

# LEARNING ACROSS ALTERNATIVES III: JUDICIAL DECISION MAKING

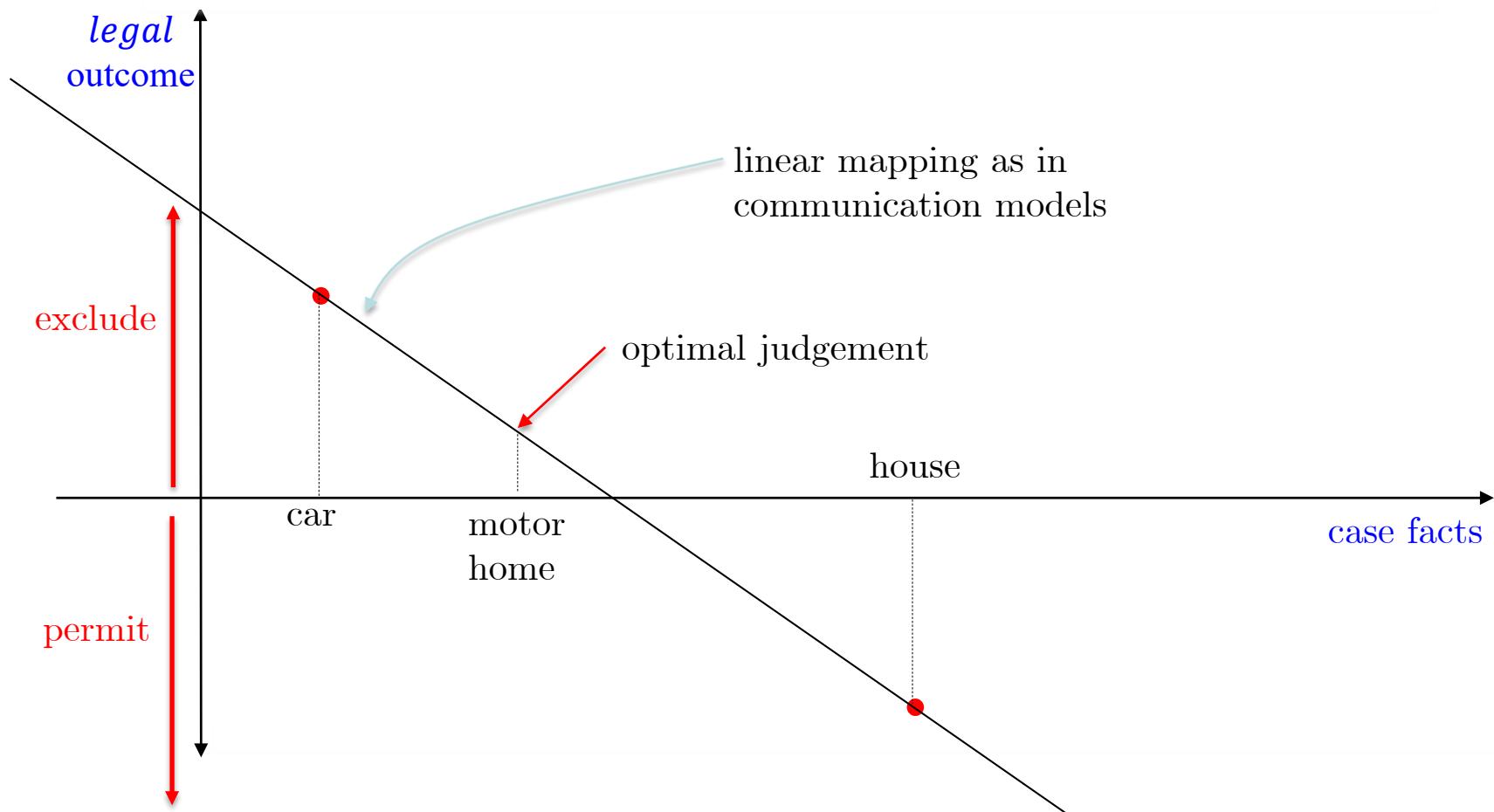
Org Econ Workshop  
October 2020

Steven Callander  
Stanford University & UNSW

## MOTIVATING EXAMPLE

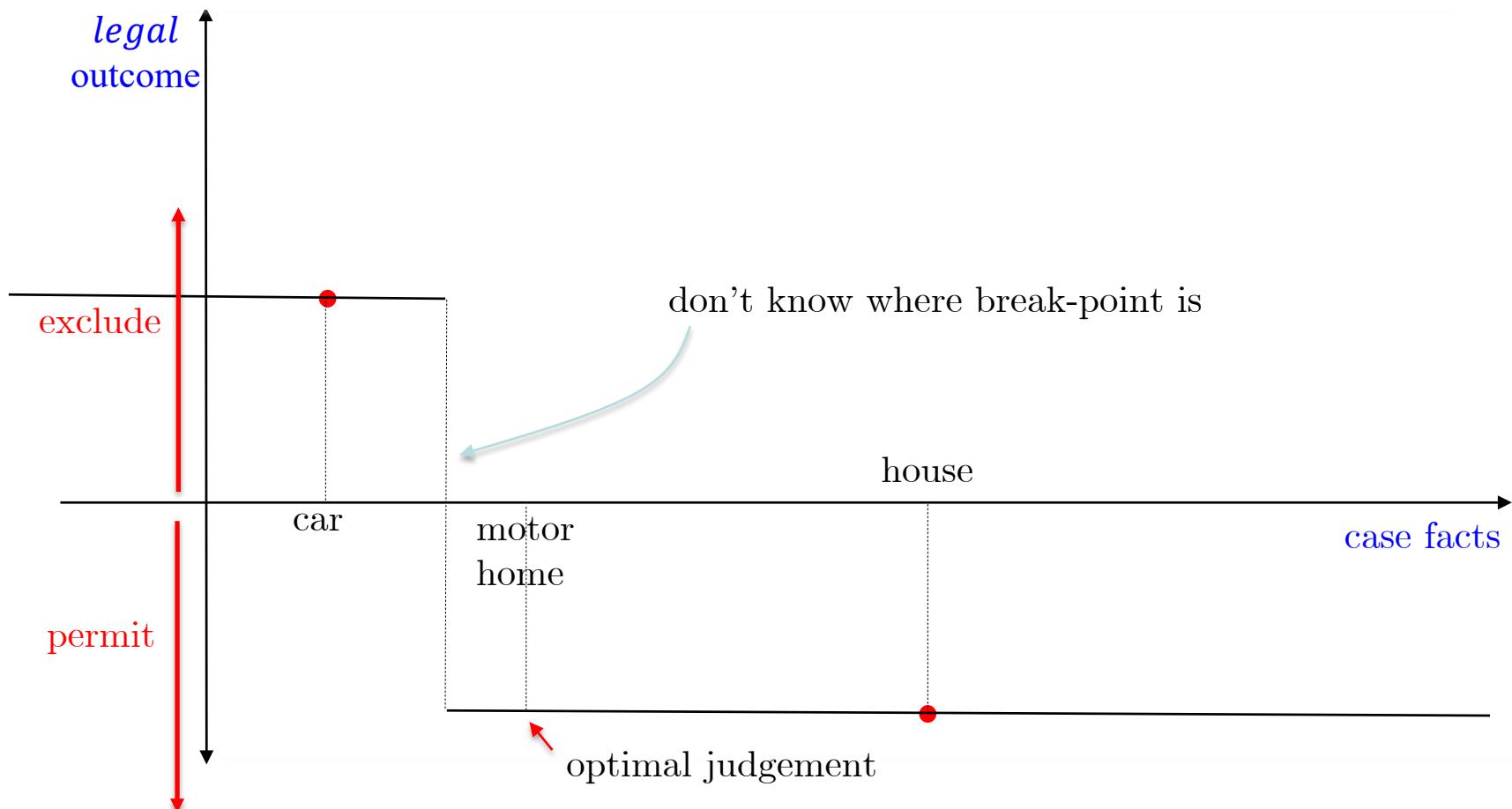
- During a murder investigation, police search a suspect's motor home and find incriminating evidence. The defense challenges the search in court. A judge must decide whether to “permit” or “exclude” the search.
- Two relevant precedents exist for this circumstance:
  - Precedent 1. Evidence from searching a suspect's car is permitted.
  - Precedent 2. Evidence from searching a suspect's home is excluded.
- How should the judge rule?
  - Is a motorhome more like a car or more like a home?
  - Is there a legal “doctrine” that guides the decision?
  - How does the judge make this decision?
- How much information can the judge extract from precedent?

