

# Modeling the Duration of Patent Examination at the European Patent Office

Dietmar Harhoff<sup>a,b</sup>, **Stefan Wagner**<sup>a</sup>

<sup>a</sup> Munich School of Management, LMU Munich.

<sup>b</sup> Centre for Economic Policy Research (CEPR), London.

E-mail: [swagner@bwl.uni-muenchen.de](mailto:swagner@bwl.uni-muenchen.de)

This research was supported by the German National Science Foundation (DFG), Sonderforschungsbereich 386 "Statistical Analysis of Discrete Structures".

---

## Contents of the Talk

- Motivation
- Institutional background of patent examination at the European Patent Office (EPO)
- Data and descriptive statistics
- Results from Bayesian multivariate survival analysis
- Summary

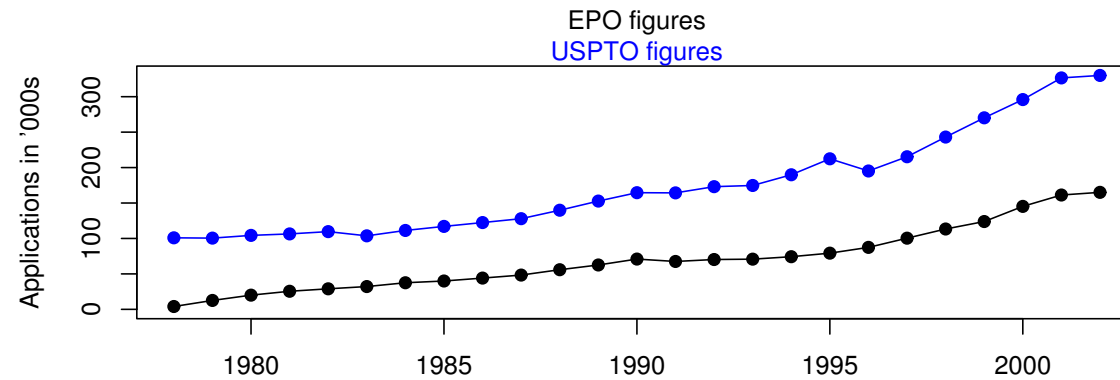
## Motivation

### Recent Developments

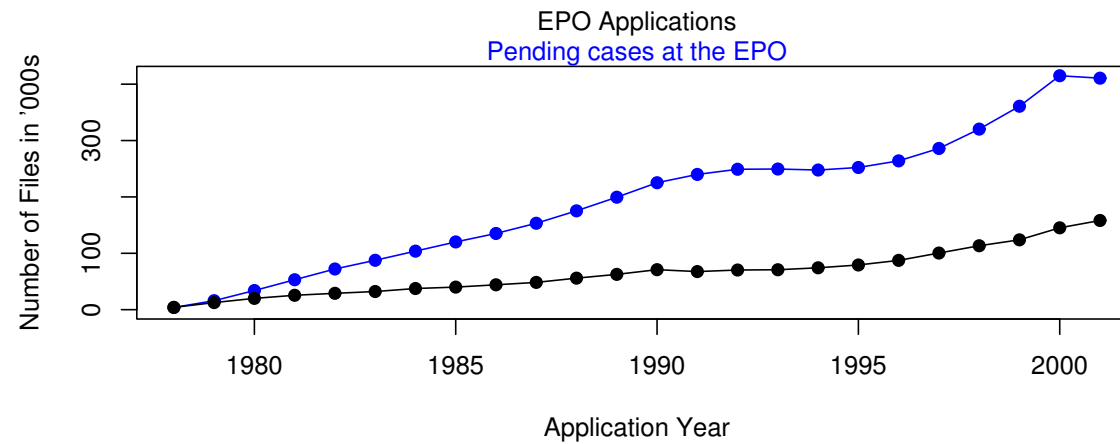
- Increase in workload due to unprecedented rise in patent applications poses new challenge to patent offices worldwide
- Average examination time of granted patents rose from 3.6 years in 1980 to more than 4.5 years during the 90s, number of pending cases exceeded 400,000 in 2000
- EPO created new project *Mastering the Workload* and put it on its high-level agenda
- First empirical analysis of decision time at the USPTO in Popp, Juhl, Johnson (2003), theoretical and empirical treatment in Régibeau, Rockett (2003)

