The University Of Sheffield.

Department Of Economics.

Sheffield Economic Research Paper Series.

Interlocking Directorships and Patenting Coordination

Michele Bernini
Georgios Efthyvoulou
Ian Gregory-Smith
Jolian McHardy
Antonio Navas

ISSN 1749-8368

SERPS no. 2014016
Published: October 2014

www.sheffield.ac.uk/economics

Interlocking directorships and patenting coordination

Michele Bernini, Georgios Efthyvoulou, Ian Gregory-Smith, Jolian McHardy, Antonio Navas *

October 2014

Abstract

The aim of this paper is to investigate the role interlocking directorships play in the patenting activities of UK companies and provide further insights into the channels through which this relationship emerges. Our empirical analysis produces three main results: first, interlocking leads to a higher number of successful patent applications; second, interlocked firms are more likely to cite each other's patents, especially around the moment of interlocking; and, third, interlocked companies tend to increase the technological similarity of their patent portfolio in the immediate period following their first interlock. To rationalise these results, we develop a theoretical model that identifies interlocking directorships as a practice that prevents property right conflicts that often arise between firms that are technologically close to each other.

JEL classification: O31 O32 D85 G30 J49
Keywords: Patents, director networks, knowledge spillovers, patent coordination

*All authors are employed in the economics department at the University of Sheffield. The authors would like to acknowledge financial support from the University of Sheffield's Departmental Research Investment Project fund. The authors would also like to thank participants at the Work, Pensions and Labour Economics Study Group (WPEG 2014) for useful comments and suggestions. Finally, the authors would like to thank Helmers, Rogers and Schautschick for providing us with access to the Oxford Firm-Level Intellectual Property Database. Corresponding author: m.bernini@sheffield.ac.uk.

1 Introduction

The allocation of resources in firms is entrusted to the wisdom of a small number of individuals who sit as directors on company boards. This concentration of decision-making is reinforced by director interlocks; that is one director at one company can sit on the board at another institution. That interlocks further concentrate decision rights to an even smaller number of individuals, who often share cultural and educational backgrounds (Mizruchi, 1996), has led to popular charges of corporate elitism (Schwartz, 1987) and restrictions on interlocking have been created to mitigate the risk of collusive behaviour (Monks and Minow, 2011).

Interlocks are an interesting phenomenon for reasons that go beyond concerns over collusion. Still, why exactly directors interlock remains unclear. Narratives of interlocking have been advanced with respect to enforcing collusive agreements (Pennings, 1980), increasing the CEO's bargaining power over the monitors of her pay and performance (Bebchuk and Fried, 2003), increasing the firm's reputation and legitimacy as perceived by providers of financial capital (Dooley, 1969; DiMaggio and Powell, 1983), and increasing the human capital of the interlocked director (Conyon and Read, 2006). Surprisingly very little has been done to understand the role they play in coordinating innovative activity or the implications for the transfer of knowledge between firms. Our working hypothesis is that interlocks play a significant role in the patenting behaviour of the firm. This stems from a rich literature that claims networks are an important source of knowledge diffusion and coordination between firms (e.g., Cohen and Levinthal, 1989; Mowery, 1990; Gemser and Wijnberg, 1995; Oerlemans *et al.*, 1998; Crépon *et al.*, 1998; Powell, 1998).

In addition to being relevant for a firm's ability to generate innovation, networks of interlocking directors may affect incentives to protect inventions with patents. Uncertainty has been identified as a key consideration in a firm's decision to patent its innovations (Lanjouw and Schankerman, 2001; Lemley and Shapiro, 2005; Heger and Zaby, 2013), and the literature suggests interlocking directorships are a way for the firm to reduce its operational and environmental uncertainty (Schoorman *et al.*, 1981; Mizruchi, 1996). For instance, directors with outside affiliations may reduce uncertainty relating to market demand, likelihood of new entrants, availability of inputs or innovations in production technology.

This paper examines the role interlocking directorships play in the patenting strategies of UK companies. We find that interlocking increases a firm's number of successful patent applications. In addition, we show that interlocked firms are more likely to cite each other around the moment of interlocking, and that after becoming interlocked, firms tend to converge in the technological classes under which their patents are classified. We use these empirical findings to inform the development of a theoretical model which identifies the conditions under which firms select interlocking and when interlocks increase the number of patents in expectation. The model also shows that interlocking will be more likely amongst firms that are close to each other in the technological space.

The empirical study that is closest to our own is a paper by Helmers *et al.* (2014). Using Indian data, these authors exploit exogenous changes in India's corporate governance framework and patent system to identify a positive impact of interlocking on patenting activity. In particular, an increase in interlocks appears to increase the propensity of the firm to patent its invention outside of India. Our findings confirm that the positive effect of interlocks on patenting holds also for the UK. However, our additional results on patent citations and technological convergence lead us to advance an alternative explanation to the one offered by Helmers *et al.* (2014). Whereas these authors interpret interlocks as a source of information on the strategic position of the firm operating in foreign markets, our results point towards the use of interlocks as a mechanism by which firms coordinate their patenting activity or transfer technological knowledge.

Our empirical and theoretical results have implications for policymakers in the field of corporate governance and intellectual property. In terms of corporate governance, opinion tends to be polarised around whether interlocked directors add value to the firm through greater levels of human capital accumulation or whether directors interlock to subvert the monitoring of their performance and their accountability to shareholders.[1] The evidence presented in our paper suggests a subtle process at work in which interlocked directors play an important role in facilitating the protection and coordination of intellectual property rights arising from innovative activity. This is also important in light of prior evidence suggesting that excessive and defensive patenting strategies by large firms impose considerable administrative costs on patenting authorities. If an interlocking director can reduce frictions arising from overlapping inventions (e.g., contested patent applications) then reducing restrictions on the number of interlocking directors may help reduce the burden on patenting authorities.

2 Data and network measures

Our main data source is the EPO PATSTAT database.[2] This database provides bibliographic information for all patents published by the major IP offices. Because our analysis is based on patent applications, it is important to notice that PATSTAT includes details for only those applications that were successful in generating patent protection. Another important feature of PATSTAT is that it identifies 'patent families', that are groups of applications referring to the same invention, which allows obtaining a more precise mapping from the number of applications to the number of distinct inventions.

Data on UK firms' directors are obtained from FAME A 2013.[3] We limit our analysis to manufacturing companies, identified on the basis of their principal economic activity (NACE). By focusing on manufacturing we map more precisely the economic units that apply for patents into those that implement a novel productive process or those that develop a patented product. On a practical perspective, our focus on manufacturing firms reduces the dataset to a manageable size even if the construction of network measures is computationally expensive. Although this sample cannot be considered representative of the entire population of UK manufacturing firms, it is arguably less skewed toward larger units than those used in most of the interlocking literature, which is generally focused only on listed companies (e.g., Helmers *et al.*, 2013; Croci and Grassi, 2013). For each firm we observe the list of current and previous directors, and their appointments and resignations dates. With this information we are able to associate each director to one or more companies over the period 1998-2012.[4] We match PATSTAT and FAME over the period 1998-2012 using strings of company names as a merging variable. Before executing the merge we standardise company names in both datasets to minimise the number of mismatches that are caused by differences in punctuation or abbreviations. Our standardization algorithm is similar to the one implemented by Helmers, Rogers and Schautschick (HRS) to match previous versions of these two datasets (Helmers and Rogers, 2011).[5]

[1]Presently, regulators in Europe and the United States have adopted a sceptical attitude towards interlocking. In the US, concerns have been raised by anti-trust authorities and Sarbannes-Oxley explicitly contains provisions against interlocking. In Europe, authorities generally adopt a more pragmatic approach via a 'comply or explain' regime. For example, an interlocked director in a listed firm in the UK violates the criteria for independence and, to comply with the Code of Best Practice, firms are prohibited from having more than half of their board as non-independent directors (Combined Code, 2012). Yet firms are free to disregard this recommendation so long as they explain their non-compliance to shareholders.

[2]We use the October 2013 version of PATSTAT.

[3]FAME A includes firms with a turnover or shareholder funds greater than 1.5 million pounds, or with profits greater than 150,000 pounds.

[4]Although FAME does not provide unique identification numbers for directors we exploit the date of birth to address cases of homonymity.

[5]Our match is very similar to the one obtained by HRS (73,914 common cases against 2,106 cases that are unmatched in our dataset but are in HRS). We also observe a large number of PATSTAT Applicant IDs that are matched in our dataset but not in HRS. These refer to patent applications filed after 2007, that are therefore excluded from HRSs' sample.

2.1 Interlocking directorships across UK companies

Interlocking directorships occur when directors sit on the boards of multiple companies within or outside the same business group. In the terminology of Social Network Analysis the matched list of companies and directors can be defined as an 'edgelist' of 'bipartite graphs', in which firm-director couples represent edges between two disjoint sets of nodes (i.e., directors and firms). Each bipartite graph is then transformed in its 'one-mode projection'; that is a network in which firms are nodes and interlocking directors are edges between nodes (König and Battiston, 2009).

In the simplest form, a network of N companies in period t can be represented by the adjacency matrix A_t; that is an $N \times N$ square matrix with entries $a_{ij} = 1$ if there is at least one director sitting at time t on the boards of firm i and j where $i \neq j$, and $a_{ij} = 0$ otherwise. We can also construct a network where we only consider connections between firms belonging to different business groups by eliminating from the original network all the within-group connections (i.e., connections between firms that share the same global ultimate owner). This network is used to construct DGA_t; that is the adjacency matrix where $a_{ij} = 1$ if there is an interlocking director between firm i and j, and if i and j belong to different groups. Lastly we obtain a third adjacency matrix SIA_t representing only edges between firms from the same 4-digit NACE industry. Adjacency matrices are then used to compute vectors of 'node degrees', whose entries record the total number of connections (ND_t) of each firm in the sample, the number of connections outside its business group ($DGND_t$), and the number of connections with firms that operate in the same 2-digit NACE industry ($SIND_t$):

$$\begin{aligned} ND_t &= A_t \times I \\ DGND_t &= DGA_t \times I \\ SIND_t &= SIA_t \times I \end{aligned}$$

where I is a $N \times 1$ column vector where each entry is equal to 1. Table 1 shows that the proportion of interlocked firms increases over the years, escalating from 43% in 1998 to 53% in 2012. There is also a decreasing trend in the proportion of interlocks outside the business group over the total number of connections, from 59% in 1998 to 33% in 2012. Instead, the proportion of connections between firms belonging to the same industry increases more slowly, passing from 30% in 1998 to 38% in 2012. The high proportion of interlocked firms, and the prevalence of intra-group connections may be explained by the composition of our sample that under-represents UK independent SMEs.

Table 1: Features of the interlocked network (1998-2012)

Year	Ratio $ND > 0$	Ratio $DGND/ND$	Ratio $SIND/ND$
1998	0.429	0.596	0.304
2003	0.471	0.519	0.332
2008	0.507	0.384	0.378
2012	0.532	0.337	0.386

Notes: The first column reports the proportion of firms with at least one connection. The second column reports the average ratio of 'out of group' connections over total connections across firms. The third columns reports the average proportion of connections with firms in the same industry over total connections across firms.

2.2 Interlocking directorships and patenting

Our first measure of innovative output is the number of patent applications $APPS_{i,t}$ filed by company i at time t. Although we retain all applications irrespective of the receiving authority, we avoid double counting by considering all documents belonging to the same patent family as a unique application.[6]. We also construct an indicator of firms patent stock $PATSTOCK_{i,t}$ following the common practice in the literature of applying a linear discount rate $\delta = 0.15$ to the cumulated stock of past applications (Griliches and Mairesse, 1984):

$$STOCK_{i,t} = (1 - \delta) \times STOCK_{i,t-1} + APPS_{i,t} \tag{1}$$

Table 2 reports the number of firms, the proportion of applicants and the proportion of firms with positive patent stock that we observe each year. The lower proportion of applicants in 2011 and 2012 is explained by the fact that PATSTAT 2013 reports only applications for which a patent has already been published. We are likely to miss some of the 2011 and 2012 applications that had not yet been published by October 2013 (i.e., when the snapshot of the patent dataset was taken) because on average 18 months are required before an eligible application translates into a publication.

