HOOVER IP²

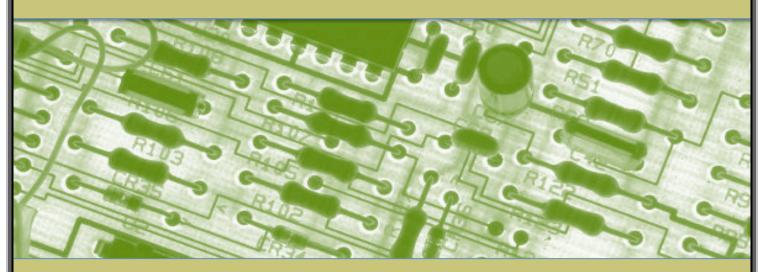## "TECHNOLOGY ENTRY IN THE PRESENCE OF PATENT THICKETS"

BRONWYN H. HALL
UC BERKELEY, NBER, IFS, AND NIESR

CHRISTIAN HELMERS
SANTA CLARA UNIVERSITY

GEORG VON GRAEVENITZ
QUEEN MARY UNIVERSITY OF LONDON, CCP, CREATE

JANUARY 5, 2017

Hoover Institution Working Group on
Intellectual Property, Innovation, and Prosperity
Stanford University

www.hooverip2.org

Technology Entry in the Presence
of Patent Thickets[*]

Bronwyn H. Hall (UC Berkeley, NBER, IFS, and NIESR)[§]

Christian Helmers (Santa Clara University)

Georg von Graevenitz (Queen Mary University of London, CCP, CREATe)

January 2017

Abstract

We analyze the effect of patent thickets on entry into technology areas by firms in the UK. We present a model that describes incentives to enter technology areas characterized by varying technological opportunity, complexity, and the potential for hold-up due to the presence of patent thickets. We show empirically that our measure of patent thickets is associated with a reduction of first time patenting in a given technology area controlling for the level of technological complexity and opportunity. Technological areas characterized by more technological complexity and opportunity, in contrast, see more entry. Our evidence indicates that patent thickets raise entry costs, which leads to less entry into technologies regardless of a firm's size.

JEL codes: O34, O31, L20, K11

Keywords: *IPR, patents, entry, technological opportunity, technological complexity, hold-up*

---

1 Introduction

The past two decades have seen an enormous increase in patent filings worldwide (Fink *et al.*, 2013). There are signs that the high level of patenting may be reducing innovation in certain technologies (Federal Trade Commission, 2011; Bessen and Meurer, 2008; Jaffe and Lerner, 2004; Federal Trade Commission, 2003). Companies drawing on these technologies face elevated legal costs of commercializing innovative products when patents that contain overlapping claims form so-called "patent thickets" (Shapiro, 2001). Patent thickets arise where individual products draw on innovations protected by hundreds or even thousands of patents, often with fuzzy boundaries. These patents belong to many independent and frequently competing firms. Patent thickets can lead to hold-up of innovations, increases in the complexity of negotiations over licenses, increases in litigation, and they create incentives to add more, often weaker patents to the patent system (Allison *et al.*, 2015). The increased transaction costs associated with patent thickets reduce profits that derive from the commercialization of innovation, and ultimately may reduce incentives to innovate.

There is a growing theoretical (Bessen and Maskin, 2009; Clark and Konrad, 2008; Farrell and Shapiro, 2008; Fershtman and Kamien, 1992) and legal literature (Chien and Lemley, 2012; Bessen *et al.*, 2011) on patent thickets. Related work analyzes firms' attempts to form patent pools to reduce hold-up (Joshi and Nerkar, 2011; Lerner *et al.*, 2007; Lerner and Tirole, 2004) and the particular challenges posed in this context by standard essential patents (Lerner and Tirole, 2013).

The existing empirical evidence on patent thickets is largely concerned with showing that they exist and measuring their density (von Graevenitz *et al.,* 2011; Ziedonis, 2004). There is less evidence on the effects patent thickets have for firms. Cockburn and MacGarvie (2011) demonstrate that patenting levels affect product market entry in the software industry. They show that a 1 per cent increase in the number of existing patents is associated with a 0.8 per cent drop in the number of product market entrants. This result echoes earlier findings by Lerner (1995) who showed for a small sample of U.S. biotech companies that first-time patenting in a given technology is affected by the presence of other companies' patents. These two papers use patent counts in narrow technological fields as a proxy for thickets; in this paper we move beyond this measure to use an indicator of the extent to which the patents actually might overlap.

Bessen and Meurer (2013) suggest that patent thickets also lead to increased litigation related to hold-up. They (and we) use the term to describe a situation where an alleged infringer faces the threat of an injunction or high licensing costs after she has sunk investment.[1] Patent thickets have remained a concern of antitrust agencies and

---

[1] "High licensing costs" refers to costs that are higher than those that would have been negotiated *ex ante* in the present of possible "invent around" before the alleged infringer sank her investment. This

regulators in the U.S. for over a decade (Federal Trade Commission, 2011, 2003; USDoJ and FTC, 2007). Reforms that address some of the factors contributing to the growth of patent thickets have recently been introduced in the U.S. (America Invents Act (AIA) of 2011)[2] and by the European Patent Office (EPO).

However, it is sometimes argued that patent thickets are a feature of rapidly developing technologies in which technological opportunities abound (Teece *et al.*, 2014). Thickets are seen as a reflection of fast technological progress that is paired with increased technological complexity (Lewis and Mott, 2013). This suggests that the increased transactions costs associated with patent thickets and the benefits of technological complexity and opportunity often coincide. There may be a trade-off between technological opportunity and growth on the one hand and increased transaction costs due to the emergence of patent thickets on the other - if the transaction costs of patenting in complex technologies are not avoidable. The challenge in assessing technologies with high levels of patenting is to develop a framework that captures the main factors that incentivize patenting and the costs and benefits thereof.

This paper contains such a framework in the form of a model that incorporates technological patenting incentives and the effects of patenting by rivals on each firm's incentives to enter and to patent. The aim of the paper is to empirically disentangle the effects of thickets and technological opportunity and complexity on entry into technology areas. Modeling patenting and entry we show that greater technological opportunity and complexity both encourage entry, while higher transaction costs associated with preexisting patent thickets reduce entry. These effects are quantified empirically using firm-level data on patenting by firms in the United Kingdom.

In a first contribution, we develop a theoretical model of entry into patenting in discrete and complex technologies. Building on previous work on patenting incentives (von Graevenitz *et al.*, 2013), we model how entry decisions are affected by the potential for hold-up due to patent thickets. The model focuses on the interaction between incumbents and new entrants through two channels: i) via legal costs associated with patent enforcement and ii) via incumbency advantages in R&D fixed costs. In contrast to previous work, we separate technological complexity and thicket density. Where previously granted patents raise thicket density, we assume that the expected legal costs of defending new patents are higher and the benefits of amassing portfolios of new patents are lower. This reflects higher potential for hold-up and leads firms to patent less and reduce entry, *ceteris paribus*. As in von Graevenitz *et al.* (2013), the model predicts that greater complexity of a technology leads to an increase of a firm's

---

possibility can arise because of either prohibitive search costs (Mulligan and Lee, 2012) or fuzzy patent boundaries or both.

[2] For further information see http://www.gpo.gov/fdsys/pkg/BILLS-112hr1249enr/pdf/BILLS-112hr1249enr.pdf

patent applications, while greater technological opportunity reduces applications.[3] The model also predicts that higher complexity and opportunity will increase entry, because competition for individual patents is reduced and the probability that entrants can establish themselves in a technology is increased. In addition, where incumbency implies that firms enjoy lower costs of R&D, incumbents patent more than entrants, which reinforces the growth of thickets.

Our second contribution consists of an empirical test of these predictions and a quantification of the importance of the conflicting effects of opportunity, complexity, and patent thickets on entry. We use patent data from both the EPO and the USPTO to construct measures of technological opportunity, technological complexity and hold-up potential and relate these to entry into new technology areas by UK firms.

To test the predictions of the model, we measure technological complexity and hold-up potential in patent thickets separately. This is an important improvement over the analysis in von Graevenitz *et al.* (2013), who conflated complexity and hold-up potential arising from existing patent portfolios. To distinguish complexity and hold-up potential, we introduce a new empirical measure of technological complexity and sharpen the definition of an existing measure of hold-up potential.[4]

Our empirical analysis of entry confirms that greater technological opportunity and complexity increase entry and that hold-up potential reduces entry substantially. While we cannot quantify the overall net welfare effect, our results indicate that patent thickets raise entry costs for large and small firms alike. To the extent that more original and radical, rather than incremental ideas come from new entrants rather than incumbents (Tushman and Anderson, 1986; Henderson, 1993), this is likely to have negative long-run consequences on innovation and product market competition. In combination with earlier results by von Graevenitz *et al.* (2013), who point to a positive correlation between patenting levels and the presence of thickets, our results suggest that any increases in transaction costs due to thickets can potentially have important dynamic effects on innovation by reducing entry.

The remainder of this paper is organized as follows. Section 2 presents a model of entry into patenting in a technology area and derives several testable predictions. Section 3 describes the data, and the empirical measurement of the key concepts in the model. Section 4 discusses our results and Section 5 provides concluding remarks.

---

[3] Opposite signs of complexity and opportunity are due to firms' efforts to maintain a constant share of patents on each opportunity, to protect their profits from patenting (von Graevenitz *et al.*, 2013).

[4] We also provide separate empirical tests of the measures we propose on the EPO patent data used by von Graevenitz *et al.* (2013) in Appendix D.

2   Theoretical Model

This section summarizes the main results of a two-stage model of entry into patenting and subsequent patenting decisions. The model shows how complexity of a technology, technological opportunity and the potential of hold-up due to patent thickets affect firms' decisions to enter into patenting in that technology.

We assume that each technology consists of a set of opportunities, each of which consists of a number of patentable facets. The more facets opportunities in a technology have, the more complex is that technology.

Greater technological opportunity attracts entrants to a technology as more avenues arise to earn a profit through application of the technology. Greater complexity of a technology also attracts entrants, because it is more likely that each entrant will be able to gain a share of the profits flowing from any one opportunity. Where multiple firms hold patents on the same opportunity, licensing negotiations or litigation ensue as firms divide the profits flowing from the opportunity. We assume that holding a larger share of patents on an opportunity is beneficial for firms in terms of licensing or litigation, but less so when thickets arise from poorly delineated patents that provide increased options to litigate. This captures the costs imposed by thickets on patentees.

The model consists of two stages: entry and patenting. We solve the model by backward induction. Analysis of the entry decision shows that greater technological opportunity and complexity increase entry into patenting, while the threat of increased legal costs arising from denser patent thickets reduces entry into patenting.

Analyzing patenting we show that denser thickets reduce patent applications, while greater technological opportunity increases applications in discrete technologies, but reduces them in complex technologies. This prediction is analogous to that in von Graevenitz *et al.* (2013). Finally, greater complexity increases patent applications.

Below we set out the notation and the main findings from the model. Most proofs and derivations are relegated to appendices A and B.

2.1   Notation and Assumptions

The key variables of the model are the complexity of a technology k, measured by F_k ($F_k \in \mathbb{R}_0^+$), the degree of technological opportunity, measured by O_k ($O_k \in \mathbb{R}_0^+$), and hold-up potential h_k. The value of all $\tilde{F}_k$ patents in an opportunity is V_k. In the simplest discrete setting this is the value of the one patent (facet) that covers each technological opportunity. In more complex technologies this is the value of controlling all patents (facets) on a technological opportunity. Firms (indexed by i) choose the number of opportunities o_i to invest in and the number of facets f_i per opportunity to patent.

In equilibrium only $\tilde{F}_k = (1 - (1 - (\hat{f}_k / F_k)^{N_0+1}))$ facets are patented, where $\hat{f}_k$ is the equilibrium number of facets chosen by applicants and N_0 is the number of firms that

have chosen a specific opportunity.[5] Since $\tilde{F}_k$ may be smaller than F_k the total value of patenting in a technology is $V(\tilde{F}_k) \le V(F_k)$. Von Graevenitz *et al.* (2013) assume that the value function $V_k(\tilde{F}_k)$ is convex in covered facets. In Appendix B.1 we discuss how this assumption is relaxed and generalize the model by introducing a concave function relating the share of patents a firm holds on an opportunity s_{ik} to the proportion of the value V_k the firm can extract through licensing and its own sales: $\Delta(s_{ik})$. This captures the benefits that a patent portfolio confers in the market for technology.

In sum, the assumptions we make on the value function and portfolio benefits are:

$$(VF): \quad V(0) = 0, \quad \frac{\partial V}{\partial \tilde{F}_k} > 0 \tag{1}$$

$$(PB): \quad \Delta(0) = 0, \quad \frac{d\Delta(s_{ik})}{ds_{ik}} > 0 \ \text{ and } \ \frac{d^2\Delta(s_{ik})}{ds_{ik}^2} < 0 \tag{2}$$

The model contains three types of patenting costs:

- Costs of R&D per opportunity, which depend on overall R&D activity in that technology area: $C_0\left(\sum_j^{N_o} o_j\right)$.

- Costs of maintaining each granted patent in force: C_a.

- Costs of coordinating R&D on *different* technological opportunities $C_c(o_i)$, where

$$\frac{\partial C_c}{\partial o_i} > 0 \tag{3}$$

These assumptions imply that R&D costs are fixed costs.[6] We allow for the endogenous determination of the level of R&D fixed costs, which rise as more opportunities are researched simultaneously by rival firms. This reflects competition for inputs into R&D, e.g. scientists and engineers that are in fixed supply in the short run (Goolsbee, 1998).

Where multiple firms own facets on an opportunity, their legal costs $L(\gamma_{ik}, s_{ik}, h_k)$ depend on the absolute number of patented facets γ_{ik}, on the share of patents per opportunity that a firm holds s_{ik}, and on the extent to which they face hold-up h_k. The first two channels capture the costs of defending a patent portfolio as the number of patents increases, while leaving scope for effects on bargaining costs that derive from the share

---

[5] The properties of N_0 are summarized in Appendix A.2.

[6] It also implies that there is no technological uncertainty. However, introducing technological uncertainty into the model does not change the main comparative statics results.

of patents owned: The hold-up parameter captures contexts in which several firms' core technologies become extremely closely intertwined. Then each firm has to simultaneously negotiate with many others to commercialize its products, which significantly raises transaction costs.

$$(LC): \quad L(\gamma_{ik}, s_{ik}, h_k), \quad \text{where} \quad \frac{\partial L}{\partial \gamma_{ik}} > 0, \frac{\partial^2 L}{\partial \gamma_{ik}^2} \geq 0, \frac{\partial L}{\partial s_{ik}} \leq 0, \frac{\partial^2 L}{\partial s_{ik}^2} \geq 0,$$

$$\frac{\partial L}{\partial h_k} > 0, \frac{\partial^2 L}{\partial \gamma_{ik} \partial h_k} > 0, \frac{\partial^2 L}{\partial s_{ik} \partial h_k} > 0 \tag{4}$$

All remaining cross partial derivatives of the legal costs function are zero.

In what follows, we use the following definitions:

$$\omega_k \equiv \frac{o_i}{O_k}, \quad \phi_k \equiv \frac{f_i}{F_k}, \quad \mu_k = \frac{\tilde{F}_k}{V(\tilde{F}_k)} \frac{\partial V(\tilde{F}_k)}{\partial \tilde{F}_k}, \quad \xi_k = \frac{s_k}{\Delta(s_k)} \frac{d\Delta(s_k)}{ds_k}, \quad \text{and} \quad \eta_k = \frac{f_i}{\tilde{F}_k} \frac{\partial \tilde{F}_k}{\partial f_i}.$$

Here ω_k is the share of opportunities each firm chooses to pursue, ϕ_k is the share of facets each firm seeks to patent per opportunity, μ_k is the elasticity of the value function *w.r.t.* the level of complexity, ξ_k is the elasticity of the benefits function Δ *w.r.t.* the share of patents each firm is granted and η_k is the elasticity of the number of covered facets *w.r.t.* the number of patent applications of each firm.

