Paper to be presented at

DRUID15, Rome, June 15-17, 2015

(Coorganized with LUISS)

Deferred Patent Examination

Dietmar Harhoff
Max Planck Institute for Innovation and Competition
Innovation and Entrepreneurship Research
dietmar.harhoff@ip.mpg.de

Ilja Rudyk
European Patent Office
Chief Economist Unit
irudyk@epo.org

Sebastian Stoll
Max Planck Institute for Innovation and Competition
Innovation and Entrepreneurship Research
sebastian.stoll@ip.mpg.de

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.

Jelcodes:L00,-

Deferred Patent Examination

May 28, 2015

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 and the German 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 and German 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; patent renewal model

JEL Classification Numbers: L00

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 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 of the patent application. Prior to 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, we provide a much richer foundation for the analysis of patent value - in particular, its realization in the very early stages of patent life - and for policy simulations than previous studies have done. Besides confirming findings of earlier studies on the distribution of patent value, we show that the private value of having just a pending application can already be substantial. Also, our model allows us to simulate the counterfactual impact of changes in the deferment period. We show that shortening the deferment period likely decreases the private value of pending applications (there is evidence that the private value of a pending application to the applicant arises from the uncertainty it creates for other users of the patent system,[1] and faster examination resolves uncertainty quicker), but increases the number of examination requests. These findings add to discussions among academics and practitioners about the introduction (respectively, reform) of deferred examination regimes.

Under a deferred examination regime, a patent office defines a certain period during which an applicant can request examination of his patent filing. In case the applicant makes this request

[1]Evidence from interviews and surveys suggests that applicants are intentionally increasing the volume and complexity of their filings and frequently delaying the examination process (using tactics like continuations or divisionals) in order to create uncertainty for other users of the system. See McGinley (2008); Opperman (2009); Harhoff and Wagner (2009); Henkel and Jell (2010).

during the pre-defined deferment period, his application enters the examination process, at the end of which it gets potentially granted. Otherwise, the application lapses. Patent offices 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 Canadian or the German patent office offer applicants a time period of up to seven years during which they can request examination. 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.[2]

The rationale behind the introduction of a deferment period was to reduce patent backlogs. In the last three decades the number of patent filings has risen substantially: partially due to increased tactical and strategic importance (Hall and Ziedonis, 2001; F.T.C., 2003; N.R.C., 2004), and partially due to lower costs and availability of patent protection (Harhoff, 2006; Guellec and de La Potterie, 2007; Bessen and Meurer, 2008). As a consequence, the patent workload has increased substantially giving reason for concern about its impact on examination quality. In reaction to a vast amount of unexamined and pending patent applications, the first system of deferred patent examination was introduced by the Dutch government on January 1, 1964. 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.

Basically, the decision about the introduction (respectively, the design) of a deferred examination system involves a trade-off between the merits of fast and that of delayed examination. The main argument for fast examination 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.

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

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). The main argument for delayed examination is that, if given additional time for assessing the value of their patents, some applicants might drop out of the examination process voluntarily. This, in turn, would reduce examination workloads. Whereas there has been some discussion on these arguments in the patent literature, there has not yet been a quantitative assessment of their relative significance based on extensive data. We fill this gap by estimating an extended patent renewal model on data from both the Canadian Intellectual Property Office (CIPO) and the German patent office (DPMA). In doing so, we add to the academic debate on deferred patent examination (respectively, from a broader perspective, on how to handle patent backlogs).

Patent renewal models describe the renewal decisions of patent applicants (respectively, holders) over a patent's life. Roughly, an applicant's renewal decision is between paying a renewal fee to keep his patent alive for one further period (and earn the corresponding returns) or letting his patent lapse. We contribute to the literature on stochastic patent renewal models (Pakes, 1986; Lanjouw, 1998; Serrano, 2006; Deng, 2011) by explicitly modelling an applicant's option to defer examination. What sets our model apart from existing models of patent renewal is that for the time a patent exists as pending application it allows for three possible decisions: to request examination, to defer examination, or to let the application lapse altogether. After examination and grant our model allows for only two decisions: to pay the renewal fees, or to let the patent lapse. We embed this choice structure 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. We allow unexamined patent applications to differ in terms of value from examined and granted applications, and we allow for the possibility that the applicant learns about the value of his patent.

In the empirical part of our paper, we use data from both the Canadian patent office (CIPO) and the German patent office (DPMA). In these systems, applicants have the option to defer examination of their patent applciations. Whereas in Germany this option existed from the outset, in the Canadian patent system it was introduced in October 1989. Between October 1989 and September 1996, the maximum deferment period in both offices was seven years. (In October 1996, the Canadian patent office shortened the maximum deferment period to five years.) We apply our patent renewal model to the Canadian and German patent data to estimate the parameters of the distribution of the private value of patents, the parameter capturing the private value of having a pending application, and the parameters governing 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. Both in the Canadian and the German patent system, learning possibilities seem to be relatively high and to 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.

2 Data on Canadian and German patent applications

We analyze data from both the Canadian and the German patent system. In both systems, patent applicants have the possibility to defer examination of their applications. Whereas in Germany the option to defer examination for up to seven years existed from the outset, in Canada it was introduced in 1989, and in 1996 the initial seven-year deferment period was reduced to five years.[3] In order to ensure a homogeneous institutional environment (and to be able to observe the full 20-year life-span of the majority of patents), for both the Canadian and the German patent system we focus our analysis on patent applications which were filed during the years 1989-1996. Since both Canada and Germany are PCT (Patent Cooperation Treaty) member states, applications which had gone through the PCT route only entered the national stages at the Canadian respectively the German patent office 30 months after their priority date (which is typically 18

[3]In a companion paper, Harhoff (2012) studies the policy reforms at the Canadian Intellectual Property Office (CIPO) in more detail.