# EASY PROBLEM WITH CANONICAL MODELS



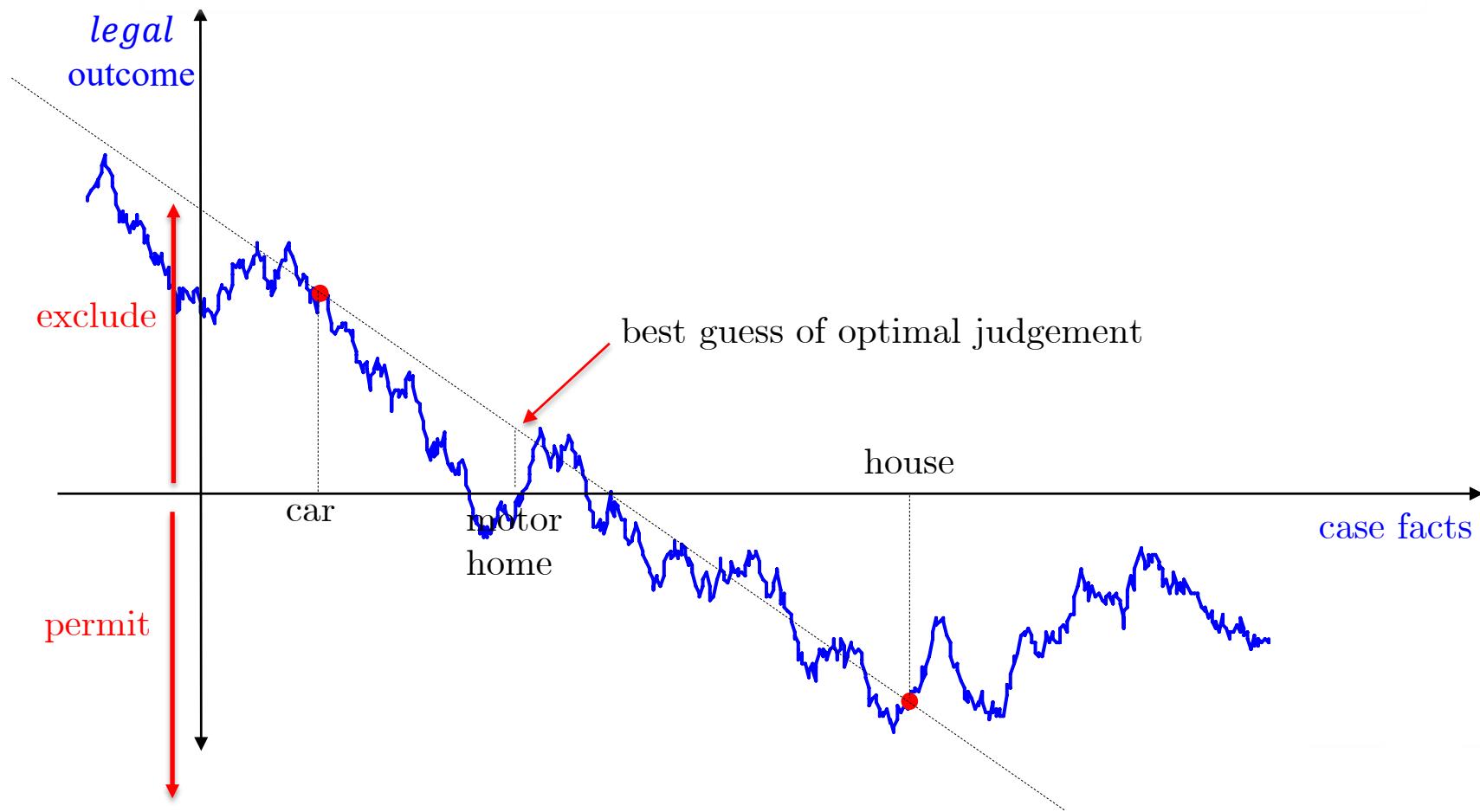
- If mapping from facts to outcomes is as in communication models, then two cases is one more than is needed to know what judge should decide.

# HARDER VERSION WITH SINGLE UNKNOWN



- Baker and Mezzetti (2012, JPE). Two precedents not enough information!
  - Can interpret as observability constraint – judges only see sign of correct outcome not value.
  - Solve optimal dynamic problem of when to hear cases and which cases to hear.
  - Results: Same facts can be adjudicated differently over time, learning converge to inefficient rules,

# JUDICIAL DECISION MAKING IN A COMPLEX WORLD



- Mapping from case facts to legal outcome is a Brownian motion.
- Any number of precedents is not enough! Infer more information from precedent.
- Hume (1748): "*From causes which appear similar we expect similar effects.*"

# A DIFFERENT TYPE OF DECISION PROBLEM

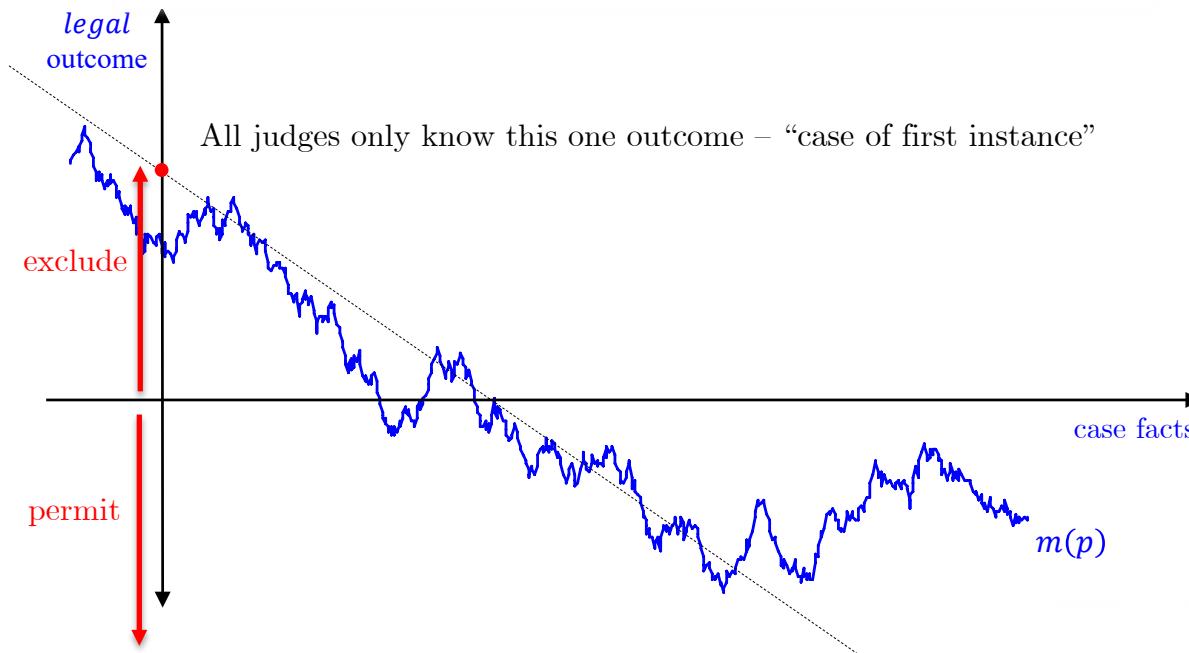
- Develop a model of judicial decision making in a complex world.
- Two levels of questions:
  1. How do judges make decisions on cases with limited information?

Note: this is a big and long-running debate in law. Do judges have a guiding theory or do they just make it up as they go?
  2. What cases should the Supreme/High court hear (grant *certiori*)? How should it shape precedent given the answer to (1)?
- This is a very different type of problem to search/exp & communication
  - Need to make a decision for each alternative and optimize over all vs. find the best alternative.
- We don't make as much progress! It is hard. (w/ Tom Clark, *APSR* 2017)
  - e.g., no forward-looking behavior, weaker characterization.
  - Arjada Bardi makes more progress on a related problem at 11am Sydney time.

# A MODEL OF A JUDICIAL HIERARCHY

- Judicial hierarchy (orgs!):
  - Fact-finding courts (lower). Can see case facts only (& precedent).
  - Law-finding courts (higher). Can see case facts & outcome.
    - (Set aside problem of identifying case facts.)
- In each period:
  - Lower courts hear continuum of cases over R (uniform – improper).
  - Higher courts can select one set of case facts to grant *certiori* at cost k.
    - Idea: higher court is resource constrained.
    - In judgement, reveals decision and outcome – this is precedent.

