EMPIRICAL ANALYSIS OF INNOVATION IN OLIGOPOLY INDUSTRIES

Lecture 3: Measuring the value of patents

CEMFI SUMMER SCHOOL – 2018 Victor Aguirregabiria (University of Toronto)

September 5, 2018

Measuring the value of patents: Outline

- 1. Introduction
- 2. Dynamic discrete choice structural models
- 3. Patent renewal models: Pakes (1986)
- 4. Trade of patents: Serrano (2018)

1. Introduction

Measuring the value of patents

- **What is the value of a patent? How to measure it?**
- The valuation of patents is important for:
  - merger & acquisition;
  - value of innovations;
  - value of patent protection;
  - using patents as collateral for loans; ...
- Different approaches to measure the value of patents, with their respective merits and limitations.

Measuring the value of patents [2]

- **[1] Hedonic approach.** Using transaction prices of traded patents together with and hedonic price regression.
  - Very few patents are traded, and there is substantial selection.
- **[2] Production function approach.** Include patents as inputs in the estimation of a PF.
  - Challenging because firms use hundreds of patents, and a patent can be used for multiple products.
- **[3] Number of citations.**
  - Very imperfect measure of the value of a patent.
- **[4] Include a firm's patents in demand of differentiated products.** Hashmi & Van Biesebroeck (RStat, 2016).

Measuring the value of patents [3]

- **[5] Case event studies and stock prices.** For public companies in the stock market, we can look at the effect on stock prices of events such as the approval or the non-renewal of a patent.
  - Limited to public companies.

- **[6] Revealed preference approach using patent renewal decisions.** Patent holders need to pay renewal fees to keep a patent alife. The observed decision of renewing or not reveals information about the firm valuation of the patent.
  - This idea can be extended to other decisions about a patent: selling, licensing, litigating.

2. Dynamic Discrete Choice Structural Models

Dynamic DIscrete Choice Structural Models

- Dynamic Discrete Choice Structural models (DDCS) have been developed and applied to study multiple topics in empirical IO:
  - Firms' investment and inventory decisions
  - Demand of durable and storable goods
  - Pricing with menu costs
  - Market entry-exit
  - Patent renwal and trade decisions, etc
- I provide here an introduction to the specification and estimation of this type of models.

Model: Choices and payoff function

- Time is discrete and indexed by $t \in \{1, 2, ..., T\}$, where T can be finite or infinite.
- And agent i makes a discrete choice $a_{it} \in \{0, 1, ..., J\}$ every period, e.g., invest or not; renew a patent or not; purchasing a variety of a differentiated products.
- This decision has implications of her current payoffs but also on future payoffs.
- The current payoff function is:

$$\Pi_{it} = \pi(a_{it}, x_{it}) + \varepsilon_{it}(a_{it})$$

with $\varepsilon_{it} = \{\varepsilon_{it}(0), \varepsilon_{it}(1), ..., \varepsilon_{it}(J)\}$.

State variables

- x_{it} and ε_{it} are vectors of state variables which are known to the agent. x_{it} is observable to the researcher and ε_{it} is unobservable.
- The agent makes the decision a_{it} to maximize her expected value:

$$E_t \left(\sum_{j=0}^{T} \beta^j \, \Pi_{i,t+j} \right)$$

  where $\beta \in (0, 1)$ is the discount factor.
- The agent knows the current state $(x_{it}, \varepsilon_{it})$ but has uncentainty about future values of these state variables.
- She knows tha stochastic process of the state variables. A Markov process:

$$p(x_{it+1}, \varepsilon_{it+1} \mid a_{it}, \; x_{it}, \; \varepsilon_{it})$$

State variable (2)

- A common structure for $p(x_{it+1}, \varepsilon_{it+1} \mid a_{it}, x_{it}, \varepsilon_{it})$ in empirical application is:

$$p(x_{it+1}, \varepsilon_{it+1} \mid a_{it}, \ x_{it}, \ \varepsilon_{it}) = f(x_{it+1} \mid a_{it}, \ x_{it}) \ g(\varepsilon_{it+1})$$

- This implies two assumptions:
  - Conditional independence: given (a_{it}, x_{it}), next period x_{it+1} does not depend on ε_{it}
  - And ε_{it} is i.i.d. over time.

- The first assumption appears naturally in many applications. The i.i.d. assumption is an important restriction in many cases.

- Without $i.i.d.$ ε's we have:

$$p(x_{it+1}, \varepsilon_{it+1} \mid a_{it}, \ x_{it}, \ \varepsilon_{it}) = f(x_{it+1} \mid a_{it}, \ x_{it}) \ g(\varepsilon_{it+1} \mid \varepsilon_{it})$$

Structural parameters / functions

- The primitives of the model are:
  - The payoff function $\pi(.)$
  - The transition probability of observable state variables: $f(.)$
  - The density function of the unobservables: $g(.)$
  - The discount factor β
- In parametric model: these functions are characterized by a finite vector of parameters θ.
- The objective of the researcher is estimating θ using the model and data on agents' actions and states $\{a_{it}, x_{it}\}$.