# Motivation

## Application Figures



## Pending Cases at the EPO



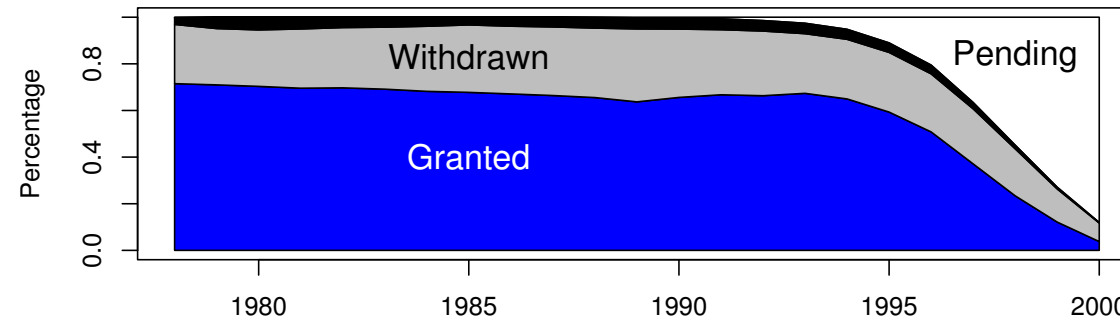
## Institutional Background

### Examination Process and Outcomes

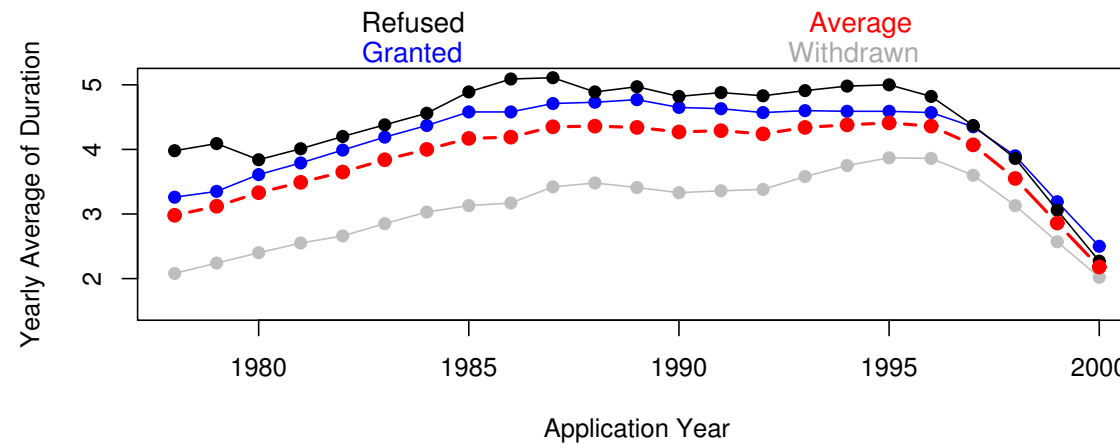
- After the filing of an application, The Hague office generates a search report describing all prior art relevant for the patentability of the invention
- Within six months after publication of the search report, the applicant has to request examination of his application, otherwise patent is deemed to be *withdrawn*
- After examination, EPO informs the applicant about patent *grant*, patent *refusal* or possibly requires the applicant to agree to changes in the claims
- Applications with no final decision available are still *pending*

