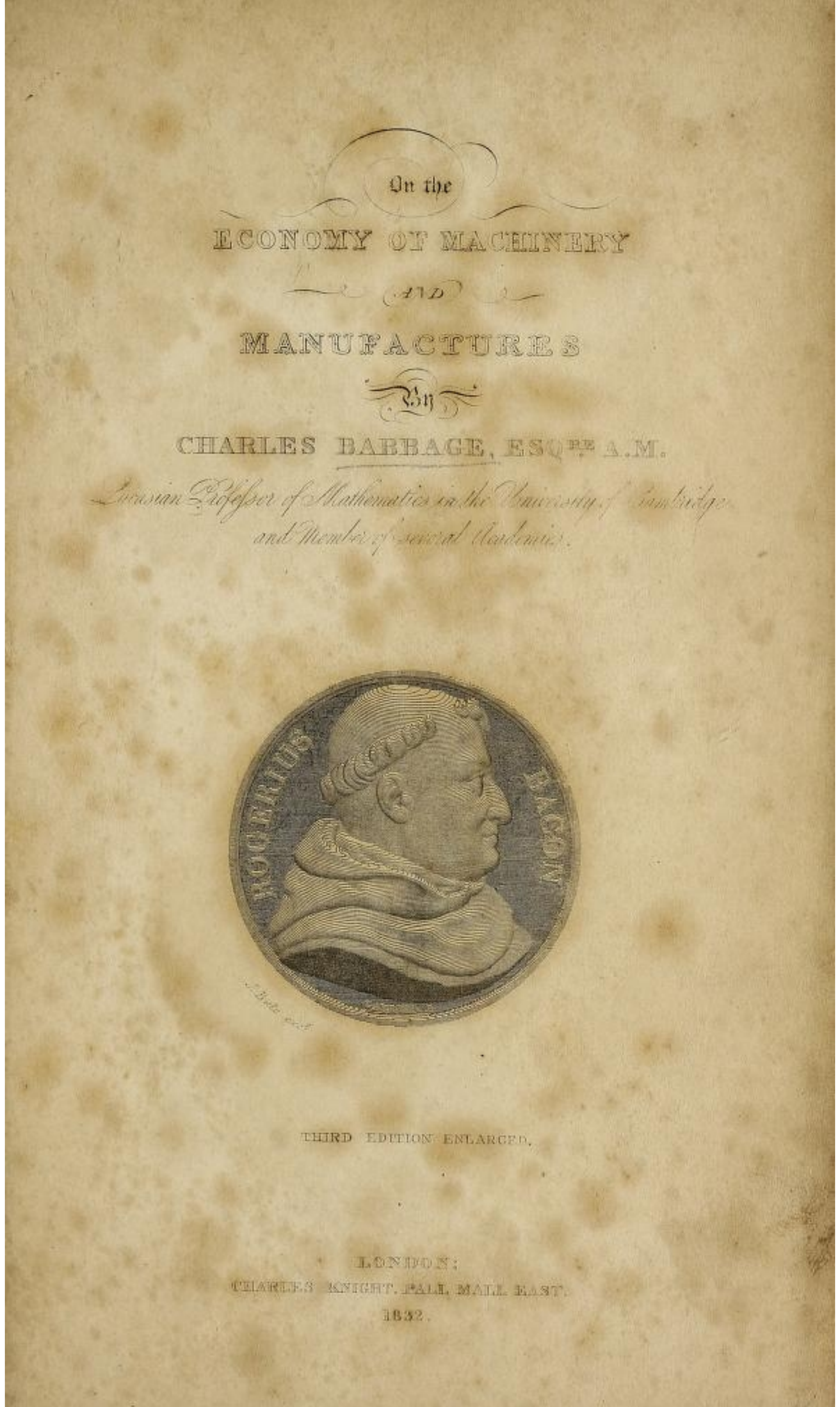
A Model of the Babbage Firm

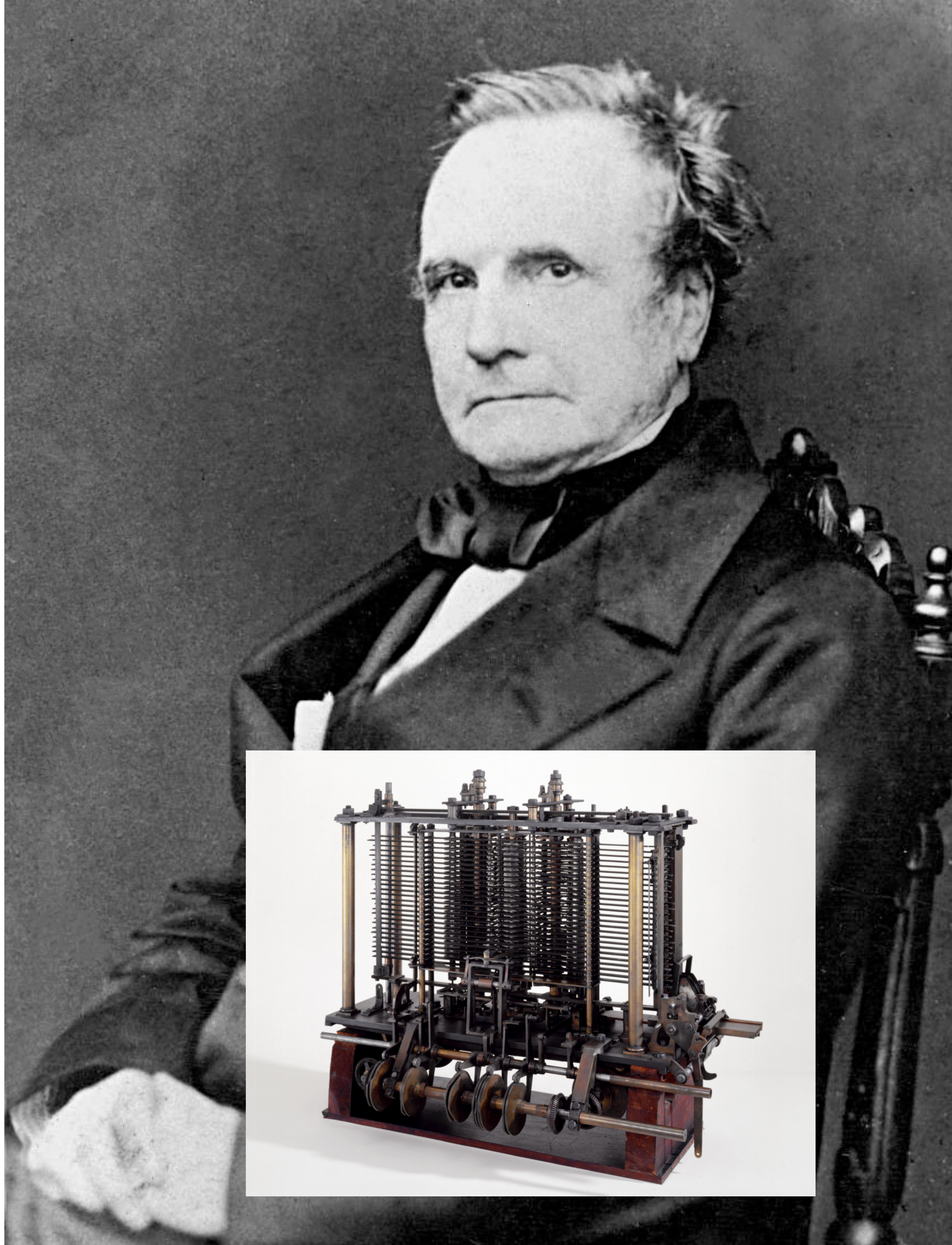
Joshua Gans

NBER Organizational Economics, 2025

Research Question: What determines the cognitive division of labour within organisations?


On the
ECONOMY OF MACHINERY
AND
MANUFACTURES
By
CHARLES BABBAGE, ESQ.RE A.M.
Lucasian Professor of Mathematics in the University of Cambridge
and Member of several Academies.

THIRD EDITION ENLARGED.
LONDON:
CHARLES KNIGHT, PALL MALL EAST.
1832.

