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
Originally published: October 2014
Updated: October 2018

www.sheffield.ac.uk/economics

Interlocking directorships and patenting coordination*

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

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. We develop a theoretical model that identifies interlocking directorships as a mechanism for resolving property rights conflicts. Our empirical analysis suggests a strong relationship between interlocking and patenting behaviour and finds that interlocking leads to a higher number of successful patent applications, particularly for those firms located in technology-intensive industries.

JEL classification: O31; O32; D85; G30; J49
Keywords: patents; director networks; patent coordination

1 Introduction

The allocation of the majority of resources in a market economy is entrusted to the wisdom of a small number of individuals who sit as directors on company boards. An important factor in the concentration of decision rights is the phenomenon of directors interlocking; that is, one director at one company can sit on the board of another institution and often at multiple institutions. Because these individuals typically share cultural and educational backgrounds (Mizruchi, 1996), this 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. 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 his or 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). In this paper, we investigate a relatively unexplored dimension in this literature, namely, how interlocking directors can have an impact on the patenting activity of firms.

†Addresses: Bank of England, Threadneedle Street, London, EC2R 8AH, United Kingdom (M. Bernini); Department of Economics, University of Sheffield, 9 Mappin Street, Sheffield, S1 4DT, United Kingdom (G. Efthyvoulou, I. Gregory-Smith, J. McHardy, A. Navas)

*The authors acknowledge financial support from the University of Sheffield's Departmental Research Investment Project fund. The authors wish to thank Sarah Brown and Maria Gil Molto for helpful comments and suggestions, and Christian Helmers, Mark Rogers and Philipp Schautschick for providing us with access to the Oxford Firm-Level Intellectual Property Database. The authors are also indebted to the participants of the 2014 Conference of the Work, Pensions and Labour Economics Study Group, the 14^{th} Conference on Research on Economic Theory and Econometrics, the TILEC Conference on Competition, Standardization, and Innovation, and a seminar at the University of Sheffield. The usual disclaimer applies.

In our theoretical model, when firms choose to patent they face uncertainty regarding whether they are actually going to enjoy the monopoly profits associated with the patent. This is because there is a possibility that their technology overlaps with a competing firm which could lead to an intellectual property (IP) rights conflict. The uncertainty regarding whether the firm is going to win in the case of conflict, with entitlement to use the patent, may deter investment in patenting. In this context, interlocking directors emerge as a solution that coordinates actions across firms in such a way that both firms are able to enjoy a part of the innovation rents. This increases patenting and indirectly innovation.

In this paper, we create a theoretical model that identifies conditions under which interlocking and patenting is the Nash equilibrium *and* the interlock results in an increase in the expected number of patents. We then test the main prediction of the theoretical model. To do so, we first construct a database with all director connections (interlocks) among UK listed companies using data from FAME over the period 1998-2012, and then merge this database with data on patenting activity obtained from PATSTAT. We obtain empirical support for our main theoretical result which is that interlocking increases patenting.

The intuition for our theoretical model arises out of a rich literature that claims that networks are an important source of 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) together with a literature that identifies uncertainty 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). This uncertainty is reinforced by the fact that involuntary patent infringement seems to be a frequent phenomenon. In a survey collected from IP managers, Cockburn and Henderson (2003) reveal that around just one third of their respondents conduct a prior art search before they start a new R&D project. Further, in the US, Cotripia and Lemley (2009) find that only a small proportion of defendants involved in cases of patent infringement have actually copied the patented technology, whilst Bessen and Meurer (2008) show that most of the defendants in cases of patent litigation are inadvertent infringers.

Additionally, while empirical evidence on IP conflicts and litigation is rare, there is survey evidence suggesting that IP conflicts and disputes are a common concern in industries where technology is frequently patented, but that only a fraction of these conflicts ever make it to court. For example, among Small and Medium Enterprises (SMEs) in the UK, Greenhalgh et al. (2010) find that around 40% of the patent holding firms in their survey had been involved in an IP dispute over a five year period and yet only 13% of the disputes ended in court. This would suggest intermediation between firms prior to litigation, consistent with the notion of interlocking directors.

The literature also suggests that firms use interlocking directorships as a way to reduce their operational and environmental uncertainty (Schoorman et al., 1981; Mizruchi, 1996). While, innovation decisions are usually the responsibility of specialist managers within the firm, board directors are expected to supervise strategic decisions involving innovation (Helmers et al., 2017). In particular, industry publications document that innovation is part of the governance responsibility of the board of directors. As a result, board directors, including outside non-executive directors, are found to shape the innovation strategy of the firm. As argued by Deschamps (2013), in companies for which innovation is critical, innovation effectiveness is added to the list of the board's auditing missions. Furthermore, the board can influence the company's innovation by reviewing the performance of the CEO and the top management team, by managing innovation risk, and by choosing a CEO with an innovation focus. Moreover, Oh and Barker (2015) find that the number of interlocking directors has a positive impact on R&D expenditures at the firm level. It is therefore reasonable to propose that interlocking directors could impact patenting also.[1]

[1] A high profile motivating case is that of Eric Schmidt, the CEO of Google who served as a director on Apple's board until August 2009. Such interlocking arrangements between Silicon Valley companies are not uncommon. So long as Schmidt was on the board of Apple, there was no litigation between these firms. However, following antitrust concerns

To the best of our knowledge this is the first paper that explores theoretically the formation of director networks and its impact on innovation. Our empirical analysis confirms that interlocks have a positive effect on patenting holds for UK firms. This is consistent with Helmers et al. (2017) who exploit exogenous changes in India's corporate governance framework and patent system to explore the relationship between interlocking and patenting activity. However, our results[2] also suggest that firms were close in the technological space prior to interlocking, which lead us to advance an alternative explanation to the one offered by Helmers et al. (2017). 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.

Our theoretical and empirical 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.[3] 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.

The paper is structured as follows. In section 2, we present our theoretical model which motivates our empirical work. Section 3 describes the data and presents a descriptive analysis before we outline our empirical strategy in section 4 and present the empirical results in section 5. Section 6 concludes.

2 Theoretical model

The intuition for our theoretical model can be understood within a framework where interlocking directorships emerge as a solution to secure intellectual property rights. According to the literature on patenting cited above, 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, the outcome of which is uncertain to the firm *ex-ante*. We identify conditions under which this uncertainty generates a sub-optimal number of patents and where interlocking directorships increase the number of patents.

and amid rumours that Schmidt's relationship with Steve Jobs had broken down, Schmidt resigned from the board of Apple. For the next four years, the two companies and various subsidiaries were involved in several patent infringement cases (see FT.com, 2009-08-03).

[2]We also find evidence that interlocked companies tend to converge in the technological classes under which their patents are classified and are more likely to cite each other around the time of interlocking. The corresponding results are reported in Appendix B1 and B2. Note that patenting occurs in the later stages of the innovative process. Therefore, since we observe firms citing each other and converging in the technological space already at the moment of interlocking we argue it is more likely that it was the technological proximity that motivated the interlock rather than interlocking increasing technological proximity through knowledge exchange.

[3]Presently, regulators in Europe and the US 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.

2.1 Model set up

Suppose there are two risk-neutral profit-maximising firms F_i, $i \in \{1, 2\}$ each with a risk-neutral utility-maximising director D_i. We model the technological distance between two firms as the Euclidean distance S: $\varepsilon = ||\boldsymbol{\rho_1} - \boldsymbol{\rho_2}||$, 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 a *UPR* but it is not a sufficient condition if another firm patents the same technology or one sufficiently close in the technology space. Each point in the technology space has a baseline rent r which can only be exploited if the firm has a *UPR* over it.[4] To allow the director to contribute to the value of the firm, we let the rent on a *UPR* technology attributed to firm i to be expanded by the firm's director according to:

$$r_i = r\tau(t) \tag{1}$$

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. For simplicity, consider the unit *UPR* to offer a continuum of symmetric opportunities for rent expansion on a unit interval. The available time t is then optimally devoted equally to each point in this unit interval, expanding it according to $\tau(\frac{t}{1})$. Director D_i has a maximum time allocation $T = 1$ which is supplied inelastically to the firm.[5] In exchange, D_i obtains a share $w \in (0, 1)$ of the profit of 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, and by what means, to try to establish a *UPR* over its new technology in order to extract the associated rent. We assume that there is an ex-ante probability $p(\varepsilon) \in [0, 1]$ that the two new technologies have an overlap, where $p'(\varepsilon) \leq 0$ (i.e., the probability of overlap in the new technologies is declining in the Euclidean distance of the firms' original technologies).[6] 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 identify the payoffs in each of the possible scenarios. If a firm is successful in attaining a *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 a *UPR* and all the symmetric opportunities within each. Hence, modifying Eq. (1), with two *UPR* technologies, director D_i expands the rent for firm i with time t according to:

$$r_i = 2r\tau\left(\frac{1}{2}\right) \tag{2}$$

Note, in line with earlier reasoning, with two *UPR*s we now have rent expansion opportunities along an interval of length 2, where the time devoted per unit length is $\frac{1}{2}$.