# JUDICIAL PREFERENCES

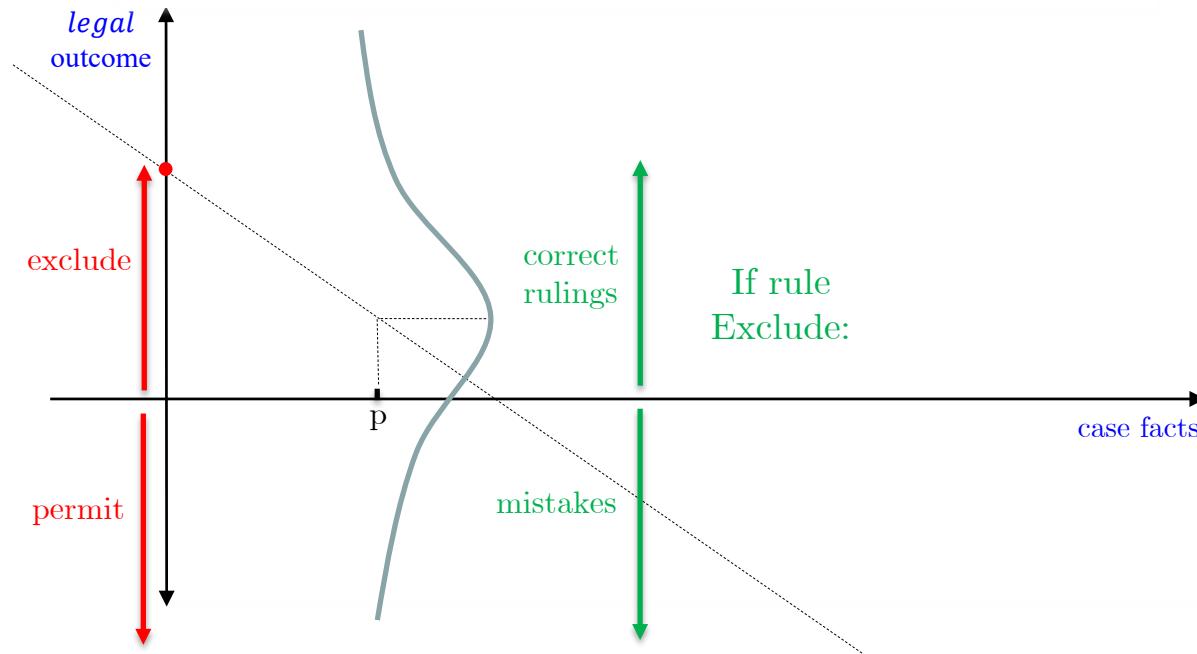


- Common preferences. Threshold is 0 between decisions:

$$u(p) = \begin{cases} |m(p)| & \text{if correct decision made} \\ -|m(p)| & \text{if incorrect decision made} \end{cases}$$

- “Close calls” vs. “slam dunks.”
- Many applications beyond courts – see Ellison and Holden (2014).
  - Very nice paper – attack similar problem with communication difficulty between levels.

# HOW DO LOWER COURTS DECIDE?



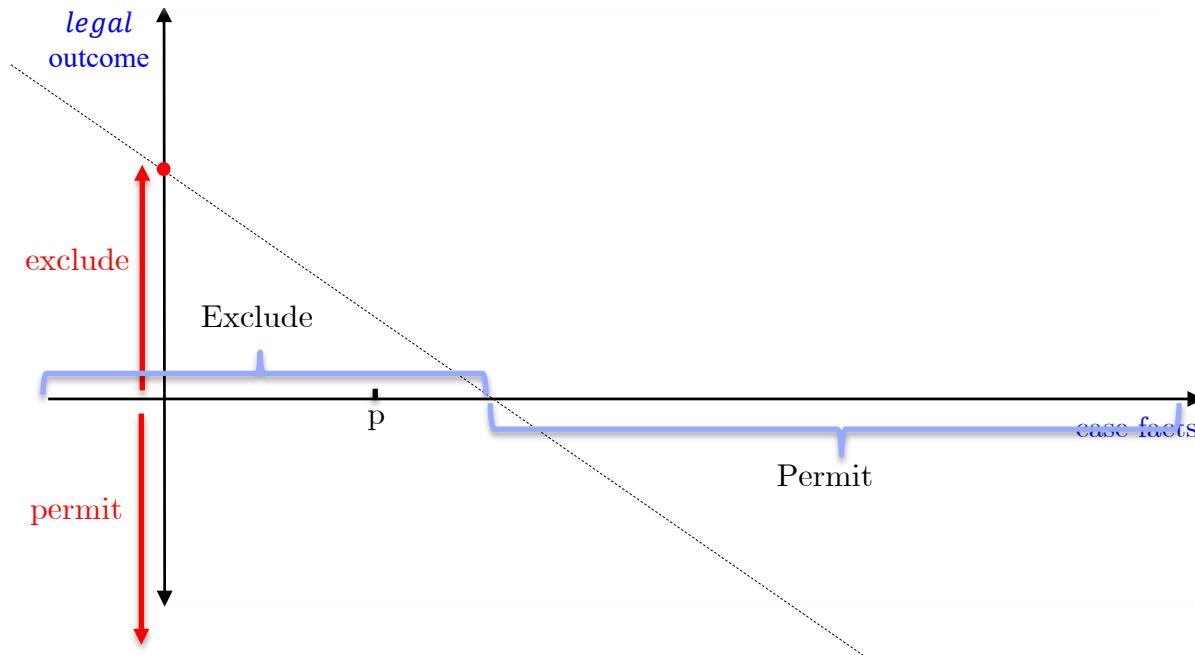
- Expected utility from ruling Exclude for case facts  $p$ :

$$u(E|p) = \int_0^{\infty} z \cdot \phi(z) dz + \int_{-\infty}^0 -z \cdot \phi(z) dz > 0 > u(P|p)$$

where  $\phi(z)$  is the pdf of the normal.

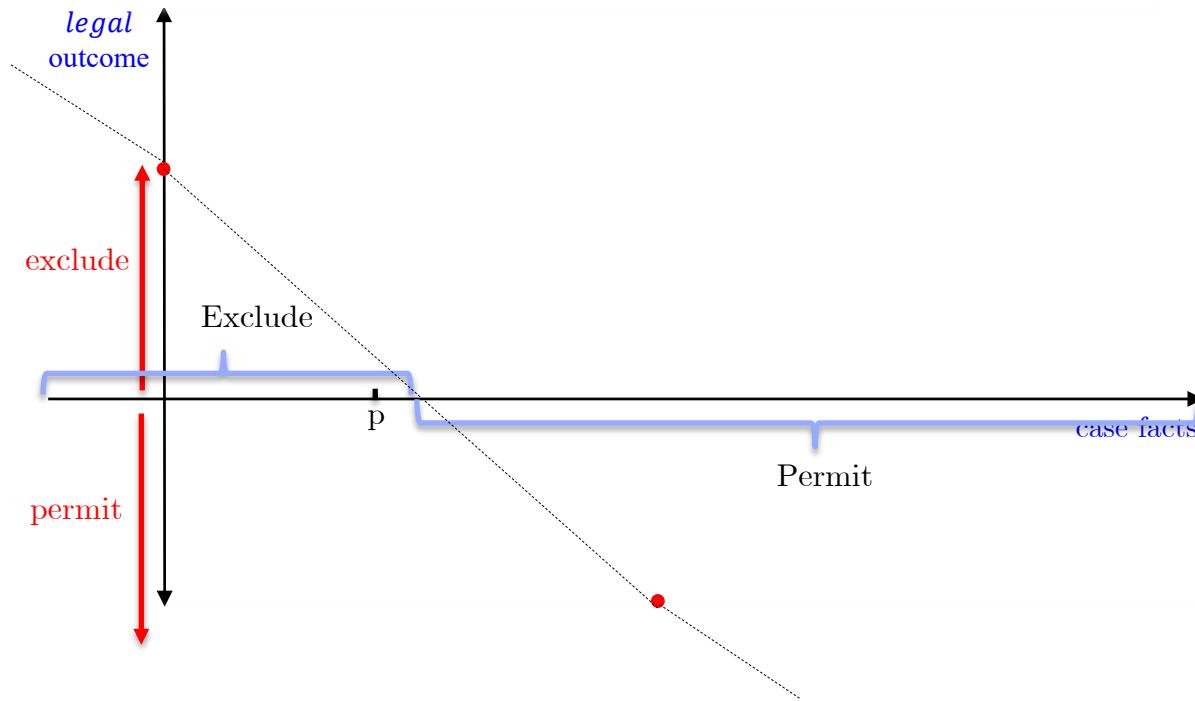
- Follow expected outcome (as variance symmetric and utility linear).
- Same holds for cases between precedents (on a Brownian bridge).

# HOW DO LOWER COURTS DECIDE?



- For “exploratory” cases, the lower court judges use precedent + theoretical knowledge (drift) to determine optimal ruling.

# HOW DO LOWER COURTS DECIDE ON A BRIDGE?



- For “exploitative” cases (between precedents), lower court judges look to the nearest precedent in either direction and take a weighted average by distance.
  - No role for drift term.
- Is this what real judges do?