Descriptive statistics for application cohorts 1989-1996:		
	Canada	Germany
Average nbr. of applications per year:	19,857	42,576
Fraction of applications for which examination was requested:	66%	63%
Fraction of requests for examination which were deferred:	64%	32%
Fraction of examined applications which got granted:	69%	60%
Average duration of examination:	4.1 yrs	2.6 yrs
Fee structure		
Renewal fees		
	(CAD)	(DEM)
Years 3-4	100	70
Year 5	100	90
Year 6	150	130
Year 7	150	180
Year 8	150	240
Year 9	150	290
Year 10	150	350
Year 11	200	470
Year 12	200	620
Year 13	200	760
Year 14	200	910
Year 15	200	1060
Year 16	400	1230
Year 17	400	1410
Year 18	400	1590
Year 19	400	1760
Year 20	400	1940
Fee for filing an application	200	100
Fee for requesting examination	400	400
Fee for publication of grant	300	150

Table 1: **Descriptive statistics and fee structures.** For the Canadian and the German patent system, the upper part of the table displays descriptive statistics for all national applications which were filed between 1989 and 1996. The lower part of the table describes the fee structure at the Canadian and the German patent office (status as of 1990). For comparison: In 1990, 1 DEM was worth 0.62 USD, and 1 CAD was worth 0.86 USD.

months from the application at the respective patent offices). In contrast, information on applications which had directly been submitted at the Canadian respectively the German patent office is available from the first day on. Therefore, both for the Canadian and the German patent office we exclude all PCT applications from the data and focus our analysis on national applications only. For all patent applications in our sample, we observe the application date and the date when examination was requested or the application was 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.

In particular, we have available data on 137,397 patent applications filed at the Canadian patent office during October 1989 and September 1996, and on 340,607 patent applications filed at the German patent office during 1989 and 1996. During the respective time periods, at both patent offices the maximum deferment period was seven years, and patent protection could be renewed for up to 20 years from application. The upper part of figure 1 gives an overview of the application cohorts 1989-1996 at both patent offices, and the upper part of table 1 offers descriptive statistics. The fractions of applications for which examination was requested are roughly similar at the Canadian and at the German office (66% in Canada and 63% in Germany), and so are the fractions of examined applications which got finally granted (69% in Canada and 60% in Germany). Apart from the yearly number of national filings, there are stark differences between the offices with regard to the circumstances of the examination process: On average, for the application cohorts considered, examination at the Canadian patent office (that is, the time between the examination request and the final grant decision) took around 4.1 years, whereas in Germany it took around 2.6 years. Also, whereas in Canada 64% of all examination requests are deferred, in Germany only 32% are deferred. Differences in the timing of the examination requests are mirrored in the hazard rates of examination requests, which are depicted in the lower part of figure 1. The hazard rate of examination requests at age t equals the number of examination requests made at application age t divided by the number of all applications still pending at age $t-1$. In the Canadian data, there is a very pronounced spike in the hazard rate at the end of the deferment period. In the German data, at the end of the deferment period there is also a sharp increase in the hazard rate, but it is less pronounced than in the Canadian data.

■ **Cost structure.** We first look at the cost structure at the Canadian patent office for applications from the cohorts 1989-1996 (compare the lower part of table 1): 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. There was one change in the nominal fee schedule which was applied to renewals starting from January 1, 2004. The 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, the fee for requesting examination to 400 CAD. A final fee of 300 CAD was due for the publication of the grant.

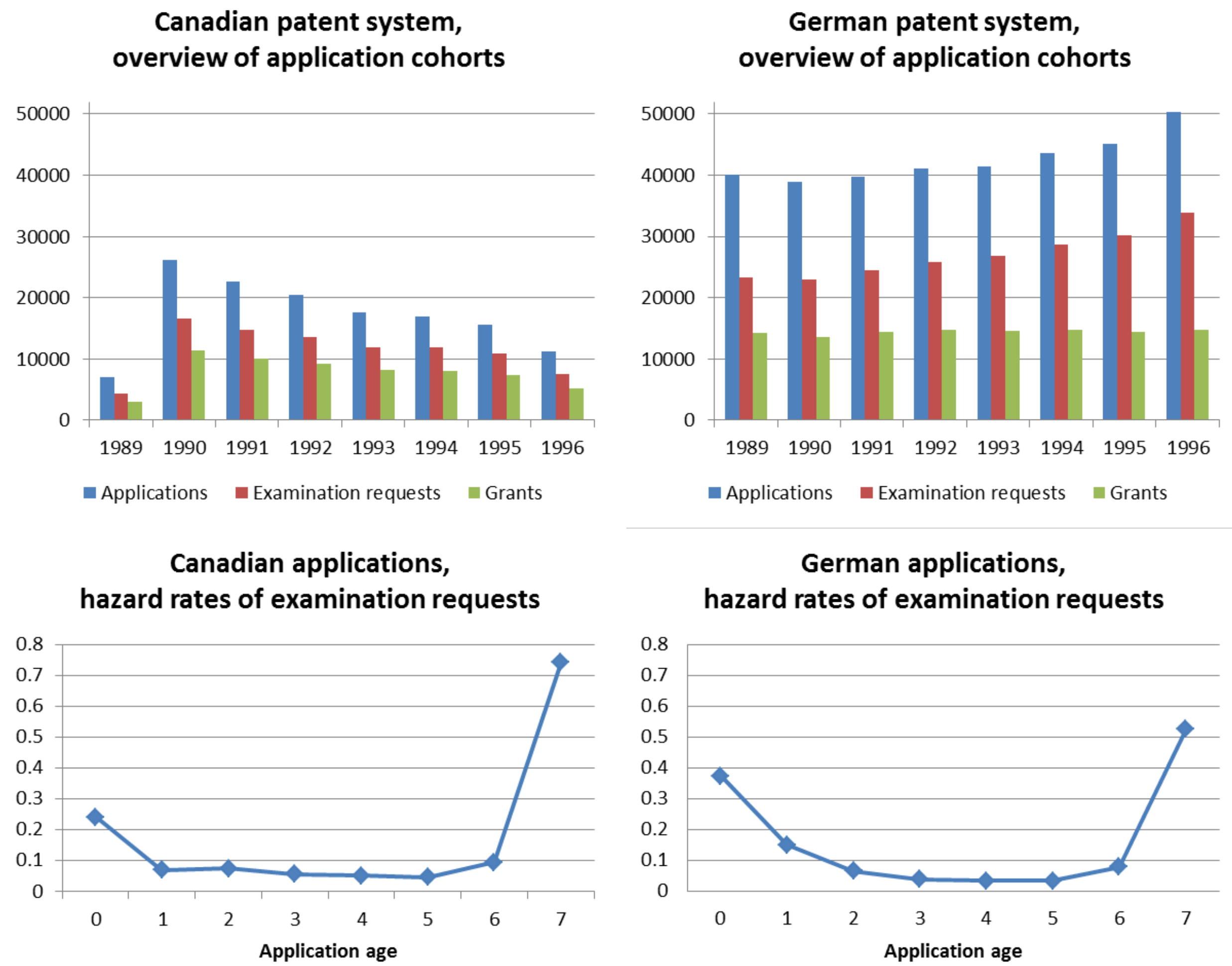

Figure 1: **Overview of application cohorts and hazard rates of examination requests.** For both the Canadian and the German patent system and the years 1989-1996, the upper two graphs show the number of applications per year, for how many of these applications examination was requested, and how many of the applications of a given year for which examination was requested were finally granted. (Note that in Canada the option to defer the examination request for up to 7 years was introduced in October 1989 and changed in October 1996. For the years 1989 and 1996, in our sample we thus only included applications which were filed between October and December 1989, respectively, January and September 1996.) The lower two graphs display the hazard rates of examination requests - that is, the fraction of pending applications of a given age for which examination is requested - for Canadian and for German applications (averaged over all application cohorts).

