How much are patents worth? This question is important because it inform public policy as to optimal patent length and design. Are patents a sufficient means of rewarding innovation?

- Q_A: value of patent at age A
- Goal of paper is to estimate Q_A using data on their renewal. Q_A is inferred from patent renewal process via a *structural model* of optimal patent renewal behavior.

■■■

BEHAVIORAL MODEL

- Treat patent renewal system as exogenous (only in Europe)
- For $a = 1, \ldots, L$, a patent can be renewed by paying the fee c_a
- Timing:
  - At age $a = 1$, patent holder obtains period revenue r_1 from patent
  - Decides whether or not to renew. If renew, then pay c_1, and proceed to age $a = 2$.
  - At age $a = 2$, patent holder obtains period revenue r_2 from patent
  - Decides whether or not to renew. If renew, then pay c_2, and proceed to age $a = 3$. And so on...
- Let V_a denote the value of patent at age a.

$$V_a \equiv \max_{t \in [a,L]} \sum_{a'=1}^{L-a} \beta^{a'} R(a + a'), \text{ where}$$
$$R(a) = \begin{cases} r_a - c_a & \text{if } t \geq a \quad \text{(when you hold onto patent)} \\ 0 & \text{if } t < a \quad \text{(after you allow patent to expire)} \end{cases} \tag{1}$$

t above denotes the age at which the agent allows the patent to expire, and is the agent's choice variable in this problem. This type of problem is called an "optimal stopping" problem.

$R(a)$ denotes the profits from the patent during the a-th year. The sequence $R(1), R(2), \ldots$ is a "controlled" stochastic process: it is inherently random, but also affect by agent's actions.

- Since the maximal age L is finite, this is a finite-horizon (nonstationary) dynamic optimization problem. Most D.O. problems that you run into fall into two camps: (i) infinite-horizon, stationary problems; (ii) finite-horizon, nonstationary problems. Stationarity just means that the value functions and optimal decision rules are *time-invariant* functions of the state variables, and only depend on time through the values that the state variables take over time.

- The state variable of this DO problem is r_a, the single-period revenue.

- Finite-horizon DO problems are solved via *backward recursion*. The value functions $\{V_1(\cdot), V_2(\cdot), \ldots, V_a(\cdot), \ldots, V_L(\cdot)\}$ are recursively related via Bellman's equation:

$$V_a(r_a) = \max \left\{ 0,\ \underbrace{r_a + \beta E\left[V_{a+1}(r_{a+1})|\Omega_a\right]}_{\equiv Q_a \text{ value of age } a \text{ patent}} - c_a \right\}.$$

RHS means you will choose to renew the patent iff $Q_a - c_a > 0$.

Ω_a: history up to age a, $= \{r_1, r_2, \ldots, r_a\}$

Expectation is over $r_{a+1}|\Omega_a$. The sequence of conditional distributions $G_a \equiv F(r_{a+1}|\Omega_a)$, $a = 1, 2, \ldots$, is an important component of the model specification. Pakes' assumptions are given in Eq. (7) of the paper:

$$r_{a+1} = \begin{cases} 0 & \text{with prob. } \exp(-\theta r_a) \\ \max(\delta r_a, z) & \text{with prob. } 1 - \exp(-\theta r_a) \end{cases} \qquad (2)$$

where density of z is $q_a = \frac{1}{\sigma_a}\exp\left[-(\gamma + z)/\sigma_a\right]$ and $\sigma_a = \phi^{a-1}\sigma,\ a = 1, \ldots, L-1$.

δ, θ, γ, ϕ, and σ are the important structural parameters of the model.

- So break down maximization problem into period-by-period problem, where each period agent decides whether or not to incur cost c_a and gain the value of the patent $Q_a = r_a +$ "option value". Option value captures the value in keeping patent alive in order to make a choice tomorrow.

■■■

Implications of model seen graphically:

- Drop out at age a if $c_a > Q_a$
- Optimal decision characterized by "cutoff points":

$$Q_a > c_a \Leftrightarrow r_a > \bar{r}_a$$

  This feature due to assumption A3.3, which ensures that Q_a is increasing in r_a (so that Q_a and c_a only cross once)

- The sequence of cutoff points $\bar{r}_a < \bar{r}_{a+1} < \cdots < \bar{r}_{L-1}$: ensured by assumption A3.4.