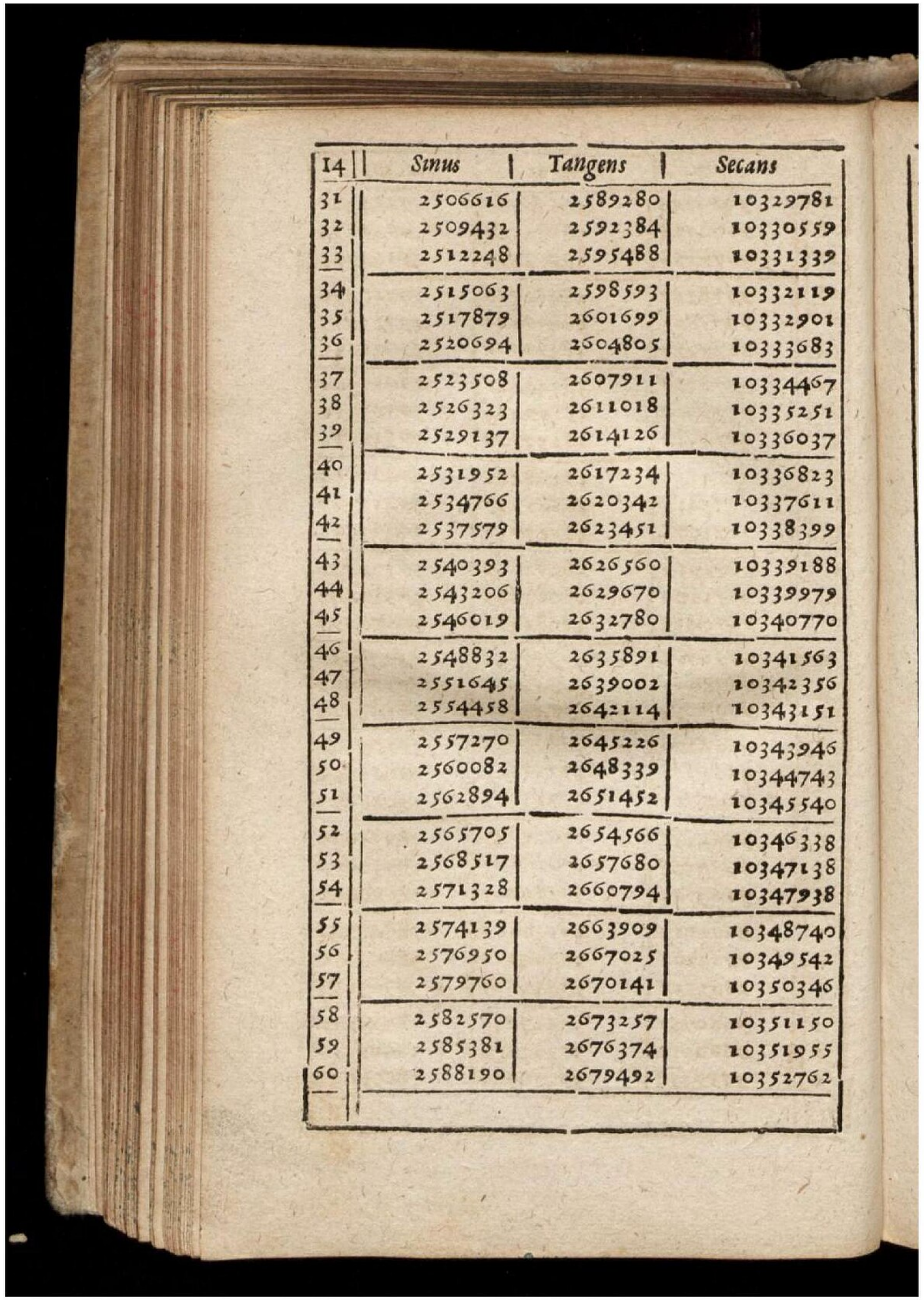


“At each increase of knowledge, as well as on the contrivance of every new tool, human labour becomes abridged.”

Charles Babbage

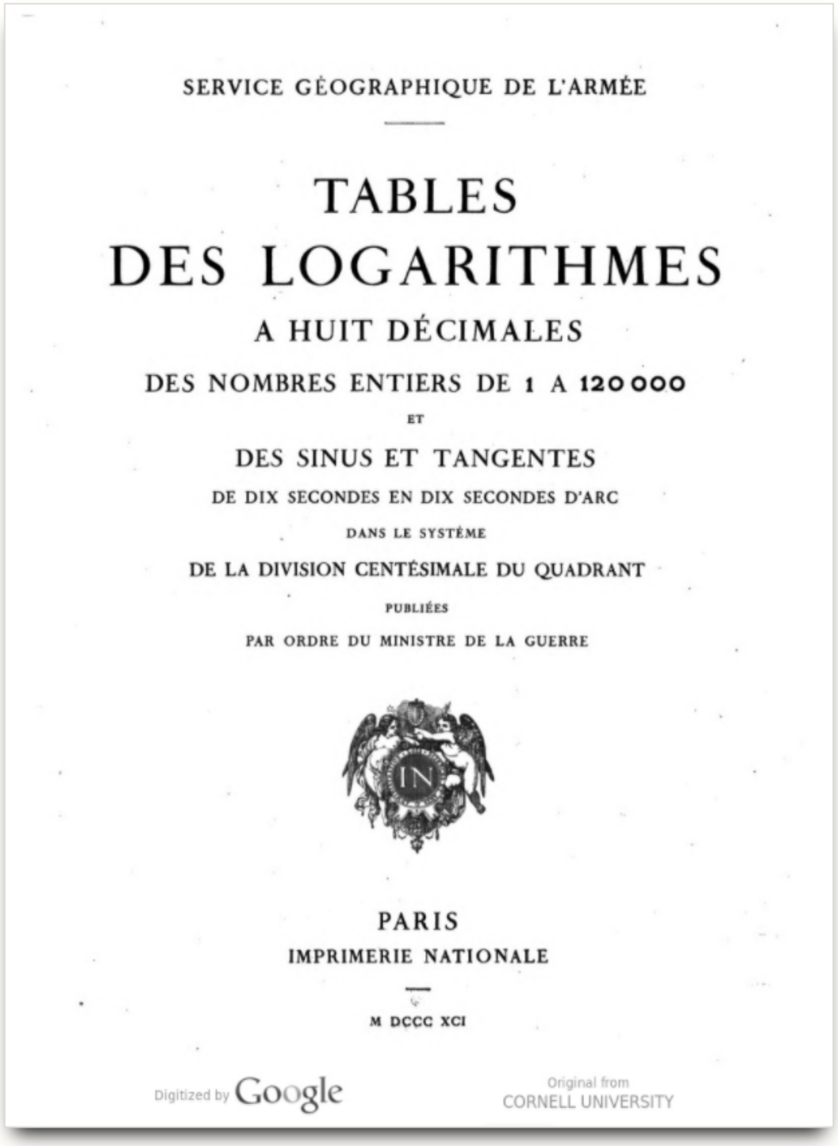
"[W]hat may, perhaps, appear paradoxical to some of our readers that the division of labour can be applied with equal success to mental as to mechanical operations, and that it ensures in both the same economy of time."

Charles Babbage



14	Sinus	Tangens	Secans
31	2506616	2589280	10329781
32	2509432	2592384	10330559
33	2512248	2595488	10331339
34	2515063	2598593	10332119
35	2517879	2601699	10332901
36	2520694	2604805	10333683
37	2523508	2607911	10334467
38	2526323	2611018	10335251
39	2529137	2614126	10336037
40	2531952	2617234	10336823
41	2534766	2620342	10337611
42	2537579	2623451	10338399
43	2540393	2626560	10339188
44	2543206	2629670	10339979
45	2546019	2632780	10340770
46	2548832	2635891	10341563
47	2551645	2639002	10342356
48	2554458	2642114	10343151
49	2557270	2645226	10343946
50	2560082	2648339	10344743
51	2562894	2651452	10345540
52	2565705	2654566	10346338
53	2568517	2657680	10347138
54	2571328	2660794	10347938
55	2574139	2663909	10348740
56	2576950	2667025	10349542
57	2579760	2670141	10350346
58	2582570	2673257	10351150
59	2585381	2676374	10351955
60	2588190	2679492	10352762

M.Prony



SERVICE GÉOGRAPHIQUE DE L'ARMÉE

TABLES
DES LOGARITHMES
A HUIT DÉCIMALES
DES NOMBRES ENTIERS DE 1 A 120000
ET
DES SINUS ET TANGENTES
DE DIX SECONDES EN DIX SECONDES D'ARC
DANS LE SYSTÈME
DE LA DIVISION CENTÉSIMALE DU QUADRANT
PUBLIÉES
PAR ORDRE DU MINISTRE DE LA GUERRE

PARIS
IMPRIMERIE NATIONALE

M DCCC XCI

Section 1: Routing

“The duty of this first section was to investigate, amongst the various analytical expressions which could be found for the same function, that which was most readily adapted to simple numerical calculation by many individuals employed at the same time. This section had **[illegible] to do with the actual numerical work**.”

No direct impact from automation

…but shift attention to innovation on routing methods

Section 2: Formulation

“This section consisted of seven or eight persons of considerable acquaintance with mathematics: and their duty was to convert into numbers the formulae put into their hands by the first section an operation of great labour. ”

Section 3: Computation

“The members of this section, whose number varied from sixty to eighty, received certain numbers from the second section, and, using nothing more than simple addition and subtraction, they returned to that section the tables in a finished state. It is remarkable that nine-tenths of this class had **no knowledge of arithmetic beyond the two first rules which they were thus called upon to exercise**, and that these persons were usually found more correct in their calculations, than those who possessed a more extensive knowledge of the subject.”

Wiped out by calculating machines

"[T]he effect of the division of labour, both in mechanical and in mental operations, is, that it enables us to purchase and apply to each process precisely that quantity of skill and knowledge which is required for it: ... we equally avoid the loss arising from the employment of an accomplished mathematician in performing the lowest processes of arithmetic."

Charles Babbage

Knowledge $\longrightarrow$ Decision-Making

Knowledge → How knowledge is acquired and processed → Decision-Making

Model of Knowledge for Decision-Making

Knowledge Division of Labour

- Garicano (2000): (knowledge) workers can devote time to learning solutions to problems.
  - “Knowing if someone knows the solution to a problem inevitably involves asking that person.”
  - Types of knowledge: routine (production floor) and exceptional (top of hierarchy)
  - Routing: work out who can solve that problem (refer it up the chain)
  - Trade-off: if people specialise in solving there are higher communication costs
  - A worker can answer questions in a domain by acquiring domain knowledge — the larger the domain, the higher the costs of acquiring knowledge.
  - If don’t know an answer, need to find someone who does but that means lost production time.
  - Division of Labour: some workers produce while others specialise in solving problems, and those specialise in knowledge.
- This paper: division of labour based on cognitive processes rather than knowledge-production-communication trade-offs.