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

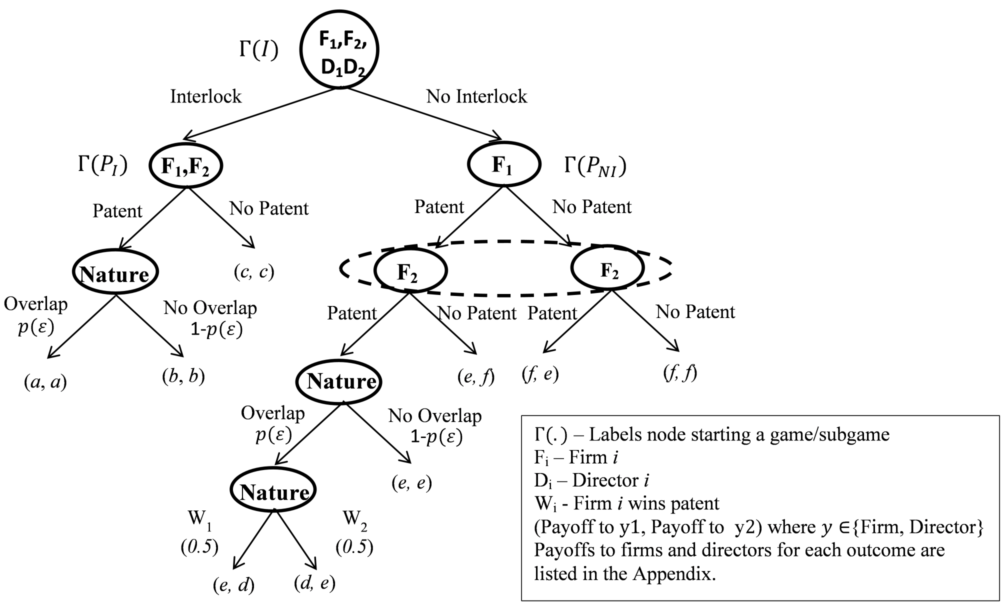
[5] A more general model that considers an elastic labor supply will not alter the main predictions of the model as what matters here is how the director distributes working time across the different activities within the firm.

[6] It could be that overlap of the innovations in the technology space are due to the innovations being complements rather than substitutes. We are grateful to Georg von Graevenitz for pointing this out to us.

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 a *UPR* on 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 $p(\varepsilon)$, as defined above. If there is no overlap, both firms have a *UPR* on their new technologies and can extract the associated rents according to Eq. (2). In the case of an overlap, the firms do not have a *UPR* on their new technologies and so cannot extract rent from them and they enter into conflict, which only one firm can win. At this stage, for simplicity, and since the firms are ex-ante symmetric, we assume that the firms obtain the *UPR* on their patent with probability $\frac{1}{2}$.[7]

The full game is characterised by Figure 1. The possible outcomes in terms of the number of new patents is $\eta \in \{0, 1, 2\}$, and in terms of the expected number of new patents is $E(\eta) \in [0, 2]$. We solve the model by backward induction. In the next subsection we start by analysing the firms' decisions to patent when the firms are not engaging in interlocking.

Figure 1: Game Tree



2.1.1 The No-Interlocking Patent Subgame $\Gamma(P_{NI})$

We begin by observing that, under the conditions of the model outlined so far, the profit of a firm with no interlocking when it patents and obtains a *UPR* over the new technology, and its profit when, instead, the new technology is not patented and exploited, are respectively:[8]

$$\pi_e = \left[2r\tau\left(\frac{1}{2}\right) - P\right](1-w), \quad \pi_f = r\tau(1)(1-w) \tag{3}$$

To characterise the possible equilibria in this subgame we also require an expression for the expected

[7] In reality firms tend to invest some resources to improve the probability of obtaining the patent. Notice that the symmetry of firms guarantees that in the case in which firms have this investment option, both firms will invest the same amount of resources and our results regarding when interlocking happens will be reinforced by including this possibility more formally. We abstract from this problem to simplify the analysis.

[8] Table A.1 in Appendix A reports the full set of payoffs for firms and directors.

payoff to a firm under no interlocking where both firms opt to patent which takes into account the possibility that the rival firm's technology may or may not be overlapping and that in the former case there might be conflict over the *UPR*. We denote this expected profit, $E(\pi_{PNI})$, where:

$$E(\pi_{PNI}) = p(\varepsilon)\left[\frac{\pi_e}{2} + \frac{\pi_d}{2}\right] + (1 - p(\varepsilon))\pi_e \tag{4}$$

Note, the term [.] in Eq. (4) is the expected profit of obtaining the patent under the existence of an overlap given both firms patent and there is no interlocking.

In the patent subgame the firm faces the subgame described in Table 1.

Table 1: Subgame $\Gamma(P_{NI})$: No Interlocking, Patent (P) versus No Patent (NP) subgame

	P_2	NP_2
P_1	$(E(\pi_{PNI}), E(\pi_{PNI}))$	(π_e, π_f)
NP_1	(π_f, π_e)	(π_f, π_f)

Payoffs to: (F_1, F_2).

The subgame described above could exhibit different Nash equilibria depending on the parameter configuration. However, there are two particular parameter sets which yield uninteresting Nash equilibria in the context of this paper, which we now seek to eliminate.

Assumption 1. *The cost of a patent, P, lies in the interval $\underline{P} < P < \overline{P}$, where:*

$$\underline{P} \equiv \frac{1}{2}\left[2r\tau\left(\frac{1}{2}\right) - r\tau(1)\right], \quad \overline{P} \equiv 2r\tau\left(\frac{1}{2}\right) - r\tau(1) \tag{5}$$

It is straightforward to show that under Assumption 1 we rule out the patent cost being so low (high) that the pure strategy Nash equilibrium for this subgame is for both firms to patent (not patent) even if it is guaranteed that the technologies will overlap i.e. $p(\varepsilon) = 1$ (will not overlap, i.e. $p(\varepsilon) = 0$).[9]

A description of the relevant equilibria for the No Interlocking subgame is provided in the following lemma:

Lemma 1. [10] *Under Assumption 1: (i) if $E(\pi_{PNI}) \geq \pi_f$, which requires $p(\varepsilon) \leq \frac{2\left(2r\tau\left(\frac{1}{2}\right) - r\tau(1) - P\right)}{2r\tau\left(\frac{1}{2}\right) - r\tau(1)}$, then there is a unique pure strategy Nash equilibrium (weak in the case of the equality) in which both firms patent with the number of new patents, $\eta = 2$, and expected firm profit is $E(\pi_{PNI})$; (ii) if $E(\pi_{PNI}) < \pi_f$, which requires $p(\varepsilon) > \frac{2\left(2r\tau\left(\frac{1}{2}\right) - r\tau(1) - P\right)}{2r\tau\left(\frac{1}{2}\right) - r\tau(1)}$, then there is a symmetric mixed strategy Nash equilibrium in which firms patent with probability, $\gamma \in (0, 1)$. The expected number of new patents is strictly less than 2, $E(\eta) = 2\gamma \in (0, 2)$, and the expected firm profit is π_f.*

Hence, from Lemma 1 the existence of uncertainty generates a situation in which both firms do not necessarily patent. In particular, if the probability of overlap is sufficiently large, the expected number of patents is strictly less than 2. The higher the probability of overlapping, the higher the likelihood the industry will be in an equilibrium with lower expected profit and a lower expected number of patents. The following subsection shows how interlocking directorships may offer an effective solution increasing the number of patents to 2.