2.2   Patenting and Entry

Firm i's profits in technology k, $\pi_{ik}(o_i, f_i, F_k, O_k, N_k, h_k)$, are a function of the number of opportunities o_i which the firm invests in, the number of facets per opportunity f_i the firm seeks to patent, the total number of patentable facets per opportunity F_k, the number of technological opportunities a technology offers O_k, the number of firms entering the technology N_k, and the degree of hold-up in that technology h_k.

In this section we analyze the following two-stage game G^*:

Stage 1: Firms enter until $\pi_{ik}(o_i, f_i, F_k, O_k, N_k, h_k) = 0$;[7]

Stage 2: Firms simultaneously choose the number of opportunities, o_i, to invest in and the number of facets per opportunity f_i to patent in order to maximize profits π_{ik}.

We solve the game by backward induction and derive local comparative statics results for the symmetric extremal equilibria of the second stage game. For the subsequent analysis it is important to note that all equilibria of this second stage game are symmetric. In case that the second stage game has multiple equilibria we focus on the

---

[7] We treat N_k as a continuous variable here, which is an abstraction that simplifies our analysis.

properties of the extremal equilibria when providing comparative statics results (Milgrom and Roberts, 1994; Amir and Lambson, 2000; Vives, 2005). Equilibrium values of the firms' choices are denoted by a hat (^) and we drop the firm specific subscripts in what follows, e.g., $\hat{\phi}_k$.

At stage two of the game each firm maximizes the following objective function:

$$\pi_{ik}(o_i, f_i) = o_i\left(V(\tilde{F}_k)\Delta(s_{ik}) - L(\gamma_{ik}, s_{ik}, h_k) - C_0(\sum_j^{N_o} o_j) - f_i p_k C_a\right) - C_c(o_i) \tag{5}$$

This expression shows that per opportunity k, the firm derives profits from its share $s_{ik} \equiv p_k f_i / \tilde{F}_k$ of patented facets, while facing legal costs L to appropriate those profits, as well as costs of R&D C_0, costs of maintaining its patent portfolio C_a, and coordination costs across opportunities C_c.

2.3 Simultaneous Entry with Multiple Facets

2.3.1 Comparative statics of patenting

We show that the second stage of this game is smooth supermodular:

Proposition 1: *The second stage patenting game, defined in particular by assumptions (VF, eq. 1), (PB, eq. 2) and (LC, eq. 4) is smooth supermodular if $\mu_k > \xi_{ik}$ and if ownership of the technology is expected to be fragmented.*

This is shown in Appendix B.1.

This result generalizes Proposition 1 derived by von Graevenitz *et al.* (2013).[8] Given this result we can show that:

Proposition 2: *The potential for hold-up in complex technologies reduces patenting incentives.*

In Appendix B.2 we show that the expected legal costs of hold-up reduce the number of opportunities that firms invest in. In addition, firms with larger portfolios are more exposed to hold-up and benefit less from the share of patents they have patented per opportunity. Both effects combine to reduce the number of facets each firm applies for.

2.3.2 Comparative statics of entry

In Appendix B.3 we show that there is a free entry equilibrium. In this equilibrium the following propositions hold:

Proposition 4: *Under free entry greater complexity of a technology increases entry.*

---

[8] Here it is no longer the case that the value function has to be increasing in the number of patented facets for supermodularity of the patenting game. We relegate further discussion of this result to Appendix B.1.

In the model, complexity has countervailing effects: first, it increases profits, because it is less likely that duplicative R&D arises making each opportunity more valuable; this clearly increases incentives to enter. Next, given the level of patent applications $\hat{f}_k$, complexity reduces the probability that each facet is patented, which reduces profits and entry incentives. Finally, complexity reduces competition for each facet, which increases the probability of patenting and increases innovation incentives. Overall we show that the positive effects outweigh the negative effects and incentives for entry rise with complexity of a technology.

To derive Proposition 4, consider how equilibrium profits are affected by the complexity of the technology F_k, the degree of technological opportunity O_k, and the potential for hold-up h_k:

$$\frac{\partial \pi(\hat{o}, \hat{f})}{\partial F_k} = \hat{o}\,\frac{\hat{s}_k}{F_k}\left(\left(\hat{\varepsilon}_{\tilde{F}_k, F_k} - \hat{\varepsilon}_{p_k, F_k}\hat{\eta}_k\right)\left[V\left(\hat{\tilde{F}}_k\right)\frac{\Delta(\hat{s}_k)}{\hat{s}_k}\left(\hat{\mu}_k - \hat{\xi}_k\right) + \frac{\partial L}{\partial \hat{s}_k}\right]\right) > 0 \tag{6}$$

$$\frac{\partial \pi(\hat{o}, \hat{f})}{\partial O_k} = \hat{o}\,\frac{\partial \widehat{N}_o}{\partial O_k}\frac{\hat{s}_k}{\widehat{N}_o}\left(\left(\hat{\varepsilon}_{\tilde{F}_k, N_o} - \hat{\varepsilon}_{p_k, N_o}\hat{\eta}_k\right)\left[V\left(\hat{\tilde{F}}_k\right)\frac{\Delta(\hat{s}_k)}{\hat{s}_k}\left(\hat{\mu}_k - \hat{\xi}_k\right) + \frac{\partial L}{\partial \hat{s}_k}\right] - \frac{\partial C_o}{\partial \widehat{N}_o \hat{o}}\frac{\widehat{N}_o \hat{o}}{\partial \hat{s}_k}\right) > 0 \tag{7}$$

$$\frac{\partial \pi(\hat{o}, \hat{f})}{\partial h_k} = -\hat{o}\,\frac{\partial L}{\partial h_k} < 0 \tag{8}$$

Proposition 4 follows from the Implicit Function theorem once we know the sign of the derivative of profits w.r.t. F. Under free entry firms' profits decrease with entry:

$$\frac{\partial N_k}{\partial F_k} = -\frac{\partial \pi}{\partial F_k} \Big/ \frac{\partial \pi}{\partial N_k} \tag{9}$$

Therefore, the Implicit Function theorem implies that the sign of the effect of complexity F on entry depends on the sign of the effect of complexity on profits.

Equation (6) shows that the effect of complexity on profits depends on the difference between the elasticities $\hat{\varepsilon}_{\tilde{F}_k, F_k}$ and $\hat{\varepsilon}_{p_k, F_k}\hat{\eta}_k$, which are derived in Appendix A.3. The elasticity $\hat{\varepsilon}_{p_k, F_k}$ is (see Appendix A.1):

$$\hat{\varepsilon}_{p_k, F_k} = \widehat{N}_o^2\,\frac{\hat{\phi}_k - \frac{1}{2}\left(1 + \frac{1}{\widehat{N}_o}\right)}{1 - \hat{\phi}_k} \tag{10}$$

This elasticity is negative for $\hat{\phi}_k < \frac{1}{2}$, which is also a precondition for supermodularity of game G*. Both terms in brackets in equation (6) are positive, when game G* is

supermodular. This implies that greater complexity raises profits and this induces entry.[9]

Proposition 5: *Under free entry greater technological opportunity increases entry.*

For any given number of entrants an increase in technological opportunity reduces competition between firms for patents. This increases firms' expected profits and increases entry.

Continuing from the proof of Proposition 4 above, by the Implicit Function theorem the sign of the derivative of profits w.r.t. technological opportunity determines the effect of technological opportunity on entry:

$$\frac{\partial N_k}{\partial O_k} = -\frac{\partial \pi}{\partial O_k} \Big/ \frac{\partial \pi}{\partial N_k} \tag{11}$$

An increase in technological opportunity increases profits and entry. In Appendix B.3 we show that the term in brackets in Equation (7) is negative under free entry. Profits increase as technological opportunity increases, because fewer firms enter per opportunity.

Proposition 6: *Under free entry the potential for hold-up reduces entry.*

An increase in the potential for hold-up raises firms' expected legal costs. This reduces expected profits and lowers potential for entry.

To derive this prediction, note that by the Implicit Function theorem the sign of the derivative of profits w.r.t. the level of hold-up in a technology area determines the effect of hold-up on entry:

$$\frac{\partial N_k}{\partial h_k} = -\frac{\partial \pi}{\partial h_k} \Big/ \frac{\partial \pi}{\partial N_k} \tag{12}$$

Hence, equation (8) shows that the effect of hold-up on entry derives from the increased legal costs that the possibility of hold-up imposes on affected firms.

2.4   Entry and Incumbency

The previous section sets out a model in which all firms entered and then invested in patents. At both stages firms' decisions were simultaneous. In Appendix B.5 we extend the model to a setting in which some firms, the incumbents, face lower costs ($C_O - \Psi$,

---

[9] When $\hat{\phi}_k \geq \frac{1}{2}$ we no longer have the assumptions necessary to show supermodularity. This situation corresponds to the case where one firm has more than half the patents in a particular technology opportunity within a technology area. Thus our results may not hold when a specific opportunity is highly concentrated. In general this will not be the case, especially at our level of empirical analysis, but it would be interesting to explore this possibility in future work.

where $\Psi > 0$) of entering opportunities. This captures the fact that incumbents have previous experience of doing R&D in a technology area. We demonstrate that our main results derived above are robust to this extension of the model. We also show that incumbents will enter more new opportunities and that this will crowd out new entrants. The increased levels of entry by incumbents raise costs of entry to new firms, lowering their numbers.

2.5 Predictions of the Model

Our model predicts how the probability of entry into patenting depends on opportunity, complexity, hold-up potential and incumbents' experience. Here we summarize these predictions, which are tested empirically below:[10]

Prediction 1: *Greater technological opportunity increases the probability of entry.*

Greater technological opportunity reduces competition for facets per opportunity, which raises expected profits and thereby attracts entry.

Prediction 2: *Greater complexity of a technology increases the probability of entry.*

Greater complexity has countervailing effects: it reduces competition per facet as well as duplicative R&D, attracting entry. It also increases the likelihood that some of a technology remains unpatented, reducing its overall value and entry. Our model shows that overall complexity increases entry.

Prediction 3: *Greater potential for hold-up reduces the probability of entry.*

Hold-up potential increases expected costs of entry, reducing it.

Prediction 4: *More experienced incumbents are more likely to enter technological opportunities new to them.*

We show that incumbency advantage raises the number of opportunities that incumbents enter. This implies that they also enter new opportunities, which they have not previously been active in. This expansion of activity by incumbents crowds out entry by new entrants.

3 Data and Empirical Model

This section of the paper describes the data we use in the empirical test of our theoretical predictions. In particular, we discuss how we measure entry, how the set of potential entrants is identified, and which measures and covariates are used.

Our empirical model is a hazard rate model of firm entry into patenting in a technology

---

[10] Note that von Graevenitz *et al.* (2013) tested predictions from a substantially more restrictive version of the model on the *level* of patent applications using data from the European Patent Office.

area as a function of technological opportunity, technological complexity, and hold-up potential that characterize a technology area. Additional firm level covariates include the age and size of firms. The models we estimate are stratified at the industry level. That is, the unit of observation for each entry hazard is a firm-technology area, but the hazard shapes and levels are allowed to vary by the industry that the firm is in. This approach recognizes that patenting propensities vary across industries for reasons that may not be technological (e.g., strategic reasons, or reasons arising from the historical development of the sector).

We use a combination of firm level data for the entire population of UK firms registered with Companies House and data on patenting at the European Patent Office and at the Intellectual Property Office for the UK. The firm data come from the data held at Companies House provided by Bureau van Dijk in their FAME database. European patent registers do not include reference numbers from company registers, nor does Bureau van Dijk provide the identification numbers used by patent offices in Europe. Linking the data from patent registers to firm register data requires matching of applicant names in patent documents and firm names in firm registers. In our work both a firm's current and previous name(s) were used for matching in order to account for changes in firm names. For more details on the matching of firm- and patent-level data see Appendix C.

Economic studies of entry are frequently hampered by the problem of identifying the correct set of potential entrants (Bresnahan and Reiss, 1991; Berry, 1992). In our case this problem is slightly mitigated by the fact that one set of potential entrants into patenting in a specific technology area consists of all those firms that currently patent in other technology areas. We complement this group of firms with a set of comparable firms from the population of UK firms that have not patented previously.

To construct the sample we deleted all firms from the data for which we have no size measure, because of missing data on assets. We select previously non-patenting firms from the population of all UK firms in two steps: 1) we delete all firms in industrial sectors with little patenting (amounting to less than 2 per cent of all patenting); and 2) we choose a sample of non-patenting firms that matches our sample of patenting firms by industry, size class, and age class. In principle, this approach will result in an endogenous (choice-based) sample. However our focus is on industry and technology area level effects rather than firm-level effects. Therefore we do not expect the sampling approach we adopt to introduce systematic biases into the estimates we report. We provide a number of robustness checks to ensure that our results are stable. These reveal that sample composition does not affect the key results we present below. All estimates are based on data weighted by the probability that a firm is in our sample.[11]

---

[11] To check this, we estimated the model with and without weights based on our sampling methodology and find little difference in the results.

The sample that results from our selection criteria is a set of firms with non-missing assets in manufacturing, oil and gas extraction and quarrying, construction, utilities, trade, and selected business services including financial services that includes all (approximately 10,000) firms applying for a patent at the EPO or UKIPO during the 2001-2009 period and another 10,000 firms that did not apply for a patent.

The definition of technology areas that we use is based on the 2008 version of the ISI-OST-INPI technology classification (denoted TF34 classes) (Schmoch, 2009). The list is shown in Table 1, along with the number of EPO and UKIPO patents applied for by UK firms with priority dates between 2002 and 2009. A comparison of the frequency distribution of patenting across the technology areas from the two patent offices shows that firms are more likely to apply for patents in Chemicals at the EPO, while Electrical and Mechanical Engineering predominate in the national patent data (see the bottom panel in Table 1).

We treat entry into each technology area as a separate decision made by firms. More than half of firms we observe patent in more than one area and 10 per cent patent in more than four. From the 20,000 firms observed, each of which can potentially enter into each one of the 34 technology areas, we obtain about 700,000 observations at risk. We cluster the standard errors by firm, so our models are effectively firm random effects models for entry into 34 technology areas. Allowing firm choices to vary by technology area is sensible under the assumption that firms' patenting strategies are contingent upon technology and industry level factors and are not homogeneous across technology areas. We confirmed the validity of this assumption through interviews with leading UK patent attorneys.

There are some technology-industry combinations that do not occur, e.g. audio-visual technology and the paper industry, telecommunications technology and the pharmaceutical industry. In order to reduce the size of the sample, we drop all technology-industry combinations for which Lybbert and Zolas (2014) find no patenting in their data and for which there was no patenting by any UK firm from the relevant industry in the corresponding technology category. This removes about 30 per cent of observations from the data. We provide a robustness check for this procedure in Table E-2 in the Appendix.

[Table 1 here]

3.1 Variables

Dependent Variable - Entry

The dependent variable is a dichotomous variable taking the value one if a firm has entered a technology area k at time t and otherwise the value zero. Entry into a technology area is measured by the first time a firm applies for a patent that is classified in that technology area, dated by the priority year of the patent.

Technological opportunity

Our first prediction from the theoretical model is that there will be more entry in technology areas with greater technological opportunity. Additional reasons that a sector may have more or less patenting include sector "size" or "breadth" and the propensity of firms to patent in particular technologies for strategic reasons or because of varying patent effectiveness in protecting inventions. To control for both technological opportunity and these other factors, we include the logarithm of the aggregate EPO patent applications in the technology sector during the year. To capture opportunity more specifically we also include the past 5-year growth rate in the non-patent (scientific publication) references cited in patents in that technology class at the EPO.[12] We have found that the growth rate in non-patent references is a better predictor of entry than the level of non-patent references, which has been used previously. Presumably the growth rate is a better indicator because it captures new or expanded technological opportunity.