Solution of the model

- Let $V(x_{it}, \varepsilon_{it})$ be the value function. The **Bellman equation**:

$$V(x_{it}, \varepsilon_{it}) = \max_{a \in \{0,1,...,J\}} \left\{ v(a, x_{it}) + \varepsilon_{it}(a) \right\}$$

where $v(a, x_{it})$ is the **choice-specific value function**:

$$\begin{aligned} v(a, x_{it}) &= \pi(a.x_{it}) \\ &+ \beta \int V(x_{it+1}, \varepsilon_{it+1}) \; p(x_{it+1}, \varepsilon_{it+1} | a, x_{it}, \varepsilon_{it}) \; dx_{it+1} \; d\varepsilon_{it+1} \end{aligned}$$

- The optimal decision rule is:

$$a_{it} = \alpha(x_{it}, \varepsilon_{it}) = \arg \max_{a \in \{0,1,...,J\}} \left\{ v(a, x_{it}) + \varepsilon_{it}(a) \right\}$$

Solution (2)

- The assumption that ε_{it} is i.i.d. implies that we can reduce the dimensionality of the DP problem.
- We can describe the solution of the problem in terms of the *integrated value function*:

$$V_{\sigma}(x_{it}) \equiv \int V(x_{it}, \varepsilon_{it})\ g(\varepsilon_{it})\ d\varepsilon_{it}$$

- This value function is the unique fixed point of the integrated Bellman equation:

$$V_{\sigma}(x_{it}) = \int \max_{a \in \{0,1,\ldots,J\}} \left\{\ v(a, x_{it}) + \varepsilon_{it}(a)\ \right\}\ g(\varepsilon_{it})\ d\varepsilon_{it}$$

where

$$v(a, x_{it}) \quad = \quad \pi(a.x_{it}) + \beta \int V_{\sigma}(x_{it+1})\ f(x_{it+1}|a, x_{it})\ dx_{it+1}$$

- As shown by Rust (1987, 1994), this integrated Bellman equation is a contraction mapping.

Solution: Logit case (3)

- The assumption of i.i.d. extreme value on ε_{it} implies a closet-form expression for the integral $\int \max\{.;.\}\ g(\varepsilon_{it})\ d\varepsilon_{it}$.

$$
\begin{aligned}
V_{\sigma}(x_{it}) &= \ln\left(\sum_{a=0}^{J} \exp\{v(a, x_{it})\}\right) \\
&= \ln\left(\sum_{a=0}^{J} \exp\left\{\pi(a.x_{it}) + \beta \int V_{\sigma}(x_{it+1})\ f(x_{it+1}|a, x_{it}) dx_{it+1}\right\}\right)
\end{aligned}
$$

Solution (4)

- If x_{it} has discrete and finite support (or we discretize it), we can represent the value function $V_\sigma(.)$ as a vector $\mathbf{V}_\sigma$ in the Euclidean space of dimension M, i.e., the number of points in the space of x_{it}.

- The integrated Bellman equation in matrix form is:

$$\mathbf{V}_\sigma = \Gamma(\mathbf{V}_\sigma) = \ln \left(\sum_{a=0}^{J} \exp \left\{ \mathbf{\Pi}(a) + \beta \ \mathbf{F}(a) \ \mathbf{V}_\sigma \right\} \right)$$

- For every choice alternative a, $\mathbf{\Pi}(a)$ is a $M \times 1$ vector with the payoffs $\pi(a.x_{it})$.

- For every choice alternative a, $\mathbf{F}(a)$ is a $M \times M$ matrix with the transition probabilities $f(x_{it+1}|a, x_{it})$.

Solution: value function iteration algorithm

- Given the integrated Bellman equation in matrix form, and given that it is a contraction, we can obtain the solution vector $\mathbf{V}_\sigma$ by iterating until convergence in this mapping.

- Let $\mathbf{V}_\sigma^0$ be an arbitrary initial value for the vector $\mathbf{V}_\sigma$.Then, at iteration $k \geq 1$ we obtain:

$$\mathbf{V}_\sigma^{k+1} = \Gamma(\mathbf{V}_\sigma^k) = \ln\left(\sum_{a=0}^{J} \exp\left\{\mathbf{\Pi}(a) + \beta\ \mathbf{F}(a)\ \mathbf{V}_\sigma^k\right\}\right)$$

- Since the Bellman equation is a contraction mapping, this algorithm always converges (regardless the initial $\mathbf{V}_\sigma^0$) and it converges to the unique fixed point.