# Institutional Background

## Distribution of Outcomes



## Average Decision Lag by Outcomes



## Determinants of Decision Making Lags

- Workload at the EPO**
- Workload defined as number of pending cases per examiner
  - Increase in workload should lead to longer examination times c.p.
- Complexity of Applications**
- Complexity of a patent can be operationalised by # of claims, # of references and # of IPC-classes
  - Growing complexity should lead to longer lags
- Statutory Provisions**
- PCT applications allow applicants to delay the completion of the examination procedure
  - Institutions like request for accelerated examination should decrease duration of examination

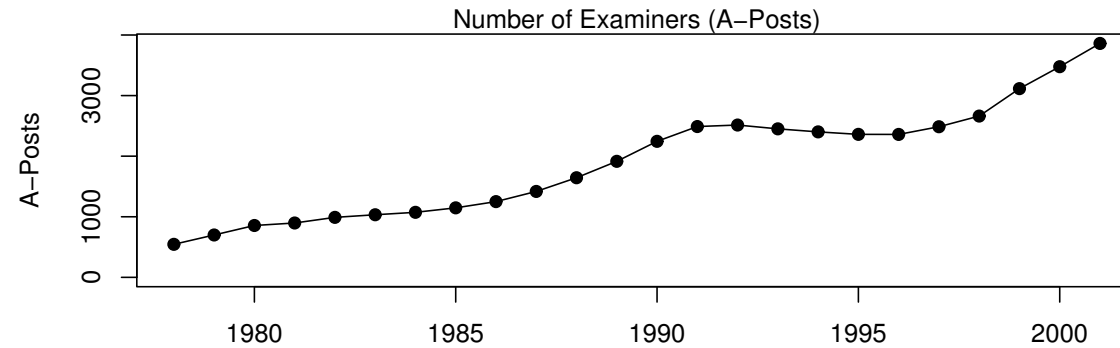
## Dataset

### Data Source

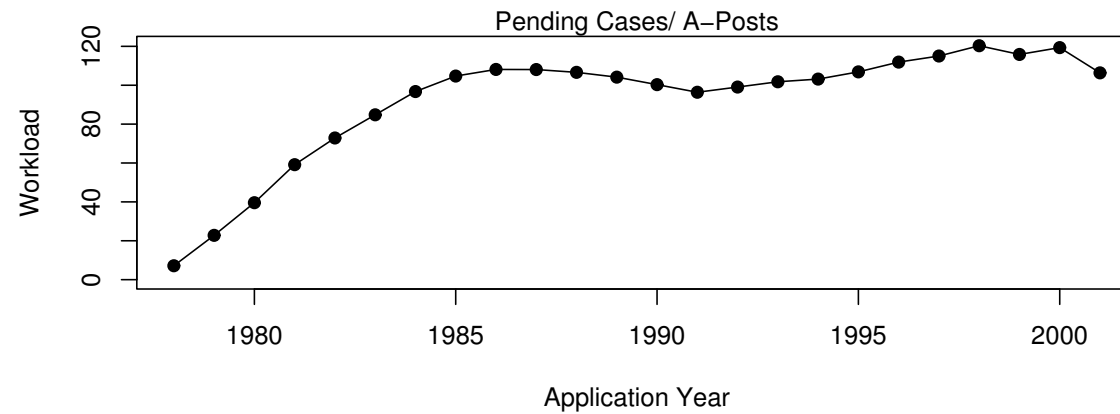
- Dataset is an image of the patent data provided by the EPO at [www.epoline.org](http://www.epoline.org) as of March 2003
- Data covers 1,266,506 patents with applications dates ranging from 1.6.1978 to 25.7.2002
- Data contains bibliographic information as well as procedural information
- Additional information was obtained from EPASYS directory excerpts and from OECD