Technology complexity

The second prediction of the theoretical model is that technological complexity increases entry, other things equal. Our interpretation of complexity is that it implies many interconnections between inventions in a particular field, rather than a series of fairly isolated inventions that do not connect to each other. To construct such a measure, we use the concept of network density applied to citations among all the patents that have issued in the particular technology area during the 10 years prior to the date of potential entry. We use citations at the U.S. patent office, both because these are richer (averaging 7 or so cites per patent during this period versus 3 for the EPO) and also to minimize correlation with the thickets measure, which is based on EPO data.[13]

The network density measure is computed as follows: in any year t, there are N_{kt} patents that have been applied for in technology area k between 1975 and year t. Each of these patents can cite any of the patents that were applied for earlier, which implies that the maximum number of citations within the technology area is given by $N_{kt}(N_{kt}-1)/2$. We count the actual number of citations made and normalize them by this quantity, scaling the measure by one million for visibility, given its small size.

Patent Thickets

The third prediction of our model is that greater potential for hold-up reduces entry. We measure the potential for hold-up in patent thickets using the triples count proposed by

---

[12] See von Graevenitz *et al.* (2013) for a more extensive discussion of this variable in the literature.

[13] It is important to emphasize that although patent offices cooperate and share search reports citations listed on U.S. patents are largely proposed by the applicant, whilst the citations listed on EPO and IPO patents are inserted by the examiner. This explains why the two measures are not highly correlated.

von Graevenitz *et al.* (2011). This is a narrower interpretation of this measure than in von Graevenitz *et al.* (2013), where it was used as a proxy for complexity and hold-up potential together.[14] In contrast, our model separates the effect of previously existing patent thickets on entry from that of technological complexity. The hold-up potential of thickets is captured by measuring how often patent applicants in a technology area simultaneously face patents that overlap with and block their own applications from firms that also block each-other's patent applications. This is more likely to happen in complex technologies, but does not necessarily arise as a result of complexity. Hence we use separate measures of complexity and hold-up potential.

The triples measure is a count of the number of fully connected triads on the set of firms' critical patent citations. At time *t* each unidirectional link between two firms *A* and *B* corresponds to one or more critical references to firm *A*'s patents in the set of patents applied for by firm *B* in the years *t, t-1* and *t-2*. These critical references are obtained from examiner search reports issued by the EPO and represent prior art that calls in question novelty and/or the inventive step of the patent application under examination.[15] Triples are then formed by groups of three firms where each firm has at least one patent that is cited as critical prior art in the search report for at least one patent held by each of the other two firms. That is, in a triple, each firm holds patents that potentially block the other firms' patents creating mutually blocking triads.

We use the same measure of triples as Harhoff *et al.* (2015), which contains all triples in each technology area. The citation data used is extracted from PATSTAT (October 2011 edition).[16] We normalize the count of triples by aggregate patenting in the same sector, so that the triples variable represents the intensity with which firms potentially hold blocking patents on each other relative to aggregate patenting activity in the technology.

The triples measure has been used in a number of papers since it was suggested by von Graevenitz *et al.* (2011). They show that counts of triples by technical area are significantly higher for technologies classified as complex than for areas classified as

---

[14] In Appendix D, we show that this confounded the separate effects of complexity and hold-up. Including the measures of complexity and hold-up potential proposed here in their empirical model, we find that the effects on patenting incentives predicted by our theoretical model for complexity (positive) and hold-up potential (negative) apply.

[15] These are the so-called X- and Y-references in EPO search reports. According to 9.2.1 of the EPO's Guidelines for Examination: "Category "Y" is applicable where a document is such that a claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other documents of the same category, such combination being obvious to a person skilled in the art. If a document explicitly refers to another document as providing more detailed information on certain features and the combination of these documents is considered particularly relevant, the primary document should be indicated by the letter "X"."

[16] Triples data was kindly provided by Harhoff et al. (2015).

discrete by Cohen *et al.* (2000). Fischer and Henkel (2012) find that the measure predicts patent acquisitions by Non-Practicing Entities. Graevenitz *et al.* (2013) use the measure to study patenting incentives in patent thickets and Harhoff *et al.* (2015) show that opposition to patent applications falls in patent thickets, particularly for patents of those firms that are caught up in thickets.

As a robustness check, we have also explored the use of duples, i.e. the count of mutual blocking relationships, to measure hold-up potential. Combining both measures in one regression leads to thorny problems of interpretation. Taken alone the measure has similar effects as the triples measure in this context.

[Table 2 here]

Covariates

It is well known that firm size and industry are important predictors of whether a firm patents at all (Bound *et al.* 1984 for U.S. data). Hall *et al.* (2013) show this for UK patenting during the period studied here. Therefore, in all of our regressions we control for firm size, industrial sector, and year of observation. We include the logarithm of the firm's reported assets and a set of year dummies in all the regressions.[17] To control for industrial sector, we stratify by industry, which effectively means that each industry has its own hazard function, which is shifted up or down by the other regressors.

We also expect the likelihood that a firm will enter a particular technology area to depend on its prior patenting experience overall, as well as its age. Long-established firms are less likely to be exploring new technology areas in which to compete. Thus we include the logarithm of firm age and the logarithm of the stock of prior patents applied for in any technology by the firm, lagged one year to avoid any endogeneity concerns. The variables on firm size and patent stock also allow us to test *Prediction 4* about the effect of incumbency advantage on entry.

3.2   Descriptive Statistics

Our estimation sample contains about 20,000 firms and 700,000 firm-TF34 sector combinations. During the 2002-2009 period there are about 10,000 entries into patenting for the first time in a technology area by these firms. Table C-1 in the appendix shows the distribution of the number of entries per firm: 2,531 enter one class, and the rest enter more than one. Table C-2 shows the population of UK firms obtained from FAME in our industries, together with the shares in each industry that have applied for a UK or European patent during the 2001-2009 period. These shares range from over 10 per cent in Pharmaceuticals and R&D Services to less than 0.1 per cent in Construction,

---

[17] The choice of assets as a size measure reflects the fact that it is the only size variable available for the majority of the firms in the FAME dataset.

Oil and Gas Services, Real Estate, Law, and Accounting.

3.3 Empirical Model

We use hazard models to estimate the probability of entry into a technology area. The models express the probability that a firm enters into patenting in a certain area conditional on not having entered yet as a function of the firm's characteristics and the time since the firm was "at risk," which is the time since the founding of the firm. In some cases, our data do not go back as far as the founding date of the firm, and in these cases the data are "left-censored." When we do not observe the entry of the firm into a particular technology sector by the last year (2009), the data is referred to as "right-censored."

In Appendix E, we discuss the choice of the survival models that we use for analysis, how to interpret the results, and present some robustness checks. We estimate two classes of failure or survival models: 1) proportional hazard, where the hazard of failure over time has the same shape for all firms, but the overall level is proportional to an index that depends on firm characteristics; and 2) accelerated failure time, where the survival rate is accelerated or decelerated by the characteristics of the firm. In the body of the paper we present results using the well-known Cox proportional hazards model stratified by industry. The effect of the stratification is that we allow firms in each of the industries to have a different distribution of the time until entry into patenting conditional on the regressors. That is, each industry has its own "failure" time distribution, where failure is defined as entry into patenting in a technology area, but the level of this distribution is also modified by the firm's size, aggregate patenting in the technology, network density, and the triples density.

Appendix Table E-1 shows exploratory regressions made using various survival models. None of the choices made large differences to the coefficients of interest, for instance results from the accelerated failure time models were similar to those of the Cox proportional hazards model, but the estimated effects are somewhat larger (shown in Table E-1).

As indicated earlier, our data for estimation are for the 2002-2009 period, but many firms have been at risk of patenting for many years prior to that. The oldest firm in our dataset was founded in 1856 and the average founding year was 1992. Because the EPO was only founded in 1978, we chose to use that year as the earliest date any of our firms is at risk of entering into patenting. That is, we defined the initial year as the maximum of the founding year and 1978. Table E-2 in the appendix presents estimates of our model using 1900 instead of 1978 as the earliest at risk year and finds little difference in the estimates.[18] We conclude that the precise assumption of the initial period is

---

[18] The main difference is in the firm age coefficient. Because the models are nonlinear, this coefficient is identified even in the presence of year dummies and vintage/cohort (which is implied by the survival

innocuous. Our assumption amounts to assuming that the shape of the hazard for firms founded between 1856 and 1978 but otherwise identical is the same during the 2002-2009 period.

4   Results

Our estimates of the model for entry into patenting are shown in Table 3. All regressions control for size, age, and industry. Both size and age are strongly positively associated with entry into patenting in a new technological area. Our indicator of technological opportunity and technology class size, the log of current patent applications in the technology class, is also positively associated with entry into that class, as predicted by our model.

Column 3 of Table 3 contains the basic result from our data and estimation, which is fully consistent with the predictions of our theoretical model: greater complexity as measured by citation network density increases the probability of entry into a technology area (Prediction 2), as does technological opportunity (Prediction 1), measured both as prior patenting in the class and as growth in the relevant science literature. Controlling for both technological opportunity and complexity, firms are discouraged from entry into areas with a greater density of triple relationships among existing firms (Prediction 3). We interpret this latter result as an indicator of the discouraging effect of hold-up possibilities or the legal costs associated with negotiation of rights or defense in the case of litigation.

We were concerned that our network density (complexity) and triples density (hold-up potential) measures might be too closely related to convey separate information, but we found that the raw correlation between these two variables was -0.001. To check for the impact of potential correlation conditional on year, industry, and the other variables, in columns 1 and 2 of Table 3 we included these two measures of complexity/thickets separately and found that although the coefficients were very slightly lower in absolute value, the results still hold, although it is clear that the aggregate class size is correlated negatively with the triples density via the denominator of the density (compare the change in the log (patents in class) coefficient between columns 1 and 2).

As we show in Appendix E, the estimated coefficients in the table are estimates of the elasticity of the yearly hazard rate with respect to the variable, and do not depend on the industry specific proportional hazard. A one standard deviation increase in the log of network density is associated with a 32 percent increase in the hazard of entry (0.13*2.78), while a one standard deviation in the log of triples density is associated

---

model formulation). However it will be highly sensitive to the assumptions about vintage due to the age-year-cohort identity.

with a 20 percent decrease in the hazard of entry (0.14*1.44). Thus the differences across these technology areas in the willingness of firms to enter them is substantial, bearing in mind that the average probability of entry is only about 1.5 per cent in this sample.

[Table 3 here]

There are fixed costs to patenting, and a firm may be more likely to enter into patenting in a new area if it already patents in another area. To test this idea, in the fourth column of Table 3, we add the logarithm of past patenting by the firm. Inline with Prediction 4, firms with a greater prior patenting history are indeed more likely to enter a new technology area – doubling a firm's past patents leads to an almost 100% higher hazard of entry.

In the last column we interact the log of assets with the log of patents, the log of network density, the growth of non-patent literature, and the log of triples density to see whether these effects vary by firm size. The results show that the network density and technological opportunity effects decline slightly with firm size. The triples density effect does not show any size relationship, suggesting that hold-up concerns affect firms of all sizes proportionately. We show this graphically in Figure 1, which overlays the coefficients as a function of firm size on the actual size distribution of our firms. From the graph one can see that the impact of aggregate patenting in a sector is higher and more variable than the impact of the network density, and that both fall to zero for the largest firms. Growth in non-patent literature is positively associated with technology entry for small firms, but negatively for large firms, suggesting the role played by the smaller firms in newer technologies based on science. Large firms seem not to be as active in these areas. Controlling for all these features of a technology, the impact of triples density is uniformly negative across firm size, which contradicts the view that the potential for hold-up discourages entry by smaller firms more than by larger firms.

4.1 Robustness

One concern we may have with the relationship between entry and the triples variable is simultaneity. That is, technology areas with lots of entry may also be prone to a higher triples density, just because of the entries. To address this possibility, we use the aggregate form of our entry regression. For each year we regress the log of the number of first time entries in each technology-industry sector combination on the characteristics of the technology class together with industry and year dummies. As instruments for the triples density, we use the median examination lag in the technology for patents applied for 5 and 6 years prior to the current year, which is long enough so that most of them will have been granted, rejected, or withdrawn. The idea is that classes with long examination lags may also be those where it is more difficult to assess patentability, leading to the hold-up potential captured by the triples proxy variable. We find that the instrumental variables regression easily passes the

specification tests for under-, weak and over-identification, justifying our choice of instruments.

Table 4 shows the results, both ordinary least squares and instrumental variables.[19] Note that we do not expect results to be identical when comparing the aggregate regressions to individual firm-level hazard rate estimations, as the functional forms of the models differ. However, the results are similar in sign to those in column 3 of Table 3, with the exception of the technological opportunity coefficient (the past five-year growth in non-patent-literature references) which has the opposite sign. For our purposes, interest centers on the coefficient of triples density. The least squares estimate of the elasticity is negative and implies a 12 per cent reduction in entry per year when the triples density increases by one standard deviation. Instrumenting this variable triples its coefficient, which suggests that our hazard rate estimates may by an underestimate of the true impact of potential hold-up on entry.

Table E-2 in the appendix explores some variations of the sample used for estimation in Table 3. Column 1 of Table E-2 is the same as column 4 of Table 3 for comparison. The first change (column 2) was to add back all the technology-industry combinations where Lybbert and Zolas (2012) find no patenting in their data and where there was no entry by any UK firm from the relevant industry into that technology category. These observations are about 20 per cent of the sample. The impact of network density on entry is weaker, but the impact of triples density and the technological opportunity variables is considerably stronger. That is, technology area-industry combinations with no patenting are also those where the technology area displays low technological opportunity.

Next we removed all the firms with assets greater than 12.5 million pounds, to check whether large firms were responsible for our findings.[20] This removed about 2 per cent of the 20,000 firms. Column 3 of Table E-2 shows that the results do not change a great deal, although they are somewhat stronger. In column 4, we removed the telecommunications technology sector from the estimation, because it is such a large triples outlier. Once again, there was little change to the estimates. The last column of Table E-2 shows the results of defining the minimum entry year as 1900. With the exception of firm age, the coefficients are nearly identical to those in column 1 of the table.

---

[19] We also estimated this model by LIML and GMM, with almost no change in the resulting coefficients (not shown).

[20] 12.5 million pounds is a cutoff based on the definition of Small and Medium-sized Enterprises (SMEs) as firms with fewer than 250 employees. We do not have employment for all our firms, so we assume that assets are approximately 50 thousand pounds per employee in order to compute this measure. For small firms only, this yields an assets cutoff of 2.5 million pounds.

5   Conclusion

Patent thickets arise for a multitude of reasons; they are mainly driven by an increase in the number of patent filings and concomitant reductions in patent quality (that is, the extent to which the patent satisfies the requirements of patentability) as well as increased technological complexity and interdependence of technological components. The theoretical analysis of patent thickets (Shapiro, 2001) and the qualitative evidence provided by the FTC in a number of reports (FTC, 2003; 2011) suggest that thickets can impose significant costs on some firms. The subsequent literature has focused on the measurement of thickets (e.g. Graevenitz *et al.* 2011; Ziedonis, 2004) and has linked thickets to changes in firms' intellectual property strategies in a number of dimensions. There is still a lack of evidence on the effect of patent thickets as well as their welfare implications at the aggregate level.

The empirical analysis of the effects of patent thickets must contend with two challenges: first, patent thickets have to be measured and secondly, effects of thickets must be separated from effects of other factors that are correlated with the growth of thickets, in particular technological complexity.

This paper confronts both challenges. We show that our empirical measure for the density of thickets captures effects of patent thickets predicted by theory. We separate the impact of patent thickets on entry from effects of technological opportunity and complexity and show that thickets reduce entry into patenting. Controlling for technological opportunity and complexity is important because both are correlated with entry into patenting and the presence of thickets. It is also worth emphasizing that our measure of thickets is purged of effects that are driven by patenting trends in particular technologies. That is, our results are not due to the level of invention and technological progress within a technology field.