Data and Estimation

- Suppose that the researcher has panel data on $\{a_{it}, x_{it}\}$ for N agents over T periods. Typically, N is large and T is small.
- The log-likelihood function of this model ad data has the following structure:

$$
\begin{aligned}
l(\theta) &= \sum_{i=1}^{N}\sum_{t=1}^{T} \ln P(a_{it}|x_{it},\theta) + \sum_{i=1}^{N}\sum_{t=1}^{T-1} \ln f(x_{it+1} \mid a_{it}, x_{it}, \theta_f) \\
&= l^{(1)}(\theta_\pi, \theta_f) + l^{(2)}(\theta_f)
\end{aligned}
$$

- Given this structure, it is convenient to estmate the structural parameters in two-steps.
  - Step 1: Estimate θ_f for the transitions likelihood $l^{(2)}(\theta_f)$. It does not require solving the model.
  - Step 2: Estimate the parameters in the payoff θ_π from the choice likelihood $l^{(1)}(\theta_\pi, \widehat{\theta}_f)$.

Nested Fixed Point Algorithm (NFXP)

- The NFXP algorithm is a gradient iterative search method to obtain the MLE of the structural parameters.
- This algorithm nests a BHHH method (outer algorithm), that searches for a root of the likelihood equations, with a value function or policy iteration method (inner algorithm), that solves the DP problem for each trial value of the structural parameters.

NFXP algorithm (2)

- The algorithm is initialized with an arbitrary vector $\hat{\boldsymbol{\theta}}_0$.
- A BHHH iteration is defined as:

$$\hat{\boldsymbol{\theta}}_{k+1} = \hat{\boldsymbol{\theta}}_k + \left(\sum_{i=1}^{N} \nabla l_i(\hat{\boldsymbol{\theta}}_k) \nabla l_i(\hat{\boldsymbol{\theta}}_k)' \right)^{-1} \left(\sum_{i=1}^{N} \nabla l_i(\hat{\boldsymbol{\theta}}_k) \right)$$

where $\nabla l_i(\boldsymbol{\theta})$ is the gradient in $\boldsymbol{\theta}$ of the log-likelihood function for individual i.

$$\nabla l_i(\boldsymbol{\theta}) = \sum_{t=1}^{T_i} \nabla \log P(a_{it} | x_{it}, \boldsymbol{\theta})$$

- To obtain this score we have to solve the DP problem. We use the Value function iterations algorithm described above.

Other estimation methods

- The main computational burden in the estimation of these models comes from the repeated solution of the dynamic programming problem. This cost increases exponentially with the dimension of the state space M and it becomes intractable rapidly.
- Different methods have been proposed to deal with this issue:
  - Hotz-Miller Conditional Choice Probailities (CCP) estimators.
  - Monte Carlo forward simulation to approximate value functions.
  - Finite dependence / Euler equations methods

3. Patent Renewal Models

Patent Renewal Models

- Pakes (1986) proposes using information on patent renewal fees together with a *Reveal Preference approach* to estimate the value of a patent.
- For the patent system studied in this paper, a patent holder should pay, every year, a renewal fee to keep her patent.
- If the patent holder decides to renew, it is because her expected value of holding the patent is greater than the renewal fee (that is publicly known).
- Therefore, observed decisions on patent renewal contain information on the value of a patent.

Patent Renewal Models (2)

- This give us only a lower or upper bound to the value of a patent in a given period of time.
- The value of a patent varies over time for multiple reaons:
  - Time to expiration
  - Technological obsolescence (new better patents)
  - New patents ideas that are complent
  - Changes in demand for the products of the patent.
- These models specify a distribution (stochastic process) for the value of patents that tries to capture these effects.
- The estimated model provides an estimate of this distribution, though not of the values of single patents.
- Using this distribution and the ages of the patents of a firm, we can obtain an estimate of the average value of the pool of patents.

Pakes (1986): Model

- Consider a patent holder who has to decide whether to renew her patent or not. We index patents by i.
- This decision should be taken at ages $t = 1, 2, ..., T$ where $T < \infty$ is the regulated term of a patent (e.g., 20 years in US, Europe, or Canada).
- Patent regulation also establishes a sequence of **Renewal Fees** $\{c_t : t = 1, 2, ..., T\}$. This sequence of renewal fees is deterministic such that a patent owner knows with certainty future renewal fees.
- The schedule $\{c_t : t = 1, 2, ..., T\}$ is typically increasing in patent age t. It may go from a few hundred dollars to a few thousand dollars.

Pakes (1986): Model [2]

- A patent generates a sequence of profits $\{\pi_{it} : t = 1, 2, ..., T\}$.
- At age t, a patent holder knows current profit π_{it} but has uncertainty about future profits $\pi_{i,t+1}$, $\pi_{i,t+2}$, ...
- The evolution of profits depends on the following factors:

(1) the initial "quality" of the idea/patent;
(2) innovations (new patents) which are substitutes of the patent and therefore, depreciate its value or even make it obsolete;
(3) innovations (new patents) which are complements of the patent and therefore, increase its value.

