Learning Across Alternatives:

Strategic Behavior in a Complicated World

Org Econ Workshop
October 2020

Steven Callander

Stanford University & UNSW

Motivating Example

- Learning from job experience.

- Consider a fresh university graduate who takes a first job at Goldman Sachs ... and hates it.

- What should she do next?
  - Take a job at Morgan Stanley? At Macquarie Bank?
  - A hedge fund? Private equity?
  - A school teacher? A farmer? ... an academic?

- From her bad experience, she will learn about other jobs.

- We learn across alternatives in many (all?) settings.

- Standard models don't capture this type of learning.
  - Multi-armed bandit model → throw out failures and begin fresh.

Plan for Lectures

- Learning across alternatives is endemic to problems studied by economic theory (or should be studied by).
  - Including problems relevant to organizations.

- I will introduce an approach to capture "learning across alternatives."
- I then apply this tool to four types of strategic/decision problems:

  1. Search & Experimentation
  2. Communication
  3. Judicial Decision Making
  4. Market Competition

- Includes many and varied applications (orgs, politics, markets, etc.).
  - These are four types of problem I have worked on.

- I will try to point to new applications and open questions.

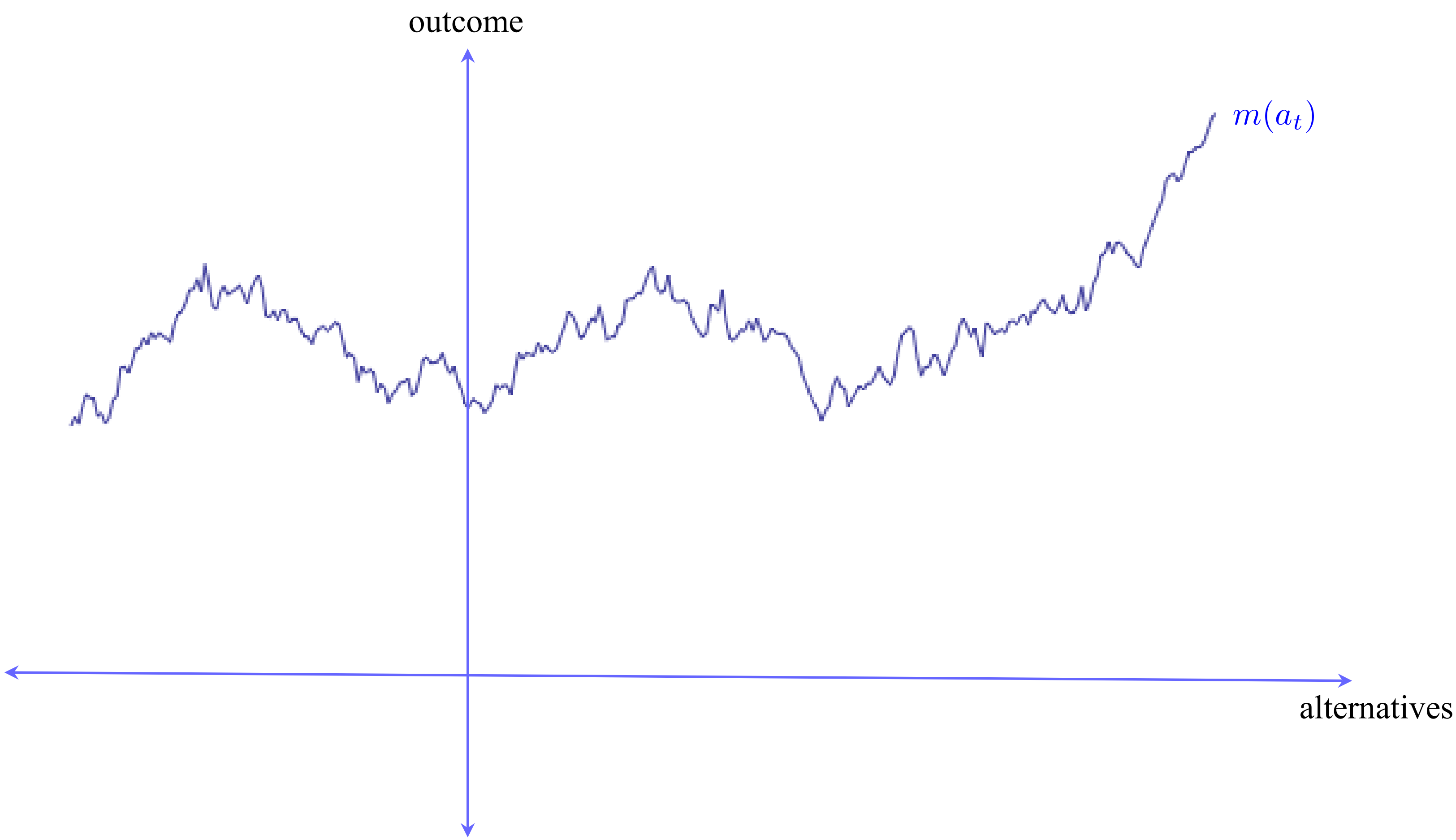
Why a Gap in Economic Theory?