To investigate the relationship between interlocking and innovative activity, we compute the proportion of patenting firms by interlocking status for each age bin. This allows us to acquire preliminary evidence on the relationship between interlocks and patenting over a firm's life cycle. Figure 1 shows that for almost all age levels, there is a greater proportion of innovators in the group of connected firms. In addition, the gap between the patenting intensity of the two groups widens over a firm's age. Looking at the patent stock (right-hand side panel), we can see that for firms with 1 year of age the difference in the proportion of firms with at least one patent between connected and unconnected firms is about 5%. This gap evolves to about 15% for firms that have been in business for over 40 years. This evidence is both consistent with the positive effect of interlocking on innovative behavior and with the greater likelihood for innovative firms to become interlocked.

Table 2: Patenting activity in the sample

Year	Num. of firms	Ratio $APPS_{i,t} > 0$	Ratio $STOCK_{i,t} > 0$
1998	13,935	0.076	0.311
1999	14,512	0.075	0.314
2000	15,114	0.076	0.320
2001	15,571	0.071	0.324
2002	15,999	0.069	0.326
2003	16,434	0.068	0.327
2004	16,579	0.061	0.327
2005	16,666	0.062	0.328
2006	16,710	0.061	0.330
2007	16,798	0.062	0.331
2008	16,723	0.057	0.332
2009	16,655	0.056	0.334
2010	16,733	0.053	0.334
2011	16,794	0.049	0.333
2012	16,567	0.045	0.335

Notes: The table reports the number of firms observed each year (column 2), the proportion of firms that fill at least one patent application (column 3), the proportion of firms with positive patent stock (column 4).

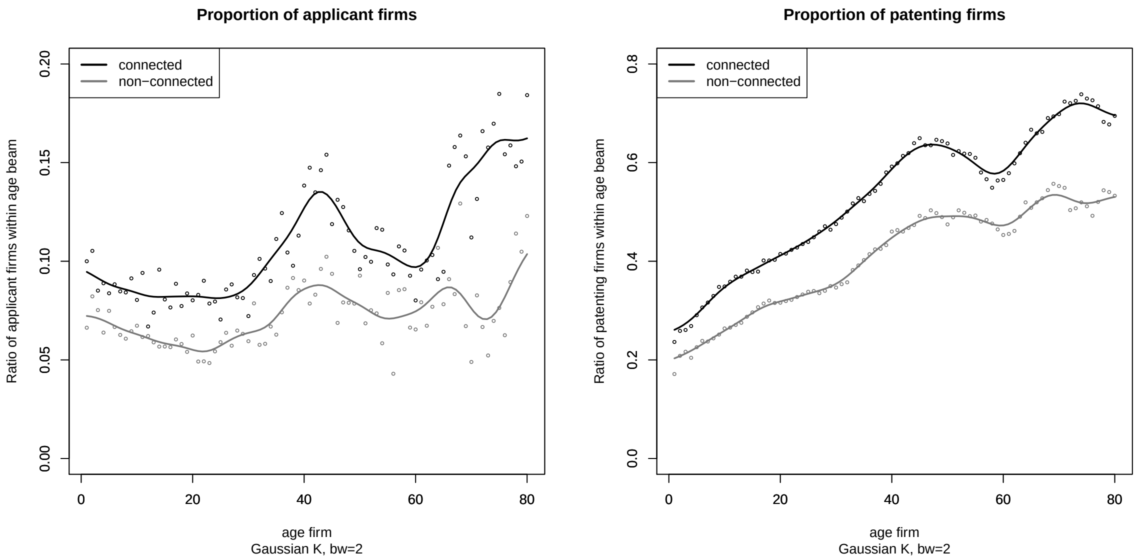
A second piece of evidence supporting a relationship between interlocks and innovative activity emerges when we graph the network of interlocked firms.[7] We find that patent applicants are often connected with other applicants. The endogeneous formation of interlocks is a possible explanation for

[6]Patent families are identified by the EPO by associating to a unique family all applications that refer to the same priority. A priority is the date of the first application to one of the patent offices.

[7]See Figure 3 in the Appendix.

this pattern, whereby innovators tend to interlock with other innovators. However, this pattern is also consistent with the presence of peer effects among connected firms, where the innovative behaviour of one company affects the innovative output of the others. In the next section we test more formally these hypotheses.

Figure 1: Patenting over firm life-cycle



Notes: The left-hand side panel plots the proportion of applicant firms ($APPS_{i,t} > 0$) by age and interlocking status ($ND_{i,t} > 0$). The right-hand side panel plots the proportion of firms with positive patent stock ($STOCK_{i,t} > 0$).

3 Econometric framework

This section describes the three empirical exercises that we conduct to investigate the relationship between interlocking directors and firms' innovative outputs. We first estimate the impact of connectedness and peer effects on patent applications. Finally we investigate the impact of interlocks on the probability that couples of firms will cite each other's patents. We then examine if there is convergence in the technological composition of firms' patent portfolios after becoming interlocked with each other.

3.1 Interlocks and patent applications

First, we describe the instrumental-variables (IV) strategy that we implement to identify: (a) the impact of interlocked directorships on a firm's innovation, and (b) the strength of peer effects across interlocked companies. Ideally, we would like to observe random connections across firms, and to measure the impact of connectedness as the difference between the expected innovative output of connected and unconnected companies, or as the different performance of companies that are randomly associated with more or less innovative partners. On the contrary, the network of interlocked firms is likely to evolve endogenously with respect to firms' innovative strategies and their accumulated knowledge. If a company's decisions to share directors with other companies are based on unobserved characteristics that are relevant for innovation, selection bias would prevent the identification of the impact of interlocks on innovative output using standard methods. Similarly, if the observed patenting activity of a firm provides a positive signal to potential partners, reverse causality may drive the positive correlation between interlocking status and innovation. Therefore, we adopt an IV strategy to estimate the following reduced form equation:

$$Y_{ist} = \gamma_1 C_{i,t-1} + \gamma_2 \Big(I(P_{i,t-1} \neq \emptyset) * \bar{Y}_{j \in P_{i,t-1}} \Big) + X'_{it}\mu + \delta_s + \delta_t + \alpha_1 \bar{Y}_{gt} + \alpha_2 \bar{Y}_{st} + \alpha_3 \bar{Y}_{ct} + u_i + \epsilon_{it} \quad (2)$$

where Y_{ist} measures the innovative performance of firm i operating in sector s at time t, X'_{it} is a vector of firm-level observable characteristics, while δ_s and δ_t are respectively 2-digit NACE industry and year effects. The terms $\bar{Y}_{gt}$, $\bar{Y}_{st}$ and $\bar{Y}_{ct}$ represent the innovative activity of companies that belong to the same business group, the same sector and the same county as firm i. These terms are introduced to control for spillovers affecting firm i's innovative activity that are not transmitted through interlocked directorship, but that are instead related to business group strategies or to technological and geographical proximity with other innovative companies. The term C_{it} is the main variable of interest and represents a firm's connectedness through interlocking directorship. The set P_{it} includes all firms interlocked with i at time t and $I(P_{it} \neq \emptyset)$ is an indicator function assuming value one if the set P_{it} is non-empty and value zero otherwise. $\bar{Y}_{j \in P_{it}}$ measures the innovative output of firm i's connections.

The parameters of interest are γ_1 and γ_2. The first one measures the direct effect of connectedness on Y_{ist}. The parameter γ_2 measures instead the peer effect generated by the innovative activities conducted by the firms interlocked with i (i.e., $\forall j \in P_{it}$). Note that this effect is present only if firm i has at least one connection. One can think of γ_2 as a specific channel through which greater connectedness, as measured by γ_1, affects innovation. Lastly, u_i is an unobserved firm-level fixed effect and ϵ_{it} is an individual firm error term.

The most serious identification problem raised by equation 2 is the omitted variable bias arising from the correlation between C_{it} and u_i. This problem occurs when firms interlock on the basis of unobservable characteristics that are relevant for innovation. Endogeneity of C_{it} may also arise because of reverse causality: if patenting companies are more attractive partners we may expect new patent applications to increase the opportunities of connections with companies seeking to capture knowledge spillovers by sharing directors. We address these problems by adopting a 2SLS estimator and two different sets of instruments for the endogenous measures of connectedness.

We first investigate the treatment effect of acquiring new connections on a firm's probability to apply for a patent. To do so, we estimate by 2SLS a linear probability model with the dummy variable $APPS_{it}$ appearing on the left-hand side of equation 2 and replace C_{it} with $add_{i,t-1}$; that is a dummy variable taking value one if the firm added at least one new connection at time $t-1$ and value zero otherwise. In the first-stage regression on $add_{i,t-1}$ we introduce two instruments excluded from the second-stage regression. These are the variables $Retire_{t-1}$ and $Hire_{t-1}$, that are respectively the ratio of retiring and newly hired directors over the total number of directors at time $t-1$. While firms that hire new directors have more opportunities to acquire new connections, the ratio of newly hired directors, unconditional on the past experience of the new hires, is not expected to impact directly on the innovative output of the company. In addition, this instrument is immune to reverse causality as it relates to connections that are acquired by the company by hiring a director that is already sitting on the board of another company. It is unrelated with connections arising from external hiring of a serving director that are more likely to arise when the company signals its knowledge stock through patenting.

Figure 1 suggests that connected and unconnected companies have a different probability of applying for patents at different age levels, and that this difference is reflected in a diverging evolution of the patent stock over their life-cycle. A second version of this model is meant to capture this long-run effect of connectedness on innovative output. In this version, we introduce the patent stock of the firm $STOCK_{it}$ as the dependent variable, and the lagged number of connections $nd_{i,t-1}$ for $C_{i,t-1}$. Because $nd_{i,t-1}$ is more persistent over time than $add_{i,t-1}$, in the first-stage we adopt two instruments that are more suitable to reflect this aspect of the endogeneous regressor: $YoungDir_{t-1}$ and $OldDir_{t-1}$. These two instruments are respectively the ratio of directors younger than 40 and the ratio of those older than 60 over the total number of directors sitting in the board. The proportion of interlocking directors by age group follows a bell-shaped distribution, with directors aged between 45 and 60 being most likely to serve multiple boards.[8] This relationship is most likely determined by the evolution of a director's reputation and connections over her career, and by her occupational choice in later life. 'Young directors' may still lack sufficient experience and reputation to be invited to sit in other companies' boards, while 'old directors' may choose to reduce the time spend sit on board meetings as they approach retirement.

The identifying assumption is that the proportion of young and old directors on the board does not affect directly a firm's patenting activities but only indirectly through the likelihood of interlocks. Some papers raised the point that a CEO's incentive to promote innovation may change over her tenure within a company and more generally over her career (e.g., Brickley *et al.*, 1999; Manso, 2011). However, it is very unlikely that the age of the CEO (and her career concerns) drive variations in our instruments because it enters their computation in the same way as the age of directors with no managerial role in the company. Hence, we argue that the age composition of the board is not a direct determinant of a firm's patenting activities, while it determines the likelihood of interlocks.

