



Bavarian Graduate Program in Economics

**BGPE Discussion Paper**

**No. 117**

**The role of patents and secrecy for  
intellectual property protection:  
theory and evidence**

**Malte Mosel**

April 2012

---

ISSN 1863-5733

Editor: Prof. Regina T. Riphahn, Ph.D.  
Friedrich-Alexander-University Erlangen-Nuremberg  
© Malte Mosel

---

# The role of patents and secrecy for intellectual property protection: theory and evidence\*

Malte Mosel<sup>†</sup>

## Abstract

Traditionally patents are seen as the gold standard for intellectual property protection. But, in line with empirical findings that secrecy is considered more important for appropriating returns, recent theories predict that firms keep their most important inventions secret. This article reconciles both opposing views in a unifying framework, accounting for the main aspects causing the discrepancy: imperfect protection, patenting costs, and simultaneous innovations. Theoretical results on the relation between patenting and innovation size are then confronted with survey data for small European firms. Using a binary size measure, we find strong support for the traditional view that firms patent their most important innovations, but a continuous size measure reveals an inverted-U relation between patenting and size, as predicted by the unifying framework.

*Keywords:* Filing fees · imitation · innovation · probabilistic patent rights · R&D · simultaneous innovation · trade secrets

*JEL-Classifications:* D22 · D23 · K11 · L16 · O34

---

\*I would like to thank Michael Pflüger, Johann Graf Lambsdorff, Dietmar Harhoff, Bohdan Kukharskyy, Philipp Ehrl, and participants of the 13th BGPE Research Workshop in Munich for helpful comments.

<sup>†</sup>Department of Economics, University of Passau, Innstraße 27, D-94032 Passau, malte.mosel@uni-passau.de, and Ludwig Maximilians University, Munich School of Management, INNO-tec, Kaulbachstraße 45, D-80539 Munich, mosel@bwl.lmu.de.

# 1 Introduction

To patent, or not to patent: that is the question arising with each new invention. Patents grant temporary exclusive rights to use an invention but require disclosure of its technical details. By relying on secrecy instead a firm can avoid compulsory disclosure but, in return, forfeits all legal grounds to prevent rivals from using the invention. Traditionally patents are seen as the key mechanism for intellectual property (IP) protection. Friedman et al. (1991: 63), e.g., summarize this view by stating that “*no rational person with a patentable invention would fail to seek a patent*”, while more recently Arundel (2001: 612) observed a “*widespread [...] bias on the part of some policy makers, economists, and jurists in favour of patents.*”<sup>1</sup> However, well-known studies find that, first, only a fraction of inventions are patented and, second, firms typically regard secrecy to be more effective for appropriation than patents.<sup>2</sup> These two facts call the ‘traditional view’ that patents dominate IP protection into question.

In their seminal work, Anton and Yao (2004) theoretically substantiate both empirical facts. They assume that inventions are heterogeneous in size, i.e., they differ regarding the extent of the technological step forward induced by them, and account for the fact that the validity of patent rights is unclear until challenged in court (‘probabilistic patent rights’). Both aspects drive the strategic decision of rivals whether or not to imitate a patented invention. Large inventions are imitated, as the prospect of making up for a large technological lead justifies risking patent infringement. For small inventions, however, the protected technological lead is too small given the infringement risk, so that no rival dares to use the information disclosed in the patent. The resulting ‘imitation deterring effect’ implies that a patent provides full protection for inventions small enough to bar rivals from imitating, while a patent protects relatively large inventions only to the extent that it withstands a challenge in court. As a result, Anton and Yao show that firms ideally patent small inventions and keep their most important ones secret. While consistent with the data, this ‘strategic view’ contradicts the traditional view that patents are the key mechanism for IP protection.

The paper presented here yet shows that this is not the end of the story. The traditional view can be brought in line with both empirical facts by considering two aspects Anton and Yao’s approach did not account for: patenting costs and the possibility that more than one firm simultaneously makes the same invention. The former aspect includes filing and renewal fees, whose significance for patenting behavior has been shown, for instance, by Pottelsberghe and Francois (2009) and Baudry and Dumont (2009). The latter aspect, brought forward by Kultti et al. (2006) and (2007), accounts for the fact that patents provide protection

---

<sup>1</sup> See for the traditional view also Nordhaus (1969), Scherer (1972), Tandon (1982), and Scotchmer (1991).

<sup>2</sup> See, e.g., Arundel/Kabla (1998), Brouwer/Kleinknecht (1999), Cohen et al. (2000), and Arundel (2001).

in situations where secrecy fails, i.e., if more than one firm discovers the same invention. Patents can thus make up for the finding that they provide less efficient protection than secrecy for a given technological lead by offering a higher probability to obtain a lead in the first place. While patenting costs may impede the use of patents, accounting for the possibility of simultaneous innovations increases their relative attractiveness as means of protection, since patents, unlike secrecy, are able to protect an invention also discovered by rivals. Based on these aspects, it can be shown that firms patent their most important inventions to ensure protection in case of simultaneous innovations and prefer secrecy for small inventions, where patenting costs relative to returns are high.<sup>3</sup> While in line with the empirical evidence outlined above, these theoretical results reestablish the traditional view and contradict the strategic view that firms keep their most important inventions secret. Answering the question which view prevails thus requires new empirical evidence.

The goal of this paper is to test the empirical validity of traditional and strategic view and, thereby, to explore the relation between propensity to patent – the probability that an invention is patented – and innovation size. To this end, we set up a simple theoretical model with heterogeneous inventions which incorporates both views as special cases and yet allows us to analyze a combination of them in an integrative model.<sup>4</sup> We focus on the key assumptions of traditional and strategic view – simultaneous innovations, patenting costs, and the imitation deterring effect of probabilistic patent rights – to replicate their predictions in the simplest set-up possible. As a result, in the case with (without) patenting costs but without (with) the imitation deterring effect, the model finds predictions consistent with the traditional (strategic) view. When including all key assumptions, the integrative model yields results which corroborate the traditional view, if we assume relatively strong patent rights manifesting in a high probability that the patent withstands a challenge in court. Yet, for relatively weak patent rights, the integrative model reveals a third, independent result: an inverted-U relation between patenting and innovation size. Intuitively, the inverted-U results from the fact that firms eschew patenting costs for very small inventions, while knowledge disclosure and the absence of the imitation deterring effect keep them from patenting large inventions. Only medium-sized inventions are patented to benefit from the higher protection patents provide in the simultaneous innovation case.

Theoretical predictions are then tested using European firm survey data extracted from the 2004 Community Innovation Survey (CIS 4) provided by Eurostat. For product innovations, it contains self-reported measures for innovation size, which are ideal to map the

---

<sup>3</sup> See also Mosel (2011b), who confirms empirical validity of the traditional view in a more general model.

<sup>4</sup> The model is a simple version of Mosel (2011b). It refrains from the intricacies involved in ensuring full empirical plausibility regarding the ‘stylized facts of patenting behavior’ outlined in Mosel (2011b).

heterogeneity of inventions assumed in the underlying theories. A binary measure of innovation size results from the information about whether within the two-year observation period the respondent introduced an innovation that is only ‘new to the firm’ or also ‘new to the market’. Besides, the data set provides the fraction of the respondent’s total turnover generated by an invention, which can be used as a continuous size measure, suitable to test the non-linear prediction of an inverted-U relation between patenting and size.

Our data set yet implies two limitations. First, unlike some small CIS 4 subsamples, it only contains information about patenting but lacks information about the use of secrecy. This can be problematic as long as secrecy and patenting are not mutually exclusive. The inherent nature and legal implementation of both mechanisms, however, justify the assumption of mutual exclusivity. A valid trade secret requires ‘reasonable measures’ to keep an invention secret, so that, once it has been patented, compulsory disclosure precludes secrecy as a complementary option.<sup>5</sup> For the opposite case, the legal situation is less clear-cut. In order to prevent inventors from keeping an invention secret and seeking a patent only in case of impending imitation, firms forfeit their right to patent an invention after a short ‘grace period’ of commercial usage. In Germany, the Czech Republic, and Hungary this grace period is six months, while most other European countries are even more strict and require absolute novelty. Before commercial usage of the invention, firms can rely on secrecy without passing up the opportunity to patent in later development stages of the invention.<sup>6</sup> However, since our CIS 4 sample solely consists of responses for *marketed* product innovations, the possibility of a complementary use of both means virtually does not apply to our data. We, therefore, consider the lack of precision resulting from assuming mutual exclusivity to be negligible and a small price to pay for a significantly larger data set.

Second, CIS 4 contains observations at firm-level, while our analysis requires information about each innovation. In order to alleviate this discrepancy, we limit the sample to firms with less than 50 employees, where the number of inventions is low enough to draw conclusions from the firm-level. This potentially limits the implications of our results. Since incurring fixed cost is harder for small firms, filing costs might exert a (compared to large firms) higher influence on the patenting decision. Yet, the fact that more than a quarter of firms in our sample are part of a larger corporate group mitigates this limitation. Albeit with due caution, this allows conclusions to be drawn about patenting behavior in general.

We use a probit model to estimate whether a firm filed a patent within the observation period or not, depending on the binary and non-binary size measures outlined above. Using

---

<sup>5</sup> See, e.g., 18 U.S.C. §1839 (3)(A) and §203 (2) StGB for the prerequisites of trade secrets.

<sup>6</sup> Hussinger (2004) finds that secrecy is important for early-stage inventions, while patents better protect IP already in the market. See also Erkal (2004).

the binary size measure, estimation results over all 14 countries suggest a strong support for the traditional view. This conclusion is robust to country-specific and, with the exception of Information & Communication Technologies (ICT), to industry-specific estimations. Using the non-binary size measure, results reveal the existence of an inverted-U relation between propensity to patent and innovation size, as predicted by the integrative model. Graphic analyses and country-specific estimations for Germany and Spain, moreover, confirm the integrative model's prediction that the patenting-size-relation exhibits an inverted-U for relatively weak property rights and becomes positive and linear for strong ones.<sup>7</sup> Hence, our findings provide new evidence in favor of traditional view and integrative model. We find no estimation result which hints at the prevalence of the strategic view. This does not call into question the seminal contribution of Anton and Yao (2004) to the debate on the role of patents and secrecy for IP protection but suggests that, only in combination with the issues raised by Kultti et al. (2007), Anton and Yao's contributions are empirically plausible.

Our findings have implications for the literature on the welfare effects of patents.<sup>8</sup> Patent regimes require disclosure in order to foster welfare-enhancing knowledge diffusion. An inverted-U relation between patenting and size limits diffusion-induced welfare effects of patents, as they only incite inventors to disclose medium-sized inventions, while the most important ones remain secret. What is more, the results affect the literature dealing with the skewness of the distribution of innovation size.<sup>9</sup> It is well known that only few inventions have major technological implications, while the impact of a vast bulk of inventions is negligible. The latter generate on average negative returns which have to be compensated by the few big inventions yielding positive ones. The riskiness of R&D investment thus depends on the fatness of the size distribution's tail, i.e., on arrival rates of large inventions. Our result of an inverted-U relation suggests that measuring the tail based on patent data might underestimate its fatness and, therewith, overstate the riskiness of R&D investment. This is because large inventions are kept secret, so that patent data understates the value of innovative activities. For the same reason, our results also affect the literature using renewal data to estimate patent values by limiting conclusions drawn from it for R&D policy.<sup>10</sup>

Besides the mentioned literature, our paper relates to Pajak (2010), who investigates a similar research question using CIS 4 data but focuses on the French part of the survey containing information about patenting *and* secrecy. He finds mixed results. For all innovative firms in his sample, the propensity to patent increases in innovation size by a significantly

---

<sup>7</sup> Germany and Spain are ideal subjects for comparison, as they obtain the most responses in the data set and differ regarding the amount of property rights protection and survey response rates.

<sup>8</sup> See, e.g., Denicolo (1996), O'Donoghue et al. (1998), and Hopenhayn et al. (2006).

<sup>9</sup> See Scherer/Harhoff (2000) and Silverberg/Verspagen (2007).

<sup>10</sup> See, e.g., Schankerman/Pakes (1986), Schankerman (1998), Lanjouw (1998), Grönqvist (2009).

larger amount than the propensity to use secrecy, which better fits to the traditional view. Yet, Pajak also finds that a higher share of innovative sales decreases an invention's probability of being patented, corroborating the strategic view that patents are rather used for small inventions. Unlike Pajak (2010), our paper theoretically integrates traditional and strategic view and, albeit at the expense of an additional variable capturing the use of secrecy included in the French part of CIS 4, checks for non-linear patenting-size-relations using a larger data set with survey results from Germany and 13 other European countries.

The paper also relates to Bhattacharya and Guriev (2006), who use a theoretical model with incomplete contracts and cumulative R&D to analyze an inventor's decision to disclose and license knowledge under both mechanisms, secrecy and patenting. Along the lines of the strategic view, they find that firms use patents to license less valuable inventions to their rivals, while more valuable knowledge is kept secret and licensed only to selected rivals. Denicolo and Franzoni (2004) endogenize the efforts to duplicate a given invention and, in line with the traditional view, find that in most circumstances it is preferable for firms to patent their inventions. Unlike the previous, our model focuses on patenting costs, simultaneous innovations, and probabilistic patent rights. Yet, our empirical approach can be applied to all theories predicting a certain patenting-size-relation, and its results are more in line with Denicolo and Franzoni (2004) than with Bhattacharya and Guriev (2006).

The paper is organized as follows. Section two outlines the theoretical model. Section three confronts the resulting predictions with the data, and section four concludes the paper.

## 2 Theoretical framework

### 2.1 The basic model

**Outline.** Consider a consumer good industry with two risk-neutral firms  $i \in \{A, B\}$  competing in quantities à la Cournot with homogeneous goods.<sup>11</sup> Without engaging in R&D, both firms are symmetric and obtain the same technology level enabling them to produce the good at marginal costs  $\bar{m} > 0$ . By engaging in R&D, each firm can try to improve its production knowledge and reduce marginal costs to  $m$ , where  $0 < m < \bar{m}$ .<sup>12</sup> The difference between  $\bar{m}$  and  $m$  measures the size of the invention. A larger invention implies a smaller  $m$  and vice versa. This heterogeneity is the key dimension for differentiating inventions with regard to the choice between secrecy and patenting as optimal means to protect them.

---

<sup>11</sup> Assuming a duopoly streamlines the analysis and suffices to include all aspects relevant in this context. Modelling damages as a fraction of the imitator's revenue, as in Anton/Yao (2004), requires assuming Cournot, as under Bertrand the imitator's revenue would be zero, so that no damages would accrue.

<sup>12</sup> We abstract from R&D costs, since they are inconsequential for the following analysis.

