GOVERNANCE AND THE EFFICIENCY OF ECONOMIC SYSTEMS
GESY

Discussion Paper No. 416

Deferred Patent Examination

Ilja Rudyk *

* University of Munich

June 2013

Financial support from the Deutsche Forschungsgemeinschaft through SFB/TR 15 is gratefully acknowledged.

Sonderforschungsbereich/Transregio 15 · www.sfbtr15.de
Universität Mannheim · Freie Universität Berlin · Humboldt-Universität zu Berlin · Ludwig-Maximilians-Universität München
Rheinische Friedrich-Wilhelms-Universität Bonn · Zentrum für Europäische Wirtschaftsforschung Mannheim

Speaker: Prof. Dr. Klaus M. Schmidt · Department of Economics · University of Munich · D-80539 Munich,
Phone: +49(89)2180 2250 · Fax: +49(89)2180 3510

Deferred Patent Examination

Ilja Rudyk*

June 2013

Abstract

Most patent systems allow applicants to defer patent examination by some time. Deferred examination was introduced in the 1960s, first at the Dutch patent office and subsequently in many other countries, as a response to mounting backlogs of unexamined patent applications. Some applicants allow the examination option to lapse and never request examination once they learn about the value of their invention. Examination loads are reduced substantially in these systems, albeit at the cost of having a large number of pending patent applications. Economic models of patent examination and renewal have largely ignored this important feature to date. We construct a model of patent application, examination and renewal in which applicants have control over the timing of examination and study the tradeoffs that applicants face. Using data from the Canadian patent office and a simulated GMM estimator, we obtain estimates for parameter values of the value distributions and of the learning process. We use our estimates to assess the value of Canadian patents as well as applications. We find that a considerable part of the value is realized before a patent is even granted. In addition, we simulate the counterfactual impact of changes in the deferment period. The estimates we obtain for the value of one additional year of deferment are relatively high and may explain why some applicants embark on delay tactics (such as continuations or divisionals) in patent systems without a statutory deferment option.

Keywords: patent, patent value, value of patent applications, patent examination, deferred patent examination

*INNO-tec, Ludwig-Maximilians-Universität (LMU), Munich.

1 Introduction

Traditionally, the literature on the economics of innovation (e.g., Pakes and Schankerman 1984; Pakes 1986; Schankerman and Pakes 1986; Lanjouw 1998) has exploited post-grant patent renewals to analyze the value of patents and its role in incentivizing R&D (Cornelli and Schankerman 1999; Scotchmer 1999). We refer to patent renewal as a patentee's decision to pay the required maintenance fees to maintain an already issued patent right. These fees are charged by the national patent offices and are due at several points in the life of a patent. However, the patent term does not start with the date a patent is granted but already with the filing date of the patent application.[1] Prior to the grant a patent application has to be examined. Additionally, most patent offices allow to defer the examination request for several years. Indeed, many patents exist longer as a pending application than as a granted patent. Nevertheless, the timing issues in the early stages of a patent's life have largely been neglected.

This paper addresses the research gap by providing a structural model of the application, examination, and renewal process in patent offices. By extending previous patent renewal models with an option to defer patent examination and by modeling examination itself in detail, we provide a much richer foundation for patent valuation and for policy simulations than previous studies have done. Under deferred examination, applicants have the option of requesting examination at some point in time. Patent offices may differ with respect to the time period during which examination can be requested as well as to the fees associated with examination and the maintenance of patent filings. While a few patent offices, notably the USPTO (US Patent and Trademark Office), follow a policy of automatic and–if possible–immediate examination, other offices such as the German patent office offer applicants a time period of up to seven years during which they can request examination.[2] The timing of examination constitutes one of the most startling institutional differences between different patent systems, but it has not received much attention so far.[3]

The framework will further allow us to contribute to the academic debate on how to handle patent backlogs. In the last three decades the number of patent filings has risen substantially: partially due to the increased tactical and strategic importance (Hall and Ziedonis 2001; F.T.C. 2003; N.R.C. 2004) and partially due to the lower costs and availability of patent

[1] Prior to 1995 the patent term in the US was 17 years following the grant date but was modified to 20 years following the patent application date.

[2] The USPTO recently announced a move towards deferred examination and to let applicants choose from three examination tracks: the examination timing as previously offered, a fast-track option for applicants seeking fast examination (similar to the option of accelerated examination at the EPO), and finally a three-year deferment option. During this time period the USPTO would not undertake any substantive examination. (Cf. http://www.uspto.gov/news/pr/2011/11-24.jsp for details.)

[3] See Harhoff (2012) for a more detailed description of deferment systems in 35 countries.

protection (Harhoff 2006; Guellec and van Pottelsberghe 2007; Bessen and Meurer 2008). As a consequence, the patent workload has increased substantially giving reason for concern about its impact on examination quality. Deferred patent examination may constitute a solution to the problem. This system was first introduced by the Dutch government on January 1, 1964 as a reaction to the vast amount of unexamined and pending patent applications. They observed that many patents lapsed already shortly after grant despite low renewal fees. The possibility to defer the examination request for up to seven years allowed the patentees to abandon applications with no commercial value without any examination. Indeed, Yamauchi and Nagaoka (2008), who try to explain the rapid increase in the number of requests for patent examination in Japan in the recent decade, conclude that one of the causes of the increase was the shortening of the period of examination requests. The workload of examiners in Japan has been increased with low quality patents.

Opting for fast examination entails a number of advantages. The main argument is that uncertainty for users of the system is reduced quickly. Both applicants and their rivals will learn soon after the filing date about the actual delineation of patent claims and possible infringement, and they may then adapt their investments accordingly. The argument that uncertainty over examination outcomes and long pendencies have negative consequences is intuitively appealing and has found some empirical support (Gans et al. 2008). Some practitioners have argued that applicants are intentionally increasing the volume and complexity of their filings, frequently delay the examination process, and thus create uncertainty for other users of the system.[4] They argue that such delay tactics should be sanctioned by patent offices.

However, delayed examination has advantages, too. Giving applicants additional time for assessing the value of their patents may lead them to drop out of the examination process voluntarily, and thus reduce examination workloads. While this effect has been discussed in the literature for some time, it has not been captured in structural models of patent examination. In comparison to the classical models of patent renewal, a model of deferred examination has to allow for three possible decisions: to request examination, to defer examination, and to let the application (respectively the granted patent) lapse altogether. We embed these three choices in a model of applicant decision-making, allowing the applicant to make optimal decisions in each period, given the information he has received so far, his knowledge of the overall distribution of patent value, and its expected evolution over time. Aside from adding an important feature to the choice set of applicants, we also employ a more detailed model of patent examination in which the applicant may drop the application after receiving a signal from the examiner. We allow unexamined patent applications to differ in terms of value from examined and granted applications. In the empirical part of our paper, we use data from

[4] See McGinley (2008), Opperman (2009), and Harhoff and Wagner (2009).

the Canadian patent office (CIPO) to estimate the parameters of the value distribution of Canadian applications and of the learning process.

Our estimates of patent value match those of earlier studies. Furthermore, results reveal that the private value of having a pending application is substantial. The returns from just having an unexamined patent application exceed the costs for keeping it in force for the majority of the applicants, even if they will never get a patent granted. The model estimates also provide insights into the learning process during the application and examination stages. Learning possibilities are relatively high and deteriorate only slowly over time.

Additionally, we employ the parameter estimates to estimate the impact of deferment on patent office workload and on the value of unexamined as well as granted applications. The policy experiments indicate that each additional year of deferment would significantly reduce the number of examination requests, and hence the workload. Also, the additional time would diminish the uncertainty about the value of inventions for which patent protection is sought, allowing for the correct decision on whether to request examination. As a consequence, the option to defer the examination request for one additional year increases the value of unexamined and granted applications.

The analysis is presented in four sections. In Section 2, we develop a structural model of deferred examination and patent renewals. Our data are described in Section 3, the estimation approach is presented in Section 4. We conduct two simulation experiments that allow us to identify the impact of changes in the deferment period on patent value and on patent office workload. Section 5 concludes with a summary and a discussion.

2 Structural Model

In this section we first describe the general setup of the model, explain the structure of the patent system, how patent applicants derive profits, and their information structure. Subsequently we describe their optimization problem and how it can be solved.

2.1 General Setup

Patent system We construct a model of patent examination and renewal in which applicants have the option to defer examination. In this section we describe the general setup of this model. Before an agent can get patent protection for his invention he first needs to file an official application at the patent office and pay the corresponding application fees C_{PO}^{Appl}. Modern patent systems require a patent to fulfill certain patentability criteria, such as novelty and inventiveness. The application is subject to a substantive examination before the patent is granted.[5] We assume that examination has to be requested by the applicant within L years from the application day.[6] This means that we allow the agent to defer the examination and the associated fees for examination C_{PO}^{Exam} for up to L years (maximum deferment term). However, deferment is not free of charge and the agent has to pay fees c_t^A ($t \in 1, ..., L$) to maintain the application pending for one more year. We assume that once examination had been requested it takes S years for the patent examiner to completely resolve the case and to provide the final decision on the patentability of the invention. If examination had been requested and the application has successfully passed the examination process the applicant can finally get the patent issued if he pays the final fee C_{PO}^{Grnt}. A patent gives the patentee the right to exclude others from using the patented invention. The patent right can be renewed for up to T years (maximum patent term) from the application date on as long as the patent owner pays the yearly renewal fees c_t^G ($t \in 1, ..., T$) for the granted patent. We assume that the maintenance fees for an application and a patent are the same $c_t^A = c_t^G = c_t$ ($t \in 1, ..., T$), and that they are non-decreasing in t.[7] If any of the fees are not paid to the patent office the application or patent expires irrevocably.

Returns The right to exclude others allows the patentee to generate non-negative returns r_t in every year the invention is protected by the patent. Since the exclusivity right is not enforceable before the patent is finally granted, we assume that the owner of a pending application is only able to realize a part $0 < q < 1$ of the returns of an already granted application, qr_t. The parameter q must be positive, since a pending application can already create value for its owner, e.g., by creating uncertainty for competitors or forming the basis for negotiations.[8]

The returns from patent protection evolve in the following way over time. The potential returns from patent protection in the first period, r_1, are drawn i.i.d. from a continuous distribution F_{IR} on a positive domain. In the next period the value from patent protection might increase or decrease depending on the information the owner obtains about his invention. The new information is represented by a growth rate $g_t \in [0, B]$ which is drawn from a distribution with the cumulative density function $F(u \mid t) = Pr\,[g_t \leq u \mid t]$. Thus, the returns in the second period are $r_2 = g_2 r_1$, and $r_t = g_t r_{t-1}$ in the following ones. Since the probability to learn how to increase the returns from patent protection should be higher for younger

[5] Registration systems without ex ante examination still exist in some countries, in particular for utility models.

[6] We assume that all decisions are made at the beginning of a year.

[7] This is exactly how the maintenance fees are structured in patent systems wich offer a deferment option. For the model to have a solution only the assumption of non-decreasing deferment fees and non-decreasing patent renewal fees is crucial. We will discuss the implications of different structures of maintenance fees in the conclusion.

[8] Patent owners are entitled to licensing fees from the day of publication. With the grant of the patent, they can also seek injunctions against potential infringers.

patents, we assume that the probability of having a high growth rate g_t decreases with a patent's maturity in the sense of first-order stochastic dominance ($F(u \mid t) \leq F(u \mid t+1)$).[9]

Before a patent is granted it has to pass an examination at the patent office. During this procedure the examiner has to verify whether the application fulfills the patentability criteria. He may require the applicant to change the patent specification, or he may even reject the application. This means that the distributions of growth rates of examined or granted patents might be different from the ones of pending applications. In the following let $g_t^A \sim F^A(u^A \mid t)$ denote the growth rate in case of a pending and unexamined patent application, and $g_t^G \sim F^G(u^G \mid t)$ in case of an examined or granted patent application. To account for cases when a patent application becomes absolutely worthless economically due to obsolescence, we assume that in every period with some probability $1-\theta$ all future returns can become a zero sequence. We allow the obsolescence rate to be the same for pending as well as granted applications. Therefore it represents the part of the uncertainty about the value of pending and granted patent applications which is not resolved even after a patent has been examined.

Agents We assume that every patent application belongs to exactly one profit maximizing agent. This means that in every period the agent always chooses the strategy with the highest expected payoff given his information structure.