At the German patent office, the fee structure is more step. Yearly renewal fees start from 70 DEM for the third patent year and rise up to 1940 DEM for the 20th patent year. The fee for filing an application amounts to 100 DEM, that for requesting examination to 400 DEM, and that for the publication of the grant to 150 DEM. There were minor changes in the fee structure, namely in 1999 and in 2001. As for the Canadian fee data, in our computations we thus use weighted averages of the fees.

In practice, besides the official fees due at the patent offices applicants incur transaction costs which arise because applicants have to transfer the official fees to the patent offices in time and, especially during the examination period, applicants have to communicate with the patent offices, which often happens via a hired patent attorney. According to information from patent attorneys, the costs an applicant incurs during the examination period in addition to the official fees amount to around 3000 CAD in Canada and to around 5000 DEM in Germany.[4]

3 Structural Model of Deferred Examination and Patent Renewal

3.1 General Setup

■ **Patent system.** 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. We assume that examination has to be requested by the applicant within L years from the application day.[5] That is, we allow the agent to defer examination (and the payment of the associated fees C_{PO}^{Exam}) for up to L years (maximum deferment term). However, deferment is not free of charge - the agent has to pay fees c_t^A $(t \in 1, ..., L)$ to maintain his 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 a 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 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

[4] In our computations, in order to roughly account for the transaction costs during the examination period, for Canada we set C_{self}^{Exam} to 3,000 CAD and for Germany we set C_{self}^{Exam} to 5000 DEM. In order to roughly account for the transaction costs arising because of the need to transfer the official renewal fees in time, for Canada we add 50 CAD to each yearly payment to the patent office and for Germany we add 100 DEM.

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

that they are non-decreasing in t.[6] If any of the fees are not paid to the patent office in due time, 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 is positive, since a pending application can already create value for its owner, for example, by creating uncertainty for competitors or forming the basis for negotiations.[7]

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 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 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 ($F(u \mid t) \leq F(u \mid t+1)$).[8]

Before a patent is granted it has to pass 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. That is, the distributions of growth rates of patents which were granted after examination 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 a 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

[6]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.

[7]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.

[8]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.

$1-\theta$ all future returns can become a zero sequence. We allow the obsolescence rate to be the same for pending as well as granted applications.

■ **Agents.** We assume that every patent application belongs to exactly one profit maximizing agent. That is, in every period the agent always chooses the strategy with the highest expected payoff given the information he has available. We make the assumption that at the beginning of a period, the growth rate $g_t^{(\cdot)}$ is revealed to the agent, so that he knows the potential returns r_t $(= g_t^{(\cdot)} r_{t-1})$ from patent protection for this period. Furthermore, we assume that he is also informed about the future distributions of the growth rate, and thus is able to build expectations on how returns will evolve in the future.

If the applicant has not yet requested examination in period t, his growth rate is drawn from $F^A(u^A \mid t)$, and he expects the growth rates of patent returns to be distributed according to $F^A(u^A \mid t' > t)$ in future periods t'. In case the applicant requests examination and receives a response on the patentability of his 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. We assume that the growth rates for the periods after grant are drawn from $F^G(u^G \mid t)$ and model the decision of the examiner on the specific patent boundaries as an exogenous and unexpected shock on the evolution of patent returns.

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).

3.2 Value Functions and the Maximization Problem

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, $E\left[\widetilde{V_K}(t+1, r_{t+1}) \mid r_t\right]$. The expectation is taken with respect to $F^G(u^G \mid t)$. With β as the discount factor between periods the value function is:

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

with

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

Since the agent's choice in every period is discrete, there exist threshold returns $\hat{r}_t^K$, $t \in \{S+1, ..., T\}$, that determine the patent owner's optimal decision. $\hat{r}_t^K$ is the minimum patent return needed for an agent to decide to keep patent protection in period t and not to let it expire. It is the solution to $\widetilde{V}^K(t, r_t) = \widetilde{V}^X(t, r_t) = 0$.[9] 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 $\hat{r}_T^K = c_T$.

■ **Examination stage.** The examination stage differs from both the application stage and the patent stage in that both its duration and outcome does not solely depend on the applicant's actions, but on the interplay between the applicant and the patent office (respectively, the examiner). For reasons of tractability we refrain from an explicit modelling of the game the applicant and the examiner play. Instead, in modeling the examination stage we follow Deng (2011) and Serrano (2006). We assume that a patent examination takes S years and that the application will be approved for grant with probability π_{Grnt}. That is, in our model once the applicant requests examination he has to continue the examination process until the final decision of the examiner on patentability, and if examination was successful, he always wants his patent to be granted. 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

[9]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.

returns from full patent protection in the future:

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

Expectations are taken with respect to $F^A(u^A \mid t)$.

■ **Application stage.** During the application stage in each period the agent can decide between three possible actions: He can request examination (E), defer the (decision on the) examination request to the next period (D), or he can withdraw his application (W). The decision to request examination can be deferred for at most L periods. That is, an agent with a pending application at the beginning of period $t = L + 1$ has to decide whether he wants to request examination (E) or to withdraw (W) his application. Given the expected revenue $\widetilde{V}^E(a, r_a)$ from requesting examination from equation (3.2), and given that $\widetilde{V}^W(t, \hat{r}_t) = 0$, we define $\hat{r}_{L+1}^E$ as the minimum patent return 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 $\widetilde{V}^E(L + 1, r_{L+1}) = \widetilde{V}^W(L + 1, r_{L+1}) = 0$.

In earlier periods ($t = 1, ..., L$), besides the possibilities to withdraw (W) the application or to request examination (E), the agent can also 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 plus the expected returns from the option of having the same choices in the next period:

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

with

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

Expectations are taken with respect to $F^A(u^A \mid t)$.