The model accounts for the possibility of simultaneous innovations implying that both firms can discover the same invention at the same time. In practice, this is not unlikely, since market requirements and technical standardization within an industry narrow down a technology's possible development paths.<sup>13</sup> To implement this notion, we follow O'Donoghue et al. (1998) and distinguish between an idea and an innovation. Ideas are common knowledge to all producers, occur during production, and outline a rough sketch of the technology's next step. In order to take this step, firms must turn the idea into an innovation containing the knowledge on how to implement it. To ensure that both firms strive for innovations of identical size (albeit with possible technical differences), we assume the idea to determine how much the technology frontier is enhanced, regardless of how it is implemented.

An innovation occurs with R&D success probability  $\phi \in [0, 1]$ .<sup>14</sup> Each firm independently innovates with probability  $\phi$  and fails to innovate with  $1 - \phi$ , so that the post-R&D market structure can take three different forms:

- i) With probability  $\phi^2$  both firms innovate, reduce their marginal costs and compete neck-to-neck at new marginal cost level  $m$  (simultaneous innovation case).
- ii) With probability  $2\phi(1 - \phi)$  one firm successfully innovates and obtains marginal costs  $m$ , while the other still produces at initial marginal cost level  $\bar{m}$  (leader-laggard case).
- iii) With probability  $(1 - \phi)^2$  both firms fail to innovate and compete neck-to-neck at  $\bar{m}$ .

Which case is realized, eventually, depends on the appropriation mechanism chosen.

**Profits and appropriation.** Innovators can choose between secrecy and patenting to protect their invention. In reality, neither of both means provides perfect protection. Leakage of information, e.g., as a result of reverse engineering or hiring away of workers, impairs the effectiveness of secrecy, while the validity of patent rights is unclear until challenged in court.<sup>15</sup> Both means differ, however, regarding their imitation deterring potential: while secrecy aims at preventing rivals from obtaining imitation facilitating information, patents inherently provide this information due to compulsory disclosure but impose the risk of having to pay infringement damages when using the information illegally. Depending on patent strength and the value of the information obtained, this risk can effectively deter rivals from imitation, in which case patent protection is perfect. Only if rivals choose to imitate despite the risk of infringement, patent rights are probabilistic and merely protect the invention in case the court adjudges the infringement claim valid. The resulting imitation

<sup>13</sup> See Kultti et al. (2006) and (2007) and Mosel (2011b).

<sup>14</sup> See Mosel (2011b) for a model with similar results, where R&D investment is endogenous.

<sup>15</sup> See Denicolo/Franzoni (2004) and Lemley/Shapiro (2005).



detering effect of probabilistic patents is crucial to the predictions of the strategic view. Our theory thus models probabilistic patent rights, if rivals choose to imitate, and perfect patent protection, if imitators eschew the risk of infringement. Yet, because qualitative results are unaffected, it refrains from also modelling imperfect protection provided by secrecy.<sup>16</sup>

Secrecy and patenting differ regarding the costs involved to use them. While secrecy might necessitate monetary incentives for employees and technical adjustments to hinder reverse engineering, patenting costs include filing fees, legal resources to manage and enforce the patent right, and renewal fees in order not to forfeit it over time. For three reasons we assume patenting to be more costly than secrecy. First, filing and renewal fees accrue at each patent office the inventor seeks protection at. Second, due to imperfect protection, patents also require monetary incentives for employees and technical adjustments to hinder imitation. Third, litigation for patent infringement and fighting rivals' countercharges by far outstrip legal resources spent for secrecy. We thus assume secrecy to be free of charge, while patenting costs are combined in  $c \geq 0$  paid upon filing, which can be interpreted as relative costs capturing the extent to which patenting is more costly than secrecy.

Producers are assumed to face demand which takes the simple unit-elastic (inverse) form  $p = \frac{\xi}{x_i + x_{-i}}$ , where  $\xi$  denotes expenditures on consumption,  $p$  is the consumer good's price, and  $x_i$  and  $x_{-i}$  with  $i \neq -i$  are the outputs produced by both firms, respectively.<sup>17</sup> Where applicable, we denote the leader by  $i$  and the laggard by  $-i$ , while keeping in mind that in cases where firms compete neck-to-neck (at the same technology level) this distinction is superfluous. Without loss of generality, we can normalize expenditures to unity and measure outputs and profits 'in terms of'  $\xi$ . Firms maximize profits which in their most general form, i.e., including the possibility of damages, equal  $\pi_i = (p - m_i) x_i + \delta p x_{-i}$  for the leader and  $\pi_{-i} = (p - m_{-i}) x_{-i} - \delta p x_{-i}$  for the laggard.  $m_i$  and  $m_{-i}$  capture marginal costs of leader and laggard, and  $\delta \in [0, 1]$  is the amount of damages per unit sold by  $-i$ .  $\delta = \eta \tau$  includes royalty rate  $\tau \in [0, 1]$  on the protected good's price and the probability  $\eta \in [0, 1]$  that a court finds the patent to be valid. Since both components result from the extent of patent protection granted,  $\delta$  can be interpreted as a measure for patent strength. While in case of  $\delta = 0$  imitation never infringes the patent,  $\delta = 1$  implies maximum patent protection, as it entitles the patentee to reap damages to the amount of the imitator's entire revenue,  $p x_{-i}$ .<sup>18</sup>

<sup>16</sup> Imperfect secrecy protection 'scales down' but does not qualitatively change results (Mosel (2011b)).

<sup>17</sup> In Cournot oligopolies, iso-elastic demand functions imply non-monotonic reaction functions, where firms' outputs are strategic substitutes for some parameter values and strategic complements for others; see Bandyopadhyay (1997). Setting up an identical maximization problem, Anton and Yao (2004) avoid this problem by assuming linear demand, but at the cost of considerably more intricate profit functions. Our model is able to replicate their results with a simple unit-elastic demand, so that we consider non-monotonic reaction functions to be a low price to pay for the significant gain of tractability.

<sup>18</sup> This profit specification follows Anton/Yao (2004). Note that whether damages are a fraction of revenue

Profits: $\pi$	Secrecy	Patenting	
		No imitation	Imitation
Leader	$\frac{\bar{m}^2}{(m+\bar{m})^2}$	$\frac{\bar{m}^2}{(m+\bar{m})^2}$	$\frac{3\delta+1}{4}$
Laggard	$\frac{m^2}{(m+\bar{m})^2}$	$\frac{m^2}{(m+\bar{m})^2}$	$\frac{1-\delta}{4}$
Neck-to-neck	0.25	0.25	0.25

Table 1: Profits under different appropriation regimes and technology gaps.

Utilizing implicit reaction functions  $m_i = \frac{(1-\delta)x_{-i}}{(x_i+x_{-i})^2}$  and  $m_{-i} = \frac{(1-\delta)x_i}{(x_i+x_{-i})^2}$ , the firms' profit-maximizing outputs equal  $x_i = \frac{(1-\delta)m_{-i}}{(m_i+m_{-i})^2}$  and  $x_{-i} = \frac{(1-\delta)m_i}{(m_i+m_{-i})^2}$ , yielding leader and laggard profits

$$\pi_i = \frac{m_{-i}^2 + 2\delta m_i m_{-i} + \delta m_i^2}{(m_i + m_{-i})^2} \quad \text{and} \quad \pi_{-i} = \frac{(1-\delta)m_i^2}{(m_i + m_{-i})^2}. \quad (1)$$

Based on (1), we can derive profits for each appropriation mechanism and imitative behavior (see table 1). If both firms compete neck-to-neck, either due to simultaneous innovation ( $m_i = m_{-i} = m$ ) or because they both fail to innovate ( $m_i = m_{-i} = \bar{m}$ ), profits become  $\pi_i = \pi_{-i} = 0.25$ . If, however, only one firm innovates, profits are more intricate. Since secrecy is assumed to provide full protection, an innovator with cost level  $m$  using secrecy manages to keep the laggard at  $\bar{m}$  and realizes  $\pi_i = \frac{\bar{m}^2}{(m+\bar{m})^2}$ , while the laggard obtains  $\pi_{-i} = \frac{m^2}{(m+\bar{m})^2}$ . The same is true for patenting, if the laggard, e.g., due to the imitation deterring effect outlined above, decides not to imitate in order to avoid infringement. Yet, if  $\delta$  is low enough not to deter the laggard from imitation, both firms produce at marginal cost  $m$ , and the imitator has to reckon paying damages to the amount of  $\delta$  on the price of each product sold, yielding  $\pi_i = \frac{3\delta+1}{4}$  and  $\pi_{-i} = \frac{1-\delta}{4}$  for leader and laggard, respectively.

**Value functions.** Innovators choose between secrecy and patenting knowing their own R&D success but without knowledge of whether the rival succeeded. They thus form expectations to trade off the value resulting from secrecy against patent value. Using success probability  $\phi$  and profits in table 1, the expected secrecy values for leader and laggard can be written

$$V_i^s = \frac{(1-\phi)\bar{m}^2}{(m+\bar{m})^2} + \frac{\phi}{4} \quad \text{and} \quad V_{-i}^s = \frac{(1-\phi)}{4} + \frac{\phi m^2}{(m+\bar{m})^2}. \quad (2)$$

Utilizing secrecy, an innovator realizes 'leader profits'  $\frac{\bar{m}^2}{(m+\bar{m})^2}$ , if the rival fails to innovate, which occurs with probability  $(1-\phi)$ . If the rival also innovates (occurring with probability

---

(as modelled here) or of the laggard's profits is irrelevant for the main intuition behind our results.

$\phi$ ), firms are neck-to-neck and the innovator's profits equal  $\pi = \frac{1}{4}$ . Similarly, a firm that fails to innovate obtains  $\pi = \frac{1}{4}$ , if the rival is unable to generate the innovation, too, while it obtains 'laggard profits'  $\frac{m^2}{(m+\bar{m})^2}$ , if the rival assumes the technological lead.

Patent values are more intricate, since profits depend on the laggard's decision whether or not to imitate and because, in the simultaneous innovation case, patents provide a fifty-fifty chance to assume the lead. That is, a patent is awarded to only one of the innovators. Since both firms are symmetric prior to R&D and innovate simultaneously, both obtain the same probability of being granted the patent:  $\frac{1}{2}$ . The leader's expected patent value can thus be written as

$$V_i^p = \begin{cases} \frac{(1-\phi)\bar{m}^2}{(m+\bar{m})^2} + \frac{\phi(m^2+\bar{m}^2)}{2(m+\bar{m})^2} - c & \text{if 'no imitation'} \\ \frac{(1-\phi)(3\delta+1)}{4} + \frac{\phi(\delta+1)}{4} - c & \text{if 'imitation'} \end{cases} . \quad (3)$$

The leader realizes the value in the first line of (3), if the laggard chooses not to imitate. It includes 'leader profits' (see table 1), if the rival fails to innovate, the simultaneous innovation case (occurring with probability  $\phi$ ), in which the innovator obtains both 'leader profits' and 'laggard profits' with probability  $\frac{1}{2}$ , and patenting costs  $c$ . The second line, besides patenting costs, comprises 'leader profits', if the rival fails to innovate but chooses to imitate, and it comprises the simultaneous innovation case, in which the innovator failing to obtain the patent chooses to use his knowledge despite the risk of infringement. In that case, occurring with probability  $\phi$ , an innovator has a fifty-fifty chance to make either  $\frac{3\delta+1}{4}$  or  $\frac{1-\delta}{4}$ .

Similarly, the expected patent value of a firm failing to innovate equals

$$V_{-i}^p = \begin{cases} \frac{(1-\phi)}{4} + \frac{\phi m^2}{(m+\bar{m})^2} & \text{if 'no imitation'} \\ \frac{(1-\phi)}{4} + \frac{\phi(1-\delta)}{4} & \text{if 'imitation'} \end{cases} , \quad (4)$$

where the first line includes 'neck-to-neck' profits (see table 1), in case the rival also fails to innovate, and 'laggard profits' without imitation, in case the rival successfully innovates. Finally, the second line comprises 'neck-to-neck profits' plus 'laggard profits', in case the laggard imitates the knowledge disclosed in the patent despite the risk of infringement.

**Timing of the model.** Once an idea occurs, each firm implements it (via an innovation) with probability  $\phi$  or fails to implement it with probability  $1-\phi$ . Firms know their own R&D success and the size of the innovation but have no knowledge of whether the rival failed to innovate or not. The model analyzes two decisions: the patenting decision and, conditional on that, the imitation decision. Each innovating firm chooses between secrecy and patenting by trading off expected patent value given by (3) against the value provided by secrecy given

by (2). If the innovator chooses to patent (or one of both innovators receives the patent right both strive for), the laggard subsequently decides whether or not to use the patented knowledge and risk infringement by weighing expected value in the ‘no imitation’ case in equation (4) against the one under ‘imitation’. The solution concept of this simple set-up is subgame-perfect equilibria, so we solve the model using backward induction, starting with the laggard’s decision whether or not to imitate.

## 2.2 The traditional view

By taking into account patenting costs and simultaneous innovations, the traditional view can be brought in line with the finding that firms regard secrecy as more effective in protecting a technological lead than patents. Secrecy can be modelled to provide more protection than patenting at less cost, but patents can compensate for these disadvantages by also providing protection in situations where secrecy fails, i.e., in case of simultaneous innovations. In order to set up a ‘traditional view specification’ of the model presented here, we assume strictly positive patenting costs  $c > 0$  and suppress the laggard’s strategic decision whether or not to imitate and risk infringement in case the leader patented his invention. To that extent, we set  $\delta = 1$  to ensure that using patented knowledge entails damages to an amount high enough to always deter imitation.<sup>19</sup>

Without the laggard’s strategic imitation decision, the only problem to solve in this specification is the innovator’s decision between secrecy and patenting, in which he trades off values  $V_i^s = \frac{(1-\phi)\bar{m}^2}{(m+\bar{m})^2} + \frac{\phi}{4}$  and  $V_i^p = \frac{(1-\phi)\bar{m}^2}{(m+\bar{m})^2} + \frac{\phi(m^2+\bar{m}^2)}{2(m+\bar{m})^2} - c$ , given by (2) and (3). Besides patenting costs, both differ regarding profits in the simultaneous innovation case occurring with probability  $\phi$ . While secrecy fails to provide protection in that case and both firm’s profits remain at the ‘neck-to-neck’ level realized prior to invention,  $\frac{1}{4}$ , patenting, despite an only fifty-fifty chance to receive the patent, generates higher profits  $\frac{m^2+\bar{m}^2}{2(m+\bar{m})^2}$ , yielding

LEMMA 1. *In the simultaneous innovation case, patenting yields higher expected profits than secrecy for each innovation size satisfying  $0 < m < \bar{m}$ .*

*Proof.* See Appendix A.1.<sup>20</sup> □

Lemma 1 enables patents to compensate for the downside of patenting, costs  $c$ , if innovation size is high enough. This turns out to be the key tradeoff driving the innovator’s decision between secrecy and patenting in the traditional view specification.

<sup>19</sup> Albeit at the expense of full empirical plausibility, assuming perfect patent rights turns out to be non-relevant for qualitative results, as probabilistic rights as in Mosel (2011b) generate identical predictions.

<sup>20</sup> Mosel (2011a) proves this feature for Cournot profits resulting from a more general CES specification.