Pakes (1986): Stochastic process of patent profits

- Pakes proposes a stochastic process that tries to capture the three forces mentioned above.

- A patent profit at the first period is a random draw from a log-normal distribution with parameters μ_1 and σ_1:

$$\ln(\pi_{i1}) \sim N(\mu_1, \sigma_1^2)$$

- After the first year, profit evolves according to the following formula:

$$\pi_{i,t+1} = \tau_{i,t+1} \ \max\left\{ \delta \ \pi_{it} \ ; \ \xi_{i,t+1} \right\}$$

- $\delta \in (0,1)$ is the depreciation rate. In the absence of unexpected shocks, the value of the patent depreciates according to the rule: $\pi_{i,t+1} = \delta \ \pi_{it}$.

Stochastic process of patent profits [2]

$$\pi_{i,t+1} = \tau_{i,t+1} \ \max\left\{ \ \delta \ \pi_{it} \ ; \ \xi_{i,t+1} \right\}$$

- $\tau_{i,t+1} \in \{0,1\}$ is a binary variable that represents that the patent becomes obsolete (i.e., zero value) due to competing innovations.
- The probability of this event is a decreasing function of profit at previous year:

$$\Pr(\tau_{i,t+1} = 0 \mid \pi_{it}, t) = \exp\{-\lambda \ \pi_{it}\}$$

- The largest is the profit of the patent at age t, the smallest is the probability that it becomes obsolete.

Stochastic process of patent profits [3]

$$\pi_{i,t+1} = \tau_{i,t+1} \ \max \left\{ \ \delta \ \pi_{it} \ ; \ \xi_{i,t+1} \right\}$$

- Variable $\xi_{i,t+1}$ represents innovations which are complements of the patent and increase its profitability.
- $\xi_{i,t+1}$ has an exponential distribution with mean γ and standard deviation $\phi^t \sigma$:

$$p(\xi_{i,t+1} \mid \pi_{it}, t) = \frac{1}{\phi^t \sigma} \exp \left\{ - \frac{\gamma + \xi_{i,t+1}}{\phi^t \sigma} \right\}$$

- If $\phi < 1$, the variance of $\xi_{i,t+1}$ declines over time (and the $E(\max \left\{ \ x \ ; \ \xi_{i,t+1} \right\})$ value declines as well).
- If $\phi > 1$, the variance of $\xi_{i,t+1}$ increases over time (and the $E(\max \left\{ \ x \ ; \ \xi_{i,t+1} \right\})$ value increases as well).

Stochastic process of patent profits [4]

- Under this specification, profits $\{\pi_{it}\}$ follow a non-homogeneous Markov process with initial density $\pi_{i1} \sim \ln N(\mu_1, \sigma_1^2)$, and transition density function:

$$f_{\varepsilon}\left(\pi_{it+1} | \pi_{it}, t\right) = \begin{cases} \exp\{-\lambda \ \pi_{it}\} & \text{if} \quad \pi_{it+1} = 0 \\ \\ \Pr\left(\xi_{it+1} < \delta\pi_{it} \mid \pi_{it}, t\right) & \text{if} \quad \pi_{it+1} = \delta\pi_{it} \\ \\ \dfrac{1}{\phi^t \sigma} \exp\left\{-\dfrac{\gamma + \pi_{it+1}}{\phi^t \sigma}\right\} & \text{if} \quad \pi_{it+1} > \delta\pi_{it} \end{cases}$$

- The vector of structural parameters is $\theta = (\lambda, \delta, \gamma, \phi, \sigma, \mu_1, \sigma_1)$.

Dynamic Decision Model

- Let $a_{it} \in \{0, 1\}$ be the decision variable that represents the event "the patent owner decides to renew the patent at age t".
- The value of not renewal ($a_{it} = 0$) is zero. The value of renewal ($a_{it} = 1$) is the current profit $\pi_{it} - c_t$ plus the expected and discounted future value.
- $V_t(\pi)$ is the value of an active patent of age t and current profit π.
- The value function is implicitly defined by the Bellman equation:

$$V_t(\pi_{it}) = \max \left\{ 0 \ ; \ \pi_{it} - c_t + \beta \int V_{t+1}(\pi_{i,t+1}) \ f_\varepsilon(d\pi_{i,t+1} \mid \pi_{it}, t) \right\}$$

with $V_t(\pi_{it}) = 0$ for any $t \geq T + 1$.

