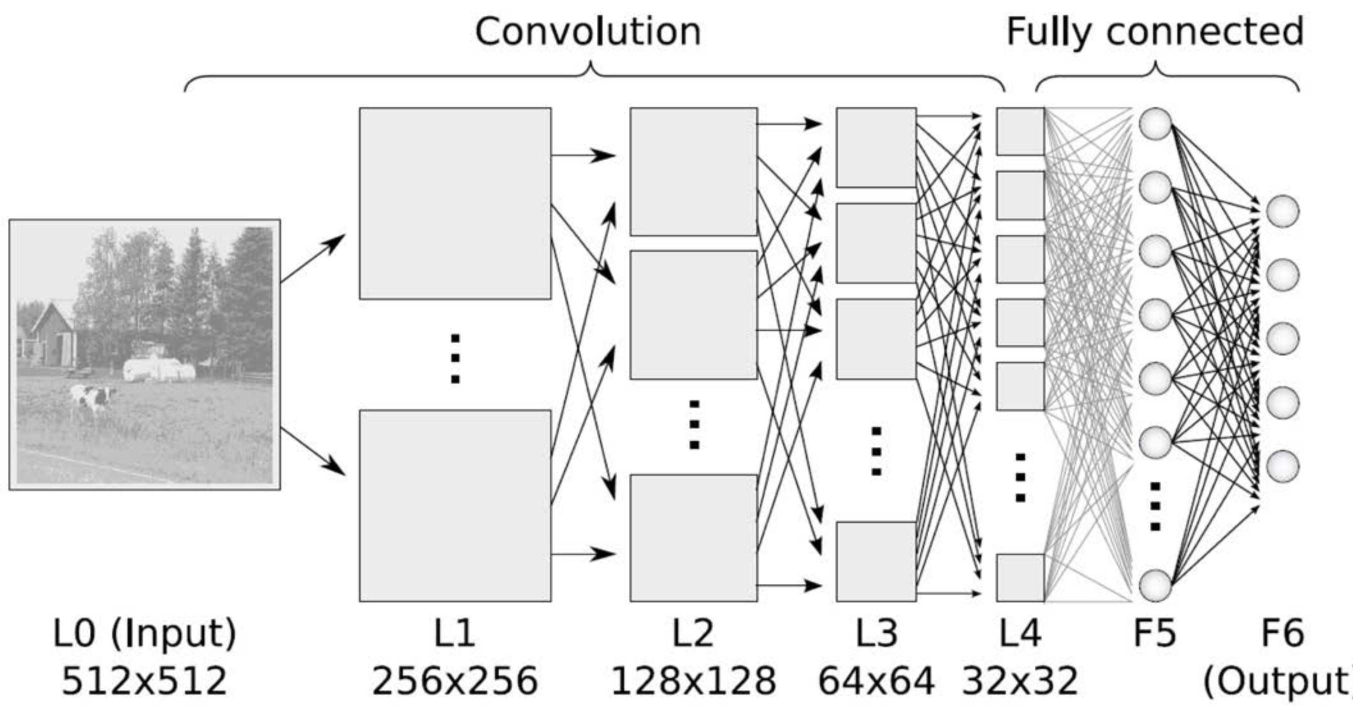
Artificial Intelligence, Scientific Discovery, and Commercial Innovation

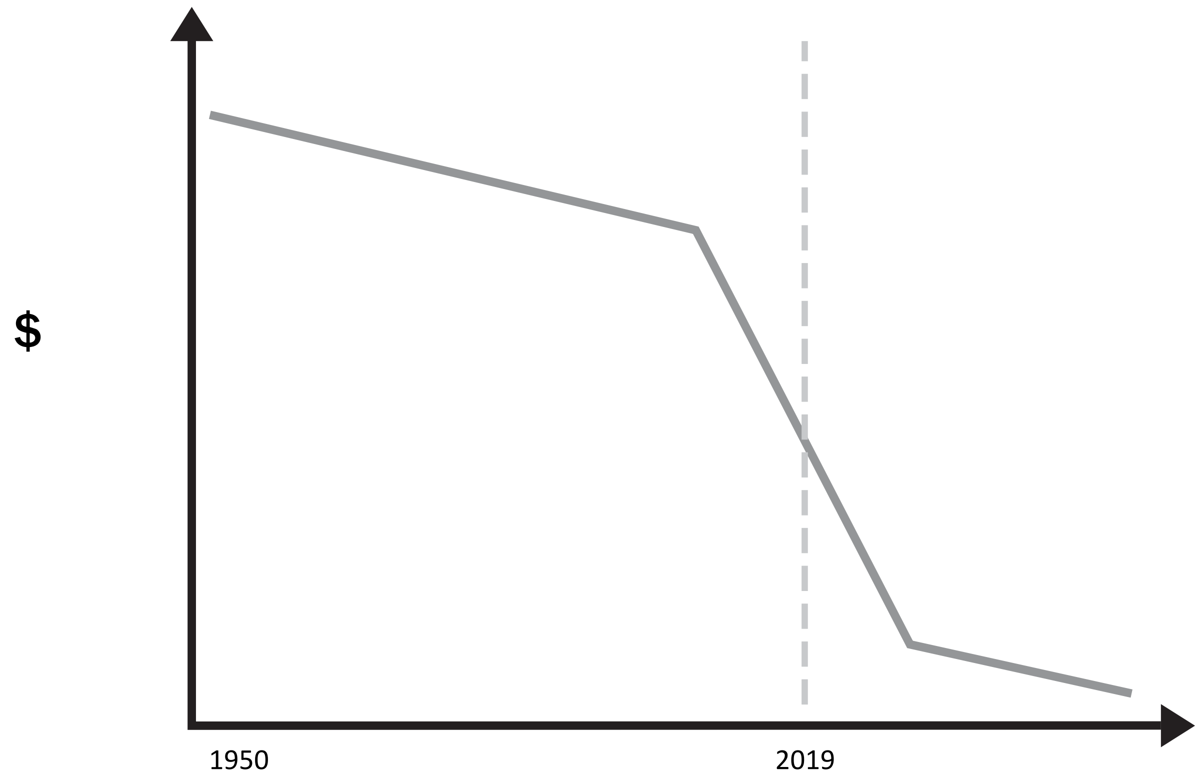
Ajay Agrawal (University of Toronto and NBER)
John McHale (National University of Ireland, Galway)
Alex Oettl (Georgia Institute of Technology and NBER)