[9]These cases are clearly uninteresting in the context of the paper since there is no scope for interlocking to increase the number of patents. In the case of $P \leq \underline{P}$, then $E(\pi_{PNI}(p(\varepsilon) = 1)) \geq \pi_f$, and the maximum number of patents is always achieved without interlocking and with $P \geq \overline{P}$, then $\pi_e \leq \pi_f$, and patenting is too expensive to be viable under any scheme.

[10]See Appendix A for a formal proof.

2.2 Introducing Interlocking

In the model (see Figure 1), firms have the possibility of interlocking in the first stage of the game - before any commitment to patenting is undertaken. For interlocking to happen, we assume that it must be incentive compatible for both directors and firms.[11] Under an interlock agreement, each firm incurs an organisational overhead, h.

The interlocked directors have a stake in both firms and 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.[12,13] 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 technologies is preserved despite the overlap and oversee the allocation of the property rights such that firms get equal shares.[14] Each director 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.[15] 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.[16]

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.

2.2.1 The Interlocking Subgame $\Gamma(P_I)$

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

$$E(\pi_{PI}) = p(\varepsilon)\pi_a + (1 - p(\varepsilon))\pi_b, \quad \pi_c = \pi_f - h(1 - w)$$

where π_a is the profit under patenting when an overlap exists and π_b is the associated profit when an overlap does not exist.

[11] 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).

[12] 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.

[13] 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.

[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. We consider the alternative mechanism in a further paper (in progress).

[15] In line with earlier reasoning, we now have a continuum of symmetric rent expansion opportunities along an interval of length $1 + \theta$, with time per unit length available given by $\frac{1-x}{1+\theta}$.

[16] 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.

Lemma 2. *The Nash equilibrium of the subgame $\Gamma(P_I)$ is for the firms to patent if $E(\pi_{PI}) > \pi_c$, and hence:*

$$p(\varepsilon)(1-w(x)-v)\left[(1+\theta)r\tau\left(\frac{1-x}{1+\theta}\right)-P-h\right]+(1-p(\varepsilon))(1-w)\left[2r\tau\left(\frac{1}{2}\right)-P-h\right] > (1-w)\left[r\tau(1)-h\right] \quad (6)$$

2.2.2 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 for interlocking to result. Hence, we now introduce the expected utility for the directors under interlocking and patenting:

$$E(U_{PI}) = p(\varepsilon)U_a + (1-p(\varepsilon))U_b \quad (7)$$

The following Lemma sets out an important condition for later analysis.

Lemma 3. *Expected utility for the directors under interlocking and patenting, $E(U_{PI})$, is greater than under no interlocking and no patenting, U_f, if:*

$$p(\varepsilon)(w(x)+v)\left[(1+\theta)r\tau\left(\frac{1-x}{1+\theta}\right)-h-P\right]+w\left\{\left[2r\tau\left(\frac{1}{2}\right)-h-P\right](1-p(\varepsilon))-r\tau(1)\right\} > 0 \quad (8)$$

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 game. More precisely, we arrive at the following proposition:

Proposition 1. *If $E(\pi_{PI}) > \pi_c$, $E(U_{PI}) > U_f$ and $E(\pi_{PNI}) < \pi_f$ 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 $2(1-\gamma) > 0$ compared to a situation in which interlocking was not an available option.*

In situations satisfying the conditions in Proposition 1, firms interlock and the expected number of patents increases after interlocking. We test this proposition in the empirical section below.

2.3 Discussion

The theoretical model above has shown that a number of new patents below the maximum of 2 can emerge from ambiguities over the rights on intellectual property. When interlocking is incentive compatible for both firms and directors (when Lemmas 2 and 3 hold), and condition (ii) in Lemma 1 also holds, interlocking and patenting is the subgame Nash equilibrium of the game by Proposition 1, resulting in an increase in the number of new patents to 2.

Lemmas 2 and 3 are more likely to hold when the size of the resolved overlapping property rights, θ, is higher, when the director time-cost of interlocking, x, is smaller, when the fixed cost of interlocking, h, is smaller and when $\tau(.)$ is more concave. We might expect these results to also be more likely as the firms are closer in the technology space so that $p(\varepsilon)$, the risk of overlap, is larger. However, although $p(\varepsilon)$ needs to be sufficiently high for the conditions of Lemma 1 (ii) to hold (under which non interlocking results in strictly less than 2 new patents) it is not generally the case that higher $p(\varepsilon)$ supports the conditions in Lemmas 2 and 3, as the direction of impact of an increase in $p(\varepsilon)$ depends on other parameters of the model.[17]

From an empirical perspective, one might suggest that the above conditions promoting Proposition 1 are more likely to be present in more technology intensive industries. In technology intensive

[17] A formal proof of the above is available upon request.

industries, we might also expect to see that the rent expansion (i.e. $\tau(.)$) is quite concave, or in another words, there are large decreasing returns associated with directors' time within each innovation. It could further be argued that directors in technology intensive industries might quickly earn experience in dealing with patenting activity resulting in lower time costs (lower x) and greater impact (higher θ). Hence, Lemmas 2 and 3 might be more likely to hold in innovation intensive industries. We will explore this possibility in the empirical part of the paper. On the other hand, condition (ii) in Lemma 1 does not depend upon θ, x, or h, but it does require that $\tau(.)$ is not too concave.

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 and exploited. However, we have also seen that introducing the option of interlocking directors to help disentangle property rights ambiguities can be incentive compatible and restore full patenting. The central result of our theoretical model is that interlocking can lead to an increase in the patenting activities of interlocked companies. Further, we suggest that the conditions supporting the hypothesis of Proposition 1 lend themselves potentially more favourably towards firms characterised by high rather than low innovation intensity. Before moving to test this prediction in the UK context, we first describe the key features of our data and network measures, and then present some descriptive analysis on the relationship between interlocking and patenting.

3 Data and network measures

Our main data source is the EPO PATSTAT database.[18] This database provides bibliographic information for all patents published by the major IP offices. An important feature of PATSTAT is that it identifies 'patent families' (groups of applications referring to the same invention) so as to allow a more precise mapping from the number of applications to the number of distinct inventions.

Data on UK firms' directors are obtained from FAME 2013.[19] We limit our analysis to manufacturing companies, as 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 data set to a manageable size and allows us to run the analysis on a personal computer, even if the construction of network measures is computationally expensive. Although this sample cannot be considered a perfect representation of the entire population of the UK manufacturing firms, it is arguably less skewed towards larger units than those used in most of the interlocking literature, which is generally focused only on listed companies (e.g., 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.[20] 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 that implemented by Helmers et al. (2011) (henceforth HRS) to match previous versions of these two datasets.[21]

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

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

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

[21] Our match is very similar to that 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 and thus are excluded from the HRS sample.

3.1 Interlocking directorships across UK Companies

Interlocking directorships occur when non-executive directors sit on the boards of multiple companies. In the terminology of Social Network Analysis the matched list of companies and directors can be defined as an 'edgelist' of a '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 to 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 its connections outside its business group ($DGND_t$), and the number of its connections with firms that operate in the same 4-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 2 shows that, for our sample, the proportion of interlocked firms increases over time, rising 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. 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 2: 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 column reports the average proportion of connections with firms in the same industry over total connections across firms.

3.2 Interlocking directorships and patenting

Our first measure of innovative output is the number of patent applications $APPS_{it}$ filed by company i at time t. Although we retain all applications irrespectively of the receiving authority, we avoid double counting by considering all documents belonging to the same patent family as a unique application.[22] We also construct an indicator of firm patent stock $STOCK_{it}$ 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):

$$STOCK_{it} = (1 - \delta) \times STOCK_{i,t-1} + APPS_{it} \tag{9}$$

Table 3 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 it takes 18 months, on average, for an eligible application to translate 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 2 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 3: Patenting activity in the sample

Year	Num. of firms	Ratio $APPS_{it} > 0$	Ratio $STOCK_{it} > 0$
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).