# WHAT DO JUDGES DO?

- Lower court decision making is rather obvious
  - (At least once you understand the Brownian motion.)
- But its simplicity is appealing: intuitive & addresses a long-standing question.
- Lawyers have had a long (heated) argument about what judges do.
  1. Legalist view. Judges use precedent and reason by analogy:

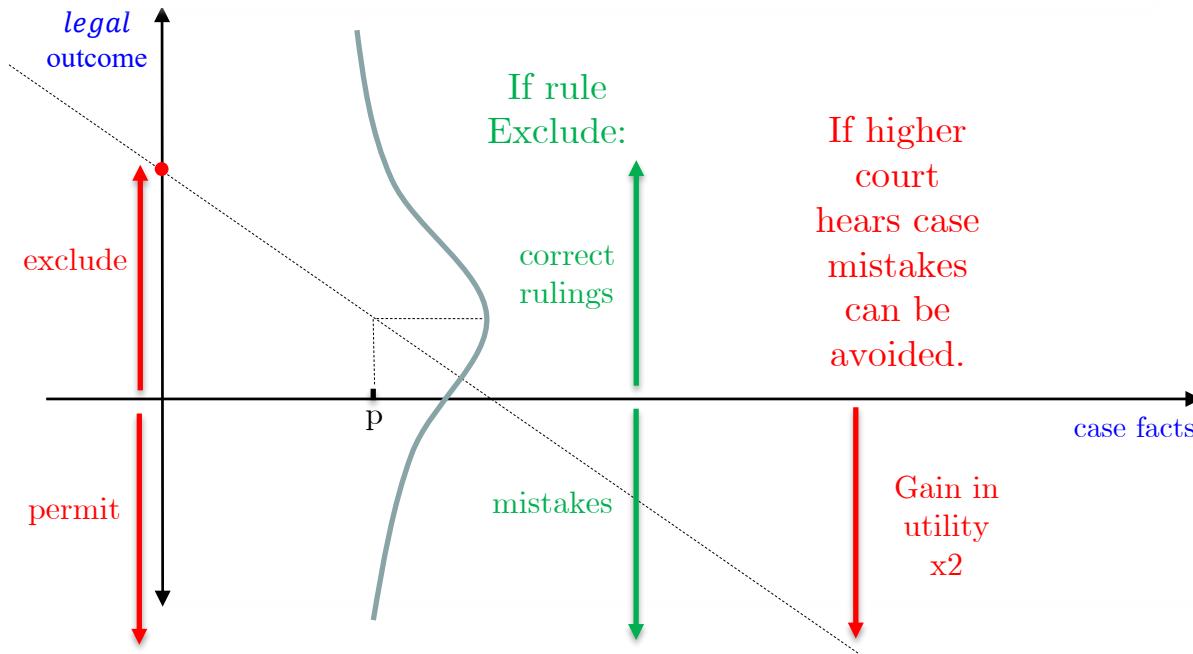
“Judges decide cases through systematic application of the external, objective sources of authority (precedent) that classically comprise the law.” (Cross 2003)
  2. Legal formalism / rule-based legal reasoning.

Judges reason deductively from legal rules. In its purest form, operates without reference to prior cases—the judge simply applies the governing legal rule to the case at hand.
- Critics & supporters of both sides.
  - “the legalist view suffers from theoretical and empirical indeterminacy.” (Cross 2003).  
“It is a fantasy.” (Alexander 1996)
  - “Analogy is a vital tool in legal reasoning.” (Murray 1992) “the life of the law has not been logic: It has been experience.” (Holmes 1881) “Lawyers are not able to explain with full specification of the theory that accounts for those beliefs. Lawyers (and almost all people) typically lack any large scale theory.” (Sunstein 1993)

# REASONING BY ANALOGY

- The Brownian motion rationalizes and reconciles these approaches.
- In the model, judges reason by analogy from the nearest precedent in either direction.
  - Lawyers refer to these as “controlling” or “guiding” precedents.
  - This is rational given the Markov property of the Brownian motion.
  - In practice, judges use a “balance test” to weight the precedents... as they do in the model.
  - Theoretical knowledge undergirds these practices – so reasoning by analogy is not atheoretical.
  - These are not mutually exclusive – it can be logic and experience contra Holmes: “the life of the law has not been logic: It has been experience.” (Holmes 1881)
- This is an organizational contribution rather than a deep one – but this type of result is a valuable contribution of theory.
  - (Perhaps one of the most valuable? There is no shortage of confusion in the world!)

# THE HIGHER COURT'S PROBLEM



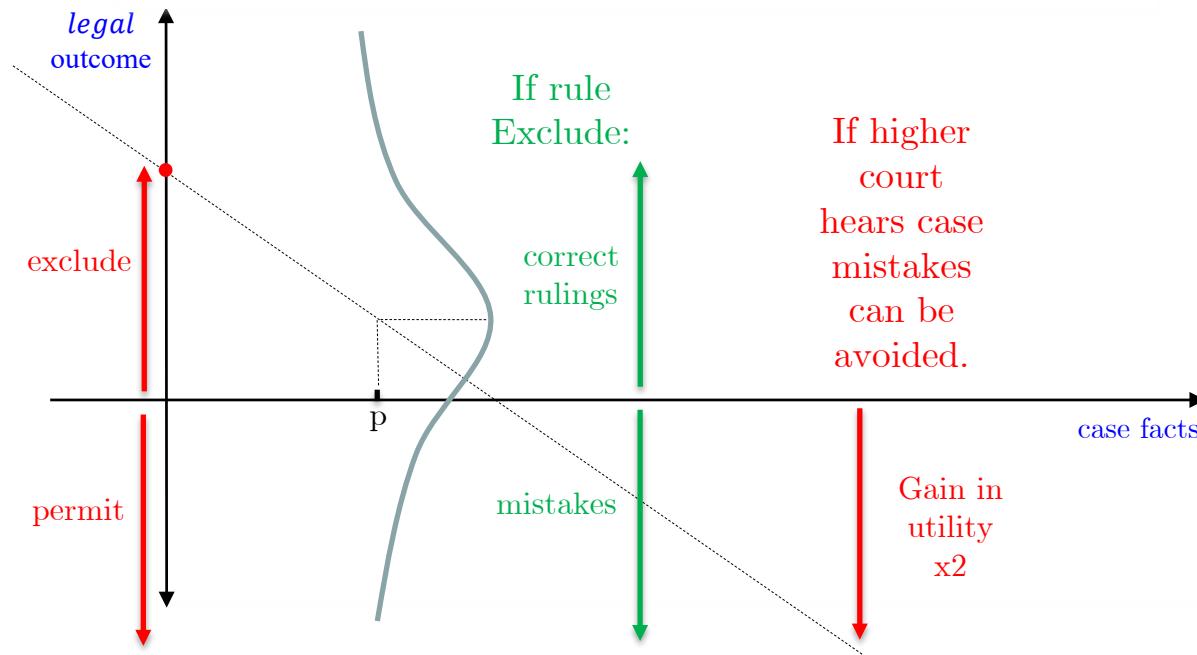
- If the Higher court hears case facts  $p$ , the expected value is:

$$U(p) = \int_0^\infty z \cdot \phi(z) dz + \int_{-\infty}^0 z \cdot \phi(z) dz = Em(p)$$

where  $\phi(z)$  is the pdf of the normal.

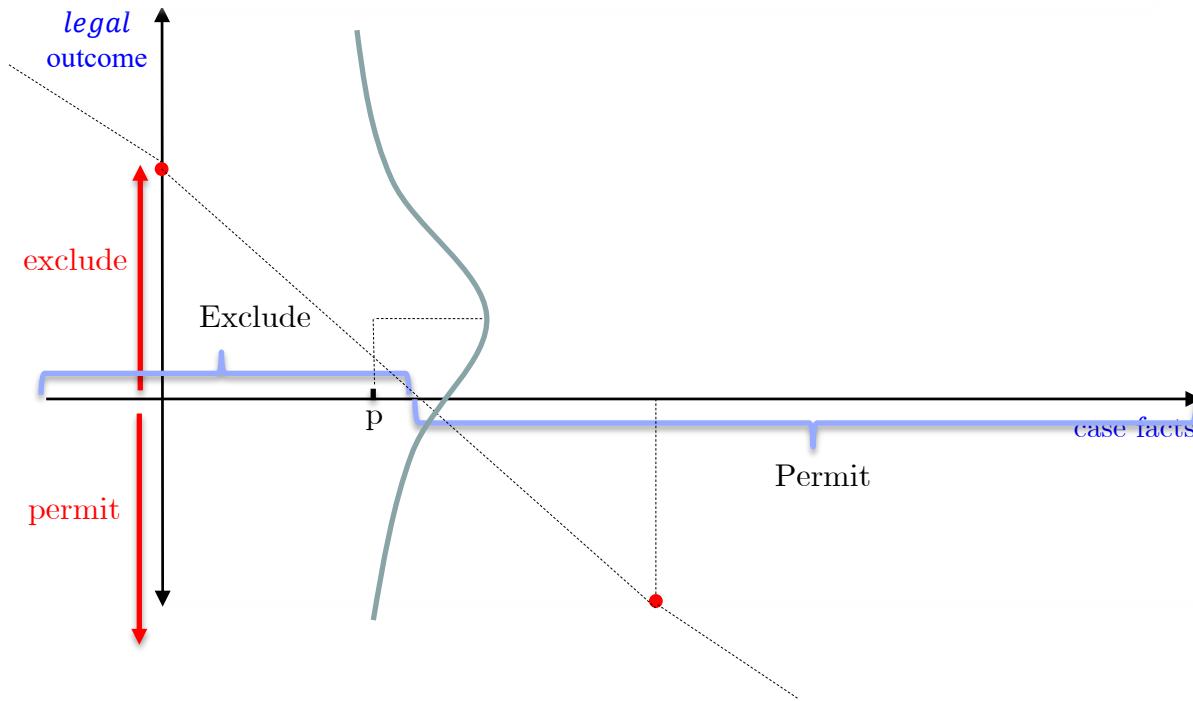
- Benefit of high court hearing a case is increasing in variance of uncertainty.
- Pushes court toward case of maximum "outcome uncertainty."
- But the cases vary in expected outcome and judges care about "error uncertainty."