Econometrica, Vol. 93, No. 2 (March, 2025), 623–659

A QUEST FOR KNOWLEDGE

CHRISTOPH CARNEHL
Dept. of Economics & IGIER, Bocconi University

JOHANNES SCHNEIDER
Department of Economics, Universidad Carlos III de Madrid

Is more novel research always desirable? We develop a model in which knowledge shapes society's policies and guides the search for discoveries. Researchers select a question and how intensely to study it. The novelty of a question determines both the value and difficulty of discovering its answer. We show that the benefits of discoveries are nonmonotone in novelty. Knowledge expands endogenously step-by-step over time. Through a dynamic externality, moonshots—research on questions more novel than what is myopically optimal—can improve the evolution of knowledge. Moonshots induce research cycles in which subsequent researchers connect the moonshot to previous knowledge.

KEYWORDS: Science of Science, Search on Brownian Path, Knowledge Creation.

1. INTRODUCTION

IN A LETTER TO FRANKLIN D. ROOSEVELT, Vannevar Bush (1945) pleads with the president to preserve freedom of inquiry by federally funding basic research—the "pacemaker of technological progress." That letter paved the way for the creation of the National Science Foundation (NSF) in 1950. The NSF today, like the vast majority of governments and scientific institutions, cherishes scientific freedom and allows academic researchers to select research projects independently.

With scientific freedom comes the responsibility to select the right research questions. However, what are the right questions? Biologist and Nobel laureate Peter Medawar (1967) famously notes that "research is surely the art of the soluble. . . . Good scientists study the most important problems they think they can solve." Finding the most important yet soluble question is nontrivial. One reason is that both importance and solubility depend on the current state of knowledge.

In this paper, we develop a microfounded model of knowledge creation through research. We conceptualize research as a costly search process that may fail. The researcher picks a question and an intensity with which to search for its answer. The cost of search depends on how well existing knowledge guides the researcher's efforts. If the researcher

Christoph Carnehl: christoph.carnehl@unibocconi.it
Johannes Schneider: jschneid@eco.uc3m.es
We thank Arjada Bardhi, Antonio Cabrales, Steven Callander, Marco Celentani, Rahul Deb, Philipp Denter, Florian Ederer, Chiara Franzoni, Alex Frug, William Fuchs, Ben Golub, Ryan Hill, Toomas Hinosaar, Nenad Kos, Jorge Lemus, Igor Letina, David Lindequist, Gerard Llobet, Ignacio Ortuño, Marco Ottaviani, Nicola Pavoni, Harry Pei, Konrad Stahl, Carolyn Stein, Armin Schmutzler, Carlo Schwarz, Adrien Vigier, Ludo Visschers, Ralph Winkler, and various audiences for comments. Johannes Schneider gratefully acknowledges financial support from Horizon Europe MSCA Project 101061192; the Agencia Estatal de Investigación (MICIU/AEI /10.13039/501100011033) through grants PID2019-111095RB-I00, PID2020-118022GB-I00, and CEX2021-001181-M; the German Research Foundation (DFG) through CRC TR 224 (Project B03); and Comunidad de Madrid, grant MAD-ECON-POL-CM H2019/HUM-5891. Funding for APC: Universidad Carlos III de Madrid (Agreement CRUE-Madroño 2024). Part of the research was carried out while Johannes Schneider was visiting the University of Bern. He thanks them for their hospitality and inspiration.

A Simplified Quest for Knowledge

Joshua S. Gans*

May 16, 2025

Abstract

This paper develops a transparent, simplified version of Carnehl and Schneider (2025)'s model of knowledge creation. Our tractable framework, which yields closed-form solutions for key welfare trade-offs, preserves the essential economic mechanisms while eliminating mathematical complexity. We derive four main insights. First, contrary to the original model's emphasis on "moonshots," our analysis demonstrates that expanding knowledge and then deepening it (the moonshot approach) is never socially optimal under direct welfare comparisons. The original model's case for moonshots relies on second-best arguments involving research costs and dynamic externalities, not on direct welfare considerations. Second, we identify a novel misalignment between private and social incentives in multidisciplinary research contexts. Even without research costs — where the original model predicts perfect alignment — researchers bridging large knowledge gaps between disciplines choose locations that create suboptimal knowledge structures. Third, we analyse how citation-based incentive systems affect knowledge creation trajectories. We show that systems privileging unique contributions over shared ones align private behaviour with social welfare objectives, while those rewarding shared contributions lead to excessive knowledge deepening. Fourth, our analysis provides precise characterisations of optimal knowledge creation paths under various initial conditions and offers clear guidance for science policy. By clarifying when interventions can address misalignments between researchers' incentives and social welfare, our simplified model offers practical insights for the design of research funding mechanisms. *Journal of Economic Literature* Classification Numbers: O31, D83, H41.

Keywords. knowledge creation, research policy, moonshot, multidisciplinary research, scientific exploration, citation incentives

*Rotman School of Management, University of Toronto and NBER. Thanks to Christoph Carnehl and Johannes Schneider for helpful discussions and Claude Sonnet 3.7 for valuable research assistance. Responsibility for all errors remains my own.

Universe of **questions**: $x \in \mathbb{R}$

Each has a unique **answer**: $y(x)$, realisations of the Brownian motion of random variable, $Y(x)$, or **truth**

Existing **knowledge**: $\mathscr{F}_t = \{(x_i, y(x_i))\}_{i=1}^{k_t}$ with $x_1 < x_2 < \cdots < x_k$

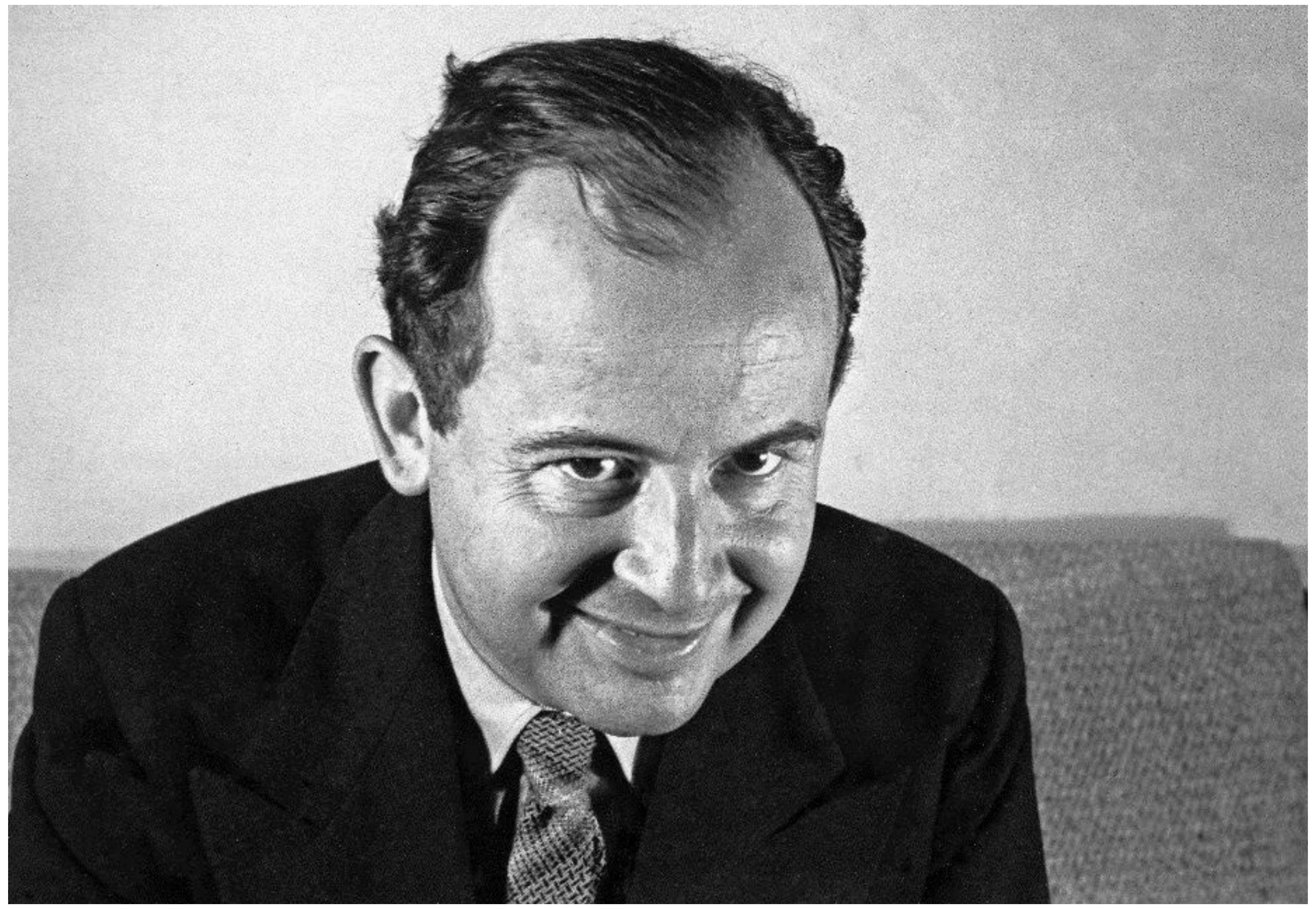
Decision-Maker (DM)

- For each question, x, DM decides whether to take an action, $a(x)$, or abstain, $\emptyset$ with payoff:

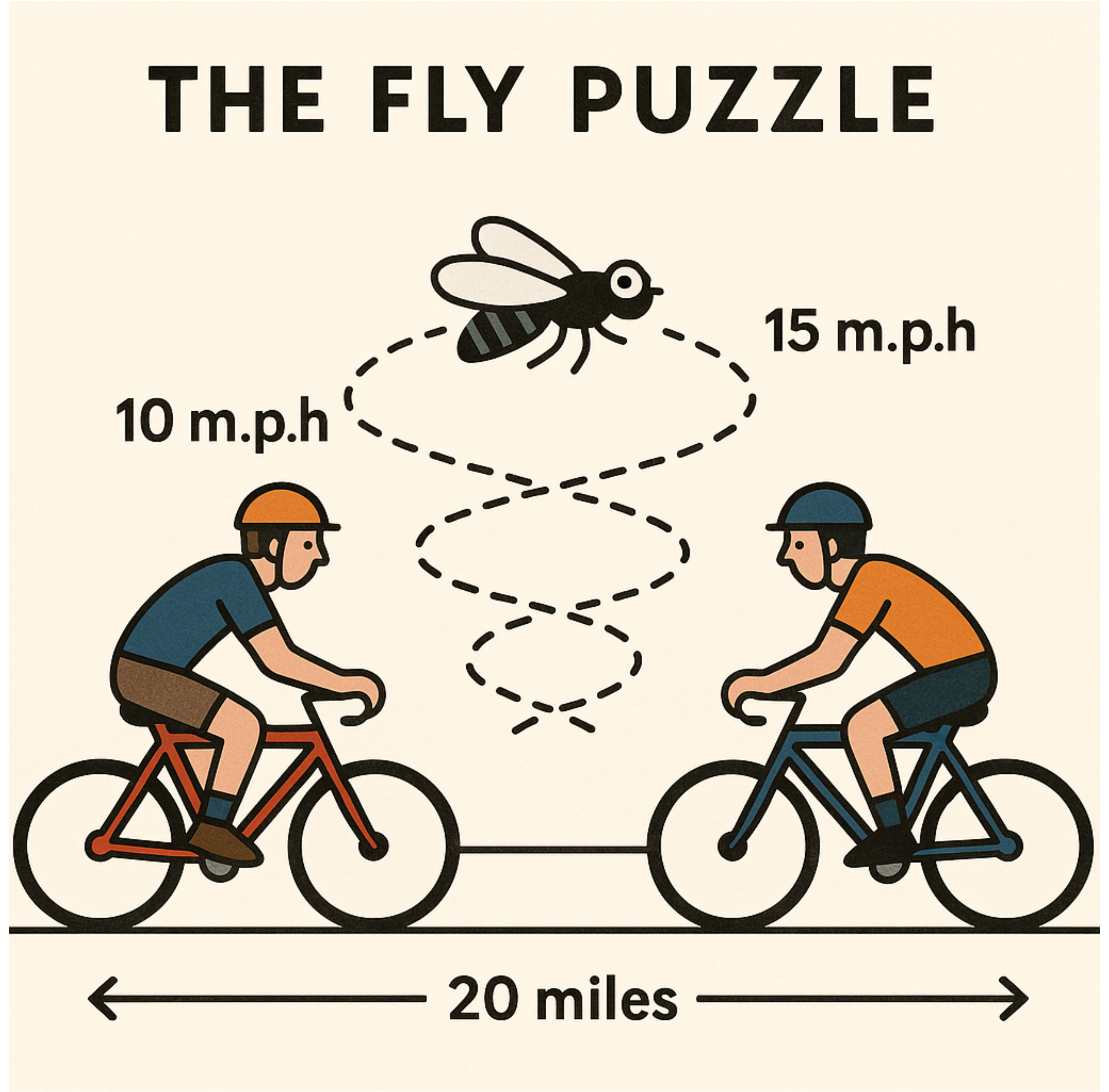
$$u(a(x), x) = \begin{cases} 0 & \text{if } a(x) = \emptyset, \\ 1 - \dfrac{(a(x) - y(x))^2}{q} & \text{if } a(x) \in \mathbb{R}\,. \end{cases}$$

- Optimum: $a^*(x) = y(x)$
- How does existing knowledge, $\mathcal{F}_t$, assist in making decisions?

Expert: someone how knows specific things and can answer questions based on related things from first principles. (Perhaps the person who created the knowledge in the first place).

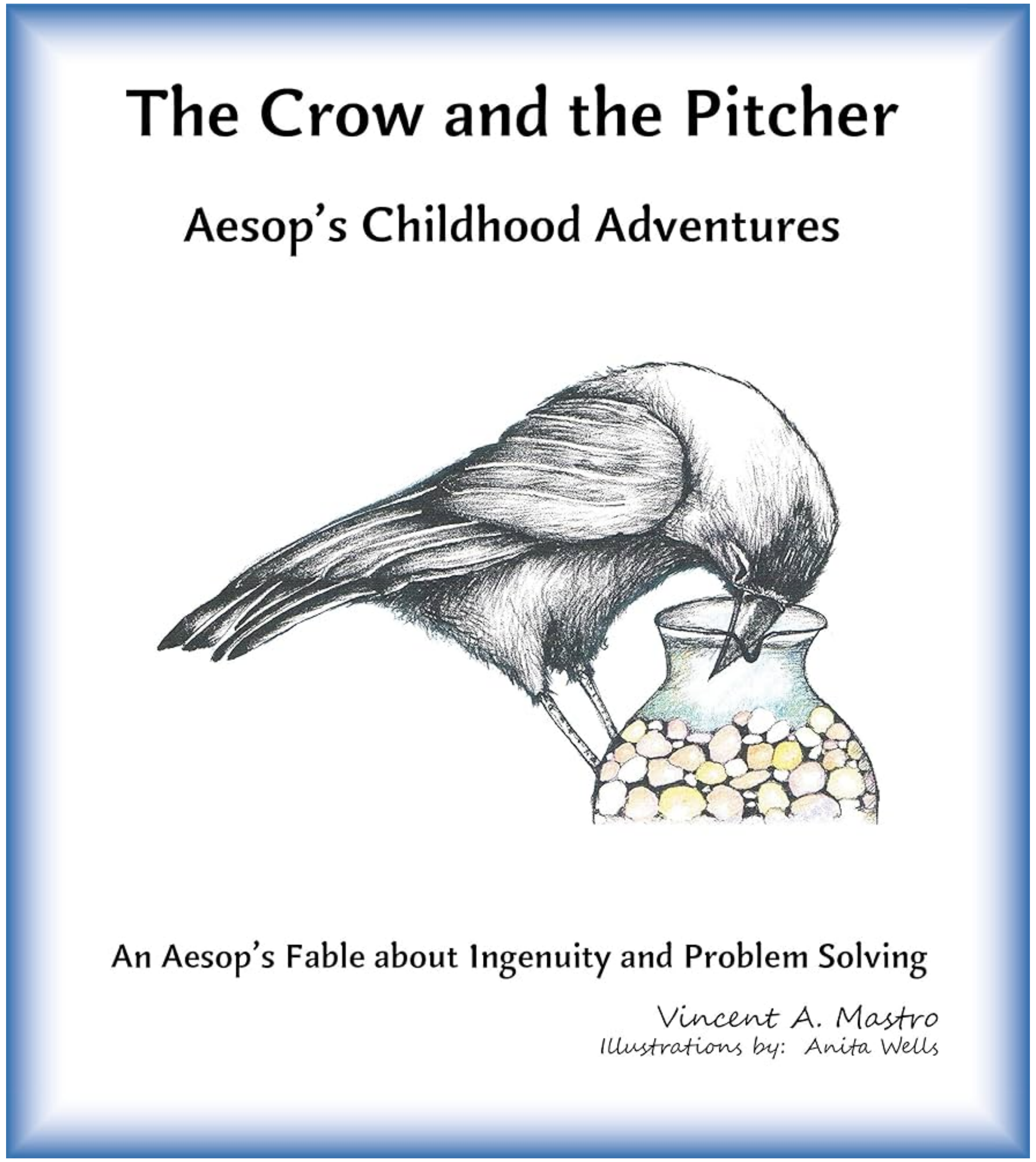


Then there is the famous fly puzzle. Two bicyclists start twenty miles apart and head toward each other, each going at a steady rate of 10 m.p.h. At the same time a fly that travels at a steady 15 m.p.h. starts from the front wheel of the southbound bicycle and flies to the front wheel of the northbound one, then turns around and flies to the front wheel of the southbound one again, and continues in this manner till he is crushed between the two front wheels. Question: what total distance did the fly cover? The slow way to find the answer is to calculate what distance the fly covers on the first, northbound, leg of the trip, then on the second, southbound, leg, then on the third, etc., etc., and, finally, to sum the infinite series so obtained. The quick way is to observe that the bicycles meet exactly one hour after their start, so that the fly had just an hour for his travels; the answer must therefore be 15 miles. When the question was put to von Neumann, he solved it in an instant, and thereby disappointed the questioner: "Oh, you must have heard the trick before!" "What trick?" asked von Neumann; "all I did was sum the infinite series." Halmos (1973), 386-287.



Routine: Take existing knowledge and make inferences to guess at answers to questions

The story concerns a thirsty crow that comes upon a pitcher with water at the bottom, beyond the reach of its beak. After failing to push it over, the bird drops in pebbles one by one until the water rises to the top of the pitcher, allowing it to drink.