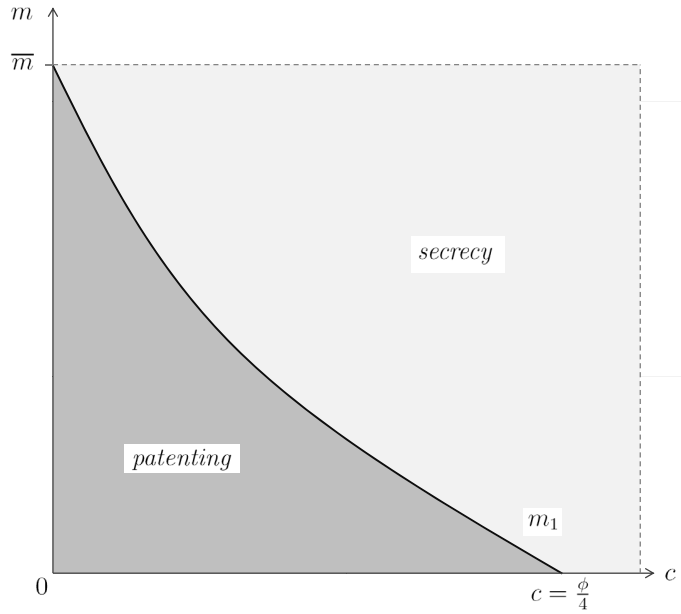


Figure 1: Patenting behavior, traditional view specification.

To see that, equate  $V_i^s$  and  $V_i^p$  to yield the size cut-off separating inventions kept secret from patented ones:

$$m_1 = \bar{m} \left( \frac{\phi + 4c - 4\sqrt{c\phi}}{\phi - 4c} \right). \quad (5)$$

Firms are indifferent between secrecy and patenting, if the innovation reduces marginal costs to  $m_1$ . For smaller innovations than this cut-off size, i.e.,  $m > m_1$ , they prefer secrecy over patents, since, relative to innovation size, the probability of simultaneous innovations,  $\phi$ , is too low and patenting costs,  $c$ , are too high to justify patenting. Large innovations, where  $m < m_1$ , on the other hand, are optimal to patent, because patenting costs are relatively small and each firm cannot afford to leave an important invention unprotected in case of simultaneous innovations (where secrecy fails). We summarize this finding in

**PROPOSITION 1.** *In the traditional view specification ( $c > 0$  and  $\delta = 1$ ), a unique cut-off  $m_1 = \bar{m} \left( \frac{\phi + 4c - 4\sqrt{c\phi}}{\phi - 4c} \right)$ , at which firms are indifferent between secrecy and patenting, exists for all  $m \in (0, \bar{m})$ ,  $\phi \in (0, 1]$  and  $c \in (0, \frac{\phi}{4})$ . Firms choose secrecy over patenting for small innovations, where  $m > m_1$ , and patent large ones, where  $m < m_1$ .*

*Proof.* See Appendix A.2. □

The optimal patenting decision in the traditional view specification is illustrated in figure 1.

### 2.3 The strategic view

The traditional view neglects the laggard's strategic decision whether or not to imitate and risk infringement by using the knowledge disclosed in a patent. Anton and Yao (2004) point

out that innovators can count on damage compensation due to infringement as an additional source of revenue in case courts rule the patent to be valid. To account for the strategic use of patents to tap alternative sources of income and for the associated imitation deterring effect of probabilistic patent rights, we set up a ‘strategic view’ specification which refrains from patenting costs ( $c = 0$ ) and rather focuses on the imitation deterring effect by assuming  $\delta \in [0, 1)$ . Due to backward induction, this implies solving the laggard’s imitation problem prior to the innovator’s decision between secrecy and patenting.

A firm failing to innovate obtains expected value  $V_{-i}^p$  given by (4), if its rival chooses to patent the invention in case of R&D success. The imitation decision thus boils down to weighing ‘no imitation’ profits,  $\frac{m^2}{(m+\bar{m})^2}$ , against ‘imitation’ profits,  $\frac{1-\phi}{4}$ , resulting in

PROPOSITION 2. *A unique imitation cut-off  $m_2 = \bar{m} \left( \frac{1-\delta+2\sqrt{1-\delta}}{3+\delta} \right)$ , at which a laggard is indifferent whether or not to imitate, exists for all  $m \in (0, \bar{m})$  and  $\delta \in [0, 1)$ . It is optimal to imitate large innovations, where  $m < m_2$ , and not to imitate small ones, where  $m > m_2$ .*

*Proof.* See Appendix A.3. □

Depending on patent strength  $\delta$ , the ‘imitation cut-off’  $m_2$  draws the line between innovations which are too small ( $m > m_2$ ) to be worth risking infringement and large ones ( $m < m_2$ ), where  $\delta$  is not high enough to deter the laggard from imitation.

Based on the imitation cut-off, the decision between secrecy and patenting breaks down into two cases: *a)* for all innovation sizes  $m > m_2$  the innovator trades  $V_i^s$  given by (2) off against  $V_i^p|_{c=0}$  in case of no imitation given by (3); and *b)* for all innovation sizes  $m < m_2$  this tradeoff includes  $V_i^s$  and  $V_i^p|_{c=0}$  in case of imitation. Case *a)* implies a simple solution, as  $\frac{(1-\phi)\bar{m}^2}{(m+\bar{m})^2} + \frac{\phi}{4} < \frac{(1-\phi)\bar{m}^2}{(m+\bar{m})^2} + \frac{\phi(m^2+\bar{m}^2)}{2(m+\bar{m})^2} \Leftrightarrow 0 < (\bar{m}-m)^2$  and, therewith,  $V_i^s < V_i^p|_{c=0}$  holds for any  $m$  within the relevant parameter range. That is, as long as the laggard always chooses not to imitate, patenting yields a higher value than secrecy for any innovation. This is due to the advantage of patents in the simultaneous innovation case (see Lemma 1), which, in this specification of the model, can be exploited without the downside of patenting costs.

In the second case, where the laggard always chooses to imitate, the innovator weighs  $V_i^s$  given by (2) against the ‘imitation’ value of  $V_i^p|_{c=0}$ , yielding ‘patenting’ cut-off

$$m_3 = \bar{m} \left( \frac{2}{\sqrt{3\delta + 1 + \frac{\delta\phi}{1-\phi}}} - 1 \right), \quad (6)$$

which separates patented innovation sizes from those optimal to be kept secret. Below this cut-off ( $m < m_3$ ), innovations are too large to justify knowledge disclosure entailed by

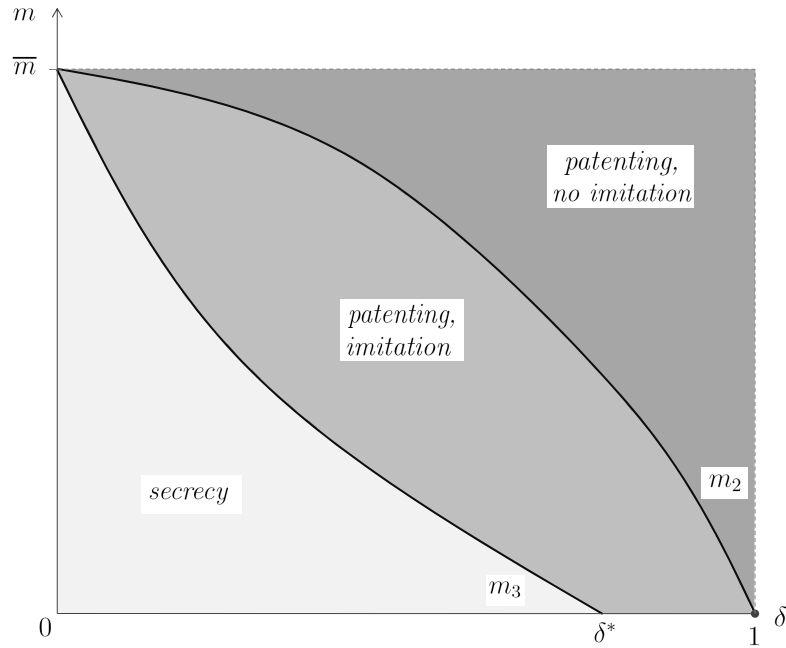


Figure 2: Patenting behavior and imitation decision, strategic view specification.

patenting, so that the innovator prefers secrecy. Neither the prospect of damage payments nor the patent's additional protection in the simultaneous innovation case (captured by  $\delta\phi$  in (6)) can incite the innovator to give away the large technological lead. By contrast, above this cut-off ( $m > m_3$ ), innovation sizes are small enough to patent the innovation and accept knowledge disclosure in light of the prospect of damage compensation. For this reason, a higher patent strength  $\delta$  strictly decreases  $m_3$  in (6): better prospects of reaping damages as an additional source of revenue makes it worthwhile to patent larger inventions.

As it turns out, for all relevant  $m \in (0, \bar{m})$ , patenting cut-off  $m_3$  given by (6) and imitation cut-off  $m_2$  from Proposition 2 do not interfere (see Appendix A.4), enabling us to write

PROPOSITION 3. *A unique cut-off  $m_3 = \bar{m} \left( \frac{2}{\sqrt{3\delta+1+\frac{\delta\phi}{1-\phi}}} - 1 \right)$  exists for all  $m \in (0, \bar{m})$ ,  $\phi \in (0, 1]$ , and  $\delta \in [0, \delta^*)$ , where  $\delta^* = \frac{3-3\phi}{3-2\phi}$ . Given imitation cut-off  $m_2$  from Proposition 2, in the strategic view specification it is optimal for innovator and laggard, respectively, to*

- a) *patent and not imitate small innovations, where  $m > m_2 > m_3$ ,*
- b) *patent and imitate medium-sized innovations, where  $m_2 > m > m_3$ , and*
- c) *keep large innovations secret, where  $m_2 > m_3 > m$ , precluding imitation as an option.*

*Proof.* See Appendix A.4. □

The strategic view predictions for patenting behavior, illustrated in figure 2, resemble those of Anton and Yao (2004).<sup>21</sup> As becomes apparent from the figure, cut-off  $m_3$  implies a patent

<sup>21</sup> Note that this paper focuses on the decision between secrecy and patenting and thus does not model the additional analysis conducted by Anton and Yao (2004) of how much knowledge to disclose.

strength  $\delta^*$ , above which patenting becomes lucrative enough for all innovation sizes to be patented. Since numerous studies show that not all inventions are patented<sup>22</sup>, Proposition 3 focuses on  $\delta \in [0, \delta^*)$  to ensure empirical plausibility. Intuitively, patenting becomes more worthwhile with a higher probability of simultaneous innovations,  $\phi$ , since, then, patents can play out their strength summarized in Lemma 1. For that reason, a higher  $\phi$  decreases  $\delta^*$  and, therewith, the empirically plausible parameter range  $\delta \in [0, \delta^*)$ .

## 2.4 An integrative model of patenting behavior

Both traditional and strategic view neglect aspects analyzed by the other, limiting their explanatory power. To make up for this shortcoming, we set up an integrative model of patenting behavior with patenting costs ( $c > 0$ ) and the imitation deterring effect of probabilistic patent rights, such that  $\delta \in [0, 1)$ . Again, using backward induction, in the following the laggard's imitation problem is solved prior to the innovator's patenting decision.

The imitation decision is identical to the one for the strategic view, since patenting costs do not affect the laggard's value functions (eq. (4)). Imitation cut-off  $m_2 = \bar{m} \left( \frac{1-\delta+2\sqrt{1-\delta}}{3+\delta} \right)$  thus carries over to the integrative model. For all patented inventions smaller than the imitation cut-off ( $m > m_2$ ), the infringement risk deters the laggard from imitation, while the laggard finds it optimal to imitate larger inventions ( $m < m_2$ ) despite what is at stake.

As before, the imitation cut-off influences the innovator's decision between secrecy and patenting. Yet, unlike in the strategic view, where the imitation deterring effect of patents implicates that all small innovations are optimal to patent, innovators in the integrative model face patenting costs and can thus no longer exploit the imitation deterring effect free of charge. As a result, not all innovations smaller than the imitation cut-off ( $m > m_2$ ) are patented. To see that, equate the innovator's secrecy value  $V_i^s$  with patent value  $V_i^p$  without imitation, yielding  $m_1 = \bar{m} \left( \frac{\phi+4c-4\sqrt{c\phi}}{\phi-4c} \right)$ , which is identical to the cut-off in the traditional view given by (5). The reason for this identity is simple. The imitation deterring effect of probabilistic patent rights implies that for small innovations (here  $m > m_2$ ) patents provide perfect protection, since it is not worthwhile for the laggard to challenge them. As long as the lagging firm chooses not to imitate, the patenting decision in the integrative model is identical to the one in the traditional view: for all  $m > m_2$ , inventions larger than  $m_1$  are patented while inventions smaller than  $m_1$  are optimal to keep secret.

For inventions larger than the imitation cut-off ( $m < m_2$ ), the patenting cut-off resembles the one for the strategic view. Yet, due to patenting costs, the innovator in the integrative model trades  $V_i^s = \frac{(1-\phi)\bar{m}^2}{(m+\bar{m})^2} + \frac{\phi}{4}$  off against  $V_i^p = \frac{(1-\phi)(3\delta+1)}{4} + \frac{\phi(\delta+1)}{4} - c$ , yielding a modified

<sup>22</sup> See, e.g., Mansfield (1986), Arundel/Kabla (1998), and Brouwer/Kleinknecht (1999).



patenting cut-off

$$m_{3,c} = \bar{m} \left( \frac{2}{\sqrt{3\delta + 1 + \frac{\delta\phi - 4c}{1-\phi}}} - 1 \right). \quad (7)$$

As long as innovations are large enough for the laggard to choose imitation, it is optimal for the innovator to patent inventions smaller than patenting cut-off  $m_{3,c}$  and hold larger ones secret. As in the previous section, this is due to the fact that innovators can benefit from a patent's advantage in the simultaneous innovation case and the possibility to reap damages in case of infringement. However, both advantages come at the cost of knowledge disclosure, which is a bigger shortcoming the larger the innovation is, so that patenting ceases to be worthwhile once the patenting cut-off is reached. By also assuming patenting costs, in this specification,  $m_{3,c}$  is reached by even smaller innovation sizes, since  $c$  makes patenting less worthwhile.<sup>23</sup> Combining this result with the previous cut-offs yields

PROPOSITION 4. *Unique cut-offs  $m_1 = \bar{m} \left( \frac{\phi + 4c - 4\sqrt{c\phi}}{\phi - 4c} \right)$ ,  $m_2 = \bar{m} \left( \frac{1 - \delta + 2\sqrt{1 - \delta}}{3 + \delta} \right)$ , and  $m_{3,c} = \bar{m} \left( \frac{2}{\sqrt{3\delta + 1 + \frac{\delta\phi - 4c}{1-\phi}}} - 1 \right)$  exist for all  $m \in (0, \bar{m})$ ,  $c \in [0, \frac{\phi}{4})$ ,  $\phi \in (0, 1]$ , and  $\delta \in [\delta_{min}, 1)$ , where  $\delta_{min} = \frac{4(1-\phi+c)}{4-3\phi} - \frac{8(1-\phi)^2 + 4(1-\phi)\sqrt{(2-\phi)^2 - 16c + 12\phi c}}{(4-3\phi)^2}$ . Moreover, as  $m_1 \leq m_4$  holds, where  $m_4$  denotes the intersection between  $m_2$  and  $m_{3,c}$ , the optimal patenting decision implies that*

- a) for all patent strengths between  $\delta_{min}$  and  $\delta^*$ , where in this case  $\delta^* = \frac{3-3\phi+4c}{3-2\phi}$ ,
  - i) small innovations ( $m > \max\{m_1, m_2\}$ ) are kept secret,
  - ii) medium-sized ones ( $\max\{m_1, m_2\} > m > m_{3,c}$ ) are patented, and
  - iii) large innovations ( $m < m_{3,c}$ ) are kept secret, while
- b) for all patent strengths between  $\delta^*$  and unity,
  - i) small innovations ( $m > \max\{m_1, m_2\}$ ) are kept secret, and
  - ii) large innovations ( $m < \max\{m_1, m_2\}$ ) are patented.