## Descriptive Statistics – Patent Office

### Number of Examiners

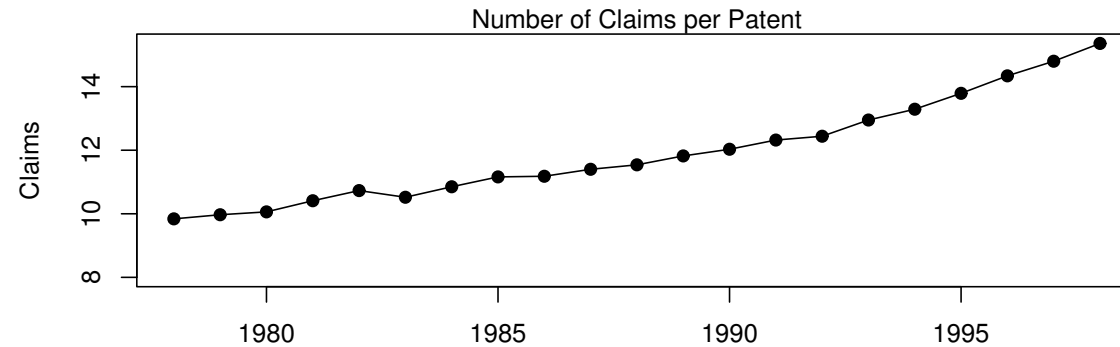


### Workload

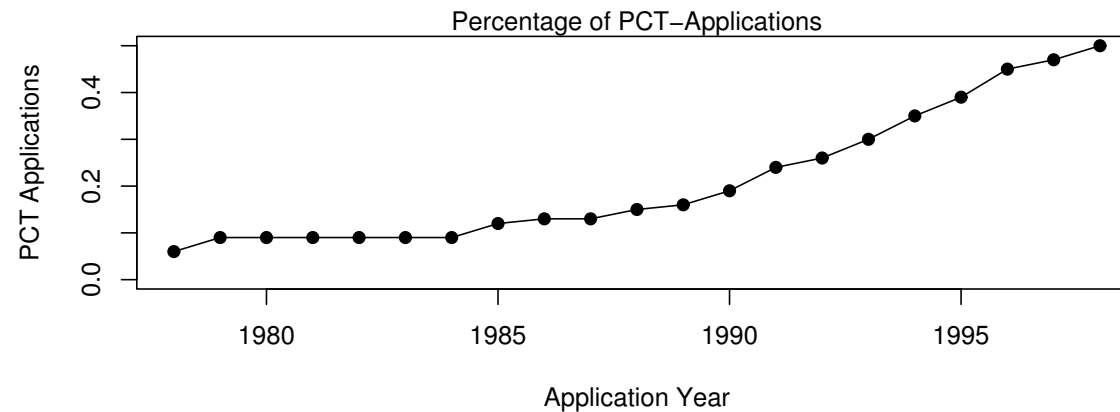


## Descriptive Statistics – Application Characteristics

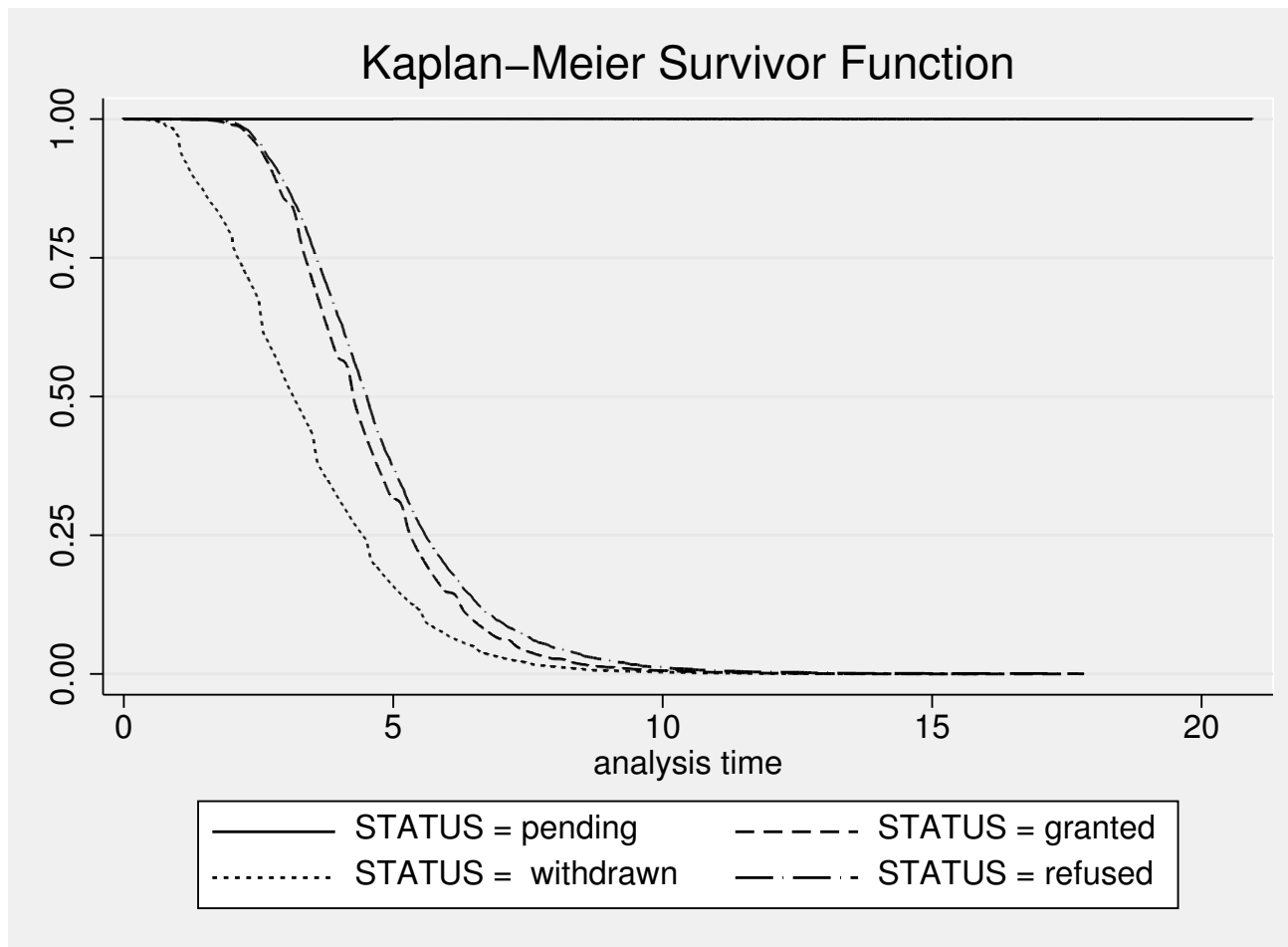
**Claims per Patent**



**Share of PCT-Applications**



## Descriptive Statistics – Kaplan-Meier Estimates



## Survival Analysis – Piecewise Exponential Model

### Piecewise Exponential Model (PEM)

- Survival model based on hazard rate  $\lambda(t)$  (instantaneous failure rate at time  $t$  given that the individual survives until  $t$ )
- $$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t}$$
- PEM divides time axis  $t$  in  $k$  equidistant intervals and all values that depend on  $t$  are assumed to be constant within these intervals
- $\lambda(t, x) = \lambda_s \cdot \exp(\eta)$  where  $t \in (a_{s-1}, a_s]$ ,  $s = 1, \dots, k - 1$  and  $\eta = f(x, \beta)$ .