Our results demonstrate that patent thickets significantly reduce entry into those technology areas in which growing complexity and growing opportunity increase the underlying demand for patent protection. These are the technology areas, which are associated most with productivity growth in the knowledge economy. However, the welfare consequences of our finding are unclear. Reduced entry into new technology areas could be welfare-enhancing: As is well known from the industrial organization literature, entry into a market may be excessive if entry creates negative externalities for active firms, for instance due to business stealing. This is likely to be true of patenting too. Furthermore, Arora *et al.* (2008) show that the patent premium does not cover the costs of patenting for the average patent (except for pharmaceuticals). These and related facts might lead one to conclude that lower entry into patenting is likely to increase welfare and that thickets raise welfare by reducing entry.

In contrast, reduced entry into patenting in new technology areas may also be welfare-reducing, for at least two reasons. First, there is the obvious argument that the benefits from more innovation may exceed any business stealing costs (as has been shown

empirically in the past by others, e.g., Bloom *et al.* 2013), so that some desirable innovation may be deterred by high entry costs. Even if this were not true, there is no reason to believe that firms that do not enter into patenting due to thickets are those we wish to deter. Given the incumbency advantage, it is likely that the failure to enter into patenting in these areas reflects less innovation by those who bring the most original ideas, that is, by those who are inventing "outside the box."

References

Allison, J.R., M.A. Lemley, and D. Schwartz, 2015. Our Divided Patent System. *The University of Chicago Law Review* 82(3): 1073-1154.

Amir, R. and V. E. Lambson, 2000. On the effects of entry in Cournot markets. *The Review of Economic Studies* 67(2): 235–254.

Arora, A., M. Ceccagnoli, and W.M. Cohen, 2008. R&D and the Patent Premium. *International Journal of Industrial Organization*, 26(5): 1153-1179.

Berry, S., 1992. Estimation of a model of entry in the airline industry, *Econometrica* 60: 889-917.

Bessen, J., and E. Maskin, 2009. Sequential Innovation, Patents, and Imitation. *RAND Journal of Economics* 40: 611-635.

Bessen, J. M. Meurer, and J. Ford, 2011. The Private and Social Costs of Patent Trolls. Boston, MA: Boston University School of Law, Law and Economics Research Paper No. 11-45.

Bessen, J. and M. Meurer, 2013. The Patent Litigation Explosion. *Loyola University Chicago Law Journal* 45(2): 401-440.

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

Bloom, N., M. Schankerman, and J. van Reenen, 2013. Idenitfying technology spillovers and product market rivalry, *Econometrica* 81 (4): 1347–1393.

Bound, J., C. Cummins, Z. Griliches, B.H. Hall, and A. Jaffe, 1984. Who does R&D and who patents? *R&D Patents and productivity*, Z. Griliches (ed.). Chicago, University of Chicago Press.

Bresnahan, T. and P. Reiss, 1991. Entry and competition in concentrated markets, *Journal of Political Economy* 99: 977-1009.

Chien, C. and M. A. Lemley, 2012. Patent Holdup, the ITC, and the Public Interest. *Cornell Law Review* 98(1).

Clark, D.J. and K. Konrad, 2008. Fragmented Property Rights and Incentives for R&D, *Management Science* 54(5): 969-981.

Cockburn, I.M. and M.J. MacGarvie, 2011. Entry and Patenting in the Software Industry. *Management Science* 57: 915–933.

Cohen, W.M., R.R. Nelson, and J.P. Walsh, 2000. Protecting their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not). Cambridge, MA: NBER Working Paper No. 7552.

Farrell, J. and C. Shapiro, 2008. How strong are weak patents? *American Economic Review* 98: 1347–1369.

Federal Trade Commission (FTC), 2003. *To Promote Innovation - The Proper Balance of Competition and Patent Law and Policy*. Washington, DC: GPO. Available at http://www.ftc.gov/reports/innovation/P040101PromotingInnovationandCompetitionrpt0704.pdf

Federal Trade Commission (FTC), 2011. *The Evolving IP Marketplace:. Aligning Patent Notice and Remedies With Competition.* Washington, DC. Available at http://www.ftc.gov/os/2011/03/110307patentreport.pdf

Fershtman, C. and M. I. Kamien, 1992. Cross licensing of complementary technologies *International Journal of Industrial Organization* 10(3): 329-348.

Fink C., M. Khan, and H. Zhou (2013). Exploring the worldwide patent surge. WIPO Economic Research Working Paper No. 12.

Fischer, T. and J. Henkel, 2012. Patent trolls on markets for technology–An empirical analysis of NPEs' patent acquisitions. *Research Policy*, 41(9): 1519-1533.

Goolsbee, A. (1998). Does Government R&D Policy Mainly Benefit Scientists and Engineers? *The American Economic Review*, 88(2): 298-302.

Graevenitz, von G., S. Wagner, and D. Harhoff, 2011. How to measure patent thickets—A novel approach. *Economics Letters* 111: 1–4.

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

Hall, B.H., C. Helmers, M. Rogers, and V. Sena, 2013. The importance (or not) of patents to UK firms. *Oxford Economic Papers* 65 (3): 603-629.

Harhoff, D., G. von Graevenitz, and S. Wagner, 2015. Conflict Resolution, Public Goods and Patent Thickets. Management Science, forthcoming.

Henderson, R., 1993. Underinvestment and Incompetence as Responses to Radical Innovation: Evidence from the Photolithographic Alignment Equipment Industry, *The RAND Journal of Economics* 24(2): 248-270.

Jaffe, A. and J. Lerner, 2004. Innovation and Its Discontents. Princeton University Press.

Joshi, A.M. and A. Nerkar, 2001. When do strategic alliances inhibit innovation by firms? Evidence from patent pools in the global optical disc industry. *Strategic*

Management Journal 32(11): 1139-1160.

Lerner, J., 1995. Patenting in the shadow of competitors. *Journal of Law and Economics* 38: 463-495.

Lerner, J. and J. Tirole, 2004. Efficient Patent Pools. *American Economic Review* 94(3): 691-711.

Lerner, J., M. Strojwas, and J. Tirole. 2007. The design of patent pools: the determinants of licensing rules, *The RAND Journal of Economics* 38(3): 610-625.

Lerner, J. and J. Tirole, 2013. Standard-Essential Patents, NBER Working Paper 19664.

Lewis, J.I.D. and R.M. Mott, 2013. The sky is not falling: Navigating the smartphone patent thicket, *WIPO Magazine*, February 2013, http://www.wipo.int/wipo_magazine/en/2013/01/article_0002.html

Lybbert, T. J. and N. J. Zolas, 2014. Getting Patents & Economic Data to Speak to Each Other: An 'Algorithmic Links with Probabilities' Approach for Joint Analyses of Patenting & Economic Activity. *Research Policy* 43(3): 530-542.

Milgrom P. and J. Roberts, 1994. Comparing equilibria. *American Economic Review* 84(3): 441–459.

Mulligan C. and T. B. Lee, 2012. Scaling the patent system. *NYU Annual Survey of American Law* 68: 289-317.

Schmoch, U., 2009. Document IPC/CE/41/5, Annex, Conception of a Technology Classification for Country Comparisons, *41th session of the IPC Committee of Experts.* 1–15.

Shapiro, C., 2001. Navigating the patent thicket: Cross licenses, patent pools, and standard setting. *Innovation policy and the economy* 1: 119–150.

Teece, D., E. Sherry, and P. Grindley, 2014. Patents and "Patent Wars" in Wireless Communications: An Economic Assessment, *Digiworld Economic Journal* 95: 85-98.

U.S. Department of Justice (USDoJ), Federal Trade Commission (FTC), 2007. *Antitrust Enforcement and Intellectual Property Rights: Promoting Innovation and Competition*, 1–217.

Vives, X., 2005. Complementarities and games: New developments. *Journal of Economic Literature* XLIII(2): 437–479.

Ziedonis, R.H., 2004. Don't Fence Me In: Fragmented Markets for Technology and the Patent Acquisition Strategies of Firms. *Management Science* 50: 804–820.

Table 1

Patenting by Fame firms on Patstat (priority years 2002-2009)

Technology categories	Weighted by #owners & #classes*			Sector shares	
	GB pats	EP pats	Total	GB pats	EP pats
Elec machinery, energy	1,321	1,101	2,422	6.1%	4.4%
Audio-visual tech	633	549	1,182	2.9%	2.2%
Telecommunications	1,181	1,206	2,386	5.5%	4.8%
Digital communication	590	732	1,323	2.7%	2.9%
Basic comm processes	302	146	447	1.4%	0.6%
Computer technology	1,481	1,302	2,783	6.8%	5.2%
IT methods for mgt	256	224	480	1.2%	0.9%
Semiconductors	269	248	518	1.2%	1.0%
Optics	392	481	873	1.8%	1.9%
Measurement	1,216	1,458	2,674	5.6%	5.8%
Analysis bio materials	132	426	557	0.6%	1.7%
Control	592	542	1,134	2.7%	2.2%
Medical technology	996	1,561	2,558	4.6%	6.3%
Organic fine chemistry	182	1,538	1,720	0.8%	6.2%
Biotechnology	193	950	1,143	0.9%	3.8%
Pharmaceuticals	277	1,876	2,153	1.3%	7.5%
Polymers	114	280	394	0.5%	1.1%
Food chemistry	88	458	547	0.4%	1.8%
Basic materials chemistry	314	1,050	1,363	1.5%	4.2%
Materials metallurgy	161	318	479	0.7%	1.3%
Surface tech coating	287	284	571	1.3%	1.1%
Chemical engineering	507	724	1,231	2.3%	2.9%
Environmental tech	296	344	640	1.4%	1.4%
Handling	996	813	1,809	4.6%	3.3%
Machine tools	428	356	784	2.0%	1.4%
Engines,pumps,turbine	887	942	1,829	4.1%	3.8%
Textile and paper mach	235	304	539	1.1%	1.2%
Other spec machines	742	623	1,365	3.4%	2.5%
Thermal process and app	410	261	671	1.9%	1.0%
Mechanical elements	1,149	854	2,002	5.3%	3.4%
Transport	1,063	930	1,993	4.9%	3.7%
Furniture, games	1,064	612	1,675	4.9%	2.5%
Other consumer goods	630	507	1,137	2.9%	2.0%
Civil engineering	2,237	960	3,196	10.3%	3.8%
Total	**21,619**	**24,959**	**46,578**		
Electrical engineering	6,032	5,508	11,540	27.9%	22.1%
Instruments	3,328	4,468	7,796	15.4%	17.9%
Chemistry	2,418	7,822	10,240	11.2%	31.3%
Mechanical engineering	5,910	5,083	10,993	27.3%	20.4%
Other Fields	3,930	2,079	6,009	18.2%	8.3%

* Weighting by owners does not affect the numbers, since they all get added back into the same cell.

Weighting by classes means that a patent in multiple TF34 sectors is downweighted in each of the sectors.

Table 2

UKIPO and EPO patents: numbers, triples and network density 2002-2009

Technology categories	Aggregate EPO patents	Number of EPO triples@	Triples per 1000 patents	US Citation network density#	Average non-patent references
Elec machinery, energy	56,714	7751	136.7	39.4	0.420
Audio-visual tech	34,131	13268	388.7	63.5	0.449
Telecommunications	62,288	27049	434.3	79.1	1.235
Digital communication	36,975	16529	447.0	178.5	1.397
Basic comm processes	10,035	2289	228.1	110.8	1.162
Computer technology	60,577	21956	362.4	54.3	1.529
IT methods for mgt	9,312	34	3.7	144.2	0.920
Semiconductors	24,544	9974	406.4	94.0	1.070
Optics	28,458	7767	272.9	58.0	0.806
Measurement	44,320	2503	56.5	45.6	1.000
Analysis bio materials	11,787	26	2.2	319.4	7.040
Control	17,612	308	17.5	112.1	0.445
Medical technology	66,062	4411	66.8	206.2	0.614
Organic fine chemistry	41,137	3993	97.1	32.9	5.253
Biotechnology	33,192	365	11.0	89.6	17.332
Pharmaceuticals	52,671	11222	213.1	76.6	6.391
Macromolecular chemistry	21,307	3722	174.7	92.7	1.236
Food chemistry	9,955	140	14.1	326.3	2.701
Basic materials chemistry	27,679	1929	69.7	84.8	1.498
Materials metallurgy	16,935	405	23.9	91.7	1.130
Surface tech coating	17,429	363	20.8	59.4	0.803
Chemical engineering	24,494	443	18.1	66.2	0.797
Environmental tech	12,708	858	67.5	206.9	0.487
Handling	30,343	252	8.3	66.9	0.137
Machine tools	24,040	508	21.1	64.0	0.191
Engines,pumps,turbine	32,602	6678	204.8	85.4	0.210
Textile and paper mach	23,145	2640	114.1	84.9	0.312
Other spec machines	29,826	319	10.7	65.7	0.422
Thermal process and app	15,290	335	21.9	146.3	0.189
Mechanical elements	32,716	1301	39.8	57.8	0.168
Transport	48,875	10929	223.6	68.3	0.203
Furniture, games	19,847	206	10.4	107.6	0.166
Other consumer goods	19,734	301	15.3	105.6	0.194
Civil engineering	28,817	171	5.9	117.1	0.150
Total	**1,025,555**	**160,945**	**156.9**	**100.3**	**#REF!**
Electrical engineering	294,575	98,850	335.6	60.4	1.042
Instruments	168,239	15,015	89.2	96.3	1.181
Chemistry	257,507	23,440	91.0	71.1	4.977
Mechanical engineering	236,836	22,962	97.0	70.1	0.227
Other Fields	68,398	678	9.9	110.8	0.167

@ Triples based on all EPO patenting, priority years 2002-2009 (see text for definition and further explanation).

Network density is 1,000,000 times the number of within technology citations between 1976 and the current year divided by the potential number of such citations.

Table 3
Hazard of entry into patenting in a TF34 Class
538,452 firm-TF34 observations with 10,665 entries (20,384 firms)

Variable	Cox Proportional Hazard Model				
Log (network density)	0.115***		0.127***	0.107***	0.184***
	(0.024)		(0.023)	(0.021)	(0.052)
Log (triples density		-0.138***	-0.139***	-0.101***	-0.100***
in class)		(0.011)	(0.011)	(0.010)	(0.023)
Log (patents in class)	0.317***	0.506***	0.545***	0.514***	0.822***
	(0.025)	(0.031)	(0.030)	(0.027)	(0.071)
5-year growth of non-	0.060***	0.084***	0.072***	-0.009	0.103*
patent refs in class)	(0.022)	(0.022)	(0.022)	(0.021)	(0.056)
Log assets	0.270***	0.270***	0.270***	0.142***	0.513***
	(0.011)	(0.011)	(0.011)	(0.013)	(0.083)
Log firm age in years	1.135***	1.135***	1.136***	0.773***	0.767***
	(0.104)	(0.104)	(0.104)	(0.130)	(0.131)
Log (pats applied for				0.836***	0.836***
by firm previously)				(0.021)	(0.021)
Log (network density)					-0.010*
* Log assets					(0.006)
Log (triples density)					-0.001
* Log assets					(0.003)
Log (patents in class)					-0.040***
* Log assets					(0.008)
Log (average NPL refs)					-0.015**
* Log assets					(0.006)
Industry dummies	*stratified#*	*stratified#*	*stratified#*	*stratified#*	*stratified#*
Year dummies	*yes*	*yes*	*yes*	*yes*	*yes*
Log likelihood	-65.96	-65.86	-65.84	-58.69	-58.67
Degrees of freedom	12	12	13	14	18
Chi-squared	1270.6	1429.1	1517.2	3465.1	3458.6

The sample is matched on size class, sector, and age class. Estimates are weighted by sampling probability.