July 2019

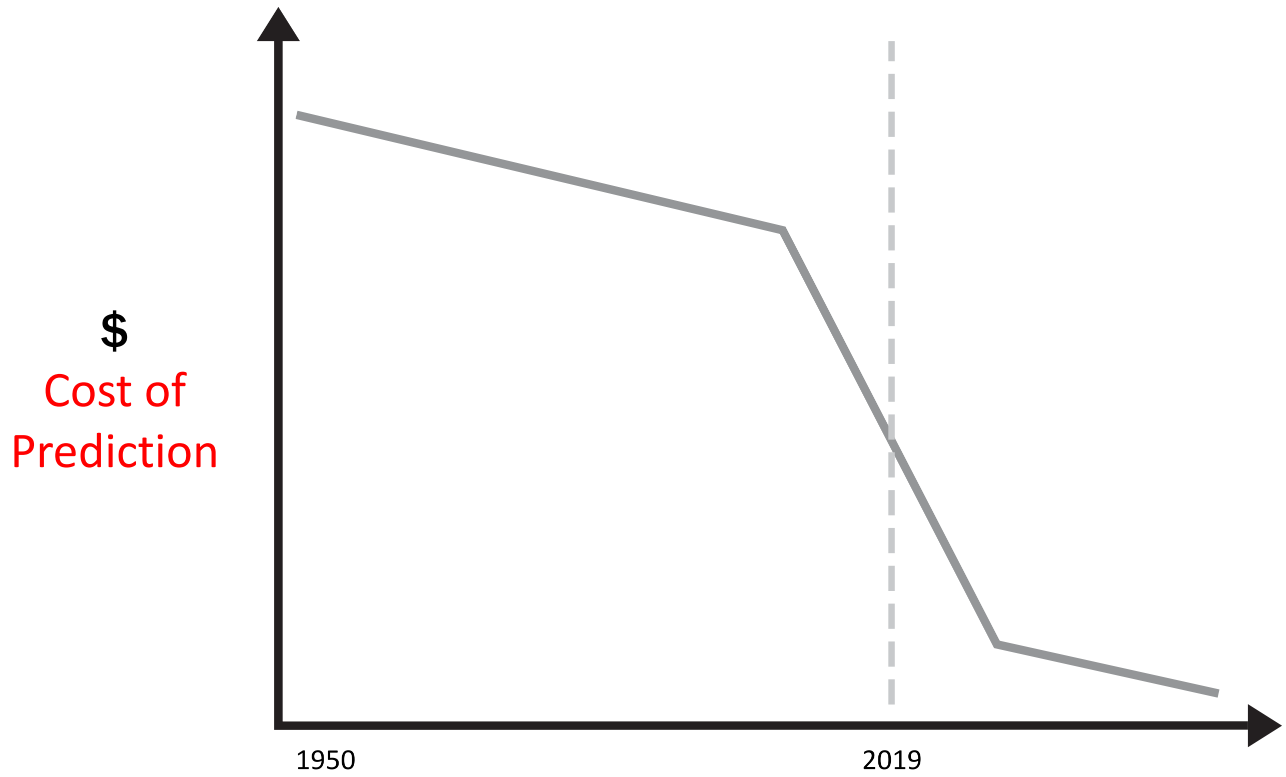


Source: https://www.ais.uni-bonn.de/deep_learning/

Artificial Intelligence



Artificial Intelligence



Prediction:

Using information that you do have to generate information that you don't have

TWO-THIRDS OF THE OPPORTUNITIES TO USE AI ARE IN IMPROVING THE PERFORMANCE OF EXISTING ANALYTICS USE CASES

In 69 percent of the use cases we studied, deep neural networks can be used to improve performance beyond that provided by other analytic techniques. Cases in which only neural networks can be used, which we refer to here as “greenfield” cases, constituted just 16 percent of the total. For the remaining 15 percent, artificial neural networks provided limited additional performance over other analytics techniques, among other reasons because of data limitations that made these cases unsuitable for deep learning.

Expanding Range of Use as Input

Expanding Range of Use as Input

Rising AI → Falling Cost of Prediction

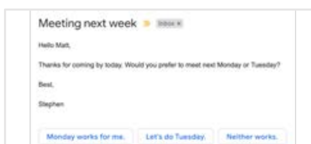
- Converting problems that historically were not considered to be AI problems into AI