At the beginning of a period the growth rate g_t^A, respectively g_t^G, is revealed to the agent, so that he knows the potential returns r_t from patent protection for this period. Furthermore, we assume that he also knows the distributions of all future growth rates, and thus is able to build expectations on how the returns will evolve in the future. Since the distributions of growth rates from patent protection are exposed to an unexpected shock during patent examination, the patent applicant has to readjust his expectations about future growth rates. We assume that this change in expectations is not anticipated by the applicant. Practically, this means that if the applicant has not yet requested application in period x, his growth rate is drawn from $F^A(u^A \mid x)$, and he expects the growth rates of patent returns to be distributed according to $F^A(u^A \mid t > x)$ in the future periods. Once he requests examination and receives a response on the patentability of the application from the patent office, he has to adjust his expectations on the evolution of returns from patent protection according to what was considered patentable by the examiner. Therefore the growth rates for subsequent periods are drawn from $F^G(u^G \mid t)$. Usually, the value of a patent application is very uncertain. It is not only uncertain whether the patented invention will have any commercial value but also whether the application can fulfill the patentability requirements. Whereas the economic

[9] Usually, the use of an invention should be determined early in a patent's life. The probability to discover new uses in later periods should accordingly be lower.

uncertainty may remain throughout the life of the patent, the latter, technical uncertainty can be resolved through examination. Hence, F^A and F^G should differ.

The fees which have to be paid to the patent office are only a part of the costs which are necessary to obtain a patent. Usually, an applicant has to invest resources in addition to the statutory fees. To properly model the choices of a representative agent during the life of a patent application we have to account for the cost of filing a patent application C_{self}^{Appl} (search, draft, translation) as well as the cost incurred during the examination proceeding C_{self}^{Exam} (negotiations with the examiner are usually conducted with the aid of a patent attorney).

2.2 Value Functions and the Maximization Problems

As described above, the life of a patent application comprises three parts:

- ▷ the application stage, in which the agent has to decide whether to apply for patent protection and, if he does, whether and when to request examination;
- ▷ the examination stage, in which the agent has to decide whether his application will be fully examined and granted, or withdrawn during the examination process;
- ▷ the patent stage, in which the agent has to decide whether to renew patent protection or to let it lapse before the expiration of its full term.

Since the model has a final horizon, the statutory patent term T, and returns are conditional only on returns in the previous age, we will see that the model can be solved recursively starting from the final age. Therefore, we continue in reverse chronological order by first analyzing the patent stage, then the examination stage, and lastly the application stage.

Patent stage If a patent is already granted at the beginning of period t the owner has to decide whether he wants to keep patent protection (K) until next period or to let it irrevocably expire (X). His choice will depend on his expected value from both strategies $\widetilde{V}^K(t, r_t)$ and $\widetilde{V}^X(t, r_t)$. The value of an expired patent is always zero, $\widetilde{V}^X(t, r_t) = 0$. The expected revenue from renewing a granted patent is the sum of current returns from patent protection r_t, less the maintenance fees c_t plus the option value of being able to renew patent protection in the next period $E\left[\widetilde{V_K}(t+1, r_{t+1}) \mid r_t\right]$.[10] With β as the discount factor between periods the value function is:

$$\widetilde{V}^K(t, r_t) = r_t - c_t + \beta\theta E\left[\widetilde{V_K}(t+1, r_{t+1}) \mid r_t\right] \tag{1}$$

[10] $\widetilde{V}^K(t, r_t)$ denotes the value function if strategy K is chosen in year t. In contrast, $\widetilde{V_K}(t, r_t)$ denotes the value function if strategy K was chosen in the previous year $t-1$ and the value maximizing strategy is chosen subsequently in year t. The optimal subsequent strategy doesn't have to be K.

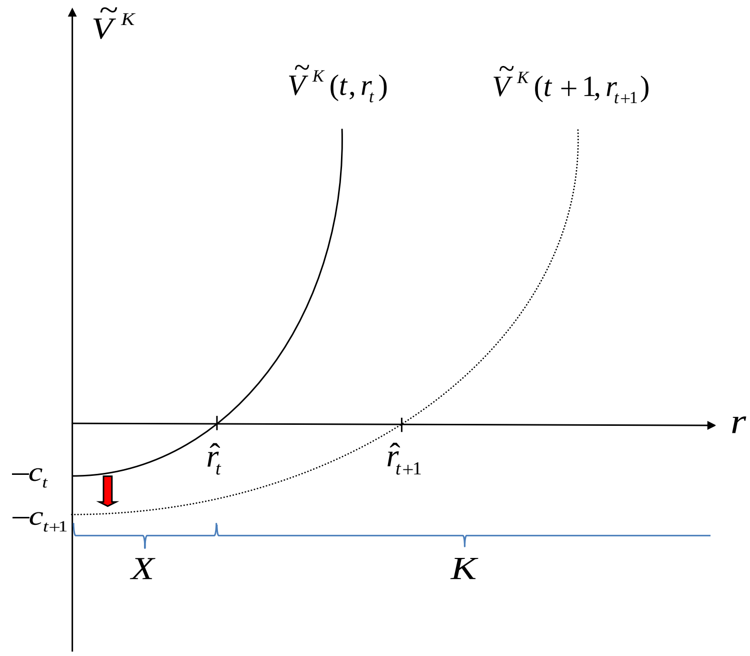

Figure 1: Value Functions and Cut-off Values - Patent Stage

with

$$\widetilde{V_K}(t+1, r_{t+1}) = max\left[\widetilde{V}^K(t+1, r_{t+1}), \widetilde{V}^X(t, r_t)\right]$$

and

$$E\left[\widetilde{V_K}(t+1, r_{t+1}) \mid r_t\right] = \int \widetilde{V_K}(t+1, u^G r_t) dF^G(u^G \mid t)$$

Since the agent's choice in every period is discrete, there exists a threshold return $\hat{r}_t$ for each period t that determines the patent owner's optimal decision (see Figure 1):

- ▷ $\left\{\hat{r}_t^K\right\}_{t=S+1}^T$: minimum patent returns needed for an agent to decide to keep (K) patent protection in period t and not to let it expire (X). This is the solution to $\widetilde{V}^K(t, r_t) = \widetilde{V}^X(t, r_t) = 0$.[11]

In period $t = T$ the option value is zero since the patent cannot be renewed anymore. Thus, the cut-off value in the last period is $\hat{r}_T^K = c_T$.

[11]The proof that $\widetilde{V}^K(t, r_t)$ is continuous and increasing in r_t, and decreasing in t can be found in Pakes (1986). These properties ensure that the sequence $\left\{\hat{r}_t^K\right\}_{t=1}^T$ exists and is increasing in t.

Examination stage We consider two alternative approaches of modeling the examination stage. Assume that a complete examination of a patent application takes S years. During this time period the examiner searches for prior art and studies the claims in the patent application. He either approves but more often objects to some or all claims. The examiner's objection will be outlined in a report or letter called a patent office action. The applicant has to respond to the examiner's objections and requirements whereupon the examiner further reconsiders and either approves or calls for further amendments. Only if the applicant has met all requirements and overcome all objections raised by the examiner the patent application will be allowed. Once the application has been allowed the applicant usually has to pay an additional granting fee for the patent to issue.

One way to model the examination stage is to look at it as a process where the applicant has the choice at the beginning of each period to continue the examination (CE) and incur the respective costs, or to withdraw his application (W) during an ongoing examination (Alternative I). Moreover, if the application has finally been approved as patentable, the applicant has to confirm the grant (G) by paying the granting fees C_{PO}^{Grnt} or he can still let it expire. As already explained above, during the examination process the expectation of how the future returns from patent protection evolve might change. We assume that the applicants adjust their distributions of future growth rates right upon the receipt of the first substantive action from the examiner.[12] This action provides new information on what is actually allowed to be granted from the examiner's perspective and is issued s ($s < S$) periods after the examination has been requested. The first action is followed by a (costly) dispute between the applicant (or representative patent attorney) and the examiner for $S-s$ remaining periods. We assume that these costs C_{self}^{Exam} are incurred in equal parts during these periods.

Assume that examination was requested in period $t = a$. We consider the grant decision in period $a+S$ first. If the applicant withdraws the examined application, then $\widetilde{V}^W(a+S, r_{a+S}) = 0$. If instead he wants the patent to be granted he will have to incur costs $c_{a+S} + C_{PO}^{Grnt}$:

$$\widetilde{V}^G(a+S, r_{a+S}) = r_{a+S} - (c_{a+S} + C_{PO}^{Grnt}) + \beta\theta E\left[\widetilde{V_K}(a+S+1, r_{a+S+1}) \mid r_{a+S}\right] \quad (2)$$

[12]Clearly, if the patent examination procedure includes several substantive actions they all may lead to an adjustment to the expectations about future returns from patent protection. Nevertheless, we do not incorporate further adjustments into the model for two reasons. First, because we do not observe whether further substantive actions have been issued, neither when they have been issued, nor their content. Second, we think it is plausible to assume that the most relevant and serious objections are outlined in the first substantive action.

with

$$E\left[\widetilde{V_K}(a+S+1, r_{a+S+1}) \mid r_{a+S}\right] = \int \widetilde{V_K}(a+S+1, u^G r_{a+S}) dF^G(u^G \mid a+S)$$

▷ $\left\{\hat{r}^G_{a+S}\right\}^L_{a=1}$: minimum patent returns needed for the agent to allow the examined application to be granted at age $t = a+S$. This is the solution to $\widetilde{V}^G(a+S, r_{a+S}) = \widetilde{V}^W(a+S, r_{a+S}) = 0$.[13]

Consider the time periods after the first substantive action has been issued, $t = a+s, ..., a+S-1$. In these periods, the patentee knows what can actually be protected by the patent. These are also the periods when the correspondence with the examiner occurs and C^{Exam}_{self} has to be paid. The applicant's options are either to withdraw the application, $\widetilde{V}^W(t, r_t) = 0$, or to continue the correspondence with the examiner:

$$\widetilde{V}^{CE}(t, r_t) = qr_t - [c_t + \frac{C^{Exam}_{self}}{S-(s+1)}] + \beta\theta E\left[\widetilde{V_{CE}}(t+1, r_{t+1}) \mid r_t\right] \tag{3}$$

with

$$\widetilde{V_{CE}}(t+1, r_{t+1}) = \begin{cases} max\left[\widetilde{V}^G(t+1, r_{t+1}), \widetilde{V}^W(t+1, r_{t+1})\right] & \text{if } t = a+S-1 \\ max\left[\widetilde{V}^{CE}(t+1, r_{t+1}), \widetilde{V}^W(t+1, r_{t+1})\right] & \text{if } t = a+s, ..., a+S-2 \end{cases}$$

and

$$E\left[\widetilde{V_{CE}}(t+1, r_{t+1}) \mid r_t\right] = \int \widetilde{V_{CE}}(t+1, u^G r_t) dF^G(u^G \mid t)$$

▷ $\left\{\hat{r}^{CE}_t\right\}^{a+S-1}_{t=a+s}$: minimum patent returns needed for the agent to continue the examination process at age $t = a+s, ..., a+S-1$. This is the solution to $\widetilde{V}^{CE}(t, r_t) = \widetilde{V}^W(t, r_t) = 0$.[14]

The remaining periods in the examination stage are the ones right after the examination request and before the first substantive action is issued, $t = a+1, ..., a+s-1$. Here, the applicant hasn't yet learned the examiner's objections and assumes that the future growth rates are drawn from $F^A(u^A \mid t)$:

$$\widetilde{V}^{CE}(t, r_t) = qr_t - c_t + \beta\theta E\left[\widetilde{V_{CE}}(t+1, r_{t+1}) \mid r_t\right] \tag{4}$$

[13] Similar to Pakes (1986) one can show that $\widetilde{V}^G(a+S, r_{a+S})$ is continuous and increasing in r_{a+S}, and decreasing in a. Therefore, the sequence $\left\{\hat{r}^G_{a+S}\right\}^L_{a=1}$ exists and $\hat{r}^G_{a+S}$ is increasing in a.