We proxy $\bar{Y}_{j \in P_{i,t-1}}$ as the average number of patent applications filed by companies that are directly interlocked with firm i, $\bar{Y}_{gt}$ as the average number of patent applications within the same business group, $\bar{Y}_{st}$ as the average number of patent applications of firms in the same 2-digit NACE industry, and $\bar{Y}_{st}$ as the average number of patent applications of firms based in the same county. Because the innovative output of interlocked firms $\bar{Y}_{j \in P_{i,t-1}}$ is not observed for firms with $nd_{it} = 0$ we will exclude this term from the model when we run regressions on the pooled sample of connected and unconnected companies. This is equal to imposing the restriction $\gamma_2 = 0$ and to identify the coefficient $\gamma_3 = \gamma_1 + \gamma_2 \left[\frac{1}{nd_t} \sum_{j \in P_{it}} Y_{jt} \right]$ for firms with a non-empty set of connected companies. In other words, when we estimate the restricted specification of the model we do not attempt to disentangle the unconditional effect of connectedness γ_1 (i.e., the effect of connectedness that does not depend on the innovative output of the connected firms). Then, we re-estimate an unrestricted specification of 2 on the sub-sample of firms with at least one connection (i.e., $nd_{it} \geq 1$). This second exercise allows us to identify the average effect of connectedness on the innovative outcome of connected firms after controlling for partners' heterogeneous innovative performance.

The endogeneity of $\bar{Y}_{j \in P_{i,t-1}}$ may depend on selection bias, if innovative companies are more likely to interlock with each other. A second problem that is often discussed in the peer-effect literature is that of 'reflection', that makes it impossible to identify the effect of peers' behaviours on individual behavior when both depend on the same set of group-level attributes (Manski, 1993). Moreover, peers' outcomes depend on individuals' outcomes generating reverse causality. We address these problems by taking advantage of the network structure of interlocking directors and we instrument the average number of patent applications among firm i's connections with the average patent stock of firm i's second-degree connections (i.e., the connections of firm i's connections that are not directly interlocked with firm i) (Bramoullé *et al.*, 2009). The identification assumption underpinning this strategy is that the characteristics of second-degree connections of firm i affects firm i's outcome only through their impact on the outcome of its first-degree connections.

3.2 Interlocks and patent citations

Ideally, we would like to compare the probability that a citation occurs between interlocked firms with the probability that it occurs between one of them and each one of all its 'placebo' interlocks, defined as those firms that are not interlocked with the target company but that are sufficiently similar to

[8] The bell-shaped relationship between a director's age and the probability she sits on more than one board emerges clearly in 4 in the Appendix.

its actual interlocks. Two issues prevent us from implementing this approach. First, the fact that we observe only a few partners for each interlocked firm does not allow the estimation of a propensity score that indicates which other companies are 'potential' partners. Second, the dataset including only 'actual' and 'placebo' interlocks may not be randomly drawn from the population of firms that may cite each others. Instead, estimation on the population of all possible firm couples is unfeasible because of the unmanageably large number of observations that must be generated. Our second best strategy is to restrict our estimation sample to the set of firm couples that cite each other at some point in time between 1998 and 2012. On this sample, we adopt a Difference-in-Differences model that identifies the causal impact of interlocks by exploiting the difference in probability of citation across couples of interlocked and non-interlocked companies between periods preceding and following the creation of an interlock. To do so we estimate the following probit model allowing for couple-specific random-effects u_{ij}:

$$Pr(C_{ij,t} = 1 | C_{ij,1998-2012} = 1) = \Phi(X'_{ij}\beta_0 + X'_j\beta_1 + X'_i\beta_2 + \sum_{s=-4}^{+4} c_{ij,t+s} + \delta_t) \tag{3}$$

where the dependent variable is the probability that firm i cites a patent of firm j at time t, conditional on observing at least one citation from i to j over the whole period. $\Phi(\cdot)$ is the cumulative probability function from the standard normal distribution. Its argument is a linear combination of the attributes of the citing and the cited firm X'_{it} and X'_{jt}, of couple-specific characteristics X'_{ij}, and includes a set of dummy variables that, for couples of interlocked companies, take value one in the period of their first connection, and in each one of the four periods preceding or following their interlock. These dummy variables take value 0 in all periods for couples of firms that do not interlock over the same period.

If interlocking directorship facilitates exchange of knowledge across companies we should expect that the difference in the probability of citation between couples of firms that interlock and those that do not interlock to be statistically insignificant in periods preceding the interlock and to be positive and significant in periods following the interlock. On the one hand, the condition $C_{ij} = 1$ ensures that we are considering only couples of firms that have the right characteristics to build on each other's knowledge. On the other hand, we cannot claim that we estimate the unconditional effect of interlocks on citation probability as we only exploit the timing of citation for identification.

3.3 Interlocks and technological convergence

In our third exercise we investigate the path of technological convergence/divergence of firms' patent portfolios as an alternative strategy to capture technological spillovers between interlocked firms. To do so we exploit the IPC technological classification of patents reported in PATSTAT to construct a time varying index of technological similarity of the patent portfolio of interlocked firms. This index is constructed as the one introduced by Schott (2004) to measure the similarity in the composition of exports across countries. Technological similarity between firm i and j is measured by the Patent Similarity Index PSI_{ijt} computed as:

$$PSI_{ijt} = \sum_{c \in I} min(s_{cit}, s_{cjt}) \tag{4}$$

where c is an index for IPC technological classes and I is the set of all classes observed in PATSTAT (i.e., defined as the first 4 characters and numbers of the IPC string as it appears on the patent application), s_{cit} and s_{cjt} are the shares of patents classified in subclass c in the portfolios of firms i and j evaluated at time t. This index ranges from 0 for complete technological dissimilarity, to 1 for complete technological similarity. If firms i and j apply for patents in more similar technological classes after getting connected, we should expect a positive effect of interlocked directorship on PSI_{ijt}.

For each couple of interlocked companies we compute this index for the whole period 1997-2012.[9] We then estimate the following model:

$$PSI_{ijt} = X'_{ij}\gamma_0 + X'_j\gamma_1 + X'_i\gamma_2 + \sum_{s=-4}^{+4} c_{ij,t+s} + \delta_t + \epsilon_{ijt} \tag{5}$$

where on the right-hand side we adopt the same specification used in the model on citations. Because we include only couples of interlocked companies in the estimation sample, this specification cannot be considered as a Dif-in-Dif model. On the contrary, identification relies on the comparison of the PSI of 'treated' firms before and after receiving the treatment (i.e., getting interlocked). Nevertheless, we can control for time-specific confounding factors by including year effects δ_t as firms get interlocked at different points in time. This model will be estimated by using a Random-Effect Tobit model to deal with the large number of 0 values in the distribution of the PSI.

4 Results

We first comment on the results obtained by estimating equation 2 on the sample including both interlocked and non-interlocked companies. The first four columns of Table 3 report the estimates of regressions on the dependent variable APP_{it}; that is a dummy variable that takes value one when the firm applies for at least one patent. In these specifications our regressor of interest is add_{t-1}; that is a dummy assuming value one if the firm has increased the number of interlocks in the previous period. The remaining columns report estimates from regressing the lagged number of interlocked connections nd_{t-1} on the patent stock of the company $STOCK_{it}$. We report first and second-stage estimates for both a 'short' specification (i.e., including only the variable of interest, the log of firms' age $log(age_t)$, industry and year fixed effects) and a 'long' specification including further controls. In the 'long' specification we control for firms' independence status $Indep_t$, its lagged capital intensity computed as the log of fixed assets over the number of employees $CapInt_{t-1}$, its size proxied by the logged number of employees $log(empl_{t-1})$, and the average number of patent applications at the 2-digit NACE industry level $\bar{Y}_{st}$, at the business group level $\bar{Y}_{gt}$, and at the UK county level $\bar{Y}_{ct}$.

The estimated coefficient of add_{t-1} in the second-stage regression on APP_{it} suggests that on average a firm increases by 14% the probability of applying for a patent by adding at least one new connection in the previous period. This effect is reduced to 6% once we control for other firm characteristics and spillovers. Because the overall proportion of applicants each year is on average 9%, the impact of increased connectedness appears economically significant, hence supporting the hypothesis that connectedness increases the expected returns or reduces the costs of patenting innovations. Over time, the higher patenting propensity of interlocked companies is reflected in the relationship between the number of firms' connections and the size of their patent stocks. The estimated coefficient of nd_{t-1} suggests that each connection increases on average the number of patents in a firm's portfolio by 0.6, and this result does not change once we estimate the 'long' specification of the model.

The Angrist-Pischke (AP) F statistics from first-stage regressions on add_{t-1} and nd_{t-1} reject the null hypothesis of weak instruments, while the Hansen J statistics does not reject the hypothesis that our instruments are uncorrelated with the second-stage errors. First-stage coefficients on $Retire_{t-1}$ and on $Hire_{t-1}$ confirm that the turnover in the board of directors positively affects the interlocking probability. As we expect, the proportion of younger and older directors on the board is negatively correlated with the number of interlocks. Therefore, we conclude that there is sufficient statistical support for our IV approach to claim that we correctly identify the positive impact of connectedness on the number of successful patent applications filed by a company.

[9] We limit this analysis to interlocked companies because the computation of this index is very time-expensive.

To measure the sensitivity of the relationship between a company's own patent stock and the patent intensity of its peers we repeat the estimation of the model on the sample of firms with at least one connection. The inclusion of the term $\bar{Y}_{j \in P_{i,t-1}}$ measuring the application intensity of firms' connections, or the average number of patents in their portfolio, is introduced to capture peer effects. Results are reported in Table 4. Once we include this term we find that the coefficient on add_{t-1} in regressions on APP_{it} is reduced to 0.05 and 0.03 respectively in 'short' and 'long' specifications of the model. In addition, the coefficient of nd_{t-1} on $STOCK_t$ is rendered insignificant. These results suggest that the impact of connectedness on a firm's patenting behaviour is conditional on the patenting activity of its connections. In other words, interlocks increase patent applications only for those firms that connect with peers that are active in patenting.

Table 3: Connectedness and patents

Dependent :	APP_{it}				$STOCK_{it}$			
Specification :	(a)		(b)		(a)		(b)	
Estim.Stage :	2nd Stage APP_{it}	1st Stage add_{t-1}	2nd Stage APP_{it}	1st Stage add_{t-1}	2nd Stage $STOCK_{it}$	1st Stage nd_{t-1}	2nd Stage $STOCK_{it}$	1st Stage nd_{t-1}
add_{t-1}	0.140***		0.063***					
	(0.010)		(0.017)					
nd_{t-1}					0.604***		0.586***	
					(0.087)		(0.218)	
$log(age_t)$	0.014***	0.003***	0.009***	-0.011***	0.325***	0.161***	0.288***	-0.061**
	(0.001)	(0.001)	(0.002)	(0.002)	(0.046)	(0.016)	(0.081)	(0.026)
$CapInt_{t-1}$			0.007***	-0.002***			0.290***	-0.006
			(0.001)	(0.001)			(0.051)	(0.015)
$Indep_t$			-0.017***	-0.024***			0.105	-0.599***
			(0.004)	(0.003)			(0.157)	(0.070)
$Log(empl_{t-1})$			0.029***	0.013***			0.759***	0.313***
			(0.002)	(0.001)			(0.119)	(0.021)
$BoardSize_t$			0.005***	0.009***			0.067	0.158***
			(0.001)	(0.001)			(0.056)	(0.011)
$\bar{Y}_{st}$			0.022***	-0.001			0.817***	0.110
			(0.006)	(0.004)			(0.265)	(0.069)
$\bar{Y}_{gt}$			0.005**	0.004***			0.245**	0.053***
			(0.002)	(0.001)			(0.102)	(0.021)
$\bar{Y}_{ct}$			0.042***	0.008			1.097***	0.052
			(0.009)	(0.006)			(0.261)	(0.084)
Excluded Instruments								
$Retire_{t-1}$		0.211***		0.231***				
		(0.006)		(0.011)				
$Hire_{t-1}$		0.358***		0.327***				
		(0.006)		(0.012)				
$YoungDir_{t-1}$						-0.660***		-0.545***
						(0.036)		(0.085)
$OldDir_{t-1}$						-1.027***		-0.801***
						(0.040)		(0.066)
Hansen J-test (p-value)	0.389		0.341		0.831		0.391	
AP F-test		3887.36		1287.89		367.93		75.34
Obs.	200,279	200,279	63,199	63,199	220,407	220,407	63,752	63,752

Notes. The table reports both first-stage and second-stage 2SLS estimation results of models on APP_{it} and $STOCK_{it}$. For each model we estimate a 'short' specification (a) including only the log of firm's age $log(age_t)$ as a control variable, and a 'long' specification (b) including the following set of firm-level controls: $CapInt_{t-1}$ is the log of firm's capital per employee, $Indep_t$ is a dummy for independent firms, $Log(empl_{t-1})$ is the log of firm's size proxied by the number of employees, $BoardSize_t$ is the number of directors in a company's board. $\bar{Y}_{st}$, $\bar{Y}_{gt}$, $\bar{Y}_{ct}$ is the average of the dependent variable across firms belonging to the same 2-digit NACE industry, the same business group of the same county. The set of excluded instruments include either $Hire_t$ and $Retire_t$ in models on APP_t, or $YoungDir_t$, $OldDir_t$ in models on $STOCK_t$. $YoungDir_t$ and $OldDir_t$ are respectively the ratio of directors under 40 years of age or above 60 in a company's board. $Hire_t$ and $Retire_t$ are respectively the ratio of retiring directors or newly hires directors in a company's board. Cluster robust standard errors are reported in parentheses (cluster unit: firm). Significance levels: *.1, **.05,***.01.