Driving

Translation


Fraud detection


Car loans



Email replies


HR


Inspection

https://commons.wikimedia.org/wiki/File:Ultrasonic_pipeline_test.jpg

Knowledge, combinations, and innovation

$$\underset{\text{Innovation}}{I} = F(\underset{\text{Knowledge}}{N})$$

Innovation **Knowledge**

Knowledge, combinations, and innovation

$$I = F(N)$$

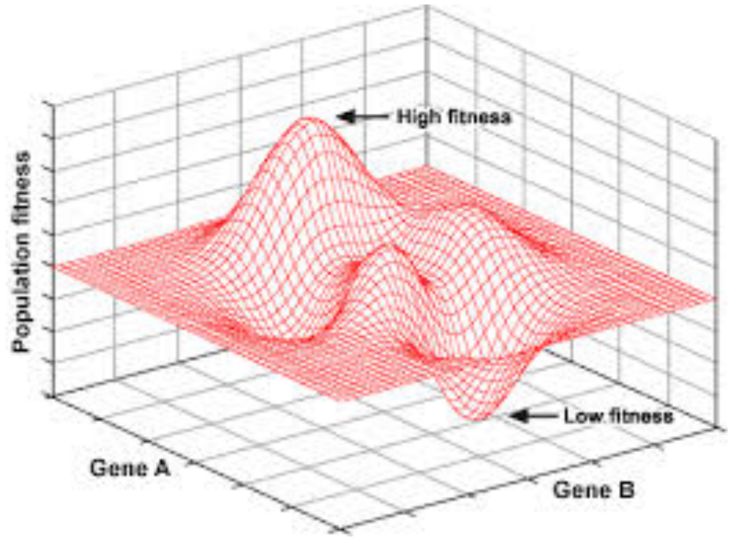
Innovation **Knowledge**

$$I = G(2^N)$$

Combinatorial view of innovation process

Conceptualisation of the innovation process

- Innovation as search over a potentially vast combinatorial search space
- Science as a map of “fitness landscapes”
- AI generates better maps



Science as a map

- Theory
- Simulation
- Data-based models
- AI

Traditional Science

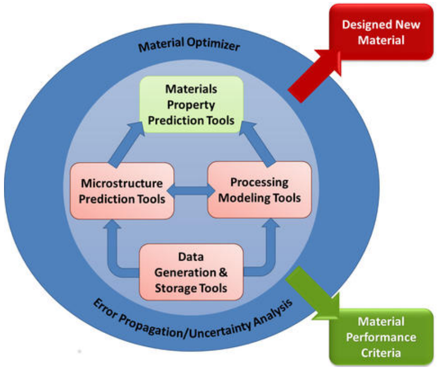
Examples of AI enhanced discovery

- New drug targets
  - AlphaFold (Google DeepMind)
    - Predict protein structures from amino asset sequences

- New small molecule drugs
  - Atomwise
    - Predict small molecule drugs that bind with target proteins

Atomwise
Better medicines faster.

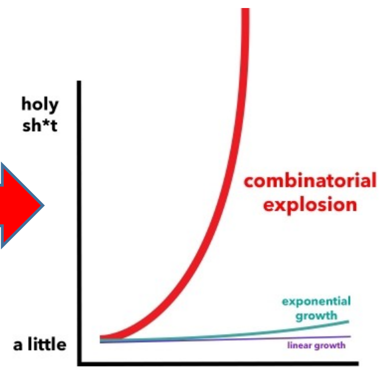
- New materials
  - Medical devices/Energy harvesting and storage
    - Predict properties of new molecules based on molecular descriptors



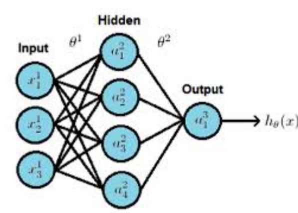
Generic workflow of science-based innovation