Solution (Backwards induction)

- We can use backwards induction to solve for the sequence of value functions $\{V_t\}$ and optimal decision rules $\{\alpha_t\}$
- Starting at age $t = T$, for any profit π:

$$V_T(\pi) = \max\{\ 0\ ;\ \pi - c_T\}$$

and

$$\alpha_T(\pi) = 1\{\ \pi - c_T \geq 0\ \}$$

- Then, for age $t < T$, and for any profit π:

$$V_t(\pi) = \max\left\{\ 0\ ;\ \pi - c_t + \beta \int V_{t+1}(\pi')\ f_\varepsilon(d\pi'|\pi, t)\ \right\}$$

and

$$\alpha_t(\pi) = 1\left\{\ \pi - c_t + \beta \int V_{t+1}(\pi')\ f_\varepsilon(d\pi'|\pi, t) \geq 0\ \right\}$$

Solution - A useful result

- Given the form of $f_{\varepsilon}(\pi'|\pi, t)$, the future and discounted expected value, $\beta \int V_{t+1}(\pi')\ f_{\varepsilon}(d\pi'|\pi, t)$, is increasing in current π.
- This implies that the solution of the DP problem can be described as a **sequence of threshold values for profits** $\{\pi_t^* : t = 1, 2, ..., T\}$ such that the optimal decision rule is:

$$\alpha_t(\pi) = 1\{\ \pi \geq \pi_t^*\ \}$$

- π_t^* is the level of current profits that leaves the owner indifferent between renewing the patent or not: $V_t(\pi_t^*) = 0$.

Solution - A useful result [2]

- These threshold values are obtained using backwards induction.
- At period $t = T$:

$$\pi_T^* = c_T$$

- At period $t < T$, π_t^* is the unique solution to the equation:

$$\pi_t^* - c_t + E\left(\sum_{s=t+1}^{T} \beta^{s-t} \ \max\{\ 0\ ;\ \pi_{t+1} - \pi_{t+1}^*\ \} \mid \pi_t = \pi_t^*\right) = 0$$

- Solving for a sequence of threshold values is much simpler that solving for a sequence of value functions.

Data

- Sample of N patents with complete (uncensored) durations $\{d_i : i = 1, 2, ...N\}$, where $d_i \in \{1, 2, ..., T + 1\}$ is patent i's duration or age at its last renewal period.
- The information in this sample can be summarized by the empirical distribution of $\{d_i\}$:

$$\widehat{p}(t) = \frac{1}{N} \sum_{i=1}^{N} 1\{d_i = t\}$$

for $t = 1, 2, ..., T + 1$

Maximum Likelihood Estimation

- The log-likelihood function of this model and data is:

$$
\begin{aligned}
l(\theta) &= \sum_{i=1}^{N} \sum_{t=1}^{T+1} 1\{d_i = t\} \ \ln \Pr(d_i = t|\theta) \\
&= N \sum_{t=1}^{T+1} \widehat{p}(t) \ \ln P(t|\theta)
\end{aligned}
$$

where:

$$
\begin{aligned}
P(t|\theta) &= \Pr\left(\pi_s \geq \pi_s^*(\theta) \text{ for } s \leq t-1, \text{and } \pi_t < \pi_t^*(\theta) \mid \theta\right) \\
&= \int_{\pi_1^*(\theta)}^{\infty} \ldots \int_{\pi_{t-1}^*(\theta)}^{\infty} \int_{0}^{\pi_t^*(\theta)} dF(\pi_1, \ldots, \pi_{t-1}, \pi_t \mid \theta)
\end{aligned}
$$

Estimation: Simulation of Probabilities

- Computing $P(t|\theta)$ involves solving an integral of dimension t. For t greater than 4 or 5, it is computationally costly to obtain the exact value of these probabilities. Instead, Pakes approximate these probabilities using Monte Carlo simulation.
- For a given value of θ, let $\{\pi_t^{sim}(\theta) : t = 1, 2, ..., T\}$ be a simulated history of profits for patent i.
- Suppose that, for a given value of θ, we simulate R **independent** profit histories. Let $\{\pi_{rt}^{sim}(\theta) : t = 1, 2, ..., T;\ r = 1, 2, ..., R\}$ be these histories.
- Then, we can approximate the probability $P(t|\theta)$ using the following simulator:

$$\tilde{P}_R(t|\theta) = \frac{1}{R}\sum_{r=1}^{R} 1\{\pi_{rs}^{sim}(\theta) \geq \pi_s^*(\theta) \text{ for } s \leq t-1, \text{and } \pi_{rt}^{sim}(\theta) < \pi_t^*(\theta$$

Estimation: Simulation-Based Estimation

- The estimator of θ (Simulated Method of Moments estimator) is the value that solves the system of T equations: for $t = 1, 2, ... T$:

$$\frac{1}{N}\sum_{i=1}^{N}\left[1\{d_i = t\} - \tilde{P}_{R,i}(t|\theta)\right] = 0$$

where the subindex i in the simulator $\tilde{P}_{R,i}(t|\theta)$ indicates that for each patent i in the sample we draw R independent histories and compute independent simulators.