■■■

ECONOMETRIC MODEL

This paper uses *aggregate data*:

- For a cohort j (which is a year in which patents are granted), observe the sequence $n(a, j)$, $a = f_j, f_j + 1, \dots, l_j - 1, l_j$: # of cohort j patents which are *not* renewed at age a.
- There is left and right censoring: only observe $\sum_{a=1}^{f_j} n(a, j)$ (all patents not renewed before age f_j) and $\sum_{a=l_j+1}^{L} n(a, j)$ (all patents not renewed after age l_j).
- Model is fully parametric, given the specification in Eq. (2) above.
- We construct likelihood of aggregate data by deriving $Prob(t_{ij} = a)$, the probability that an individual patent i from cohort j is renewed up to age a:

$$\begin{aligned} Prob(t_{ij} = a) &= Prob\left(r_a < \bar{r}_a, r_{a-1} > \bar{r}_{a-1} \dots r_1 > \bar{r}_1\right) \\ &= \int_{-\infty}^{\bar{r}_a} \int_{\bar{r}_{a-1}}^{\infty} \cdots \int_{\bar{r}_1}^{\infty} f(r_a, \dots, r_1) dr_1 \cdots dr_a \equiv \pi(a; \vec{c}_j) \end{aligned} \tag{3}$$

  where $f(r_a, \dots, r_1)$ denotes the joint density of revenues (as implied by Eq. (2), and $\vec{c}_j$ is the fee schedule in place for cohort j patents.

  Similarly:

$$Prob(t_{ij} \le f_j) = \sum_{a=1}^{f_j} \pi(a; \vec{c}_j) \equiv \mathbf{A}$$

$$Prob(t_{ij} \ge l_j) = 1 - \sum_{a=1}^{l_j} \pi(a; \vec{c}_j) \equiv \mathbf{B}$$

- Hence, the log-likelihood function for the aggregate dataset is (ω denotes the parameters)

$$l(\omega) = \sum_{j=1}^{J} \sum_{a=f_j}^{l_j} \log \tilde{\pi}(a; \vec{c}_j) n(a, j) \tag{4}$$

where

$$\tilde{\pi}(a;\vec{c}_j) = \begin{cases} \pi(a;\vec{c}_j) & \text{if } f_j < a < l_j \\ \mathbf{A} & \text{if } a \leq f_j \\ \mathbf{B} & \text{if } a \geq l_j \end{cases}$$

■■■

ESTIMATION DETAILS

Nested algorithm:

- Inner loop: at current parameter values $\hat{\omega}$, need to solve the DO problem in order to obtain the optimal cutoff values $\bar{r}_1, \ldots, \bar{r}_{L-1}$. (We will discuss this in more detail.)
- Outer loop: for these cutoff values, evaluate the log-likelihood function (4). This is a very complicated integral, so we evaluate this by ????? (you should know by now!)

■■■

COMPUTATIONAL DETAILS: Inner loop

The most difficult part of the estimation procedure is to solve the DO problem, for a given set of parameter values $\hat{\omega}$. We need to do this in order to recover the optimal cutoff values $\bar{r}_1, \ldots, \bar{r}_{L-1}$, which are the bounds of integration which appear in the likelihood function.

By solving the DO problem, we mean solving for the value functions $V_1(r;\hat{\omega}), \ldots, V_{L-1}(r;\hat{\omega})$. This is done by the procedure of "numerical backward recursion". There are many ways to do this (and many surveys of available methods; e.g. Rust (1996)); I describe the simplest way, the method of *discretization.*

Assume that, for each age L, the return r takes values in $[0,\bar{R}]$. Consider an M-point grid of points $\vec{r} \equiv \{r_1 < r_2 < \cdots < r_M\}$, within the interval $[0,\bar{R}]$. We refer to $\vec{r}$ as a "discretization" of the continuous "state space" $[0,\bar{R}]$. The point is that we compute the values functions $V_1(r;\hat{\omega}), \ldots, V_{L-1}(r;\hat{\omega})$ only at points $r \in \vec{r}$, and then approximate the functions at all other points via interpolation.