[14] Similar to Pakes (1986) one can show that $\widetilde{V}^{CE}(t, r_t)$ is continuous and increasing in r_t, and decreasing in t as well as a. Therefore, $\left\{\hat{r}^{CE}_t\right\}^{a+S-1}_{t=a+S}$ must exist and $\hat{r}^{CE}_t$ is increasing in a as well as t.

with

$$\widetilde{V_{CE}}(t+1, r_{t+1}) = max\left[\widetilde{V}^{CE}(t+1, r_{t+1}), \widetilde{V}^{X}(t+1, r_{t+1})\right]$$

and

$$E\left[\widetilde{V_{CE}}(t+1, r_{t+1}) \mid r_t\right] = \int \widetilde{V_{CE}}(t+1, u^A r_t) dF^A(u^A \mid t)$$

▷ $\left\{\hat{r}_t^{CE}\right\}_{t=a+1}^{a+s-1}$: minimum patent returns needed for the agent to continue the examination process at age $t = a+1, ..., a+s-1$. This is the solution to $\widetilde{V}^{CE}(t, r_t) = \widetilde{V}^{W}(t, r_t) = 0$.[15]

Therefore, the expected revenue from requesting examination in period $t = a$, $\widetilde{V}^E(a, r_a)$, is:

$$\widetilde{V}^E(a, r_a) = \begin{cases} qr_a - (c_a + C_{PO}^{Exam} + C_{PO}^{Appl} + C_{self}^{Appl}) + \\ +\beta\theta E\left[\widetilde{V_{CE}}(a+1, r_{a+1}) \mid r_a\right] & \text{if } a = 1 \\ qr_a - (c_a + C_{PO}^{Exam}) + \\ +\beta\theta E\left[\widetilde{V_{CE}}(a+1, r_{a+1}) \mid r_a\right] & \text{if } a = 2, ..., L+1 \end{cases} \quad (5)$$

The traditional way of modeling the examination stage (Deng 2007; Serrano 2011) is to assume that a patent examination takes S years and at the end of these periods the application will be approved for grant with probability π_{Grnt} or rejected with probability $1 - \pi_{Grnt}$. This means that once the applicant requests examination he has to continue the examination process until the final decision of the examiner on the patentability, and if the examination was successful, he always wants his patent to be granted. The agent might only withdraw his application during the examination if the invention becomes obsolete, or its protection commercially worthless. According to this view the expected value of requesting examination in year $t = a$, $\widetilde{V}^E(a, r_a)$ comprises the expected returns from having a pending application minus all expected examination costs $K = C_{PO}^{Exam} + C_{self}^{Exam} + C_{PO}^{Grnt}$ and maintenance fees, plus the expected returns from a pending application and the expected returns from full patent protection in the future:

[15]Similar to Pakes (1986) one can show that $\widetilde{V}^{CE}(t, r_t)$ is continuous and increasing in r_t, and decreasing in t as well as a. Therefore, $\left\{\hat{r}_t^{CE}\right\}_{t=a+S}^{a+S-1}$ must exist and $\hat{r}_t^{CE}$ is increasing in a as well as t.

$$\widetilde{V}^E(a,r_a) = \begin{cases} qr_a + \\ +\beta\theta qE(r_{a+1}|r_a) + ... + (\beta\theta)^{S-1} qE(r_{a+S-1}|r_a) - \\ -(c_a + C_{PO}^{Exam} + C_{PO}^{Appl} + C_{self}^{Appl}) - \\ -\beta\theta c_{a+1} - ... - (\beta\theta)^{S-1}(c_{a+S-1} + \frac{C_{self}^{Exam}}{S-(s+1)}) + \\ +(\beta\theta)^S \pi_{Grnt} \left(E\left[\widetilde{V_K}(a+S, r_{a+S}) \mid r_a\right] - C_{PO}^{Grnt}\right) & \text{if } a = 1 \\ \\ qr_a + \\ +\beta\theta qE(r_{a+1}|r_a) + ... + (\beta\theta)^{S-1} qE(r_{a+S-1}|r_a) - \\ -(c_a + C_{PO}^{Exam}) - \\ -\beta\theta c_{a+1} - ... - (\beta\theta)^{S-1}(c_{a+S-1} + C_{self}^{Exam}) + \\ +(\beta\theta)^S \pi_{Grnt} \left(E\left[\widetilde{V_K}(a+S, r_{a+S}) \mid r_a\right] - C_{PO}^{Grnt}\right) & \text{if } a = 2, ..., L+1 \end{cases} \quad (6)$$

with

$$E(r_{a+S-1}|r_a) = \int \cdots \int (u_{a+1}^A \cdot ... \cdot u_{a+S-1}^A \cdot r_a) dF^A(u_{a+1}^A \mid a) ... dF^A(u_{a+S-1}^A \mid a+S-2)$$

and

$$E\left[\widetilde{V_K}(a+S, r_{a+S}) \mid r_a\right] =$$
$$= \int \cdots \int \widetilde{V_K}(a+S, u_{a+1}^A \cdot ... \cdot u_{a+S}^A \cdot r_a) dF^A(u_{a+1}^A \mid a) ... dF^A(u_{a+S}^A \mid a+S-1)$$

Application stage During the application stage the potential applicant has to decide first whether he wants to file a patent application and, once he has decided to file an application, whether and when to request examination. The decision to request examination can be deferred for at most L periods. This means that an agent who still holds a pending application in the beginning of period $t = L+1$ has to decide whether he finally wants to request examination (E) or to withdraw (W) it completely. Given the expected revenues from requesting examination $\widetilde{V}^E(a, r_a)$ from equations (5) or (6) and with $\widetilde{V}^W(t, \hat{r}_t) = 0$ we define:

- ▷ $\hat{r}_{L+1}^E$: minimum patent returns needed for the agent to request an examination (E) and not to withdraw (W) the application in period $t = L+1$. This is the solution to $\widetilde{V}^E(L+1, r_{L+1}) = \widetilde{V}^W(L+1, r_{L+1}) = 0$.[16]

In earlier periods, $t = 1, ..., L$, the owner of a pending application has three options. Besides the possibilities to withdraw (W) the application and to request examination (E) he can also

[16] One can easily show that $\widetilde{V}^E(L+1, r_{L+1})$ is continuous and increasing in r_{L+1}, such that $\hat{r}_{L+1}^E$ exists.

choose to defer (D) the decision to the next period. The expected value of the third option $\widetilde{V}^D(t, r_t)$ consists of the returns from having a pending application in this period, minus the deferment fees and the expected returns from the option of having the same choices in the next period:

$$\widetilde{V}^D(t, r_t) = \begin{cases} qr_t - (c_t + C_{PO}^{Appl} + C_{self}^{Appl}) + \\ +\beta\theta E\left[\widetilde{V_D}(t+1, r_{t+1}) \mid r_t\right] & \text{if } t = 1 \\ qr_t - c_t + \beta\theta E\left[\widetilde{V_D}(t+1, r_{t+1}) \mid r_t\right] & \text{if } t = 2, ..., L \end{cases} \tag{7}$$

with

$$\widetilde{V_D}(t+1, r_{t+1}) = \begin{cases} max\left[\widetilde{V}^E(t+1, r_{t+1}), \widetilde{V}^W(t+1, r_{t+1})\right] & \text{if } t = L \\ max\left[\widetilde{V}^E(t+1, r_{t+1}), \widetilde{V}^D(t+1, r_{t+1}), \widetilde{V}^W(t+1, r_{t+1})\right] & \text{if } t = 1, ..., L-1 \end{cases}$$

and

$$E\left[\widetilde{V_D}(t+1, r_{t+1}) \mid r_t\right] = \int \widetilde{V_D}(t+1, u^A r_t \mid r_t) dF^A(u^A \mid t)$$

Since at this stage the applicant has three options, in every period $t = 1, ..., L$ there exist two threshold values that determine the optimal choices (see Figure 2):

- ▷ $\left\{\hat{r}_t^D\right\}_{t=1}^L$: minimum patent returns needed for the agent to defer the decision (D) at age t and not let it expire (W). This is the solution to $\widetilde{V}^D(t, r_t) = \widetilde{V}^W(t, r_t) = 0$.[17]
- ▷ $\left\{\hat{r}_t^E\right\}_{t=1}^L$: minimum patent returns needed for the agent to request an examination (E) instead of deferring the decision (D) at age t. This is the solution to $\widetilde{V}^E(t, r_t) = \widetilde{V}^D(t, r_t)$.[18]

According to the study by Henkel and Jell (2010), the two main motives behind the decision to defer examination are to create uncertainty for competitors and to gain time for evaluation of the commercial value. Both motives are incorporated in our model. The value of creating uncertainty in the marketplace is incorporated in the returns from patent protection that can be already realized through a pending application, qr_t. Given that the potential returns from patent protection are not high enough to request examination the

[17]Similar to Pakes (1986) one can show that $\widetilde{V}^D(t, r_t)$ is continuous and increasing in r_t so that $\left\{\hat{r}_t^D\right\}_{t=1}^L$ must exist. For $t = 2, ..., L$, $\widetilde{V}^D(t, r_t)$ is decreasing in t. Therefore, the sequence of cut-off values $\left\{\hat{r}_t^D\right\}_{t=2}^L$ is increasing in t. In $t = 1$, $\hat{r}_1^D$ might be higher than in the subsequent periods, since the applicant has to incur additional costs for the application $(C_{PO}^{Appl} + C_{self}^{Appl})$.

[18]Since $\widetilde{V}^E(t, r_t)$ and $\widetilde{V}^D(t, r_t)$ are continuous in r_t, so must be $\widetilde{V}^E(t, r_t) - \widetilde{V}^D(t, r_t)$. The proof that $\widetilde{V}^E(t, r_t) - \widetilde{V}^D(t, r_t)$ is increasing in r_t can be found in Appendix A.1. Thus, the sequence $\left\{\hat{r}_t^E\right\}_{t=1}^L$ must exist.

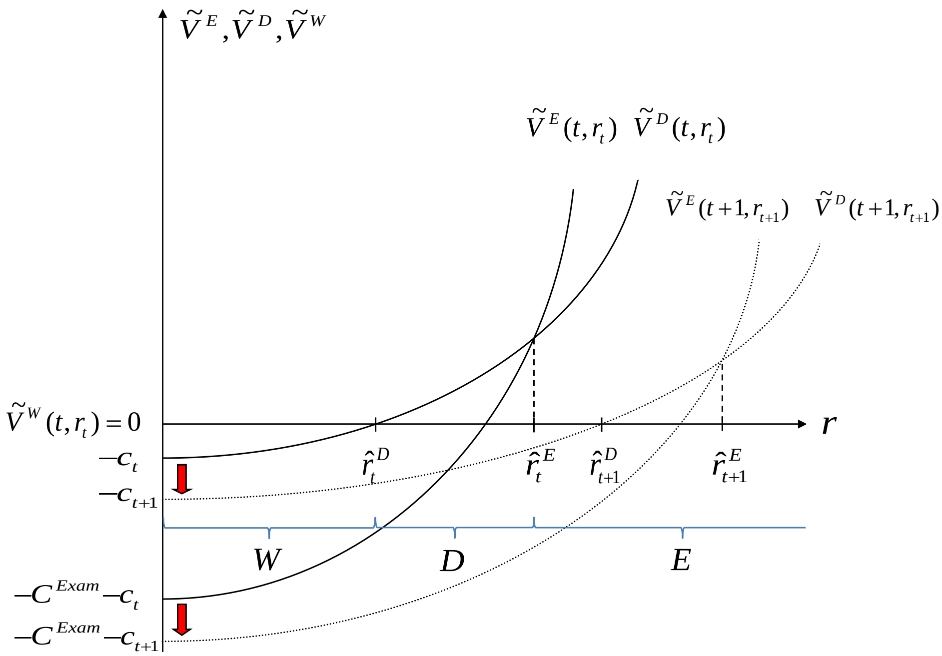

Figure 2: Value Functions and Cut-off Values - Application Stage prior to Examination Request

applicant will choose to defer examination either if qr_t is high or the option value of future returns is high enough ($\triangleq$ gain time for evaluation).

Since the problem described above is finite and returns in one period depend only on returns realized in the previous period, the model can be solved for the sequences of the cut-off values by backward recursion.[19]

3 Data

For our estimations we are using data on Canadian patent applications.[20] In particular, we analyze 211,550 patent applications filed between October 1, 1989 and September 30, 1996 with information from years 1989-2008. Patent protection could be renewed for up to 20 years from the application, $T = 20$. On October 1, 1989 renewal fees had been introduced for the first time, and until September 30, 1996 the examination request had to be made within 7 years from the filing date of the application, $L = 7$. Since Canada is a PCT (Patent Cooperation Treaty) member, applications which had gone through the PCT route only entered the national stage at the CIPO (Canadian Intellectual Property Office) 30 months after their priority date (which is typically 18 months from the application at the Canadian patent office). In turn, information on applications which had directly been submitted at the CIPO is available from the first day on. Furthermore, a different fee schedule applies for international applications to get examined. Therefore, we exclude all PCT applications from the data and use only 137,397 CIPO patent applications. Besides the date of the application, for each application we observe the date when examination was requested or the application withdrawn. In case examination had been requested, the data include information about the grant date or the date of withdrawal during examination. For granted patents we also observe when the patent owner stopped paying the renewal fees and the patent lapsed.