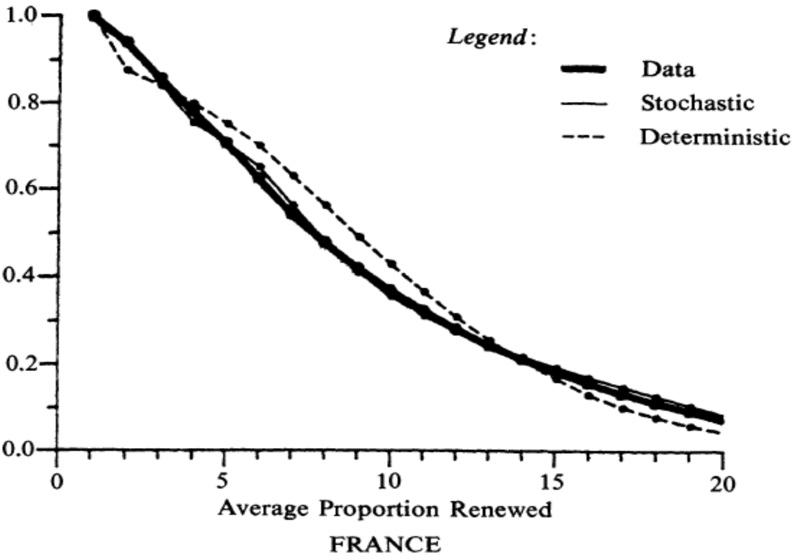
- **Effect of simulation error.** Note that $\tilde{P}_{R,i}(t|\theta)$ is unbiased such that $\tilde{P}_{R,i}(t|\theta) = P(t|\theta) + e_i(t,\theta)$, where $e_i(t,\theta)$ is the simulation error. Since the simulation errors are independent random draws:

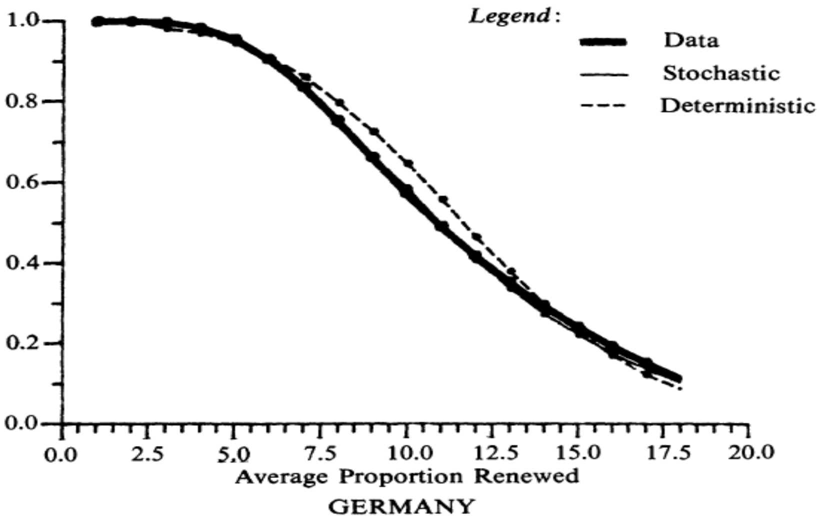
$$\frac{1}{N}\sum_{i=1}^{N} e_i(t,\theta) \rightarrow_p 0 \quad \text{and} \quad \frac{1}{\sqrt{N}}\sum_{i=1}^{N} e_i(t,\theta) \rightarrow_d N(0, V_R)$$

The estimator is consistent an asymptotically normal for any R. The variance of the estimator declines with R.

Identification

- Since there are only 20 different values for the renewal fees $\{c_t\}$ we can at most identify 20 different points in the probability distribution of patent values.
- The estimated distribution at other points is the result of interpolation or extrapolation based on the functional form assumptions on the stochastic process for profits.
- It is important to note that the identification of the distribution of patent values is NOT up to scale but in dollar values.
- For a given patent with age t, all what we can say is that: if $a_{it} = 0$, then $V_{it} < V(\pi_t^*)$; and if $a_{it} = 1$, then $V_{it} \geq V(\pi_t^*)$.


1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
Legend:
Data
Stochastic
Deterministic
Average Proportion Renewed
FRANCE


1.0
0.8
0.6
0.4
0.2
0.0
Legend:
Data
Stochastic
Deterministic
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
Average Proportion Renewed
GERMANY

Empirical Questions

- The estimated model can be used to address important empirical questions.
- **Valuation of the stock of patents.** Pakes uses the estimated model to obtain the value of the stock of patents in a country.
- According to the estimated model, the value of the stock of patents in 1963 was \$315 million in France, \$385 million in UK, and \$511 in Germany.
- Combining these figures with data on R&D investments in these countries, Pakes calculates rates of return of 15.6%, 11.0% and 13.8%, which look like quite reasonable.