Estimates suggest that if peers' application intensities (i.e., the average number of applications among a company's connections) increases by one application the probability that a company applies for a patent increases by 3%. In the long run this effect has an important impact on connected firms' patent stocks, as we find that if peers' patent intensities (i.e., the average number of patents held by the peers of a company) increases by one patent on average the patent stock of the company increases by 0.6. This result explains the divergence in the patent stocks of connected and unconnected companies as they age over time, that we observed in Figure 1.

Table 5 reports the estimates obtained by regressing model 3 on a sample that includes all couples of citing and cited firms that we could identify in our database. This model investigates the relationship between interlocks and patenting by exploiting the timing of citation. In other words, we compare the probability of citation between two firms over the years preceding and following the creation of an interlock. Standard errors of the estimated coefficients are relatively large, due to the small number of interlocks in the sample (see Table 9 in the Appendix). The low precision of the point estimates suggests a qualitative interpretation of the results.

The three columns of Table 5 report the coefficients obtained by estimating the model on the whole sample, on the sample including only couples of firms belonging to the same business group, and on the sample including only couples of firms belonging to different business groups. Comparing the results obtained across these samples we may infer whether the relationship between interlocks and patenting is the same when connections are created in the presence of ownership ties between companies.[10] In the model we also control for the number of patents in the patent stock of the cited company $PatStock_{j,t}$, for the number of applications of the citing company $PatCount_{i,t}$, for the age of the two firms, and for their common belonging to the same 2-digit NACE industry. We include a dummy equal to one in the period of the first interlock between the two firms $lock_{ij,t}$, a set of dummies $lock_{ij,t-s}$ assuming value one in one of the s periods preceding the interlock, and a set of dummies $lock_{ij,t+s}$ assuming value one in one of the s periods following the interlock.[11]

By estimating the model on the whole sample, we find a statistically significant increase in the probability of citation in the period when the first interlock is created and in the following period. However, the coefficient on $lock_{ij,t+1}$ is only significant at the 10% level. Point estimates for the periods preceding the interlock are not significant at the 5% level. When we estimate the model on the split samples of firm couples that belong to the same or to different business groups, we find that the previous results are confirmed only for citations occurring between firms belonging to different groups.

We conclude this section by presenting the results for the analysis on technological convergence between interlocked companies. Table 6 reports the results of regressions on the Patent Similarity Index (PSI). The estimation sample now includes all couples of firms for which we observe interlocking in the period 1998-2012. Identification now relies on the comparison of the PSI in period preceding and following the interlock. On the whole sample, we find evidence of technological convergence starting one period before the first interlock. Coefficients on the dummies for the period of the interlock $lock_{ij,t}$ and for later periods are positive and significant at the 1% level, and they suggest that the PSI increases monotonically starting from the year before the interlock. Similarly to what we found for citations, there is no evidence of technological convergence for couples of interlocked companies that belong to the same business group. On the contrary, between couples of firms belonging to different groups, it appears that technological convergence starts later on, two years after the occurrence of the first interlock.

5 Theoretical model

In this section we present a theoretical model that rationalises our empirical results. We show that our main results can be explained within a framework in which interlocking directorships emerge as a substitute for costly litigation for securing intellectual property rights. According to evidence in the literature, patenting is beset by uncertainty because firms may have competing claims on a new technology. Firms that are closer technologically are more likely to enter into property rights conflicts. These conflicts could be resolved through costly litigation, the outcome of which is uncertain to the firm *ex-ante*. We identify conditions under which this uncertainty generates a sub-optimal number

[10] To compare the coefficients obtained on the three samples within a unique regression we also estimated a model including interaction terms between the variables of interest and a dummy taking value one if both the citing and the cited firm belong to the same business group. Results are in line with those reported in Table 5 and they are available upon request from the corresponding author.

[11] We report the specification including 4 dummies for periods preceding and following the interlock. We also run the regression including longer or shorter timing structures around the interlock, and the results are insensitive to different specifications.

Table 4: Connectedness and peer effects

	$(APP_{it} \| nd_{it} > 0)$		$(STOCK_{it} \| nd_{it} > 0)$	
	(1)	(2)	(3)	(4)
add_{t-1}	0.050***	0.032**		
	(0.010)	(0.016)		
nd_{t-1}			0.178	0.216
			(0.143)	(0.237)
$\bar{Y}_{j\in P_{i,t-1}}$	0.034***	0.037***	0.662***	0.718***
	(0.004)	(0.007)	(0.116)	(0.181)
$log(age_t)$	0.020***	0.013***	0.494***	0.383***
	(0.002)	(0.003)	(0.081)	(0.109)
$CapInt_{t-1}$		0.008***		0.284***
		(0.002)		(0.062)
$Log(empl_{t-1})$		0.029***		0.905***
		(0.003)		(0.129)
$Indep_t$		-0.026***		-0.281
		(0.008)		(0.278)
$BoardSize_t$		0.005***		0.056
		(0.001)		(0.058)
$\bar{Y}_{st}$		0.013		0.491
		(0.009)		(0.346)
$\bar{Y}_{gt}$		-0.006		-0.032
		(0.004)		(0.162)
$\bar{Y}_{ct}$		0.049***		1.047***
		(0.014)		(0.370)
nace FE	Yes	Yes	Yes	Yes
year FE	Yes	Yes	Yes	Yes
Hansen J-test (p-value)	0.292	0.381	0.811	0.270
AP F-test (add_{t-1} or nd_{t-1})	3560.41	1414.19	187.30	76.97
AP F-test ($\bar{Y}_{j\in P_{i,t-1}}$)	378.06	156.01	936.02	333.00
Obs.	101,114	37,900	110,735	38,303

Notes. The table reports second-stage 2SLS estimates of models on APP_{it} and $STOCK_{it}$ on firms with at least one interlock $nd_t > 0$. For each model we estimate both a 'short' specification including only the log of a firm's age $log(age_t)$ as a control variable, and a 'long' specification including the following set of firm-level controls: $CapInt_{t-1}$ is the log of a firm's capital per employee, $Indep_t$ is a dummy for independent firms, $Log(empl_{t-1})$ is the log of firm's size proxied by the number of employees, $BoardSize_t$ is the number of directors in a company's board. $\bar{Y}_{st}$, $\bar{Y}_{gt}$, $\bar{Y}_{ct}$ is the dependent variable averaged across firms belonging to the same 2-digit NACE industry, the same business group or the same county. The set of excluded instruments include $Hire_t$, $Retire_t$ and $\bar{Y}_{j\in P_{i,t-2}}$ in models on APP_t. And $YoungDir_t$, $OldDir_t$ and $\bar{Y}_{j\in P_{i,t-2}}$ in models on $STOCK_t$. The instrument $\bar{Y}_{j\in P_{i,t-2}}$ is respectively the average number of patent applications (in models on APP_t), or the average number of patents in firms' portfolio (in models on $STOCK_t$) computed across the second-degree connections of the company. $YoungDir_t$ and $OldDir_t$ are respectively the ratio of directors under 40 years of age or above 60 in a company's board. $Hire_t$ and $Retire_t$ are respectively the ratio of retiring directors or newly hires directors in a company's board. Cluster robust standard errors are reported in parentheses (cluster unit: firm). Significance levels: *.1, **.05,***.01.

Table 5: Probability of citation

	Whole sample	Same business group	Different business groups
$lock_{ij,t-4}$	0.305*	0.400	-0.077
	(0.183)	(0.247)	(0.349)
$lock_{ij,t-3}$	0.057	-0.104	0.048
	(0.191)	(0.269)	(0.305)
$lock_{ij,t-2}$	0.168	0.144	-0.070
	(0.159)	(0.210)	(0.292)
$lock_{ij,t-1}$	0.138	0.073	-0.055
	(0.159)	(0.223)	(0.260)
$lock_{ij,t}$	0.341**	0.002	0.508**
	(0.143)	(0.217)	(0.205)
$lock_{ij,t+1}$	0.289*	-0.041	0.447**
	(0.153)	(0.230)	(0.220)
$lock_{ij,t+2}$	0.205	0.049	0.170
	(0.161)	(0.241)	(0.233)
$lock_{ij,t+3}$	-0.014	-0.208	-0.002
	(0.180)	(0.258)	(0.264)
$lock_{ij,t+4}$	0.050	-0.049	-0.041
	(0.194)	(0.290)	(0.280)
$PatCount_{i,t}$	0.195***	0.307***	0.189***
	(0.007)	(0.030)	(0.007)
$PatStock_{j,t}$	0.036***	0.077***	0.036***
	(0.003)	(0.020)	(0.003)
$SameIndustry_{ij}$	0.215***	0.234***	0.187***
	(0.022)	(0.081)	(0.023)
$Log(age_{i,t})$	-0.054***	-0.079**	-0.052***
	(0.008)	(0.039)	(0.008)
$Log(age_{j,t})$	-0.007	-0.020	-0.004
	(0.009)	(0.041)	(0.009)
Couples	4,864	188	4,676
Obs.	64,647	2,442	62,205

Notes. The table reports Random-Effect Probit models where the panel unit is set at the level of each firms' couple. The estimation sample includes only firms that cite each others' patents at some point in time during the period 1998-2012. We report a set of dummies assuming value one in the year the couple interlocks ($lock_{ij,t}$) or in each of the four year before and after the interlock. These dummy takes value 0 for firm couples that do not interlock over this period of time. Estimation is conducted separately on the whole sample (column 1), on the sample including only couples of firms belonging to the same business group (column 2) and on that including only couples of firms from different business groups (column 3). Robust standard errors are reported in parentheses. Significance levels: *.1, **.05,***.01.

of patents. In those situations, we show that, under certain conditions, interlocking directorships increase the number of patents.

5.1 Model set up

Consider that there are two risk-neutral profit-maximising firms F_i, $i \in \{1,2\}$, and one risk-neutral utility-maximising director D_i per firm. We model the technological distance between two firms as the Euclidean distance S: $e = ||\boldsymbol{\rho_1} - \boldsymbol{\rho_2}||$, where S is a continuous and finite n-dimensional space $S \subset \mathbb{R}^n$ and $\boldsymbol{\rho_i}$ is the location of the existing technology for firm i (assumed to be common knowledge). Each firm has an "unambiguous property right" (UPR) over its existing technology. The *UPR* is defined as a situation in which no other firm can claim property rights over that technology. In our framework, patenting is required to achieve an *UPR* but it is not a sufficient condition if another firm patents the same technology. Each point in the technology space has a baseline rent r which can only be exploited if the firm has an *UPR* over it.[12] To allow the director to contribute to the value of the firm, we let the rent on an *UPR* technology attributed to firm i to be expanded by the firm's director according to:

$$r_i = r\tau(t) \tag{6}$$

where $\tau(.) > 1$, $\tau'(.) > 0$ and $\tau''(.) < 0$ and t is the quantity of time devoted by the director to expanding the profit opportunities associated with that technology. Director D_i has a maximum time

[12]Considering an alternative scenario in which firms can also imperfectly exploit new technologies without establishing an *UPR* will reduce the incentives for firms to patent but it will not alter the qualitative results that we develop below.