# THE HIGHER COURT'S PROBLEM



- The maximum value for a single case is any case with an expected outcome of 0.
  - Regardless of the variance.
- So should the court hear a case of maximum “error uncertainty”?
- No → Informational spillover.

# INFORMATIONAL SPILLOVERS

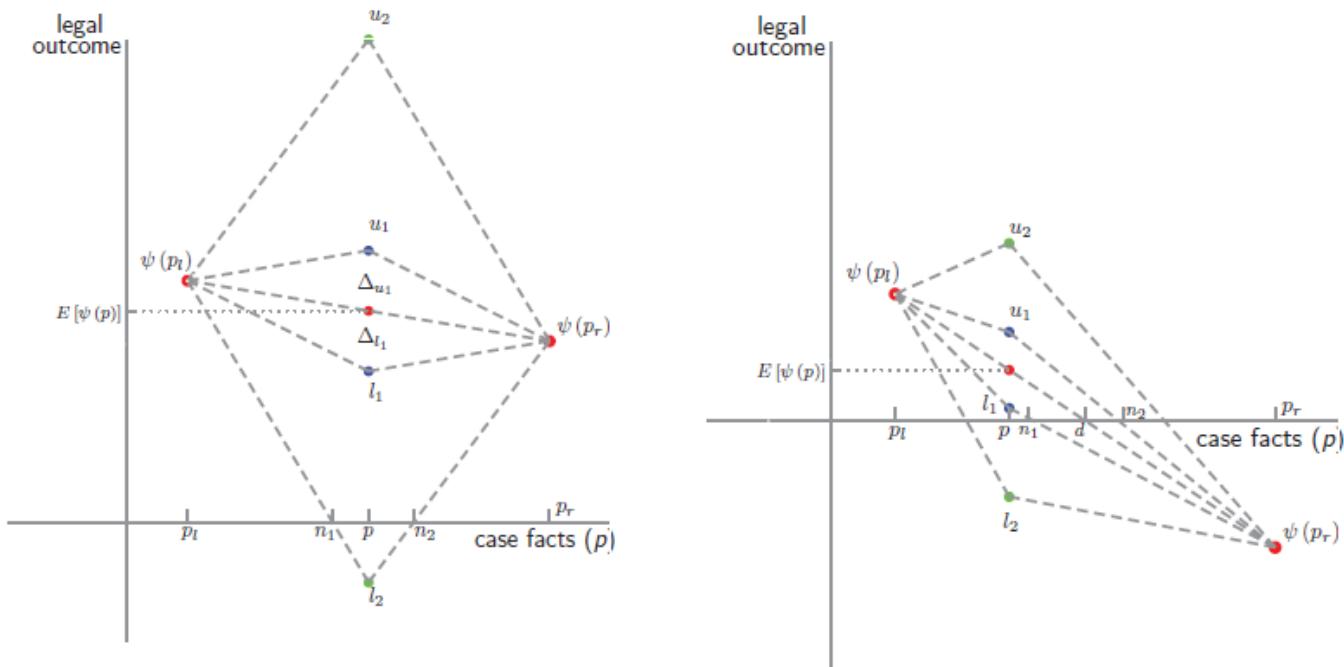


- For “exploitative” cases, hearing case facts  $p$  changes beliefs over all cases between the guiding precedents.
- All else equal, variance reduction is good. The total variance is reduced the most by the case of maximum “outcome uncertainty” – i.e., the case in the middle of the bridge.
- In choosing a case, the Higher Court needs to evaluate the broader impact.
  - Integrate over possible outcomes of all case facts that are affected.