Since the applicant has three options, in every period $t = 1, ..., L$, there exist two threshold values that determine the optimal choices: First, $\{\hat{r}_t^D\}_{t=1}^L$, which is the minimum patent return needed for the agent to defer the decision (D) at age t and not let it expire (W). It is the solution to $\widetilde{V}^D(t, r_t) = \widetilde{V}^W(t, r_t) = 0$. Second, $\{\hat{r}_t^E\}_{t=1}^L$, which is the minimum patent return needed for the agent to request an examination (E) instead of deferring the decision (D) at age t. It is the solution to $\widetilde{V}^E(t, r_t) = \widetilde{V}^D(t, r_t)$.

According to 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 gain of time for evaluation is reflected in the option value of future returns, 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.

Since the problem described above is finite and returns in one period only depend on returns realized in the previous period, the model can be solved for the sequences of the cut-off values by backward recursion. See Appendix A.2 for a sketch of how the model is solved recursively.

4 Structural Estimation of Patent Renewal Model

4.1 Estimation Strategy

We use a simulated minimum distance estimator as developed by McFadden (1989) and Pakes and Pollard (1989) for the estimation.[10] 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

[10] McFadden (1989) and Pakes and Pollard (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.

vector of structural parameters ω. As proposed in Lanjouw (1998), in order to determine the vector ω_0 of the true parameters we fit the hazard probabilities derived from our model to the hazard probabilities we observe in the data. Each structural 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 $j = E, D, K$, and the distribution of returns in each age, r_t, which in turn determine the hazard probabilities. This allows identification of the model parameters. Although in theory a solution to the structural model (that is, 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 distance estimator (SGMM) $\hat{\omega}_N$. The estimator is the argument that minimizes the norm of the distance between the vector of observed and simulated hazard proportions. We use a weighting matrix $A(\omega)$ to improve the efficiency of the estimator:

$$A(\omega)\left\|h_N - \eta_N(\omega)\right\| \text{ with } \hat{\omega}_N = arg\min_{\omega} A(\omega)\left\|h_N - \eta_N(\omega)\right\| \tag{1}$$

h_N is the vector of sample hazard proportions, $\eta_N(\omega)$ is the vector of simulated hazard proportions. $A(\omega) = diag\sqrt{\underline{n}/N}$ 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.

The simulation of hazard probabilities for a parameter set ω proceeds as follows: We first have to calculate the sequences of the cut-off values $\left\{\hat{r}_t^j\right\}$ with $j = E, D, K$. 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 values for all periods we perform simulations. In each simulation we take N pseudo random draws from the distribution of initial returns and N draws from the distributions of the growth rates g_t^A and g_t^G (t=1,...,20). Afterwards, we pass the draws of the initial returns 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 the simulation is inserted into the objective function (1). The objective function is minimized using a two step approach: We use a 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.

We 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 (that is, the percentage of granted patents that expire in a given year out of those granted and renewed in the previous period).

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 application 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. The maximum patent term in Canada and Germany is 20 years, and for the application cohorts in our sample the average duration of examination (S) was four years in Canada and three years in Germany. Thus, we calculate HR_X for 15 periods for Canada and for 16 periods for Germany.

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

4.2 Stochastic Specification

As in previous patent renewal studies,[11] we assume that the initial returns r_1 of the applications are lognormally distributed:[12]

$$log(r_1) \sim Normal(\mu_{IR}, \sigma_{IR}) \tag{2}$$

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

[11] See for example Pakes (1986).

[12] 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.

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 cases when the applicant is able to "learn" about how to increase the returns above the minimum growth rates. We assume that $\phi^A < 1$, meaning that the probability of getting higher returns will decrease with age t.

We model the evolution of the growth rates during the patent stage, g_t^G, in a more static way. We assume that learning possibilities disappear once the patent has been examined. This means that uncertainty about future returns from patent protection 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 at $\beta = 0.95$. Furthermore, since the first weighted hazard rate of expiration (HR_X for period 5 for Canada, HR_X for period 4 for Germany) is 0.0465 (respectively, 0.0393), we set $\theta = 0.9535$ (respectively, $\theta = 0.9607$). Year 5 (respectively, year 4) is when the first patent applications are granted, and we assume that the earliest granted patents do not expire in the first year after grant because of too high renewal fees (which are only 100 CAD for the fifth year in Canada and 70 DEM for the fourth year in Germany), but because of obsolescence.

With q as the fraction of returns an application generates already before grant we are left with seven structural parameters to be estimated: q, μ_{IR}, σ_{IR}, $\phi^A, \sigma_0^A, \delta^A$, δ^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

Analogous to other patent renewal models, the model parameters are identified by the cost structure and the parametrization 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 (where the first maintenance fees are due), 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, higher variance σ_{IR} has the opposite effect.

The parameter q, which represents the fraction of returns which can be realized before grant, is mainly identified by variation in $HR_D(t)$ and $HR_E(t)$. 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 has the opposite effect.

The distribution of the growth rates of returns from pending applications is fully determined by ϕ^A, σ_0^A, and δ^A. The parameters ϕ^A and σ_0^A have a similar impact on the three hazard rates. Higher values of the parameters go along with higher hazard rates of examination requests in all years and lower hazard rates of deferment (except for the third year, where it leads to an increase in HR_D). Different values of ϕ^A and σ_0^A have also an impact on the curve of the hazard rates of expiration. Higher values produce a concave curve such that the hazard rates decline or remain constant for older patents. Lower values produce a convex curve with increasing hazard proportions for older patents. 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 rates of examination requests for the years 2-7, whereas higher values of the latter imply decreasing hazard rates for the same years (and vice versa). δ^A determines from what year on the hazard rates of expiration, $HR_X(t)$, start to exceed the rate of obsolescence. Furthermore, a higher depreciation rate (that is, a lower δ^A) decreases $HR_E(t)$ in all years. It also decreases $HR_D(t)$ for periods 3 to 7, when maintenance fees are due, but increases it in the first two periods. δ^G determines the evolution of returns of already examined patent applications. Therefore δ^G does neither impact $HR_D(t)$ nor $HR_E(t)$, and is identified by variation in $HR_X(t)$ only.

4.4 Structural Parameter Estimates and Model Fit

■ **Structural parameter estimates.** Table 2 presents the results of our structural estimation. 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.[13] We estimated our patent renewal model separately on the Canadian and the German data. Accordingly,

[13] Note that the parametric bootstrap method is computationally very demanding and time-consuming, and due to time constraints in this version of our paper we can only present standard errors for the Canadian data.