Coefficients for the hazard of entry into a patenting class are shown.

Standard errors are clustered on firm. *** (**) denote significance at the 1% (5%) level.

Time period is 2002-2009 and minimum entry year is 1978. Sample is UK firms with nonmissing assets, all patenting firms and a matched sample of non-patenting firms

Estimates are stratified by industry - each industry has its own baseline hazard.
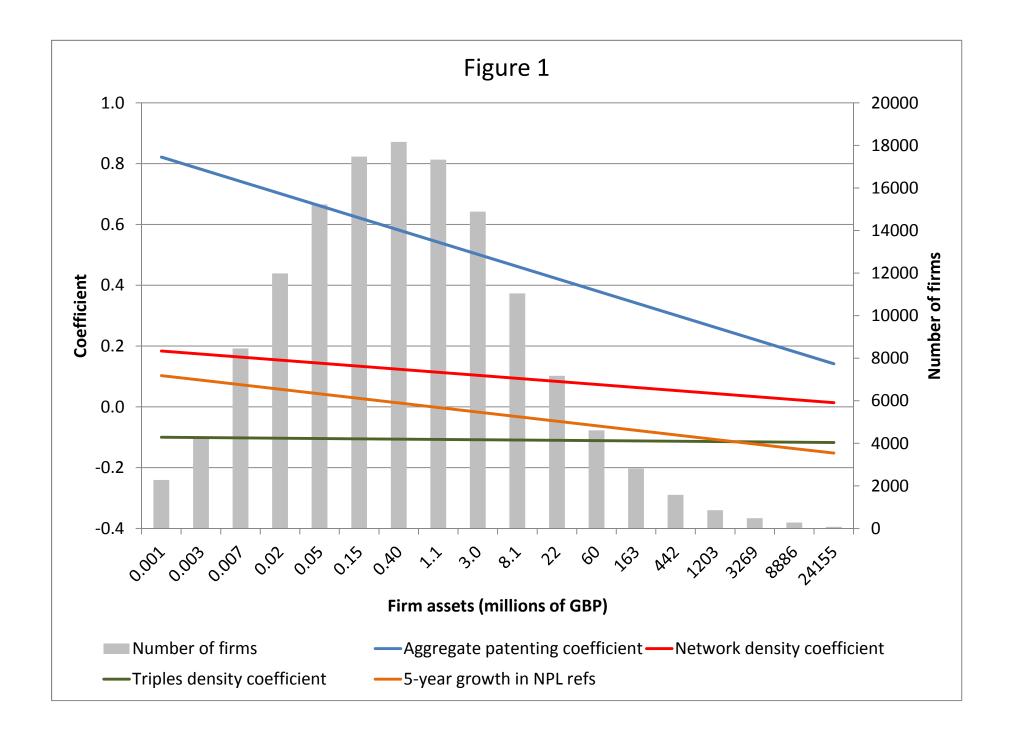
Table 4
Aggregate regressions for entry into patenting classes 2001-2009
9 years*33 tech classes*26 sectors = 7722 observations

Dependent variable	Log number of first time entries by a firm into class by sector					
	OLS			IV#		
	Coef.	*s.e.**		*Coef.*	*s.e.**	
Log (US network density)	0.018	0.029		0.055	0.032	*
Log (triples density)	-0.079	0.010	***	-0.251	0.042	***
Log (patent apps in class)	0.247	0.030	***	0.469	0.059	***
Past 5 year growth in NPL refs	-0.099	0.015	***	-0.101	0.016	***
R-squared		0.478			0.389	
Standard error		0.564			0.608	

*Standard errors are clustered on class and industrial sector (which allows free correlation over time).

Instruments are lag 5 and 6 median exam duration for patents in the class.

Log of triples density is treated as endogenous in the IV and LIML estimates

Figure 1

Appendix

A   Previous Results

To help the reader this appendix summarizes a number of results derived by Graevenitz et al. (2013) as well as some additional results that are useful.

A.1   The Probability of Patenting a Facet

The probability p_k that a firm obtains a patent on a facet is:

$$p_k(\mathbf{f}_{/k}, F, N_O(O, \mathbf{o}_{/k}, N)) = \sum_{j=0}^{N_O} \frac{1}{j+1} \binom{N_O}{j} \prod_{l=0}^{N_O-j} \left(1 - \frac{f_l}{F_k}\right) \prod_{m=N_O-j}^{N_O} \frac{f_m}{F_k} \quad . \tag{A.1}$$

Then, the expected number of patents a firm owns when it applyies for f_i facets is $\gamma_k \equiv p_k f_i$.

For the comparative statics of entry stage it is useful to know that the elasticity of p_k w.r.t. F is negative if $\hat{\phi}_k < \frac{1}{2}$:

$$\frac{\partial p_k}{\partial F_k} = \sum_{i=0}^{N_O} \frac{1}{i+1} \binom{N_O}{i} (1 - \hat{\phi}_k)^{N_O-i} \hat{\phi}_k^i (-1) \left(\frac{N_O}{F_k} - \frac{N_O - i}{F - \hat{f}}\right) \tag{A.2}$$

Then the elasticity ϵ_{p_k, F_k} is:

$$\epsilon_{p_k, F_k} = N_O^2 \frac{\left(\hat{\phi}_k - \frac{1}{2}(1 + \frac{1}{N_O})\right)}{1 - \hat{\phi}_k} \tag{A.3}$$

A.2   The Expected Number of Rival Investors

The expected number of rival firms N_O that undertake R&D on the same technology opportunity as firm i can be expressed as a sum of products:

$$N_O = \sum_{j=0}^{N} j \binom{N}{j} \prod_{l=0}^{N-j} (1 - \omega_l) \prod_{m=N-i}^{N} \omega_m. \tag{A.4}$$

Graevenitz et al. (2013) show that N_O is increasing in ω_n, where $n \in \{l, m\}$.

In the second stage equilibrium N_O can be rewritten as:

$$N_O = \sum_{j=0}^{N} j \binom{N}{j} (1 - \hat{\omega}_k)^{(N-j)} \hat{\omega}_k^j. \tag{A.5}$$

Incumbency Advantage

In the case in which there are incumbents and entrants the expected number of rival firms N_O has to rewritten slightly. To do this define:

$$\omega_n^I \equiv o_i^I / O \qquad\qquad \omega_n^E \equiv o_i^E / O \qquad\qquad \text{(A.6)}$$

We assume that in a previous period N^p firms entered and of these a fraction λ are still active. Then the expected number of rival firms $\tilde{N}_O$ that undertake R&D on the same technology opportunity as firm i is:

$$\tilde{N}_O = \sum_{j=0}^{\lambda N^P} j \binom{\lambda N^P}{j} (1 - \hat{\omega}_k^J)^{(N-j)} (\hat{\omega}_k^J)^j + \sum_{j=0}^{N} j \binom{N}{j} (1 - \hat{\omega}_k^E)^{(N-j)} (\hat{\omega}_k^E)^j. \qquad \text{(A.7)}$$

A.3   The Expected Number of Facets Covered

In the second stage equilibrium the expected number of facets covered through the joint efforts of all firms investing in a technological opportunity is:

$$\tilde{F}_k = F\left[1 - (1 - \hat{\phi}_k)^{(N_O+1)}\right] \qquad\qquad \text{(A.8)}$$

The derivative of this expression with respect to F is positive:

$$\frac{\partial \tilde{F}_k}{\partial F_k} = 1 - \left(1 - \hat{\phi}_k\right)^{N_O} \left(1 + \hat{\phi}_k N_O\right) \geq 0 . \qquad\qquad \text{(A.9)}$$

The elasticities of $\tilde{F}_k$ with respect to f_k and F are:

$$\hat{\eta}_k = \frac{\hat{\phi}_k (1 - \hat{\phi}_k)^{N_O}}{1 - (1 - \hat{\phi}_k)^{(N_O+1)}} \qquad\qquad \text{(A.10)}$$

$$\hat{\epsilon}_{\tilde{F}_k, F_k} = \frac{1 - (1 - \hat{\phi}_k)^{(N_O)} (1 + \hat{\phi}_j N_O)}{1 - (1 - \hat{\phi}_k)^{(N_O+1)}} . \qquad\qquad \text{(A.11)}$$

which shows that $1 \geq \epsilon_{\tilde{F}_k F_k} \geq 0$ as the denominator in the fraction is always greater than the numerator. It is useful to observe that the upper bound of the elasticity $\hat{\eta}_k$ is decreasing in N_O. To see this note that the elasticity can be expressed as:

$$\hat{\eta}_k = \frac{(1 - \hat{\phi}_k)^{N_O}}{(N_o + 1)\left(1 - \hat{\phi}_k \frac{N_o}{2!} + \hat{\phi}_k^2 \frac{N_o(N_O-1)}{3!} ...\right)} . \qquad\qquad \text{(A.12)}$$

This shows that the upper bound of the elasticity decreases in N_O: $\lim_{\hat{\phi}_k \to 0} \eta_k = 1 / (N_O + 1) \leq 1$. Here we use the binomial expansion of $(1 - \hat{\phi}_k)^{N_o+1}$. The expression also shows that the lower bound of $\eta_{k|\hat{\phi}_k=1}$ is zero.

B Results

This appendix contains derivations for the propositions set out in Section 2 of the paper.

B.1 Supermodularity of the Second Stage Game

This section sets out the main results needed to show that the second stage of game G* is supermodular.

Consider the first order conditions that determine the equilibrium number of facets ($\hat{f}$) and technological opportunities ($\hat{o}$):

$$\frac{\partial \pi_{ik}}{\partial o_i} = V\Delta(s_{ik}) - L(\gamma_{ik}, s_{ik}) - C_o(\Sigma_{j=1}^{N_o} o_j) - \gamma_{ik}C_a - \frac{\partial C_c}{\partial o_i} = 0 \quad , \tag{B.13}$$

$$\frac{\partial \pi_{ik}}{\partial f_i} = \frac{o_i p_k}{\tilde{F}_k} \left(\left[V\mu_k\eta_{ik}\frac{\Delta(s_{ik})}{s_{ik}} - \tilde{F}_k\left(\frac{\partial L}{\partial \gamma_{ik}} + C_a\right) \right] + \left[V\frac{d\Delta}{ds_{ik}} - \frac{\partial L}{\partial s_{ik}} \right](1-\eta_{ik}) \right) = 0 \quad . \tag{B.14}$$

Now, consider the cross-partial derivatives which must be positive, if the second stage game is supermodular. First, we derive the cross partial derivative with respect to firms' own actions:

$$\frac{\partial^2 \pi_{ik}}{\partial o_i \partial f_i} = \frac{p_k}{\tilde{F}_k} \left(\left[V\mu_k\eta_{ik}\frac{\Delta(s_{ik})}{s_{ik}} - \tilde{F}_k\left(\frac{\partial L}{\partial \gamma_{ik}} + C_a\right) \right] + \left[V\frac{d\Delta}{ds_{ik}} - \frac{\partial L}{\partial s_{ik}} \right](1-\eta_{ik}) \right) = 0 \quad . \tag{B.15}$$

This expression corresponds to the first order condition (B.14) for the optimal number of facets.

Now consider effects of rivals' actions on firms' own actions:

$$\frac{\partial^2 \pi_k}{\partial o_i \partial o_m} = \frac{\partial \tilde{F}_k}{\partial o_m}\frac{s_{ik}}{\tilde{F}_k}\left[V\frac{\Delta}{s_{ik}}(\mu_k - \xi_{ik}) + \frac{\partial L}{\partial s_{ik}} \right] + \frac{\partial p_k}{\partial o_m}\frac{f_i}{\tilde{F}_k}\left[\left(V\frac{d\Delta}{ds_{ik}} - \frac{\partial L}{\partial s_{ik}}\right) - \tilde{F}_k\left(\frac{\partial L}{\partial \gamma_{ik}} + C_a\right) \right] \tag{B.16}$$

$$- \frac{\partial C_o}{\partial \Sigma_{j=1}^{N_o} o_j},$$

$$\frac{\partial^2 \pi_k}{\partial o_i \partial f_m} = \frac{\partial \tilde{F}_k}{\partial f_m}\frac{s_{ik}}{\tilde{F}_k}\left[V\frac{\Delta}{s_{ik}}(\mu_k - \xi_{ik}) + \frac{\partial L}{\partial s_{ik}} \right] + \frac{\partial p_k}{\partial f_m}\frac{f_i}{\tilde{F}_k}\left[\left(V\frac{d\Delta}{ds_{ik}} - \frac{\partial L}{\partial s_{ik}}\right) - \tilde{F}_k\left(\frac{\partial L}{\partial \gamma_{ik}} + C_a\right) \right], \tag{B.17}$$

$$\frac{\partial^2 \pi_k}{\partial f_i \partial o_m} = \frac{\partial \tilde{F}_k}{\partial o_m}\left[\frac{\partial V}{\partial \tilde{F}_k} + \frac{\partial^2 V}{\partial \tilde{F}_k^2}\tilde{F}_k\eta_{ik} - \frac{\partial L}{\partial \gamma_{ik}} - C_a + \frac{\partial^2 L}{\partial s_{ik}^2}\frac{s_{ik}}{\tilde{F}_k}(1-\eta_{ik}) \right]$$
$$+ \frac{\partial \eta_{ik}}{\partial o_m}\left(V\frac{\Delta}{s_{ik}}(\mu_k - \xi_{ik}) + \frac{\partial L}{\partial s_{ik}} \right) - \frac{\partial p_k}{\partial o_m}\left[\frac{\partial^2 L}{\partial \gamma_{ik}^2}f_i + \frac{\partial^2 L}{\partial s_{ik}^2}\frac{f_i}{\tilde{F}_k}(1-\eta_{ik}) \right], \tag{B.18}$$

$$\frac{\partial^2 \pi_k}{\partial f_i \partial f_m} = \frac{\partial \tilde{F}_k}{\partial f_m}\left[\frac{\partial V}{\partial \tilde{F}_k}\frac{\Delta}{s_{ik}}\left(\xi_{ik}(1-\epsilon_{\tilde{F}_k,f}) + \epsilon_{\tilde{F}_k,f} \right) + \frac{\partial^2 V}{\partial \tilde{F}_k^2}\tilde{F}_k\eta_{ik}\frac{\Delta}{s_{ik}} - \frac{\partial L}{\partial \gamma_{ik}} - C_a + \frac{\partial^2 L}{\partial s_{ik}^2}\frac{s_{ik}}{\tilde{F}_k}(1-\eta_{ik}) \right]$$
$$+ \frac{\partial \eta_{ik}}{\partial f_m}\left(V\frac{\Delta}{s_{ik}}(\mu_k - \xi_{ik}) + \frac{\partial L}{\partial s_{ik}} \right) - \frac{\partial p_k}{\partial f_m}\left[\frac{\partial^2 L}{\partial \gamma_{ik}^2}f_i + \frac{\partial^2 L}{\partial s_{ik}^2}\frac{f_i}{\tilde{F}_k}(1-\eta_{ik}) \right] \quad . \tag{B.19}$$

The second stage game is supermodular, if the equations (B.16)-(B.19) are non-negative. The follow-

3

ing results show that the conditions noted in Section 2 above must hold simultaneously if the game is supermodular.

Using the first order condition (B.14), which will hold for any interior equilibrium, it can be shown that:

$$\left[\left(V\frac{d\Delta}{ds_{ik}} - \frac{\partial L}{\partial s_{ik}}\right) - \tilde{F}_k\left(\frac{\partial L}{\partial \gamma_{ik}} + C_a\right)\right] = -\eta_{ik}\left(V\frac{\Delta}{s_{ik}}(\mu_k - \xi_{ik}) + \frac{\partial L}{\partial s_{ik}}\right). \tag{B.20}$$

If $\left(V\frac{\Delta}{s_{ik}}(\mu_k - \xi_{ik}) + \frac{\partial L}{\partial s_{ik}}\right) > 0$, then the second term in the cross-partial derivatives (B.16) and (B.17) is the product of two negative expressions, and then equation (B.17) is positive. Equation (B.16) is also positive in a free entry equilibrium: the negative term at the end is less than the negative term in the derivative of profits w.r.t. N_k in Section 2, which is otherwise the same as equation (B.16): $\frac{\partial C_o}{\partial N_o\hat{o}}\hat{o} > \frac{\partial C_o}{\partial N_o\hat{o}}$.