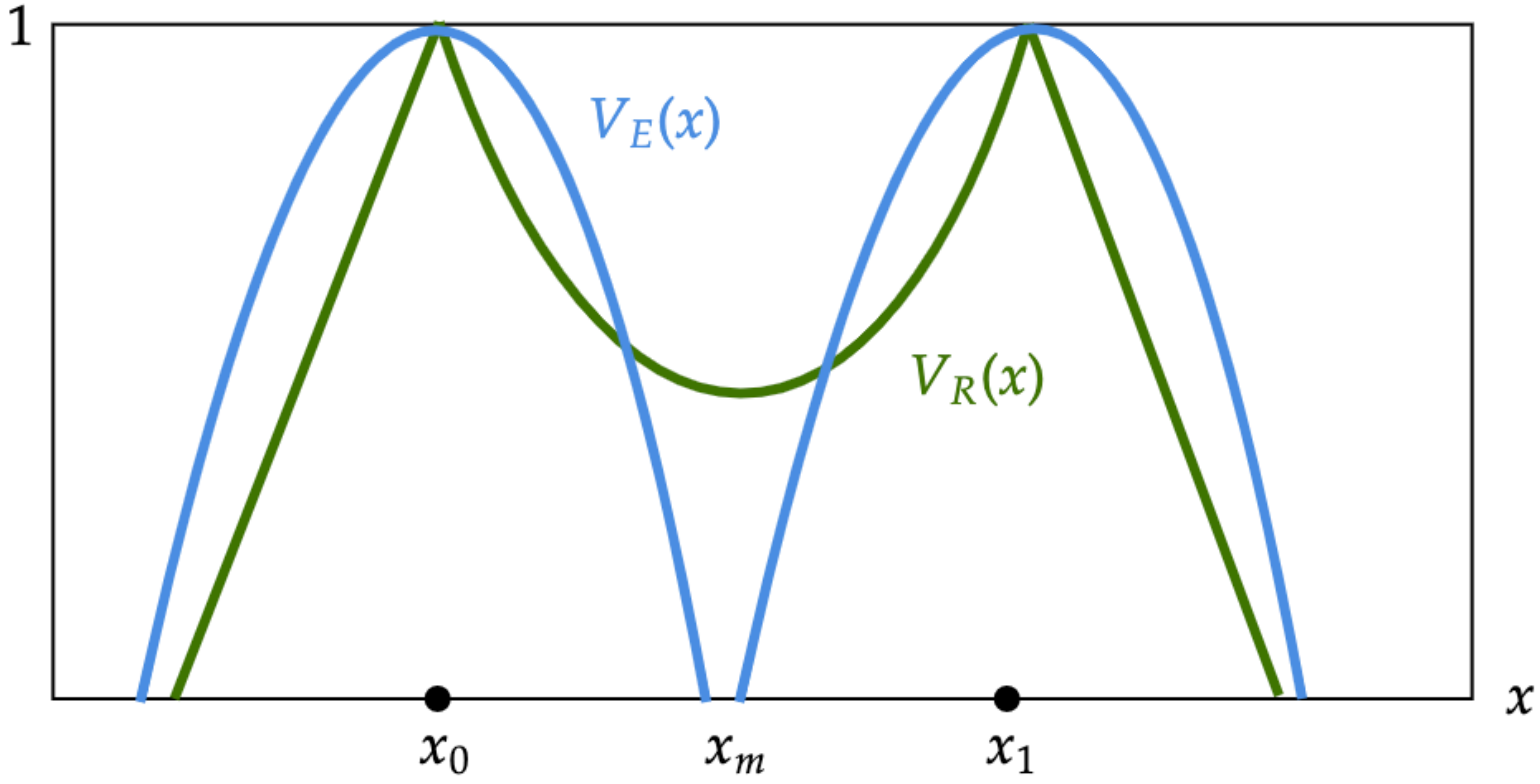


Using existing knowledge, **experts** (who were responsible for the knowledge in the first place) can create (tacit) solutions to problems within a domain

Using existing knowledge, **routine workers** make conjectures about answers

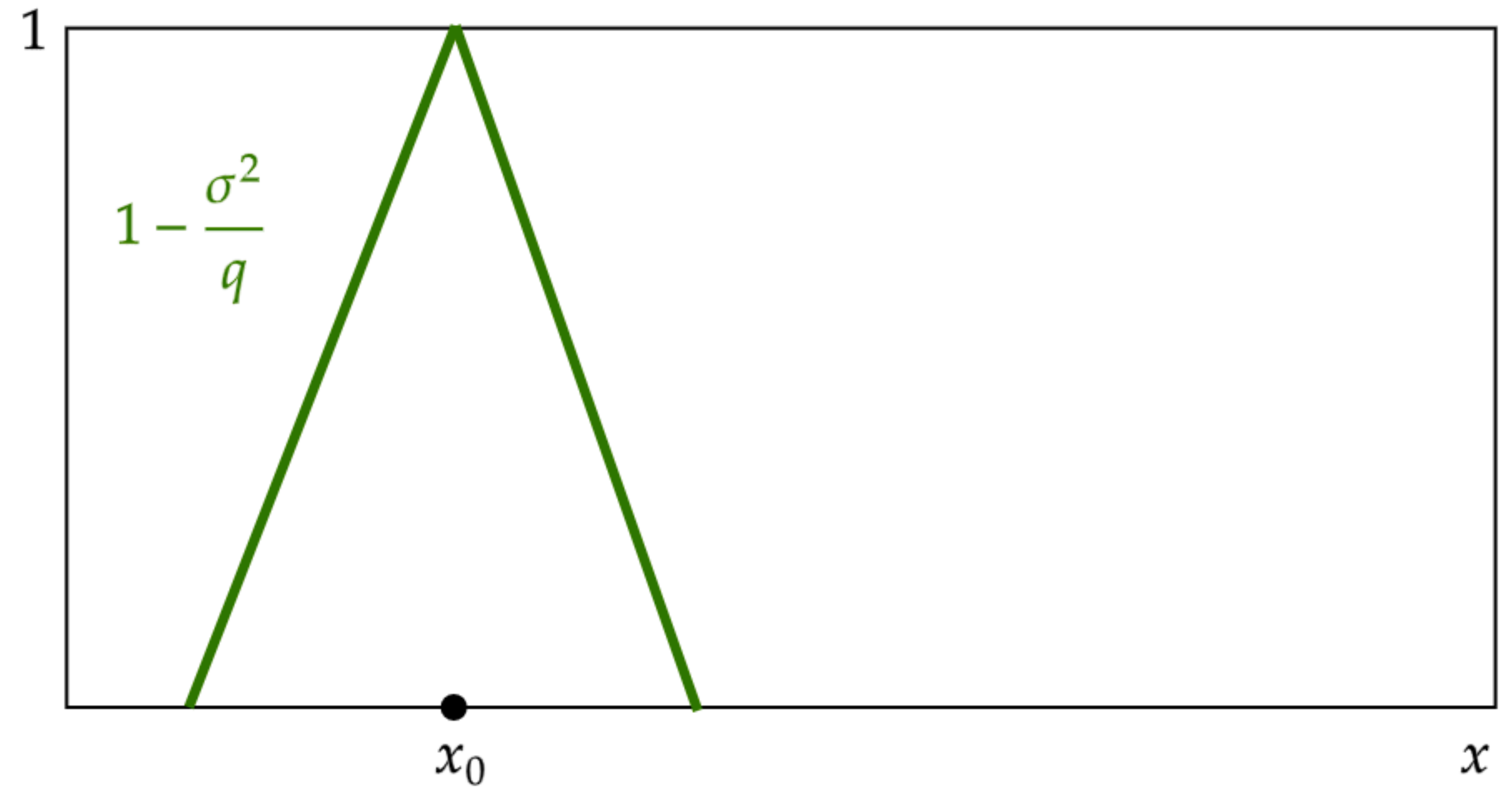
Two Ways of Processing Knowledge for Decisions

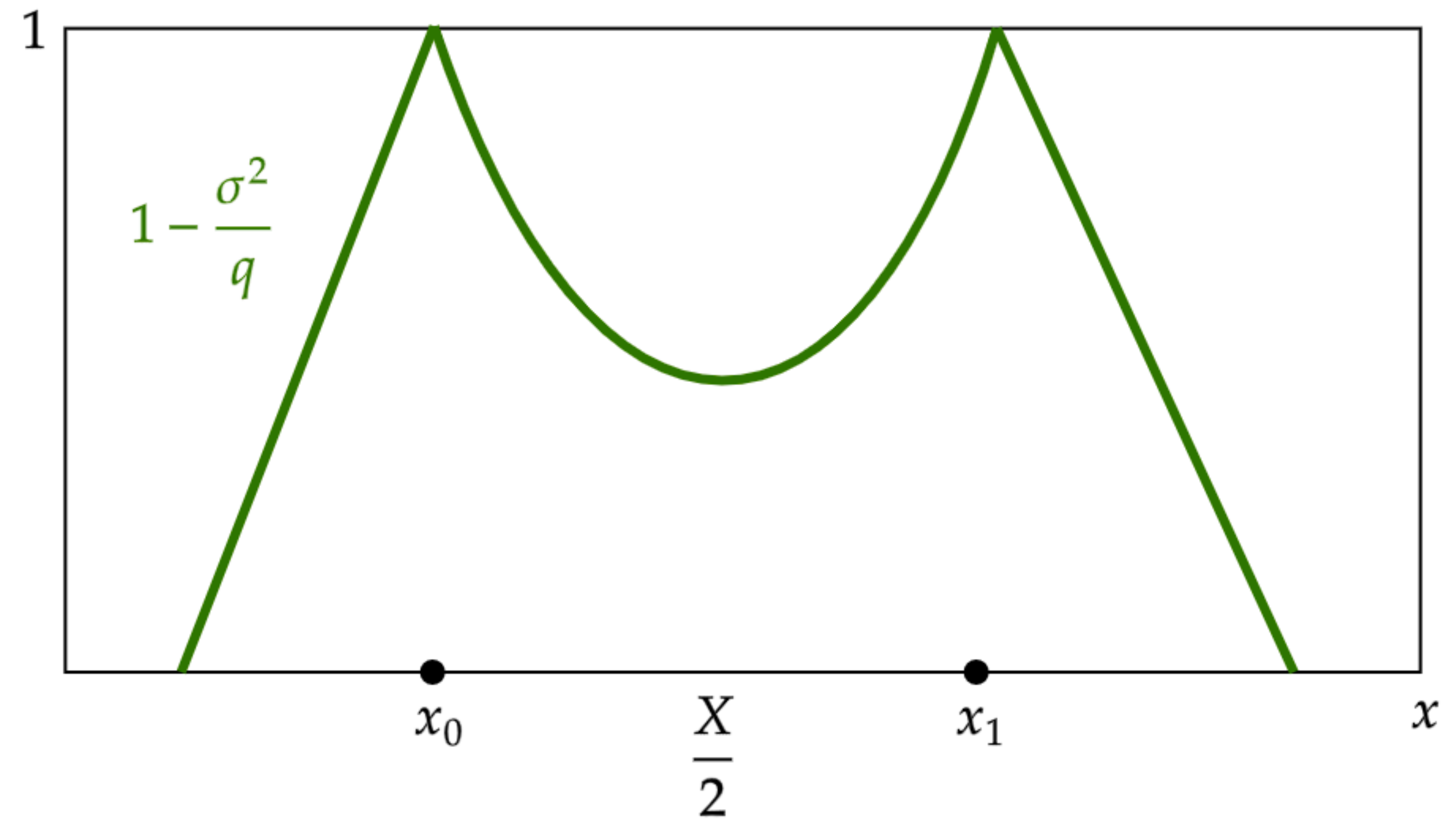
- New question at $x = x_0 + s$
- **Experts**: can use an existing knowledge point, work from first-principles, and find a solution at cost, $V_E(s) = \max\{1 - \eta s^2, 0\}$
- **Routines**: can interpolate a conjecture from existing knowledge with expected payoff, $V_R(s) = \max\{1 - \frac{\sigma^2(s)}{q}, 0\}$ with $\sigma^2(s) = s$ or $\frac{s(X-s)}{X}$ (where X is the distance between two nearest points).



