EC5230 - Industrial Organisation

Lecture 4 - Patents and Intellectual Property Rights

Gerhard Riener

University of St Andrews

University of St Andrews

Patents and Intellectual Property Rights

Where we left off: the innovation problem

Last week we established three facts about private incentives to innovate:

1. Replacement effect – an incumbent monopolist under-invests because innovation partly cannibalises its own rents ($\Delta\pi^m < \Delta\pi^{pc}$)
2. Appropriability gap – without protection, imitation dissipates rents and the innovator's WTP falls to zero
3. Private vs social value – the social planner values a cost reduction at $\Delta W > \Delta\pi$, so private incentives are generically too weak

These problems share a common root: innovators cannot capture enough of the surplus they create.

Today's question: can intellectual property rights close this gap

Where we left off: the innovation problem

and at what cost?

Start by briefly recapping the three facts from Lecture 3 — students should recognise these immediately. Emphasise that the replacement effect, the appropriability gap, and the private-vs-social wedge all point in the same direction: without intervention, private R&D incentives are too weak.

Pose the question on the slide as a genuine open problem: IPRs are one response, but they come with their own costs. This sets up the central trade-off that runs through the entire lecture.

IPRs: economic problem

University of St Andrews

Object of analysis

IPRs make a non-rival good partially excludable, creating private incentives for costly innovation.

- Knowledge is (partly) non-rival and often hard to exclude
- Without protection, imitation can dissipate rents $\Rightarrow$ weak private incentive to incur fixed R&D cost
- Policy objective: provide incentives for invention while limiting static distortions in product markets
- Discussion: Why can't we just rely on first-mover advantage?

IPRs: economic problem

Stress the public-good nature of knowledge: non-rival means my use doesn't diminish yours; hard to exclude means patents are an artificial mechanism to restore excludability. Ask students: "Why can't we just rely on first-mover advantage?" – this motivates the need for legal protection when imitation is cheap (as in pharma, where reverse-engineering a molecule is far cheaper than discovering it).

The policy objective on the slide is the thread that connects every model in this lecture.

Learning objectives

- Explain why non-rival ideas can lead to underinvestment without protection
- Use the ideas model to compare the private investment condition $\pi\nu T \geq F$ to the planner objective
- Understand how length and breadth jointly shape the incentive-distortion trade-off
- Explain why patent races can generate socially excessive duplication

i Today: roadmap

1. Patents as incentives: the central trade-off
2. Patent length in the ideas model (screening and welfare)
3. Breadth, endogenous R&D, and patent races

Read through the learning objectives quickly – these are the exam-relevant skills. The roadmap gives students a mental scaffold: we move from the conceptual trade-off, to a formal model (Scotchmer), to design instruments (breadth vs length), and finally to

Learning objectives

strategic interaction (patent races). Flag that the investment condition $\pi\nu T \geq F$ will reappear multiple times.

IPRs: instruments

University of St Andrews

- Main IPR types:
  - → Patents
  - → Trademarks
  - → Copyrights
  - → Design rights
- Key design dimensions (patents):
  - → Length (duration)
  - → Breadth (scope)
  - → Geographical coverage
  - → Transferability (sale, licensing)

Give quick real-world anchors: patents last 20 years from filing (but effective life is shorter after regulatory approval in pharma); trademarks can last indefinitely if renewed; copyright is life + 70 years in most jurisdictions; design rights vary. The key point is that

IPRs: instruments

patents are the main IPR instrument we study because their finite duration creates the central trade-off between incentives and deadweight loss.

IPRs: central trade-off

- Patents create temporary market power $\Rightarrow$ static deadweight loss
- Stronger protection (longer/broader) typically:
  - $\rightarrow$ increases expected private returns to R&D
  - $\rightarrow$ increases the static distortion during protection
- Questions:
  - $\rightarrow$ What is an optimal length and breadth?
  - $\rightarrow$ How does competition in R&D (patent races) affect efficiency?

This slide states the fundamental tension that students must internalise: patents create deadweight loss (static cost) but generate incentives for innovation (dynamic benefit). Emphasise that "stronger" protection is not unambiguously good — it depends on where we are on the trade-off.

IPRs: central trade-off

Pose the two questions as the organising questions for the rest of the lecture. We will answer the length question with Scotchmer and Shy, and the race question in the final section.

Application: Pharmaceuticals

- Fixed development cost F is large (clinical trials), while marginal cost is low
- Patent protection creates temporary market power, but expiry enables generic entry
- Breadth maps to how close a substitute can be without infringing; races map to multiple labs pursuing the same target

Pharmaceuticals are the canonical example because (i) fixed costs are enormous (clinical trials can exceed $1bn), (ii) marginal costs are low (pill production), and (iii) patent expiry triggers immediate generic entry with large price drops. This makes every parameter in the models we study directly interpretable.

We will return to this application throughout: F = clinical trial cost, ν = health benefit flow, breadth = how different a molecule must be, races = multiple labs targeting the same disease.