Combinations

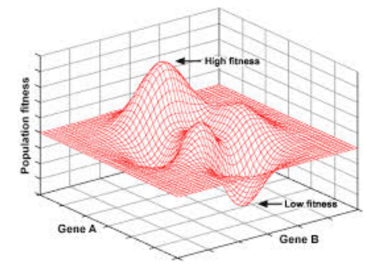
Knowledge

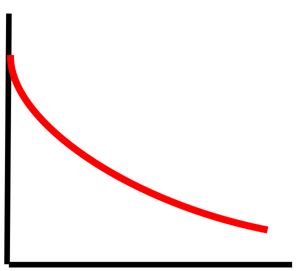


AI

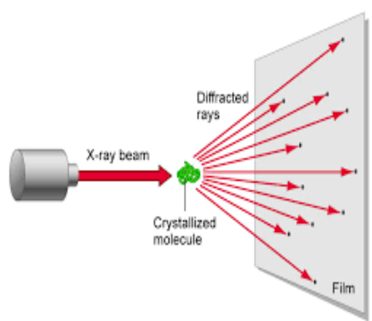


Fitness Landscape





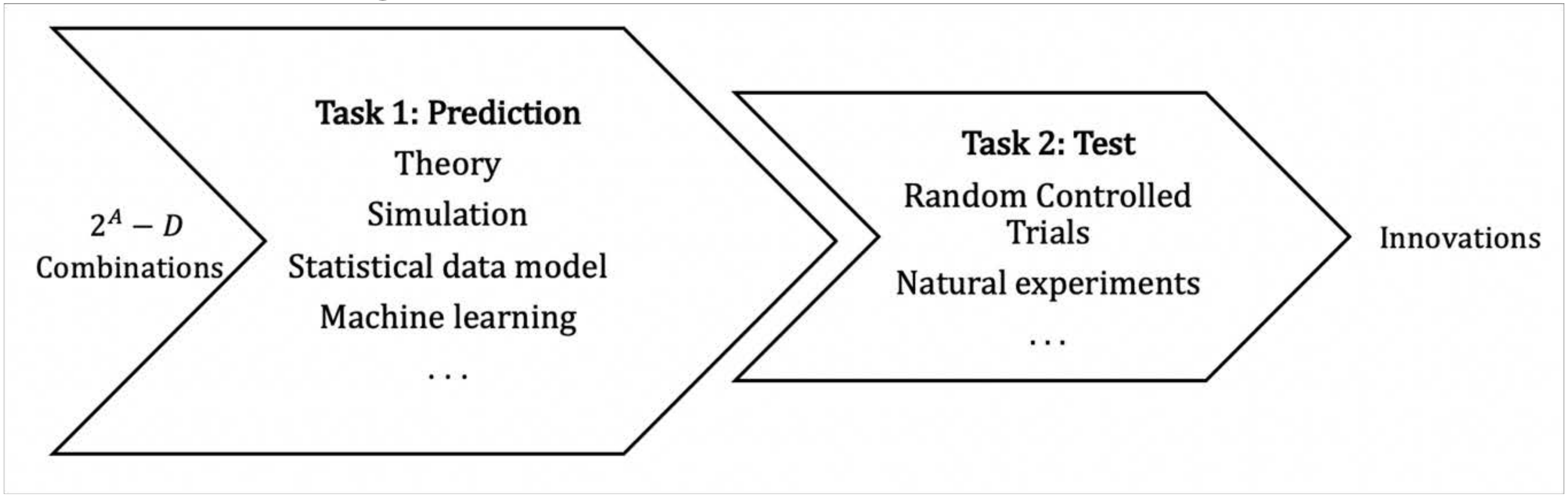
Ranking



Experiments

Innovation

Figure 5: Generic Workflow of the Two-Task Model



Search space

- Number of potential combinations: 2^A-D
  - A is # ideas the scientist has to combine into new ideas
  - D is # of observations on prior successes and failures

- Known
  - Scientist knows that G successes exist to be found
  - Share of combinations that will be a success: G/(2^A-D)

Modelling AI-aided innovation

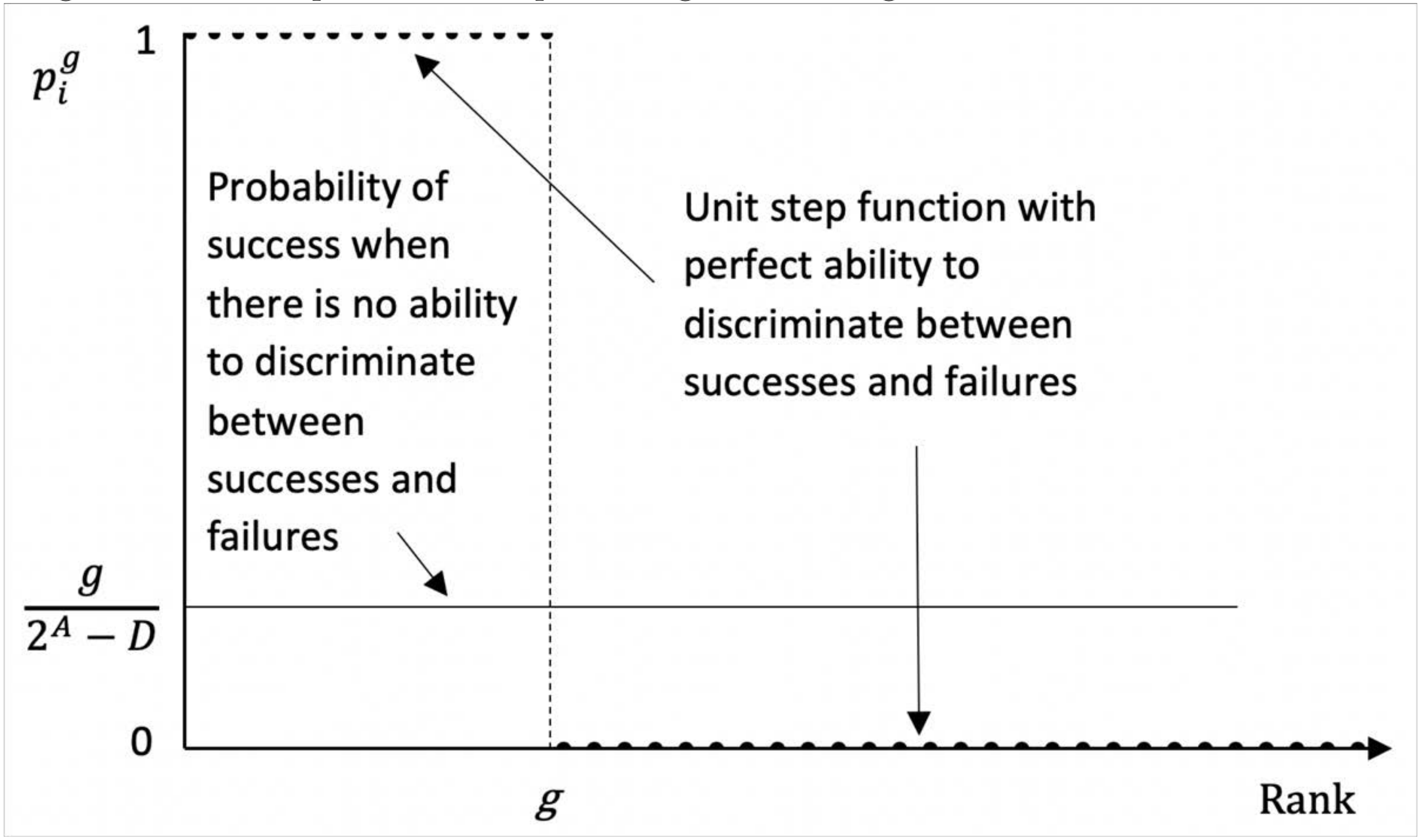
- A baseline model of exhaustive neighborhood search