- Economic theory is hard!
- Theory only captures a fraction of what is important in the real world.
  - This is easily (and often) forgotten.
- The technique of theory: Simplify, simplify, simplify.
- The simplest way to capture incomplete information is a model with a single unknown.
  - E.g., strategic communication of Crawford-Sobel (1982), Milgrom (1981).
  - The expert knows one thing the non-expert doesn't. (Lecture 2)
- An impressive and powerful apparatus has been built on top of this.
  - Dynamics, strategic play, organizational structure, etc.
- I go back and reexamine this simplifying assumption.
  - Show that it matters & suggest a way forward.
  - While imposing my own simplifications/limitations (trade-off).

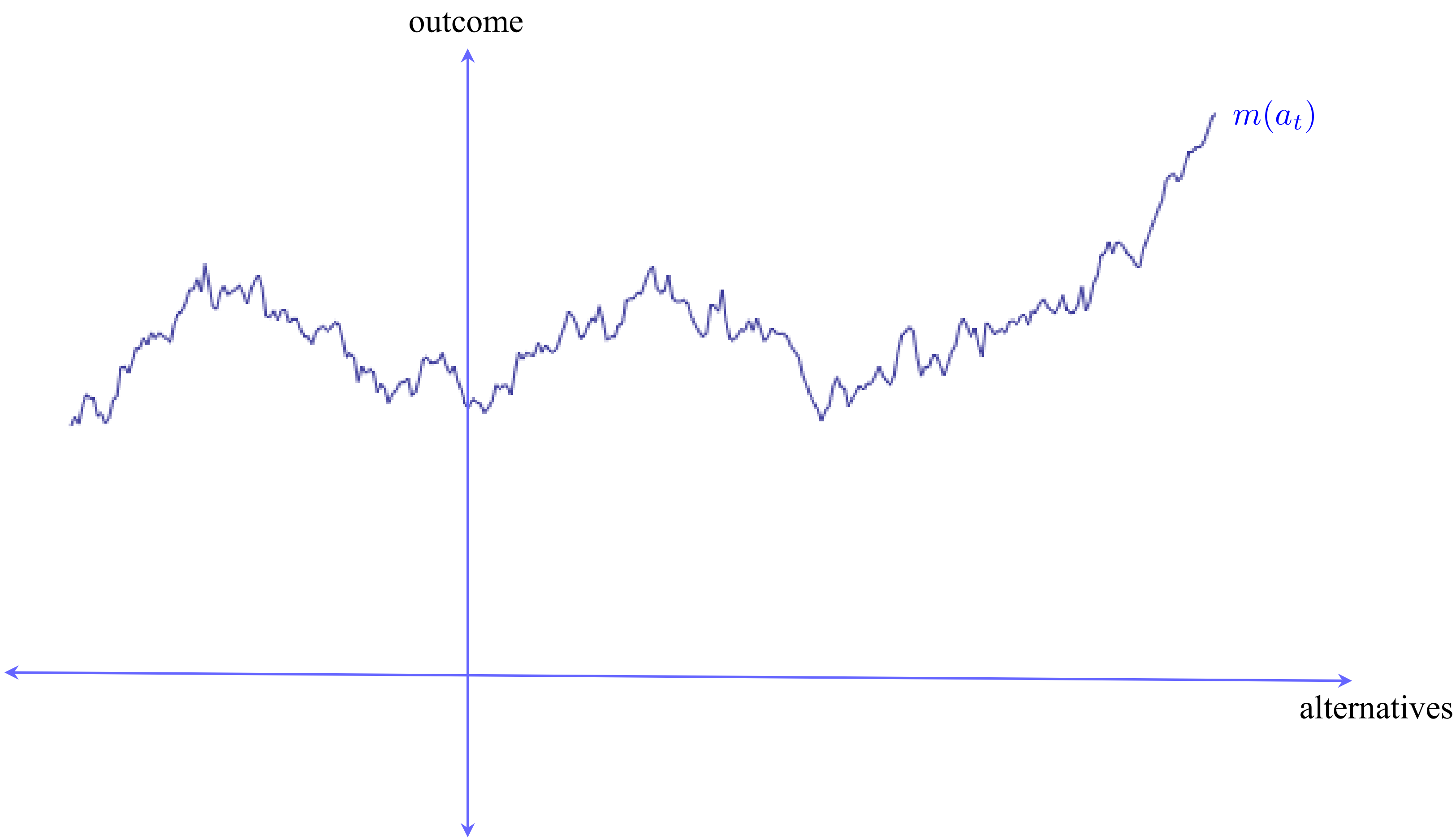
Enriching the Information Environment

- We can add more information in many ways.
- What did the job searcher learn from her experience at Goldman Sachs?
- Two properties of the decision problem stand out:
    1. Some sense of distance between alternatives.
    2. Correlation of outcomes is increasing in distance.
        - Goldman Sachs vs. hedge fund vs. school teacher.
        - Alternatives further from what you know are riskier.
- Place the alternatives along a line, R^1.
- Outcomes are given by a function $m: R \to R$.
    - What is the function m? Can it capture desiderata 1 and 2?