Scotchmer (2006): Ideas model

Ideas model: primitives

University of St Andrews

Following Scotchmer (2006)

- An "idea" is a pair (ν, F)
- ν: per-period consumer surplus under competitive supply (value parameter)
- F: fixed cost to develop the idea into an innovation (R&D cost)
- Interpretation
  - $\rightarrow$ ν captures the size of social gains from making the idea usable
  - $\rightarrow$ F is the up-front resource cost required for development

Application: a new drug

Map (ν, F) to an innovation with large fixed R&D cost F and a flow of benefits ν that is partially appropriable during patent protection.

Ideas model: primitives

Scotchmer's ideas model is deliberately simple – the pair (ν, F) is sufficient to characterise an idea's social value and development cost. Stress that ν is consumer surplus under competitive supply (i.e., post-patent), not profit. The firm cannot capture ν directly; it captures $\pi\nu$ per period during patent life.

For the pharma application: F is the total clinical development cost, ν is the per-period health benefit to patients measured as consumer surplus if the drug were priced at marginal cost.

Ideas model (Scotchmer): social value under discounting

Assume social value lasts forever and the product is competitively supplied.

- Per-period social value: ν
- Discounted social value:

$$\sum_{t=1}^{\infty} \frac{1}{(1+r)^t}\nu = \frac{\nu}{r}$$

- Interpretation
  - $\rightarrow$ Discount rate r reduces the present value of long-run benefits
  - $\rightarrow$ Longer-lived benefits (lower r) raise the social value of an idea

Walk through the geometric series derivation step by step: $\sum_{t=1}^{\infty} \rho^t \nu = \nu \cdot \rho/(1-\rho) = \nu/r$ where $\rho = 1/(1+r)$. This is the infinite-horizon present value of a perpetual flow ν.

Ideas model (Scotchmer): social value under discounting

Key intuition: even though each period contributes ν, the present value is finite because of discounting. Lower r (more patient society) implies higher social value — which strengthens the case for investing in ideas.

Ideas model (Scotchmer): private returns and deadweight loss

University of St Andrews

- Firm's per-period private profit under patent: $\pi\nu$ where $0 < \pi < 1$
- Patent profit for discounted length T:

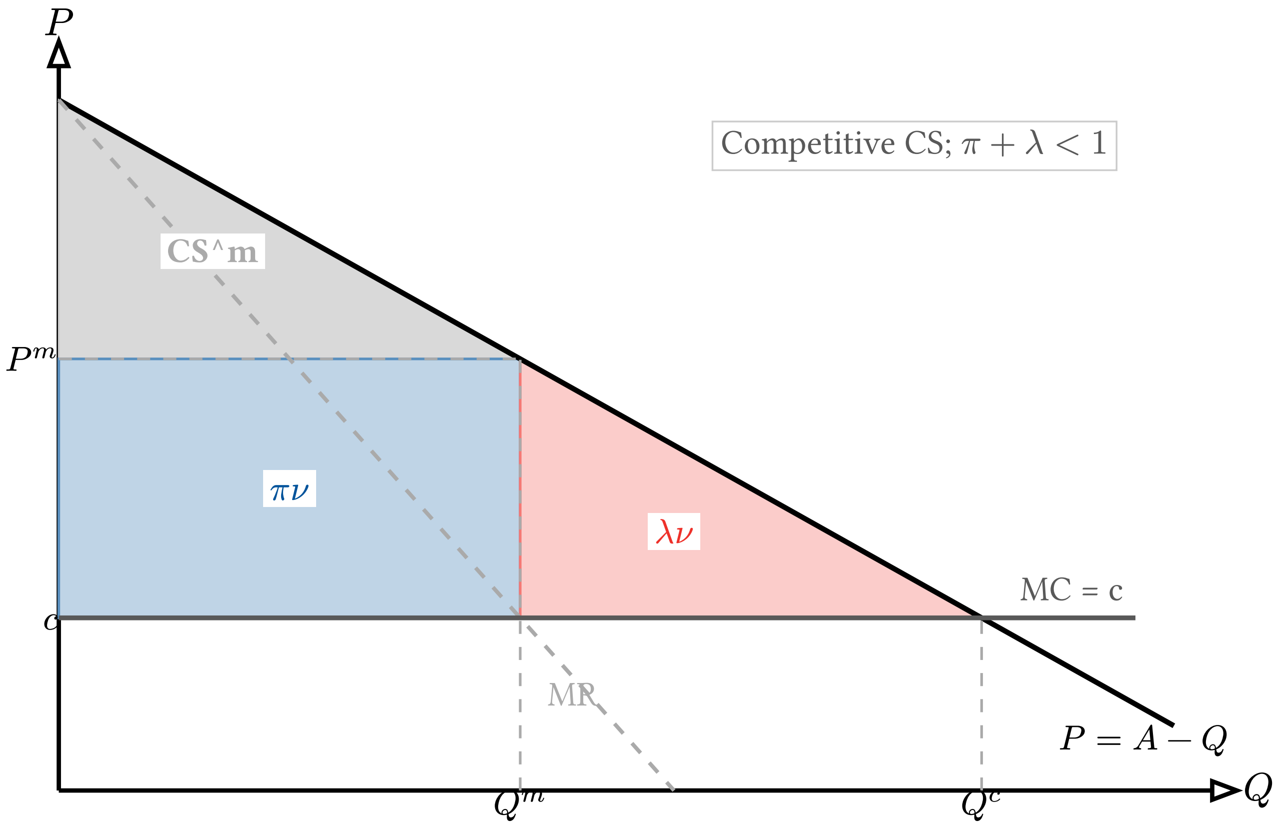
$$\pi\nu T$$

- Per-period deadweight loss: $\lambda\nu \Rightarrow$ patent DWL: $\lambda\nu T$

Interpretation of π and λ

- π is a reduced-form "appropriability" parameter
- λ captures the static distortion created by protection

Interpretation of π and λ



Interpretation of π and λ

The two parameters π and λ are crucial reduced-form objects. The diagram shows the standard monopoly decomposition: $\pi\nu$ is the profit rectangle, $\lambda\nu$ is the Harberger triangle. Note that $\pi + \lambda < 1$ because consumer surplus under monopoly (CS^m) is positive.

Ask: "What happens as demand becomes more elastic?" – π falls and λ rises, making patents more costly relative to the incentive they provide. This motivates the later discussion of breadth.

Discounted patent length: T

- Let τ be the undiscounted duration (in periods)
- Define discounted duration:

$$T = \int_0^{\tau} e^{-rt}\, dt = \frac{1 - e^{-r\tau}}{r}$$

- Discrete-time approximation used in many models:

$$T \approx \sum_{t=1}^{\tau} \frac{1}{(1+r)^t}$$

- Interpretation
  - $\rightarrow$ T is increasing in τ but bounded as $\tau \to \infty$ when $r > 0$

Discounted patent length: T

This slide introduces the discounted length T which converts calendar duration τ into present-value terms. Derive: $T = \int_0^\tau e^{-rt}\,dt = [-e^{-rt}/r]_0^\tau = (1 - e^{-r\tau})/r$. As $\tau \to \infty$, $T \to 1/r$, so even infinite patent life has finite present value.

This matters because it means the marginal benefit of extending patent duration diminishes – an extra year of protection is worth less the further into the future it falls. This is key for the optimal-length result.

Optimal patent length: innovating firm

- Patent gives discounted net profit:

$$\pi \nu T - F$$

- Firm invests if $\pi \nu T \geq F$
- Interpretation
  - $\rightarrow$ Higher ν or larger π reduces the minimum protection needed for investment
  - $\rightarrow$ Higher F requires longer (or stronger) protection to break even

The break-even condition $\pi \nu T \geq F$ is the private investment threshold. Rearranging: $T \geq F/(\pi \nu)$. This gives the minimum discounted patent length needed to induce investment.

Optimal patent length: innovating firm

Emphasise that this is a necessary condition only – it tells us when the firm is willing to invest, not whether investment is socially desirable. The planner's problem on the next slide adds the welfare side.

Optimal patent length: social planner

- Discounted net social value (invention made):

$$\frac{\nu}{r} - \lambda \nu T - F$$

- Interpretation
  - $\rightarrow$ Planner values the full flow benefit ν, but counts DWL during protection
  - $\rightarrow$ Optimal design trades off inducing investment against static costs

The planner's net social value if the idea is developed is $\nu/r - \lambda\nu T - F$. The first term is the perpetual benefit, the second is the DWL during patent life, and the third is the development cost. The planner wants to set T just large enough to satisfy the firm's break-even condition, and no larger – every extra unit of T costs $\lambda\nu$ in deadweight loss without additional benefit once investment is already induced.

Optimal patent length: social planner

This is why the planner's optimal T exactly solves $\pi\nu T = F$ for marginal ideas, i.e., $T^* = F/(\pi\nu)$.

Optimal length: heterogeneity and screening (intuition)

- If inventions differ in (ν, F), "one-size-fits-all" length is not generally optimal
- Comparative static (holding other parameters fixed):
  - $\rightarrow$ more elastic demand / stronger substitution $\Rightarrow$ higher DWL λ $\Rightarrow$ shorter protection
  - $\rightarrow$ higher development cost F $\Rightarrow$ longer protection to ensure investment

With heterogeneous ideas, uniform patent length cannot be first-best: some ideas need long protection (high F/ν) while others need little. A single T either under-protects high-cost ideas or over-protects low-cost ones.

This is the screening problem: the planner cannot observe (ν, F) directly and must use a single instrument. In practice, patent offices partially address this through examination standards and varying breadth across technology classes, but the fundamental tension remains.

Example: ideas A and B

- Idea A: $(\nu_A = 5, F_A = 10)$
- Idea B: $(\nu_B = 2, F_B = 20)$

Let $T = 20, \pi = \frac{1}{2}, \lambda = \frac{1}{4}, r = \frac{1}{3}$.