More specifically:

- Start from the final period L. Note that

$$V_L(r_L;\hat{\omega}) = r_L, \ \forall r_L \tag{5}$$

  because there are no profits and no chance of renewal after age L.

- Now go back to period $L-1$:

$$\begin{aligned} V_{L-1}(r_{L-1};\hat{\omega}) &= \max\left\{0, r_{L-1} + E_{r_L|r_{L-1}} V_L(r_L;\hat{\omega}) - c_{L-1}\right\} \\ &= \max\left\{0, r_{L-1} + E_{r_L|r_{L-1}} r_L - c_{L-1}\right\} \end{aligned} \tag{6}$$

  where $E_{r_L|r_{L-1}} r_L$ can be calculated based on the parametric assumptions of the model (Eq. 7 in the paper).

  Hence, for each $r \in \vec{r}$, you can calculate $V_{L-1}(r;\hat{\omega})$, resulting in an M-vector of numbers $V_{L-1}(r_1;\hat{\omega}), \ldots, V_{L-1}(r_M;\hat{\omega})$. Then you interpolate between the points, in order to obtain an approximation of the function $\tilde{V}_{L-1}(r;\hat{\omega})$, for every $r \in [0, \bar{R}]$.

  Note: this function will be zero for $r \in [0, \bar{r}_{L-1}(\hat{\omega})]$, where $\bar{r}_{L-1}(\hat{\omega})$ is the value of the cutoff point $\bar{r}_{L-1}$, at the current parameter values.

- Now go to period $L-2$:

$$V_{L-2}(r_{L-2};\hat{\omega}) = \max\left\{0, r_{L-2} + E_{r_{L-1}|r_{L-2}} \tilde{V}_{L-1}(r_{L-1};\hat{\omega}) - c_{L-2}\right\} \tag{7}$$

  where the conditional expectation of $r_{L-1}|r_{L-2}$ is given by Eq. 7 in the paper, and $\tilde{V}_{L-1}(\cdots)$ denotes the value function for period $L-1$, which was approximated in the previous step.

  As before, calculate $V_{L-2}(r_{L-2};\hat{\omega})$ for the discretized state space $r \in \vec{r}$, and then interpolate to obtain the approximation $\tilde{V}_{L-1}(r;\hat{\omega})$. Let $\bar{r}_{L-2}(\hat{\omega})$ denote the cutoff value for period $L-2$.

- Repeat this procedure for all periods $L-3, L-4, \ldots, 1$, to obtain $\tilde{V}_a(r_a;\hat{\omega})$ and $\bar{r}_a(\hat{\omega})$ for $a = 1, \ldots, L-1$.

After obtain the cutoff points $\bar{r}_a(\hat{\omega})$ for $a = 1, \ldots, L-1$, we need to simulate the likelihood function $l(\omega)$. We consider how to simulate $\pi(a; \vec{c})$.

For $s = 1, \ldots, S$ (where S is the number of simulation draws):

- Draw a sequence of returns $r_1^s, \ldots, r_L^s$, according to the parametric specification, in Eq. 7 of the paper. (ie. start by drawing r_1^s, then draw r_2^s given r_1^s, then draw r_3^s given r_2^s, etc.
- Given this sequence of returns, figure out the dropout age t^s, which equals the first a at which $r_a^s < \bar{r}_a(\hat{\omega})$.

Then, for all $a = 1, \ldots, L-1$, you can approximate

$$\pi(a; \vec{c}) \approx \frac{1}{S} \sum_{s=1}^{S} \mathbf{1}(t^s = a).$$

Then you can perform simulated maximum likelihood using the above approximation in Eq. (4) above.

References

Pakes, A. (1986): "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks," *Econometrica*, 54(4), 755–84.

Rust, J. (1996): "Numerical Dynamic Programming in Economics," in *Handbook of Computational Economics*, ed. by D. K. H. Amman, and J. Rust. North Holland.