Turning to equations (B.18) and (B.19) we can show that:

$$\frac{\partial \eta_{ik}}{\partial o_m} = \frac{\partial^2 \tilde{F}_k}{\partial f_i \partial o_m}\frac{f_i}{\tilde{F}_k} - \frac{\partial \tilde{F}_k}{\partial f_i}\frac{\partial \tilde{F}_k}{\partial o_m}\frac{f_i}{\tilde{F}_k^2} = -\tilde{F}_k^{-1}\frac{\partial \tilde{F}_k}{\partial o_m}\left(\frac{\phi_k}{1-\phi_k} + \eta_{ik}\right) \tag{B.21}$$

$$\frac{\partial \eta_{ik}}{\partial f_m} = \frac{\partial^2 \tilde{F}_k}{\partial f_i \partial f_m}\frac{f_i}{\tilde{F}_k} - \frac{\partial \tilde{F}_k}{\partial f_i}\frac{\partial \tilde{F}_k}{\partial f_m}\frac{f_i}{\tilde{F}_k^2} = -\tilde{F}_k^{-1}\frac{\partial \tilde{F}_k}{\partial f_m}\left(\frac{\phi_k}{1-\phi_k} + \eta_{ik}\right) \tag{B.22}$$

This result allows us to rewrite equations (B.18) and (B.19) as follows:

$$\frac{\partial^2 \pi_k}{\partial f_i \partial o_m} = \frac{1}{\tilde{F}_k}\frac{\partial \tilde{F}_k}{\partial o_m}\left[\left(V\frac{\Delta}{s_{ik}}(\mu_k - \xi_{ik}) + \frac{\partial L}{\partial s_{ik}}\right)\left(1 - 2\eta_{ik} - \frac{\phi}{1-\phi}\right) + \frac{\partial^2 V}{\partial \tilde{F}_k^2}\tilde{F}_k\eta_{ik} + \frac{\partial^2 L}{\partial s_{ik}^2}\frac{s_{ik}}{\tilde{F}_k}(1 - \eta_{ik})\right]$$
$$- \frac{\partial p_k}{\partial o_m}\left[\frac{\partial^2 L}{\partial \gamma_{ik}^2}f_i + \frac{\partial^2 L}{\partial s_{ik}^2}\frac{f_i}{\tilde{F}_k}(1 - \eta_{ik})\right], \tag{B.23}$$

$$\frac{\partial^2 \pi_k}{\partial f_i \partial f_m} = \frac{1}{\tilde{F}_k}\frac{\partial \tilde{F}_k}{\partial f_m}\left[\left(V\frac{\Delta}{s_{ik}}(\mu_k - \xi_{ik}) + \frac{\partial L}{\partial s_{ik}}\right)\left(1 - 2\eta_{ik} - \frac{\phi}{1-\phi}\right) + \frac{\partial^2 V}{\partial \tilde{F}_k^2}\tilde{F}_k\eta_{ik} + \frac{\partial^2 L}{\partial s_{ik}^2}\frac{s_{ik}}{\tilde{F}_k}(1 - \eta_{ik})\right]$$
$$- \frac{\partial p_k}{\partial f_m}\left[\frac{\partial^2 L}{\partial \gamma_{ik}^2}f_i + \frac{\partial^2 L}{\partial s_{ik}^2}\frac{f_i}{\tilde{F}_k}(1 - \eta_{ik})\right]. \tag{B.24}$$

Given assumptions (VF) and (LC) these two equations will be positive if $\left(V\frac{\Delta}{s_{ik}}(\mu_k - \xi_{ik}) + \frac{\partial L}{\partial s_{ik}}\right) > 0$ and $\left(1 - 2\eta_{ik} - \frac{\phi}{1-\phi}\right) > 0$. We analyze each condition in more detail next:

1. Given our assumptions on the legal cost function (LC, eq. 4) the condition $\left(V\frac{\Delta(\hat{s}_k)}{\hat{s}_k}(\mu_k - \hat{\xi}_k + \frac{\partial L}{\partial \hat{s}_k}\right) > 0 \Leftrightarrow V\frac{\Delta(\hat{s}_k)}{\hat{s}_k}(\mu_k - \hat{\xi}_k > -\frac{\partial L}{\partial \hat{s}_k}$ implies that $\mu_k > \hat{\xi}_k$. The elasticity of the value function w.r.t. additional covered patents must exceed the elasticity of the portfolio benefits function w.r.t. the share of patents held by the firm. This condition is less restrictive than the assumption in Graevenitz et al. (2013) that $\mu_k > 1$, since we are assuming that $\hat{\xi}_k < 1$.

2. $(1 - 2\hat{\eta}_k) - \frac{\hat{\phi}_k}{1-\hat{\phi}_k} > 0 \Leftrightarrow (1 - 2\hat{\phi}) > (1 - \hat{\phi})^{(N_O+1)}$. This holds for any $\hat{\phi}_k < \frac{1}{2}$ and N_o sufficiently large. These restrictions imply a setting in which the ownership of patents belonging to each opportunity is fragmented amongst many firms. It is more likely to arise if the technology is highly complex, otherwise the condition that $\hat{\phi}_k < \frac{1}{2}$ is less likely to hold.

4

In Appendix B.4 we derive the conditions under which the equilibrium of game G^* is unique. If there is a unique solution to the optimization problem of the firm at which profits are maximized, then this requires that $\partial^2 \pi_k / \partial \hat{f}^2 < 0$. The restrictions that i) $\mu_k < 1$ and ii) the share of overall profits which the firm obtains is decreasing at the margin in the share of patents the firm holds ($\partial^2 \Delta / \partial \hat{s}_k^2 < 0$) ensure that there is always such a unique interior solution.

In this game G^* the comparative statics of patenting are the same as in the main model analyzed in Graevenitz et al. (2013). Specifically we can show that the following effects hold in this game:

$$\frac{\partial^2 \pi}{\partial o_i \partial F_k} > 0, \; \frac{\partial^2 \pi}{\partial f_i \partial F_k} > 0, \; \frac{\partial^2 \pi}{\partial o_i \partial O_k} < 0, \; \frac{\partial^2 \pi}{\partial f_i \partial O_k} < 0 \tag{B.25}$$

This implies that complexity of the technology increases firms' patent applications while increased technological opportunity reduces firms' patenting applications.

B.2 Effect of Hold-up on Patenting

Here we show that Proposition 2 holds. Consider the following cross-partial derivatives for the effects of higher legal costs L due to hold-up:

$$\frac{\partial^2 \pi_k}{\partial \hat{o} \partial h_k} = -\frac{\partial L(\hat{\gamma}_k, \hat{s}_k, h_k)}{\partial h_k} < 0 \quad , \tag{B.26}$$

$$\frac{\partial^2 \pi_k}{\partial \hat{f} \partial h_k} = -\frac{\hat{o} p_k}{\tilde{F}_k} \left(\tilde{F}_k \left(\frac{\partial^2 L}{\partial \hat{\gamma}_k \partial h_k} \right) + \frac{\partial^2 L}{\partial \hat{s}_k \partial h_k} (1 - \hat{\eta}_k) \right) < 0 \quad . \tag{B.27}$$

The first of these conditions shows that the expected legal costs of hold-up reduce the number of opportunities a firm invests in, in equilibrium. The second condition shows that firms with larger portfolios will be more exposed to hold up and will benefit less from the share of patents they have patented per opportunity. Both of these effects reduce the number of facets each firm applies for.

B.3 Free Entry Equilibirum

Proposition 3

There is a free entry equilibrium at which the marginal entrant can just break even, if R&D fixed costs per opportunity (C_o) increase in the number of entrants.

In a free entry equilibrium it must be the case that the following conditions hold:

$$\pi_k(\hat{o}_k, \hat{f}_k, \hat{N}_k) > 0 \qquad \wedge \qquad \pi_k(\hat{o}_k, \hat{f}_k, \hat{N}_k + 1) < 0 \; . \tag{B.28}$$

The effect of entry on profits at the first stage of game G^* can be shown to be:

$$\frac{\partial \pi(\hat{o}, \hat{f})}{\partial N_k} = \hat{o} \frac{\partial N_O}{\partial N_k} \left(\frac{\hat{s}_k}{\tilde{F}_k} \frac{\partial \tilde{F}_k}{\partial N_O} \left[V \mu \frac{\Delta(\hat{s}_k)}{\hat{s}_k} - \left(V(\tilde{F}_k) \frac{\partial \Delta}{\partial \hat{s}_k} - \frac{\partial L}{\partial \hat{s}_k} \right) \right] \right)$$

$$+\frac{\partial p_k}{\partial N_O}\frac{\hat{f}}{\tilde{F}_k}\left[\left(V(\tilde{F}_k)\frac{\partial\Delta}{\partial\hat{s}_k}-\frac{\partial L}{\partial\hat{s}_k}\right)-\tilde{F}_k\left(\frac{\partial L}{\partial\hat{\gamma}_k}+C_a\right)\right]-\frac{\partial C_o}{\partial N_o\hat{o}}\right) .\quad\text{(B.29)}$$

This expression can be further simplified:

$$\frac{\partial\pi(\hat{o},\hat{f})}{\partial N_k} = \hat{o}\epsilon_{N_O,N}\frac{\hat{s}_k}{N}\left(\left(\epsilon_{\tilde{F}_k,N_o}-\epsilon_{p_k,N_o}\hat{\eta}_k\right)\left[V\frac{\Delta(\hat{s}_k)}{\hat{s}_k}\left(\mu_k-\hat{\xi}_k\right)+\frac{\partial L}{\partial\hat{s}_k}\right]-\frac{\partial C_o}{\partial N_o\hat{o}}\frac{N_o\hat{o}}{\hat{s}_k}\right) .\quad\text{(B.30)}$$

The first two terms in brackets in this derivative are positive and so is the third term. We can show that the limits of $\epsilon_{\tilde{F}_k,N_o}$ and $\epsilon_{\tilde{F}_k,f}$ in N_o are both zero. Therefore the above derivative is negative as long as the R&D fixed costs per opportunity are increasing in N_o. This is the condition set out in Proposition 3.

B.4 Uniqueness of the second stage equilibrium

We show that stage 2 of game G* is supermodular. This implies that there exists at least one equilibrium of the stage game. An alternative way of deriving existence of the second stage equilibrium for game G* is to analyze the conditions under which the Hessian of second derivatives of the profit function (H_π) is negative semidefinite. This matrix consists of four derivatives of which only one leads to additional restrictions on the model.

It is easy to see that $\frac{\partial^2\pi}{\partial o_i{}^2}<0$ due to the coordination costs $C_c(o_i)$ and the restrictions we impose with assumption (FVC). The two cross-partial derivatives are both zero in equilibrium - refer to equation B.15.

Therefore, the only expression that remains to analyze is $\frac{\partial^2\pi}{\partial f_i{}^2}$.

$$\frac{\partial^2\pi}{\partial f_i{}^2} = \frac{o_i p_k}{\tilde{F}_k}\left[\frac{\partial^2 V}{\partial\tilde{F}_k{}^2}\left(\frac{\partial\tilde{F}_k}{\partial f_i}\right)^2\frac{\Delta\tilde{F}_k}{p_k}+2\frac{\partial V}{\partial\tilde{F}_k}\frac{\partial\tilde{F}_k}{\partial f_i}\frac{d\Delta}{ds_{ik}}(1-\eta_{ik})+V\frac{d^2\Delta}{s_{ik}^2}\frac{p_k}{\tilde{F}_k}(1-\eta_{ik})^2\right.$$
$$\left.-\frac{\partial^2 L}{\partial\gamma_{ik}{}^2}p_k^2-\frac{\partial^2 L}{\partial s_{ik}{}^2}\left(\frac{\partial s_{ik}}{\partial f_i}\right)^2-2\left(V\frac{d\Delta}{ds_{ik}}-\frac{\partial L}{\partial s_{ik}}\right)\frac{(1-\eta_{ik})\eta_{ik}}{f_i}\right] .$$

This can be further simplified:

$$\frac{\partial^2\pi}{\partial f_i{}^2} = \frac{o_i p_k}{\tilde{F}_k}\left[\frac{\partial^2 V}{\partial\tilde{F}_k{}^2}\left(\frac{\partial\tilde{F}_k}{\partial f_i}\right)^2\frac{\Delta\tilde{F}_k}{p_k}+V\frac{d^2\Delta}{s_{ik}^2}\frac{p_k}{\tilde{F}_k}(1-\eta_{ik})^2-\frac{\partial^2 L}{\partial\gamma_{ik}{}^2}p_k^2\right.$$
$$\left.-\frac{\partial^2 L}{\partial s_{ik}{}^2}\frac{p_k}{\tilde{F}_k}(1-\eta_{ik})^2-2\left(V\frac{\Delta}{s_{ik}}\xi_{ik}(1-\mu_k)-\frac{\partial L}{\partial s_{ik}}\right)\frac{(1-\eta_{ik})\eta_{ik}}{f_i}\right] .\quad\text{(B.31)}$$

If we impose the restriction that the second derivative of the value function is negative and that the elasticity of the value function, $\mu_k<1$, then the first and the last terms in the above expression are negative. The sign of the second term in the expression depends on $sign\{\frac{\partial^2\Delta}{\partial s_{ik}{}^2}\}$, which we will assume is negative. The third and fourth terms in the above expression are negative given the conditions imposed on the legal cost function above.

6

B.5 Entry and Incumbency

In this section we analyze a game in which incumbents have lower costs of entry and demonstrate that our main predictions are robust.

We assume that a fraction λ $(0 < \lambda < 1)$ of the previously active N^P firms remain as incumbents. The firms enter until the marginal profit from entry is reduced to zero.

Objective Functions

First, consider the objective functions of incumbents and entrants and the patenting game they are involved in. We analyze this game and show when it is supermodular.

Given symmetry of technological opportunities (Assumption S) the expected value of patenting for entrant and incumbent firm's in a technology area k is:

$$\pi_{ik}^I(o_i^I, f_i^I) = o_i^I \left(V(\tilde{F}_k)\Delta(s_{ik}^I) - L(\gamma_{ik}^I, s_{ik}^I) - \left(C_o\left(\sum_{j=1}^{N^P\lambda - 1 + N^E} o_j \right) - \Psi \right) - f_i^I p_k C_a \right) - C_c(o_i^I) \quad .$$

(B.32)

$$\pi_{ik}^E(o_i^E, f_i^E) = o_i^E \left(V(\tilde{F}_k)\Delta(s_{ik}^E) - L(\gamma_{ik}^E, s_{ik}^E) - C_o\left(\sum_{j=1}^{N^P\lambda + N^E - 1} o_j \right) - f_i^E p_k C_a \right) - C_c(o_i^E) \quad .$$

(B.33)

Define a game G^E in which:

- There are $N^P\lambda$ incumbent firms and the number of entrants, N^E, is determined by free entry.

- Entrants and incumbents simultaneously choose the number of technological opportunities $o_i^I, o_i^E \in [0, O^n]$ and the number of facets applied for per opportunity $f_i^I, f_i^E \in [0, F^n]$. Firms' strategy sets S_n are elements of R^4.

- Firms' payoff functions π_{ik}, defined at (B.32,B.33), are twice continuously differentiable and depend only on rivals' aggregate strategies.

- Assumptions (VF, eqn. 1) and (LC, eqn. 4) describe how the expected value and the expected cost of patenting depend on the number of facets owned per opportunity.