- Tasks:
  - $\rightarrow$ Which ideas are privately profitable ($\pi\nu T \geq F$)?
  - $\rightarrow$ Which ideas have positive discounted net social value $(\nu/r - \lambda\nu T - F \geq 0)$?
- Interpretation
  - $\rightarrow$ With these parameter values, private profitability need not coincide with positive net social value: patent protection can induce investment even when net welfare is negative

Work through the numbers on the board:

Example: ideas A and B

Idea A ($\nu_A = 5, F_A = 10$): Private profit $= \pi\nu T = 0.5 \times 5 \times 20 = 50 \geq 10$ – privately profitable. Net social value $= \nu/r - \lambda\nu T - F = 5/0.333 - 0.25 \times 5 \times 20 - 10 = 15 - 25 - 10 = -20$ – socially negative.

Idea B ($\nu_B = 2, F_B = 20$): Private profit $= 0.5 \times 2 \times 20 = 20 \geq 20$ – just profitable. Net social value $= 2/0.333 - 0.25 \times 2 \times 20 - 20 = 6 - 10 - 20 = -24$ – also socially negative.

The punchline: both ideas are privately profitable but socially wasteful at these parameter values. This illustrates how patent protection can induce "too much" innovation from a welfare perspective when DWL is high.

Breadth: product space

Patent breadth: product space (definition)

- Breadth determines how close a substitute can be without infringing
- Reduced-form implication:
  - $\rightarrow$ Narrower breadth $\Rightarrow$ more close substitutes enter
  - $\rightarrow$ Broader breadth $\Rightarrow$ fewer close substitutes enter
- Interpretation
  - $\rightarrow$ Allowing close substitutes increases the elasticity of demand faced by the patent holder

i Application: close substitutes

In pharmaceuticals, breadth maps to whether a close substitute can launch during protection without infringing.

Patent breadth: product space (definition)

Breadth is harder to formalise than length because it operates in product space rather than time. The key reduced-form idea: broader patents exclude more potential substitutes, which makes the patent holder's residual demand less elastic and allows higher prices.

In pharma: if a patent on molecule X is broad, then molecules X', X" that are chemically similar cannot be sold. If narrow, "me-too" drugs can enter. The Hatch-Waxman Act in the US and supplementary protection certificates in the EU are real-world breadth instruments.

Breadth and demand elasticity (intuition)

- If close substitutes are allowed:
  - $\rightarrow$ residual demand becomes more elastic
  - $\rightarrow$ equilibrium price is lower (all else equal)
- If substitutes are excluded (broader patent):
  - $\rightarrow$ residual demand is less elastic
  - $\rightarrow$ equilibrium price is higher (all else equal)
- Discussion: If you were a patent holder, would you prefer broad–short or narrow–long protection?

This is pure intuition – no formal model yet. The mechanism: more substitutes $\longrightarrow$ more elastic residual demand $\longrightarrow$ lower equilibrium price $\longrightarrow$ lower per-period profit π. Conversely, excluding substitutes raises π but also raises λ (DWL) because the monopoly distortion is larger when demand is less elastic.

Breadth and demand elasticity (intuition)

University of St Andrews

Ask students: “If you were a patent holder, would you prefer broad-short or narrow-long protection?” This previews the next slide’s trade-off.

Breadth–length trade-off (given a target value)

Assume the "correct" expected private value of protection is fixed.

- Regimes (product space):
  - → Broad–short: $(\hat{T}, \hat{\pi}_1 + \hat{\pi}_2)$
  - → Narrow–long: $(\tilde{T}, \tilde{\pi}_1)$
- Broad patent yields higher per-period profit (includes infringing market):

$$\hat{T}(\hat{\pi}_1 + \hat{\pi}_2) = \tilde{T}\tilde{\pi}_1$$

Breadth–length trade-off (continued)

University of St Andrews

- Therefore:

$$\hat{T} < \tilde{T}$$

- Interpretation
  - $\rightarrow$ Broad protection can be paired with shorter duration to deliver the same incentive level

The key equation $\hat{T}(\hat{\pi}_1 + \hat{\pi}_2) = \tilde{T}\tilde{\pi}_1$ says that if we hold the firm's expected discounted profit constant, a broader patent (which yields higher per-period profit by including the infringing market) can achieve the same incentive with a shorter duration.

The notation: $\hat{\pi}_1$ is profit from the patented good, $\hat{\pi}_2$ is profit from the infringing substitute (both captured under broad protection), and $\tilde{\pi}_1$ is profit from the patented good alone under narrow protection. Since $\hat{\pi}_1 + \hat{\pi}_2 > \tilde{\pi}_1$, we need $\hat{T} < \tilde{T}$.

Which regime is better?