The Outcome Function

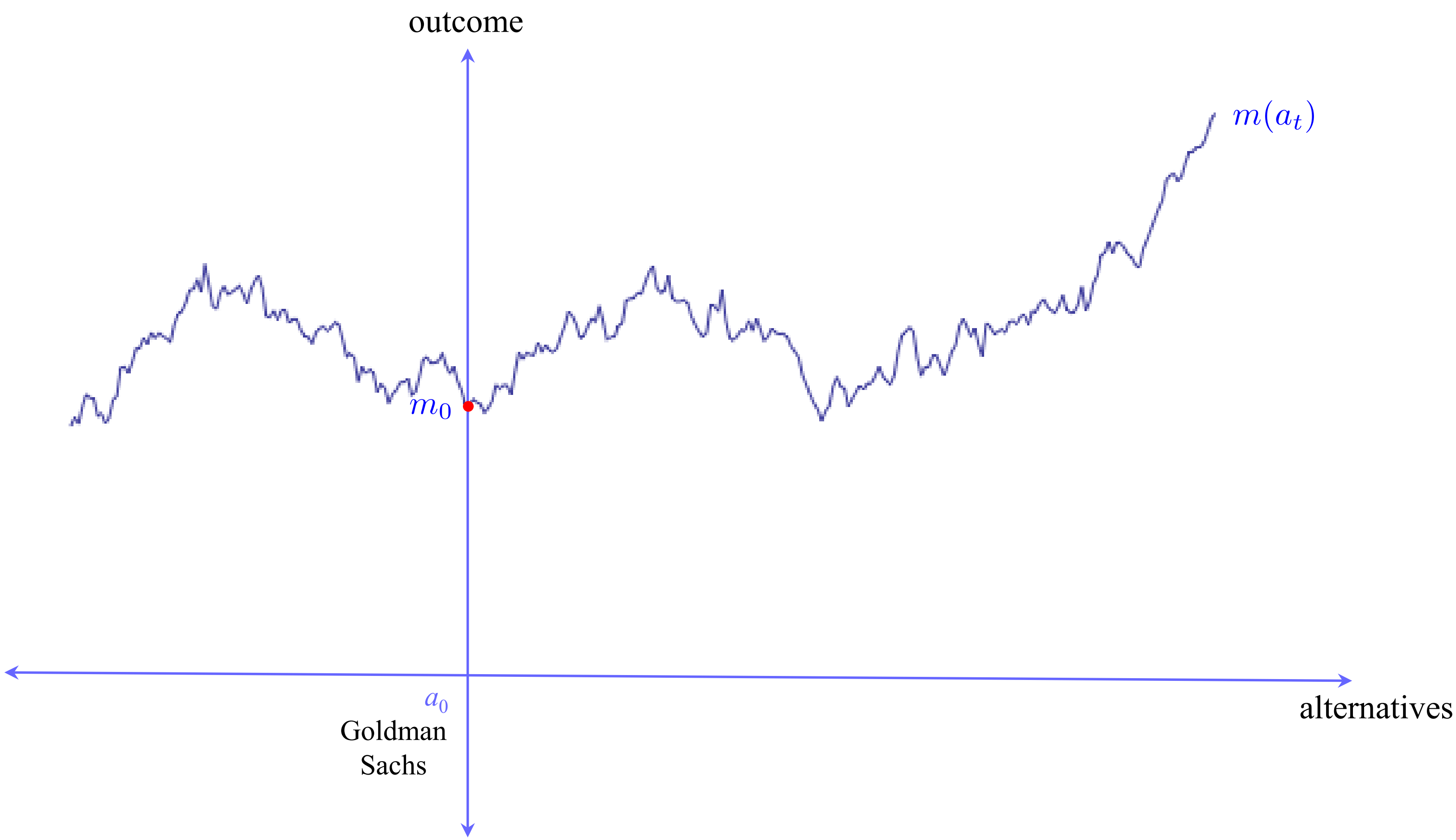


The Outcome Function

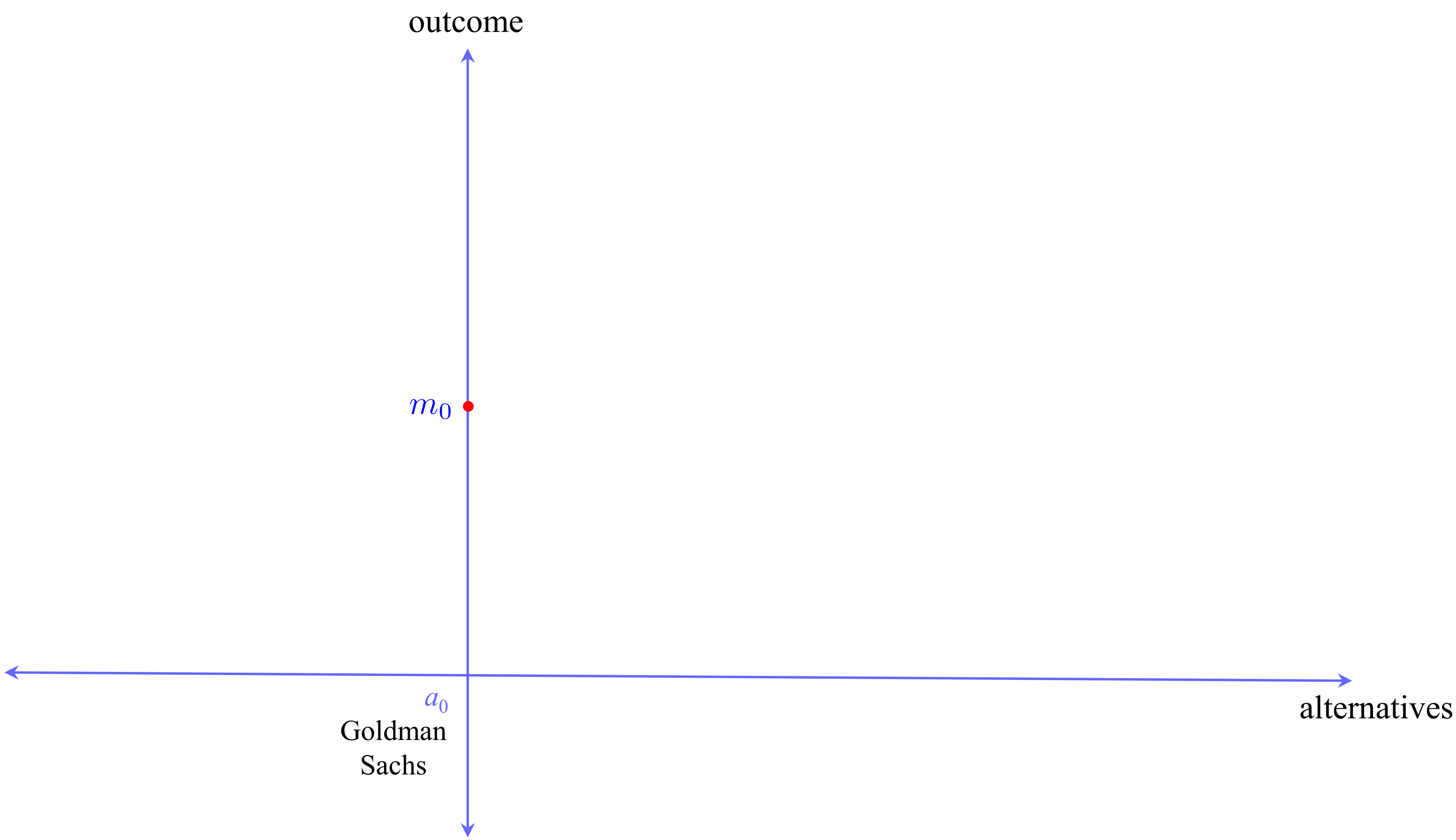


- Represent the mapping as the realized path of a Brownian motion.
  - With drift $\mu > 0$ and variance $\sigma^2 > 0$
- The path is not evolving through time – evolves through alternative space.