# CHOOSING A CASE TO HEAR

- The

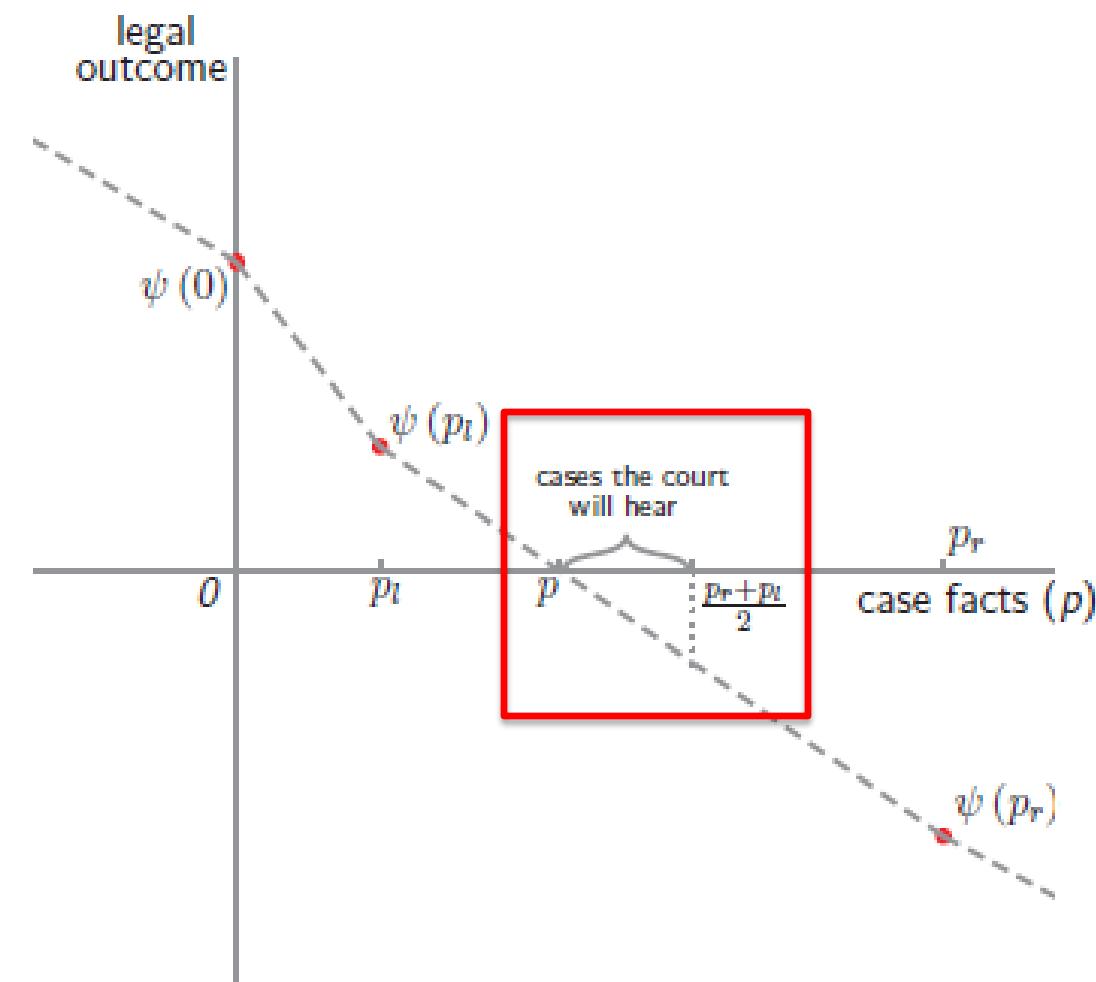
**FIGURE 3. The Source of Value**



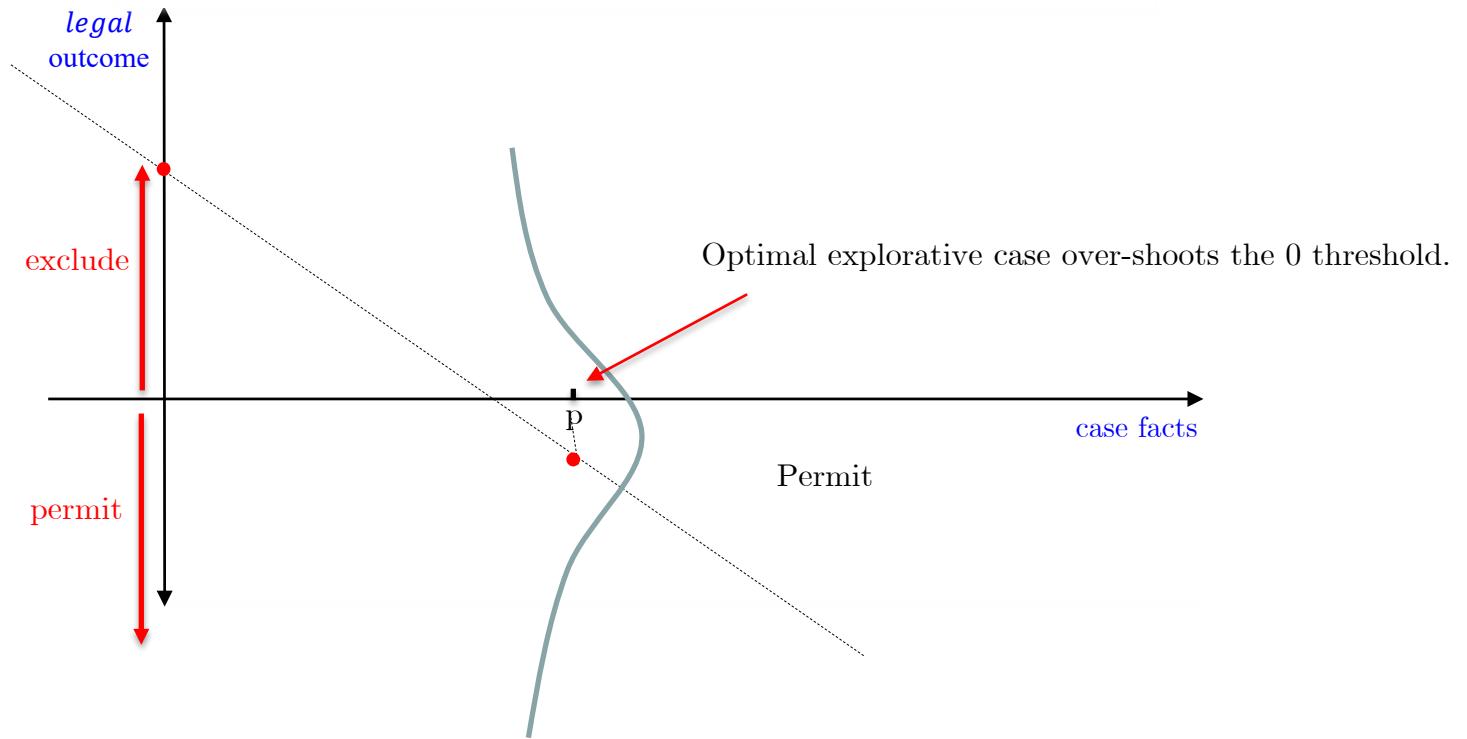
*Notes:* Each panel shows possible scenarios that could result if a new exploitative case  $p$  is heard. Precedents exist at points  $p_l$  and  $p_r$ . The new case  $p$  may have a legal outcome anywhere on  $\mathbb{R}$ , and the points  $u_1$ ,  $u_2$ ,  $l_1$ , and  $l_2$  illustrate possible realized legal outcomes. The dashed lines show expected legal outcomes for other new cases  $p' \neq p$  after case  $p$  is heard.

## CHOOSING CASES TO HEAR II

**FIGURE 4. Cases the Higher Court will Hear Among Exploitative and Standard Cases**



# OPTIMAL EXPLORATIVE CASE



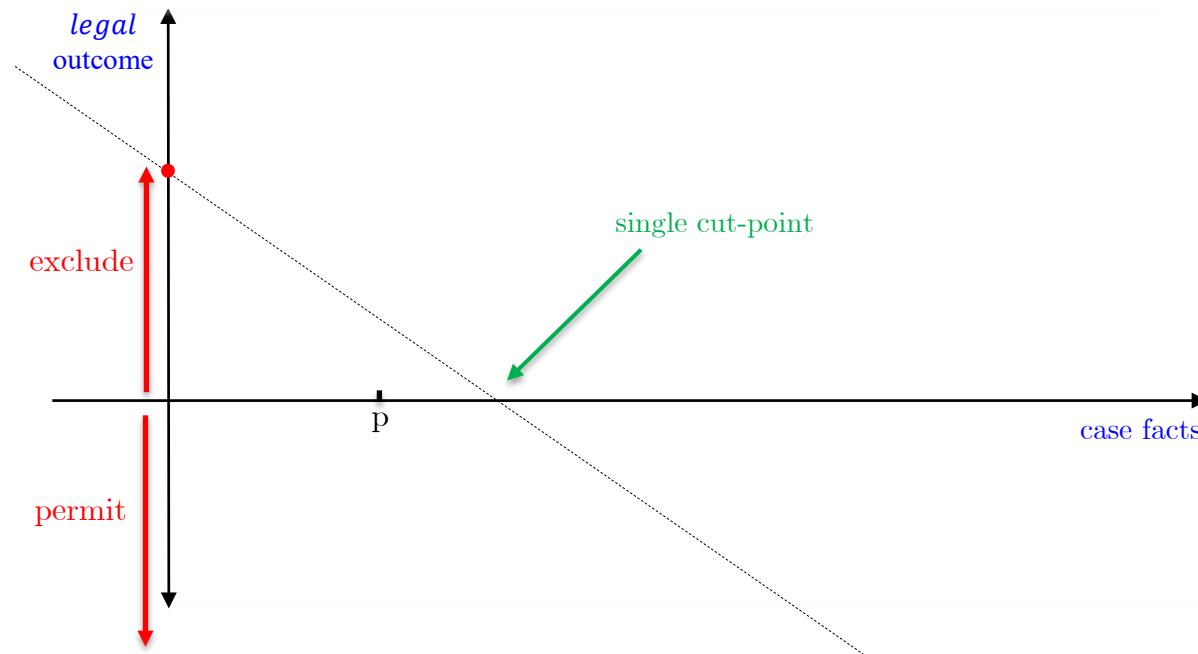
- For explorative cases (only one guiding precedent), the optimal case of the Higher court to hear has an expected outcome less than zero (for a guiding precedent with outcome above zero).

# THE DYNAMIC PATH OF JUDICIAL PRECEDENT

- Precedent accumulates over time following no fixed pattern of cases
  - E.g., exploitative → explorative → standard → non-standard.
  - Depends on outcomes realized.
- If cost to higher court of hearing a case is positive,  $k > 0$ , learning will stop with probability one.
- **Proposition 5:** With probability one, the Higher Court stops hearing new cases in finite time.
- Learning is not complete.
- Unbounded degree of “outcome uncertainty” remains.
- But “error uncertainty” → 0.

# DOCTRINAL COMPLEXITY

- How complex is the law?
- How many cut-points in the space of case facts are sufficient to fully prescribe optimal behavior by the Lower Courts?
- Define this as “Doctrinal Complexity.”
- At the start of play, with only the “case of first instance”, doctrinal complexity = 1.