Parameter estimates:

	Canadian data:		German data:	
	Point estimate	(s.e.)	Point estimate	(s.e.)
β (fixed)	0.9500	-	0.9500	
θ (fixed)	0.9535	-	0.96066	
μ_{IR}	5.9015	(0.0491)	8.5626	
σ_{IR}	1.8865	(0.0222)	1.3155	
q	0.7307	(0.0032)	0.6647	
ϕ^A	0.9659	(0.0011)	0.8947	
σ_0^A	1.4090	(0.0238)	2.1566	
δ^A	0.8400	(0.0101)	0.5714	
δ^G	0.9363	(0.0026)	0.8944	
Size of Sample	137,427		340,607	

Table 2: **Estimates of the structural model parameters.** The table shows the estimates of the structural parameters of our patent renewal model, both for the Canadian and the German patent application data.

the left part of table 2 shows estimates of the structural model parameters for Canada, and the right part shows estimates for Germany. For each structural parameter, the estimates derived from the Canadian data and those derived from the German data are of the same order of magnitude. We will discuss the specific implications of the our estimates of the structural model parameters in the next section.

■ **Fit of the model.** To get an indication of how well the estimated model fits the data, we compare the simulated hazard rates with the sample hazard rates. The graphs in figure 2 compare simulated and sample hazard rates (averaged over all application cohorts). The graphs on the left compare these hazard rates for the Canadian data, that on the right for the German data. With regard to the hazard proportions of examination requests and deferments ($HR_E(t)$ and $HR_D(t)$), one can see that the simulated hazard proportions fit the sample hazard proportions rather well.

With regard to the hazard rates of expiration, however, our model performs not as well: In general, it underpredicts the hazard rates of expiration - the exceptions being two kinks in the application of our model on the Canadian data and the last three periods in its application on the German data. This might be due to two reasons: First, as the focus of our analysis lies on the early stages of the life of a patent application, in order to make our model tractable and computationally feasible we deliberately chose a simple specification of the late patent stage with relatively few degrees of freedom, which, in turn, renders the model less adaptive in the patent stage. Second, the agents in our model might be more sensitive to the development of the official renewal fees than

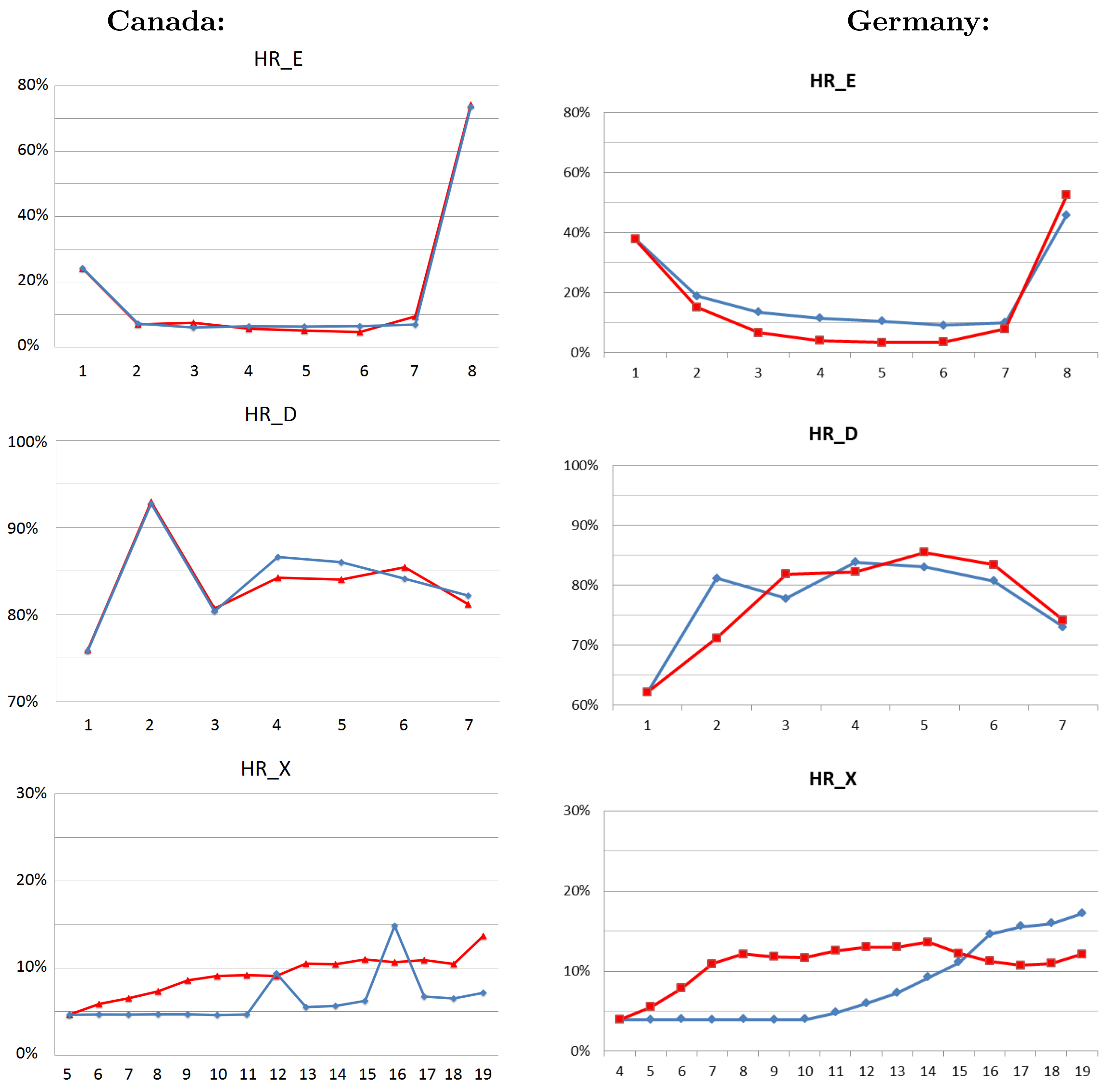

Figure 2: **Model fit: comparison of simulated (blue) and sample (red) hazard proportions.** The graphs compare simulated and sample hazard rates (both for Canadian and German data and averaged over all application cohorts). From top to bottom the graphs compare the development of the hazard proportions of examination requests (HR_E), deferments (HR_D) and expirations (HR_X) over patent age. The red lines show the hazard proportions we observe in the data, the blue lines show the hazard proportions as simulated by our model.

the "real world agents" are (because the official renewal fees might only be a part of actually much higher "real costs" of patent renewal). In this direction points the observation that the simulated hazard rates somehow mirror the respective developments of the renewal fees: For Canada, the large kink in year 16 coincides with the year when the official renewal fees almost double and then stay the same for the following years. For Germany, the absolute increase in renewal fees rises with patent age, which is mirrored in the strong increase in the simulated hazard rates of expiration at the end of patent life. The sample hazard rates of expiration, in contrast, seem to be less sensitive