- A two-task model
  - Task 1: Prediction
  - Task 2: Testing

- Introduce AI as a "shock" to the prediction task

- A multi-task model
  - The bottleneck problem
  - AI as a complement and substitute to R&D labor

Figure 1: Unit Step Function Representing the Ranking Function for the Ground Truth

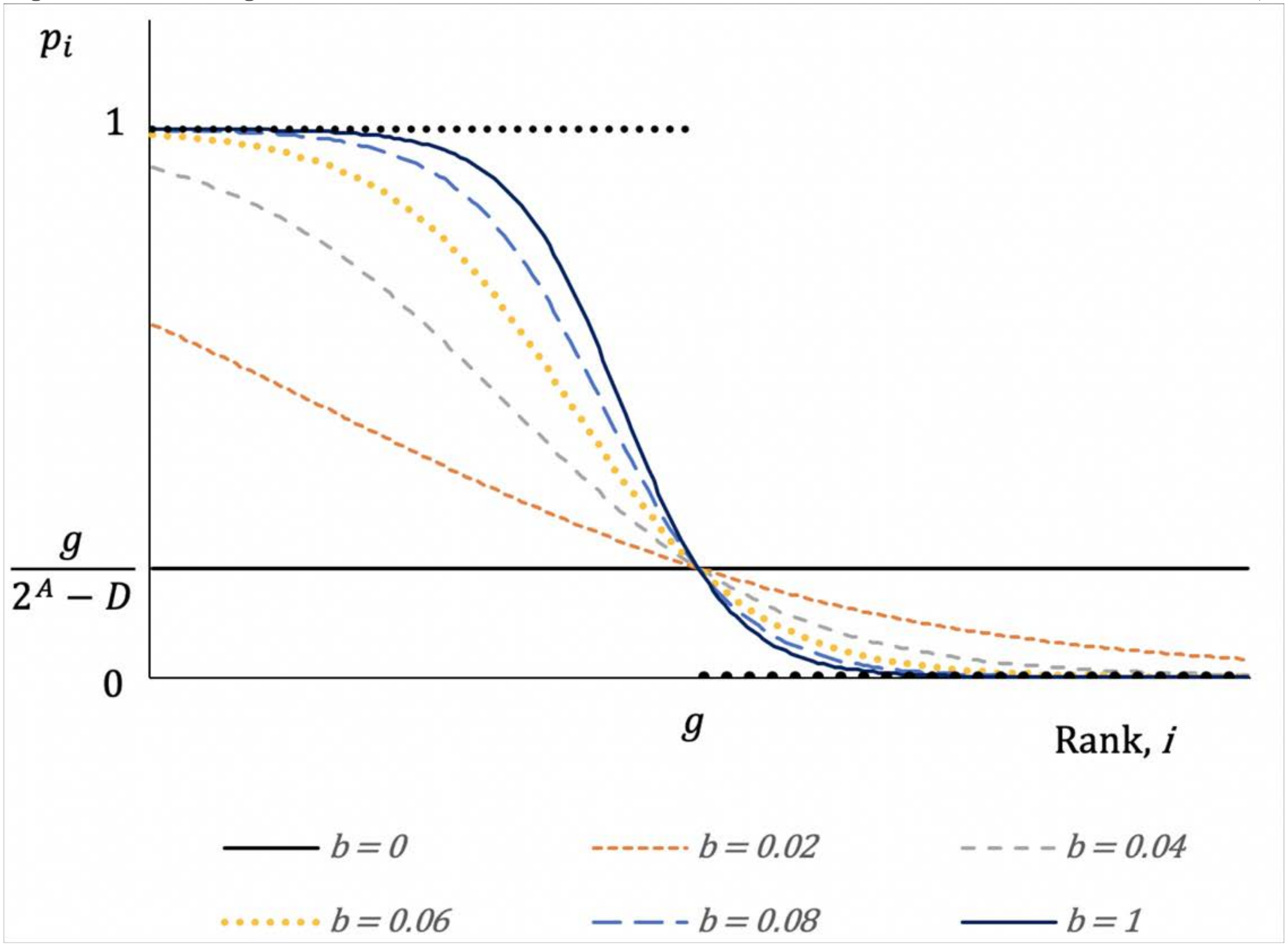


Ranking function

$$p_r = \frac{1}{1 + Ke^{b(r-G)}}$$

- Functional form:
  - Probability of discovering a success when the prediction model has zero discriminating power = $G/(2^A-D)$
  - Approaches ground truth as the model approaches perfect discrimination
    - Approach ground truth as b -> ∞

Figure 2: Ranking Function Curves for Different Values of the Discrimination Parameter, b