University of St Andrews

- The best regime depends on substitution patterns:
  - → substitution between the patented good and an infringing substitute
  - → substitution between these goods and the rest of consumption
- Interpretation
  - → Broad–short: more sensitive to outside substitution (pricing alignment across many goods)
  - → Narrow–long: more sensitive to within-category substitution

This is the punchline of the breadth section: there is no universal answer. If the patented good and its substitute are close substitutes for each other but poor substitutes for other goods (e.g., two statins), broad-short may generate less DWL. If they are good substitutes for outside goods, narrow-long may be preferable.

Which regime is better?

The practical implication: optimal patent design is industry-specific. Pharma (where substitutes are close) may warrant different breadth-length combinations than software (where substitutes are more distant).

Optimal patent length with endogenous R&D (Shy)

University of St Andrews

Shy (1995) model: setup

- Demand: $P(Q) = a - Q$
- Process innovation reduces marginal cost from c to $c - x$
- R&D effort x costs $R(x)$
- Two-stage game:
  1. Regulator chooses patent duration τ
  2. Firm chooses x to maximise discounted profit
- Objects
  - $\rightarrow$ Choice variables: x (firm), τ (regulator)
  - $\rightarrow$ Parameters: a, c, r

The Shy model endogenises R&D effort x, unlike Scotchmer where the idea (ν, F) is exogenous. Here the firm chooses how much cost reduction to pursue, and the regulator

Shy (1995) model: setup

sets patent duration τ anticipating this choice. This is a Stackelberg game with the regulator as leader.

Linear demand $P = a - Q$ keeps the algebra tractable. The key modelling choice: process innovation reduces marginal cost, so the gain from R&D is a cost advantage. This connects back to the drastic/non-drastic innovation distinction from Lecture 3.

Firm's choice of x given τ

Firm solves:

$$\max_x \Pi(x;\tau) = \sum_{t=1}^{\tau} \rho^{t-1}\pi(x) - R(x),$$

$$\text{where} \quad \rho = \frac{1}{1+r}$$

$$\sum_{t=1}^{\tau} \rho^{t-1} = \frac{1-\rho^\tau}{1-\rho}$$

Assume:

- per-period profit: $\pi(x) = (a-c)\,x$
- cost: $R(x) = \frac{x^2}{2}$

Firm's choice of x given τ

University of St Andrews

Walk through the firm's problem: $\max_x \sum_{t=1}^{\tau} \rho^{t-1}(a-c)x - x^2/2$. The per-period profit $(a-c)x$ is linear in x (simplified from the full monopoly profit for tractability). The R&D cost $x^2/2$ is convex, ensuring an interior solution.

The geometric sum $\sum \rho^{t-1} = (1-\rho^{\tau})/(1-\rho)$ converts the finite-horizon discounted sum into a single multiplier. This is the discrete-time analogue of the T we defined earlier. FOC: $(1-\rho^{\tau})/(1-\rho) \cdot (a-c) = x$, giving the next slide's result.

Induced innovation level

FOC implies:

$$x^I(\tau) = \frac{1-\rho^\tau}{1-\rho}(a-c)$$

- Comparative statics:
  - $\rightarrow$ x^I increases with τ
  - $\rightarrow$ x^I increases with a and decreases with c
  - $\rightarrow$ x^I increases with ρ (decreases with r)
- Interpretation
  - $\rightarrow$ Longer protection raises the marginal benefit of R&D because profits are earned for more discounted periods

The induced innovation level $x^I(\tau) = [(1-\rho^\tau)/(1-\rho)](a-c)$ is the firm's best response to any given patent duration.

Induced innovation level

University of St Andrews

Comparative statics to emphasise: (i) $\partial x^I / \partial \tau > 0$ because ρ^τ is decreasing in τ; (ii) $\partial x^I / \partial a > 0$ because larger markets raise the return to cost reduction; (iii) $\partial x^I / \partial r < 0$ (equivalently $\partial x^I / \partial \rho > 0$) because more patient firms value future profits more.

Note that x^I is bounded above: as $\tau \to \infty$, $x^I \to (a - c)/(1 - \rho)$. Infinite patent life still yields finite R&D effort.

Shy model: planner's choice of patent duration (statement)

University of St Andrews

Planner chooses τ trading off: - higher induced innovation $x^I(\tau)$ - static deadweight loss under monopoly pricing during protection

- Result (as in Shy): optimal duration is finite, $T^* < \infty$
- Interpretation
  - $\rightarrow$ Marginal benefit of longer protection (higher induced $x^I(\tau)$) eventually falls below the marginal cost (additional monopoly distortion during protection)

The planner's problem: choose τ to maximise total welfare, which includes consumer surplus, firm profit, and the cost of monopoly distortion during the patent period. The result $T^* < \infty$ follows because the marginal benefit of extending τ (more R&D, hence lower cost) is diminishing (since x^I is concave in τ), while the marginal cost (one more period of monopoly DWL) is approximately constant.

Shy model: planner's choice of patent duration (statement)

Connect to Scotchmer: the Shy model endogenises what Scotchmer took as given — the level of innovation — but reaches the same qualitative conclusion: finite, interior patent duration is optimal.