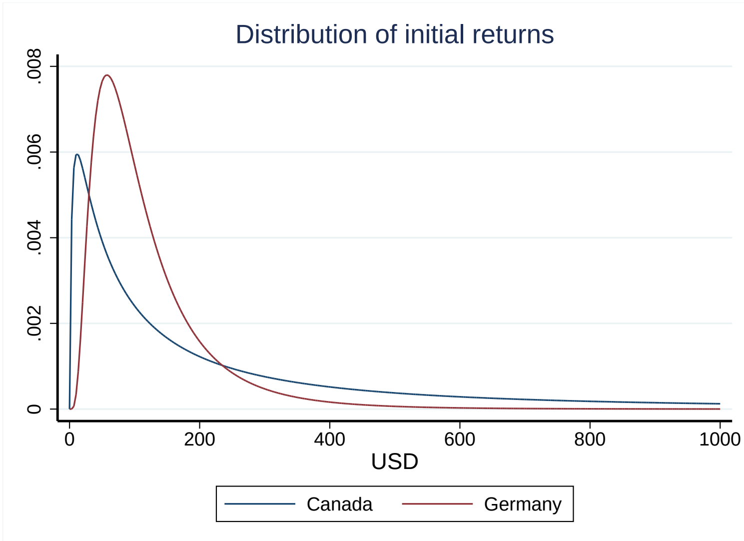

Figure 3: **Distribution of initial returns.** The blue graph displays the distribution of initial returns of Canadian applications, the red graph displays the distribution of initial returns of German applications. In order to make returns comparable, the initial returns of both Canadian and German applications are given in 1989 USD.

to the developments of the fee structures.

5 Implications of the Parameter Estimates and Simulations

5.1 Direct Implications of the Structural Parameter Estimates

We now discuss the implications of our estimates of the structural model model parameters. The parameters μ_{IR} and σ_{IR} determine the distribution of initial returns, which is assumed to follow a log-normal distribution with these parameters as mean and standard deviation. Figure 3 displays the distributions of the initial returns of Canadian and German applications. For Canadian applications, the mean initial return is 546 USD, and the median initial return is 151 USD. For German applications, the mean initial return is 119 USD, and the median initial return is 94 USD. That is, the distribution of the initital value of Canadian applications is more heavily skewed to the right than that of German applications.

The parameters ϕ^A, σ_0^A, and δ^A determine the evolution of returns during the application stage and the examination stage. More concretely, the parameters ϕ^A and σ_0^A determine the probability with which an applicant learns that the value of his application does not decrease by $1 - \delta^A$ but only by an amount smaller than $1 - \delta^A$ or even increases. The left part of figure 4 displays

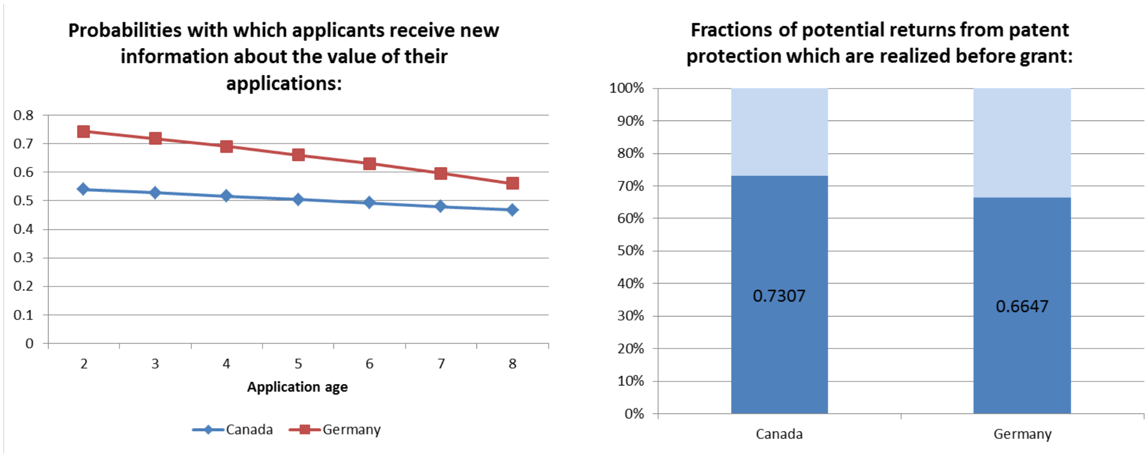

Figure 4: **Learning Probabilities and fractions of potential returns from patent protection which are realized before grant.** For Canadian and for German applications, the left graphs depict the development of the probability with which an applicant learns that the value of his application does not decrease by $1-\delta^A$ but instead decreases by a smaller amount or increases. The bar-graphs on the right-hand side display the fractions of potential returns from patent protection that can be realized by the applicant before his patent application is finally granted.

the development of this probability over the age of an application for Canadian and for German applications. For both Canadian and German applications the probabilities with which applicants receive new information about the values of their applications decline only slowly over application age. Throughout the application stage, the probability to receive new information about application value is significantly higher for German than for Canadian applications. The rate with which the yearly return of Canadian applications decreases given that the applicant does not receive new information about the value of his application is 16.0%, that for German applications is 42.9%. The parameter δ^G determines the evolution of returns during the patent stage. The rate with which the value of a patent depreciates, $1-\delta^G$, is roughly similar for Canadian and German patents. For Canada it is 6.4%, for Germany it is 10.6%

The parameter q determines the fraction of potential returns from patent protection that can already be realized before the patent is finally granted. Both for Canadian and for German applications that fraction is relatively high, being 73.1% for Canadian applications and 66.5% for German applications (compare also the right-hand side of figure 4). That is, although applicants have not yet gained full patent rights, they seem to be able to profit from just having a pending application.

	All Applications:	Applications which were granted:	Applications which were not granted:
Canada:			
50th %ile	1,447	10,425	125
75th %ile	11,061	27,211	1,420
90th %ile	34,522	67,205	5,739
95th %ile	66,301	121,086	11,839
99th %ile	245,937	417,826	47,819
99.9th %ile	1,157,131	1,801,356	271,763
Mean	17,470	34,579	3,086
Germany:			
50th %ile	18,550	50,986	6,970
75th %ile	54,855	121,900	17,829
90th %ile	145,962	259,700	40,354
95th %ile	248,888	407,093	64,883
99th %ile	664,673	1,006,894	166,076
99.9th %ile	2,091,539	3,226,428	503,182
Mean	62,598	117,803	17,679

Table 3: **Simulated patent (application) value distributions.** Based on our estimation results given in table 2, we simulated the value distributions of Canadian and German patent applications. The table shows different percentiles and the means of the simulated value distributions of the sample of all applications, the subsample of applications which were granted later on, and the subsample of applications which were not granted. All values are in 1989 USD.