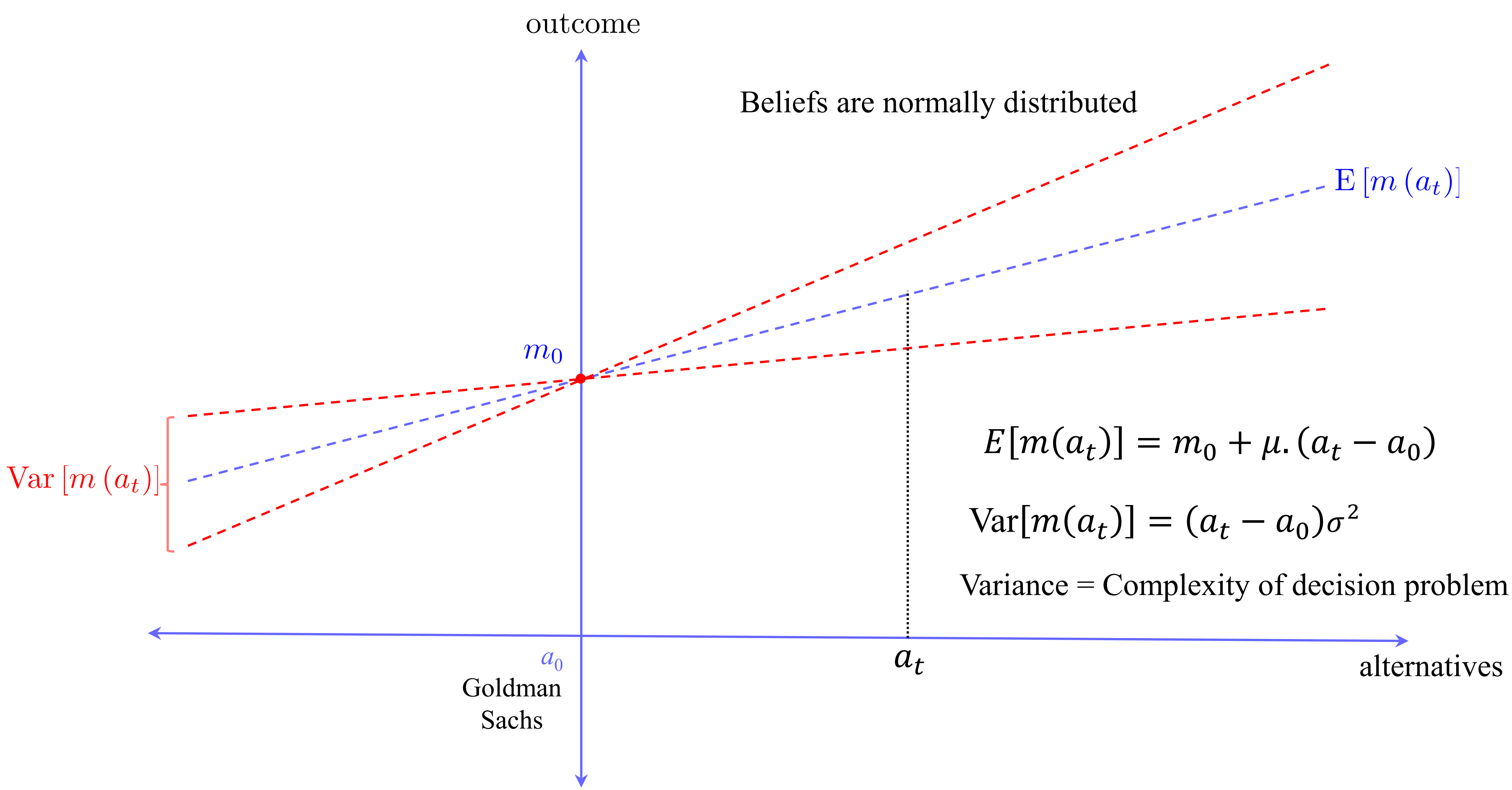
Why a Brownian Motion?



Why a Brownian Motion?