Patent races

Symmetric patent race: Setup

University of St Andrews

- Two symmetric firms may incur a fixed cost f to establish a research division
- Success probability: p (per firm)
- Payoffs:
  - $\rightarrow$ monopoly profit if sole innovator: π^m
  - $\rightarrow$ duopoly profit if both succeed: π^d

i Application: parallel R&D programs

Think of multiple labs racing to develop the same drug/vaccine: each pays a fixed setup cost and succeeds with some probability.

We now shift from single-firm models to strategic interaction. The patent race setup is a simultaneous-move game: each firm decides whether to invest f (enter the race) or not. The key feature is "winner takes all" – if only one firm succeeds, it gets π^m; if both succeed, they split the market (π^d each).

Symmetric patent race: Setup

The welfare benchmarks W^m and W^d include consumer surplus, which is higher under duopoly ($CS^d > CS^m$). This creates a wedge between private and social incentives: firms care about profit, the planner cares about total welfare.

Pharma application: multiple labs racing to develop a COVID vaccine, each investing billions with uncertain success.

Symmetric patent race: Welfare benchmarks

- Welfare benchmarks (post-innovation welfare):
  - → one research division: $W^m = \pi^m + CS^m$
  - → two research divisions: $W^d = 2\pi^d + CS^d$
- Assumption (for the comparison):
  - → $CS^d > CS^m$ (more competition in the product market raises consumer surplus)

Patent race: game structure

University of St Andrews

⚠ Duplication incentive

"Winner-takes-all" payoffs can create privately excessive entry into R&D when firms ignore duplication costs.

- Two firms simultaneously choose Invest (I) or Not Invest (NI)
- Each firm succeeds independently with probability p
- Payoffs depend on market structure:
  - $\rightarrow$ Unique success $\Rightarrow$ monopoly profit π^m
  - $\rightarrow$ Both succeed $\Rightarrow$ duopoly profit $\pi^d < \pi^m$

Set up the game carefully before jumping to the equilibrium condition. Emphasise that success is probabilistic and independent – this is what creates the duplication problem. If success were certain ($p = 1$), the game reduces to a standard entry game. The

Patent race: game structure

University of St Andrews

interesting case is $p < 1$, where the second firm provides "insurance" against the first firm's failure but also wastes f if both succeed.

Patent race: Nash equilibrium condition

University of St Andrews

- If rival invests, my expected payoff from I:

$$p(1-p)\pi^m + p^2\pi^d - f$$

- (I, I) is a Nash equilibrium if:

$$f \leq p(1-p)\pi^m + p^2\pi^d \equiv f_2^{priv}$$

- Interpretation
  - $\rightarrow$ First term: I succeed, rival fails $\Rightarrow$ monopoly
  - $\rightarrow$ Second term: both succeed $\Rightarrow$ duopoly

Derive the condition: if the other firm invests, my expected payoff from investing is $p(1-p)\pi^m + p^2\pi^d - f$. The first term is the probability I succeed and the rival fails (I get monopoly profit); the second term is both succeed (duopoly profit). I invest if this exceeds zero, i.e., $f \leq p(1-p)\pi^m + p^2\pi^d$.

Patent race: Nash equilibrium condition

If the other firm does not invest, my expected payoff from investing is $p\pi^m - f$, so I invest if $f \leq p\pi^m$. For (I, I) to be a NE, the first condition must hold (neither wants to deviate from both investing).

Patent race: social optimum condition

- It is socially optimal to have one research division rather than two if:

$$f \geq p(1-2p)W^m + p^2 W^d \equiv f_2^{publ}$$

- Interpretation
  - $\rightarrow$ The planner compares expected welfare under one vs two research divisions, counting duplication cost f

The planner's condition: two divisions are socially optimal over one if the expected welfare gain from the second lab exceeds its cost. Expected welfare with two labs: $p^2 W^d + 2p(1-p)W^m + (1-p)^2 \cdot 0$. With one lab: pW^m. The net gain from the second lab must exceed f.

Patent race: social optimum condition

After simplification, two divisions are socially preferred if $f \leq p(1-2p)W^m + p^2 W^d \equiv f_2^{publ}$. Note: when $p > 1/2$, the term $p(1-2p)W^m$ is negative, reflecting that with high success probability, the second lab mostly duplicates what the first would have achieved.

Socially excessive R&D (region)

- Socially excessive duplication occurs when:

$$f_2^{publ} < f < f_2^{priv}$$

- Interpretation:
  - → Firms overinvest when the negative externality on rivals' profits outweighs the consumer-surplus gain from having two innovators

The excessive duplication region $f_2^{publ} < f < f_2^{priv}$ exists whenever $f_2^{priv} > f_2^{publ}$. This happens because private firms count only their own profit (business stealing from the rival), while the planner counts total welfare (including consumer surplus that partially offsets the duplication cost).

The intuition: each firm ignores two externalities – (i) the negative externality on the rival's profit (business stealing), and (ii) the positive externality on consumers (more

Socially excessive R&D (region)

competition). When (i) dominates (ii), private entry is excessive. This is the classic common-pool problem applied to R&D.