5.2 Simulation of the Patent Value Distribution

We use the estimated structural parameters to simulate the value distributions of Canadian and German patent applications. For each simulation, we take 250,000 pseudo-random draws from the respective initial distributions and pass them through the model using the estimated parameter values for the respective patent system. We calculate the net value of a patent application as the discounted present value of the stream of returns less the discounted maintenance fees. For the time an application is pending we multiply the return in the respective period by q, and in case of an examination request we subtract the discounted costs incurred for examination.

For both the Canadian and the German patent system, table 3 presents the simulated value distributions for the sample of all patent applications, the subsample of applications which were granted, and the subsample of applications which were not granted. All monetary values are in 1989 USD. According to our simulations, the value of an average German patent application is around three to four times that of an average Canadian patent application. Similar to previous renewal studies, we find that in both patent systems the value distributions are highly skewed: For the sample of all applications, in Canada the median value is 1,447 USD and the mean value is 17,470 USD. In Germany, the median value is 18,550 USD and the mean value is 62,598 USD.

Unsurprisingly, in both patent systems there is a huge difference between the value of applications which were granted and that which were not granted. In Canada, the average value of a granted patent application is 34,579 USD, and the average value of an application which got not granted amounts only to around 9% of that (3,086 USD). In Germany, the average value of a granted patent application is 117,803 USD, and that of an application which got not granted amounts to around 15% of that (17,679 USD).[14]

Additionally, we are able to report what part of the total value of a patent is generated before grant. On average, an owner of a Canadian patent is able to realize 50.4% of overall patent value already during the application and examination stages. An owner of a German patent is able to realize 41.6% of overall patent value already during the application and the examination stages.

5.3 Counterfactual Analysis: Value of the Deferment Option

In this section we use the estimated structural parameters of the Canadian patent system to shed light on the role of the option to defer the examination request. We simulate the value distributions of patent applications for two counterfactual patent systems: one which allows deferment for up to six years, and one for up to five. Comparison of the patent value distributions 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 shows the simulated numbers of examination requests at different application ages for varying lengths of the deferment period (5 years, 6 years and 7 years). In general, the overall number of examination requests increases in case the deferment period is shortened. It increases by 4.13% if we reduce the length of the deferment period by one year, and by 8.19% if we reduce it by two years.

[14]Note that with respect to their order of magnitude our findings are in line with results of previous patent renewal studies for other patent systems. For US patents, Serrano (2006) finds a median value of around 19,000 USD, and a mean value of around 59,000 USD. For EPO patents, Deng (2011) finds a median value of around 36,000 USD, and a mean value of around 505,000 USD. (All values in 1989 USD.)

Age	$L=5$	$L=6$	$L=7$
1	58,914	58,091	57,460
2	14,226	14,007	13,846
3	11,383	11,142	10,986
4	9,464	9,095	8,872
5	8,361	7,835	7,562
6	76,998	7,096	6,715
7		64,959	6,055
8			54,267
$\sum$	179,346	172,225	165,763

Table 4: **Simulated number of examination requests at different application ages for varying lengths of the deferment period.** The table shows the simulated number of examination requests for the years (from the application date) during which examination can be requested. The numbers were simulated for different lengths of the deferment period (5 years, 6 years, 7 years).

This is consistent with the results of Yamauchi and Nagaoka (2008), who analyzed the effects of the shortening of the deferment period in Japan from 7 to only 3 years: While the number of patent applications remained rather stable, the number of requests for patent examinations increased significantly. As with our simulated results, the difference in the number of examination requests was most pronounced 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 examination requests are made in the last year, whereas in the patent system with a maximum deferment period of seven years only 32.74% of all examination requests are made in the last year. The explanation is that applicants are given additional time to evaluate their invention and unveil the uncertainty surrounding it. The additional two years in which deferment is possible permit two types of corrections: First, some applicants who, given a shorter deferment period, would have requested examination without being sure about the value of their application now might learn that their application is actually of low value and thus refrain from the costly examination request. Second, some applicants with high-value applications who, given a shorter deferment period, might not have received a value signal triggering the examination request before the end of the deferment period now might receive a respective signal before the end of the deferment period (and, consequently, request examination before the end of the deferment period).

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

	7 → 5 years	7 → 6 years
All applications		
Change in total value	-5.49%	-2.56%
Applications which were granted		
Change in total value	-4.81%	-2.25%
Change in mean value	-11.99%	-5.93%
Applications which were not granted		
Change in total value	-11.86%	-5.56%
Change in mean value	-5.37%	-2.34%

Table 5: **Changes in the private value of patents in case the deferment period is shortened.** The table shows the simulated changes in the private value of patents in case the deferment period is shortened. We simulated these changes for a decrease of the deferment period from 7 to 5 years, and for a decrease from 7 to 6 years. The table shows the respective changes for the sample of all applications, the subsample of applications which got granted, and the subsample of applications which did not get granted.

the distribution are not affected by the extended deferment period. However, 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 in the fifth (or sixth) year since they have to decide before this information is revealed to them. As presented in table 5, the total 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 on average by 2.34% (respectively, 5.37%) in case the the deferment period is shortened by one (respectively, two) years. 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 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 and German 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 actually gets granted. We estimate that an owner of a patent application which is not yet granted but pending at the Canadian patent office is able to realize 73% of the returns he would generate if he had full patent protection, and an owner of a patent application pending at the German patent office is able to realize 66% of these returns. As a consequence, most of the Canadian and German patent applications which have never been granted have a positive discounted value. With regard to the structure of the process of how patent applicants learn about the value of their applications, we find that whereas for German applications the probability to receive new information about application value during the application stage is higher than for Canadian applications, both for Canadian and German applications the learning process is rather slow, meaning that during the early life of an application the probability that applicants receive new information about the value of their application is on a constantly high level.

Furthermore, our model allows us to simulate the effects a reduction of the maximum deferment period from seven to five years has. 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 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%. 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 their applications would have become worthless shortly after. As a consequence, we estimate a considerable negative impact on the value of unexamined as well as granted patents.

Our results add to the discussion on the optimal institutional organization of patent systems. 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) 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 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 a 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.