Beliefs

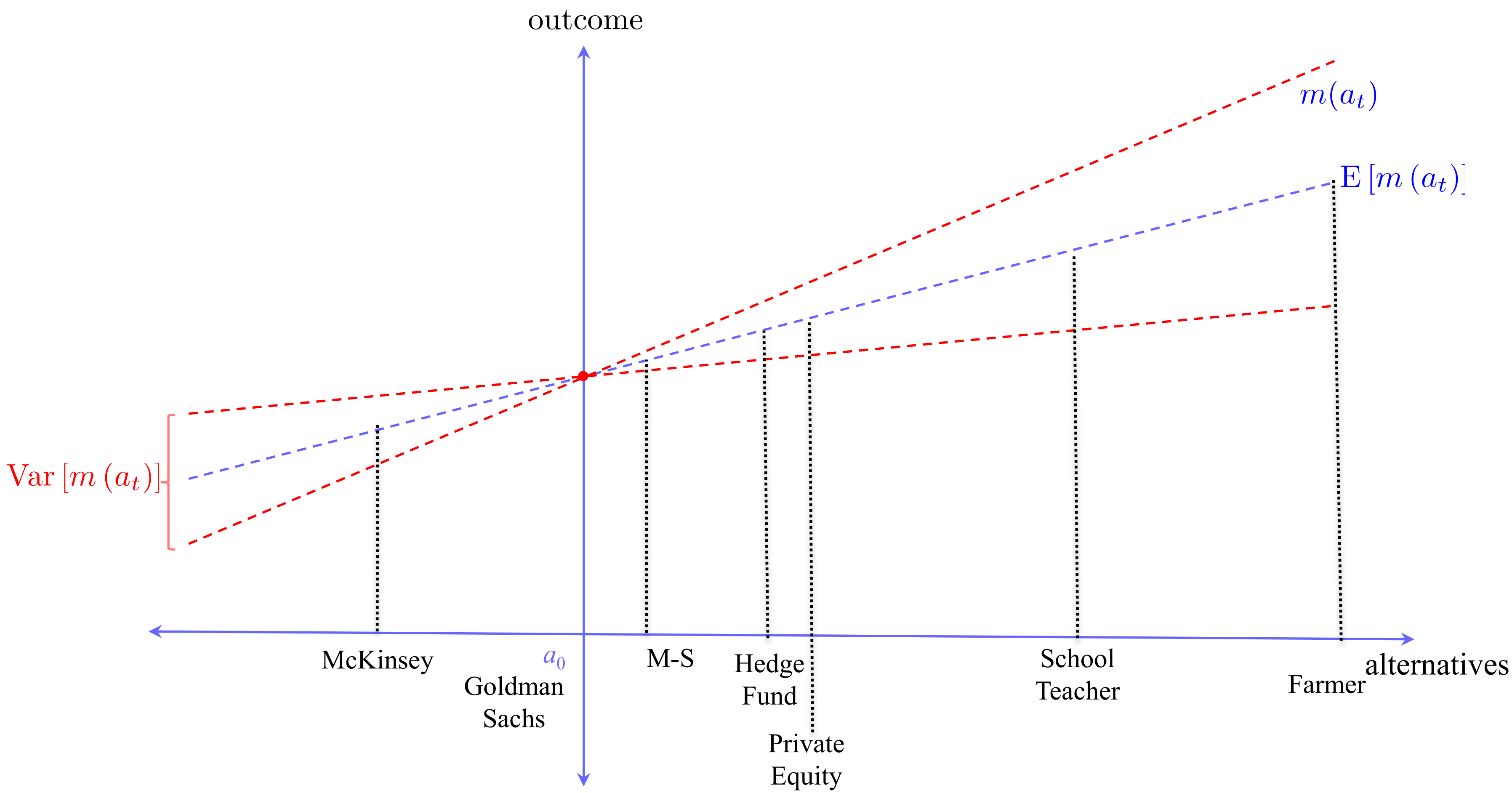


$$E[m(a_t)] = m_0 + \mu.(a_t - a_0)$$