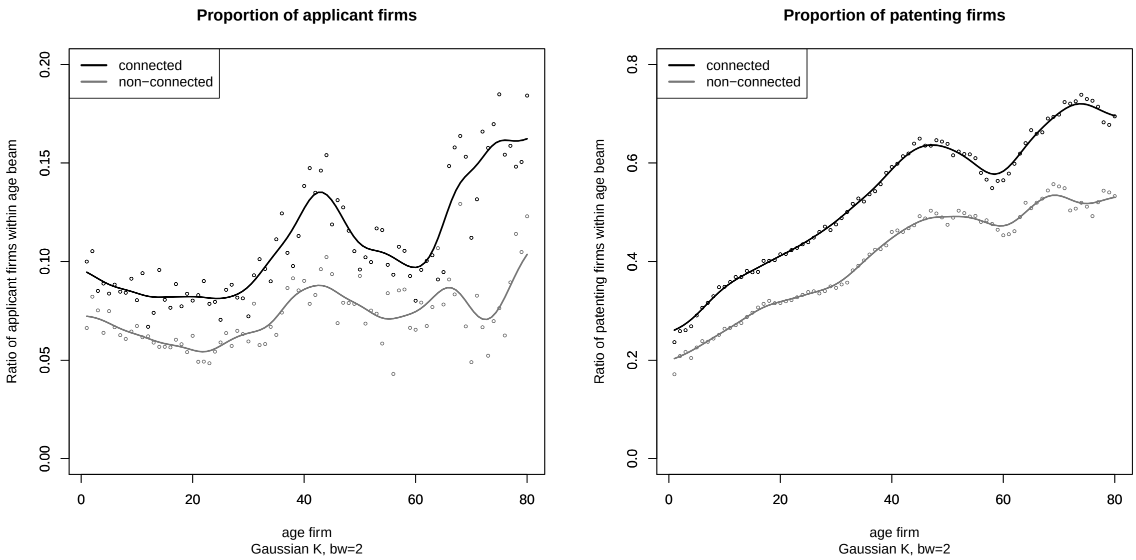
A second piece of evidence supporting a relationship between interlocks and innovative activity emerges when we graph the network of interlocked firms (see Figure B.1 in Appendix B3). We find that patent applicants are often connected with other applicants. The endogeneous formation of interlocks is a possible explanation for this pattern, whereby innovators tend to interlock with other innovators.

[22] 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.

However, this pattern is also consistent with the presence of peer effect among connected firms, where the innovative behaviour of one company affects the innovative output of the others.

Finally, in order to investigate the existence of technological spillovers between interlocked firms, we employ data on patent citations and the technological composition of firms' patent portfolios, and perform two empirical exercises; see Appendices B1 and B2 for a full discussion. The results obtained reveal that interlocked companies are more likely to cite each other, especially around the time of interlocking, and tend to exhibit a high degree of technological similarity of their patent portfolio in the immediate period following their first interlock.

Figure 2: Patenting over firm life-cycle



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

4 Econometric framework

We now proceed to test the main proposition of our theoretical model. Our instrumental variables (IV) strategy aims at identifying: (i) the impact of interlocked directorships on firm innovation; and, (ii) 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. However, the network of interlocked firms is likely to evolve endogenously with respect to firms' innovative strategies and their accumulated knowledge. If companies' 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. 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. To address our research question, we consider the following reduced form equation:

$$Y_{ist} = \gamma_1 C_{i,t-1} + \gamma_2 \Big(I(P_{i,t-1} \neq \emptyset) * \bar{Y}_{j \in P_{i,t-1}} \Big) + X'_{it}\mu + \delta_s + \delta_t + \beta_1 \bar{Y}_{gt} + \beta_2 \bar{Y}_{st} + \beta_3 \bar{Y}_{ct} + u_i + \epsilon_{it} \quad (10)$$

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 2-digit NACE industry and year effects, respectively. 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 directorships, but 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 the firm's connectedness through interlocking directorships. 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 the firm i's connections.

The parameters of interest are γ_1 and γ_2. The first measures the direct effect of connectedness on Y_{ist}, whereas the second measures 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 firm innovation. Lastly, u_i captures unobserved firm-level fixed effects and ϵ_{it} is an individual firm error term.

The most serious identification problem raised by Eq. 10 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. Indeed, our theoretical model suggests firms may seek connections with those firms who regularly patent closely related technologies in order to avoid property right conflicts. 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 of applying for a patent. To do so, we treat Eq. 10 as a linear probability model and estimate it by 2SLS, where the dependent variable is $DAPPS_{it}$ (a dummy variable assuming value one when the firm applies for at least one patent), and the main variable of interest, $C_{i,t-1}$, is captured by $add_{i,t-1}$; that is, a dummy variable assuming 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 $Retire_{i,t-1}$ and $Hire_{i,t-1}$, representing the ratio of retiring and newly hired directors, respectively, 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 from hiring a director that is already sitting in the board of another company. It is thus unrelated with connections arising from the external hiring of a serving director, which is more likely to arise when the company signals its knowledge stock through patenting. As shown in Table 4, while the two instruments are highly correlated with the endogenous variable $add_{i,t-1}$, the correlation coefficients with the dependent variable $DAPPS_{it}$ and the innovation outcomes of firms in the same business group, sector and county (as captured by the $\bar{Y}$ terms) have very small values.

Figure 2 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 specification further below is meant to capture this long-run effect of connectedness on innovative output. In this version, we introduce the firm's patent stock $STOCK_{it}$ as the dependent variable, and the lagged number of connections $nd_{i,t-1}$ as the main variable of interest (i.e., $C_{i,t-1}$ in Eq. 10). Because $nd_{i,t-1}$ is more persistent over time than $add_{i,t-1}$, we now employ two instruments that are more suitable to reflect this aspect of the

Table 4: Pairwise correlation matrix

	$DAPP_{it}$	$add_{i,t-1}$	$Retire_{i,t-1}$	$Hire_{i,t-1}$	$\bar{Y}_{st}$	$\bar{Y}_{gt}$	$\bar{Y}_{ct}$
$DAPP_{it}$	1.0000						
$add_{i,t-1}$	0.0363	1.0000					
$Retire_{i,t-1}$	0.0411	0.2061	1.0000				
$Hire_{i,t-1}$	0.0207	0.2665	0.3089	1.0000			
$\bar{Y}_{st}$	0.0899	0.0199	0.0445	0.0372	1.0000		
$\bar{Y}_{gt}$	0.0632	0.0399	0.0496	0.0291	0.1311	1.0000	
$\bar{Y}_{ct}$	0.0698	0.0110	0.0307	0.0124	0.1098	0.0655	1.0000
	$STOCK_{it}$	$nd_{i,t-1}$	$JuniorDir_{i,t-1}$	$SeniorDir_{i,t-1}$	$\bar{Y}_{st}$	$\bar{Y}_{gt}$	$\bar{Y}_{ct}$
$STOCK_{it}$	1.0000						
$nd_{i,t-1}$	0.0721	1.0000					
$JuniorDir_{i,t-1}$	-0.0255	-0.0592	1.0000				
$SeniorDir_{i,t-1}$	-0.0244	-0.0822	-0.1486	1.0000			
$\bar{Y}_{st}$	0.1299	0.0201	-0.0065	-0.0373	1.0000		
$\bar{Y}_{gt}$	0.1304	0.0705	-0.0137	-0.0388	0.1311	1.0000	
$\bar{Y}_{ct}$	0.0632	0.0007	0.0103	-0.0318	0.1098	0.0655	1.0000

endogeneous regressor: $JuniorDir_{i,t-1}$ and $SeniorDir_{i,t-1}$. These two variables represent the ratio of directors younger than 40 and the ratio of those older than 60, respectively, over the total number of directors sitting on 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 on multiple boards.[23] This relationship is most likely determined by the evolution of a director's reputation and connections over their career, and by their occupational choice in later life. 'Junior directors' may lack sufficient experience and reputation to be invited to sit on other companies' boards, while 'senior directors' may choose to reduce the time spent sitting in board meetings as they approach retirement.

The identifying assumption is that the proportion of junior and senior directors on the board affects a firm's patenting activity only indirectly through the likelihood of interlocks. Some studies raise the point that a CEO's incentive to promote innovation may change over their tenure within a company and more generally over their career (e.g., Brickley et al., 1999; Manso, 2011). However, it is very unlikely that the age of the CEO (and their 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 activity, while it determines the likelihood of interlocks. As also shown in Table 4, the correlation coefficients between the instruments $JuniorDir_{i,t-1}$ and $SeniorDir_{i,t-1}$ and the endogenous variable $nd_{i,t-1}$ are about three times larger, in absolute value, than the correlation coefficients between the two instruments and the dependent variable $STOCK_{it}$.