Assume that $\eta \in [\frac{1}{4q^2}, \frac{1}{q^2}]$ so that the expert is **dominates** the routine worker for a single knowledge point, but their expertise is **localised**.

Payoff from Routine DM





$(X < 4q)$

Expert's Second Role: Create Knowledge (Exploration)

- Instead of answering a question in a period, the expert can engage in knowledge creation — that is, research towards a new $(x_{\text{new}}, y(x_{\text{new}}))$
- The manager must consider (a) whether to allocate the expert to immediate decisions and, if not, (b) where the expert should create new knowledge?
- Consider the expert role allocation. Let

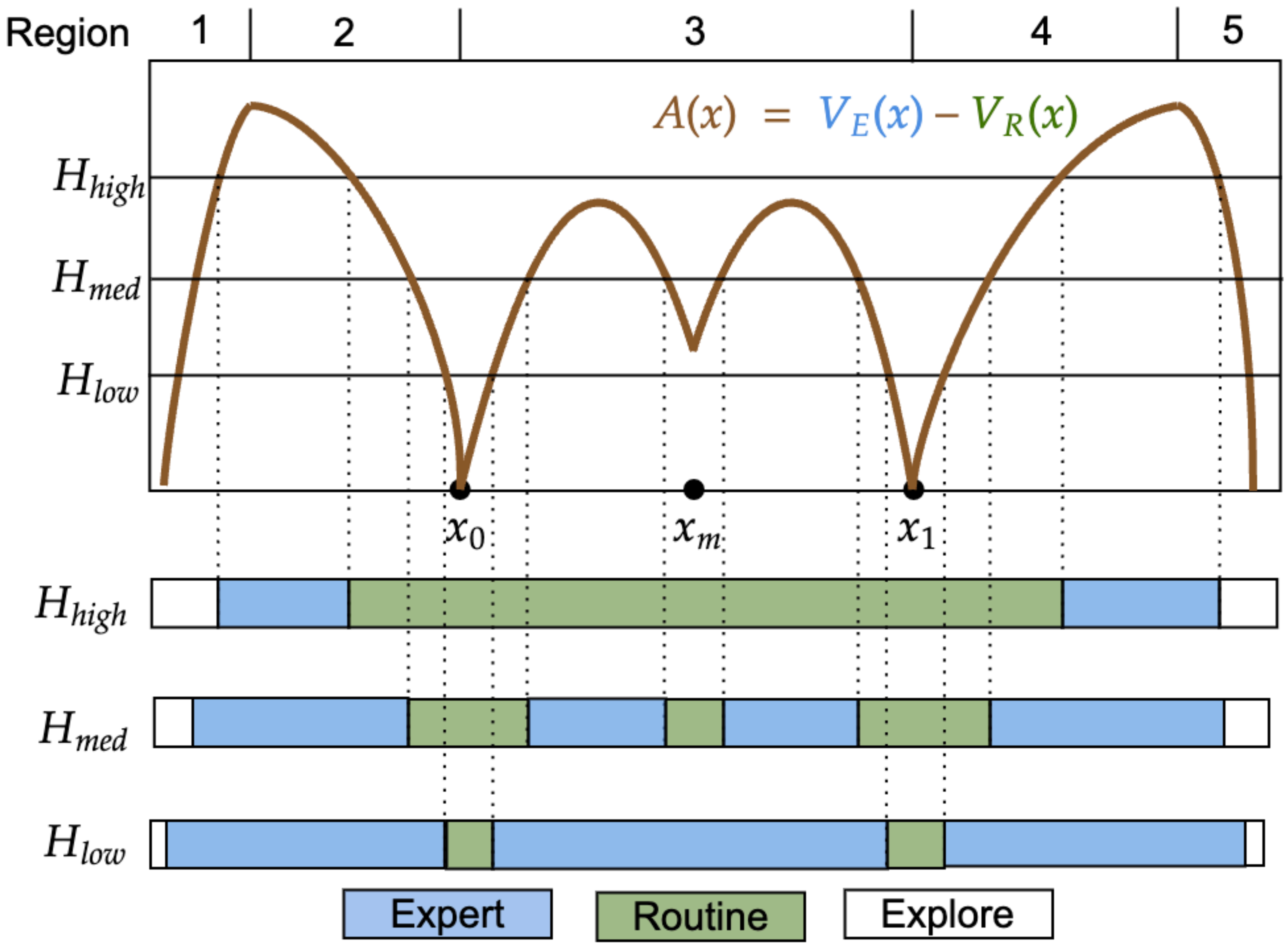
$$H^*(\mathcal{F}_t) = \max_{x_{\text{new}}} \left\{ \delta \left[W(\mathcal{F}_t \cup \{(x_{\text{new}}, y(x_{\text{new}}))\}) - W(\mathcal{F}_t) \right] \right\}$$

be the **hurdle level** (that is, the optimal NPV of the payoff from new knowledge)

Then, upon observing the current question, $x(t)$, allocate the expert to research if and only if:

$$V_E(x(t)) - V_R(x(t)) \leq H^*(\mathcal{F}_t)$$

Cognitive Division of Labour



Well, this is kinda hard …

- In order to calculate the cognitive division of labour at any t, we need to know $H^*(\mathscr{F}_t)$.
- That requires solving for the optimal knowledge point that might be added for each $\mathscr{F}_t$.
- Which requires knowing the allocation of future labour for any new knowledge point.
- Taking into account the random arrival of questions that determines whether experts exploit or explore in any period, the probability (ρ) of which is endogenous.
- And because knowledge accumulates, the dynamic programming problem is non-stationary.
- Nonetheless, as new knowledge points are added $H^*(\mathscr{F}_t)$ will not increase with time.
- I'm either gonna need an expert or …

… make a simplifying approximation …

$$H_t^{SVA} = \frac{\delta}{1-\delta}\frac{\Delta^*(t)}{L}$$

where $\Delta^*(t)$ is the expected improved performance from **one** optimally placed new knowledge point

- **Greedy**: only static improvement
- **Separation**: no need to consider future prob of exploration, $\rho(t)$
- **Bias**: $H_t^{SVA} < H_t^*$ early (under-states exploration) and $H_t^{SVA} > H_t^*$ (over-states exploration) later