As one can see in Figure 3, on average, about 30,000 applications were submitted each year at the Canadian IP Office.[21] In 1990 almost 69% of all Canadian patent applications took the national application route through the CIPO, of which 74% have requested examination within the deferment period. In 1995 the portion of applications taking the national route decreased to 38% with almost 90% requesting examination. The average grant rate, defined as the percentage of applications that have successfully gone through a patent examination

[19]See Appendix A.2 for a sketch how the model is solved recursively.

[20]In a companion paper Harhoff (2012) studies the policy reforms at the Canadian Intellectual Property Office (CIPO) in more detail. Canada switched in 1989 from a US-style system with publication at grant to a seven-year deferment system with publication of the unexamined application after 18 months. In 1996, CIPO reduced the deferment period to five years.

[21]For cohorts 1989 and 1996 only patents filed between October and December 1989, respectively January and September 1996, were affected by the change to the patent system.

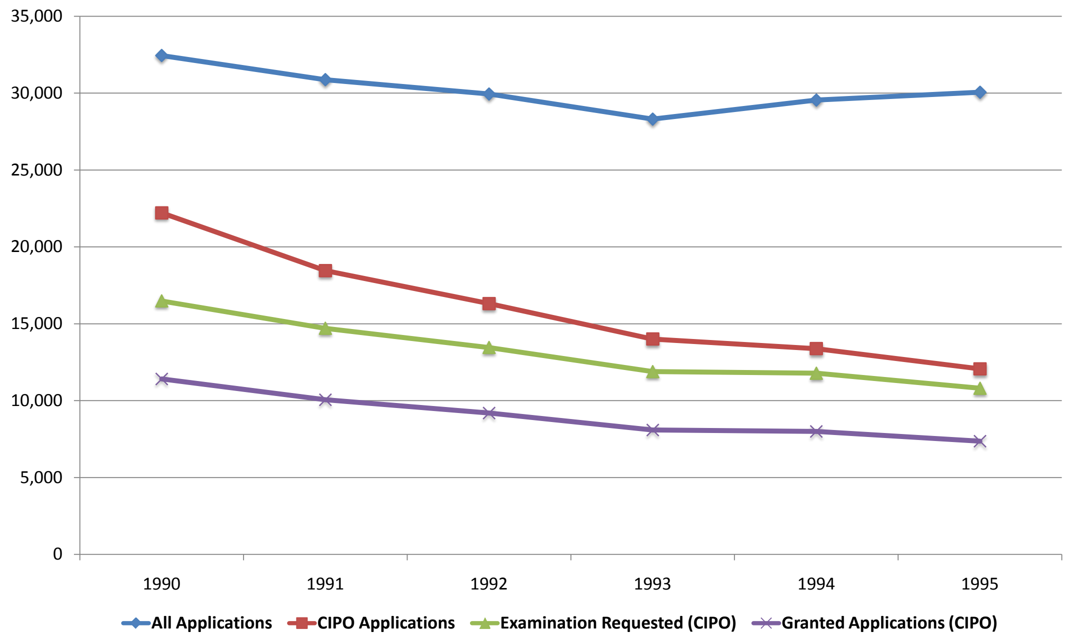

Figure 3: Canadian Patent Applications by Filing Year (1989-1995)

out of those that had actually requested examination was 68.71%. [22] According to the CIPO Annual Reports, 80% percent of applications with a request for examination were waiting less than 2 years for a first substantive examination action (including all known objections to patentability). On average, a patent was granted about 4 years after examination had been requested. Hence, we set $s = 2$ and $S = 4$.

Cost structure The maintenance fees at the CIPO for pending applications, as well as patents, were zero in the first two years, 100 CAD$ for years 3-5, 150 CAD$ for years 6-10, 200 CAD$ for years 11-15, and 400 CAD$ for years 16-20.[23] There was one change in the nominal fee schedule which was applied to renewals starting from January 1, 2004. The

[22]There was some variation within cohorts. The grant rate lied slightly above 70% and remained rather stable for applications for which examination has been requested within the deferment period. Only when examination had been requested at the end of the deferment period the grant rate dropped to 64%. Nevertheless we maintain the assumption that there is no selection into grant rates throughout the paper in order to not overcomplecate the model. Grant rates may also vary across technologies and applicant types (see Schankerman 1998). However, in this paper we aggregate over all non-PCT applications and maintain a common grant rate.

[23]Indeed, the fee structure is different for small and large applicants. By CIPO's definition, a small applicant is an entity that employs 50 or fewer employees or that is a university. Small applicants are offered a 50% reduction on application and maintenance fees. Unfortunately, we are not able to distinguish between small and large entities. Nevertheless, small applications consist of less than 15% of total applications.

maintenance fees were increased by 50 CAD$ for the years 6-20. To ease the computational burden, we have used the weighted average of the maintenance fees before and after the change in the fee structure for estimation. The fee for filing an application amounted to 200 CAD$ and 400 CAD$ for requesting examination for the cohorts under consideration. A final fee of 300 CAD$ was due for the publication of the grant.[24]

As already mentioned above, it would be incorrect to assume that the decisions made during the application and examination stage depend solely on the statutory patent fees. To have a rough estimate we use the information on the costs of filing a patent application in Canada provided by Canadian law firms.[25] According to this information the costs of examination range from 750 CAD$ to 7,500 CAD$ depending on the complexity and the number of arguments put forward by the examiner. The filing costs may have even a higher variation depending on its length, whether it requires translation from other languages, and whether the applicant does a search to find out whether anyone else has already thought of the idea to be patented. We decided to set C_{self}^{Exam} to 3,000 CAD$ such that $C_{PO}^{Exam} + C_{self}^{Exam} + C_{PO}^{Grnt} = 250 + 3000 + 450 = 3700$ CAD$.

4 Estimation

4.1 Estimation Strategy

We use a simulated minimum distance estimator developed by McFadden (1989) and Pakes and Polland (1989) for the estimation.[26] In the first step we assign a stochastic specification to our structural model by making functional form assumptions which in turn will depend on a vector of parameters ω. In order to determine the vector ω_0 of the true parameters we fit the hazard probabilities derived from the theoretical model to the true hazard proportions as proposed in Lanjouw (1998). Each parameter has a different effect on the structure of the sequences of the cut-off values derived from the model, $\left\{\hat{r}_t^j\right\}$ with $j = E, D, CE, G, K$ and the distribution of returns in each age, r_t, which in turn determine the hazard probabilities. This allows the identification of the model parameters. Although in theory a solution to the structural model, i.e., the sequences $\left\{\hat{r}_t^j\right\}$ can be found analytically, this is hardly possible in practice due to the complexity of the model. Thus, we use a weighted simulated minimum

[24] To make the cost structure more realistic we have added 50 CAD$ to each payment due to the patent office. Usually patent attorneys charge their clients for these money transfer services or the applicant has at least to invest time for the completion of the respective forms. Since these costs can vary a lot we regard 50 CAD$ as a lower bound.

[25] See for example http://www.valuetechconsulting.com/cost.php, last accessed December 2012.

[26] McFadden (1989) and Pakes and Polland (1989) provide conditions required to ensure the consistency and asymptotic normality of the estimator. Pakes (1986) and Lanjouw (1998) show that the required conditions hold for our type of model.

distance estimator (SGMM) $\hat{\omega}_N$. The estimator is the argument that minimizes the norm of the distance between the vector of true and simulated hazard proportions. We use a weighting matrix $A(\omega)$ to improve the efficiency of the estimator:

$$A(\omega)\left\| h_N - \eta_N(\omega) \right\| \text{ with } \hat{\omega}_N^* = arg\,\underset{\omega}{min}\, A(\omega)\left\| h_N - \eta_N(\omega) \right\| \tag{8}$$

▷ h_N is the vector of sample or true hazard proportions.

▷ $\eta_N(\omega)$ is the vector of simulated hazard proportions (predicted by the model).

▷ $A(\omega) = diag\sqrt{\underline{n}/N}$ is the weighting matrix. $\underline{n}$ is the vector of the number of patents in the sample for the relevant age-cohort. N is the sample size.

In order to calculate the simulated hazard rates for a parameter set ω we first have to calculate the sequences of the cut-off values $\left\{\hat{r}_t^j\right\}$ with $j = E, D, CE, G, K$. To do so we proceed recursively by first determining the value functions in the last period and calculating the corresponding cut-off values. Subsequently, with these cut-off values, we calculate the value functions in the second last period and proceed recursively in the same manner until the first period. Once we have calculated the cut-off functions for all periods we perform five simulations. In each simulation we take $3 \cdot N$ pseudo random draws from the initial distribution and exactly the same amount of draws from each distribution of the growth rates g_t^A and g_t^G. Afterwards we pass the initial draws through the stochastic process, compare them with the corresponding cut-off values, and calculate the hazard proportions for all years. The vector of hazard rates from each simulation is then averaged over the five simulation draws and inserted in the objective function (8). The objective function is then minimized using a two step approach. We use global optimization algorithms in the first step and a Nelder-Mead-type local optimization search algorithm to find the local minimum in the second step.[27]

We will fit three types of hazard proportions: (1) HR_E, the percentage of applications for which examination was requested, (2) HR_D, the percentage of applications which were deferred to the next period in a given year out of those that had been deferred in the previous period, and (3) HR_X, the hazard proportion of expired patents. There are two possible ways to calculate HR_X depending on the way we model the examination stage. According to the traditional view (Version 1 assuming $\pi_{Grnt} < 1$), HR_X^1 is the percentage of granted patents that expire in a given year out of those granted and renewed in the previous period. But if we explicitly model the examination stage, then HR_X^2 (Version 2 with $\pi_{Grnt} = 1$) is the percentage of all granted patents and applications under examination that expire in a given

[27]MATLAB (matrix laboratory) is a numerical computing environment developed by MathWorks. Since the objective function is supposed to be non-smooth we apply the Simulated Annealing algorithm and the Genetic algorithm in the first step. Both are probabilistic search algorithms (see description of the Global Optimization Toolbox for MATLAB). The Nelder-Mead-type search algorithm implemented in MATLAB is called *fminsearch*.

year out of those applications that have already requested examination and patents which are still alive.

We decided to use only the traditional way of modeling the examination stage for the estimation. The reason is that estimation of both alternatives requires us to assume the same duration of patent examination S for all applications. If we the way of modeling the examination stage where we do not distinguish between patents under examination and already granted patents (Alternative I) we will get biased results. In reality examination patents can be examined within 2 years or examination can even take more than 10 years meaning that S is rather heterogenous. Assuming a constant duration of examination of four years for all patents thus leads to simulated hazard rates $HR_X^2(t)$ whose composition of patents still under examination and already granted patents would mismatch the composition of the real hazard rates. Assume that examination was requested in the third period and the patent has already been granted after 2 years. This will reduce the hazard rate in the fifth year $HR_X^2(5)$. Since by assumption such a low duration of examination is not possible, the model trying to fit this hazard rate will be adjusted by allowing more applications which requested examination in the first year to be granted or more patents to be renewed. A similar reasoning applies to patents which were examined longer than 4 years and not granted.[28] This kind of bias is avoided if we use the traditional way of modeling the examination stage (Alternative II), since the hazard rates which we use for the estimation only include patents which are already granted $HR_X = HR_X^1$. They do not depend on the duration of the patent examination.

Since for the applications in our data the maximum deferment period was 7 years, we calculate HR_D for 7 periods and HR_E for 8 periods for each of the seven cohorts. The decision to request examination can be made anytime within the 7 years period. Therefore we assign all requests which were made within the first 6 months past the filing date of the application to the first period and all requests which were made in the following 12 months to the second period, and so forth.[29] The maximum patent term in Canada was 20 years but since we only observe events before the end of 2008 the vector $HR_X^1(t)$ consists of 15 entries for cohort 1989 and 8 for cohort 1996 (beginning with period 5).

Furthermore, we do not consider the application decision for our final estimation. The reason is that the estimation results, especially the parameters of the initial distribution, will highly depend on the costs of filing an application. Since we do not observe these costs

[28]To avoid this kind of bias one could restrict the sample to applications which had never requested examination and applications which had requested examination but were either granted only after 4 years or were dropped less than 4 years after the request. But we refrained from sub-sampling the data in this way, since this approach could introduce an even stronger bias and considerably restrict the validity of our results.