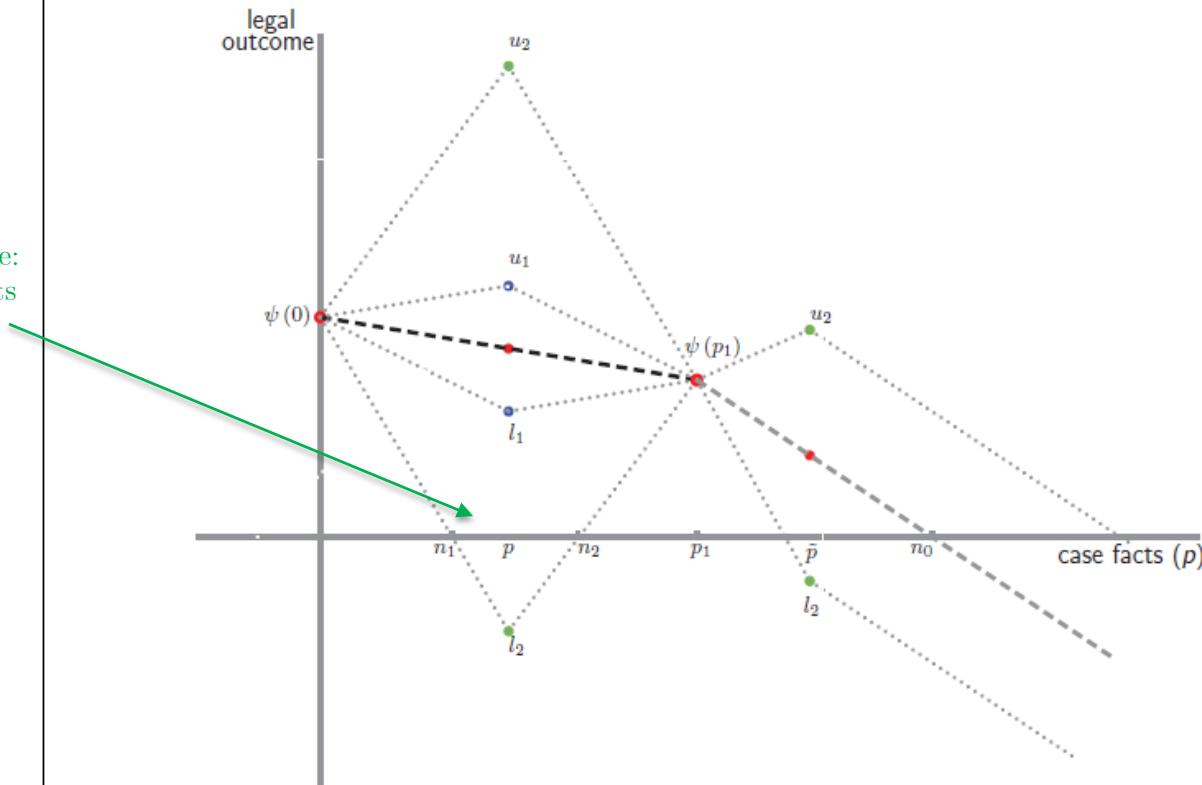


# DOCTRINAL COMPLEXITY

- Doctrinal Complexity can only increase (in this model).
- Evolution of doctrinal complexity depends on the type of cases heard by the Higher Court.

**FIGURE 5. Example of Cases the Higher Court can Consider as Precedent Accumulates**

Non-standard case:  
guiding precedents  
have same  
judgement.



*Notes:* The precedents are the first period precedent,  $(p_1, \psi(p_1))$ , and the initial precedent  $(0, \psi(0))$ . The other marked points correspond to potential legal outcomes for cases  $p$  and  $\bar{p}$  that may be heard in the second period and the dotted lines corresponds to new expected legal outcomes that result.

# DOCTRINAL COMPLEXITY

- Doctrinal Complexity can only increase (in this model).
- Evolution of doctrinal complexity depends on the type of cases heard by the Higher Court.

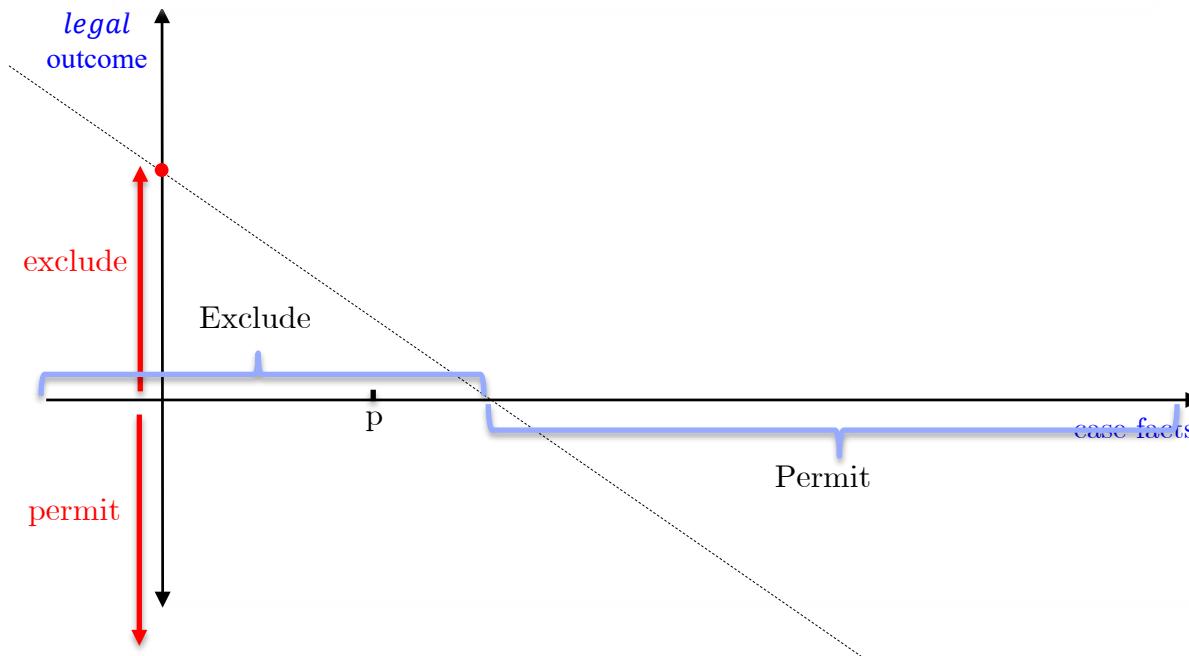
**Proposition 4:** Doctrinal complexity weakly increases over time.

- When the Higher Court hears a nonstandard case, doctrinal complexity increases by two with probability between 0 and 1/2, otherwise it remains unchanged.
  - When the Higher Court hears a standard case, doctrinal complexity does not change. Nevertheless, the doctrinal cut point changes with probability one.
- 
- Lax (2007) limits attention to “proper rules” with a single cut-point.
    - This is equivalent to doctrinal rules of 1 – the simplest possible doctrine.
  - Increases in Doctrinal Complexity are what lawyers call “carve outs”

# WHAT IS LEGAL DOCTRINE?

- “A **legal doctrine** is a framework, set of rules, procedural steps, or test, often established through precedent in the common law, through which judgments can be determined in a given legal case. A doctrine comes about when a judge makes a ruling where a process is outlined and applied, and allows for it to be equally applied to like cases. When enough judges make use of the process soon enough it becomes established as the *de facto* method of deciding like situations.”
- What does this mean?
- Lax (JOP, 2012): “*In my terminology, a doctrine or rule is any logical partitioning of cases into equivalence classes (yes vs. no, winner vs. loser, legitimate search vs. unreasonable search, etc.).*”
- A “doctrine” seems to be something more than what Lax suggests.

# DRIFT AS LEGAL DOCTRINE

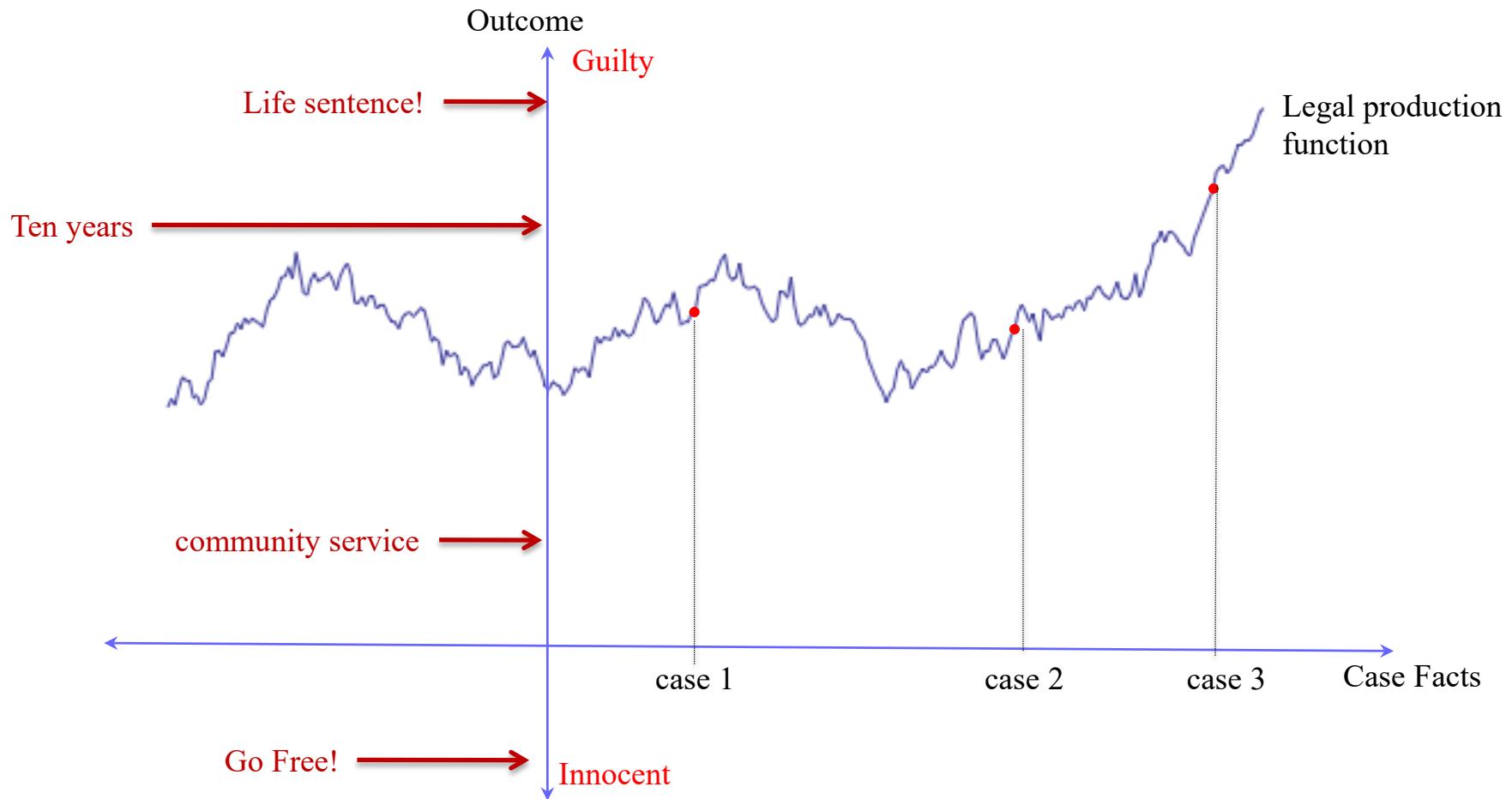


- For “exploratory” cases, the lower court judges use precedent + theoretical knowledge (drift) to determine optimal ruling..
- If the Lower Courts do not have the theoretical knowledge – do not know drift – then the Higher Court needs to communicate more than just outcomes.
  - Not necessary for exploitative cases (between two precedents).
- Is this legal doctrine?