## Survival Analysis – Specification Issues

### Model Specification

- All independent variables are treated as time-invariant
- Competing risk models are estimated in order to characterize differences in the processes leading to either a withdrawal of the application by the applicant, a refusal or a grant of the patent by the EPO
- Purely parametric specifications with  $\eta = X'\beta$  are compared to semi-parametric specifications with  $\eta = \sum f_i(x_i) + X'_{-i}\beta_{-i}$

## Survival Analysis – Estimation Issues

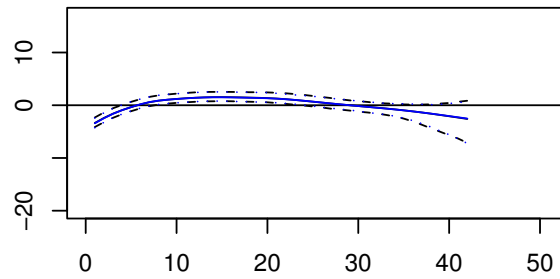
### Bayesian Estimation Framework

- Estimation is carried out in a fully Bayesian framework employing Markov-Chain-Monte-Carlo (MCMC) simulation techniques using *BayesX*
- Non-linear effects are estimated using a Bayesian version of P-Spline regression as introduced by Lang & Brezger (2002)

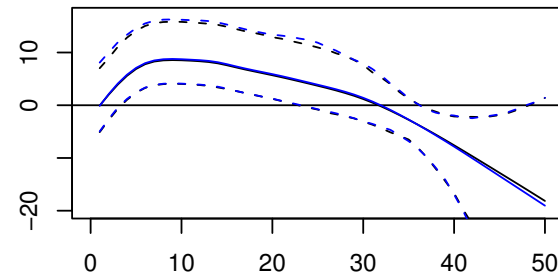
### Sampling Issues

- Multivariate analysis restricted to patents with application dates prior to February 1998
- Due to computational intensity of the applied simulation approach, 0.5% sub-sample (4,608 patents) is drawn from the population