Optimal number of tests

$$MV_r^e = p_r \pi \qquad MC_r = c.$$

Optimal number of tests

$$MV_r^e = p_r \pi \qquad MC_r = c.$$

$$p_{r^*} = \frac{c}{\pi}.$$

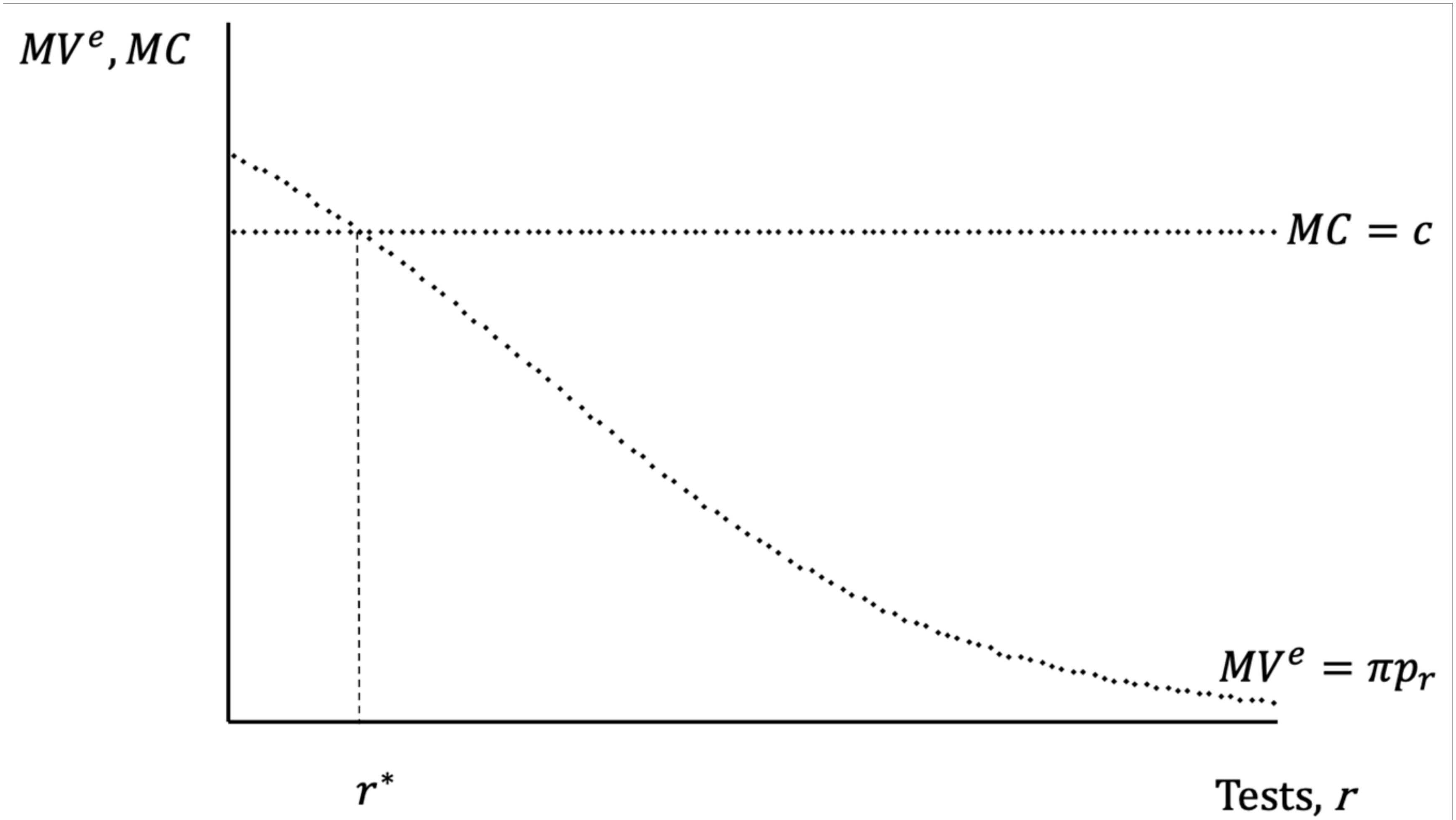
Optimal number of tests

$$MV_r^e = p_r \pi \qquad MC_r = c.$$