Table 6: Technological convergence

	Whole sample (1)	Same business group (2)	Different business groups (3)
$lock_{ij,t-4}$	-0.013***	-0.007	-0.016***
	(0.003)	(0.006)	(0.004)
$lock_{ij,t-3}$	-0.005	0.005	-0.012***
	(0.003)	(0.005)	(0.004)
$lock_{ij,t-2}$	0.001	0.004	-0.005
	(0.003)	(0.005)	(0.003)
$lock_{ij,t-1}$	0.009***	0.009*	0.003
	(0.003)	(0.005)	(0.003)
$lock_{ij,t}$	0.011***	0.002	0.002
	(0.003)	(0.004)	(0.003)
$lock_{ij,t+1}$	0.016***	0.006	0.006
	(0.003)	(0.005)	(0.004)
$lock_{ij,t+2}$	0.015***	0.000	0.013***
	(0.003)	(0.005)	(0.005)
$lock_{ij,t+3}$	0.018***	0.002	0.022***
	(0.004)	(0.005)	(0.005)
$lock_{ij,t+4}$	0.021***	0.008	0.020***
	(0.004)	(0.005)	(0.006)
$SameIndustry_{ij}$	0.098***	0.006	0.137***
	(0.004)	(0.011)	(0.007)
Couples	10,182	3,326	7,984
Obs.	152,730	42,004	110,726

Notes. The table reports Random-Effect Tobit estimates on the PSI. The estimation is repeated on the whole sample of firm-couples that interlock over the period 1998-2012 (column 1), the sample of firm couples belonging to the same business group (column 2), and the sample of interlocking firm couples from different business groups (column 3). Robust standard errors are reported in parentheses. Significance levels: *.1, **.05,***.01.

allocation $T = 1$ which it supplies inelastically to the firm.[13] In exchange D_i obtains a share $w \in (0, 1)$ of the profit of its firm, i.

The main focus of the paper is to explore how interlocking directorships impact upon the patenting decisions of firms. Consequently, we start from an initial situation in which firms have already discovered a new technology, $\boldsymbol{\rho}_i^* \in S$. We assume that each firm's discovery of a patentable new technology is common knowledge but that the location of that technology in the space is unknown for both firms. The problem for each firm is then to decide whether to try to, and by what means, establish *UPR* over its new technology in order to extract the associated rent. We assume that there is an ex-ante probability $p(e) \in [0, 1]$ that the two new technologies have an overlap, where $p'(e) \leq 0$ (i.e. the probability of overlap in the new technologies is declining in the Euclidian distance of the firms' original technologies).[14] For simplicity, we assume that the probability of the new technologies overlapping with either of the initial technologies is zero.

In order to analyse the optimal behaviour of each of the agents, we need to obtain the profits/rents in each of the scenarios. If a firm is successful in attaining an *UPR* on its new technology then it earns baseline rent r (in addition to that derived from the initial technology). We assume that the productivity of the director's time in expanding the profit opportunities of each of the technologies is identical and alongside the assumed concavity of $\tau(.)$ the director's time is optimally redistributed equally across all technologies with *UPR*. Hence, modifying Eq. (6), with two *UPR* technologies, director D_i expands the rent with time t according to:

$$r_i = 2r\tau\left(\frac{t}{2}\right) \tag{7}$$

We assume that firms incur a fixed cost, P, of applying for a patent. Once the firm invests this fixed cost, the patent is assigned to the firm. If only one firm applies for a patent it gains *UPR* on

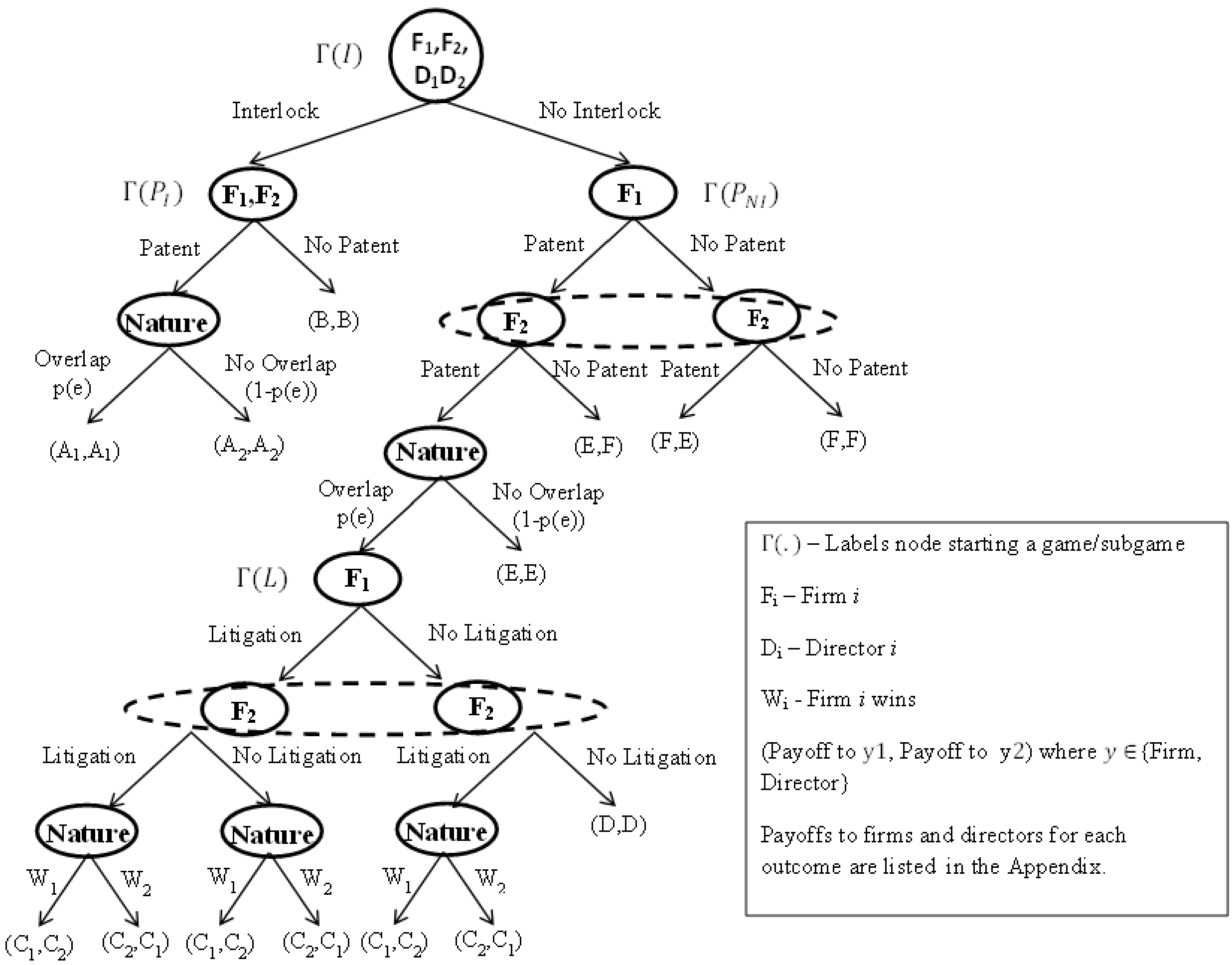
[13] A more general model that considers an elastic labor supply will not alter the main predictions of the model as what matters here is how the director distributes working time across the different activities within the firm.

[14] Interlocking directors could have a more direct impact on the innovative behaviour of the firm by altering the incentives to undertake innovations. Both mechanisms are complex and worthy of study separately. Consequently, in this paper we focus on the former while in a further paper (in progress) we outline the second mechanism.

its new technology. However, if both firms have obtained a patent for their technology the market (Nature) reveals if there is an overlap - with the probability of overlap being $p(e)$, as defined above. If there is no overlap, both firms have *UPR* on their new technologies and can extract the associated rents according to Eq. (7). In the case of an overlap, the firms do not have *UPR* on their new technologies and so cannot extract rent from them. In such cases the firms can try and resolve the property rights issue by litigation. Firms' decisions to litigate are simultaneous and independent. If one or both firms decide to litigate then each faces a cost L and the property rights carrying base line rent r are assigned to firm i with probability $\frac{1}{2}$. The losing firm is not able to extract rent from its new technology. If, upon patenting, Nature reveals an overlap and neither firm litigates then neither firm has *UPR* on its new technology hence no rent can be extracted from these technologies.

The full game is characterised by Figure 2. The possible outcomes in terms of the expected number of new patents (n) in this game is $n \in [0, 2]$, where 2 represents the case where both firms patent with certainty. In the next subsection we will focus on the subgame $\Gamma(P_{NI})$ and therefore we will study the firms' incentives to patent when their unique option to solve this conflict is to enter into litigation. We solve the model by backward induction. The following assumption guarantees that in a situation

Figure 2: Game Tree



in which there is no overlap and consequently no conflict, firms would profitably be able to patent and exploit their new technologies.

Assumption 1. *[Patent Viability] Firms strictly gain from patenting the new technology if there is no overlap, hence:*

$$\pi_E > \pi_F \quad \Rightarrow \quad \nabla \equiv \left[2r\tau\left(\frac{1}{2}\right) - r\tau(1)\right] > P$$

where π_E and π_F, defined in Table 10 in the Appendix, are, respectively, the firm's profit when it patents and obtains an UPR over the new technology and the firm's profit when, instead, the new technology is not patented and exploited. Given the bracketed expression on the left hand side of the inequality will appear frequently in the model we abbreviate it by ∇. This is the full innovation rent.

Two further expressions feature repeatedly throughout model, for which we introduce the following notational shorthand:

$$\Delta_1 \equiv \frac{2[p(e)L + P]}{2 - p(e)} \quad \Delta_2 \equiv \frac{P}{1 - p(e)}$$

Notice that Δ_1 and Δ_2 are increasing in $p(e)$ or in another words, decreasing in e: a decrease in the distance between the firms increases both Δ_1 and Δ_2.

5.2 Litigation to solve the patent conflict

5.2.1 The Litigation Subgame $\Gamma(L)$

In this subgame the firms decide non cooperatively whether to start litigation, conditional on both having patented and a property rights conflict arising (Nature having determined there is an overlap in the new technologies). The subgame is characterised by the matrix in Table 7, where

$$E(\pi_L) = \frac{1}{2}(\pi_{C_1} + \pi_{C_2}) = \frac{1}{2}(\pi_E + \pi_F - P(1-w)) - L(1-w)$$

$$\pi_{NL} = \pi_F - P(1-w)$$

$\frac{1}{2}(\pi_{C_1} + \pi_{C_2})$ is the expected profits in the case of patenting and litigating, π_{C_1} is the profit if the firm's wins the patent and π_{C_2} is the profit if the firm loses the patent.

Table 7: Subgame $\Gamma(P_L)$: No Interlocking with Patenting and Overlapping - Litigation (L) versus No Litigation (NL)

	L_2	NL_2
L_1	$(E(\pi_L), E(\pi_L))$	$(E(\pi_L), E(\pi_L))$
NL_1	$(E(\pi_L), E(\pi_L))$	(π_{NL}, π_{NL})

In this subgame, two cases can arise depending upon the relative size of $E(\pi_L)$ and π_{NL}. The results are summarised in Lemma 1.