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 $\bar{Y}_{ct}$ 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 exclude this term from the model when we run regressions on the pooled sample of connected and unconnected companies. This is equivalent 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, we first estimate the restricted specification of Eq. 10, where 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), and then estimate the unrestricted specification of Eq. 10

[23] The relationship between these instruments and the number of firm interlocks is evident when we look at Figure B.2 in Appendix B3.

on the sub-sample of firms with at least one connection (i.e., $nd_{it} \geq 1$). The latter exercise allows us to identify the average effect of connectedness on the innovative outcome of connected firms after controlling for partners' heterogeneous innovative performances.

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', which makes it impossible to identify the effect of peers' behaviour on individual behaviour 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 instrumenting the average number of patent applications among firm i's connections with the average patent stock of the firm's second degree connections; that is, the connections of firm i's connections that are not directly interlocked with firm i. The identification assumption underpinning this strategy is that the characteristics of second-degree connections of firm i affect the firm's outcome only through their impact on the outcome of its first-degree connections[24].

5 Results

We first comment on the results obtained by estimating Eq. 10 on the sample that includes both interlocked and non-interlocked companies. The first four columns of Table 5 report the estimates of regressions of the dummy $add_{i,t-1}$ (taking value one if the firm added at least one new connection at time $t-1$) on the dummy $DAPP_{it}$ (taking value one when the firm applies for at least one patent). The remaining four columns report estimates from regressing the lagged number of interlocked connections $nd_{i,t-1}$ 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 a firm's age $log(age)_{it}$, industry and year fixed effects) and a 'long' specification with additional controls[25]. Specifically, in the 'long' specification, we also control for a firm's lagged capital intensity computed as the log of fixed assets over the number of employees $CapInt_{i,t-1}$, its independence status $Indep_{it}$, its lagged size proxied by the log of the number of employees $log(empl)_{i,t-1}$, the number of directors in its board $BoardSize_{it}$, and the average number of patent applications at the 2-digit NACE industry level $\bar{Y}_{st}$, the business group level $\bar{Y}_{gt}$, and the UK county level $\bar{Y}_{ct}$.

The estimated coefficient of $add_{i,t-1}$ in the second stage regression on $DAPP_{it}$ suggests that, adding at least one new connection in the previous period increases the probability of applying for a new patent by 14 percentage points (on average). This effect is reduced to 6 percentage points once we control for other firm characteristics and spillover effects. 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 a firm's connections and the size of their patent stocks. The estimated coefficient of $nd_{i,t-1}$ suggests that each connection increases a firm's number of patents by 0.6 (on average), and this result does not change once we estimate the 'long' specification of the model [26].

[24]This is similar in spirit to the estimating strategy proposed by Bramoullé et al. (2009). Ideally, we would obtain an estimating equation driven by exogenous variation in characteristics of the firm's peers. However, in the absence of other instruments our best strategy is to instrument first degree connections with second degree connections.

[25]We do not have information on R&D expenditure, a potentially relevant variable for patenting activity. However, to the extent that R&D expenditure is correlated with firm size, our estimates are unlikely to be severely biased from this omission

[26]OLS estimates (available on request) return coefficients of the same sign but smaller in magnitude in comparison to the IV results. This suggests the direction of selection bias is downwards.

Table 5: Connectedness and patents

Dependent :	$DAPP_{it}$				$STOCK_{it}$			
Specification :	(a)		(b)		(a)		(b)	
Estim.Stage :	2nd Stage $DAPP_{it}$	1st Stage $add_{i,t-1}$	2nd Stage $DAPP_{it}$	1st Stage $add_{i,t-1}$	2nd Stage $STOCK_{it}$	1st Stage $nd_{i,t-1}$	2nd Stage $STOCK_{it}$	1st Stage $nd_{i,t-1}$
$add_{i,t-1}$	0.140***		0.063***					
	(0.010)		(0.017)					
$nd_{i,t-1}$					0.604***		0.586***	
					(0.087)		(0.218)	
$log(age)_{it}$	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)
$CapInt_{i,t-1}$			0.007***	-0.002***			0.290***	-0.006
			(0.001)	(0.001)			(0.051)	(0.015)
$Indep_{it}$			-0.017***	-0.024***			0.105	-0.599***
			(0.004)	(0.003)			(0.157)	(0.070)
$log(empl)_{i,t-1}$			0.029***	0.013***			0.759***	0.313***
			(0.002)	(0.001)			(0.119)	(0.021)
$BoardSize_{it}$			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								
$Retire_{i,t-1}$		0.211***		0.231***				
		(0.006)		(0.011)				
$Hire_{i,t-1}$		0.358***		0.327***				
		(0.006)		(0.012)				
$JuniorDir_{i,t-1}$						-0.660***		-0.545***
						(0.036)		(0.085)
$SeniorDir_{i,t-1}$						-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 $DAPP_{it}$ and $STOCK_{it}$. For each model we estimate a 'short' specification including only the log of the firm's age $log(age)_{it}$ as a control variable, and a 'long' specification including the following set of firm-level controls: $CapInt_{i,t-1}$ is the log of the firm's capital per employee at time $t-1$, $Indep_{it}$ is a dummy for independent firms, $log(empl)_{i,t-1}$ is the log of the firm's size proxied by the number of employees at time $t-1$, $BoardSize_{it}$ is the number of directors on a company's board. $\bar{Y}_{st}$, $\bar{Y}_{gt}$, $\bar{Y}_{ct}$ capture the average of the dependent variable across firms belonging to the same 2-digit NACE industry, the same business group and the same county, respectively. The set of excluded instruments include either $Hire_{it}$ and $Retire_{it}$ in models on $DAPP_{it}$, and $JuniorDir_{it}$ and $SeniorDir_{it}$ in models on $STOCK_{it}$. $JuniorDir_{it}$ and $SeniorDir_{it}$ are respectively the ratio of directors under 40 years of age or above 60 on a company's board. $Hire_{it}$ and $Retire_{it}$ are respectively the ratio of retiring directors or newly hired directors on a company's board. Cluster robust standard errors are reported in parentheses (cluster unit: firm). Significance levels: *.1, **.05,***.01.

The Angrist-Pischke (AP) F statistics from first stage regressions on $add_{i,t-1}$ and $nd_{i,t-1}$ reject the null hypothesis of weak instruments, while the Hansen J statistics do not reject the hypothesis that our instruments are uncorrelated with the second stage errors.[27] First stage coefficients on $Retire_{i,t-1}$ and $Hire_{i,t-1}$ confirm that the turnover in the board of directors affects positively the interlocking probability. As we expect, the proportion of younger and older directors in the board is negatively correlated with the number of interlocks. Therefore, there is sufficient statistical support to claim that we correctly identify the positive impact of connectedness on the number of successful patent applications filed by a company.

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 6. Once we include this term, we find that the coefficient on $add_{i,t-1}$ in the regressions on $DAPP_{it}$ is reduced to 0.05 and 0.03 respectively in the 'short' and 'long' specifications of the model. In addition, the coefficient of $nd_{i,t-1}$ on $STOCK_{it}$ is rendered insignificant. These results suggest that, the impact of connectedness on a firm's patenting behaviour

[27]To provide further support for this, we estimate the first stage equations by random effects and then augment the specifications with the dependent variable of the second stage equations. As shown in Table B.4 (Appendix B3), the dependent variable (both in the current period and the previous period) enters the specifications statistically insignificantly and leaves the estimates on the instruments virtually unchanged. This suggests that the error terms of the first stage regressions are uncorrelated with the innovative output.

Table 6: Connectedness and peer effects

	$(DAPP_{it}\|nd_{it} > 0)$		$(STOCK_{it}\|nd_{it} > 0)$	
	(1)	(2)	(3)	(4)
$add_{i,t-1}$	0.050***	0.032**		
	(0.010)	(0.016)		
$nd_{i,t-1}$			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)
$log(age)_{it}$	0.020***	0.013***	0.494***	0.383***
	(0.002)	(0.003)	(0.081)	(0.109)
$CapInt_{i,t-1}$		0.008***		0.284***
		(0.002)		(0.062)
$log(empl)_{i,t-1}$		0.029***		0.905***
		(0.003)		(0.129)
$Indep_{it}$		-0.026***		-0.281
		(0.008)		(0.278)
$BoardSize_{it}$		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 ($add_{i,t-1}$ or $nd_{i,t-1}$)	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 $DAPP_{it}$ and $STOCK_{it}$ for firms with at least one interlock $nd_t > 0$. The set of excluded instruments include $Hire_{it}$, $Retire_{it}$ and $\bar{Y}_{j\in P_{i,t-2}}$ in models on $DAPP_{it}$, and $JuniorDir_{it}$, $SeniorDir_{it}$ and $\bar{Y}_{j\in P_{i,t-2}}$ in models on $STOCK_{it}$. The instrument $\bar{Y}_{j\in P_{i,t-2}}$ is the average number of patent applications (in models on $DAPP_{it}$) or the average number of patents in a firm's portfolio (in models on $STOCK_{it}$), computed across the second degree connections of the company. Significance levels: *.1, **.05,***.01. See also notes for Table 5.