Firms' payoffs depend on their rivals' aggregate strategies because the probability of obtaining a patent on a given facet is a function of all rivals' patent applications. Note that the game is symmetric within the two groups of firms as it is exchangeable in permutations of the players. This implies that symmetric equilibria exist, if the game can be shown to be supermodular (Vives, 2005).[1]

---

[1] Note also that only symmetric equilibria exist as the strategy spaces of players are completely ordered.

First order conditions for game G^E:

$$\frac{\partial \pi_{ik}^I}{\partial o_i^I} = V\Delta(s_{ik}) - L(\gamma_{ik}, s_{ik}) - \left(C_o \left(\sum_{j=1}^{N^P \lambda - 1 + N^E} o_j \right) - \Psi \right) - \gamma_{ik} C_a - \frac{\partial C_c}{\partial o_i^I} = 0 \quad , \tag{B.34}$$

$$\frac{\partial \pi_{ik}^I}{\partial f_i^I} = \frac{o_i^I p_k}{\tilde{F}_k} \left(\left[V\mu\epsilon_{\tilde{F}_k f_i} \frac{\Delta(s_{ik})}{s_{ik}} - \tilde{F}_k \left(\frac{\partial L}{\partial \gamma_{ik}} + C_a \right) \right] + \left[V\frac{\partial\Delta(s_{ik})}{\partial s_{ik}} - \frac{\partial L}{\partial s_{ik}} \right] (1 - \eta_{ik}) \right) = 0 \quad , \tag{B.35}$$

$$\frac{\partial \pi_{ik}^E}{\partial o_i^E} = V\Delta(s_{ik}) - L(\gamma_{ik}, s_{ik}) - C_o \left(\sum_{j=1}^{N^P \lambda + N^E - 1} o_j \right) - \gamma_{ik} C_a - \frac{\partial C_c}{\partial o_i^E} = 0 \quad , \tag{B.36}$$

$$\frac{\partial \pi_{ik}^E}{\partial f_i^E} = \frac{o_i^E p_k}{\tilde{F}_k} \left(\left[V\mu\epsilon_{\tilde{F}_k f_i} \frac{\Delta(s_{ik})}{s_{ik}} - \tilde{F}_k \left(\frac{\partial L}{\partial \gamma_{ik}} + C_a \right) \right] + \left[V\frac{\partial\Delta(s_{ik})}{\partial s_{ik}} - \frac{\partial L}{\partial s_{ik}} \right] (1 - \eta_{ik}) \right) = 0 \quad . \tag{B.37}$$

Proposition 7

In game G^E the equilibrium number of facets chosen by incumbents and entrants is the same: $\hat{f}^I = \hat{f}^E$.

We show in Appendix A.2 that in the game with incumbents the number of rivals per opportunity $\tilde{N}_O$ becomes a function of both $\hat{o}^I, \hat{o}^E$. The first order conditions determining $\hat{f}^I, \hat{f}^E$ both depend on the total number of entrants per technological opportunity $\tilde{N}_O$ and so both on $\hat{o}^I, \hat{o}^E$. This is the only way in which rivals' choices of the number of opportunities to pursue enter these first order conditions[2]. Therefore the two conditions are identical and Proposition 7 holds.

Proposition 8

The second stage of game G^E is smooth supermodular under the same conditions as game G^*. Comparative statics results for game G^* also apply to game G^E.

The first order conditions characterizing the game with incumbents and entrants are identical to those for the game without incumbents as long as $\Psi = 0$. As this variable is a constant it does not enter into the second order conditions which we analyze to establish supermodularity and which underpin the comparative statics predictions in Propositions 4-6.

Proposition 9

In the second stage of game G^E incumbents enter more technological opportunities, if they have a cost advantage in undertaking R&D ($\Psi > 0$).

The first order conditions determining the equilibrium number of opportunities chosen by incumbents and entrants are identical if firms R&D fixed costs per opportunity are the same ($\Psi = 0$). Therefore $\hat{o}_{|\Psi=0}^I = \hat{o}^E$. As the cost advantage of incumbents in undertaking R&D grows this increases the number of opportunities chosen by incumbents:

$$\frac{\partial^2 \pi_{ik}^I}{\partial o^I \partial \Psi} = 1 > 0 \quad . \tag{B.38}$$

---

[2]Clearly the factors outside the brackets in equations (B.35), (B.37) also depend on these variables, but these do not affect the equilibrium values of f_i^E, f_j^I.

Proposition 10

In the second stage of game G^E the number of entrants decreases as the cost advantage of incumbents increases.

Due to the supermodularity of the second stage game, increases in incumbents' choices of the number of opportunities to invest in will raise the number of opportunities entrants invest in as well as the numbers of facets entrants and incumbents seek to patent in equilibrium. The increases in $\hat{o}^I$ and $\hat{o}^E$ will raise the fixed costs of entry into new opportunities, C_o, which then reduces entry.

C   Data

Our analysis relies on an updated version of the Oxford-Firm-Level-Database, which combines information on patents (UK and EPO) with firm-level information obtained from Bureau van Dijks Financial Analysis Made Easy (FAME) database (for more details see Helmers et al. (2011) from which the data description in this section draws).

The integrated database consists of two components: a firm-level data set and IP data. The firm-level data is the FAME database that covers the entire population of registered UK firms. The original version of the database, which formed the basis for the update carried out by the UKIPO, relied on two versions of the FAME database: FAME October 2005 and March 2009. The main motivation for using two different versions of FAME is that FAME keeps details of inactive firms (see below) for a period of four years. If only the 2009 version of FAME were used, intellectual property could not be allocated to any firm that has exited the market before 2005, which would bias the matching results. FAME is available since 2000, which defines the earliest year for which the integrated data set can be constructed consistently. The update undertaken by the UKIPO used the April 2011 version of FAME. However, since there are significant reporting delays by companies, even using the FAME 2011 version means that the latest year for which firm-level data can be used reliably is 2009.

FAME contains basic information on all firms, such as name, registered address, firm type, industry code, as well as entry and exit dates. Availability of financial information varies substantially across firms. In the UK, the smallest firms are legally required to report only very basic balance sheet information (shareholders' funds and total assets). The largest firms provide a much broader range of profit and loss information, as well as detailed balance sheet data including overseas turnover. Lack of these kinds of data for small and medium-sized firms means that our study focuses on total assets as a measure of firm size and growth.

The patent data come from the EPO Worldwide Patent Statistical Database (PATSTAT). Data on UK and EPO patent publications by British entities were downloaded from PATSTAT version April 2011. Due to the average 18 months delay between the filing and publication date of a patent, using the April 2011 version means that the patent data are presumably only complete up to the third quarter in 2009. This effectively means that we can use the patent data only up to 2009 under the caveat that it might be somewhat incomplete for 2009. Patent data are allocated to firms by the year in which a firm applied for the patent.

Since patent records do not include any kind of registered number of a company, it is not possible to

merge data sets using a unique firm identifier; instead, applicant names in the IP documents and firm names in FAME have to be matched. Both a firm's current and previous name(s) were used for matching in order to account for changes in firm names. Matching on the basis of company names requires names in both data sets to be 'standardized' prior to the matching process in order to ensure that small (but often systematic) differences in the way names are recorded in the two data sets do not impede the correct matching. For more details on the matching see Helmers et al. (2011).

Table C-1: Number of TF34 sectors entered between 2002 and 2009

Number of sectors	Number of firms	Number of entries
1	2,531	2,531
2	1,347	2,694
3	647	1,941
4	271	1,084
5	155	775
6	71	426
7	45	315
8	29	232
9	20	180
10	14	140
11	4	44
12	2	24
13	3	39
14	0	0
15 or more	13	240
Total	**5,152**	**10,665**

Table C-2: Sample Population of Firms, by industry

	2-digit SIC[3] Industry	Number of firms	Number of patenters	Share patenting 2001-2009	Number of patents
1	Basic metals	2,836	52	1.83%	231
2	Chemicals	3,834	246	6.42%	126
3	Electrical machinery	2,948	281	9.53%	727
4	Electronics & instruments	9,298	561	6.03%	444
5	Fabricated metals	24,681	606	2.46%	70
6	Food, beverage, & tobacco	8331	102	1.22%	29
7	Machinery	9,365	608	6.49%	313
8	Mining, oil&gas	83,491	15	0.02%	96
9	Motor vehicles	2,337	117	5.01%	22
10	Other manufacturing	94,952	1362	1.43%	150
11	Pharmaceuticals	1,008	105	10.42%	551
12	Rubber & plastics	6,094	398	6.53%	590
13	Construction	295596	372	0.13%	59
14	Other transport	3,292	89	2.70%	6274
15	Repairs & retail trade	128,266	251	0.20%	2324
16	Telecommunications	14,348	133	0.93%	2096
17	Transportation	60,837	75	0.12%	621
18	Utilities	12,880	75	0.58%	428
19	Wholesale trade	138,398	728	0.53%	3608
20	Business services	689,942	1639	0.24%	6757
21	Computer services	177,319	716	0.40%	1132
22	Financial services	183,042	219	0.12%	4930
23	Medicalservices	38,424	103	0.27%	1419
24	Personal services	94,791	196	0.21%	1194
25	R&D services	7,915	713	9.01%	168
26	inactive	37,525	271	0.72%	5673
	Total	**2,131,750**	**10,033**	**0.47%**	**40,032**

Table C-3: Entry into techology area 2002-2009

	Numbers			Shares		
Technology	Total patenting in sector by GB firms	First time patenter	Pat'd previously in another tech.	Total entry	First time patenter	Pat'd previously in another tech
Elec machinery, energy	1763.7	214	250	26.3%	12.1%	14.2%
Audio-visual tech	788.3	148	192	43.1%	18.8%	24.4%
Telecommunications	1874.4	146	179	17.3%	7.8%	9.5%
Digital communication	1054.0	92	141	22.1%	8.7%	13.4%
Basic comm processes	256.5	21	93	44.4%	8.2%	36.3%
Computer technology	2167.2	291	251	25.0%	13.4%	11.6%
IT methods for mgt	283.2	117	157	96.7%	41.3%	55.4%
Semiconductors	347.2	38	118	44.9%	10.9%	34.0%
Optics	584.7	55	130	31.6%	9.4%	22.2%
Measurement	1765.3	226	269	28.0%	12.8%	15.2%
Analysis bio materials	339.0	39	111	44.3%	11.5%	32.7%
Control	712.1	165	241	57.0%	23.2%	33.8%
Medical technology	1668.4	184	209	23.6%	11.0%	12.5%
Organic fine chemistry	1569.4	36	83	7.6%	2.3%	5.3%
Biotechnology	701.6	41	99	20.0%	5.8%	14.1%
Pharmaceuticals	1700.7	54	80	7.9%	3.2%	4.7%
Polymers	224.4	27	112	61.9%	12.0%	49.9%
Food chemistry	492.7	36	87	25.0%	7.3%	17.7%
Basic materials chemistry	1020.2	85	144	22.4%	8.3%	14.1%
Materials metallurgy	360.8	54	109	45.2%	15.0%	30.2%
Surface tech coating	400.7	77	195	67.9%	19.2%	48.7%
Chemical engineering	842.7	142	213	42.1%	16.9%	25.3%
Environmental tech	446.6	106	166	60.9%	23.7%	37.2%
Handling	1326.3	274	290	42.5%	20.7%	21.9%
Machine tools	577.8	106	180	49.5%	18.3%	31.2%
Engines,pumps,turbine	1443.4	82	160	16.8%	5.7%	11.1%
Textile and paper mach	442.0	77	137	48.4%	17.4%	31.0%
Other spec machines	847.4	180	234	48.9%	21.2%	27.6%
Thermal process and app	455.7	105	159	57.9%	23.0%	34.9%
Mechanical elements	1445.9	223	317	37.3%	15.4%	21.9%
Transport	1288.2	213	236	34.9%	16.5%	18.3%
Furniture, games	1239.2	288	223	41.2%	23.2%	18.0%
Other consumer goods	788.1	194	246	55.8%	24.6%	31.2%
Civil engineering	2045.8	463	255	35.1%	22.6%	12.5%
Total	**33263.6**	**4599**	**6066**	**32.1%**	**13.8%**	**18.2%**

D   Robustness Tests on Network Density and Triples

This section examines how the addition of the network density measure to the analysis undertaken in Graevenitz et al. (2013) changes the coefficient and sign of the triples measure reported there. They examine how complexity, technological opportunity and other variables affect the number of patents firms apply for. Graevenitz et al. (2013) do not distinguish between hold-up and complexity in their model and use the triples measure to capture complexity. The model presented in this paper separates the effects of hold-up and complexity and predicts that hold-up will reduce firms' patenting incentives, while complexity raises these. In this paper, the network density measure is introduced as a measure of technological complexity, while we argue that the triples measure captures hold-up.

The exercise undertaken in this appendix is a validation of these two measures in light of the updated model we present in this paper. The evidence provided is based on two data sets: first we report regression results obtained by adding the network density measure to the data used by (Graevenitz et al., 2013), second we report results obtained from a new dataset. This dataset covers the same period as that used by (Graevenitz et al., 2013), but it is based on the same more recent technology area classification as that used in this paper. Furthermore, it is based on the same measure of triples as that used in this paper.

The models presented below are system GMM models which include a lagged dependent variable. We demonstrate that regardless of how the data are constructed the triples measure reduces patenting while the network density measure increases patenting efforts in our data. This supports our view that the triples measure is a measure of hold-up.

In the results presented below we instrument potentially endogenous variables using lagged values. Exogeneity of the instruments is tested using difference in Hansen tests. We instrument the lagged dependent variable and its interaction with fourth order lags. All other variables are instrumented with third order lags or higher. We include only year and area dummies in the levels equations as it is likely that the fixed effects are correlated with differences in the remaining explanatory variables.

Instrument sets are collapsed in order to reduce the number of instruments used. Throughout we rely on the Hansen test to determine whether instruments are exogenous. Where the statistic indicated that this was not the case we rejected the models. We report only those models that were not rejected by the test for which the lagged dependent variable was within the range one would expect from estimation of OLS models with the same specification.

D.1   Sample and Definition of Variables

The sample used for both tables below consists of all firms that have at least one hundred patent applications at EPO across all teachnology areas between 1987 and 2002 and who have applied for patents in at least three years in the sample period in a technology area.

The two tables below include a number of variables that we do not use in this paper other than here. We briefly discuss these variables next:

Dependent variable   In both tables below the dependent variable is the logarithm of the number of patents each firm has applied for in a technology area and year. To deal with missing values arising

from firms not having patent applications in some years we add one to all patent counts before taking the logarithm.

Triples count　　In Table D-1 below we use the triples count employed by Graevenitz et al. (2013). They count how often firm triples arise, such that each firm in a triple holds patents that are cited as limiting one or more patent applications submitted by each of the other two firms. Their measure of triples is constructed using only the ten most frequently cited firms in each applicant's patent portfolio in any area and year. In Table D-2 we use the same triples count as in this paper, i.e. the restriction to the most frequently cited firms is removed.

Fragmentation　　The fragmentation measure used here is based on Ziedonis (2004). The measure is based only on critical references and captures the concentration of prior art cited in the patent portfolio of a firm in a year and area.

Large / Relative Size　　In Table D-1 below we use a dummy variable that is one for all firms above the median firm by size of patent portfolio in each area and year. In Table D-2 we capture relative size by measuring the size of each firm's patent portfolio by area up to a given year relative to the total number of all firms' patent applications in that area and year.

D.2　Data used by Graevenitz et al. (2013)

The results presented here are based on adding the network density measure of complexity constructed for this paper from US patent data to the data used by Graevenitz et al. (2013). They study the determinants of the level of patent applications at EPO.