Empirical Questions

- **Factual policies.** The estimated model shows that a very important part of the observed between-country differences in patent renewal can be explained by differences in policy parameters (i.e., renewal fees and maximum length).
- **Counterfactual policy experiments.** The estimated model can be used to evaluate the effects of policy changes (in renewal fees and/or in maximum length) which are not observed in the data.

Lanjow (REStud, 1999)

- Estimates the private value of patent protection for four technology areas—computers, textiles, combustion engines, and pharmaceuticals - using new patent data for West Germany, 1953-1988.
- The model takes into account that patentees must pay not only renewal fees to keep their patents but also legal expenses to enforce them.
- The dynamic structural model takes into account the potential need to prosecute infringement.
- Results show that the aggregate value of protection generated per year is on the order of 10% of related R&D expenditure.

4. Trade of Patents

Trade of Patents: Serrano (2018)

- The sale of patents is an incentive to invest in R&D, especially for small firms.
- This market can generate social gains by reallocating patent rights from innovators to firms that may be more effective in using, commercializing, or enforcing these rights.
- There are also potential social costs, if the acquiring firms can exercise more market power.
- Serrano (IER, 2018) investigates the value of trading patents by estimating a structural model that includes renewal and trading decisions.

Data

- Panel of patents granted to U.S small firms (no more than 500 employees) in the period 1988-1997 (15% of patents granted to firms).
- In the U.S. patent system, the patent holder needs to pay renewal fees to maintain the patent only at ages 5, 9, and 13 years.
- Fee increases with age: $c_{13} > c_9 > c_5$.
- Serrano (2000) constructs the dataset with renewals and transfers/sales.
- Working sample: 54,840 patents from 10 granting cohorts (1988 to 1997) followed from granting period until 2001 or not renewal.

Renewal and trading frequencies

- Probability that a patent is traded (between renewal dates):
  - higher if previously untraded.
  - decreases with age.
- Probability of patent expiration (at renewal dates)
  - lower for previously traded.
  - increase over time.

Renewal and trading frequencies

A-1: Percentage of Active Small Business Patents Traded and

Age	All	Not Previously Traded	Previously Traded (Years since last trade)	
			Any Year	One year
A. Probability that an active patent is traded				
2	2.99	2.85	7.47	7.47
7	2.81	2.46	4.79	6.63
11	2.51	2.13	3.77	2.55
B. Probability that an active patents is allowed to expire				
5	17.2	17.7	12.7	6.2
9	25.6	26.6	21.4	11.6
13	25.5	26.6	22.5	14.1

Model: Key features

- The transfer/sale of a patent involves a transaction cost.
- This transaction cost creates a selection effect: patents with higher per period returns are more likely to be traded.
- This selection effect explains the observed pattern that previously traded patents are:
  - more likely to be traded;
  - less likely to expire

Model: Returns

- At age t, a patent has:
  - an **internal return** for the current patent owner, x_t;
  - a potential **external return** for the best alternative user, y_t.
- There is an "improvement factor", g_t^e, that relates external and internal returns:
$$y_t = g_t^e \; x_t$$
- g_t^e is $i.i.d.$ with a truncated (at zero) exponential distribution: $\gamma^e \equiv \Pr(g_t^e = 0)$, and σ^e is the mean of the exponential.

Model: Returns [2]

- Initial (internal) returns: $\log(x_1) \sim N(\mu, \sigma_R^2)$.
- Next period returns:

$$x_{t+1} = \begin{cases} g_t^i \ x_t & \text{if not traded at age } t \\ g_t^i \ y_t & \text{if traded at age } t \end{cases}$$

- g_t^i is a random variable with a truncated (at zero) exponential distribution: $\gamma^i \equiv \Pr(g_t^i = 0)$, and σ_t^i is the mean of this exponential, and $\sigma_t^i = \phi^t \ \sigma_0^i$, with $\phi \in (0, 1)$.
- This implies that x_{t+1} follows a first order Markov process.
- Remember that there is a lump-sum transaction cost, τ. It is assumed that is paid by the buyer.

Model: Renewal and Sale decisions

- Let $V_t(x_t, y_t)$ be the value of a patent with age t, current internal and external returns x_t and y_t, resp.

$$V_t(x_t, y_t) = \max \left\{ 0,\ V_t^K(x_t, y_t),\ V_t^S(x_t, y_t) \right\}$$

$V_t^K(x_t, y_t) =$ value of keeping; $V_t^S(x_t, y_t) =$ value of selling.

- And for $t \leq T = 17$:

$$\begin{aligned} V_t^K(x_t, y_t) &= x_t - c_t + \beta\, \mathbb{E}\left[V_{t+1}(x_{t+1}, y_{t+1}) \mid x_t, y_t, a_t = K \right] \\ V_t^S(x_t, y_t) &= x_t - c_t - \tau + \beta\, \mathbb{E}\left[V_{t+1}(x_{t+1}, y_{t+1}) \mid x_t, y_t, a_t = S \right] \end{aligned}$$

with $V_{T+1}^K = V_{T+1}^S = 0$.