*Proof.* See Appendix A.5. □

The predictions of the integrative model are illustrated in figure 3. It becomes apparent that, due to patenting costs, for all  $\delta < \delta_{min}$  it is optimal for innovators to solely rely on secrecy, even if the laggard chooses not to imitate. Since, as outlined above, at least some inventions are patented, Proposition 4 focuses on the empirically plausible cases  $\delta \in [\delta_{min}, \delta^*)$  and  $\delta \in [\delta^*, 1)$ . The latter implies a prediction which resembles the traditional view: small inventions ( $m > \max\{m_1, m_2\}$ ) are ideally kept secret, while large ones ( $m < \max\{m_1, m_2\}$ )

<sup>23</sup> It immediately follows from (7) that a higher  $c$  increases  $m_{3,c}$ .

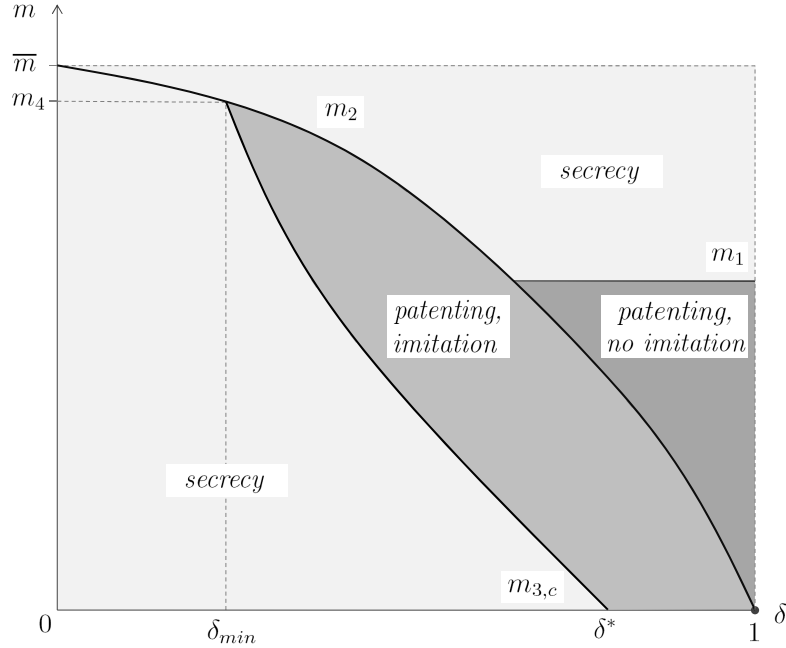


Figure 3: Patenting behavior, integrative model.

are optimal to patent, as for all  $\delta > \delta^*$  patent strength is high enough to compensate for compulsory disclosure and patenting costs. The former case's prediction is a mixture between traditional and strategic view: for all  $\delta \in [\delta_{min}, \delta^*)$ , small innovations and large inventions ( $m > \max\{m_1, m_2\}$  and  $m < m_{3,c}$ ) are kept secret, while medium-sized ones ( $m_{3,c} < m < \max\{m_1, m_2\}$ ) are optimal to patent. In both cases, the cut-off identifying small innovations is the higher of either  $m_1$  or  $m_2$ . For qualitative results the exact location of the intersection of  $m_1$  and  $m_2$  is irrelevant, as long as  $m_1$  is positive and smaller than the intersection of  $m_2$  and  $m_{3,c}$  (corresponding to point  $[m_4, \delta_{min}]$  in figure 3). Both requirements for  $m_1$  are ensured by the parameter restrictions in Proposition 4.

### 3 Empirical implementation

Table 2 summarizes the results of the different specifications of the model. It becomes apparent that a simple modification of the underlying assumptions regarding filing costs and patent strength fundamentally changes theoretical predictions. Three different predictions have been derived. Firstly, traditional view and a specification of the integrative model predict that large inventions are patented and small ones are kept secret (prediction I). Secondly, the strategic view reaches the opposite result where firms patent small innovations and keep large ones secret (prediction II). Thirdly, the extant specification of the integrative model predicts that it is optimal to keep small and large inventions secret and patent medium-sized ones (prediction III). The question which one of these predictions prevails is an empirical

	Specification		no.	Predictions		
	<i>filing costs</i>	<i>patent strength</i>		<i>small</i>	<i>medium</i>	<i>large</i> → invention size
Traditional	$c > 0$	$\delta = 1$	I	Secrecy	-	Patenting
Strategic	$c = 0$	$\delta \in [0, \delta^*)$	II	Patenting	-	Secrecy
Integrative	$c > 0$	$\delta \in [\delta^*, 1)$	I	Secrecy	-	Patenting
		$\delta \in [\delta_{min}, \delta^*)$	III	Secrecy	Patenting	Secrecy

*Note:*  $\delta^*$  and  $\delta_{min}$  measure the amount of patent protection which, depending on the specification, ensures empirical plausibility, i.e., the fact that patenting *and* secrecy occurs.

Table 2: Theoretical predictions: traditional view, strategic view, and integrative model.

one. For that reason, the remainder of the paper deals with an empirical analysis testing them using data from the European Community Innovation Survey 4.

### 3.1 Data

The CIS 4 firm survey was conducted under the auspices of Eurostat over the period of 2002 to 2004. Under the guidelines of the OECD Oslo Manual, in all E.U. member states plus Norway, Switzerland, and Iceland, national statistical offices sent a widely standardized questionnaire to a sample of representative firms.<sup>24</sup> Data for 15 out of all participating countries are available at Eurostat as anonymized micro-data via CD-Rom. The results for 14 of these countries (69,152 firm responses) constitute the basis for the following analysis.<sup>25</sup>

CIS 4 provides the rare opportunity to test the impact of innovation size on patenting behavior, but at the cost of two shortcomings. A minor one results from the fact that the variables used as size measures are only available for product innovations, forcing us to eliminate all process innovations from the sample. Yet, since for our theoretical results there is no difference between reducing marginal production costs of a given consumer good and increasing the quality of a good produced at constant cost, this shortcoming of the data set should not limit its applicability for testing the theoretical set-up outlined above.

A more important shortcoming of using CIS 4 data in the context of patenting behavior is that firm-level observations lack information about the appropriation mechanism chosen for each innovation.<sup>26</sup> As firms might generate more than one innovation within the two-

<sup>24</sup> Despite the same guidelines, each country was free to determine additional questions, participation requirement (voluntary or mandatory), the medium used for data collecting, and sample size.

<sup>25</sup> Italy obtains other size classes for the number of employees. As this distinction is crucial to our model and as the number of extant observations is sufficient, we use data for all 15 countries but Italy.

<sup>26</sup> The main statistical unit is the firm defined as “*the smallest combination of legal units that is an organisational unit producing goods or services, which benefits from a certain degree of autonomy in*

year observation period, it is impossible to entirely disentangle which innovation made the firm file a patent application. In a similar study, Pajak (2010) addresses this problem by focussing on small firms, where the number of patentable inventions can be expected to be smaller than for large firms. He uses data from the French ‘Enquete Annuelle Entreprises’ for the year 2005 to show that, as a matter of fact, firms with less than 50 (30) employees file on average only 1.27 (0.95) patent applications, while the average increases to over 8 applications for firms with more than 500 employees. Following Pajak (2010), we thus restrict the analysis to the firm size class with the least number of employees (less than 50 in 2004) in order to approximate the ideal case that a firm’s decision to patent is driven by one underlying innovation. This leaves us with a sample of 7,185 firm responses (henceforth: ‘large sample’), described in more detail in table 4 in Appendix B.<sup>27</sup> To further alleviate the multi-innovation problem, we also identify a ‘small sample’ of firms that refrained from using means of IP protection hinting at the presence of innovations not qualifying for utility patents. These include copyrights, design patents, or trademarks.<sup>28</sup> Table 4 compares the number of responses for the ‘large’ and ‘small sample’ by country. On the one hand, using the ‘small sample’ further reduces the possibility of a multi-innovation observation and increases the focus on patentable inventions. On the other, the ‘small sample’ neglects patentable innovations which are compatible with the other means of IP protection. To consider both approaches, results of both samples are reported to check for robustness.

CIS 4 provides information about patenting behavior, measured by binary variable  $y_p$ , which equals one, if the respondent filed a patent application within the period of observation, and zero, if he did not. It can be used as the dependent variable mapping the decision between secrecy and patenting in the model. The reason for this is twofold. For one, if the invention is commercially used, mutual exclusivity of both means is ensured by the underlying legal provisions of both means of protection: valid trade secrets require ‘reasonable measures’ to keep an invention secret, ruling out patents due to compulsory disclosure, while inventors forfeit the right to patent after a short ‘grace period’ of commercial usage, as patentability requires absolute novelty.<sup>29</sup> Since our sample consists of *marketed* product innovations, the fact that in earlier stages of development both means can be complementary is irrelevant given the underlying data set. For another, alternative means of IP protection either concern non-patentable inventions and are accounted for by separate binary variables in the sample (e.g., trademarks or copyrights), or they concern patentable inventions (i.e.,

---

*decision making, especially for the allocation of its current resources.*”; see CIS 4 documentation.

<sup>27</sup> Besides large firm size classes, we eliminate responses with missing information in key variables.

<sup>28</sup> Design patents protect non-technical designs that are not protectable by ‘normal’ utility patents.

<sup>29</sup> See 18 U.S.C. §1839 (3)(a) and §203 (2) StGB.

lead time advantages), but their use is highly correlated and oftenly combined with secrecy.<sup>30</sup>

CIS 4 also provides us with two self-reported measures of innovation size: a binary variable separating inventions which are only ‘new to the firm’ from those which are also ‘new to the market’, and the percentage share of total turnover generated by an invention that is ‘new to the market’.<sup>31</sup> The former distinguishes inventions with a small technological impact on the market from those with a large impact and is thus suitable for testing predictions I and II, where IP protection only differs for small and large inventions. The latter variable measures the significance of the invention for the firm, which presumably increases in innovation size. Since the fraction of turnover generated by an innovation can also result from other factors than size (e.g., commercial success of the line of business introducing the invention), this size measure can be problematic. However, unlike a binary variable, it provides enough information to also test prediction III, where small, medium-sized, and large innovations differ with regard to the IP protection mechanism chosen. While keeping in mind to treat results with due care, we thus also use the percentage share of innovative turnover as a size measure and check for robustness using the binary alternative.

The survey also includes several control variables: respondents’ total turnover and R&D expenditures in € in 2004; the market scale captured by four binary variables that equal one, respectively, if the firm sells its products ‘regionally’, ‘nationally’, ‘E.U.-wide’, or ‘outside of the E.U.’; country and industry fixed effects; and a binary variable capturing whether the firm is part of a larger corporate group. The latter aspect might influence patenting behavior, as a subsidiary can be forced to reallocate the IP protection decision to the parent to centralize efforts. The market scale might influence patenting behavior, as firms active worldwide face a larger diversity of competitors and are more likely to face a rival with a similar invention. Industry fixed effects consist of five categories (‘ICT’, ‘Chemicals & Plastics’, ‘Machinery & Metal Products’, ‘Electronics’, and ‘Miscellaneous’) aggregated manually based on the E.U. ‘NACE’ classification of economic activities in order to cope with the low number of observations in some countries.<sup>32</sup> All variables are summarized in table 5 in Appendix B.

Finally, note that CIS 4 comprises a considerable country heterogeneity in survey response rates, because answering the questionnaire was voluntary in some countries (e.g., Germany and Belgium) and mandatory in most others (see table 4). To account for non-responses, the data set provides a weight, computed by each national statistical office via stratification

---

<sup>30</sup> It is hard to think of cases where ‘lead time advantages’ are effective in appropriating returns without the simultaneous use of secrecy to protect the lead temporarily by hindering reverse engineering.

<sup>31</sup> An invention that is only new to the firm, for instance, results from inventing around an existing patent, i.e., a product serves the same purpose as an existing product but is based on a technological solution which is patentable on its own, as it is different enough not to violate the existing patent.

<sup>32</sup> For a detailed classification, see Commission Regulation No. 1450/2004, Annex section 2.

over industries and firm sizes. Since non-responses can be due to certain firm characteristics, using the weight entails the risk of scaling up self-selection bias. The following estimation results are thus reported with and without the weight in order to check for robustness.

### 3.2 Descriptive statistics

To get an idea of how patenting behavior relates to our two measures of innovation size, let us define a ‘patent ratio’  $\mathcal{P} = \frac{\text{responses where } y_p=1}{\text{all responses}}$ , which captures the average propensity to patent an invention over all responses. If the traditional view holds, we expect the patent ratio to increase in innovation size, while, if the strategic view holds, the patent ratio should be smaller the more significant the invention. Figures 4 and 5 illustrate patent ratios by country and by industry for small and large inventions based on the binary size measure. Without exception, large inventions obtain higher patent ratios than small inventions, providing us with a first hint at the prevalence of the traditional view.