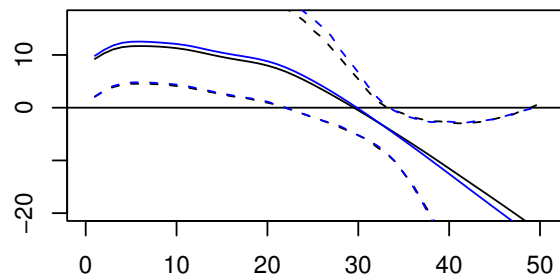
## Results – Hazard Rate Estimates



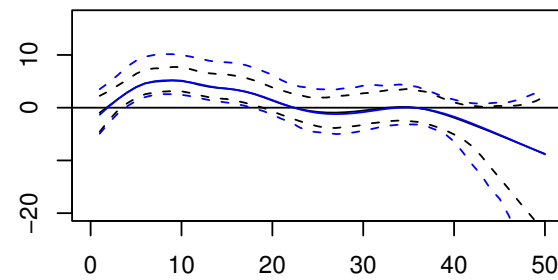
(a) Pooled Risks



(b) Granted Patents



(c) Application withdrawn



(d) Refused

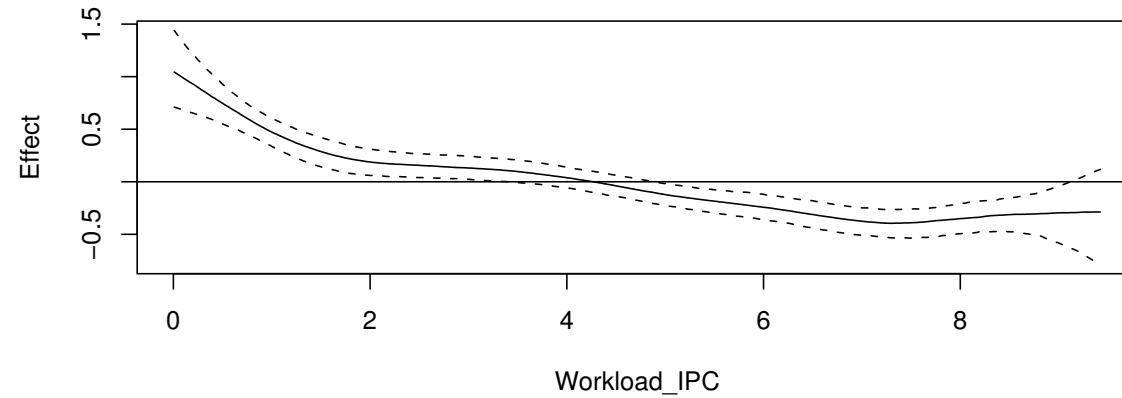
## Results – Parameter Estimates

Outcomes	Pooled Outcomes		Granted		Refused		Withdrawn	
	Parametric	Semi-Parametric	Parametric	Semi-Parametric	Parametric	Semi-Parametric	Parametric	Semi-Parametric
INTERCEPT	-1.2097*	-2.7667*	-9.5317*	-9.9423*	-9.737*	-12.7693*	-14.7382*	-17.8812*
BWD TOT	-0.0020	-0.0004	-0.0250*	-0.0238	0.0170	0.0135	-0.0292*	-0.0285*
BWD_NP_TOT	-0.087*	-0.0856*	-0.0032	-0.0018	-0.0140	-0.0076	-0.0838*	-0.0839*
SHARE X	-0.3961*	-0.3815*	-0.7370*	-0.7533*	0.0913	0.1146	0.5814*	0.5950*
SHARE A	0.2469*	-0.2167*	0.0684	0.0475	-0.0304	-0.018	-0.3231*	-0.3048*
ORIGINALITY	-0.1165	-0.0893	0.0746	0.0589	-1.0165	-0.9540	-0.1307	-0.0971
GENERALITY	0.0860	0.0569	0.3447*	0.2005*	-0.1074	-0.1589	-0.2594	-0.2289
PCTAPPL	-0.3583*	-0.3464*	-0.0142	-0.0472	0.3610	0.4826*	-0.4456*	-0.4028*
<b>RACCEXM</b>	<b>0.0973</b>	<b>0.1055</b>	<b>0.8313*</b>	<b>0.8540*</b>	-1.736	-1.6441	<b>-1.1370*</b>	<b>-1.1392*</b>
<b>IPC_TOT</b>	<b>-0.0413*</b>	<b>-0.0429*</b>	<b>-0.0420*</b>	<b>-0.0409*</b>	-0.0669	-0.0585	<b>0.0293</b>	<b>0.0265</b>
<b>WORKLOAD</b>	<b>-0.1247*</b>		<b>-0.0895*</b>		<b>-0.0105</b>		<b>0.0073</b>	
FWD 5YRS	-0.0193		0.0820*		-0.0906		-0.2409*	
DESG	-0.0048		-0.0115*		-0.0019		0.0148*	
<b>CLAIMS</b>	<b>-0.0131*</b>		<b>-0.0104*</b>		-0.0158		0.0042	
DIC	31,295.7	31,248.8	24,304.3	24,297.2	2,974.0	2,971.7	14,951.9	14,934.0
Observations	4,608		4,608		4,608		4,608	
Exits	4,287		2,832		194		1,261	

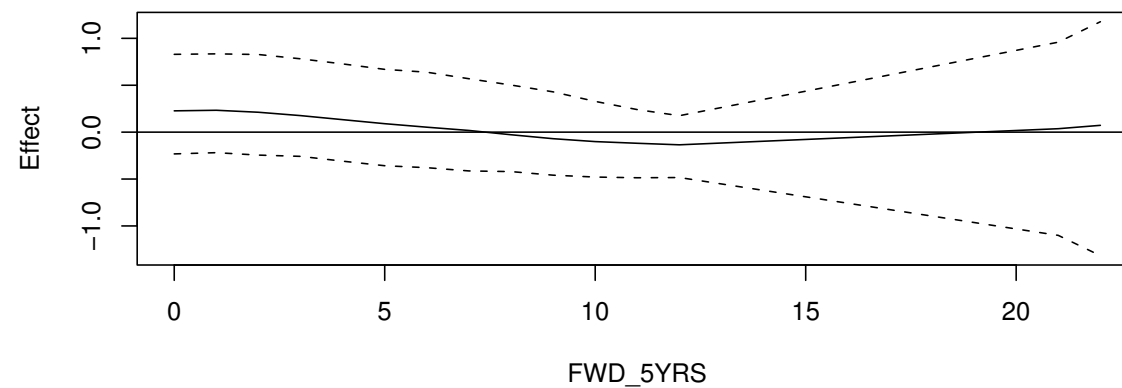
\* 95% credible region excludes 0.