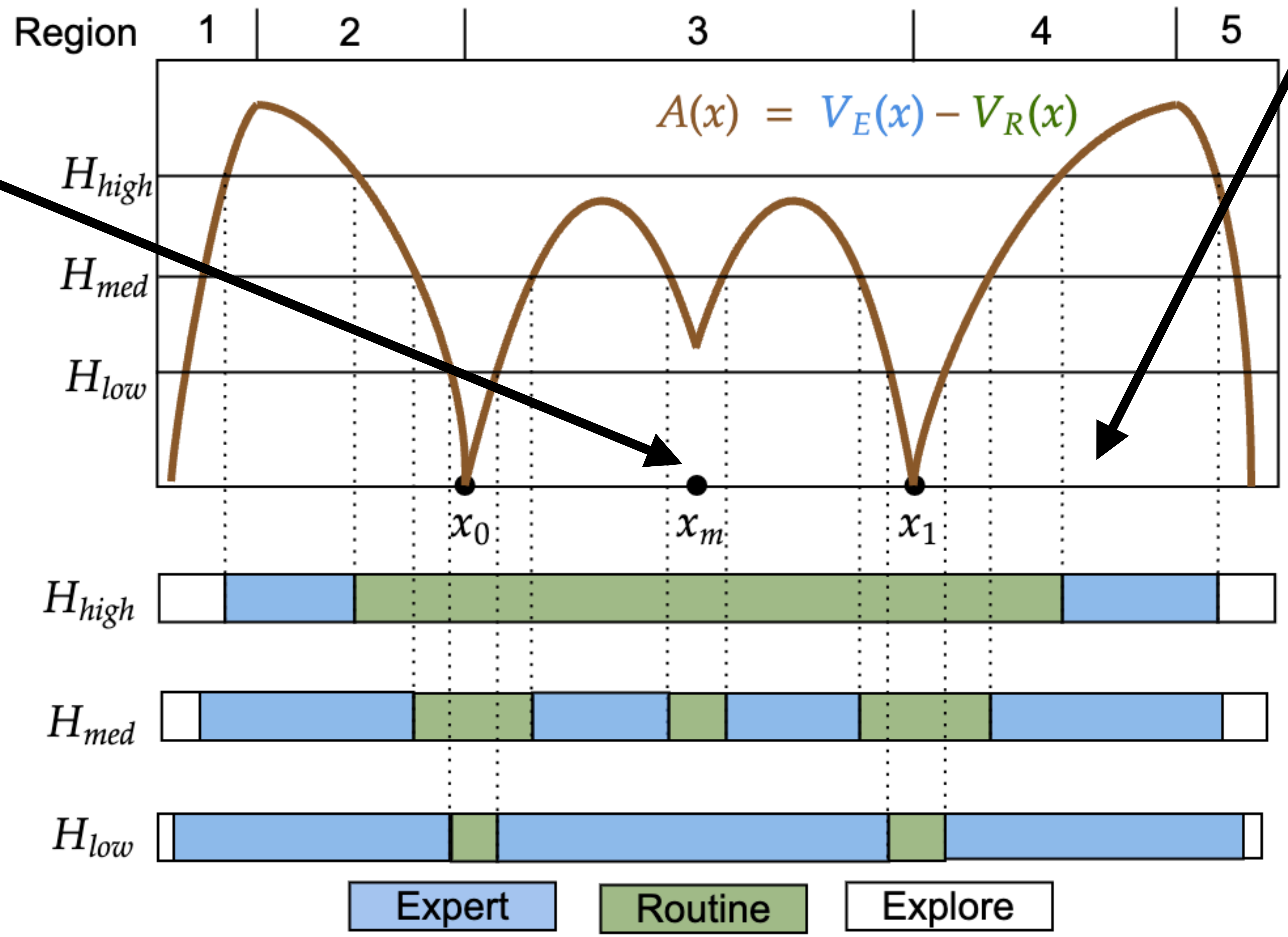
Tails versus Gaps

New Gap Point

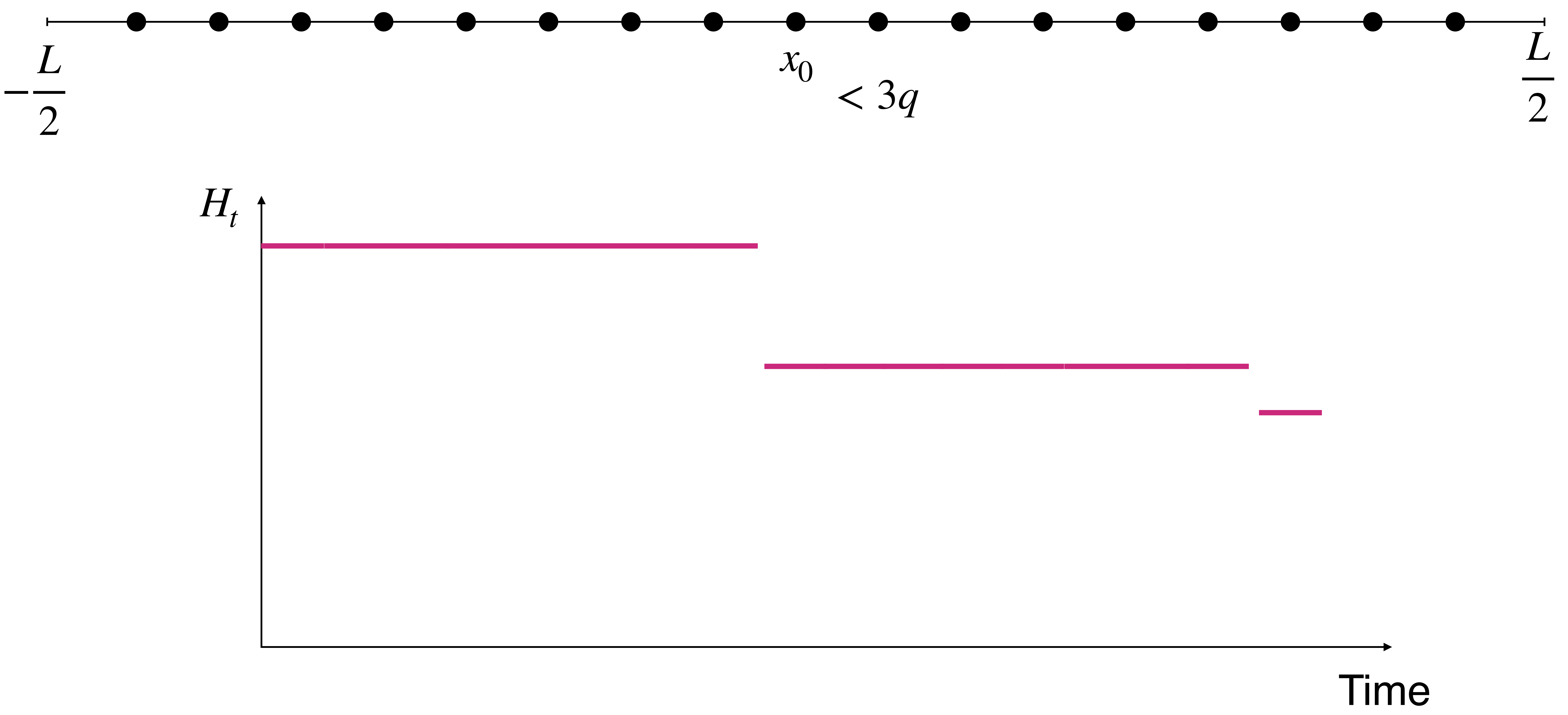
Smaller bridge
New anchor

New Tail Point

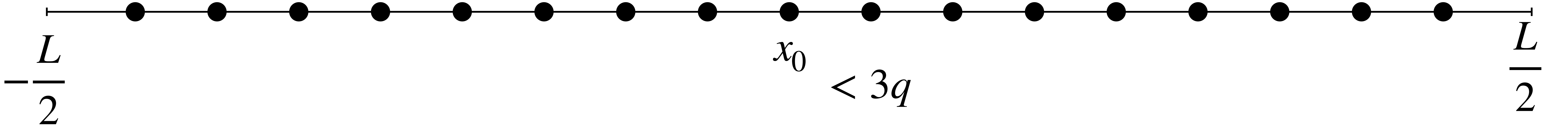
New bridge
New anchor

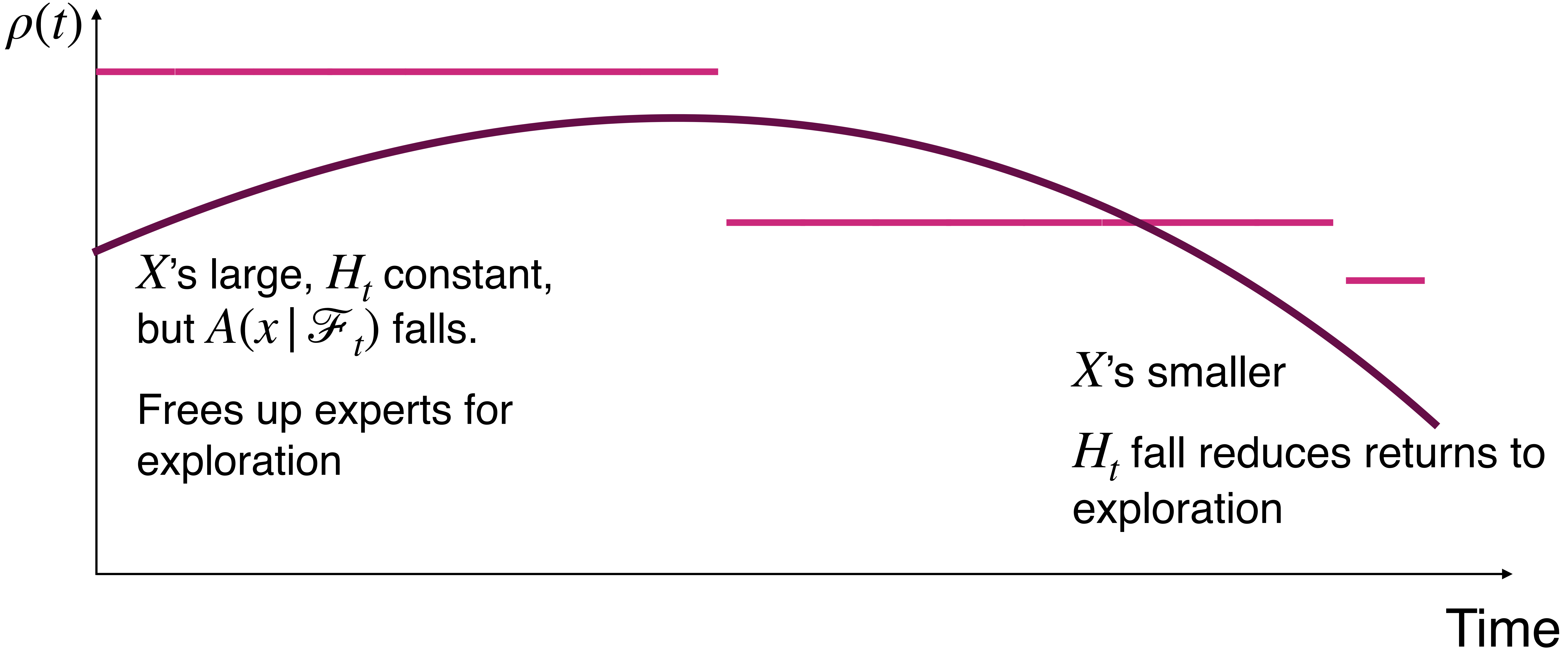


Evolution of H_t



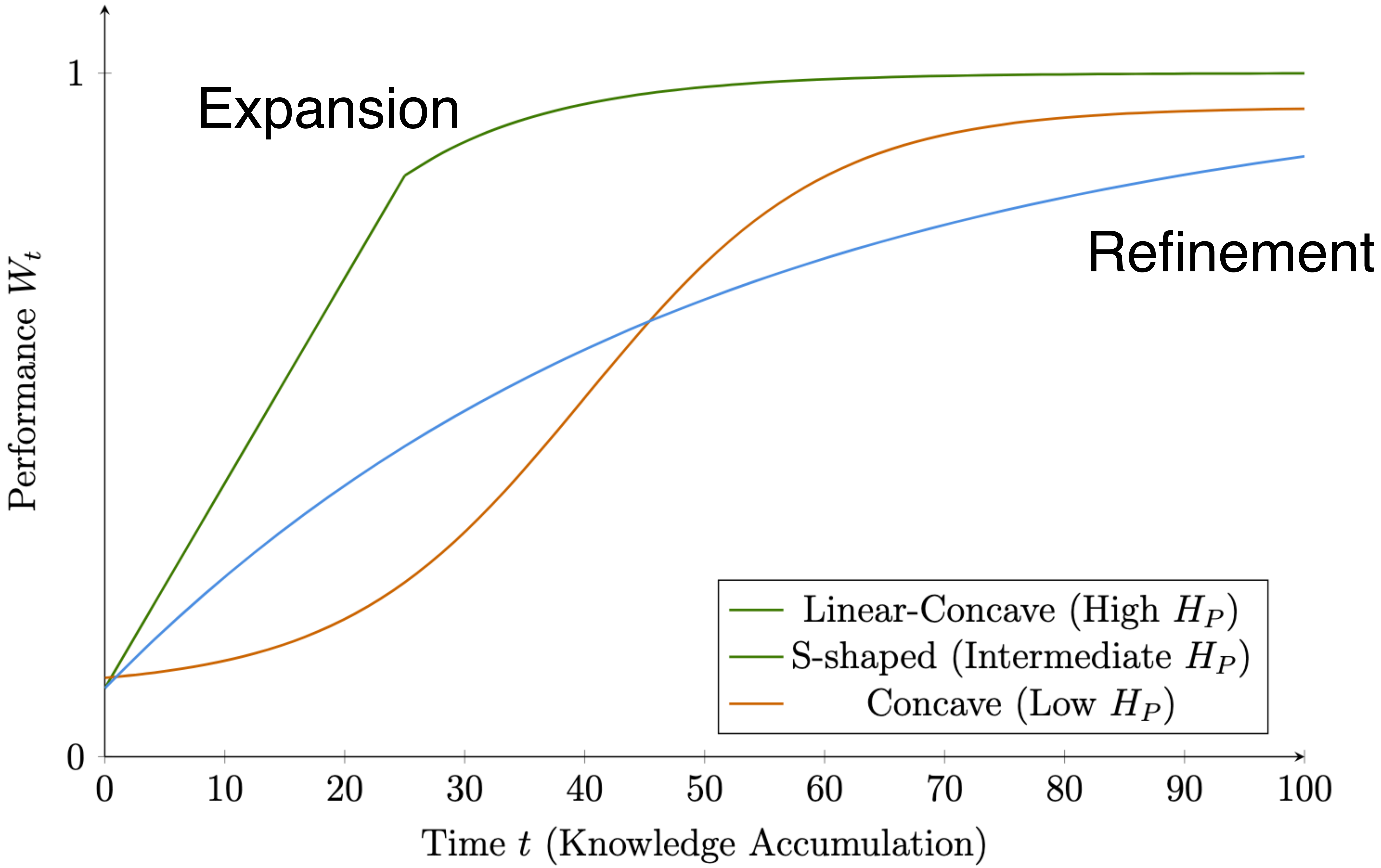
Evolution of $\rho(t)$





Performance Over Time

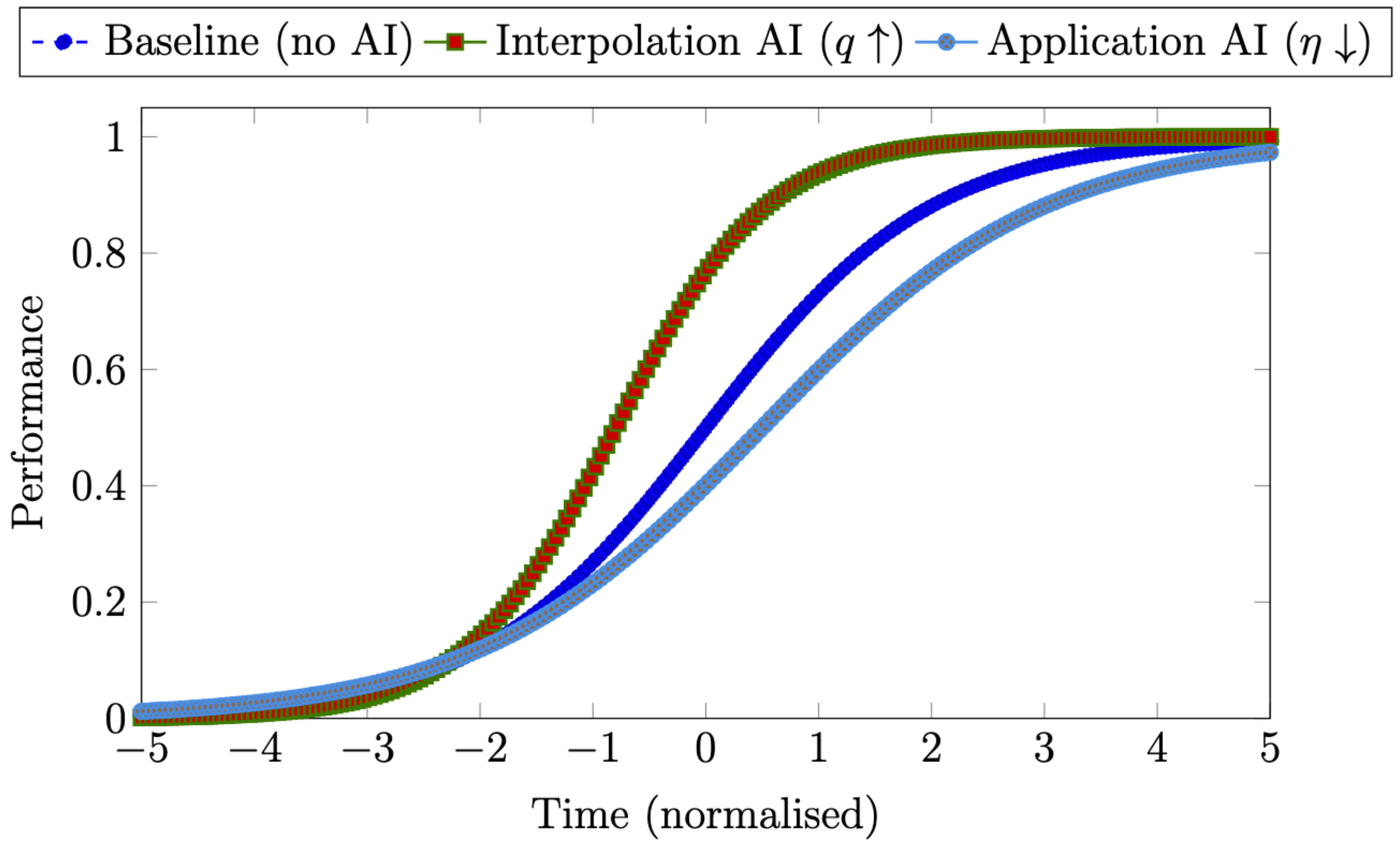
$$G(t) = \mathbb{E}[W_{t+1} - W_t] = \rho(t) \cdot H_t$$



Impact of AI

AI can impact on operational productivity of experts ($\downarrow \eta$) or routines ($\uparrow q$)

- Interpolation AI ($\uparrow q$), reduces $A(x)$ and increases H_t and $\rho(t)$, accelerating growth
- Application AI ($\downarrow \eta$), increases $A(x)$ and increases H_t but decreases $\rho(t)$, which may reduce growth
- Interpolation AI more important for start-ups while Application AI is more important for mature firms.



Shifts in Question Domain

Suppose that the question domain shifts by B to $[-\frac{L}{2}+B, \frac{L}{2}+B]$, what is the impact on nascent versus mature firms?

- Nascent firms continue to invest in knowledge points in new area. (Normal trajectory)
- Mature firms lose the value of some knowledge (proportional to B) but face renewed incentives for exploration in the 'open domain.' But expands from $\frac{L}{2}$ in discrete increments (if B large). Substantial move from exploitation to exploration with ρ close to 1.

Thus, the mature firm grows faster than the nascent firm, contrary to many *inertia* theories. In this model, knowledge is valuable if it is relevant, as it weakens the trade-off between exploitation and exploration in response to a shock.

- Very reliant on uniform distribution logic. If non-uniform, other things may occur.

Conclusions

- Division of labour:
  - Role assignment in each period between knowledge workers who use existing knowledge to make predictions for decision-making (routines) versus those who work at solving problems close to existing knowledge points (experts) as they are also the people who created the knowledge point initially.
  - Chief tension: using experts to “put out fires” (exploitation) versus invest in new knowledge (exploration). Depends on current question location (operational expert advantage) and current knowledge set (prospective returns to new knowledge).
  - Conclusion: division of labour evolves over time with exploration favoured initially towards evolving to a situation where experts exploit at little cost (S-curve)
- Applications:
  - AI: what impact does it have on decision-making roles (depends on type of AI) and how does this impact on organisational dynamics (startups versus mature firms)
  - Adaptability: what happens when the set of questions shifts? Current model shows that knowledge is valuable and allows faster exploration of new, open areas.