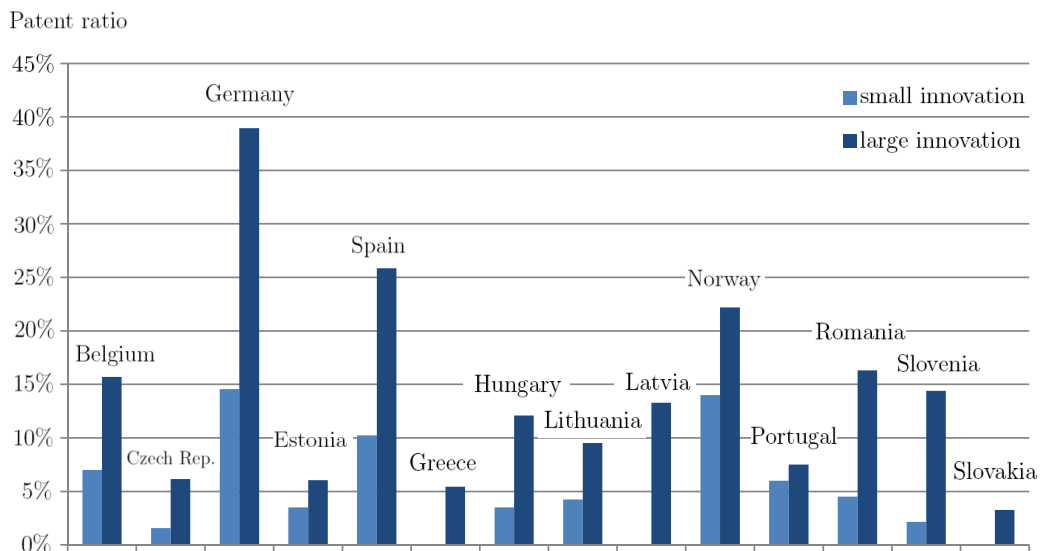


Figure 4: Average patent ratio for the large sample, by country.

Using the non-binary size measure, we can also account for the prediction of an inverted-U relation between size and the propensity to patent. Figure 6 shows the average patent ratio for each quintile of the fraction of innovative turnover by country. Only six countries in the large sample obtain observations over the whole range of this size measure. Four of them (Belgium, Czech Rep., Estonia, and Spain) seem to be consistent with the prediction of an inverted-U relation between patenting and size (lhs of fig. 6), while the patent ratios for two countries (Germany and Norway) suggest that the propensity to patent is the highest for large innovations, as predicted by the traditional view (rhs of fig. 6). Both results are well in

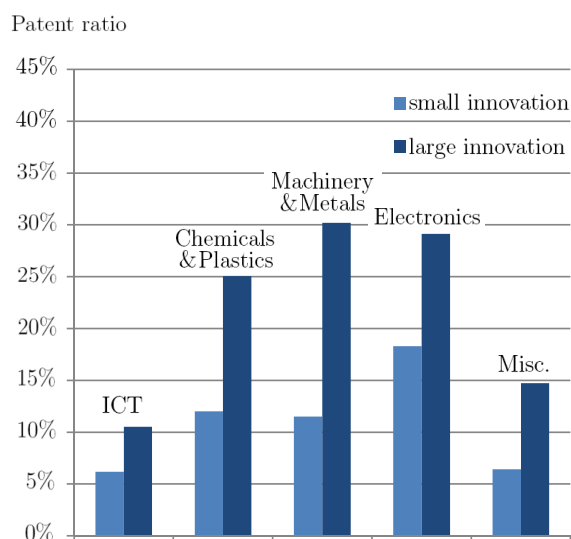


Figure 5: Average patent ratio for the large sample, by industry.

line with the integrative model. According to the 2007 International Property Rights Index (IPR-Index), Norway and Germany in fact rank first and third among 70 countries regarding the strength of property rights granted. Their IPR-Index scores (8.3 and 8.1 on a scale of 0 to 10, where 10 is the strongest protection possible) differ considerably from the other four countries, among which Belgium reaches the highest score (6.8) followed by Spain with 6.5. The lowest IPR-Index score (4.7) relates to the inverted-U curve in figure 6, whose peak is located the farthest to the right: the one of the Czech Republic.<sup>33</sup> The graphic analysis thus corroborates predictions I and III of the integrative model.

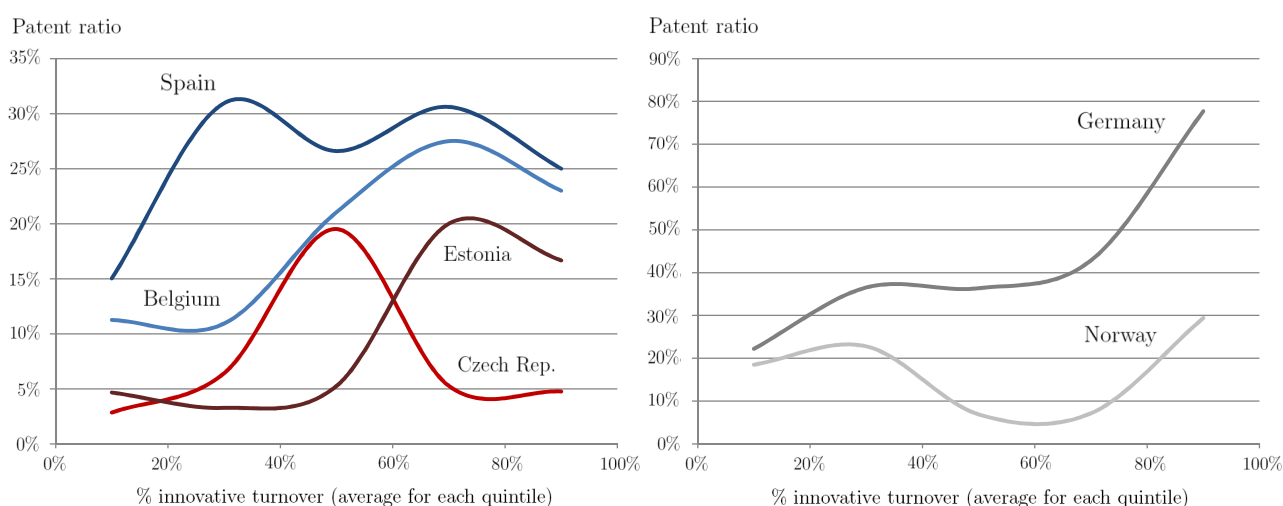


Figure 6: Average patent ratio for quintiles of percentage share of innovative turnover, *left*: Countries with IPR-Index < 7.0, *right*: Countries with IPR-Index > 8.0.

<sup>33</sup> The score for Estonia (5.8) is taken from the 2011 IPR-Index due to a missing observation. The 2007 IPR-Index Report is provided by UNDP at [http://www.undp.org/legalempowerment/pdf/PRA\\_Interior\\_LowRes.pdf](http://www.undp.org/legalempowerment/pdf/PRA_Interior_LowRes.pdf) and the 2011 Report at <http://www.internationalpropertyrightsindex.org>.

The explanatory power of the graphic analysis, however, is limited by the fact that outliers may drive patent ratios, especially for large fractions of innovative turnover, where the number of responses is fairly limited. Moreover, it neglects other factors influencing the propensity to patent potentially interfering with the impact of innovation size on patenting. The following econometric analysis aims at making up for this neglect.

### 3.3 Specifications and estimation results

Patenting behavior is captured by dependent variable  $y_p \in \{0, 1\}$ . We thus estimate a binary response model in the form of a probit regression and check for robustness using a linear probability model. In a first specification, we use the binary measure of innovation size to test

$$y_p = \alpha_0 + \alpha_1 \text{LargeInnovation} + \alpha_2 \text{Controls} + \varepsilon, \quad (8)$$

where  $\varepsilon$  is the error term, and  $\alpha_2$  is a vector of coefficients for the vector of the aforementioned control variables. The strategic view predicts that  $\alpha_1 < 0$  (prediction II), while traditional view and integrative model (prediction I) suggest that  $\alpha_1 > 0$ .<sup>34</sup>

In the second specification of the model, we use the non-binary size measure, the fraction of turnover generated by an innovation that is ‘new to the market’, to test for non-linearities in the relation between patent propensity and innovation size:

$$y_p = \beta_0 + \beta_1 (\% \text{InnovTurnover}) + \beta_2 (\% \text{InnovTurnover})^2 + \beta_3 \text{Controls} + \varepsilon, \quad (9)$$

where, again,  $\varepsilon$  is the error term, and  $\beta_3$  is a vector of coefficients for the control variables. The strategic view predicts that  $\beta_1 < 0$  and  $\beta_2$  is non-significant or also smaller than zero (prediction II), while  $\beta_1 > 0$  with a positive or non-significant  $\beta_2$  would corroborate the traditional view and the integrative model in case of strong patent rights (prediction I). Finally, according to the integrative model with relatively weak patent rights (prediction III), we should have  $\beta_1 > 0$  and  $\beta_2 < 0$ .

Table 3 reports probit estimation results for both specifications over all countries in the sample. The first specification suggests that the propensity to patent increases in innovation size, as all coefficients for the binary size measure are positive and significant at the 1%-level. In case of specification (1a), for instance, the coefficient is +0.541, which, given the negative constant -1.785, implies a propensity to patent of 8.48% for a large invention compared to 1.65% for a small one.<sup>35</sup> Since the extant coefficients of specifications (1a) and (1b) are

<sup>34</sup> Pajak (2010) uses a similar set-up. Recall that we cannot test prediction III using a binary size measure.

<sup>35</sup> We use the standard normal distribution to calculate the percentages as implied by probit regressions.



<i>All countries</i>								
Sample size	(1a)		(1b)		(2a)		(2b)	
	Large	Small	Large	Small	Large	Small	Large	Small
			weighted	weighted			weighted	weighted
<i>Innovation size</i>								
LargeInnovation	0.541*** (12.377)	0.488*** (7.472)	0.532*** (4.789)	0.475*** (3.202)				
% InnovTurnover					2.172*** (8.517)	1.680*** (4.273)	1.958*** (3.238)	2.586*** (2.832)
(% InnovTurnover) <sup>2</sup>					-1.675*** (-5.661)	-1.132** (-2.481)	-1.401* (-1.818)	-2.525** (-2.367)
<i>Firm controls</i>								
Turnover <sup>†</sup>	-0.002 (-0.929)	-0.002 (-0.689)	-0.001 (-0.524)	-0.007 (-0.960)	-0.001 (-0.821)	-0.002 (-0.617)	-0.001 (-0.572)	-0.006 (-0.830)
R&D expenditures <sup>†</sup>	0.002 (1.466)	0.003* (1.752)	0.002 (1.535)	0.002* (1.692)	0.002 (1.407)	0.002* (1.756)	0.002 (1.287)	0.002* (1.703)
Corporate group	0.038 (0.757)	0.033 (0.446)	-0.123 (-1.008)	0.037 (0.212)	0.040 (0.816)	0.018 (0.251)	-0.127 (-1.043)	0.022 (0.126)
<i>Market scale</i>								
Local	-0.156** (-2.296)	-0.236** (-2.424)	-0.254* (-1.806)	-0.354* (-1.813)	-0.153** (-2.246)	-0.223** (-2.294)	-0.276** (-1.977)	-0.339* (-1.712)
National	0.103* (1.655)	-0.001 (-0.007)	0.222 (1.545)	0.264 (1.231)	0.123** (1.980)	0.023 (0.260)	0.239* (1.701)	0.257 (1.202)
European	0.168*** (3.428)	0.070 (0.945)	0.317*** (2.656)	0.164 (0.886)	0.196*** (4.024)	0.107 (1.441)	0.343*** (2.895)	0.188 (1.030)
International	0.257*** (5.182)	0.269*** (3.567)	0.326*** (2.917)	0.538*** (3.336)	0.271*** (5.509)	0.282*** (3.768)	0.338*** (3.024)	0.551*** (3.425)
<i>Fixed effects</i>								
Industry f.e.	✓	✓	✓	✓	✓	✓	✓	✓
Country f.e.	✓	✓	✓	✓	✓	✓	✓	✓
Constant	-1.785*** (-13.302)	-1.693*** (-9.137)	-1.991*** (-7.532)	-2.114*** (-5.572)	-1.677*** (-12.679)	-1.592*** (-8.728)	-1.834*** (-7.209)	-2.041*** (-5.613)
Observations	6732	4864	6732	4864	6732	4864	6732	4864

t-values in parantheses; significance levels: \* p<.10, \*\* p<.05, \*\*\* p<.01; † in millions of €.

Table 3: Probit estimation results for large and small samples with all 14 countries.

of similar size, the results suggest a strong support for prediction I (traditional view and integrative model). This finding is robust to using the weight to control for non-responses and to estimating the linear model instead of a probit regression (table 6 in Appendix B).

Industry-specific estimations using the binary size measure, reported in table 7 in Appendix B, to a large extent also corroborate prediction I. With the exception of ICT, industry results find that the propensity to patent increases in size, albeit in case of electronics the coefficient's p-value is only 16.6%. For ICT, however, the data suggests that patenting is less important, as the size coefficient is negative, albeit not statistically significant. This is well in line with the literature accounting for the cumulative nature of the 'new technologies', the importance of interoperability standards for ICT industries, and their high dependence on intangible assets, calling the usefulness of patents for ICT into question.<sup>36</sup>

Estimation results for the second specification of the model corroborate prediction III that firms keep small and large inventions secret and patent medium-sized ones. Table 3 indicates a positive and highly significant coefficient for the fraction of innovative turnover and a negative and significant coefficient for its square. In case of the small sample, sign and significance level of the coefficients are robust to using the weight to correct for non-responses.

<sup>36</sup> See, e.g., Scotchmer (1991), Heller/Eisenberg (1998), Bessen/Hunt (2007), and Hall/MacGarvie (2010).

But even results for the non-weighted large sample, where the squared size measure only has a p-value of 6.9%, hint at the prevalence of prediction III. For instance, specification (2a) predicts that an increment of the fraction of innovative turnover by 20 percentage points can have an ambivalent effect: raising the non-binary size measure from 20% to 40% increases the propensity to patent from 7.17% to 12.51%, while raising the size measure from 80% to unity reduces the propensity to patent from 14.32% to 9.89%. These results suggest the existence of an inverted-U relation between patenting and innovation size. Again, they are robust to estimating the linear model instead of probit regressions (see table 6).

The findings of both specifications 1 and 2 are not contradictory, as the size measure used in specification 2 provides sufficient information to test for functional forms of the patenting-size-relation that specification 1 cannot account for. Both results are well in line with the integrative model, whose theoretical prediction depends on the strength of patent protection. Since countries differ regarding the protection granted, testing the non-linear patenting-size-relation for all countries depending on a measure for IP protection would undoubtedly improve informative value of the econometric analysis. Unfortunately, most countries in the sample provide insufficient observations to analyze country-specific implications of the model. For that reason, we focus on two countries that provide enough data to be tested separately and exhibit a considerable difference in property rights protection (measured by the IPR-Index): Germany and Spain. Both also exhibit vast differences in survey response rates, as the response to CIS 4 was mandatory in Spain and voluntary in Germany (see table 4), so that a comparison potentially accounts for heterogeneity in that regard, too. Table 8 in Appendix B reports the results for both countries. As predicted by the integrative model, the estimation for Spain (IPR-Index: 6.5) finds significant coefficients supporting the prediction of an inverted-U relation between patenting and size (prediction III). For Germany (IPR-Index: 8.1), however, results draw a slightly different picture. The coefficients' signs hint at an inverted-U relation in line with prediction III, but only 'small sample' results are significant at the 10% level. Yet, the binary size measure is statistically significant for all sample sizes, suggesting that the German data can be better explained by the traditional view. Either way, estimation results find strong support for the integrative model.