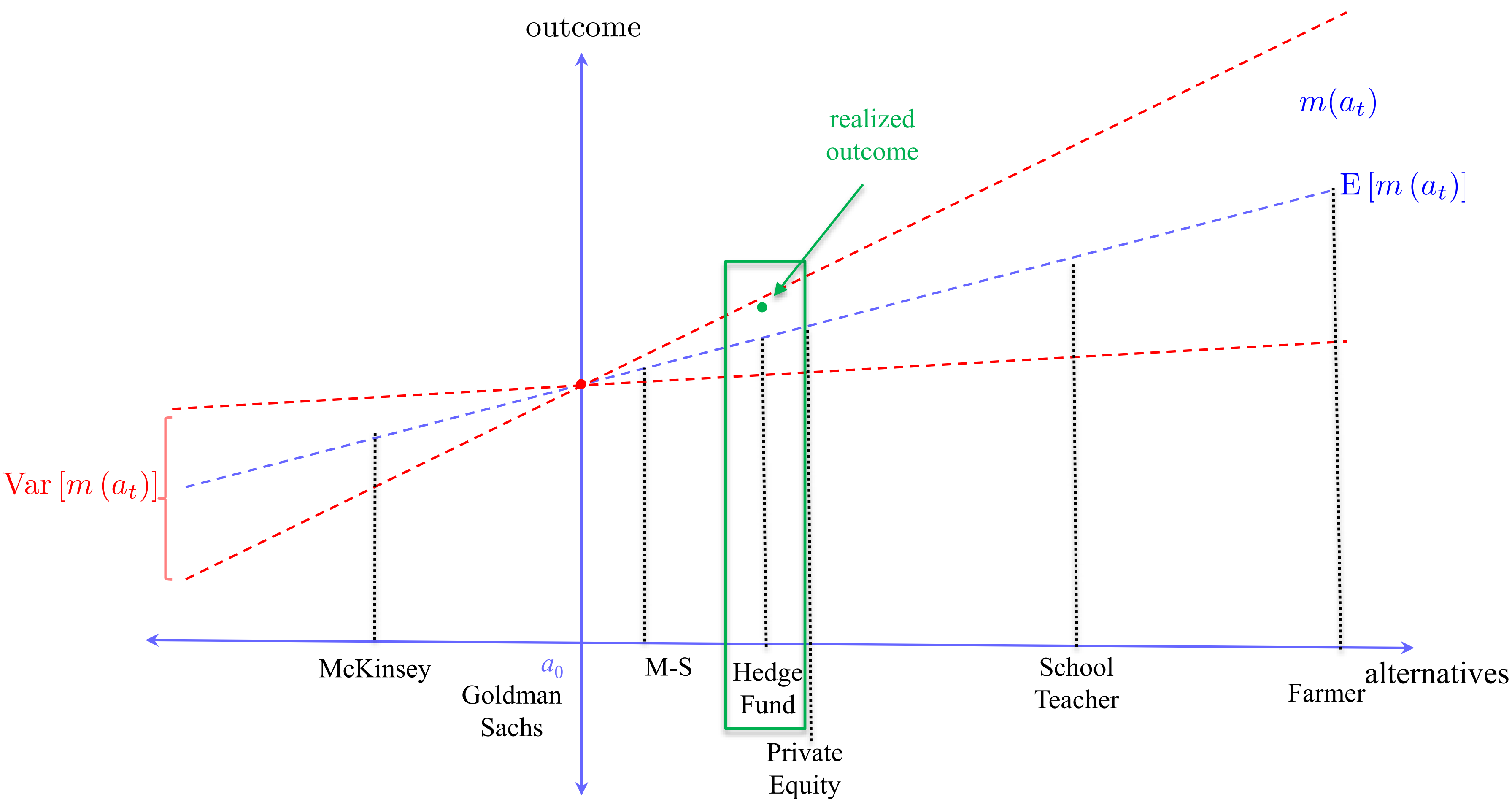
$$\mathrm{Var}[m(a_t)] = (a_t - a_0)\sigma^2$$