Numerical check: duplication region

Let $p = 0.8, \pi^m = 10, \pi^d = 3, W^m = 15, W^d = 16.$

- Private two-division threshold:

$$\begin{aligned} f_2^{priv} &= p(1-p)\pi^m + p^2\pi^d \\ &= 0.8(0.2)(10) + 0.64(3) = 3.52 \end{aligned}$$

- Social two-division threshold:

$$\begin{aligned} f_2^{publ} &= p(1-2p)W^m + p^2 W^d \\ &= 0.8(-0.6)(15) + 0.64(16) = 3.04 \end{aligned}$$

- If $f = 3.2$, then $f_2^{publ} < f < f_2^{priv}$: two research divisions are privately viable but socially excessive.
- Discussion: What happens to this region as p increases?

Numerical check: duplication region

University of St Andrews

Walk through the arithmetic step by step:

$$f_2^{priv} = 0.8 \times 0.2 \times 10 + 0.64 \times 3 = 1.6 + 1.92 = 3.52.$$

$$f_2^{publ} = 0.8 \times (-0.6) \times 15 + 0.64 \times 16 = -7.2 + 10.24 = 3.04.$$

So the duplication region is $f \in (3.04, 3.52)$. At $f = 3.2$: privately both firms want to invest (3.2 < 3.52), but the planner would prefer only one lab (3.2 > 3.04). The welfare loss is the wasted $f = 3.2$ from the redundant lab, net of the small probability-weighted gain from having a backup.

Ask: “What happens to this region as p increases?” – it widens, because high p makes duplication more wasteful.

Summary and next week

Summary

- Patents trade off dynamic incentives against static distortions (deadweight loss during protection)
- In the ideas model, investment requires $\pi\nu T \geq F$, while welfare accounts for ν/r and the DWL term $\lambda\nu T$
- Breadth and length can be substitutes in delivering a given private incentive level (broad−short vs narrow−long)
- Patent races can generate socially excessive duplication when private entry incentives exceed social benefits

Next week: Multi stage gamess

- Commitment and first-mover advantage (Stackelberg)
- Subgame perfect equilibrium and backward induction

Summary and next week

University of St Andrews

- Strategic delegation (Vickers)

Recap the four main results: (1) patents trade static DWL for dynamic incentives; (2) the ideas model gives a clean break-even condition $\pi\nu T \geq F$ and shows the planner sets T to just satisfy it; (3) breadth and length are substitutes in delivering incentives; (4) patent races can produce excessive duplication.

Preview next week: we move from innovation policy to strategic commitment. Stackelberg leadership and strategic delegation (Vickers) are about how firms use observable commitments to shift equilibrium outcomes — a different type of strategic interaction from the simultaneous-move games we've studied so far.

Case Study: COVID-19 vaccine patent race

History: mRNA as a platform (Storz 2022)

University of St Andrews

- mRNA vaccines look modular, but rely on layers of upstream patents (modified nucleosides, delivery/LNPs, manufacturing, formulations)
- Race was not only time-to-market, but also time-to-priority:
  - → who files first on enabling technologies vs. product-specific
  - → how broad claims interact with follow-on entrants and variants
- Post-success phase: fast diffusion of know-how ⟶ ex post disputes over who owns key enabling steps
  (licensing demands, litigation threats, bargaining over royalties)

i IO lens

Treat the vaccine as an innovation stack: competition happens across layers (upstream enablement) and across products (downstream vaccines).

History: mRNA as a platform (Storz 2022)

Use Storz (2022) as the narrative backbone: mRNA "speed" is enabled by prior platform advances, which is exactly why IP becomes contentious after the fact. Emphasise: in platform industries, "who invented the product" is not the same as "who owns the bottleneck patents."

Problem 1: patent-race incentives interact with safety (Kim 2020)

University of St Andrews

- Standard patent-race logic: winner-takes-most rents $\longrightarrow$ strong incentives to accelerate R&D effort
- With vaccines, firms choose two margins:
  1) investment in speed (inventing first)
  2) investment in safety (reducing side-effect risk)
- Liability regime matters:
  - $\rightarrow$ If the first inventor captures rents but also bears liability, private incentives can tilt toward too much investment in both speed and safety under strict liability (relative to the social optimum in the model)
- Empirical complication:

Problem 1: patent-race incentives interact with safety (Kim 2020)

→ "Race outcomes" depend on legal institutions (liability, indemnification, compensation schemes), not only patent length/breadth

⚠ **Identification challenge**

Observed "R&D intensity" bundles patent-race incentives with the liability/indemnity environment—hard to attribute outcomes to IP alone.

Anchor to Kim (2020): the distinctive feature is the joint choice of "invent" and "make safe," and how liability rules re-rank incentives. Connect back to your race model: in vaccines, the prize is not just monopoly profit; it's profit net of expected liability and reputational costs.