## 4 Conclusion

This paper sets up a unifying framework between the traditional view that patents are the most important means of IP protection and the strategic view that firms keep their most important inventions secret. It incorporates both view's theoretical predictions as special

cases and combines them in an integrative model. For countries with strong patent rights, the integrative model yields the same prediction as the traditional view, while for countries with relatively weak patent rights, it predicts an independent third result: an inverted-U relation between propensity to patent and innovation size. Predictions are then confronted with recent European firm survey data, which, for firms with less than 50 employees, contains a binary and a non-binary measure of innovation size. Estimations using the binary size measure find strong support for the traditional view (with the exception of ICT industries). Estimations using the non-binary size measure reveal an inverted-U relation between patenting and innovation size. Both these results are well in line with the integrative model and shed new light on the role of patents and secrecy for different innovation sizes.

Our findings have an impact on several strands of the literature. First, an inverted-U relation between patenting and invention size limits patents' ability to foster welfare-enhancing knowledge diffusion via compulsory disclosure, since it implies that firms patent and disclose only medium-sized inventions and keep large ones secret. The non-linearity of patenting behavior in size thus affects the literature on the welfare effects of patents. Second, it also affects analyses of the riskiness of R&D investments based on estimations of the distribution of innovation sizes. The riskiness of a portfolio of R&D investments depends on the fatness of the size distribution's tail covering the probability that large inventions occur. The inverted-U relation suggests that measuring the tail based on patent data might underestimate its fatness and, therewith, overstate the riskiness of R&D investment, as it predicts that firms rather keep large inventions secret. Third, our results might also affect the literature using patent renewal data to estimate patent values. Since the inverted-U relation suggests that patent data understates the value of innovative activities, it limits the conclusions that can be drawn from this literature for R&D policy.

Finding evidence for two distinct predictions of the integrative model does not limit its empirical validity. On the contrary, the graphic analysis suggests that, as implied by the integrative model, the amount of property rights granted by a country determines which prediction prevails. Country-specific estimations for Germany and Spain corroborate this finding. Yet, for the extant countries in the data set we can neither confirm nor refute this finding, as too few of them obtain enough observations over the whole range of the non-binary size measure. While we use the best data available to us to-date, further investigation of this particular aspect of the integrative model would require a larger data set. Moreover, due to data constraints, our estimations are based on a sample of product innovations for firms with less than 50 employees. Scrutinizing whether the conclusions carry over to a larger sample of innovations by all kinds of firms is also a valuable topic of further research.

# Appendices

## Appendix A

### A.1

*Objective.* Show that in the simultaneous innovation case profits from secrecy,  $\frac{\phi}{4}$ , are strictly smaller than profits from patenting,  $\frac{\phi(m^2+\bar{m}^2)}{(m+\bar{m})^2}$ .

*Proof.* We prove  $\frac{\phi}{4} < \frac{\phi(m^2+\bar{m}^2)}{(m+\bar{m})^2}$  by contradiction, supposing that  $\frac{\phi}{4} \geq \frac{\phi(m^2+\bar{m}^2)}{(m+\bar{m})^2}$ . Rearranging this expression yields  $(m+\bar{m})^2 \geq 2(m^2+\bar{m}^2)$  and  $0 \geq (\bar{m}-m)^2$ , which, due to the exponent, is a contradiction for all  $m < \bar{m}$ , which is implied by the definition of an innovation.  $\square$

### A.2

*Objective.* Show that  $m_1 = \bar{m} \left( \frac{\phi+4c-4\sqrt{c\phi}}{\phi-4c} \right)$  follows from  $V_i^s = V_i^p$ , exists, and is unique for all  $c \in (0, \frac{\phi}{4})$  and  $m \in (0, \bar{m})$ .

*Proof.* From  $V_i^s = V_i^p$  immediately follows that  $\frac{\phi}{4} = \frac{\phi(m^2+\bar{m}^2)}{2(m+\bar{m})^2} - c$ , where, for further reference, we denote the lhs by  $\mathcal{V}^s$  and the rhs by  $\mathcal{V}^p$ . Rearranging the equation yields  $m = \bar{m} \left( \frac{\phi+4c \pm 4\sqrt{c\phi}}{\phi-4c} \right)$ . The expression includes two cut-offs (due to the  $\pm$ ), but only one of them exists within the assumed parameter range. Since  $\phi \in (0, 1]$ , we have that  $4c > 4\sqrt{c\phi}$ , so that the numerator in the brackets is positive in both cases. It follows from the denominator that, for  $m$  to be positive, it must hold that  $0 < c < \frac{\phi}{4}$ . Moreover, since for an innovation  $m < \bar{m}$  always holds, the expression in brackets must be smaller than unity, so that  $\frac{\phi+4c \pm 4\sqrt{c\phi}}{\phi-4c} < 1$  and  $8c \pm 4\sqrt{c\phi} < 0$ . It follows from the second expression that we can rule out the solution adding the square root, as it violates the inequation, leaving us with one solution that satisfies  $m < \bar{m}$ , if  $8c - 4\sqrt{c\phi} < 0$  holds. This is the case for all  $0 < c < \frac{\phi}{4}$ , postulating this parameter restriction to ensure existence of the solution

$$m_1 = \bar{m} \left( \frac{\phi + 4c - 4\sqrt{c\phi}}{\phi - 4c} \right). \quad (5)$$

Uniqueness of this solution requires monotonicity of  $\mathcal{V}^p$ . To this end, it suffices that the first derivative of  $\mathcal{V}^p$  w.r.t.  $m$  is strictly negative, as  $\mathcal{V}^s$  is constant in  $m$ . From  $\mathcal{V}^p = \frac{\phi(m^2+\bar{m}^2)}{2(m+\bar{m})^2} - c$  it follows that  $\frac{\partial \mathcal{V}^p}{\partial m} = \frac{\phi m}{(m+\bar{m})^2} - \frac{\phi(m^2+\bar{m}^2)}{(m+\bar{m})^3}$ , so that  $\frac{\partial \mathcal{V}^p}{\partial m} < 0$  holds, if  $\frac{\phi m}{(m+\bar{m})^2} - \frac{\phi(m^2+\bar{m}^2)}{(m+\bar{m})^3} < 0$  and  $m < \bar{m}$ . Hence, the solution is unique as long as the innovation decreases marginal costs of production, which is implied by its definition.  $\square$

### A.3

*Objective.* Show that  $m_2 = \bar{m} \left( \frac{1-\delta+2\sqrt{1-\delta}}{3+\delta} \right)$ , which follows from equating both expressions in (4), exists and is unique for all  $\delta \in [0, 1)$  and  $m \in (0, \bar{m})$ .

*Proof.* Equating both expressions in (4) gives us  $\frac{1-\delta}{4} = \frac{m^2}{(m+\bar{m})^2} \Leftrightarrow m = \bar{m} \left( \frac{1-\delta \pm 2\sqrt{1-\delta}}{3+\delta} \right)$ . Only the solution adding the square bracket yields an imitation cut-off within the relevant parameter range  $m \in (0, \bar{m})$ , ruling out the expression of  $m$  with a negative square bracket. To see that, write  $m_2 < 0 \Leftrightarrow 1 - \delta + 2\sqrt{1-\delta} < 0$  and  $\delta^2 + 2\delta < 3$ . The latter expression holds for all  $\delta \in [0, 1)$ , so that

$$m_2 = \bar{m} \left( \frac{1 - \delta + 2\sqrt{1-\delta}}{3 + \delta} \right) \quad (10)$$

is the only relevant solution within assumed parameter range. Uniqueness is ensured by the fact that  $\frac{1-\delta}{4}$  is monotonic in  $\delta$  and  $\frac{m^2}{(m+\bar{m})^2}$  is monotonic in  $m$  for all  $m \in (0, \bar{m})$ . To see the latter, take the first derivative of  $\frac{m^2}{(m+\bar{m})^2}$  w.r.t.  $m$ ,  $\frac{2m}{(m+\bar{m})^2} - \frac{2m^2}{(m+\bar{m})^3}$ , and evaluate its sign. It immediately follows from  $\frac{2m\bar{m}}{(m+\bar{m})^3} > 0$  that  $\frac{m^2}{(m+\bar{m})^2}$  is monotonic in  $m$ , so that (10) is a unique solution.  $\square$

#### A.4

*Objective.* Show that  $m_3 = \bar{m} \left( \frac{2}{\sqrt{3\delta+1+\frac{\delta\phi}{1-\phi}}} - 1 \right)$  exists and is unique. Moreover, show that  $m_2 > m_3$  holds for all  $m \in (0, \bar{m})$  and that the empirically plausible case corresponds to  $\delta \in [0, \delta^*)$ .

*Proof.*  $m_3$  follows from  $V_i^s = V_i^p|_{c=0}$  in case of imitation, yielding  $\frac{(1-\phi)\bar{m}^2}{(m+\bar{m})^2} + \frac{\phi}{4} = \frac{(1-\phi)(3\delta+1)}{4} + \frac{\phi(\delta+1)}{4}$   
 $\Leftrightarrow \frac{4\bar{m}^2}{(m+\bar{m})^2} = 3\delta + 1 + \frac{\delta\phi}{1-\phi} \Leftrightarrow m_3 = \bar{m} \left( \frac{\pm 2}{\sqrt{3\delta+1+\frac{\delta\phi}{1-\phi}}} - 1 \right)$ , where we can rule out the negative expression in brackets, since it would yield a negative  $m$  due to  $\delta \geq 0$  and  $\phi \in (0, 1]$ . The resulting  $m_3$  above is unique. This follows from the second equation of the above algebraic transformation: the lhs is monotonic in  $m$ , because  $\frac{\partial \left( \frac{4\bar{m}^2}{(m+\bar{m})^2} \right)}{\partial m} = -\frac{8\bar{m}^2}{(m+\bar{m})^3} < 0$ ; the rhs is monotonic in  $\delta$  and  $\phi$ , as follows immediately from  $3\delta + 1 + \frac{\delta\phi}{1-\phi}$ .

Note that, in case no innovation occurs and patents provide no protection,  $\phi = 0$  and  $\delta = 0$ ,  $m_3$  naturally equals  $\bar{m}$ . For large  $\phi$  and  $\delta$ , on the other hand,  $m_3 = \bar{m} \left( \frac{2}{\sqrt{3\delta+1+\frac{\delta\phi}{1-\phi}}} - 1 \right)$  can become negative, as the fraction on its rhs can become too small to compensate the  $-1$ . A positive  $m_3$  thus requires  $\frac{2}{\sqrt{3\delta+1+\frac{\delta\phi}{1-\phi}}} > 1$ , which can be rearranged to yield the limiting amount of patent strength  $\delta^* = \frac{3-3\phi}{3-2\phi}$  ensuring empirical plausibility (i.e., the fact that not all innovations are patented).

Finally, we prove that within the range  $m \in (0, \bar{m})$  the imitation cut-off is always greater than the patenting cut-off: i.e.,  $m_2 > m_3$  or  $\frac{1-\delta+2\sqrt{1-\delta}}{3+\delta} > \frac{2}{\sqrt{3\delta+1+\frac{\delta\phi}{1-\phi}}} - 1$ , which can be rearranged to  $4 + 4\delta^2 + \frac{\delta^2\phi}{1-\phi} < 8\delta + 5\frac{\delta\phi}{1-\phi} + 4\sqrt{1-\delta} \left( 3\delta + 1 + \frac{\delta\phi}{1-\phi} \right)$ . Let us split the latter expression in two terms and prove each one separately. The first expression focuses on the term  $4\delta^2 + \frac{\delta^2\phi}{1-\phi} < 4\delta + \frac{\delta\phi}{1-\phi} \Leftrightarrow \delta < 1$ , which holds, since we assume  $\delta \in [0, \delta^*)$ , where  $\delta^* = \frac{3-3\phi}{3-2\phi}$  can at most be unity. We can thus get rid of these terms and focus on the extant ones, giving us  $4 < 4\delta \left( 1 + \frac{\phi}{1-\phi} \right) + 4\sqrt{1-\delta} \left( 3\delta + 1 + \frac{\delta\phi}{1-\phi} \right)$ . Since  $\phi \in (0, 1]$ , the inequation becomes an even stricter condition, if we subtract the positive term  $4\frac{\delta\phi}{1-\phi}$ , so that  $m_2 > m_3$  holds a fortiori for  $4 < 4\delta + 4\sqrt{1-\delta} \left( 3\delta + 1 + \frac{\delta\phi}{1-\phi} \right)$ . Rearranging yields  $1 - \delta < \left( 3\delta + 1 + \frac{\delta\phi}{1-\phi} \right)^2$  and, therewith,

$0 < 9\delta^2 + 7\delta + \frac{\delta\phi}{1-\phi} \left(2 + 6\delta + \frac{\delta\phi}{1-\phi}\right)$ . This inequation is satisfied for any  $\delta \in [0, \delta^*)$  and  $\phi \in (0, 1]$ , with the exception of zero patent strength,  $\delta = 0$ , where both cut-offs intersect, i.e.,  $m_2 = m_3$ . However, in this case ( $\delta = 0$ ) both also equal  $\bar{m}$ , which has been ruled out by the definition of an innovation:  $m \in (0, \bar{m})$ . It follows that  $m_2 > m_3$  holds within the defined parameter range.  $\square$

## A.5

*Objective.* Show that, *i*),  $m_{3,c} = \bar{m} \left( \frac{2}{\sqrt{3\delta+1+\frac{\delta\phi-4c}{1-\phi}}} - 1 \right)$  exists and is unique, *ii*), the intersection between  $m_2$  and  $m_{3,c}$  corresponds to  $m_4 = \bar{m} \left( \frac{4-3\phi-4c+2\sqrt{(2-\phi)^2-16c+12\phi c}}{8-5\phi+4c} \right)$ , *iii*), for this intersection,  $m_4 \geq m_1$  always holds, and *iv*), patent strengths  $\delta^*$  and  $\delta_{min}$  in this case are given by  $\delta^* = \frac{3-3\phi+4c}{3-2\phi}$  and  $\delta_{min} = \frac{4(1-\phi+c)}{4-3\phi} - \frac{8(1-\phi)^2+4(1-\phi)\sqrt{(2-\phi)^2-16c+12\phi c}}{(4-3\phi)^2}$ .

*Proof.* From  $V_i^s = V_i^p$ , given by (2) and (3) for the imitation case, it follows that  $\frac{(1-\phi)\bar{m}^2}{(m+\bar{m})^2} + \frac{\phi}{4} = \frac{(1-\phi)(3\delta+1)}{4} + \frac{\phi(\delta+1)}{4} - c$ , which can be solved for  $m$  yielding  $m_{3,c} = \bar{m} \left( \frac{\pm 2}{\sqrt{3\delta+1+\frac{\delta\phi-4c}{1-\phi}}} - 1 \right)$ . Similar to Appendix A.4, the  $m_{3,c}$  given above follows from the fact that a negative square root would yield a negative  $m_{3,c}$ , due to  $\delta \geq 0$ ,  $c \geq 0$ , and  $\phi \in (0, 1]$ , which has been ruled out by  $m \in (0, \bar{m})$ .  $m_{3,c}$  is unique, as the expression in the second line is monotonic in  $c$ . Its monotonicity in the extant variables follows along the lines of Appendix A.4.