Lemma 1. [15] *If $E(\pi_L) \geqslant \pi_{NL}$, hence $\nabla \geqslant 2L$, then one or both firms play litigation. Otherwise, $\nabla < 2L$, and both firms play no litigation.*

Notice that with $E(\pi_L) \geqslant \pi_{NL}$, we have three pure strategy Nash equilibria, one in which both firms litigate and another two in which only one of the firms litigates.[16] The main reason is that when one firm litigates, the other firm becomes indifferent between litigation or not because it faces the same expected profits and the same costs (the firm needs to incur the same fixed cost L to defend itself). From our point of view, all three equilibria are identical since payoffs for both firms are identical and also the expected number of new patents, the focus of our study, is identical too. Notice, also, that

[15] Where nontrivial, proofs are reported in the Appendix.

[16] In the exceptional case in which the equality holds any pair of strategies constitute a pure strategy Nash equilibrium. As in the case above all have identical payoffs.

litigation will be profitable for firms when it costs less (lower L) or when $\tau(.)$ is relatively concave exhibiting decreasing returns to scale in the director's effort so it is convenient to split the director's time across as many technologies as possible.

5.2.2 The Patent Subgame $\Gamma(P_{NI})$

The next step is to examine the firm's decision to patent.

Litigation is optimal $E(\pi_L) \geqslant \pi_{NL}$

In this subsection we will focus on the case where litigation in subgame $\Gamma(L)$ is optimal. The associated payoff matrix is given by Table 8(a), where:

$$E(\pi_{PL}) = p(e)E(\pi_L) + (1 - p(e))\pi_E = \left(1 - \left(\frac{p(e)}{2}\right)\right)\pi_E + \frac{p(e)}{2}\left[\pi_F - P(1 - w)\right] - p(e)L(1 - w) \quad (9)$$

Lemma 2 summarises the results for this scenario which depend upon the relative size of $E(\pi_{PL})$ and π_F.

Lemma 2. *If litigation is optimal ($E(\pi_L) \geqslant \pi_{NL}$), then we can distinguish between two situations i) if $E(\pi_{PL}) > \pi_F$, (hence $\nabla > \Delta_1$) then there is a unique pure strategy Nash equilibrium in which both firms patent and litigate with the expected number of patents equal to 2 and expected profit $E(\pi_{PL})$; (ii) if $E(\pi_{PL}) < \pi_F$ (hence $\nabla < \Delta_1$) then there is mixed strategy symmetric Nash equilibrium in which firms patent with probability, γ_L.[17] The expected number of patents is $2\gamma_L \in (0, 2)$ and expected profit is π_F.*

Notice that patenting (and litigation) will happen, ceteris paribus, when the cost of patenting (litigation) is relatively low and/or and the probability of an overlap is relatively low. In other words, if firms are initially close technologically, so that a probability of an overlap is high, firms will only patent on a random basis, looking for cases in which the other firm does not patent.

Corollary 1. *The expected number of patents under the hypothesis of Lemma 2(i), $2\gamma_L$, is decreasing in P and L and increasing in e and the concavity of $\tau(.)$.*

Litigation is not optimal ($E(\pi_L) < \pi_{NL}$)

When litigation becomes relatively expensive, the expected profits of patenting in subgame $\Gamma(P_{NI})$ are given by:

$$E(\pi_{PN}) = p(e)\pi_{NL} + (1 - p(e))\pi_E = p(e)(\pi_F - P(1 - w)) + (1 - p(e))\pi_E \quad (10)$$

and the subgame is characterised by the payoff matrix in Table 8(b).

Table 8: Subgame $\Gamma(P_{NI})$: No Interlocking, Patent (P) versus No Patent (NP)

(a) $\pi_{NL} \leqslant E(\pi_L)$

	P_2	NP_2
P_1	$(E(\pi_{PL}), E(\pi_{PL}))$	(π_E, π_F)
NP_1	(π_F, π_E)	(π_F, π_F)

(b) $\pi_{NL} > E(\pi_L)$

	P_2	NP_2
P_1	$(E(\pi_{PN}), E(\pi_{PN}))$	(π_E, π_F)
NP_1	(π_F, π_E)	(π_F, π_F)

[17] See the proof of Lemma 2 in the Appendix for the definition of γ_L.

Lemma 3. *If litigation is not optimal ($E(\pi_L) < \pi_{NL}$), two situations can arise i) if $E(\pi_{PN}) > \pi_F$, (hence $\nabla > \Delta_2$), then the unique pure strategy Nash equilibrium is both firms patenting and without litigation the expected number of patents is equal to 2 and expected profit is $E(\pi_{PN})$; ii) if $E(\pi_{PN}) < \pi_F$ (hence $\nabla < \Delta_2$) then there exists a mixed strategy symmetric Nash equilibrium in which firms patent with probability, γ_N[18]. The expected number of patents in this case is given by $2\gamma_N \in (0, 2)$ and expected profit is π_F.*

Corollary 2. *The expected number of patents under the hypothesis of Lemma 3(ii), $2\gamma_N$, is decreasing in P and increasing in e and the concavity of $\tau(.)$.*

5.2.3 Summary of the results without interlocking

Following from Lemmas 2 and 3, Table 11, in the Appendix, provides a categorisation of the outcomes of the subgame $\Gamma(P_{NI})$ according to the relative size of P, L and $p(e)$. Since the primary concern of the theoretical model is to help understand the circumstances under which the availability of interlocking may increase the number of patents, Corollary 3 identifies, from results presented in Lemmas 2 and 3, four classes of parameter values which result in Nash equilibria for the subgame $\Gamma(P_{NI})$ with expected number of patents being strictly less than 2. Hence, these represent situations where the expected number of patents may be increased under interlocking.

Corollary 3. *The following is a subcategory of cases in subgame $\Gamma(P_{NI})$ which result in $E(n) < 2$ with expected profit π_F (U_F):*
***Case 1 - High patent cost** , ($P > 2L$, **with** $\nabla < \Delta_1$) yields a mixed strategy Nash equilibrium in which firms patent with probability γ_L and litigate under an overlap and $E(n) = 2\gamma_L \in (0, 2)$.*
***Case 2 - Intermediate patent cost gains high** ($2(1 - p(e))L < P < 2L$, **with** $2L < \nabla < \Delta_1$) yields a mixed strategy Nash Equilibrium in which both firms patent with probability γ_L and litigate under an overlap and $E(n) = 2\gamma_L \in (0, 2)$.*
***Case 3 - Intermediate patent cost gains low** ($2(1 - p(e))L < P < 2L$, **with** $\nabla < 2L$) yields a mixed strategy Nash equlibrium in which firms patent with probability γ_N and they do not litigate and $E(n) = 2\gamma_N \in (0, 2)$.*
***Case 4 - Low patent cost** ($P < 2(1-p(e))L$, **with** $\nabla < \Delta_2$) yields a mixed strategy Nash equilibrium in which firms patent with probability γ_N and they do not litigate and $E(n) = 2\gamma_N \in (0, 2)$.*

In the next subsection we introduce the firm's decision to interlock and we initially explore the conditions under which, being interlocked, it is profitable to patent. We will show that for each of the four cases described in Corollary 3, where $E(n) < 2$, there is a parameter configuration in which both firms and directors find it profitable to interlock and patent both technologies. In these cases, in the absence of interlocking, firms definitely patent less. Consequently, in these circumstances, allowing the firms the option of interlocking increase firms' patenting behaviour.

5.3 Introducing Interlocking

Consider now that firms have the possibility of interlocking in the first-stage (before any commitment to patenting is undertaken). For interlocking to happen we assume that it must be incentive compatible for both directors and both firms.[19] Under an interlock agreement, each firm incurs an organisational overhead, f.

[18] See the proof of Lemma 3 in the Appendix for the definition of γ_N.

[19] The idea that the decision to interlock has to be incentive compatible with all four players, including the directors, reflects the observation that firms are not completely in control of what interlocks its directors choose to engage with as suggested, for instance, by setting up remuneration schemes to disincentivise excessive interlocking (e.g. see Conyon and Read, 2006).

As the interlocked directors have a stake in both firms, they are able to coordinate in order to reach the optimal patenting policy for both firms. The interlocked firms may decide to pursue both patents or neither.[20],[21] Further, in the case that patenting is selected and there is an overlap in the new technologies, the interlocking directors can ensure that a proportion $\theta \in (0, 1]$ of the rents associated with the new technology is preserved despite the overlap and oversee the allocation of the property rights such that firms get equal shares. Each directors does this at time cost x, leaving $1 - x$ for other rent-expanding work on the resulting $1 + \theta \in (1, 2)$ worth of *UPR* technology for their home firm. In return for interlocking work, director D_i earns a share $v \in (0, w)$ of the profit of firm j but the home firm reduces the share of profit it pays its own director $w(x) < w$ reflecting the reduced time that is spent expanding rents for firm i.[22]

We solve the model again by backward induction. In the next subsection we focus on the decision of patenting, conditional on being interlocked. In addition, we will show under certain conditions, interlocking promotes more patenting than its best alternative.

5.3.1 The Interlocking Subgame $\Gamma(P_I)$

We turn to the firms' decisions to patent. Once firms and directors have agreed to interlock, the decision regarding whether or not to patent falls to the firms. Firms cooperatively decide between patenting and not patenting both new technologies with expected profits, respectively:

$$E(\pi_{IP}) = p(e)\pi_{A_1} + (1 - p(e))\pi_{A_2}, \quad \pi_B = \pi_F - f(1 - w)$$

where π_{A_1} is the profit when an overlap exists and π_{A_2} is the profit when an overlap does not exist.

Lemma 4. *The Nash equilibrium of the subgame $\Gamma(P_I)$ is for the firms to patent if $E(\pi_{IP}) > \pi_B$, for which a sufficient condition is $E(\pi_{IP}) > \pi_F$, and hence:*

$$\begin{aligned}
&(1 - w)(1 - p(e))\left[2r\tau\left(\frac{1}{2}\right) - r\tau(1) - f - P\right] + p(e)(1 - w(x))\left[(1 + \theta)r\tau\left(\frac{1 - x}{1 + \theta}\right) - f - P\right] \\
&> vp(e)\left[(1 + \theta)r\tau\left(\frac{1 - x}{1 + \theta}\right) - f - P\right] \qquad (11)
\end{aligned}$$

where the right hand side of the inequality is the remuneration of the rival firm's interlocking director weighted by the probability that there is overlapping.

5.3.2 The Interlocking Decision

Up until now the relative payoffs of the directors have not featured in decision-making, but of course in the decision whether to interlock or not, all parties have to be in favour. Hence, we now introduce

[20] For modelling convenience, we assume that under an interlock only symmetric outcomes are feasible - hence we rule out the scenario in which one firm patents and the other does not. If the game were repeatedly played across different pairs of new technologies it would be possible to imagine a scenario in which one firm might forgo property rights on its new technology in one play of the game knowing the interlocking directors will ensure it is allocated the next one.

[21] In practice interlocking takes the form of a director from one firm (the interlocking director) sitting on the board of another firm rather than each firm committing a director to interlocking. We assume the latter for symmetrical expedience without meaningfully affecting the nature of the results.