[29]A few recording dates for the examination request exceeded 7.5 years. We assigned these decisions to the 8th period.

and they tend to vary considerably across patents, incorporating this decision might bias the estimation results.[30]

4.2 Stochastic Specification

Initial returns As in previous patent renewal studies[31] we assume that the initial returns r_1 of all applications are lognormally distributed where $F_{IR}(\mu_{IR}, \sigma_{IR})$ is a normal distribution with mean μ_{IR} and variance σ_{IR}:[32]

$$log(r_1) \sim Normal(\mu_{IR}, \sigma_{IR}) \tag{9}$$

Distributions of the growth rates For the growth rates during the years before the patent has been examined we follow a similar stochastic specification as in Pakes (1986). We assume that the realized growth of returns is the maximum between the minimum growth rate δ^A, and a growth rate which is drawn from an exponential distribution v^A with variance $\sigma_t^i = (\phi^i)^{t-1}\sigma_0^i$: $g_t^A = max(\delta^A, v^A)$. The second growth rate, v^A , represents the cases when the applicant is able to "learn" about how to increase the returns above the minimum growth rates. We also assume that $\phi^A < 1$ such that the probability of getting higher returns will decrease with age t. The overall distribution of g_t^A can be represented as follows:

$$g_t^A \sim F^A(u^A \mid t) = \begin{cases} 1-\theta & \text{if } 0 \leq u^A < \delta^A \\ 1-\theta+\theta(1-exp(-\frac{u^A}{\sigma_t^A})) & \text{if } \delta^A \leq u^A \end{cases} \tag{10}$$

We model the evolution of the growth rates during the patent stage, $F^G(u^G \mid t)$, in a more static way. We assume that learning possibilities disappear once the patent has been examined.[33] This means that uncertainty about future returns from patent protection

[30] To our knowledge Deng (2011) is the only one who has estimated a dynamic stochastic patent renewal model incorporating the application decision for European patent filings. She has only considered the statutory application and granting fees at the European Patent Office (EPO) for estimation and disregarded the costs of drafting and translating EPO patent applications. These costs usually exceed the statutory fees and vary considerably across technology areas as well as applicant types.

[31] See for example Pakes (1986).

[32] Lanjouw (1998) was the only one who deviated from this assumption. She assumed that the initial returns of patents applications is zero and its value only evolves over time.

[33] We have also estimated a competing model where we explicitly allowed for learning opportunities during the patent stage. This dynamic model provided a somewhat better fit to the data since the evolution of returns during the patent stage was now determined by three parameters $\phi^G, \sigma_0^G, \delta^G$ instead of a single one δ^G. The increased fit could be fully attributed to adjustments in the simulated $HR_X(t)$. All other parameters remained in a narrow range of the presented model. Apart from the value distributions, which have become more skewed due to the additional learning opportunities, all results presented below, in particular the qualitative ones, continue to hold. The reason why we have chosen the more static model for the following analysis is that the identification of $\phi^G, \sigma_0^G, \delta^G$ relies solely on the variation in $HR_X(t)$ and the variation in the patent renewal fees. We cannot fully exclude offsetting effects between these three parameters.

completely disappears after the grant. The evolution of returns is then fully deterministic and they depreciate at a constant rate δ^G.

To ease the computational burden we fix the discount rate $\beta = 0.95$. Furthermore, since the first weighted hazard rate of expiration (HR_X for period 5) is 0.0465, we set $\theta = 0.9535$. Year 5 is when the first patent applications are granted. Therefore, we find it plausible to assume that patents from a cohort which are granted first expire in the first year after grant because of obsolescence and not because of too high renewal fees. The maintenance fees for the fifth year amount to only 100 CAD$. Since the fraction of applications which were granted out of those that had requested examination was 68.71%, the probability that an application which has not become obsolete during the examination process will be successfully granted is: $\pi_{Grnt} = \frac{0.6871}{\theta^{S-1}} = \frac{0.6871}{0.9535^3} = 79.26\%$.

With q as the fraction of returns an application can generate already before grant we are left with seven structural parameters to be estimated:

$$q,\ \mu_{IR}, \sigma_{IR},\ \phi^A, \sigma_0^A, \delta^A,\ \delta^G$$

These parameters altogether determine the structure of the sequences of the cut-off values derived above, $\hat{r}_t^j$ with $j = E, D, K$, and the distribution of returns in each period, r_t.

4.3 Identification

Like in other patent renewal models the parameters are identified by the cost structure and the non-linearity of the model. Different parameter values imply different cut-off value functions which in turn imply different hazard rates.

In particular, the parameters μ_{IR} and σ_{IR} determine the mean and variance of the initial distribution of returns. Both have an effect on all three sequences of hazard rates. Variation in σ_{IR} results in changes in $HR_E(t)$ in the first and last year, and changes in $HR_D(t)$ in the first and third year, but leaves the hazard rates in the other years rather unchanged. In contrast, variation in μ_{IR} changes $HR_E(t)$ and $HR_D(t)$ in all years. Interestingly, whereas higher values of μ_{IR} result in a higher hazard rate of deferment in the third period $HR_D(3)$, the period when the first maintenance fees are due, higher variance σ_{IR} has the opposite effect.

The parameter q, which represents the fraction of the returns from patent protection which can be realized with an unexamined patent application, is mainly identified by the variation in $HR_D(t)$, and especially in $HR_E(t)$. A higher q raises the hazard rates of deferment for all years almost constantly whereas it increases the hazard proportion of requesting examination only in the last, eighth year and decreases them for years 1 to 7. A lower q would have the opposite effect.

The distributions of growth rates of returns from pending applications $F^A(u^A \mid t)$ are fully determined by ϕ^A, σ_0^A, and δ^A. The parameters ϕ^A and σ_0^A have similar impact on all three hazard rates. Higher values of parameters go along with higher hazards of examination request in all years and lower hazard of deferment except for the third year, $HR_D(3)$, where it leads to an increase. Different values of ϕ^A and σ_0^A have also an impact on the curve of the hazards of expiration. Higher values produce a concave curve such that the hazard proportions decline or remain constant for older patents. Lower values produce a convex curve with increasing hazard proportions for older patents. This is the consequence of constant maintenance fees for the years 16-20. Nevertheless, there is a difference between variation in ϕ^A and variation in σ_0^A. Higher values of the first parameter imply increasing hazard proportions of examination request for the years 2-7 whereas higher values of the latter imply decreasing hazards for the same years and vice versa. δ^A, together with the other two parameters determine from what year on the hazard proportions of expiration, $HR_X(t)$, start to exceed the rate of obsolescence. Furthermore, a higher depreciation rate, i.e., lower δ^A, decreases $HR_E(t)$ in all years. It also decreases $HR_D(t)$, but only for periods 3 to 7, when maintenance fees are due, but increases $HR_D(t)$ in the first two periods.

δ^G determines the evolution of returns of already examined patent applications, $F^G(u^G \mid t)$. Therefore δ^G does neither impact $HR_D(t)$ nor $HR_E(t)$, and is identified by the variation in $HR_X(t)$ only.

As with other renewal models, the main caveat of our estimates is the sensitivity to the functional forms assumed for the distribution of returns. As Lanjouw (1998) notes: "Unlike the patents which are dropped, and which thereby indicate that they have expected returns at that point bounded above by the renewal fee, there is no information in the data which directly identifies an upper bound on the returns generated by patents which renew until the statutory term. The value of the patents in this group is identified indirectly by the functional form assumptions, together with the fact that the potential for high returns in the future influences renewal decisions in the early years."

Parameter	Estimates	(s.e.)
β (fixed)	0.9500	-
θ (fixed)	0.9535	-
μ_{IR}	5.9015	(0.0491)
σ_{IR}	1.8865	(0.0222)
q	0.7307	(0.0032)
ϕ^A	0.9659	(0.0011)
σ_0^A	1.4090	(0.0238)
δ^A	0.8400	(0.0101)
δ^G	0.9363	(0.0026)
Age-Cohort Cells	212	
Size of Sample	137,427	
Size of Simulation	412,281	
$Var_{All}(h_N)$	0.117316	
$MSE_{All}{}^{\dagger}$	0.000855	
$1 - MSE_{All}/Var_{All}(h_N)$	99.27%	
$Var_E(h_N)$	0.050834	
MSE_E	0.000115	
$Var_D(h_N)$	0.002848	
MSE_D	0.000154	
$Var_X(h_N)$	0.000586	
MSE_X	0.000619	

†MSE is the sum of squared residuals divided by the number of age-cohort cells.

Table 1: Parameter Estimates

4.4 Estimation Results

The estimation results are presented in Table 1.[34]

Fit of the model To get an indication of how well the estimated model fits the data, we compare the simulated with the sample hazard proportions. Furthermore, we also report how much of the variability in the sample hazard proportions can be explained by the model. Figures 4-6 show the simulated and sample hazard rates from the pooled data. By looking at the hazard proportions of examination requests and declarations, $HR_E(t)$ and $HR_D(t)$, one can see that there are no major deviations between the empirical and simulated moments. The model seems to capture all sharp increases as well as decreases. The mean square errors (MSE_E and MSE_D) are low compared to the variance in the actual hazard proportions ($Var_E(h_N)$ and $Var_D(h_N)$). Only 5.41%, respectively 0.23% of the variance in the actual

[34] A sketch of how the value functions and the cut-off values have been calculated can be found in Appendix A.2. We are using a parametric bootstrap method to obtain the standard errors as described in Appendix A.3.

hazard proportions is not accounted for by the model. For the hazards of expiration, however, the model overpredicts the hazard proportions for the years 12 and 16, and underpredicts them in all others. Consequently, the mean square error, MSE_X, is high compared to the variance. Why the model performs poorly in explaining the variation in $HR_X(t)$ may lie in the assumptions we have made regarding the cost structure and the duration of examination. The kink in year 16 coincides with the year when the official renewal fees almost double and then stay the same for the following years. However, the real costs associated with patent renewal might be much higher such that the official renewal fees represent only a fraction of them. This might explain why we do not observe a kink in the actual hazard proportions.[35] The jump in year 12 is due to our assumption that examination takes exactly 4 years for all applications and applicants always proceed the examination unless the application becomes completely worthless. According to our model, owners of patents of lower economic value defer examination until the last deferment period, and then decide whether to request it. If they request examination the patent will be granted exactly 4 years later. However, for many of these patents the value will have depreciated such that the renewal fees in year 12 will exceed the expected returns. In reality, the duration of examination is heterogeneous and the decision whether to continue examination might be endogenous as well. Therefore, the patent lapses the model predicts for year 12 are allocated around this year in the actual data. Furthermore, we have assumed that the examination costs are the same for all applicants. However, the actual examination costs should differ across applicants. Applicants with patents of lesser economic value should have requested examination earlier than predicted by the model if their examination costs were low enough. Applicants with higher examination costs should have postponed the examination request or even dropped the application although their applications were relatively valuable. Therefore, we observe higher hazard proportions of expiration especially for younger patents in the sample compared to the ones predicted by the model.[36] Nevertheless, the overall MSE_{All} is very low, suggesting that our estimated model fits the data well and is able to explain 99.27% of the overall variation.

Estimated parameters We now turn to the discussion of the estimated model parameters.[37] The initial distribution of returns is determined by μ_{IR} and σ_{IR}, and implies a mean initial potential return from patent protection for all applications of 2,155 CAD$ (122 CAD$)

[35] Another possible explanation is that the assumption of a constant rate of obsolescence for all granted patent applications might be unrealistic. An increasing rate of obsolescence for older patents might provide a better fit for the progression of the hazard proportions of expiration but would make calculations and identification more difficult.

[36] One possible way to alleviate this bias is to assume that the costs of examination are proportional to the duration of examination. The examination costs would then simply be a function of the duration of examination making them heterogeneous across applicants. However, the problem arises how to assign a duration to applications for which examination has never been requested, or which have never been granted.

[37] All monetary values are in units of 2002 CAD$. Standard errors are reported in parenthesis.