## Results – Non-linear Estimates

### Workload

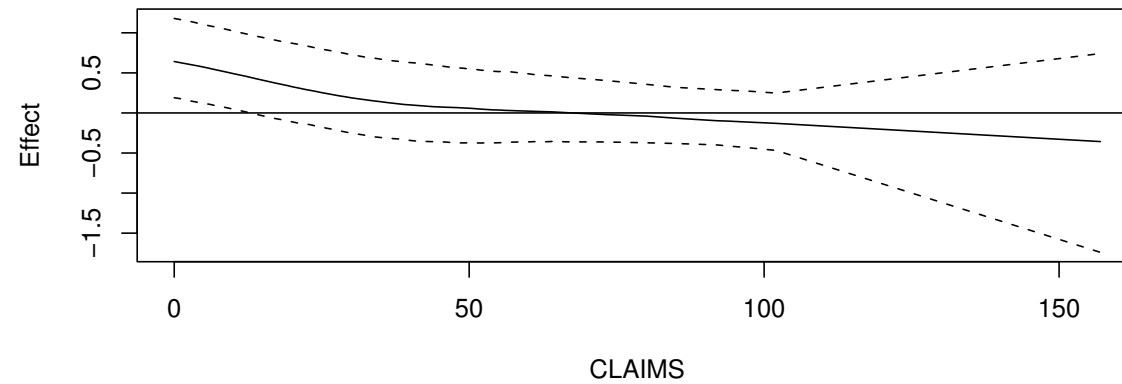


### Citations received within five Years

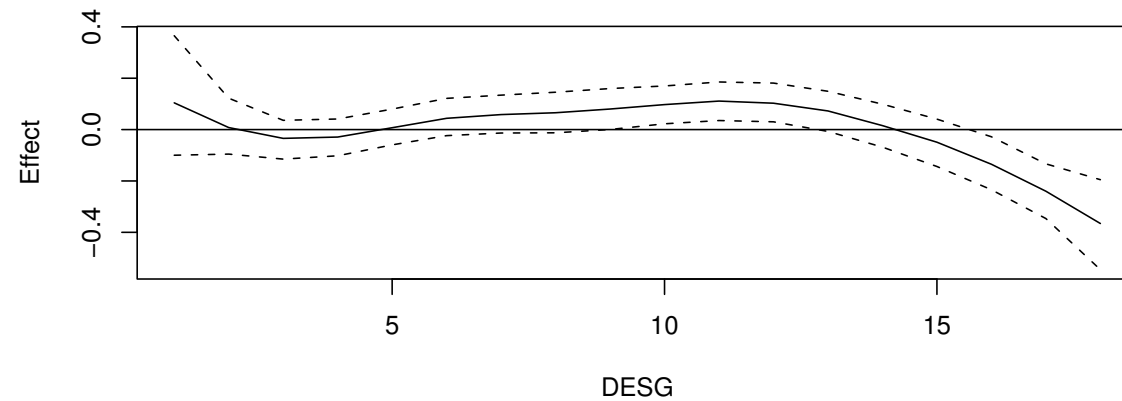


## Results – Non-linear Estimates

**Number of Claims**



**Number of Designated States**



---

## Summary

- Empirical findings are in line with theoretical assumptions:
  - Decisions on more complex patents require more time than decisions on average patents
  - Workload is significant determinant of decision lags
  - Findings from competing risk specifications reflect endogenous behaviour of applicants
- Moderate non-linearities in some independent variables detected

Thank You!