[22] This represents a simplification of the remuneration and incentive system of interlocking directors, excluding, amongst other things, any human capital gains from interlocking (see for example Conyon and Read, 2006), but preserves the essential properties i.e. there is an opportunity cost to the home firm and its director of interlocking since it decreases the time available for directors to expand rents and earn their respective share of associated profit. Note, that taking into account human capital effects for interlocking directors, for example through a concavity-preserving monotonic transformation of $\tau(.)$, would, ceteris paribus, promote interlocking to both firms and directors.

the expected utility for the directors under interlocking and patenting:

$$E(U_{IP}) = p(e)U_{A_1} + (1 - p(e))U_{A_2} \tag{12}$$

The following Lemma sets out an important condition:

Lemma 5. *Expected utility for the directors under interlocking and patenting, $E(U_{IP})$, is greater than under no interlocking and no patenting, π_F, if:*

$$p(e)(w(x) + v)\left[(1 + \theta)r\tau\left(\frac{1 - x}{1 + \theta}\right) - f - P\right] + w\left\{\left[2r\tau\left(\frac{1}{2}\right)\right](1 - p(e)) - r\tau(1)\right\} > 0 \tag{13}$$

The main lesson from our model is that under certain scenarios patenting activity under interlocking is higher than without interlocking, conditional on interlocking being the optimal strategy in the new game. More precisely, we arrive at the following proposition:

Proposition 1. *If $E(\pi_{IP}) > \pi_F$ and $E(U_{IP}) > U_F$ and if litigation is optimal (case 1 and case 2 in Corollary 3) then the Nash equilibrium for game $\Gamma(I)$ is for firms and directors to interlock and patent, yielding an increase in the expected number of new patents of $2(1 - \gamma_L) > 0$. If $E(\pi_{IP}) > \pi_F$ and $E(U_{IP}) > U_F$ and litigation is not optimal (case 3 and case 4 in Corollary 3) then the Nash equilibrium for game $\Gamma(I)$ is for firms and directors to interlock and patent, yielding an increase in the expected number of new patents of $2(1 - \gamma_N) > 0$. In both cases we have an increase in the number of patents relative to the expected number that would prevail in the Nash equilibrium without interlocks.*

5.4 Discussion

The theoretical model above has shown that a number of patents below 2 can emerge from ambiguities over the rights on intellectual property. When firms are technologically close to each other, the probability of overlapping increases, widening the intervals, discussed in Corollary 3, where the number of patents is smaller than 2. In these cases interlocking emerge as an alternative option that results in an increase in the number of patents. Although a fall in e strictly reduces the expected rewards to interlocking, this option is still more profitable than the alternatives in Corollary 3 if Lemmas 4 and 5 hold. In turn, both lemmas are more likely to hold when the gains to interlocking are higher (θ), when the director time cost x is smaller, when fixed cost of interlocking f and patenting P are smaller, and when $\tau(.)$ is more concave. Indeed, so long as the payment v required to incentivise the directors to interlock is small, then the inequality will hold even if e (distance) is small (i.e. when the probability of overlap, $p(e)$, is large).

To conclude, we have shown that where firms wish to patent and exploit viable new technologies, but face a risk of property rights overlapping with rivals, then viable technologies may not be patented. However, we have also seen that where circumstances favour a number of patents less than 2, introducing the option of interlocking directors to help disentangle property rights ambiguities increases the number of patents. A key role is played by the technological distance between firms.

6 Conclusion

This paper provides some new insights into the role of the interlocking directors for patenting activity. In particular, it contributes to the literature in four main aspects. First, we use data from about 70,000 firms in the UK over the period 1998-2012 to investigate the impact of connectedness and peer effect on patent applications. Second, we employ information on patent citations to explore the extent to which interlocking firms cite each other's work. Third, we examine whether interlocking directorship leads to convergence in the technological composition of firms' patent portfolios. Finally, we develop a theoretical model that rationalises our empirical results and explores the conditions under which firms decide to engage in interlocking, leading to an increase in patenting activity in the industry.

The overall picture that emerges from our empirical analysis suggests that interlocking and patenting behaviour are strongly positively related. Specifically, we find that adding at least one new connection increases the probability of applying for a patent in the next year up to 14%. In addition, the impact of connectedness on firms' patenting behaviour appears to be conditional on the patenting activity of the firm's connection: a rise in peers' patent intensity by one application increases the probability of applying for a patent in the next year by 3%. Concerning the relationship between interlocks, citations and technological convergence two key results emerge: first, interlocked firms are more likely to cite each other, especially around the moment of interlocking; and second, interlocked companies tend to exhibit high degree of technological similarity of their patent portfolio in the immediate period following their first interlock. To provide an explanation of these results, we develop a formal framework that identifies interlocking directorships as an alternative to litigation for resolving property right conflicts that often arise between firms that are technologically close to each other. In particular, we argue that interlocking directors can prevent such conflicts by allocating appropriate rents to the interlocked companies.

From a policy point of view, our results emphasise the role of interlocking directorates as a one of the driving forces behind higher patenting activity and innovation performance. Therefore, adopting a positive stance with respect to interlocking could have a positive impact on innovation and reduce patent wars. Furthermore, to the extent that interlocking directors can act as key players in facilitating the protection of intellectual property rights and mitigate frictions arising from overlapping inventions, reducing restrictions on the number of interlocking directors may help alleviate large administrative costs for patenting authorities.

Empirical Appendix

Figure 3: Patent applications across interlocked firms in 1998

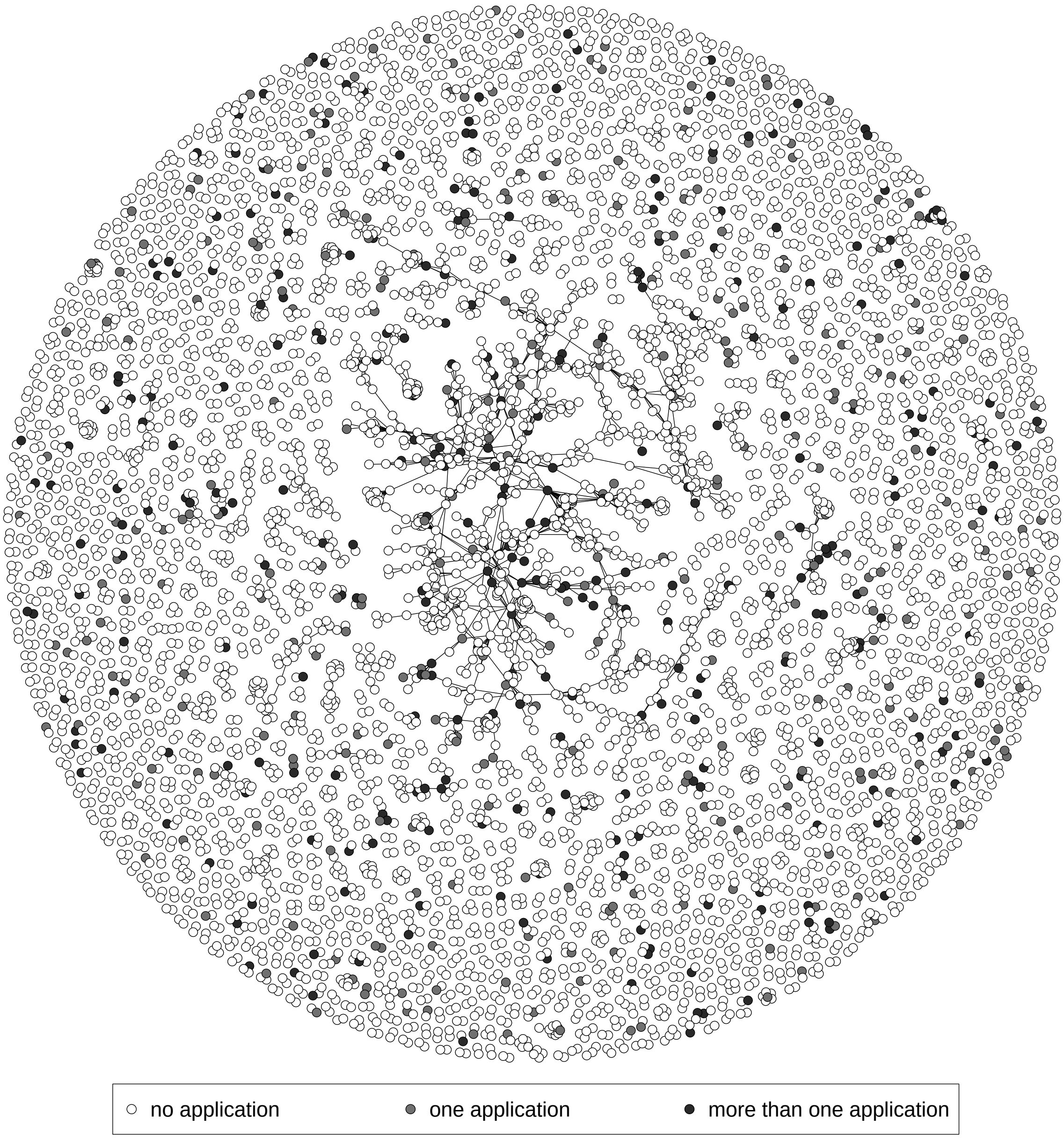
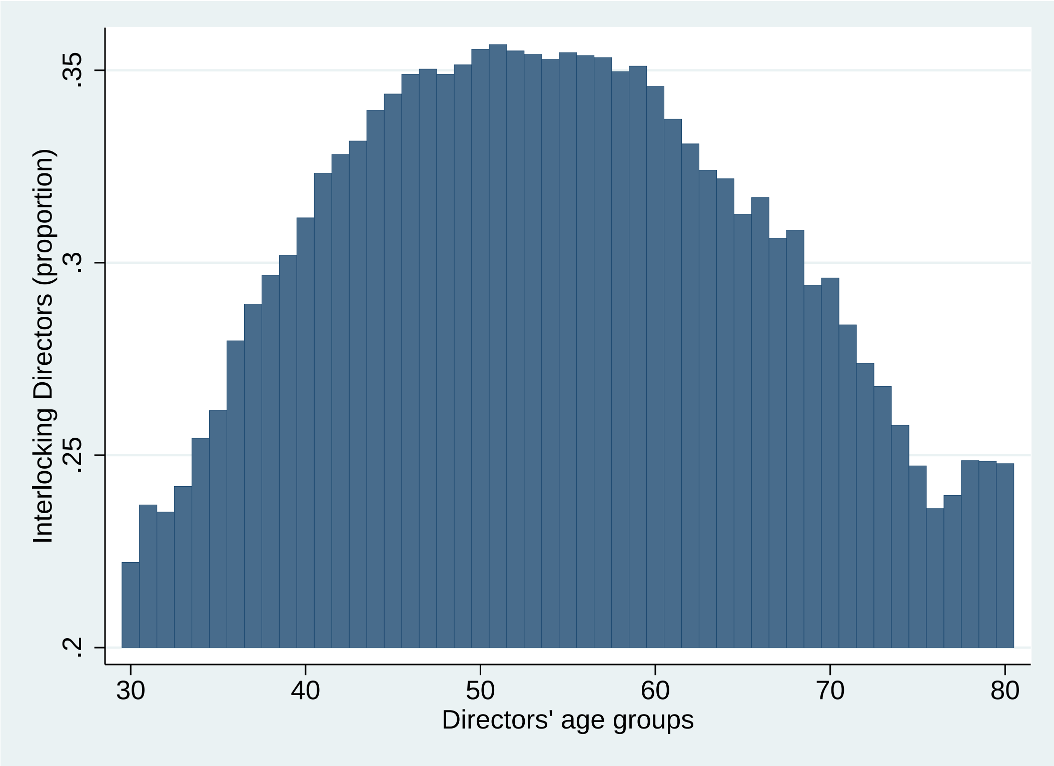


Figure 3 shows the network of firms connected by interlocking directors in 1998. Each circle in the figure represents one company. Interlocked companies are clustered together or they are connected by black lines. The round shape of the graph is produced by the Fruchterman-Reingold algorithm that optimises the position of the nodes in the space.[23] Different colors are associated to firms with different number of patent applications in 1998.

[23]This graph has been created using the package igraph available for R.

Figure 4: Inteterlocks and directors' age



Notes: Each bar represents the proportion of directors of each age that serve in more than one company.

Table 9: Interlocks and citations

	Not interlocked	Interlocked
Not citing	57,383	1,988
Citing	7,337	405
Column Total	64,720	2,393

Notes: The table reports the number of firm couples / year observations retained in the estimation sample. The model is estimated on an unbalanced panel including on average 4,500 firm couples per year over the period 1998-2012.