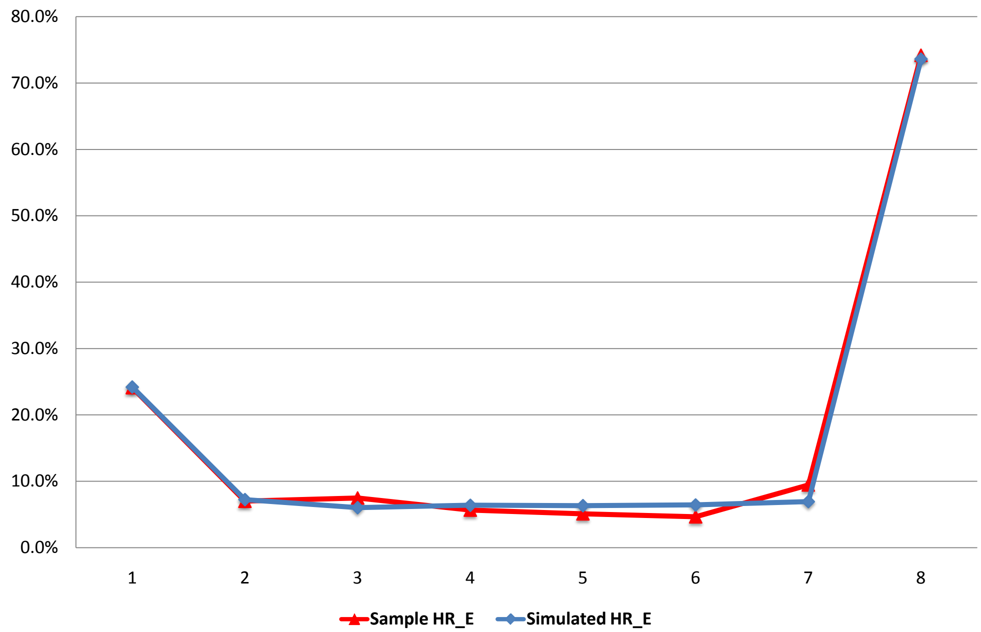

Figure 4: Simulation vs. Sample Hazard Proportions $HR_E(t)$

and a median value of 365 CAD$ (17 CAD$). The parameters ϕ^A, σ_0^A, and δ^A determine the evolution of returns during the application stage. The implication of these parameters, especially of ϕ^A being close to 1, is that before an application is examined the applicants expect high and slowly decreasing learning opportunities. If an applicant is not able to learn how to increase the returns from his patent application the next years returns depreciate by 16%. In Table 2 we see that 53.9% of pending patent applications in the second year and still 46.8% in the eighth year are able to increase the potential returns from patent protection and defy depreciation. Interestingly, although learning opportunities for Canadian patent applications diminish with age, they do it at a much slower pace as estimated for granted patent applications by previous patent renewal studies. For example, Pakes (1986) reports that learning is over by age 5 for German patents. Lanjouw (1998) estimates a similar speed of learning. This shows that the uncertainty underlying pending patent applications is high and is resolved only slowly over time.

The parameter q, which was defined as the fraction of potential returns from patent protection that can already be realized before the patent is finally granted, is estimated to be 73.1%. Although the applicant practically has not yet gained the right to enforce his right to exclude others, he is able to profit from having a pending application. This means that even

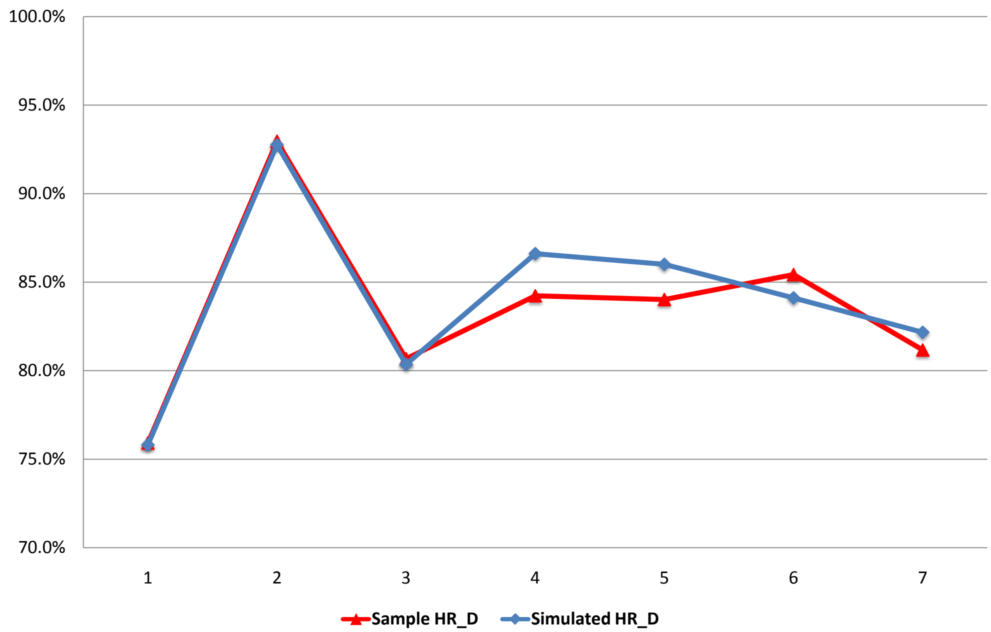

Figure 5: Simulation vs. Sample Hazard Proportions $HR_D(t)$

Age	2	3	4	5	6	7	8
$Pr(g_t^A \geq \delta^A)$	53.95%	52.78%	51.61%	50.41%	49.21%	47.99%	46.76%
(s.e.)	(0.81%)	(0.79%)	(0.77%)	(0.75%)	(0.73%)	(0.71%)	(0.69%)

Table 2: Learning Possibilities During the Application Stage

though he might never receive a patent on his invention, the realized value might still exceed the expenses. Since we assumed that there are no learning possibilities for already examined patent applications, the returns from full patent protection depreciate at $1 - \delta^G = 6.39\%$ per year.

5 Implications

Value of Canadian patent applications In this section we use the estimated parameters to calculate the value distributions of Canadian patent applications for the 1989 cohort. We simulate the patent system by taking 250,000 pseudo-random draws from the initial distributions and passing them through the model using the estimated parameter values. Then, the net value of protection defined as the discounted present values of the streams of returns less the discounted maintenance fees was calculated for each simulated application.

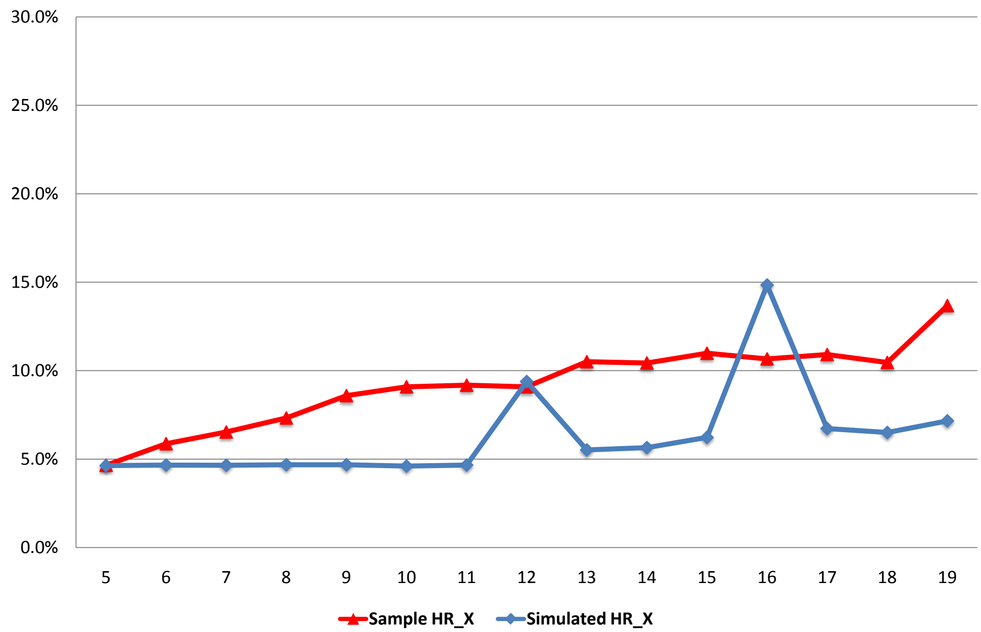

Figure 6: Simulation vs. Sample Hazard Proportions $HR_X(t)$

In case the application was still pending we multiplied the return in the respective period by q and in case examination had been requested we subtracted the discounted costs incurred for examination.

Table 3 presents the simulated value distributions for all patent applications, for applications which have been granted, and for applications which have not been granted. All monetary values are in 2002 CAD$. Similar to previous renewal studies, we find that the value distributions are highly skewed. The median simulated application value is 2,132 CAD$, whereas the mean value is 25,743 CAD$. Less than 10% of all applications are worth more than 50,870 CAD$ and less than 0.1% are worth more than 1,705,073 CAD$.

Unsurprisingly, there is a huge difference between patents and not granted applications. The average value of a patent is 50,954 CAD$. 50% are worth more than 15,361 CAD$ and 1% even more than 615,681 CAD$. These numbers confirm the results of previous patent renewal studies for other countries (Serrano 2006 for the USA; Deng 2007 for EPO patent applications). Patent applications which have never been granted were worth 4,547 CAD$ on average. Interestingly, the median value is positive with 184 CAD$.

Additionally, we are able to report what part of the value is generated before and what part after a patent has been granted. It seems that on average a patent owner is able to realize

Percentile	All Applications	Patents Overall Value	Patents Before Grant†	Not Granted Applications
50	2,132	15,361	42.15%	184
(s.e.)	(213)	(794)	(0.54%)	(36)
75	16,299	40,096	66.06%	2,093
(s.e.)	(894)	(1,915)	(0.36%)	(203)
90	50,870	99,029	88.97%	8,457
(s.e.)	(2,535)	(4,721)	(0.51%)	(701)
95	97,697	178,425	100.00%	17,445
(s.e.)	(4,844)	(9,284)	(-)	(1,281)
99	362,397	615,681	100.00%	70,463
(s.e.)	(21,352)	(40,859)	(-)	(5,076)
99.9	1,705,073	2,654,362	100.00%	400,452
(s.e.)	(125,165)	(225,414)	(-)	(30,818)
Mean Value	25,743	50,954	50.38%	4,547
(s.e.)	(1,536)	(2,961)	(0.43%)	(393)

†Calculated as the fraction of returns which accrued before the patent had been granted.
We did not subtract any costs to avoid negative numbers.

Table 3: Value Distributions for Cohort 1989 in 2002 CAD$

50.38% of the overall value already during the application and examination stages. Only less than 50% of all patents have realized more than 67.85% of the overall value during the patent stage. These are mostly patents which have requested examination very early. Owners of more than 5% of granted patent applications had only been able to accrue value during the application and examination stages. These patents became worthless shortly before or after they had been granted.

Withdrawn or not granted patent applications account for 54.33% of all patent applications. According to the simulation results, the owners of these applications do not make losses on average. Since applicants can profit from a pending application and realize 73.07% of potential returns from patent protection already before the patent issues, even the 50th percentile is positive. The other reason why we observe positive values for not granted patent applications is that some of them have become obsolete or failed examination in spite of having generated high returns in the past.

Value of deferment Now, we use the estimated parameters to shed light on the role of the possibility to defer the examination request. We calculate the vectors of cut-off values for two additional patent systems: one which allows deferment for up to six years, and one for up to five. Using the same simulated cohort of applications as in the previous section with a patent system which allows deferment for up to seven years, we calculate and compare the

Age	$L = 5$	$L = 6$	$L = 7$
1	58,914 (262)†	58,091 (250)	57,460 (240)
2	14,226 (199)	14,007 (201)	13,846 (202)
3	11,383 (116)	11,142 (128)	10,986 (128)
4	9,464 (89)	9,095 (99)	8,872 (115)
5	8,361 (88)	7,835 (94)	7,562 (108)
6	76,998 (292)	7,096 (96)	6,715 (102)
7		64,959 (276)	6,055 (97)
8			54,267 (270)
$\sum$	179,346 (352)	172,225 (338)	165,763 (340)

†Standard errors in parenthesis.

Table 4: Examination Requests

value distributions across the patent systems. This allows us to calculate the option value of the possibility to defer the examination request for one additional year. Furthermore, we also compare the number of total examination requests and assess the implications of different lengths of deferment for the patent office's workload.

Table 4 presents the simulated numbers of patent examination requests for the years examination can be requested. The table shows that the overall number of requests increases if we shorten the deferment period. It will increase by 4.13% (0.083%) if we reduce the period of request for examination by one year and by 8.19% (0.150%) if we reduce it by two years. This is consistent with the analysis by Yamauchi and Nagaoka (2008) who observed a significant increase in the number of requests for patent examinations in Japan, while the number of patent applications remained rather stable. They show empirically that one of the major causes of the increase was the shortening of the deferral period from 7 to only 3 years. The difference is highest in the last period in which examination can be requested.

In the patent system with a maximum deferment period of five years 46.93% of all examinations were requested in the last year, whereas in the patent system with a maximum deferment period of seven years only 32.74%. The explanation is that applicants are given additional time to evaluate their invention and unveil the uncertainty surrounding it. The additional deferment period permits two types of corrections. First, it allows those applications that become obsolete or are exposed to value depreciation in the following year not to request costly examination. Second, applicants may learn that their inventions are capable of generating higher returns and still request examination.