Next, let us derive the intersection of  $m_{3,c}$  and  $m_2$ . The latter results from  $\frac{m^2}{(m+\bar{m})^2} = \frac{1-\delta}{4}$ , which instead for  $m$  can be solved for  $\delta$ , yielding  $\delta_2 = \frac{\bar{m}^2+2m\bar{m}-3m^2}{(m+\bar{m})^2}$ . Similarly,  $V_i^s = V_i^p$  above can be solved for  $\delta$ , which gives us  $\delta_3 = \frac{4\bar{m}^2(1-\phi)}{(3-2\phi)(m+\bar{m})^2} + \frac{4c-1+\phi}{3-2\phi}$ . By equating  $\delta_2$  and  $\delta_3$ , yielding  $(\bar{m}^2 + 2m\bar{m} - 3m^2)(3 - 2\phi) = 4\bar{m}^2(1 - \phi) + 4c(m + \bar{m})^2 - (1 - \phi)(m + \bar{m})^2$  and  $0 = m^2 + m \left( \frac{2\bar{m}(4c-4+3\phi)}{8-5\phi+4c} \right) + \frac{\bar{m}^2(4c-\phi)}{8-5\phi+4c}$ , we can solve for  $m$  to obtain

$$m_4 = \bar{m} \left( \frac{4 - 3\phi - 4c \pm 2\sqrt{(2-\phi)^2 - 16c + 12\phi c}}{8 - 5\phi + 4c} \right). \quad (11)$$

In order to see that only a positive square root yields a positive  $m_4$ , firstly note that the denominator is positive due to  $c \in [0, \frac{\phi}{4})$  and  $\phi \in (0, 1]$ . Secondly,  $m_4$  becomes negative, if the numerator due to a negative square root is also negative. To see that, let us square the numerator and write  $(4 - 3\phi - 4c)^2 < 4(4 - 4\phi + \phi^2 - 16c + 12\phi c)$  or  $0 > \phi(5\phi - 8 - 24c) + c(32 + 16c)$ . Due to  $c \in [0, \frac{\phi}{4})$ , filing cost can be rewritten as  $c = \frac{\phi-\epsilon}{4}$ , where  $\epsilon \in (0, \phi]$  is a dummy variable accounting for all possible values for  $c$ . Inserting this in the inequation yields  $4\phi + \epsilon < 8$ , which for the assumed parameter range holds true. It follows that adding the square root in (11) is the only relevant solution for  $m_4$ . Note that we can use the same method to prove that the square root exists for the given parameter range, i.e., the expression under the square root is positive: with  $c = \frac{\phi-\epsilon}{4}$ , we have that  $\sqrt{(2-\phi)^2 - 16c + 12\phi c} = \sqrt{(2-2\phi)^2 + 4\epsilon - 3\phi\epsilon}$ , which is positive due to  $\phi \in (0, 1]$  and  $\epsilon \in (0, \phi]$ . Since with  $c = \frac{\phi-\epsilon}{4}$ ,  $4 - 3\phi - 4c = 4 - 4\phi + \epsilon > 0$ ,  $m_4$  in (11) is positive for the relevant

parameter range in case the square root is added.

For Proposition 4 to give a clear-cut distinction between patent strengths, which are too small to enable patenting, and empirically plausible patent strengths, where secrecy *and* patenting occurs, it is necessary to show that  $m_4 \geq m_1$ . If this inequality holds, the limiting patent strength for empirical plausibility,  $\delta_{min}$ , corresponds to  $m_4$  instead of the intersection between  $m_1$  and  $m_2$ , so that  $\delta_{min}$  can be derived (see below) by inserting  $m_4$  into  $\delta_2$ . Hence, showing that  $m_4 \geq m_1$  enables us to give a clear lower border for empirically plausible  $\delta$ : the patent strength corresponding to  $m_4$ .

Using expressions (5) and (11),  $m_4 \geq m_1$  can be rewritten to

$$(4 - 4c - 3\phi + 2\sqrt{(2 - \phi)^2 - 16c + 12\phi c})(\phi - 4c) \geq (\phi + 4c - 4\sqrt{\phi c})(8 - 5\phi + 4c),$$

where the inequality is unaffected by the elimination of denominators of  $m_4$  and  $m_1$ , as both are positive due to  $\phi \in (0, 1]$  and  $c \in [0, \frac{\phi}{4}]$ . Further rearranging yields

$$12\phi c + \phi^2 + \sqrt{\phi c}(16 - 10\phi + 8c) + \phi\sqrt{\cdot} \geq 24c + 2\phi + 4c\sqrt{\cdot},$$

where for notational simplicity we used  $\sqrt{\cdot} = \sqrt{(2 - \phi)^2 - 16c + 12\phi c}$ . Let us denote the rhs and lhs of the expression by  $\mathcal{A}$  and  $\mathcal{B}$ , respectively, so that  $\mathcal{A} = 24c + 2\phi + 4c\sqrt{\cdot}$  and  $\mathcal{B} = 12\phi c + \phi^2 + \sqrt{\phi c}(16 - 10\phi + 8c) + \phi\sqrt{\cdot}$ . For  $m_4 \geq m_1$  to hold, it suffices to show that  $\mathcal{B} \geq \mathcal{A}$ . This condition is fulfilled with equality for the limiting cases of filing costs  $c = 0$  and  $c \rightarrow \frac{\phi}{4}$ . In the former case, we have that  $\mathcal{A} = 2\phi = \mathcal{B}$ , while in the latter case both expressions equal  $8\phi + \phi\sqrt{\cdot}$ . For the remaining parameter values,  $\mathcal{B}$  is greater than  $\mathcal{A}$ , if *a*), at least for very small filing costs  $\frac{\partial \mathcal{B}}{\partial c} > \frac{\partial \mathcal{A}}{\partial c}$  holds, and *b*), the second derivatives of both expressions are strictly smaller than zero. These conditions suffice to prove  $m_4 \geq m_1$ , because with  $\mathcal{A}$  and  $\mathcal{B}$  intersecting at  $c = 0$  and  $c \rightarrow \frac{\phi}{4}$ , condition *b*) ensures that there are no other intersections, so that if condition *a*) shows that  $\mathcal{B} > \mathcal{A}$  for small values of  $c$ , this must be the case for all other  $c$  within the range. In other words, with strictly negative second derivatives, neither  $\mathcal{A}$  nor  $\mathcal{B}$  has an inflection point, so that the derived intersections must be the only ones. Together with  $\frac{\partial \mathcal{B}}{\partial c}|_{c=0} > \frac{\partial \mathcal{A}}{\partial c}|_{c=0}$ , this suffices to show that  $m_4 \geq m_1$ .

Let us first show that  $\frac{\partial \mathcal{B}}{\partial c}|_{c=0} > \frac{\partial \mathcal{A}}{\partial c}|_{c=0}$  holds. The first derivatives are  $\frac{\partial \mathcal{A}}{\partial c} = 24 + 4\sqrt{\cdot} + \frac{c(24\phi - 32)}{\sqrt{\cdot}}$  and  $\frac{\partial \mathcal{B}}{\partial c} = \frac{\phi(8 - 5\phi + 4c)}{\sqrt{\phi c}} + 8\sqrt{\phi c} + \frac{\phi(6\phi - 8)}{\sqrt{\cdot}} + 12\phi$ . In order to avoid a division by zero when setting  $c = 0$ , we simply scale up both derivatives by  $\sqrt{\phi c}$ , yielding  $\frac{\partial \mathcal{A}}{\partial c} \sqrt{\phi c} = 24\sqrt{\phi c} + 4\sqrt{\cdot} \sqrt{\phi c} + \frac{c(24\phi - 32)\sqrt{\phi c}}{\sqrt{\cdot}}$  and  $\frac{\partial \mathcal{B}}{\partial c} \sqrt{\phi c} = \phi(8 - 5\phi + 4c) + 8\phi c + \frac{\phi(6\phi - 8)\sqrt{\phi c}}{\sqrt{\cdot}} + 12\phi\sqrt{\phi c}$ . Evaluating both at  $c = 0$  gives us  $\frac{\partial \mathcal{A}}{\partial c}|_{c=0} = 0 < \phi(8 - 5\phi) = \frac{\partial \mathcal{B}}{\partial c}|_{c=0}$ , which holds true of all  $\phi \in (0, 1]$  and  $c \in [0, \frac{\phi}{4}]$ .

Next, we prove condition *b*) requiring that  $\frac{\partial^2 \mathcal{A}}{\partial c^2} < 0$  and  $\frac{\partial^2 \mathcal{B}}{\partial c^2} < 0$ . From  $\frac{\partial^2 \mathcal{A}}{\partial c^2} = \frac{4(12\phi - 16)}{\sqrt{\cdot}} - \frac{c(12\phi - 16)^2}{(\sqrt{\cdot})^3}$  immediately follows that it is strictly smaller than zero, as  $12\phi < 16$  for all  $\phi \in (0, 1]$  and  $\sqrt{\cdot} > 0$  has been shown above. To also show that  $\frac{\partial^2 \mathcal{B}}{\partial c^2} = -\frac{\phi^2(8 - 5\phi + 4c)}{2(\sqrt{\phi c})^3} + \frac{8\phi}{\sqrt{\phi c}} - \frac{\phi(6\phi - 8)^2}{(\sqrt{\cdot})^3} < 0$ , we rearrange this expression to  $\phi(5\phi - 8 + 12c)(\sqrt{\cdot})^3 - 8(3\phi - 4)^2(\sqrt{\phi c})^3 < 0$ . Since filing costs

by assumption range between zero and a quarter of  $\phi$ , we further simplify using the aforementioned  $c = \frac{\phi - \epsilon}{4}$ , where again  $\epsilon \in (0, \phi]$ . Hence, the inequation becomes

$$0 > \phi(8\phi - 8 - 3\epsilon) [(2 - 2\phi)^2 + 4\epsilon - 3\phi\epsilon]^{\frac{3}{2}} - (3\phi - 4)^2 [\phi(\phi - \epsilon)]^{\frac{3}{2}}.$$

Given  $\phi \in (0, 1]$  and  $\epsilon \in (0, \phi]$ , this inequation holds, as the first expression on the rhs is negative (due to  $8\phi < 8 + 3\epsilon$ ) and is multiplied by a positive expression (due to  $4 > 3\phi$  and the squared term  $(2 - 2\phi)^2$ ). The second to last bracket in the inequation is positive due to the even exponent, while the last one is either positive (for all  $\epsilon \in (0, \phi)$ ) or zero (for  $\epsilon = \phi$ ). In either case, both terms are weakly smaller than zero, so that the inequality holds. Together with condition *a*) and the limiting cases ( $c = 0$  and  $c \rightarrow \frac{\phi}{4}$ ) above, it follows that  $m_4 \geq m_1$  holds.

Finally, we can show that  $\delta^* = \frac{3-3\phi+4c}{3-2\phi}$  immediately follows from equating  $V_i^s = V_i^p$  (imitation case) and evaluating them at  $m = 0$ , which after rearranging yields  $\frac{\bar{m}^2}{(0+\bar{m})^2} + \frac{c}{1-\phi} = \frac{3\delta+1}{4} + \frac{\delta\phi}{4(1-\phi)}$ , resulting in the above given  $\delta^*$ . In order to derive  $\delta_{min}$ , we insert  $m_4$  given by (11) (with a positive square root) into  $\delta_2 = \frac{\bar{m}^2+2m\bar{m}-3m^2}{(m+\bar{m})^2}$  derived earlier in this section:

$$\delta_{min} = \frac{(8 - 5\phi + 4c)^2 + (16 - 10\phi + 8c)(4 - 3\phi - 4c + 2\sqrt{\cdot}) - 3(4 - 3\phi - 4c + 2\sqrt{\cdot})^2}{(12 - 8\phi + 2\sqrt{\cdot})^2}.$$

This can be rearranged to yield  $\delta_{min} = \frac{4(1-\phi+c)}{4-3\phi} - \frac{8(1-\phi)^2+4(1-\phi)\sqrt{(2-\phi)^2-16c+12\phi c}}{(4-3\phi)^2}$  as ought to be shown.  $\square$

## Appendix B

	Response rate	No. of responses large sample			No. of responses small sample		
		No patent filed	Patent filed	Total	No patent filed	Patent filed	Total
Belgium	39% <sup>v</sup>	418	59	477	349	33	382
Czech Rep.	74%	582	29	611	449	6	455
Estonia	78%	351	18	369	292	5	297
Germany	41% <sup>v</sup>	445	147	592	319	58	377
Greece	n/a	169	6	175	121	1	122
Hungary	76%	217	19	236	194	11	205
Latvia	96% <sup>m</sup>	41	6	47	36	1	37
Lithuania	93% <sup>m</sup>	102	8	110	88	2	90
Norway	95% <sup>m</sup>	549	125	674	429	63	492
Portugal	74% <sup>m</sup>	498	37	535	395	6	401
Romania	79% <sup>m</sup>	422	42	464	347	20	367
Slovakia	72%	95	2	97	71	1	72
Slovenia	87% <sup>m</sup>	116	13	129	99	3	102
Spain	91% <sup>m</sup>	2,195	474	2,669	1,690	119	1,809
Total		6,200	985	7,185	4,879	329	5,208

<sup>v</sup> and <sup>m</sup> mark voluntary and mandatory questionnaires, respectively.

Table 4: CIS 4 response rates and number of responses by country (only product innovations and firms with less than 50 employees). *Sources:* CIS 4 data set provided by Eurostat; [www.medibtikar.eu/IMG/ppt/UNESCOAmman\\_Session2\\_4LT.ppt](http://www.medibtikar.eu/IMG/ppt/UNESCOAmman_Session2_4LT.ppt); OECD (2009): “Innovation in firms - a microeconomic perspective”, Annex A.



<i>All countries</i>		Large sample			Small sample			Description
Variable	Obs	Mean	Min-Max	Obs	Mean	Min-Max		
Patent filed <sup>b</sup>	7,185	.137	0-1	5,208	.063	0-1	One, if firm filed a patent application within the observation period.	
Large innovation <sup>b</sup>	7,185	.529	0-1	5,208	.484	0-1	One, if firm generated an innovation that is 'new to the market'.	
% InnovTurnover	7,185	.124	0-1	5,208	.109	0-1	Percentage share of turnover generated by an innovation that is 'new to the market'.	
Turnover	7,130	4.828	0-6,614	5,169	5.177	0-6,614	A firm's total turnover of 2004 in millions of €.	
R&D exp.	6,972	1.568	0-729	5,035	1.729	0-729	A firm's total R&D expenditures in 2004 in millions of €.	
Corporate group <sup>b</sup>	7,185	.256	0-1	5,208	.247	0-1	One, if firm belongs to a corporate group.	
Local <sup>b</sup>	7,020	.703	0-1	5,089	.706	0-1	One, if firm sold its goods in regional market between 2002 and 2004.	
National <sup>b</sup>	7,128	.795	0-1	5,167	.774	0-1	One, if firm sold its goods in national market between 2002 and 2004.	
E.U.-wide <sup>b</sup>	7,129	.497	0-1	5,163	.795	0-1	One, if firm sold its goods within the E.U. between 2002 and 2004.	
International <sup>b</sup>	7,075	.277	0-1	5,122	.244	0-1	One, if firm sold its goods in markets outside the E.U. between 2002 and 2004.	
Industry f.e. <sup>b</sup>	Five categories: 1. 'Information & Communication Technologies': nace classes '64' and '72'; 2. 'Chemicals & Plastics': nace classes '23', '24' and '25'; 3. 'Machinery & metal products': nace classes '27', '28' and '29'; 4. 'Electronics': nace classes '30', '31', '32', '33', '40'; 5. 'Miscellaneous': extant nace classes in data set. Note that CIS 4 contains only a limited number of nace classes.							
Country f.e. <sup>b</sup>	One for all observations in a particular country, 'Belgium' used as reference.							