Theoretical Appendix

Table 10: Payoffs to firms, π_m, and directors, U_m, at outcomes $m \in \{A_1, A_2, B, C_1, C_2, D, E, F\}$

m	Firm Profit (π_m)	Director Utility (U_m)
A_1	$\left\{(1+\theta)r\tau\left(\frac{1-x}{1+\theta}\right) - P - f\right\}(1 - w(x) - v)$	$\left[(1+\theta)r\tau\left(\frac{1-x}{1+\theta}\right) - P - f\right](w(x) + v)$
A_2	$\pi_E - f(1-w)$	$\left[2r\tau\left(\frac{1}{2}\right) - P - f\right]w$
B	$\pi_F - f(1-w)$	$(r\tau(1) - f)w$
C_1	$\pi_E - L(1-w)$	$\left[2r\tau\left(\frac{1}{2}\right) - P - L\right]w$
C_2	$\pi_F - (L+P)(1-w)$	$[r\tau(1) - P - L]w$
D	$\pi_F - P(1-w)$	$(r\tau(1) - P)w$
E	$\left[2r\tau\left(\frac{1}{2}\right) - P\right](1-w)$	$\left[2r\tau\left(\frac{1}{2}\right) - P\right]w$
F	$r\tau(1)(1-w)$	$r\tau(1)w$

Table 11: A categorisation of the outcomes of the subgame $\Gamma(P_{NI})$ according to the relative size of P, L and $p(e)$

Panel A: Expected number of patents < 2)	
Case 1: High patent cost ($P > 2L$)	
$\nabla < \Delta_1$	yields a mixed strategy Nash equilibrium in which firms patent with probability γ_L and litigate under an overlap and $E(n) = 2\gamma_L \in (0, 2)$.
Case 2: Intermediate patent cost gains high ($2(1-p(e))L < P < 2L$)	
$2L < \nabla < \Delta_1$	yields a mixed strategy Nash equilibrium in which both firms patent with probability γ_L and litigate under an overlap and $E(n) = 2\gamma_L \in (0, 2)$.
Case 3: Intermediate patent cost gains low ($2(1-p(e))L < P < 2L$)	
$\nabla < 2L$	yields a mixed strategy Nash equlibrium in which firms patent with probability γ_N and they do not litigate and $E(n) = 2\gamma_N \in (0, 2)$.
Case 4: Low patent cost ($P < 2(1-p(e))L$)	
$\nabla < \Delta_2$	yields a mixed strategy Nash equilibrium in which firms patent with probability γ_N and they do not litigate and $E(n) = 2\gamma_N \in (0, 2)$.
Panel B: Expected number of patents = 2)	
High patent cost ($P > 2L$)	
$\nabla > \Delta_1$	yields a unique pure strategy Nash equilibrium in which both firms patent and litigate under an overlap with $n = 2$.
Intermediate patent cost ($2(1-p(e))L < P < 2L$)	
$\nabla > \Delta_1$	yields a unique pure strategy Nash equilibrium in which both firms patent and litigate under an overlap with $n = 2$.
Low patent cost ($P < 2(1-p(e))L$)	
$\nabla > 2L$	yields a unique pure strategy Nash equilibrium in which both firms patent and litigate under an overlap with $n = 2$.
$\Delta_2 < \nabla < 2L$	yields a unique pure strategy Nash equilibrium in which both firms patent and not litigate with $n = 2$.

Proof to Lemma 2. First, that $E(\pi_{PL}) > \pi_F$ implies $\nabla > \Delta_1$ follows from straightforward manipulation of Eq. (9) and π_F in Table 10. By hypothesis $E(\pi_N) \geq \pi_{NL}$ hence by Lemma 1, in the case of both firms patenting, litigation ensues. Hence Table 8(a) is the relevant matrix for this game. Under the hypothesis of (i) $E(\pi_{PL}) > \pi_F$. Given $\pi_E > E(\pi_{PL})$, it follows that this game has a unique asymmetric pure strategy Nash equilibria with both firms patenting and earning expected profit $E(\pi_{PL})$. Under the hypothesis of (ii) $E(\pi_{PL}) < \pi_F$. Since $\pi_E > \pi_F$ it follows that this game has a mixed strategy Nash equilibria. The expected profit for firm i playing patent with probability γ_i is given by $E(\pi_i(\gamma_i)) = \gamma_i\gamma_j E(\pi_L) + \gamma_i(1-\gamma_j)\pi_E + (1-\gamma_i)\pi_F$. Differentiating with respect to γ_i, setting equal to zero and solving, recognising symmetry, $\gamma_L = \gamma_i = \gamma_j$:

$$\gamma_L = \frac{\pi_E - \pi_F}{\pi_E - E(\pi_{PL})} = \frac{2r\tau\left(\frac{1}{2}\right) - r\tau(1) - P}{\frac{p(e)}{2}\left[2r\tau\left(\frac{1}{2}\right) - r\tau(1) + 2L\right]}$$

The expected number of patents in this subgame is then $2\gamma_L^2 + 2\gamma_L(1-\gamma_L) + 0(1-\gamma_L)^2 = 2\gamma_L$. Expected firm profit is given by:

$$E(\pi(\gamma_L)) = \frac{1}{2}\gamma_L^2 p(e)(\pi_{C_1} + \pi_{C_2}) + \gamma_L(1 - p(e)\gamma_L)\pi_E + (1-\gamma_L)\pi_F = \pi_F$$

□

Proof to Lemma 3. The proof is as for Lemma 2 replacing $E(\pi_{PL})$ with $E(\pi_{PN})$, with Table 8(b) being the relevant matrix for this game and in the case of both firms patenting there is no litigation.

The expected number of patents in this subgame is then $\gamma_N = \frac{2r\tau\left(\frac{1}{2}\right) - r\tau(1) - P}{rp(e)\left[2\tau\left(\frac{1}{2}\right) - \tau(1)\right]}$

□

Proof to Proposition 1. By Corollary 3 under the conditions of cases 1-4 expected firm profit and director utility in the Nash equilibrium of subgame $\Gamma(P_{NI})$ are, respectively, π_F and U_F, and the associated expected number of new patents is $2\gamma_L$ for cases 1 and 2 and $2\gamma_N$ for cases 3 and 4. If in addition, by hypothesis, $E(\pi_{IP}) > \pi_F$, then by Lemma 4 patenting is the unique Nash equilibrium of the subgame $\Gamma(P_I)$ since $E(\pi_{IP}) > \pi_F > \pi_B$, and hence under interlocking firms will opt to patent generating 2 new patents. Hence under the conditions of cases 1-4 the relevant choice for directors and firms is between an expected profit (utility) of π_F (U_F) under no interlock and $E(\pi_{IP})$ ($E(U_{IP})$) under interlocking, with the former exceeding the latter by hypothesis, interlocking and patenting is the unique pure stragegy Nash equilibrium. □

References

Bebchuk, Lucian Arye and Fried, Jesse (2003). Executive compensation as an agency problem. *Journal of Economic Perspectives*, **17**(3), 71–92.

Bramoullé, Yann, Djebbari, Habiba, and Fortin, Bernard (2009, May). Identification of peer effects through social networks. *Journal of Econometrics*, **150**(1), 41–55.

Brickley, James A., Linck, James S., and Coles, Jeffrey L. (1999, June). What happens to CEOs after they retire? New evidence on career concerns, horizon problems, and CEO incentives. *Journal of Financial Economics*, **52**(3), 341–377.

Cohen, W.M. and Levinthal, D.A. (1989). Innovation and learning: the two faces of R&D. *The Economic Journal*, **99**(397), 569–596.

Combined Code (2012). The Combined Code on Corporate Governance. Technical report, Financial Reporting Council, London.

Conyon, Martin J and Read, Laura E (2006). A model of the supply of executives for outside directorships. *Journal of Corporate Finance*, **12**(3), 645–659.

Crépon, B., Duguet, E., and Mairesse, J. (1998). Research, innovation, and productivity: an econometric analysis at the firm level. *Economics of Innovation and New Technology*, **7**(3), 115–158.

Croci, Ettore and Grassi, Rosanna (2013). The economic effect of interlocking directorates in Italy: new evidence using centrality measures. *Computational and Mathematical Organization Theory*, 1–24.

DiMaggio, Paul J. and Powell, Walter W. (1983). The iron cage revisited: institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, **48**(1), 147–160.

Dooley, Peter C (1969). The interlocking directorate. *The American Economic Review*, 314–323.

Gemser, G. and Wijnberg, N.M. (1995). Horizontal networks, appropriability conditions and industry life cycles. *Journal of Industry Studies*, **2**(2), 129–140.

Griliches, Zvi and Mairesse, Jacques (1984, April). Productivity and R&D at the Firm Level. In *R & D, Patents, and Productivity*, NBER Chapters, pp. 339–374. National Bureau of Economic Research, Inc.

Heger, Diana and Zaby, Alexandra K. (2013, July). The heterogeneous costs of disclosure and the propensity to patent. *Oxford Economic Papers*, **65**(3), 630–652.

Helmers, Christian, Manasa, Patnam, and Raghavendra, Rau P. (2013, August). Do board interlocks increase innovation? Evidence from natural experiments in India . Mimeo.

Helmers, Christian, Patnam, Manasa, and Rau, P Raghavendra (2014). Do board interlocks increase innovation? evidence from natural experiments in india. *Working Paper*.

Helmers, Christian and Rogers, Mark (2011, April). Intellectual Property at the Firm-Level in the UK: The Oxford Firm-Level Intellectual Property Database. Economics Series Working Papers 546, University of Oxford, Department of Economics.

König, D. Michael and Battiston, Stefano (2009). From graph theory to models of economic networks. a tutorial. In *Networks, Topology and Dynamics* (ed. A. Naimzada, S. Stefani, and A. Torriero), Volume 613, Lecture Notes in Economics and Mathematical Systems, pp. 23–63. Springer Berlin Heidelberg.

Lanjouw, Jean O and Schankerman, Mark (2001, Spring). Characteristics of Patent Litigation: A Window on Competition. *RAND Journal of Economics*, **32**(1), 129–51.

Lemley, Mark A. and Shapiro, Carl (2005, Spring). Probabilistic Patents. *Journal of Economic Perspectives*, **19**(2), 75–98.

Manski, Charles F (1993, July). Identification of Endogenous Social Effects: The Reflection Problem. *Review of Economic Studies*, **60**(3), 531–42.

Manso, Gustavo (2011, October). Motivating Innovation. *Journal of Finance*, **66**(5), 1823–1860.

Mizruchi, Mark S. (1996). What do interlocks do? An analysis, critique, and assessment of research on interlocking directorates. *Annual Review of Sociology*, **22**, pp. 271–298.

Monks, Robert and Minow, Nell (2011). *Corporate Governance, Fifth Edition.* Blackwell Publishing.

Mowery, D.C. (1990). The development of industrial research in u.s. manufacturing. *The American Economic Review*, **80**(2), 345–349.

Oerlemans, L., Meeus, M., and Boekema, F. (1998). Do networks matter for innovation? the usefulness of the economic network approach in analysing innovation. *Tijdschrift voor Economische en Sociale Gcografie*, **89**(3), 298–309.

Pennings, J M (1980). *Interlocking Directorates.* Jossey-Bass.

Powell, W.W. (1998). Learning from collaboration: knowledge and networks in the biotechnology and pharmaceutical industries. *California Management Review*, **40**(3), 228–240.

Schoorman, F David, Bazerman, Max H, and Atkin, Robert S (1981). Interlocking directorates: A strategy for reducing environmental uncertainty. *Academy of Management Review*, **6**(2), 243–251.

Schwartz, Michael (1987). *The Structure of power in America: the corporate elite as a ruling class.* Homles and Meier.