The effect on the value distributions of the simulated patent cohort is consistent with the effect on the number of examination requests. Since examination is requested early for patents which are known to be valuable as early as at the application date, patents in the top percentiles of the distribution are not affected by the extended deferment period. However,

change in the	$5 \leftarrow 7$ years	$6 \leftarrow 7$ years
value/ all applications	-5.49%	-2.56%
(s.e.)	(0.168%)	(0.080%)
value/ patents	-4.81%	-2.25%
(s.e.)	(0.152%)	(0.081%)
mean value/ patents	-11.99%	-5.93%
(s.e.)	(0.398%)	(0.178%)
value/ not granted applications	-11.86%	-5.56%
(s.e.)	(0.561%)	(0.283%)
mean value/ not granted applications	-5.37%	-2.34%
(s.e.)	(0.660%)	(0.321%)

Table 5: Option Value of Deferment

a shorter deferment period reduces the effect of both correction mechanisms. Many owners of applications with initially low returns are deprived of additional time to reevaluate the value of their inventions and to request examination in case they would have discovered a way to increase them. Besides, applications which devaluate in the sixth or the seventh year nevertheless request examination since they have to decide before this information is revealed to them. As presented in Table 5 the value of all patents decreases by 2.25% if we shorten the deferment period by one year and by 4.81% if we shorten it by two years. Since the number of examination requests increases if we reduce the maximum deferment period, the decrease in the average patent value is even higher.

The value of applications which have been withdrawn or have failed examination decreases by 2.34%, respectively 5.37%, on average. More applicants request examination and incur costs for the examination if the deferment period is shortened. Consequently, the value of all patent applications in the cohort falls by even 5.56%, respectively 11.86%.

6 Conclusion

The model developed in this paper is the first to embed the option of deferred patent examination in the context of stochastic optimization. We utilize the rich information from deferment and renewal actions to estimate parameters of the value distribution of Canadian patent applications and granted patents, as well as of the associated learning process. Knowledge of these parameters allows us to perform two simulation experiments and to study the impact of the timing of examination on the patent office's capacity problem as well as on the value of unexamined and granted patents.

Our first main finding is that a substantial part of the value from patent protection is generated in the time before a patent gets actually granted. We estimate that already during the application and examination stages an owner of a pending patent application is able to realize 73.09% of the returns he would generate if he had full patent protection. As a consequence, the majority of Canadian patent applications which have never been granted have a positive discounted value. Furthermore, the learning process of the value of applications which are still pending is much slower compared to the one of granted patents studied in previous literature.

In addition, our model allows us to simulate a change in the patent system from a seven years to a five years deferment period. This experiment is particularly interesting considering that the maximum deferment period in the Canadian patent system was indeed shortened from seven to five years for applications filed on or after October 1, 1996. The simulation experiment resulted in an increase in the number of examination requests which have to be dealt with at the Canadian patent office by 8.19%.[38] The applicants were deprived of time necessary to reduce the uncertainty associated with the value of their inventions. Therefore, many applications which would have turned out to be valuable in the future were withdrawn. Even more applicants decided to request examination and incur the corresponding costs although they would have become worthless shortly after. We estimate a considerable negative impact on the value of unexamined as well as granted patents, as a consequence.

Although we have used data on Canadian patent applications the results ought to be valid for other patent systems as well. A possibility to defer the examination request does not only reduce the patent office's workload but also acts as a quality control mechanism. Applicants seeking patent protection for inventions with highly uncertain value have the possibility to defer the examination request after the uncertainty has been resolved. Nevertheless, delayed examination may create additional possibilities for applicants to act strategically and increase uncertainty in the marketplace. However, to constrain strategic behavior by applicants most countries allow third parties to request examination and impose a fee on this activation right to prevent abuses. Concern has also been expressed that deferred patent examination could potentially increase patent filing rates. Indeed, our estimates show that for many patent applications returns realized during the application stage have been high enough to cover even the application costs. Nevertheless, an increase in the number of high-quality filings should not give cause for concern. To avoid an increase in poor-quality filings one could either try to raise the quality threshold for initial filings, involve third parties, or use the deferment fee structure as an additional policy instrument weeding out such applications.

The literature on the optimal renewal fees starting with Cornelli and Schankerman (1999)

[38]In reality the number of patent examination requests for Canadian applications has increased by about 6.9% after the reform. The deviation from the estimated percentage number may arise out of three reasons: 1) The simulated results apply to 1989 applications. 2) We do not take potential consequences of the policy intervention for the application filing decision into account. 3) We have used data on applications which were filed directly with the Canadian patent office for estimation and excluded all PCT applications.

and Scotchmer (1999) has identified the renewal structure as a direct revelation mechanism. Assuming heterogeneity in R&D productivity across firms and information asymmetry on the part of the government, optimal patent renewal fees should be low in early years and rise sharply with patent length. Baudry and Dumont (2009) arrive at the same conclusion for the welfare optimizing "one profile fits all" renewal fees. Nevertheless, both studies disregard the fact that in many patent systems the patent renewal fees constitute only a part of the total statutory fees. Our framework incorporates application fees, deferment fees, as well as patent renewal fees. Assuming a particular welfare function which relates the deadweight loss to the private value from patent protection, one could try to determine the welfare optimizing cost structure taking into account the interplay between the different types of fees. For example Cornelli and Schankerman (1999) report that the optimal patent renewal schedule should be more sharply graduated if there is post-patent learning, compared to the case when there is no uncertainty about the value of the invention. If we assume that applicants defer examination because they are highly uncertain about the value of their invention, then welfare could be increased by applying different schemes to deferment fees and patent renewal fees. Thus, the model developed in this paper provides the suitable framework for tackling these research questions in a more comprehensive manner.

References

Baudry, M., and B. Dumont (2009): "A bayesian real option approach to patents and optimal renewal fees," Working Paper.

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

Cornelli, F., and M. Schankerman (1999): "Patent renewals and R&D incentives," *The RAND Journal of Economics*, 30(2), 197–213.

Deng, Y. (2011): "A dynamic stochastic analysis of international patent application and renewal processes," *International Journal of Industrial Organization*, 29(6), pp. 766–777.

Efron, B., and R. J. Tibshirani (1993): *An introduction to the bootstrap*. Chapman & Hall, New York.

F.T.C. (2003): "To promote innovation - the proper balance of competition and patent law and policy," Discussion paper, GPO.

Gans, J., D. Hsu, and S. Stern (2008): "The impact of uncertainty intellectual property rights on the market for ideas: Evidence from patent grant delays," *Management Science*, 54(5), 982–997.

Guellec, D., and B. de La Potterie (2007): *The economics of the European patent system*. Oxford University Press Oxford.

Hall, B., and R. Ziedonis (2001): "The patent paradox revisited: an empirical study of patenting in the US semiconductor industry, 1979-1995," *RAND Journal of Economics*, 32(1), pp. 101–128.

Harhoff, D. (2006): "The role of patents and licenses in securing external finance for innovation," in *Advancing Knowledge and the Knowledge Economy*, ed. by J. Cantwell, A. Gambardella, and O. Granstrand, vol. Kahin, B. and Foray, D., pp. 331–350. MIT Press.

——— (2012): "Deferred examination - Searching for a double dividend," Working Paper.

Harhoff, D., and S. Wagner (2009): "The duration of patent examination at the European Patent Office," *Management Science*, 55(12), 1969–1984.

Henkel, J., and F. Jell (2010): "Patent Pending - Why faster isn't always better," *Available at SSRN 1738912*.

Lanjouw, J. O. (1998): "Patent protection in the shadow of infringement: Simulation estimations of patent value," *The Review of Economic Studies*, 65(4), pp. 671–710.

Lev, B. (2001): *Intangibles: Management, measurement, and reporting*. Brookings Inst Press.

McFadden, D. (1989): "A method of simulated moments for estimation of discrete response models without numerical integration," *Econometrica: Journal of the Econometric Society*, 57(6), pp. 995–1026.

McGinley, C. (2008): "Taking the heat out of the global patent system," *Intellectual Asset Management*, 31, pp. 24–29.

N.R.C. (2004): "A patent system for the 21st century. Board on science, technology, and economic policy (2004)," Discussion paper, Washington, DC: National Academies Press.

Opperman, C. (2009): "The elephant in the room," *Intellectual Asset Magazine*, July/August, pp. 53–59.

Pakes, A. (1986): "Patents as options: Some estimates of the value of holding European patent stocks," *Econometrica*, 54(4), pp. 755–784.

Pakes, A., and D. Pollard (1989): "Simulation and the asymptotics of optimization estimators," *Econometrica*, 57(5), pp. 1027–1057.

Pakes, A., and M. Schankerman (1984): "The rate of obsolescence of knowledge, research gestation lags, and the private rate of return to research resources," in *Patents, R&D, and Productivity*, ed. by Z. Griliches, pp. 73–88. The University of Chicago Press.

Schankerman, M., and A. Pakes (1986): "Estimates of the value of patent rights in European countries during the post-1950 period," *The Economic Journal*, 96(384), pp. 1052–1076.

Scotchmer, S. (1999): "On the optimality of the patent renewal system," *The RAND Journal of Economics*, 30(2), pp. 181–196.

Serrano, C. (2006): "The market for intellectual property: Evidence from the transfer of patents," Ph.D. thesis, University of Minnesota.

Yamauchi, I., and S. Nagaoka (2008): "Complementary reforms of patent examination request system in Japan," IIR Working Paper.

A Appendix

A.1 Proof that $\widetilde{V}^E(t,r_t) - \widetilde{V}^D(t,r_t)$ is increasing in r_t

The proof is done by induction for the traditional way of modeling the examination stage. The proof for the alternative way of modeling the examination stage is identical to setting $\pi_{Grnt} = 1$.

We know that $\widetilde{V}^E(t,r_t)$, $\widetilde{V}^D(t,r_t)$, $\widetilde{V}^K(t,r_t)$ are continuous and increasing in r_t. Assume that $q < \pi_{Grnt}$ and without loss of generality that the examination period takes only one year ($S = 1$).

Consider period $t = a = L$ first:

$$\widetilde{V}^E(L,r_L) = qr_L - (c_L + C_{PO}^{Exam} + C_{self}^{Exam} + \beta\theta\pi_{Grnt}C_{PO}^{Grnt}) + \beta\theta\pi_{Grnt}E\left[\widetilde{V_K}(L+1,r_{L+1}) \mid r_L\right];$$

$$\widetilde{V}^D(L,r_L) = qr_L - c_L + \beta\theta E\left[\widetilde{V_D}(L+1,r_{L+1}) \mid r_L\right];$$

$$\widetilde{V}^K(L+1,r_{L+1}) = r_{L+1} - c_{L+1} + \beta\theta E\left[\widetilde{V_K}(L+2,r_{L+2}) \mid r_{L+1}\right].$$

We show that $\widetilde{V}^E(L,r_L) - \widetilde{V}^D(L,r_L) + (C_{PO}^{Exam} + C_{self}^{Exam} + \beta\theta\pi_{Grnt}C_{PO}^{Grnt})$ is increasing in r_L:

$$\begin{aligned}
&\widetilde{V}^E(L,r_L) - \widetilde{V}^D(L,r_L) + (C_{PO}^{Exam} + C_{self}^{Exam} + \beta\theta\pi_{Grnt}C_{PO}^{Grnt}) = \\
&= \beta\theta\left\{\pi_{Grnt}E\left[\widetilde{V_K}(L+1,r_{L+1}) \mid r_L\right] - E\left[\widetilde{V_D}(L+1,r_{L+1}) \mid r_L\right]\right\} = \\
&= \beta\theta\{\pi_{Grnt}E\left[max\left(\widetilde{V}^K(L+1,r_{L+1}), \widetilde{V}^X(L+1,r_{L+1})\right) \mid r_L\right] - \\
&\quad -E\left[max\left[\widetilde{V}^E(L+1,r_{L+1}), \widetilde{V}^W(L+1,r_{L+1})\right] \mid r_L\right]\} =,
\end{aligned}$$

and since deferring examination is not possible in $t = L+1$,

$$= \beta\theta\{E\left[max\left(\pi_{Grnt}\widetilde{V}^K(L+1,r_{L+1}),0\right) \mid r_L\right] - E\left[max\left(\widetilde{V}^E(L+1,r_{L+1}),0\right) \mid r_L\right]\}.$$

Since $r_{L+1} = g_{L+1}r_L$ it suffices to show that $\pi_{Grnt}\widetilde{V}^K(L+1,r_{L+1}) - \widetilde{V}^E(L+1,r_{L+1})$ is increasing in $r_{L+1} \geq 0$:

$$\pi_{Grnt}\widetilde{V}^K(L+1,r_{L+1}) - \widetilde{V}^E(L+1,r_{L+1}) =$$

$$= \pi_{Grnt}r_{L+1} - \pi_{Grnt}c_{L+1} + \pi_{Grnt}\beta\theta E\left[\widetilde{V_K}(L+2,r_{L+2}) \mid r_{L+1}\right] - qr_{L+1} +$$

$$+(c_{L+1} + C_{PO}^{Exam} + C_{self}^{Exam} + \beta\theta\pi_{Grnt}C_{PO}^{Grnt}) - \beta\theta\pi_{Grnt}E\left[\widetilde{V_K}(L+2,r_{L+2}) \mid r_{L+1}\right] =$$

$$= \pi_{Grnt} r_{L+1} - q r_{L+1} + \pi_{Grnt} \beta\theta E\left[\widetilde{V_K}(L+2, r_{L+2}) \mid r_{L+1}\right] -$$

$$-\beta\theta\pi_{Grnt} E\left[\widetilde{V_K}(L+2, r_{L+2}) \mid r_{L+1}\right] + (c_L + C_{PO}^{Exam} + C_{self}^{Exam} + \beta\theta\pi_{Grnt} C_{PO}^{Grnt}) - \pi_{Grnt} c_{L+1} =$$

$$= (\pi_{Grnt} - q) r_{L+1} + C_{PO}^{Exam} + C_{self}^{Exam} + \beta\theta\pi_{Grnt} C_{PO}^{Grnt}) + (1 - \pi_{Grnt}) c_{L+1} \ ,$$

and increasing in $r_{L+1} \geq 0$.