Column E in Table D-1 below is reported by Graevenitz et al. (2013) and is presented here as a reference point. Models K,L,M contain new results. Model K replicates model E closely, differences are likely due to updates to the code used to estimate these models. Model L is like model K with the network density measure added. In model M we adjust the set of instruments to obtain a model with a lagged dependent variable that is significantly below 1 at the 5% level at the mean of the triples variable.

These results show that adding the network density measure to the data does not change the sign or significance of the other variables reported in Table D-1. Network density itself has a positive and significant effect on patenting in models L and M. We would expect to see this, if this measure captures complexity.

Adding network density does have an important and not immediately obvious effect. The range of values of non-patent references for which an increase in the hold-up measure (triples) reduces patenting incentives is larger in model M than model E at the mean of the patent count: in model E non-patent references must lie beyond 1.22 for an increase in triples to have a negative effect on patenting incentives, while in model M non-patent references beyond 1.1 have the same effect. Similarly the range of values of the patent count for which an increase in the hold-up measure (triples) reduces patenting incentives is larger in model M than model E at the mean of non patent references. This shows that the triples measure

contained in this data has a negative effect on patenting incentives in sufficiently complex technologies. We would expect to see this, if triples is a measure of hold-up.

Table D-1: GMM Models for Patent Applications - Old Data

Variable	SGMM **E**	SGMM **K**	SGMM **L**	SGMM **M**
log Patentcount$_{t-1}$	0.749***	0.879***	1.020***	0.976***
	(0.093)	(0.099)	(0.113)	(0.106)
log Patentcount$_{t-1}$ × Triples	−0.017***	−0.017***	−0.013***	−0.012***
	(0.003)	(0.002)	(0.002)	(0.002)
Non Patent References (NPR)	1.553***	1.863***	1.648***	1.389***
	(0.254)	(0.252)	(0.240)	(0.183)
NPR × Triples	−0.036***	−0.038***	−0.033***	−0.028***
	(0.006)	(0.005)	(0.005)	(0.004)
NPR × Triples × Large	0.007***	0.006***	0.005***	0.005***
	(0.002)	(0.001)	(0.001)	(0.001)
NPR × Large	−0.366***	−0.386***	−0.339***	−0.340***
	(0.081)	(0.062)	(0.051)	(0.043)
Fragmentation	−0.474**	−0.521**	−0.543**	−0.490***
	(0.170)	(0.182)	(0.174)	(0.129)
Fragmentation × Triples	0.006	0.009*	0.005	0.004
	(0.006)	(0.004)	(0.004)	(0.004)
Triples	0.055***	0.052***	0.046***	0.039***
	(0.010)	(0.007)	(0.007)	(0.005)
Areas	0.096***	0.084***	0.049**	0.050**
	(0.012)	(0.010)	(0.017)	(0.016)
Large	0.342**	0.476***	0.427***	0.424***
	(0.117)	(0.105)	(0.091)	(0.072)
Network Density			0.002**	0.002**
			(0.001)	(0.001)
Year dummies	YES	YES	YES	YES
Primary area dummies	YES	YES	YES	YES
Constant	−1.443***	−1.846***	−2.017***	−1.772***
	(0.319)	(0.267)	(0.270)	(0.237)
N	173448	173448	173448	173448
m1	−10.860	−10.470	−9.318	−9.634
m2	4.739	6.307	6.2	6.305
m3	.896	.509	−.071	−.302
Hansen	10.988	2.017	4.454	12.506
p-value	.052	.569	.216	.052
Degrees of freedom	5	3	3	6

* p<0.05, ** p<0.01, *** p<0.001

1. Asymptotic standard errors, asymptotically robust to heteroskedasticity are reported in parentheses

2. m1-m3 are tests for first- to third-order serial correlation in the first differenced residuals.

3. Hansen is a test of overidentifying restrictions. It is distributed as χ^2 under the null of instrument validity, with degrees of freedom reported below.

4. In all cases GMM instrument sets were collapsed and lags were limited.

D.3   New Data at 34 Area Level

Here we present results based on an updated dataset of patenting in Europe that is based on PATSTAT, October 2014, but covers the same range of years (1987-2002), for better comparability with the data presented in the previous section.

Table D-2: GMM Models for Patent Applications - New Data

Variable	SGMM E	SGMM K	SGMM L
$\log$ Patentcount$_{t-1}$	$0.454***$	$0.799***$	$0.808***$
	(0.111)	(0.080)	(0.082)
Non Patent References (NPR)	$-0.014***$	$-0.014***$	$-0.014***$
	(0.001)	(0.001)	(0.001)
Triples	$-0.091***$	$-0.100***$	$-0.117***$
	(0.016)	(0.011)	(0.012)
Network Density	-0.216	0.195	$0.217*$
	(0.129)	(0.101)	(0.107)
Fragmentation	$0.899***$	$0.658***$	$0.804***$
	(0.155)	(0.093)	(0.085)
Areas	$0.020*$	$0.016**$	-0.001
	(0.009)	(0.006)	(0.006)
Relative Size	$-0.010***$	0.000	-0.000
	(0.003)	(0.003)	(0.002)
Constant	-0.033	$-0.208***$	$-0.149**$
	(0.045)	(0.052)	(0.046)
N	168066	168066	168066
m1	-7.598	-12.948	-13.08
m2	3.871	9.964	9.896
m3	-1.450	$.538$	$.055$
Hansen	3.49	3.79	2.1
p-value	$.48$	$.29$	$.55$
Degrees of freedom	4	3	3

* p<0.05, ** p<0.01, *** p<0.001

1. Asymptotic standard errors, asymptotically robust to heteroskedasticity are reported in parentheses

2. m1-m3 are tests for first- to third-order serial correlation in the first differenced residuals.

3. Hansen is a test of overidentifying restrictions. It is distributed as χ^2 under the null of instrument

validity, with degrees of freedom reported below.

4. In all cases GMM instrument sets were collapsed and lags were limited.

The main difference between the data used here and the older data used by Graevenitz et al. (2013) is that we now rely on a more recent, slightly finer specification of the number of technology areas: the current classification contains 34 rather than 30 areas. In addition, the triples measure we use now captures all triples and not just those affecting each firm and its ten closest technology rivals as in the earlier data. Due to the larger number of areas we now exclude slightly more patentees when applying the criterion that a firm must have at least one hundred patents in a technology area and must have at least three years of patent activity in an area to be included in the analysis.

Table D-2 demonstrates that the predicted negative effects of triples and non-patent references are present in all specifications we report. We also show that network density is either not significant or positive and significant. The model in which the measure is positive and significant is our prefered model, due to the low instrument count and the better test of overidentifying restrictions.

Overall, both sets of models demonstrate that an interpretation of the triples measure as a measure of hold-up only, and of the network density as a measure of complexity, is consistent with the effects we observe in the data on patent applications presented here.

E   Estimating Survival Models

This appendix gives some further information about the various survival models we estimated and the robustness checks that were performed. We estimated two general classes of failure or survival models: 1) proportional hazard, where the hazard of failure over time has the same shape for all firms, but the overall level is proportional to an index that depends on firm characteristics; and 2) accelerated failure time, where the survival rate is accelerated or decelerated by the characteristics of the firm. We transform (2) to a hazard rate model for comparison with (1), using the usual identity between the probability of survival to time t and the probability of failure at t given survival to $t-1$. The first model has the following form:

$$Pr\left(i \text{ first patents in } j \text{ at } t \,|\, i \text{ has no patents in } j \,\forall s < t, X_i\right)$$

$$h(X_i, t) = h(t) \exp\left(X_i, \beta\right)$$

where i denotes a firm, j denotes a technology sector, and t denotes the time since entry into the sample. $h(t)$ is the baseline hazard, which is either a non-parametric or a parametric function of time since entry into the sample. The impact of any characteristic x on the hazard can be computed as follows:

$$\frac{\partial h(X_i, t)}{\partial x_i} = h(t) \exp\left(X_i, \beta\right)\beta \quad \text{or} \quad \frac{\partial h(X_i, t)}{\partial x_i} \frac{1}{X_i, t} = \beta \tag{E.39}$$

Thus if x is measured in logs, β measures the elasticity of the hazard rate with respect to x. Note that this quantity does not depend on the baseline hazard $h(t)$, but is the same for any t. We use two choices for

$h(t)$: the semi-parametric Cox estimate and the Weibull distribution pt^{p-1}. By allowing the Cox $h(t)$ or p to vary freely across the industrial sectors, we can allow the shape of the hazard function to be different for different industries while retaining the proportionality assumption.

In order to allow even more flexibility across the different industrial sectors, we also use two accelerated failure time models, the log-normal model and the log-logistic model. These have the following basic form:

$$\text{Log-normal} : S(t) = 1 - \Phi \left\{ \frac{\ln(t) - \mu}{\sigma} \right\} \tag{E.40}$$

$$\text{Log-logistic} : S(t) = \frac{1}{1 + (\lambda t)^{\frac{1}{\gamma}}} \tag{E.41}$$

where $S(t)$ is the survival function and $\lambda_i = \exp(X_i \beta)$. We allow the parameters σ (log-normal) or γ (log-logistic) to vary freely across industries (j). That is, for these models, both the mean and the variance of the survival distribution are specific to the 2-digit industry. In the case of these two models, the elasticity of the hazard with respect to a characteristic x depends on time and on the industry-specific parameter (σ or γ), yielding a more flexible model. For example, the hazard rate for the log-logistic model is given by the following expression:

$$h(t) = \frac{-d \log S(t)}{dt} = \frac{\lambda^{\frac{1}{\gamma}} t^{\frac{1}{\gamma} - 1}}{\gamma \left(1 + (\lambda t)^{\frac{1}{\gamma}} \right)} \tag{E.42}$$

From this we can derive the elasticity of the hazard rate with respect to a regressor x:[4]

$$\frac{\partial \log h_{ij}(t)}{\partial x_i} = \frac{-\beta}{\left(1 + (\lambda t)^{\frac{1}{\gamma}} \right)} \tag{E.43}$$

One implication of this model is therefore that both the hazard and the elasticity of the hazard with respect to the regressors depend on t, the time since the firm was at risk of patenting. We sample the firms during a single decade, the 2000s, but some of the firms have been in existence since the 19th century. This fact creates a bit of a problem for estimation, because there is no reason to think that the patenting environment has remained stable during that period. We explored variations in the assumed first date at risk in Tables E-1(1978) and E-2 (1900), finding that the choice made little difference. Accordingly, we have used a minimum at risk year of 1978 for estimation in the main table in the text.

---

[4]We assume that x is in logarithms, as is true for our key variables, so this can be interepreted as an elasticity.

Table E-1
Hazard of entry into patenting in a TF34 Class - Comparing models
538,452 firm-TF34 observations with 10,665 entries (20,384 firms)

Variable	**Proportional hazard**		**AFT**	
	Cox PH	Weibull	Log logistic	Log normal
Log (network density)	$0.108***$	$0.111***$	$0.308***$	$0.247***$
	(0.021)	(0.021)	(0.096)	(0.040)
Log (triples density	$-0.100***$	$-0.098***$	$-0.511***$	$-0.258***$
in class)	(0.010)	(0.010)	(0.071)	(0.024)
Log (patents in class)	$0.513***$	$0.513***$	$2.177***$	$1.095***$
	(0.027)	(0.027)	(0.304)	(0.089)
5-year growth of non-patent	(0.013)	-0.001	-0.077	-0.039
refs in class	(0.022)	(0.021)	(0.084)	(0.038)
Log assets	$0.149***$	$0.130***$	$0.529***$	$0.198***$
	(0.013)	(0.013)	(0.082)	(0.024)
Log (pats applied for	$0.848***$	$0.860***$	$5.685***$	$3.973***$
by firm previously)	(0.021)	(0.019)	(0.917)	(0.368)
Industry dummies	stratified	stratified	stratified	stratified
Year dummies	yes	yes	yes	yes
Log likelihood	-58.8	$-96,051.0$	$-114,127.0$	$-111,662.1$
Degrees of freedom	13	38	38	38
Chi-squared	3183.2	4560.1	168.4	300.1

All estimates are weighted estimates, weighted by sampling probability. For the Cox and Weibull

models, coefficients shown are elasticities of the hazard w.r.t. the variable. For the log-logistic, -beta is shown.

*** (**) denote significance at the 1% (5%) level.

Time period is 2002-2009 and minimum entry year is 1978. Sample is all UK firms with nonmissing assets.

AFT - Accelerated Failure Time models

\# Estimates are stratified by industry - each industry has its own baseline hazard.

Table E-2

Hazard of entry into patenting in a TF34 Class - Robustness

Variable	(1)	(2)	(3)	(4)	(5)	Big firms
Log (network density)	0.107***	−0.008	0.114***	0.107***	0.108***	0.390***
	(0.021)	(0.019)	(0.024)	(0.021)	(0.021)	(0.145)
Log (triples density	−0.101***	−0.183***	−0.114***	−0.103***	−0.103***	−0.099*
in class)	(0.010)	(0.008)	(0.011)	(0.009)	(0.009)	(0.055)
Log (patents in class)	0.514***	0.623***	0.605***	0.520***	0.518***	0.390**
	(0.027)	(0.024)	(0.031)	(0.027)	(0.027)	(0.157)
5-year growth of non-	−0.009	−0.196***	−0.002	−0.012	−0.012	−0.125
patent refs in class	(0.021)	(0.020)	(0.025)	(0.021)	(0.021)	(0.135)
Log assets	0.142***	0.138***	0.186***	0.139***	0.146***	0.303
	(0.013)	(0.013)	(0.018)	(0.013)	(0.013)	(0.233)
Log firm age in years	0.773***	0.588***	0.739***	0.778***	0.021	0.710
	(0.130)	(0.145)	(0.155)	(0.131)	(0.246)	(0.672)
Log (lagged firm-level)	0.836***	0.947***	0.954***	0.840***	0.878***	0.512***
patent stock	(0.021)	(0.019)	(0.033)	(0.021)	(0.021)	(0.132)
Industry dummies	yes	yes	yes	yes	yes	yes
Observations	538,452	692,038	452,313	523,547	538,452	5,655
Firms	20,384	20,384	17,993	20,384	20,384	255
Entries	10,665	10,665	8,149	10,340	10,665	219
Entry rate	1.98%	1.54%	1.96%	1.97%	1.98%	3.87%
Log likelihood	−58.69	−54.60	−40.59	−56.94	−53.81	−0.21
Degrees of freedom	14	14	14	14	14	14
Chi-squared	3465.1	5065.8	1688.8	3459.1	3137.4	51.3

All estimates are weighted estimates, weighted by sampling probability. Coefficients shown are negative of the estimates (larger coefficient increases entry probability).

*** (**) denote significance at the 1% (5%) level.

Time period is 2002-2009. Sample in (1) is all UK firms with nonmissing assets.

Log-logistic model stratified by industry.

(1) Estimates from Table 3, for comparison.

(2) Observations for tech sectors of firms whose industry has no such patenting (Lybbert and Zolas, 2014) and where there is no entry by any UK firm in that industry are dropped.

(3) SMEs: firms with assets > 12.5 million GBP removed.

(4) The Telecom tech sector is removed.

(5) The minimum founding year is 1900 instead of 1978.

References

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, 521–563.

HELMERS, C., M. ROGERS, AND P. SCHAUTSCHICK (2011): "Intellectual Property at the Firm-level in the UK: The Oxford Firm-level Intellectual Property Database," Tech. Rep. DP 546.

LYBBERT, T. J. AND N. J. ZOLAS (2014): "Getting Patents and Economic Data to Speak to each other: An Algorithmic Links with Probabilities Approach for Joint Analyses of Patenting and Economic Activity," *Research Policy*, 43, 530–542.

VIVES, X. (2005): "Complementarities and Games: New Developments," *Journal of Economic Literature*, XLIII, 437–479.

ZIEDONIS, R. H. (2004): "Don't Fence Me In: Fragmented Markets for Technology and the Patent Acquisition Strategies of Firms," *Management Science*, 50, 804–820.