Problem 2 public funding $\longrightarrow$ private appropriation (Florio 2022) + pooling (Billette de Villemeur et al. 2023)

- mRNA COVID vaccines built on substantial public-sector science + funding, but patent ownership is largely private
  $\longrightarrow$ distributional conflict: who should capture returns? (Florio 2022)
- This creates at wedge:
  - $\rightarrow$ ex ante: strong incentives to race
  - $\rightarrow$ ex post: fragmented rights $\longrightarrow$ bargaining, "hold-up" risk, and politically salient access constraints
- Proposed "third way": patent pooling to expand access while preserving incentives (Billette de Villemeur et al., SSRN 2023)
  - $\rightarrow$ pool aggregates relevant rights and offers standardized licences (reducing transaction costs)
  - $\rightarrow$ goal: maximize access subject to participation/incentive constraints

Problem 2 public funding ⟶ private appropriation (Florio 2022) + pooling (Billette de Villemeur et al. 2023)

⟶ positioned as an alternative to: (i) full waiver, (ii) fully fragmented bilateral licensing

i Design margin

Pooling is a market design response: reduce IP fragmentation/frictions without eliminating IP rents entirely.

Use Florio (2022) to motivate why purely “private property” framing is politically unstable when taxpayers funded upstream knowledge. Then introduce pooling (Billette de Villemeur et al.) as an IO mechanism: coordination device that can mitigate anticommons/hold-up and accelerate diffusion. Tie back to your lecture: this is breadth/ transferability in practice (licensing rules + governance).

Problem: fragmented IP after the innovation race

University of St Andrews

- COVID-19 vaccines are not single inventions but bundles of complementary technologies:
  - $\rightarrow$ platform patents (e.g. mRNA chemistry, delivery systems)
  - $\rightarrow$ formulation and manufacturing patents
  - $\rightarrow$ downstream production know-how
- Result: multiple entities hold indispensable rights
  $\Rightarrow$ downstream producers must negotiate many licenses to manufacture at scale.
- IO interpretation:
  - $\rightarrow$ complements create multiple marginalization
  - $\rightarrow$ bargaining frictions and legal uncertainty raise effective marginal costs

Problem: fragmented IP after the innovation race

University of St Andrews

Key tension

The patent race may solve innovation speed ex ante, but can generate coordination failures ex post.

This slide bridges from the patent-race section. Stress that the race produced innovation quickly, but ownership is fragmented afterward. Use language students know: complements, vertical structure, multiple margins. This sets up why a market-design solution becomes relevant.

Proposed solution: a non-profit patent pool ("third way")

University of St Andrews

- The paper proposes a non-profit patent pool that bundles all required licenses into a single contract (a "one-stop shop").
- Pool objective:
  - → maximize access / quantities
  - → while ensuring patent holders are not worse off than under separate licensing (participation constraint).
- Compared regimes in the model:
  1. separate, non-cooperative licensing (multiple margins)
  2. for-profit pool (joint profit maximization)
  3. non-profit pool (access maximization)
- Economic mechanism:

Proposed solution: a non-profit patent pool (“third way”)

University of St Andrews

→ internalizes complementarities
→ lowers royalty stack
→ reduces transaction costs and hold-up risk.

i IO insight

Pooling converts many complementary monopolies into a single coordinated pricing problem.

Emphasise: this is not a patent waiver. It preserves incentives by maintaining patent-holder profits relative to decentralized licensing. Relate explicitly to Shapiro (patent thickets) and multiple marginalization from vertical IO.

How the solution fits – and what it does not solve

University of St Andrews

What it addresses well

- Fragmented ownership $\longrightarrow$ bundling reduces bargaining frictions
- Access problem $\longrightarrow$ lower final prices, higher quantities
- Welfare gains increase with the number of patent holders.

What remains unresolved

- Does not model the ex ante patent race:
  - $\rightarrow$ incentives for speed vs safety
  - $\rightarrow$ duplication of R&D effort
- Does not resolve manufacturing capacity or tacit know-how constraints by itself

Lecture takeaway

- Patent races solve the innovation problem.
- Patent pools address the coordination problem after innovation.

How the solution fits — and what it does not solve

University of St Andrews

- Policy design must separate:
  - → incentives to invent (dynamic efficiency)
  - → mechanisms for diffusion and global access (static + distributional efficiency).

This is the synthesis slide. Position pooling as a downstream governance mechanism rather than an innovation-incentive mechanism. This helps students see that different policy tools target different margins. End with a strong conceptual distinction: race vs diffusion.

References

Kim, Jeong-Yoo. 2020. "Patent Races for COVID-19 Vaccines and Liability Rules." Asian Journal of Law and Economics 11 (3). https://doi.org/10.1515/ajle-2020-0042.

Scotchmer, Suzanne. 2006. Innovation and Incentives. MIT Press.

Shy, Oz. 1995. Industrial Organization: Theory and Applications. MIT Press.

Storz, Ulrich. 2022. "The COVID-19 Vaccine Patent Race." Nature Biotechnology 40 (7): 1001–4. https://doi.org/10.1038/s41587-022-01376-1.