F.T.C. (2003). To promote innovation - the proper balance of competition and patent law and policy. Technical report, 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 J. Cantwell, A. Gambardella, and O. Granstrand (Eds.), *Advancing Knowledge and the Knowledge Economy*, Volume Kahin, B. and Foray, D., pp. 331–350. MIT Press.

Harhoff, D. (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.

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). Technical report, 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 Z. Griliches (Ed.), *Patents, R&D, and Productivity*, 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:

$$\begin{aligned}
\widetilde{V}^E(L,r_L) &= qr_L - (c_L + C_{PO}^{Exam} + C_{self}^{Exam} + \beta\theta\pi_{Grnt}C_{PO}^{Grnt}) + \beta\theta\pi_{Grnt}E\left[\widetilde{V_K}(L+1,r_{L+1}) \mid r_L\right], \\
\widetilde{V}^D(L,r_L) &= qr_L - c_L + \beta\theta E\left[\widetilde{V_D}(L+1,r_{L+1}) \mid r_L\right], \\
\widetilde{V}^K(L+1,r_{L+1}) &= r_{L+1} - c_{L+1} + \beta\theta E\left[\widetilde{V_K}(L+2,r_{L+2}) \mid r_{L+1}\right].
\end{aligned}$$

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:

$$\begin{aligned}
\widetilde{V}^E(L,r_L) &- \widetilde{V}^D(L,r_L) + (C_{PO}^{Exam} + C_{self}^{Exam} + \beta\theta\pi_{Grnt}C_{PO}^{Grnt}) = \\
&= \beta\theta\left\{\pi_{Grnt}E\left[\widetilde{V_K}(L+1,r_{L+1}) \mid r_L\right] - E\left[\widetilde{V_D}(L+1,r_{L+1}) \mid r_L\right]\right\} = \\
&= \beta\theta\{\pi_{Grnt}E\left[max\left(\widetilde{V}^K(L+1,r_{L+1}), \widetilde{V}^X(L+1,r_{L+1})\right) \mid r_L\right] \\
&\quad -E\left[max\left[\widetilde{V}^E(L+1,r_{L+1}), \widetilde{V}^W(L+1,r_{L+1})\right] \mid r_L\right]\}, \\
&= \beta\theta\{E\left[max\left(\pi_{Grnt}\widetilde{V}^K(L+1,r_{L+1}),0\right) \mid r_L\right] - E\left[max\left(\widetilde{V}^E(L+1,r_{L+1}),0\right) \mid r_L\right]\}.
\end{aligned}$$

The last equality holds because deferring examination is not possible in $t = L+1$. 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 $r_{L+1} \geq 0$:

$$\begin{aligned}
&\pi_{Grnt}\widetilde{V}^K(L+1,r_{L+1}) - \widetilde{V}^E(L+1,r_{L+1}) = \\
= \; &\pi_{Grnt}r_{L+1} - \pi_{Grnt}c_{L+1} + \pi_{Grnt}\beta\theta E\left[\widetilde{V_K}(L+2,r_{L+2}) \mid r_{L+1}\right] - qr_{L+1} \\
&+(c_{L+1} + C_{PO}^{Exam} + C_{self}^{Exam} + \beta\theta\pi_{Grnt}C_{PO}^{Grnt}) - \beta\theta\pi_{Grnt}E\left[\widetilde{V_K}(L+2,r_{L+2}) \mid r_{L+1}\right] \\
= \; &(\pi_{Grnt} - q)r_{L+1} + C_{PO}^{Exam} + C_{self}^{Exam} + \beta\theta\pi_{Grnt}C_{PO}^{Grnt}) + (1 - \pi_{Grnt})c_{L+1}.
\end{aligned}$$

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

$$\begin{aligned}
\widetilde{V}^E(t,r_t) &= qr_t - (c_t + C_{PO}^{Exam} + C_{self}^{Exam} + \beta\theta\pi_{Grnt}C_{PO}^{Grnt}) + \beta\theta\pi_{Grnt}E\left[\widetilde{V_K}(t+1,r_{t+1}) \mid r_t\right], \\
\widetilde{V}^D(t,r_t) &= qr_t - c_t + \beta\theta E\left[\widetilde{V_D}(t+1,r_{t+1}) \mid r_t\right], \\
\widetilde{V}^K(t+1,r_{t+1}) &= r_{t+1} - c_{t+1} + \beta\theta E\left[\widetilde{V_K}(t+2,r_{t+2}) \mid r_{t+1}\right].
\end{aligned}$$

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:

$$\begin{aligned}
& \widetilde{V}^E(t,r_t) - \widetilde{V}^D(t,r_t) + (C_{PO}^{Exam} + C_{self}^{Exam} + \beta\theta\pi_{Grnt}C_{PO}^{Grnt}) \\
= \; & \beta\theta\{\pi_{Grnt}E\left[\widetilde{V_K}(t+1,r_{t+1}) \mid r_t\right] - E\left[\widetilde{V_D}(t+1,r_{t+1}) \mid r_t\right]\} \\
= \; & \beta\theta\{E\left[max\left(\pi_{Grnt}\widetilde{V}^K(t+1,r_{t+1}),0\right) \mid r_t\right] - E\left[max\left(\widetilde{V}^E(t+1,r_{t+1}),\widetilde{V}^D(t+1,r_{t+1}),0\right) \mid r_t\right]\}.
\end{aligned}$$

Since

$$\begin{aligned}
& \pi_{Grnt}\widetilde{V}^K(t+1,r_{t+1}) - \widetilde{V}^E(t+1,r_{t+1}) = \\
& \quad (\pi_{Grnt} - q)r_{t+1} + C_{PO}^{Exam} + C_{self}^{Exam} + \beta\theta\pi_{Grnt}C_{PO}^{Grnt}) + (1-\pi_{Grnt})c_{t+1}
\end{aligned}$$

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:

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

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

Cut-off value:[15]
$\widetilde{V}_8^E(8,r_8) = 0 \Rightarrow \hat{r}_8^E$ (respectively $\hat{g}_8^E = \frac{\hat{r}_8^E}{r_7}$).

Period 7
If examination is requested:

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

If examination is deferred:

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

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:

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

If examination is deferred:

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

Cut-off value:
$\widetilde{V}_6^D(6,r_6) = 0 \Rightarrow \hat{r}_6^D$;
$\widetilde{V}_6^E(6,r_6) = \widetilde{V}_6^D(6,r_6) \Rightarrow \hat{r}_6^E$.

Period 5
...

[15]Deferment is not possible anymore.