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. Qualitatively, the 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 percentage points. In the long run, this effect has an important impact on connected firms' patent stocks, as we find that, if peers' patent intensity (i.e., the average number of patents held by the peers of a company) increases by one patent, the patent stock of the company increases, on average, by 0.6. This result explains the divergence in the patent stocks of connected and unconnected companies as they age over time, as observed in Figure 2.

Our theoretical model also suggests that the interlocking-patenting relation is stronger in firms where technology is more important in their business. To examine this hypothesis, we restrict the sample to include only the most technology-intensive industries[28] and estimate the same regression set-up. The results, displayed in Table B.5 (Appendix B3), indicate that when we focus on these industries (which constitute one-third of our sample), the impact of interlocking on patenting activities is much more pronounced, especially in the regressions on $STOCK_{it}$. Specifically, the estimated coefficient on $nd_{i,t-1}$ suggests that each connection increases a firm's number of patents by almost 2; that is, three times more than for the full sample of industries.

[28]Following BIS (2011), we include the following industries: (i) chemicals and chemical products; (ii) basic pharmaceutical products and pharmaceutical preparations; (iii) computer, electronic and optical products; (iv) electrical equipment; (v) other manufacturing (musical instruments, medical and dental instruments and supplies, sports goods, games and toys, etc).

6 Conclusion

This paper provides new insights into the role of interlocking directors for patenting activity. In particular, it contributes to the literature in two main aspects. First, we develop a formal framework that identifies interlocking directorships as a mechanism for resolving property rights conflicts that arise between innovating firms. In particular, we argue that interlocking directors can prevent such conflicts by allocating appropriate time resource to the interlocked companies. Second, we use data from about 70,000 firms in the UK over the period 1998-2012 to investigate the impact of connectedness and peer effects on patent applications.

Consistent with our main theoretical proposition, we find that adding at least one new connection increases the probability of applying for a patent in the next year by up to 14 percentage points. In addition, the impact of connectedness on a firm's patenting behaviour appears to be conditional on the patenting activity of the firm's connections: a rise in peer patenting intensity by one application increases the probability of applying for a patent in the next year by 3 percentage points. These results are stronger in technology intensive industries.

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. Of course, we do not claim that patent coordination is the only reason why firms interlock and there may be more sinister forces at play. Nevertheless, adopting a more positive stance than present with respect to interlocking could have a positive impact on innovation and reduce welfare reducing patent wars. Indeed, the patenting 'thicket' literature has found that some firms respond to property rights conflicts by flooding patenting authorities with suspect applications (von Graevenitz et al., 2013). 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. Another implication arising from patents thickets is that our main empirical result, interlocking directors increase patenting, could be viewed as a conservative lower bound estimate. As our theoretical model argues that interlocking directors facilitate coordination, marginal thicket-like patents might diminish under interlocking. However, our current model abstracts from thickets and hence we leave testing this proposition for future research[29].

[29]We are thankful to an anonymous referee for highlighting the implications of patent thickets for our theoretical and empirical results

A. Theoretical Appendix

Table A.1: Payoffs to firms, π_m, and directors, U_m, at outcomes $m \in \{a, b, c, d, e, f\}$

m	Firm Profit (π_m)	Director Utility (U_m)
a	$\left\{(1+\theta)r\tau\left(\frac{1-x}{1+\theta}\right) - P - h\right\}(1 - w(x) - v)$	$\left[(1+\theta)r\tau\left(\frac{1-x}{1+\theta}\right) - P - h\right](w(x) + v)$
b	$\pi_e - h(1-w)$	$\left[2r\tau\left(\frac{1}{2}\right) - P - h\right]w$
c	$\pi_f - h(1-w)$	$(r\tau(1) - h)w$
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	$r\tau(1)(1-w)$	$r\tau(1)w$

Proof to Lemma 1. (i) In the case of $E(\pi_{PNI}) \geq \pi_f$, since, by Assumption 1, $\pi_e > E(\pi_{PNI})$, it follows that this game has a (weak in the case of the equality) unique symmetric pure strategy Nash equilibrium with both firms patenting, $\eta = 2$, and earning expected profit $E(\pi_{PNI})$. (ii) In the case of $E(\pi_{PNI}) > \pi_f$, substituting using Eqs. (3) and (4), implies the following must hold:

$$p(\varepsilon) > \frac{2\left(2r\tau\left(\frac{1}{2}\right) - r\tau(1) - P\right)}{2r\tau\left(\frac{1}{2}\right) - r\tau(1)} \tag{A.i}$$

Under Assumption 1, $E(\pi_{PNI}) < \pi_f$ yields a game with a mixed strategy Nash equilibria. The expected profit for firm i playing patent with probability γ_i is given by $E(\pi_i(\gamma_i)) = \gamma_i\gamma_j E(\pi_{PNI}) + \gamma_i(1-\gamma_j)\pi_e + (1-\gamma_i)\pi_f$. Differentiating with respect to γ_i, setting equal to zero and solving, recognising symmetry, $\gamma = \gamma_i = \gamma_j$, yields:

$$\gamma = \frac{\pi_e - \pi_f}{\pi_e - E(\pi_{PNI})} = \frac{2r\tau\left(\frac{1}{2}\right) - r\tau(1) - P}{\frac{p(\varepsilon)}{2}\left[2r\tau\left(\frac{1}{2}\right) - r\tau(1)\right]} \tag{A.ii}$$

Notice that under Assumption 1 and applying the inequality Eq. (A.i) in Eq. (A.ii), γ must lie in the open interval: $\gamma \in (0, 1)$.

Next, the expected number of patents in this subgame is then:

$$E(\eta) = 2\gamma^2 + 2\gamma(1-\gamma) + 0(1-\gamma)^2 = 2\gamma$$

whereupon, given $\gamma \in (0, 1)$, then $E(\eta) \in (0, 2)$. Hence, the expected number of patents is strictly less than 2.

Finally, expected firm profit is given by:

$$E(\pi(\gamma)) = \gamma^2\left(E(\pi_{PNI} - \pi_e)\right) + \gamma(\pi_e - \pi_f) + \pi_f = \pi_f$$

□

B. Empirical Appendix

In Appendices B1 and B2, we consider the relationship between interlocks and patent citations and subsequently the path of technological convergence of firms' patent portfolios. Our main motivation behind this exercise is to offer more evidence to support our theoretical model about patent coordination, from an alternative explanation based on technological spillovers as a source for interlocking and subsequent increase in innovation activities. Since we do not have a very clean source of identification for these exercises, we have placed this material in the appendix. Nevertheless, we feel that these exercises are informative and are, at least, descriptively consistent with our theoretical framework where interlocking is used as a device to coordinate patenting behaviour.

Appendix B1: 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 firms that are not interlocked with the target company, but are sufficiently similar to 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 a random draw from the population of firms that may cite each other. Instead, estimation on the population of all possible firm couples is not feasible 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}:

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

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} and couple-specific characteristics X'_{ij}, and includes a set of dummies that, for couples of interlocked companies, assume value one in the period of their first connection and in each one of the four periods preceding or following their interlock. These dummies assume value 0 in all periods for couples of firms that do not interlock over the same period.

If interlocking directorships facilitate 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 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 the citation probability, as we only exploit the timing of citation for identification.

