.833	1.163	0
		Importance of copyright (0-3)	0.872	1.128	0
		Importance of confidentiality agreements (0-3)	1.461	1.224	2
		Importance of secrecy (0-3)	1.347	1.163	1
		Importance of complexity (0-3)	1.113	1.079	1
		Importance of lead time (0-3)	1.483	1.174	2
		Importance of patents relative to secrecy (-3 to 3)	-0.328	0.799	-0.13
BSD		R&D spending (10,000s of GB pounds)**	2.181	4.15	1.57
CIS		D (firm does R&D this period)	0.458	0.498	0
BSD	lnage	Log age	2.721	0.681	2.83
		Age in years	17.287	9.469	16
	lnemply	Log employment	-2.662	1.643	-3.01
		Employment (1000s)	0.379	1.611	0.05
CIS	fincon	D (financial constraints)	0.399	0.489	0
BSD	d_export	D (export status)	0.228	0.419	0
	group	D (member of a group)	0.236	0.424	0
	d_f_own	D (foreign ownership)	0.343	0.475	0
UKIPO/OHIM	d_tm	D (has a UK or EU trademark)	0.089	0.285	0
CIS	q810	Turnover share from prods new to market (%)	6.208	15.541	0
	q820	Turnover share from prods new to firm, not to mkt (%)	7.005	15.711	0
	prodnF	D (product innov. new to firm, not to mkt)	0.306	0.461	0
	procnF	D (process innov. new to firm, not to mkt)	0.397	0.489	0
	prodnM	D (product innov. new to market)	0.398	0.489	0
	procnM	D (process innov. new to market)	0.162	0.369	0
Patstat		Log (average #IPCs per pat)	0.093	0.508	0
		Log (backward cites per pat)	0.132	0.624	0
		Log (forward cites per pat)	0.056	0.391	0
		Log (NPL cites per pat)	0.035	0.288	0

* These are computed across the sample at the UK SIC 3-digit level.

** For the 5964 nonzero observations only.

*** CIS data refer to 3-year period (CIS 3 1998-2000, CIS 4 2002-2004, CIS 5 2004-2006); data from BSD, UKIPO, OHIM, Patstat available annually and collpased to match 3-year CIS reference period. The patent and trademark data are at the enterprise level. In case of multi-establishemnt companies, CIS and BSD data were aggregated at the enterprise level.

Table A-4

Correlation matrix (10093 observations)

	d_patent	qFIP	qRIP	qIIP	sFIP	sIIP	lnage	lnemply	d_export	group	d_f_own	d_tm	fincon
d_patent	1												
qFIP	0.2567	1											
qRIP	0.2789	0.9475	1										
qIIP	0.1724	0.6527	0.5576	1									
sFIP	0.1587	0.3591	0.3383	0.3173	1								
sIIP	0.1635	0.3251	0.301	0.3453	0.9359	1							
lnage	0.0654	0.0369	0.0635	0.0103	0.1043	0.092	1						
lnemply	0.2178	0.235	0.2362	0.1668	0.1113	0.1008	0.2815	1					
d_export	0.1722	0.1597	0.1718	0.1237	0.1251	0.1298	0.1862	0.4961	1				
group	0.1943	0.1601	0.1569	0.1203	0.08	0.07	0.1728	0.4594	0.2402	1			
d_f_own	0.0341	0.0326	0.0435	0.0282	0.036	0.0482	-0.0022	0.1472	0.1881	0.0223	1		
d_tm	0.2722	0.2161	0.2284	0.121	0.095	0.0749	0.092	0.2455	0.1602	0.2018	-0.0003	1	
fincon	-0.0179	0.1377	0.117	0.1496	0.0095	0.0096	-0.0631	-0.0537	-0.0267	-0.0389	-0.0228	-0.0209	1
q810	0.0724	0.1341	0.1089	0.1844	0.0713	0.0829	-0.1289	-0.0662	-0.0229	-0.01	0.0134	0.0271	0.0507
q820	-0.0111	-0.0062	-0.0094	-0.0049	0.0196	0.021	-0.11	-0.08	-0.0715	-0.0257	-0.0172	-0.0063	0.0211
prodnF	-0.0798	-0.0964	-0.0911	0.1261	0.0089	-0.0148	-0.0004	-0.062	-0.1058	-0.0206	-0.0632	-0.046	-0.015
procnF	-0.0283	-0.0647	-0.0711	-0.043	-0.0339	-0.0268	-0.0473	0.059	0.0351	0.0016	0.0188	-0.016	-0.0158
prodnM	0.1502	0.2629	0.2417	0.2951	0.1669	0.1747	-0.0145	0.0705	0.0561	0.0691	0.0273	0.0987	0.0254
procnM	0.081	0.1019	0.0832	0.1479	0.0146	0.0276	-0.0324	0.0801	0.0548	0.0509	0.0097	0.0427	0.0433

	q810	q820	prodnF	procnF	prodnM	procnM
q810	1					
q820	-0.0386	1				
prodnF	-0.2657	0.3565	1			
		-	-			
procnF	-0.1322	0.0277	0.0621	1		
		-	-	-		
prodnM	0.4915	0.0666	0.5405	0.2001	1	
		-	-	-		
procnM	0.1505	0.0519	0.1969	0.3582	0.2064	1

See the previous table for definitions of the variables.