*b* marks binary variables.

Table 5: Descriptive statistics of large and small sample.

<i>All countries</i>	(1a)		(1b)		(2a)		(2b)	
Sample size	Large	Small	Large	Small	Large	Small	Large	Small
			weighted	weighted			weighted	weighted
<i>Innovation size</i>								
LargeInnovation	0.106*** (12.670)	0.054*** (7.334)	0.099*** (4.837)	0.044*** (2.890)				
% InnovTurnover					0.439*** (7.439)	0.186*** (3.607)	0.378*** (3.027)	0.248** (2.295)
(% InnovTurnover) <sup>2</sup>					-0.330*** (-4.482)	-0.113* (-1.704)	-0.281* (-1.655)	-0.252** (-2.105)
<i>Firm controls</i>								
Turnover <sup>†</sup>	-0.000*** (-4.018)	-0.000*** (-3.463)	-0.000 (-0.989)	-0.000 (-1.436)	-0.000*** (-3.834)	-0.000*** (-3.073)	-0.000 (-0.634)	-0.000 (-1.468)
R&D expenditures <sup>†</sup>	0.000 (1.549)	0.000 (1.377)	0.000 (1.259)	0.000 (1.177)	0.000 (1.339)	0.000 (1.374)	0.000 (1.102)	0.000 (1.352)
Corporate group	0.007 (0.689)	0.003 (0.405)	-0.024 (-1.089)	-0.007 (-0.362)	0.009 (0.906)	0.004 (0.454)	-0.024 (-1.074)	-0.007 (-0.390)
<i>Market scale</i>								
Local	-0.024* (-1.895)	-0.026** (-2.221)	-0.040 (-1.464)	-0.034 (-1.319)	-0.023* (-1.770)	-0.024** (-2.112)	-0.044 (-1.581)	-0.034 (-1.352)
National	0.019* (1.925)	-0.000 (-0.022)	0.032* (1.657)	0.012 (0.751)	0.022** (2.240)	0.002 (0.211)	0.033* (1.744)	0.012 (0.778)
European	0.032*** (3.516)	0.006 (0.830)	0.055** (2.528)	0.010 (0.568)	0.037*** (3.994)	0.009 (1.198)	0.058*** (2.672)	0.011 (0.614)
International	0.064*** (5.479)	0.041*** (3.940)	0.074*** (2.750)	0.076*** (2.953)	0.271*** (5.509)	0.282*** (3.768)	0.077*** (2.829)	0.078*** (3.006)
<i>Fixed effects</i>								
Industry f.e.	✓	✓	✓	✓	✓	✓	✓	✓
Country f.e.	✓	✓	✓	✓	✓	✓	✓	✓
Constant	0.000 (0.006)	0.053** (2.240)	-0.013 (-0.290)	0.032 (0.812)	0.017 (0.645)	0.061*** (2.587)	0.012 (0.260)	0.038 (0.974)
Observations	6732	4864	6732	4864	6732	4864	6732	4864

Robust s.e.; t-values in parantheses; significance levels: \*  $p < .10$ , \*\*  $p < .05$ , \*\*\*  $p < .01$ ; † in millions of €.

Table 6: OLS estimation results for large and small samples over all 14 countries.

<i>All countries</i>	ICT	Chemicals &Plastics	Machinery & Metals	Electronics	Misc.
	weighted	weighted	weighted	weighted	weighted
<i>Innovation size</i>					
LargeInnovation	-0.409 (-1.513)	0.960*** (4.246)	0.583** (2.403)	0.334 (1.385)	0.586*** (3.545)
<i>Firm controls</i>					
Turnover <sup>†</sup>	0.010 (1.338)	0.004 (0.462)	0.001 (0.566)	-0.003** (-2.314)	-0.001 (-0.224)
R&D expenditures <sup>†</sup>	0.003* (1.837)	0.122 (1.143)	0.003** (2.132)	-0.015 (-1.465)	0.002 (0.847)
Corporate group	-0.376 (-1.039)	0.096 (0.362)	-0.314 (-1.087)	0.017 (0.067)	-0.130 (-0.710)
<i>Market scale</i>					
Local	-0.407 (-0.926)	-0.297 (-0.924)	-0.666* (-1.930)	-0.211 (-0.728)	-0.144 (-0.696)
National	-0.339 (-0.945)	0.158 (0.361)	0.518 (1.563)	0.245 (0.818)	0.297 (1.590)
European	0.524 (1.627)	0.105 (0.392)	0.013 (0.051)	0.561** (2.058)	0.323* (1.954)
International	0.409 (1.211)	-0.195 (-0.766)	0.589** (2.171)	-0.162 (-0.655)	0.429*** (2.659)
<i>Fixed effects</i>					
Industry f.e.	✓	✓	✓	✓	✓
Country f.e.	✓	✓	✓	✓	✓
Constant	-0.607 (-0.787)	-1.308* (-1.950)	-1.478*** (-2.626)	-1.587*** (-2.707)	-2.290*** (-5.962)
Observations	701	599	824	616	3865

t-values in parantheses; significance levels: \* p<.10, \*\* p<.05, \*\*\* p<.01; † in millions of €.

Table 7: Probit estimation results over all 14 countries by industry.

<i>Country</i> Specification Sample Size	<i>Spain</i>				<i>Germany</i>			
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
	Large weighted	Small weighted	Large weighted	Small weighted	Large weighted	Small weighted	Large weighted	Small weighted
<i>Innovation size</i>								
LargeInnovation	0.644*** (4.045)	0.559*** (2.863)			0.414** (2.133)	0.391* (1.714)		
% InnovTurnover			2.568*** (3.202)	2.733** (2.337)			1.832 (1.326)	3.452* (1.951)
(% InnovTurnover) <sup>2</sup>			-1.955* (-1.737)	-2.747* (-1.936)			-1.778 (-1.052)	-4.260* (-1.865)
<i>Firm controls</i>								
Turnover <sup>†</sup>	-0.003 (-0.401)	-0.008 (-0.700)	-0.002 (-0.469)	-0.007 (-0.711)	-0.001 (-0.478)	-0.057 (-1.603)	-0.001 (-0.573)	-0.059* (-1.748)
R&D expenditures <sup>†</sup>	0.186* (1.931)	0.419*** (3.006)	0.187* (1.941)	0.436*** (3.171)	0.254 (0.767)	0.438 (1.134)	0.267 (0.823)	0.449 (1.199)
Corporate group	-0.460** (-2.208)	0.329 (1.070)	-0.450** (-2.150)	0.309 (1.027)	-0.102 (-0.543)	-0.036 (-0.156)	-0.120 (-0.623)	-0.043 (-0.187)
<i>Market scale</i>								
Local	-0.009 (-0.042)	-0.039 (-0.178)	-0.031 (-0.154)	-0.079 (-0.376)	-0.339* (-1.904)	-0.545** (-2.361)	-0.367** (-2.065)	-0.538** (-2.278)
National	0.001 (0.004)	0.588*** (3.077)	0.058 (0.241)	0.623*** (3.374)	0.376* (1.710)	0.272 (0.769)	0.363* (1.675)	0.218 (0.612)
European	0.577*** (2.994)	0.624** (2.329)	0.576*** (3.054)	0.615** (2.412)	0.356* (1.682)	0.226 (0.839)	0.368* (1.721)	0.223 (0.814)
International	0.120 (0.612)	-0.323 (-1.044)	0.196 (0.968)	-0.246 (-0.838)	0.367* (1.858)	0.831*** (3.270)	0.365* (1.825)	0.826*** (3.247)
<i>Fixed effects</i>								
Industry f.e.	✓	✓	✓	✓	✓	✓	✓	✓
Country f.e.	✓	✓	✓	✓	✓	✓	✓	✓
Constant	-1.777*** (-5.325)	-3.249*** (-11.169)	-1.702*** (-5.527)	-3.141*** (-11.791)	-1.735*** (-6.587)	-1.863*** (-4.459)	-1.624*** (-6.366)	-1.796*** (-4.581)
Observations	2669	1809	2669	1809	592	377	592	377

t-values in parantheses; significance levels: \* p<.10, \*\* p<.05, \*\*\* p<.01; † in millions of €.

Table 8: Probit estimation results for Spain and Germany.

## References

- Anton, J.J./Yao, D.A. (2004): Little patents big secrets: managing intellectual property, *RAND Journal of Economics*, 35(1), 1-22.
- Arundel, A. (2001): The relative effectiveness of patents and secrecy for appropriation, *Research Policy*, 30(4), 611-624.
- Arundel, A./Kabla, I. (1998): What percentage of innovations are patented? - Empirical estimates for european firms, *Research Policy*, 27(2), 127-141.
- Bandyopadhyay, S. (1997): Demand elasticities, asymmetry and strategic trade policy, *Journal of International Economics*, 42(1-2), 167-177.
- Baudry, M./Dumont, B. (2009): A Bayesian real option approach to patents and optimal renewal fees, *mimeo*, Université de Nantes.
- Bessen, J./Hunt, R.M. (2007): An empirical look at software patents, *Journal of Economics & Management Strategy*, 16(1), 157-189.
- Bhattacharya, S./Guriev, S. (2006): Patents vs. trade secrets: knowledge licensing and spillover, *Journal of the European Economic Association*, 4(6), 1112-1147.
- Brouwer, E./Kleinknecht, A. (1999): Innovative output, and a firm's propensity to patent. An exploration of CIS micro data, *Research Policy*, 28(6), 615-624.
- Cohen, W./Nelson, R./Walsh, J. (2000): Protecting their intellectual assets - Why US manufacturing firms patent or not, *NBER working paper* 7552.
- Denicolo, V. (1996): Patent races and optimal patent breadth and length, *Journal of Industrial Economics*, 44(3), 249-265.
- Denicolo, V./Franzoni, L.A. (2004): Patents, secrets, and the first-inventor defense, *Journal of Economics & Management Science*, 13(3), 517-538.
- Erkal, N. (2004): On the interaction between patent policy and trade secret policy, *Australian Economic Review*, 37(4), 427-435.
- Friedman, D.D./Landes, M./Posner, R.A. (1991): Some economics of trade secret law, *Journal of Economic Perspectives*, 5(1), 61-72.
- Grönqvist, C. (2009): The private value of patents by patent characteristics: evidence from Finland, *Journal of Technology Transfer*, 34(2), 159-168.
- Hall, B./MacGarvie, M. (2010): The private value of software patents, *Research Policy*, 39(7), 994-1009.
- Heller, M.A./Eisenberg, R.S. (1998): Can patents deter innovation? The anticommons in biomedical research, *Science*, 280(5364), 698-701 .
- Hopenhayn, H./Llobet, G./Mitchell, M. (2006): Rewarding sequential innovators: Prizes, patents, and buyouts, *Journal of Political Economy*, 114(6), 1041-1068.
- Hussinger, K. (2004): Is silence golden? – Patents versus secrecy at the firm level, *Economics of Innovation and New Technology*, 15(8), 735-752.
- Kultti, K./Takalo, T./Toikka, J. (2006): Simultaneous model of innovation, secrecy, and patent policy, *American Economic Review*, 96(2), 82-86.
- Kultti, K./Takalo, T./Toikka, J. (2007): Secrecy vs. patenting, *RAND Journal of Economics*, 38(1), 22-42.

- Lanjouw, J.O. (1998): Patent protection in the shadow of infringement: Simulation estimations of patent value, *Review of Economic Studies*, 65(225), 671-710.
- Lemley, M.A./Shapiro, C. (2005): Probabilistic patents, *Journal of Economic Perspectives*, 19(2), 75-98.
- Mansfield, E. (1986): Patents and innovation - An empirical study, *Management Science*, 32(2), 173-181.
- Mosel, M. (2011a): Competition, imitation, and R&D productivity in a growth model with industry-specific patent protection, *Review of Law and Economics*, 7(2), 601-652.
- Mosel, M. (2011b): Big patents, small secrets: how firms protect inventions when R&D outcome is heterogeneous, *BGPE Discussion Paper Series*, No. 105.
- Nordhaus, W.D. (1969): *Invention, growth, and welfare - A theoretical treatment of technological change*, Cambridge Univ. Press, Cambridge Mass.
- O'Donoghue, T./Scotchmer, S./Thisse, J. (1998): Patent breadth, patent life, and the pace of technological progress, *Journal of Economics & Management Strategy*, 7(1), 1-32.
- Pajak, S. (2010): Do innovative firms rely on big secrets? An analysis of IP protection strategies with the CIS 4 survey, *mimeo*, available at: [https://editorialexpress.com/cgi-bin/conference/download.cgi?db\\_name=JEI2010&paper\\_id=35](https://editorialexpress.com/cgi-bin/conference/download.cgi?db_name=JEI2010&paper_id=35).
- Pottelsberghe de la Potterie, B. van/Francois, D. (2009): The cost factor in patent systems, *Journal of Industry Competition and Trade*, 9(4), 329-355.
- Schankerman, M. (1998): How valuable is patent protection? Estimates by technology field, *RAND Journal of Economics*, 29(1), 77-107.
- Schankerman, M./Pakes, A. (1986): Estimates of the value of patent rights in European countries during the post-1950 period, *Economic Journal*, 96(384), 1052-1076.
- Scherer, F.M. (1972): Nordhaus' theory of optimal patent life: a geometric reinterpretation, *American Economic Review*, 62(3), 422-427.
- Scherer, F.M./Harhoff, D. (2000): Technology policy for a world of skew-distributed outcomes, *Research Policy*, 29(4-5), 559-566.
- Scotchmer, S. (1991): Standing on the shoulders of giants: cumulative research and the Patent Law, *Journal of Economic Perspectives*, 5(1), 29-41.
- Silverberg, G./Verspagen, B. (2007): The size distribution of innovations revisited: an application of extreme value statistics to citation and value measures of patent significance, *Journal of Econometrics*, 139(2), 318-339.
- Tandon, P. (1982): Optimal patents with compulsory licensing, *Journal of Political Economy*, 90(3), 470-486.