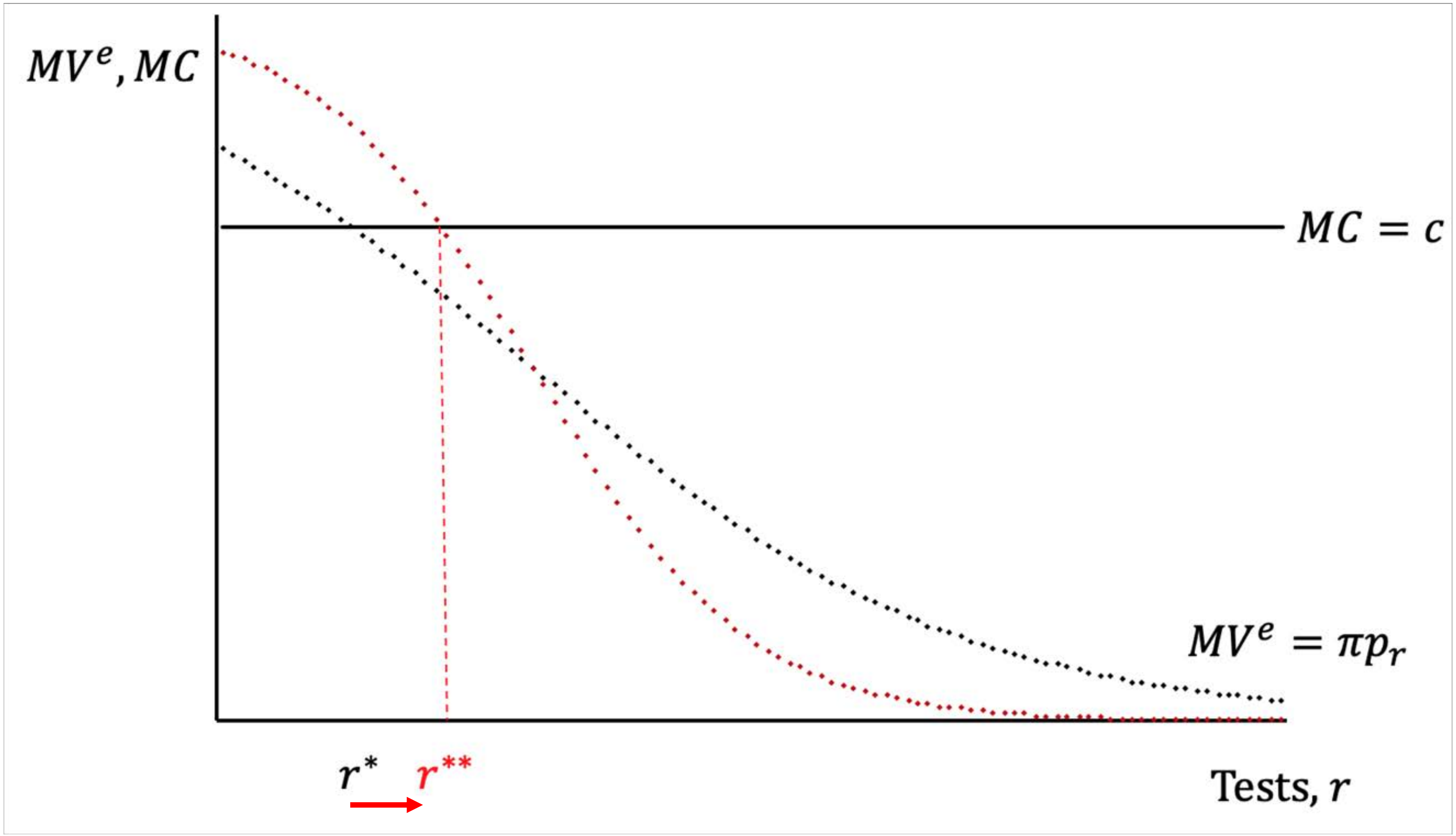
$$p_{r^*} = \frac{c}{\pi}.$$

$$\frac{1}{1 + \left(\frac{2^{A-D-G}}{G}\right) e^{b(r^* - G)}} = \frac{c}{\pi}$$

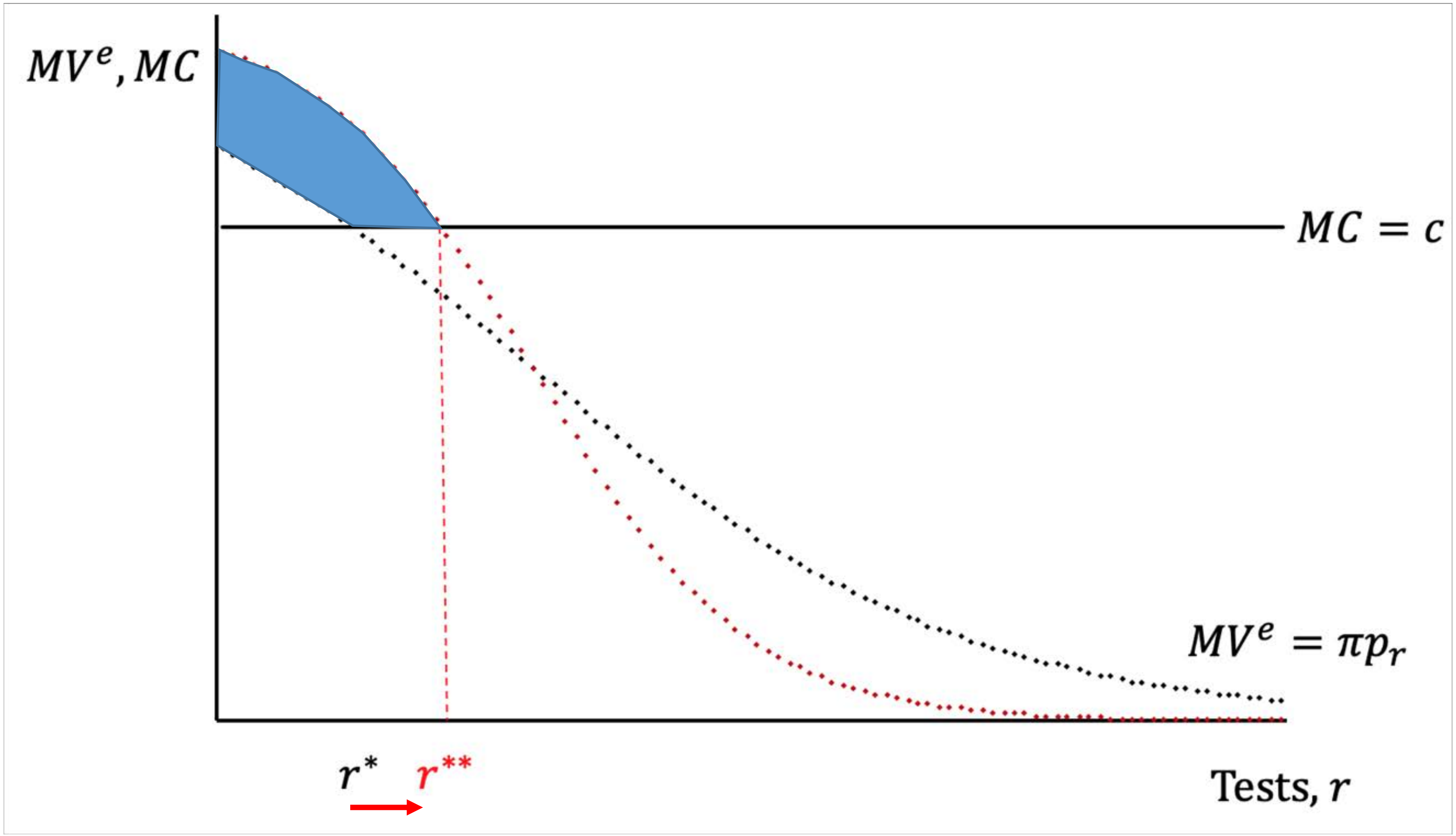
$$\Rightarrow r^* = G - \frac{\ln\left(\frac{2^{A-D-G}}{G}\right) - \ln(\frac{\pi}{c} - 1)}{b}$$