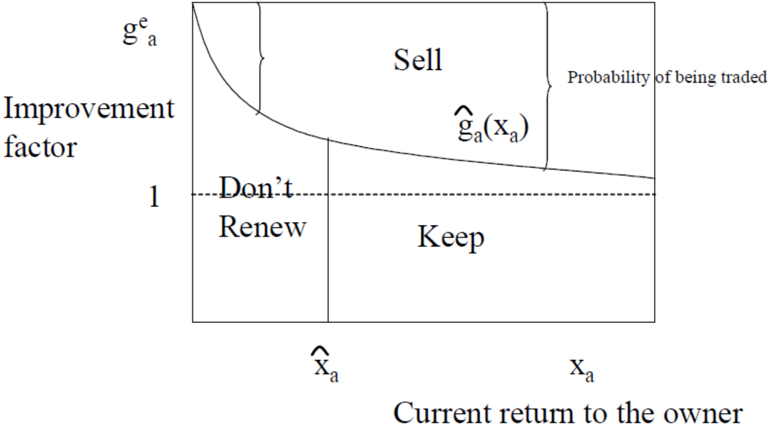
Model: Optimal decision rule

- Lemma 1: $V_t(x_t, y_t)$ is weakly increasing in x_t and y_t, and weakly decreasing in t.
- Proposition 1. There are two threshold values: $x_t^*(\theta)$ that depends on age and structural parameters, and $g_t^*(x, \theta)$, that depends on age, internal return, and parameters, such that the optimal decision rule a_t is:

$$a_t = \begin{cases} S & \text{if} \quad g_t^e \geq g_t^*(x_t, \theta) \\ \\ K & \text{if} \quad g_t^e < g_t^*(x_t, \theta) \quad \text{and} \quad x_t \geq x_t^*(\theta) \\ \\ 0 & \text{if} \quad g_t^e < g_t^*(x_t, \theta) \quad \text{and} \quad x_t < x_t^*(\theta) \end{cases}$$

Model: Optimal decision rule

Figure 1: Optimal choices of a Patent Holder



Identification and Estimation

- Method: Simulated method of moments.
- Moments describing the history of trading and renewal decisions of patent owners.

- (1) probability that an active patent is traded at different ages conditional on having been previously traded, and conditional on not having been previously traded.
- (2) probability that an active patent is allowed to expire at different renewal dates conditional on having been previously traded, and conditional on not having been previously traded.

- A total of 186 moments.

Parameter estimates

- Transaction cost: \$5,850, about one-third of the average return at age 1 (8% of the average value at age 1).
- On average, internal growth of returns is greater than external.

Parameter estimates

Table 1: Parameter Estimates

Description (Parameter)	Estimate[a]
A. Patent initial returns	
Mean parameter of the Lognormal Initial Distribution (μ)	8.4179 ($4.0*10^{-2}$)
Std. Deviation parameter of the Lognormal Initial Distribution (σ_R)	1.6911 ($1.2*10^{-2}$)
B. Internal growth of returns	
Depreciation factor (δ)	0.8917 ($5.2*10^{-3}$)
Not obsolescence (γ^i)	0.9673 ($6.5*10^{-3}$)
Internal Growth of Returns (σ^i)	0.4450 ($3.6*10^{-3}$)
Upside opportunities (ϕ)	0.5941 ($6.1*10^{-3}$)
C. Market for patents and transaction costs	
Transaction cost (τ)	5,850.1 (50.49)
Mean External Growth of Returns (σ^e)	0.3745 ($2.9*10^{-3}$)
Proportion of unsuccessful transfers (γ^e)	0.0385 ($4*10^{-4}$)
Random transfers (ε)	0.0059 ($4*10^{-4}$)
Size of sample	54,840
Simulations in the estimation	164,520
MSE[b]	$3.784*10^{-4}$

Evaluating the value of the market for patents

- The possibility of trading patents has two types of the effects on the value of the pool of patents:
  - a direct causal effect due to the reallocation to an owner with higher returns;
  - a selection effect, through the renewal decisions (renewal decision is different with and without the possibility of trading).
- Serrano measures these two sources of value.

Evaluating the value of the market for patents

- (1) Total effect on the value of patents:

- 50% of the total value of patents.
- Only 23% of patents are sold, but the value of a traded patent is 3 times higher than untraded patent ($173,668 vs. $54,960).

- (2) Direct gains from trade (from reallocation)

- accounts for 10% of the total value of the traded patents.
- The distribution of the gains from trade is very skewed.

Counterfactual: Reducing transaction cost

- Lowering transaction cost by 50% (from \$5,850 to \$2,925).
- It raises the proportion of patents traded by 6 percentage points: from 23.1% to 29.6%.
- It boosts the gains from trade (reallocation) by an additional 8.7%.
- It increases the total value of the patent market by 3%.