Table B.1 reports the corresponding estimates. Standard errors of the estimated coefficients are relatively large due to the small number of interlocks in the sample (see Table B.2). The low precision of the point estimates suggests a qualitative interpretation of the results. The three columns of Table B.1 report the coefficients obtained by estimating the model on the whole sample, on the sample

Table B.1: 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)
$lock_{ijt}$	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)
$PatCount_{it}$	0.195***	0.307***	0.189***
	(0.007)	(0.030)	(0.007)
$PatStock_{jt}$	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)
$Log(age)_{it}$	-0.054***	-0.079**	-0.052***
	(0.008)	(0.039)	(0.008)
$Log(age)_{jt}$	-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 the results of random effects probit estimation where the panel unit is set at the level of each firm 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 ($lock_{ijt}$) or in each of the four years before and after the interlock. These dummies assume 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.

Table B.2: 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.

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.[30] In the model we also control for the number of patents in the patent stock of the cited company $PatStock_{jt}$, the number of applications of the citing company $PatCount_{it}$, the age of the two firms, and their 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 $lock_{ijt}$, a set of dummies $lock_{ij,t-s}$ assuming value one in one of the s periods

[30] To compare the coefficients obtained on the three samples within the same specification, we also estimate a model including interaction terms between the variables of interest and a dummy assuming 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 B.1 and are available upon request.

preceding the interlock, and a set of dummies $lock_{ij,t+s}$ assuming value one in one of the s periods following the interlock.[31]

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 different business groups, we find that the previous results are confirmed only for citations occurring between firms belonging to different groups. We interpret this evidence as supporting the hypothesis that interlocks serve as information channels or coordination mechanisms across firms only in the absence of other organizational linkages between them.

Appendix B2: Interlocks and technological convergence

We now turn to investigate the path of technological convergence of firms' patent portfolios as an alternative strategy to capture technological spillovers between interlocked firms. To construct a time varying index of technological similarity of the patent portfolio of interlocked firms, we exploit the IPC technological classification of patents reported in PATSTAT. The 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:

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

where c is an index for IPC technological classes and I is the set of all classes observed in PATSTAT (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.[32] We then estimate the following model:

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

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 difference-in-differences model. On the contrary, identification relies on the comparison of the PSI of 'treated' firms before and after receiving the treatment (that is, 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 is estimated using a random-effect tobit model to deal with the large number of 0 values in the distribution of the PSI.

Table B.3 reports the corresponding results. 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 $lock_{ijt}$ and for later periods are positive and significant at the 1% level, and suggest that the PSI increases monotonically starting from the year before the interlock. Similarly to what we found

[31]We report the specification including four dummies for periods preceding and following the interlock. Running the regressions including longer or shorter timing structures around the interlock produces very similar results.

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

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; that is, two years after the occurrence of the first interlock.

Table B.3: Technological convergence

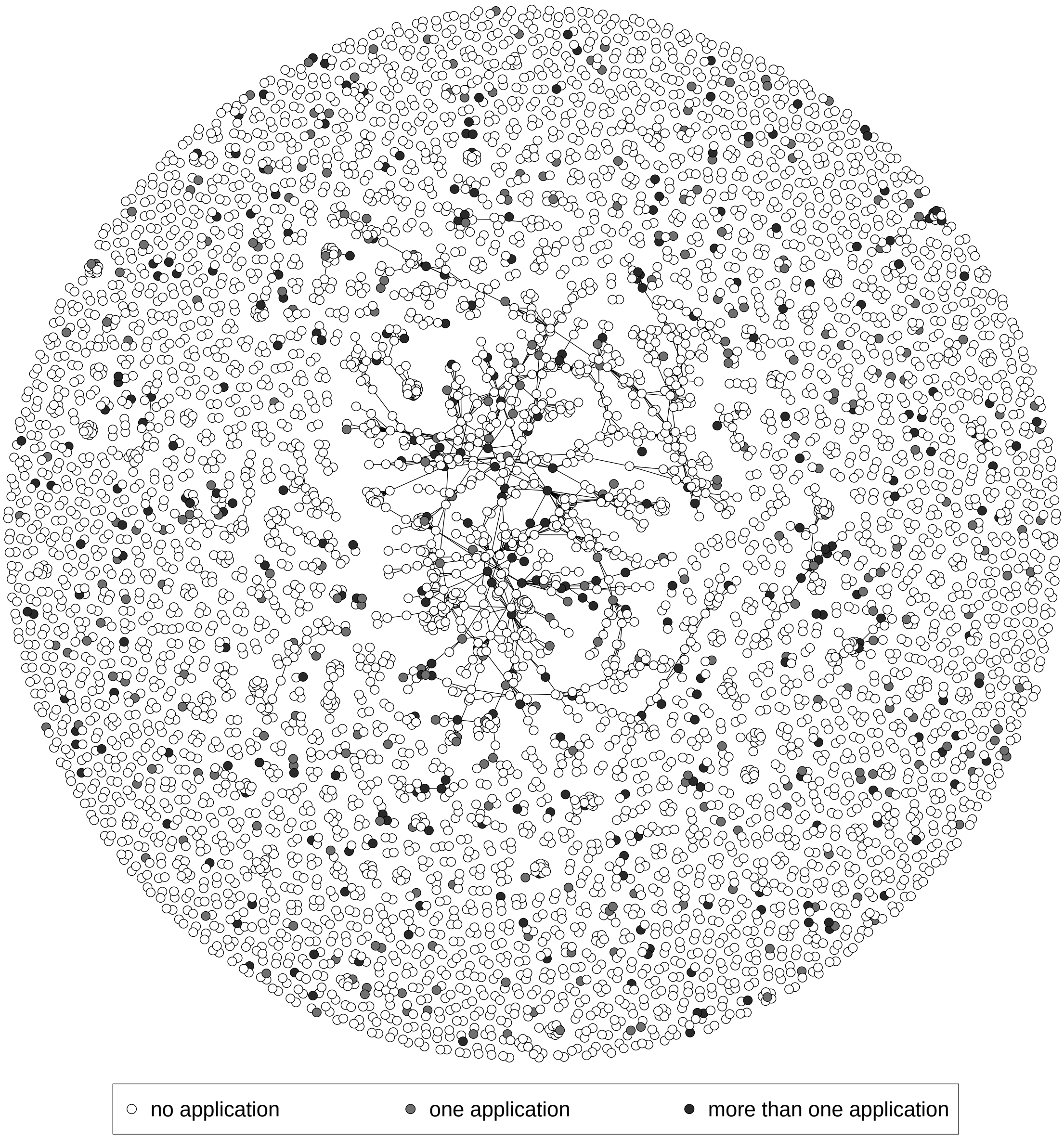
	Whole sample (1)	Same business group (2)	Different business groups (3)
$lock_{ij,t-4}$	-0.013***	-0.007	-0.016***
	(0.003)	(0.006)	(0.004)
$lock_{ij,t-3}$	-0.005	0.005	-0.012***
	(0.003)	(0.005)	(0.004)
$lock_{ij,t-2}$	0.001	0.004	-0.005
	(0.003)	(0.005)	(0.003)
$lock_{ij,t-1}$	0.009***	0.009*	0.003
	(0.003)	(0.005)	(0.003)
$lock_{ijt}$	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 effects 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.

Appendix B3: Tables and Figures

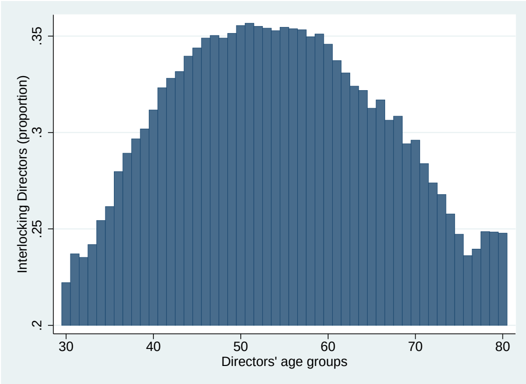
Figure B.1 shows the network of firms connected by interlocking directors in 1998. Each circle in the figure represents one company. Interlocked companies are clustered together or they are connected by black lines. The round shape of the graph is produced by the Fruchterman-Reingold algorithm that optimises the position of the nodes in the space.[33] Different colours are associated to firms with different number of patent applications in 1998.

Figure B.1: Patent applications across interlocked firms in 1998



[33]This graph has been created using the package igraph available on R.