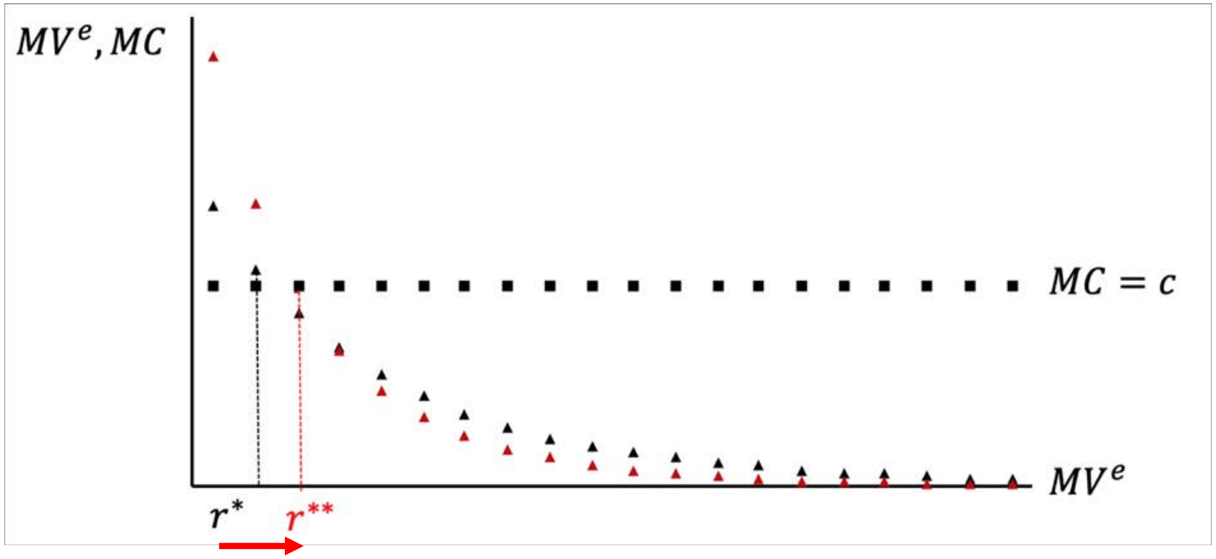
(a) Determination of the Optimal Number of Tests



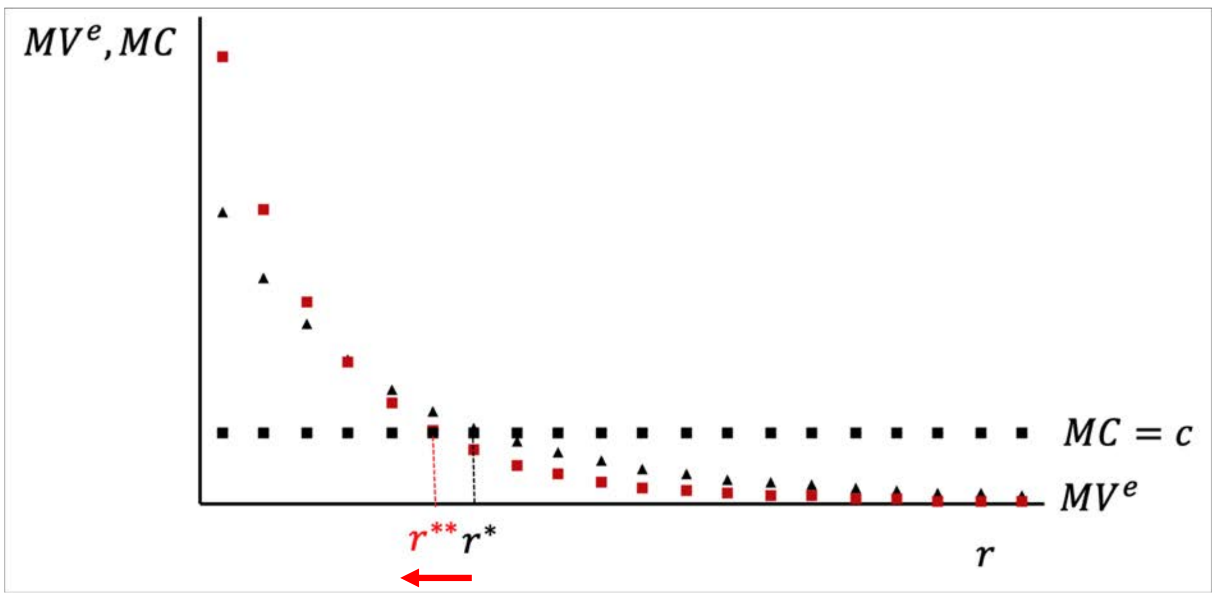
(b) Impact of an Improvement in the Prediction Model on the Optimal Number of Tests



(b) Impact of an Improvement in the Prediction Model on the Optimal Number of Tests



(b) Impact of an Improvement in the Prediction Model when the Innovator has a Single Innovation Target and the Crossover Probability is below c



(c) Impact of an Improvement in the Prediction Model when the Innovator has a Single Innovation Target Crossover Probability is above c

Next steps

- Theory
  - Multi-task innovation process
    - Substitute or complement to R&D labour
  - Closed-loop innovation processes
    - Choosing the next experiment
  - Endogenous growth from data spillovers
    - Optimal number and type of tests also considers the value of data spillovers (success/failure feedback data)

- Empirics
  - Testable hypothesis
    - AI increases the productivity of the innovation process
      - Look for exogenous variation in access to AI

Thank you