# CONCLUSIONS

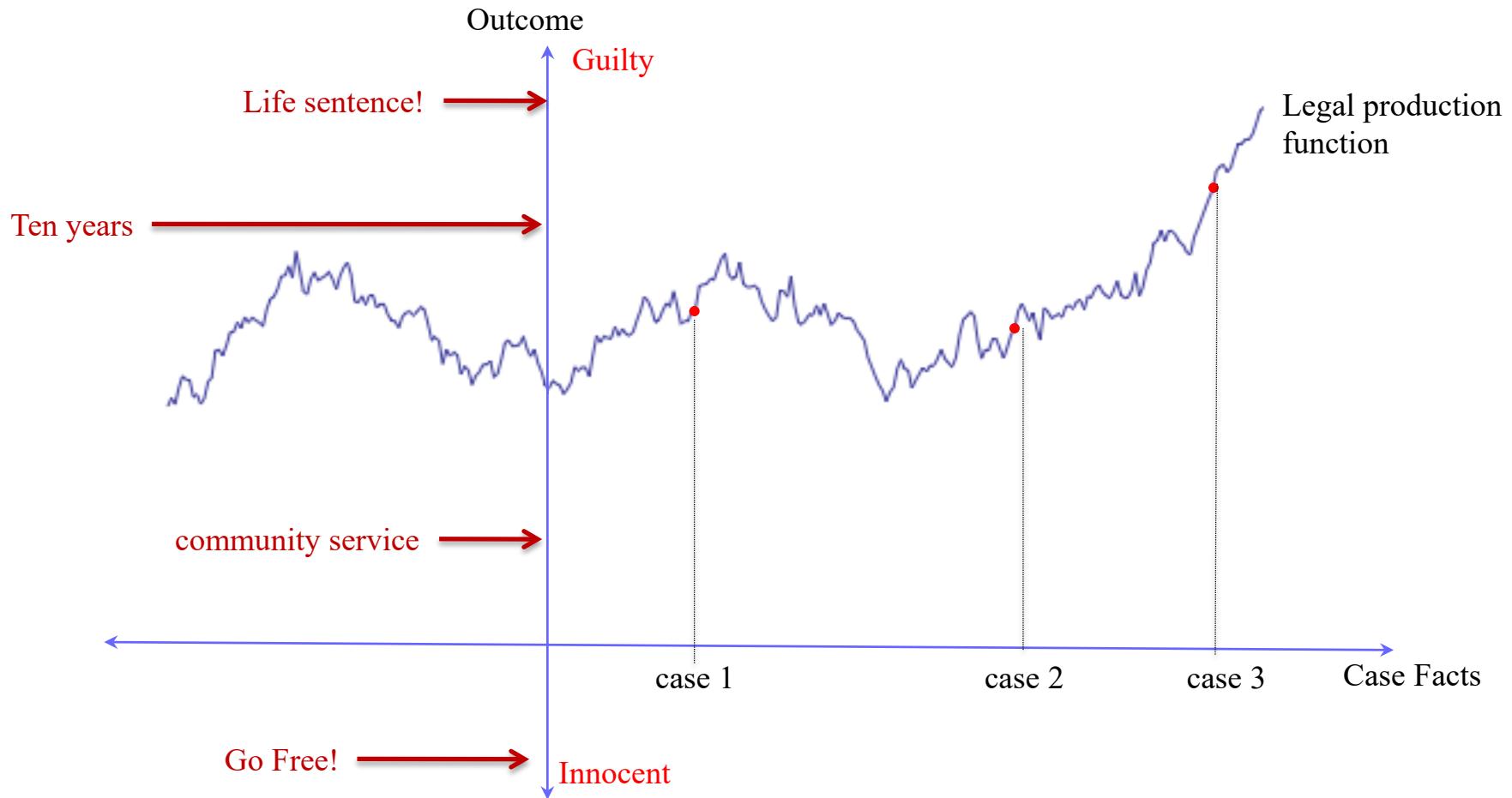
- Judges have a hard problem to solve.
- Why judge made law rather than (elected) legislature made law?
- Argument: Because we don't know what we want to do until we see it.
  - It is very difficult to specify complete laws ex ante.
  - (Not only because of the difficulty in writing complete contracts.)
- Legislatures leave it to judges to work it out.
  - So let's have a model of judges working it out (like this one!)
  - What does this mean for the laws legislatures should write? How much leeway should they leave for judges to work out? How much should legislation decree rules vs. discretion?
- + much more enrichment and insight into the model we have.

# BINARY OUTCOMES



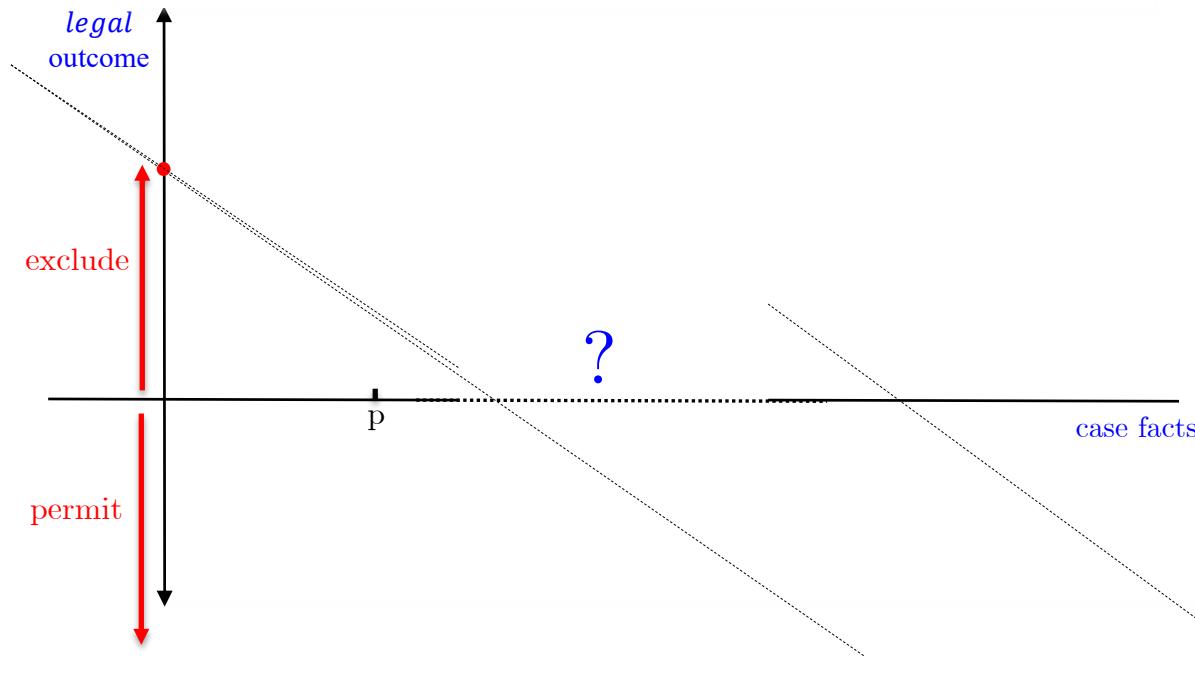
- Justices often have multiple rulings beyond win/lose. E.g., jail sentence for crime.
- How does that affect decision making?

# DIFFERENT PREFERENCES



- Justices do not always share the same preference for legal outcomes.
- Check out the U.S. Supreme Court fights!

# DISCOVERY BEYOND THE OUTCOME MAPPING



- Idea! Hear a new case and discover space in the case facts dimension.
- Is that what happened with the motor home case?
- Privacy ... now on internet. New issues that weren't anticipated.