Figure B.2: Interlocks and directors' age



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

Table B.4: Connectedness and instruments

$Dependent:$	add_{it}			nd_{it}		
	(1)	(2)	(3)	(4)	(5)	(6)
$Retire_{it}$	0.161***	0.161***	0.161***			
	(0.011)	(0.011)	(0.011)			
$Hire_{it}$	0.384***	0.384***	0.384***			
	(0.012)	(0.012)	(0.012)			
$JuniorDir_{it}$				-0.153***	-0.153***	-0.154***
				(0.059)	(0.059)	(0.059)
$SeniorDir_{it}$				-0.139***	-0.139***	-0.135***
				(0.042)	(0.042)	(0.042)
$log(age)_{it}$	-0.010***	-0.010***	-0.010***	-0.088***	-0.089***	-0.055**
	(0.002)	(0.002)	(0.002)	(0.023)	(0.023)	(0.023)
$CapInt_{it}$	-0.002**	-0.002**	-0.002**	-0.012	-0.012	-0.014
	(0.001)	(0.001)	(0.001)	(0.012)	(0.012)	(0.012)
$Indep_{it}$	-0.026***	-0.026***	-0.026***	-0.796***	-0.794***	-0.799***
	(0.004)	(0.004)	(0.004)	(0.063)	(0.063)	(0.063)
$log(empl)_{it}$	0.012***	0.011***	0.011***	0.107***	0.105***	0.102***
	(0.001)	(0.001)	(0.001)	(0.017)	(0.017)	(0.018)
$BoardSize_{it}$	0.015***	0.015***	0.015***	0.222***	0.222***	0.220***
	(0.001)	(0.001)	(0.001)	(0.008)	(0.008)	(0.008)
$\bar{Y}_{st}$	-0.004	-0.004	-0.004	-0.047	-0.048	-0.051
	(0.004)	(0.004)	(0.004)	(0.033)	(0.033)	(0.032)
$\bar{Y}_{gt}$	0.002**	0.002**	0.002**	-0.003	-0.004	-0.003
	(0.001)	(0.001)	(0.001)	(0.007)	(0.007)	(0.007)
$\bar{Y}_{ct}$	-0.004	-0.004	-0.004	-0.096*	-0.097*	-0.112**
	(0.005)	(0.005)	(0.005)	(0.050)	(0.050)	(0.048)
$DAPP_{it}$		0.006				
		(0.005)				
$DAPP_{i,t-1}$			0.007			
			(0.005)			
$STOCK_{it}$					0.002	
					(0.002)	
$STOCK_{i,t-1}$						0.003
						(0.002)

Notes: The Table reports random effects estimates of the first stage equations on add_{it} and nd_{it}, before and after adding the dependent variable of the second stage equations ($DAPP_{it}$ and $STOCK_{it}$, respectively) as control. All regressions include industry and year effects. Cluster robust standard errors are reported in parentheses (cluster unit: firm). Significance levels: *.1, **.05,***.01.

Table B.5: Connectedness and patents: technology-intensive industries

$Dependent:$	$DAPP_{it}$				$STOCK_{it}$			
$Specification:$	(a)		(b)		(a)		(b)	
$Estim.Stage:$	2nd Stage $DAPP_{it}$	1st Stage $add_{i,t-1}$	2nd Stage $DAPP_{it}$	1st Stage $add_{i,t-1}$	2nd Stage $STOCK_{it}$	1st Stage $nd_{i,t-1}$	2nd Stage $STOCK_{it}$	1st Stage $nd_{i,t-1}$
$add_{i,t-1}$	0.166***		0.070**					
	(0.017)		(0.029)					
$nd_{i,t-1}$					1.945***		1.213*	
					(0.385)		(0.719)	
$log(age)_{it}$	0.017***	0.001	0.010**	-0.011***	1.208***	0.134***	1.089***	-0.033
	(0.003)	(0.002)	(0.005)	(0.003)	(0.199)	(0.030)	(0.323)	(0.048)
$CapInt_{i,t-1}$			0.012***	-0.002			0.667***	0.010
			(0.003)	(0.002)			(0.191)	(0.030)
$Indep_{it}$			-0.027**	-0.019**			0.105	-0.724***
			(0.011)	(0.008)			(0.678)	(0.109)
$log(empl)_{i,t-1}$			0.037***	0.011***			2.306***	0.208***
			(0.004)	(0.002)			(0.394)	(0.037)
$BoardSize_{it}$			0.010***	0.013***			0.182	0.186***
			(0.002)	(0.001)			(0.217)	(0.019)
$\bar{Y}_{st}$			0.032***	0.003			2.198**	0.251**
			(0.012)	(0.007)			(0.950)	(0.105)
$\bar{Y}_{gt}$			0.004	0.003**			0.312	0.052*
			(0.003)	(0.001)			(0.200)	(0.030)
$\bar{Y}_{ct}$			0.056***	-0.012			3.462***	-0.171
			(0.019)	(0.010)			(1.054)	(0.149)
Excluded Instruments								
$Retire_{i,t-1}$		0.213***		0.271***				
		(0.011)		(0.020)				
$Hire_{i,t-1}$		0.416***		0.353***				
		(0.011)		(0.021)				
$JuniorDir_{i,t-1}$						-0.510***		-0.505***
						(0.069)		(0.168)
$SeniorDir_{i,t-1}$						-1.099***		-1.082***
						(0.072)		(0.123)
Hansen J-test (p-value)	0.381		0.541		0.944		0.479	
AP F-test		2739.05		892.41		260.87		78.59
Obs.	65,681	65,681	21,189	21,189	72,151	72,151	21,360	21,360

See notes for Table 5. Technology-intensive manufacturing industries include: (i) chemicals and chemical products; (ii) basic pharmaceutical products and pharmaceutical preparations; (iii) computer, electronic and optical products; (iv) electrical equipment; (v) other manufacturing.

References

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

Bessen, J. and M. J. Meurer (2008). *Patent failure: How judges, bureaucrats and lawyers put innovators at risk*. Princeton, NJ: Princeton University Press.

BIS (2011). The distribution of innovation activity across UK industries. Technical report, Department for Business Innovation & Skills.

Bramoullé, Y., H. Djebbari, and B. Fortin (2009). Identification of peer effects through social networks. *Journal of Econometrics 150*(1), 41–55.

Brickley, J. A., J. S. Linck, and J. L. Coles (1999). 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.

Cockburn, I. and R. Henderson (2003). Survey results from the 2003 intellectual property owners association on strategic management of intellectual property. Technical report.

Cohen, W. and D. Levinthal (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, M. J. and L. E. Read (2006). A model of the supply of executives for outside directorships. *Journal of Corporate Finance 12*(3), 645–659.

Cotripia, C. A. and M. A. Lemley (2009). Copying in patent law. *North Carolina Law Review*, 1421–1466.

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

Croci, E. and R. Grassi (2013). The economic effect of interlocking directorates in Italy: new evidence using centrality measures. *Computational and Mathematical Organization Theory 20*(1), 89–112.

Deschamps, J.-P. (2013). Governing innovation in practice – the role of the board of directors. Online article in InnovationManagement.se, 2013-05-21.

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

Dooley, P. C. (1969). The interlocking directorate. *The American Economic Review 59*(3), 314–323.

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

Greenhalgh, C., J. Philips, R. Pitkethly, M. Rogers, and J. Tomalin (2010). Intellectual property enforcement in smaller UK firms. Technical report, Strategy Advisory Board for Intellectual Property Policy (SABIP).

Griliches, Z. and J. Mairesse (1984). Productivity and R&D at the Firm Level. In *R & D, Patents, and Productivity*. National Bureau of Economic Research, Inc.

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

Helmers, C., M. Patnam, and P. R. Rau (2017). Do board interlocks increase innovation? evidence from a corporate governance reform in india. *Journal of Banking & Finance 80*(Supplement C), 51 – 70.

Helmers, C., M. Rogers, and P. Schautschick (2011). Intellectual property at the firm-level in the UK: The Oxford firm-level intellectual property database. *Department of Economics Discussion Papaer Series 546*.

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

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

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

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

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

Mizruchi, M. 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, R. and N. Minow (2011). *Corporate Governance, Fifth Edition*. Blackwell Publishing.

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

Oerlemans, L., M. Meeus, and F. Boekema (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.

Oh, W.-Y. and V. L. Barker (2015). I like what i see: CEO outside directorship and strategic imitation in R&D investment. Unpublished working paper, University of Calgary.

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

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

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

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

von Graevenitz, G., S. Wagner, and D. Harhoff (2013). Incidence and growth of patent thickets: The impact of technological opportunities and complexity. *The Journal of Industrial Economics 61*(3), 521–563.