Now, consider periods $t = 1, ..., L-1$:

$$\widetilde{V}^E(t, r_t) = q r_t - (c_t + C_{PO}^{Exam} + C_{self}^{Exam} + \beta\theta\pi_{Grnt} C_{PO}^{Grnt}) + \beta\theta\pi_{Grnt} E\left[\widetilde{V_K}(t+1, r_{t+1}) \mid r_t\right];$$

$$\widetilde{V}^D(t, r_t) = q r_t - c_t + \beta\theta E\left[\widetilde{V_D}(t+1, r_{t+1}) \mid r_t\right];$$

$$\widetilde{V}^K(t+1, r_{t+1}) = r_{t+1} - c_{t+1} + \beta\theta E\left[\widetilde{V_K}(t+2, r_{t+2}) \mid r_{t+1}\right].$$

Assume that $\widetilde{V}^E(t+1, r_{t+1}) - \widetilde{V}^D(t+1, r_{t+1})$ is increasing in r_{t+1}.

We have to show that with this assumption,

$\widetilde{V}^E(t, r_t) - \widetilde{V}^D(t, r_t) + (C_{PO}^{Exam} + C_{self}^{Exam} + \beta\theta\pi_{Grnt} C_{PO}^{Grnt})$ is increasing in r_t.

$$\widetilde{V}^E(t, r_t) - \widetilde{V}^D(t, r_t) + (C_{PO}^{Exam} + C_{self}^{Exam} + \beta\theta\pi_{Grnt} C_{PO}^{Grnt}) =$$

$$= \beta\theta\{\pi_{Grnt} E\left[\widetilde{V_K}(t+1, r_{t+1}) \mid r_t\right] - E\left[\widetilde{V_D}(t+1, r_{t+1}) \mid r_t\right]\} =$$

$$= \beta\theta\{E\left[max\left(\pi_{Grnt}\widetilde{V}^K(t+1, r_{t+1}), 0\right) \mid r_t\right] -$$

$$-E\left[max\left(\widetilde{V}^E(t+1, r_{t+1}), \widetilde{V}^D(t+1, r_{t+1}), 0\right) \mid r_t\right]\}.$$

Since $\pi_{Grnt}\widetilde{V}^K(t+1, r_{t+1}) - \widetilde{V}^E(t+1, r_{t+1}) =$

$= (\pi_{Grnt} - q) r_{t+1} + C_{PO}^{Exam} + C_{self}^{Exam} + \beta\theta\pi_{Grnt} C_{PO}^{Grnt}) + (1 - \pi_{Grnt}) c_{t+1}$ is increasing in r_{t+1},

and by assumption $\widetilde{V}^E(t+1, r_{t+1}) - \widetilde{V}^D(t+1, r_{t+1})$ is increasing in r_{t+1},

then $\pi_{Grnt}\widetilde{V}^K(t+1, r_{t+1}) - \widetilde{V}^D(t+1, r_{t+1})$ must also be increasing in r_{t+1}.

Given that r_{t+1} is increasing in r_t, $\widetilde{V}^E(t, r_t) - \widetilde{V}^D(t, r_t)$ must also be increasing in r_t and the proof is complete.

A.2 Value Functions and Cut-off Values

Here, we present a sketch of how the value functions and the cut-off values can be calculated assuming $L = 7$, $S = 1$ and K expected examination costs. The presentation is general and not restricted to a specific type of stochastic specification or assumptions concerning the examination stage.

Periods 2-20 if patent is already granted: $(g_t = g_t^G \sim F_t^G(u_T^G))$

Period 20:

$\widetilde{V}^K(20, r_{20}) = r_{20} - c_{20} = g_{20}^G r_{19} - c_{20};$

Cut-off value:

$\widetilde{V}^K(20, r_{20}) = 0 \Rightarrow \hat{r}_{20}^K = c_{20}$ (respectively $\hat{g}_{20}^K = \frac{c_{20}}{r_{19}^G}$).

Period 19:

$$\widetilde{V}^K(19, r_{19}) = r_{19} - c_{19} + \theta\beta \int_{\hat{g}_{20}^K}^{B} \widetilde{V}^K(20, r_{20}) dF_{20}^G(u_{20}^G) =$$

$$= r_{19} - c_{19} + \theta\beta \int_{\frac{c_{20}}{r_{19}^G}}^{B} (u_{20}^G r_{19} - c_{20}) dF_{20}^G(u_{20}^G);$$

Cut-off value:

$\widetilde{V}^K(19, r_{19}) = 0 \Rightarrow \hat{r}_{19}^K$ (respectively $\hat{g}_{19}^K = \frac{\hat{r}_{19}^K}{r_{18}}$).

Period 18:

$$\widetilde{V}^K(18, r_{18}) = r_{18} - c_{18} + \theta\beta \int_{\hat{g}_{19}^K}^{B} \widetilde{V}^K(19, r_{19}) dF_{19}^G(u_{19}^G) =$$

$$= r_{18} - c_{18} + \theta\beta \int_{\hat{g}_{19}^K}^{B} \left[u_{19}^G r_{18} - c_{19} + \theta\beta \{ \int_{\hat{g}_{20}^K}^{B} (u_{20}^G r_{19} - c_{20}) dF_{20}^G(u_{20}^G) \} \right] dF_{19}^G(u_{19}^G) =$$

$$= r_{18} - c_{18} + \theta\beta \int_{\frac{\hat{r}_{19}^K}{r_{18}^G}}^{B} \left[u_{19}^G r_{18} - c_{19} + \theta\beta \{ \int_{\frac{c_{20}}{u_{19}^G r_{18}^G}}^{B} (u_{20}^G u_{19}^G r_{18} - c_{20}) dF_{20}^G(u_{20}^G) \right] dF_{19}^G(u_{19}^G);$$

Cut-off value:

$\widetilde{V}^K(18, r_{18}) = 0 \Rightarrow \hat{r}_{18}^K$ (respectively $\hat{g}_{18}^K = \frac{\hat{r}_{18}^K}{r_{17}}$).

Period 17:

...

Periods 1-8 in case of a pending application: $(g_t = g_t^A \sim F_t^A(u_t^A))$

Since the shock on the growth rate during the examination is unexpected one also has to calculate $\widetilde{V}_t^K(t, r_t)$ and $\hat{r}_t^K$ (respectively $\hat{g}_t^K$), $t \in 2, ..., 20$ assuming $g_t = g_t^A \sim F_t^A(u_t^A)$ since these are the returns the applicants expect to receive in future periods. These value functions and cut-off values are then used for the calculation of value functions of the applicant.

Period 8

if examination is requested:

$$\widetilde{V}^E(8, r_8) = qr_8 - K - c_8 + \theta\beta\pi_{Grnt} \int_{\hat{g}_9^K}^{B} \widetilde{V}^K(9, r_9) dF_9^A(u_9^A) =$$

$$= qr_8 - K - c_8 + \theta\beta\pi_{Grnt} \int_{\frac{\hat{r}_9^K}{r_8}}^{B} \widetilde{V}_9^K(9, g_9^A r_8) dF_9^A(u_9^A);$$

cut-off value:[39]

$\widetilde{V}_8^E(8, r_8) = 0 \Rightarrow \hat{r}_8^E$ (respectively $\hat{g}_8^E = \frac{\hat{r}_8^E}{r_7}$).

Period 7

if examination is requested:

$$\widetilde{V}_7^E(7, r_7) = qr_7 - K - c_7 + \theta\beta\pi_{Grnt} \int_{\frac{\hat{r}_8^K}{r_7}}^{B} \widetilde{V}_8^K(8, g_8^A r_7) dF_8^A(u_8^A);$$

if examination is deferred:

$$\widetilde{V}_7^D(7, r_7) = qr_7 - c_7 + \theta\beta\pi_{Grnt} \int_{\frac{\hat{r}_8^E}{r_7}}^{B} \widetilde{V}_8^E(8, g_8^A r_7) dF_8^A(u_8^A);$$

cut-off values:

$\widetilde{V}_7^D(7, r_7) = 0 \Rightarrow \hat{r}_7^D$ (respectively $\hat{g}_7^D = \frac{\hat{r}_7^D}{r_6}$);

$\widetilde{V}_7^E(7, r_7) = \widetilde{V}_7^D(7, r_7) \Rightarrow \hat{r}_7^E$ (respectively $\hat{g}_7^E = \frac{\hat{r}_7^E}{r_6}$).

Period 6

if examination is requested:

$$\widetilde{V}_6^E(6, r_6) = qr_6 - K - c_6 + \theta\beta\pi_{Grnt} \int_{\frac{\hat{r}_7^K}{r_6}}^{B} \widetilde{V}_7^K(7, g_7^A r_6) dF_6^A(u_6^A);$$

if examination is deferred:

$$\widetilde{V}_6^D(6, r_6) = qr_6 - c_6 + \theta\beta\pi_{Grnt} \int_{\frac{\hat{r}_7^E}{r_6}}^{B} \widetilde{V}_7^E(7, g_7^A r_6) dF_7^A(u_7^A) + \int_{\frac{\hat{r}_7^D}{r_6}}^{\frac{\hat{r}_7^E}{r_6}} \widetilde{V}_7^D(7, g_7^A r_6) dF_7^A(u_7^A);$$

Cut-off value:

$\widetilde{V}_6^D(6, r_6) = 0 \Rightarrow \hat{r}_6^D;$

[39]Deferment is not possible anymore.

$\widetilde{V}_6^E(6, r_6) = \widetilde{V}_6^D(6, r_6) \Rightarrow \hat{r}_6^E.$

Period 5

...

A.3 Parametric Bootstrap

Since we do not know the empirical distribution of the observed hazard rates we will apply a parametric bootstrap method to estimate the standard errors of the parameters ω. Instead of simulating bootstrap samples that are i.i.d. from the empirical distribution as it is done in non-parametric bootstrap methods we simulate bootstrap samples that are i.i.d. from the estimated parametric model. Following Efron and Tibshirani (1993) we apply the following bootstrap algorithm:

1. Use the estimated parameters $\hat{\omega}^*$ and generate a random sample of N patent applications.

2. Simulate the decisions resulting from the model specification and obtain the sequence of pseudo hazard rates $\eta(\hat{\omega}^*)$.

3. Minimize the loss function in 8 using $\eta(\hat{\omega}^*)$ instead of h_N and obtain $\hat{\omega}_b^*$.

4. Repeat the steps 1.-3. B times.

5. Calculate the parametric bootstrap estimate of standard error:

$$\hat{se}_B = \left\{ \frac{\sum_{b=1}^{B} \left[\hat{\omega}_b^* - \hat{\omega}^*(.)\right]^2}{(B-1)} \right\}^{\frac{1}{2}}, \text{ where } \hat{\omega}^*(.) = \frac{\sum_{b=1}^{B} \hat{\omega}_b^*}{B}$$