Variance = Complexity of decision problem

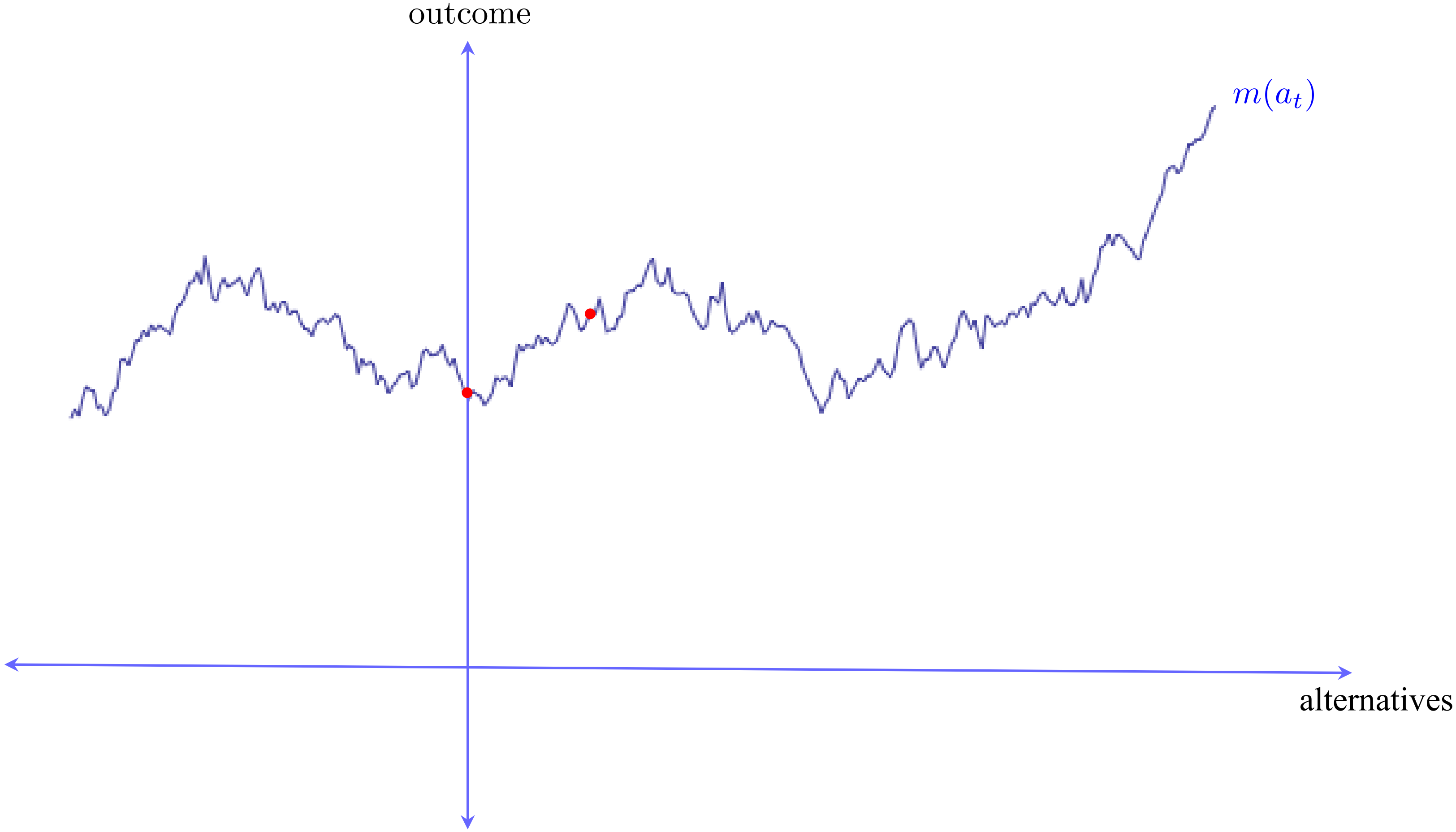
Learning by Trial and Error



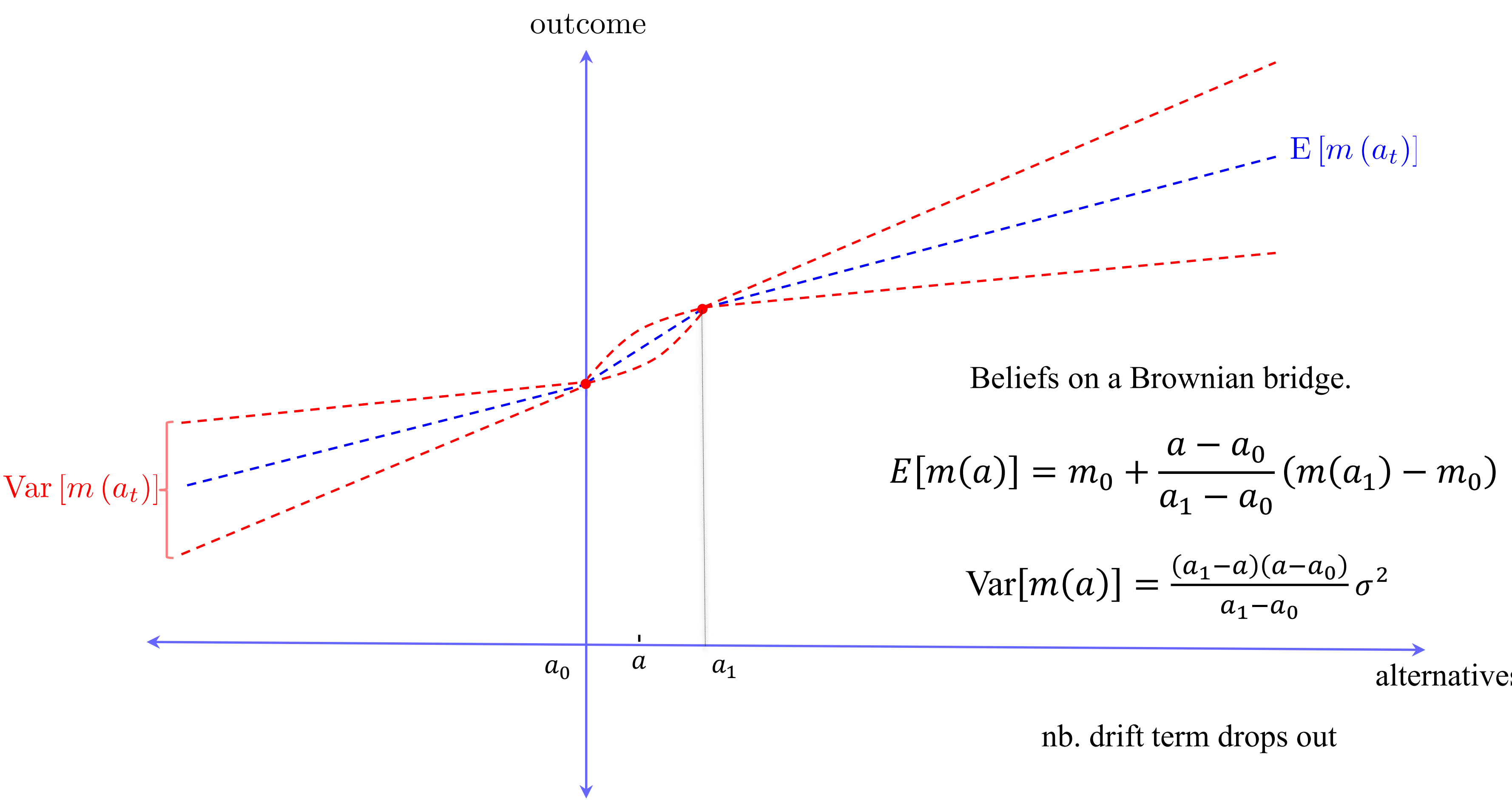
Learning by Trial and Error



Belief Updating



Belief Updating



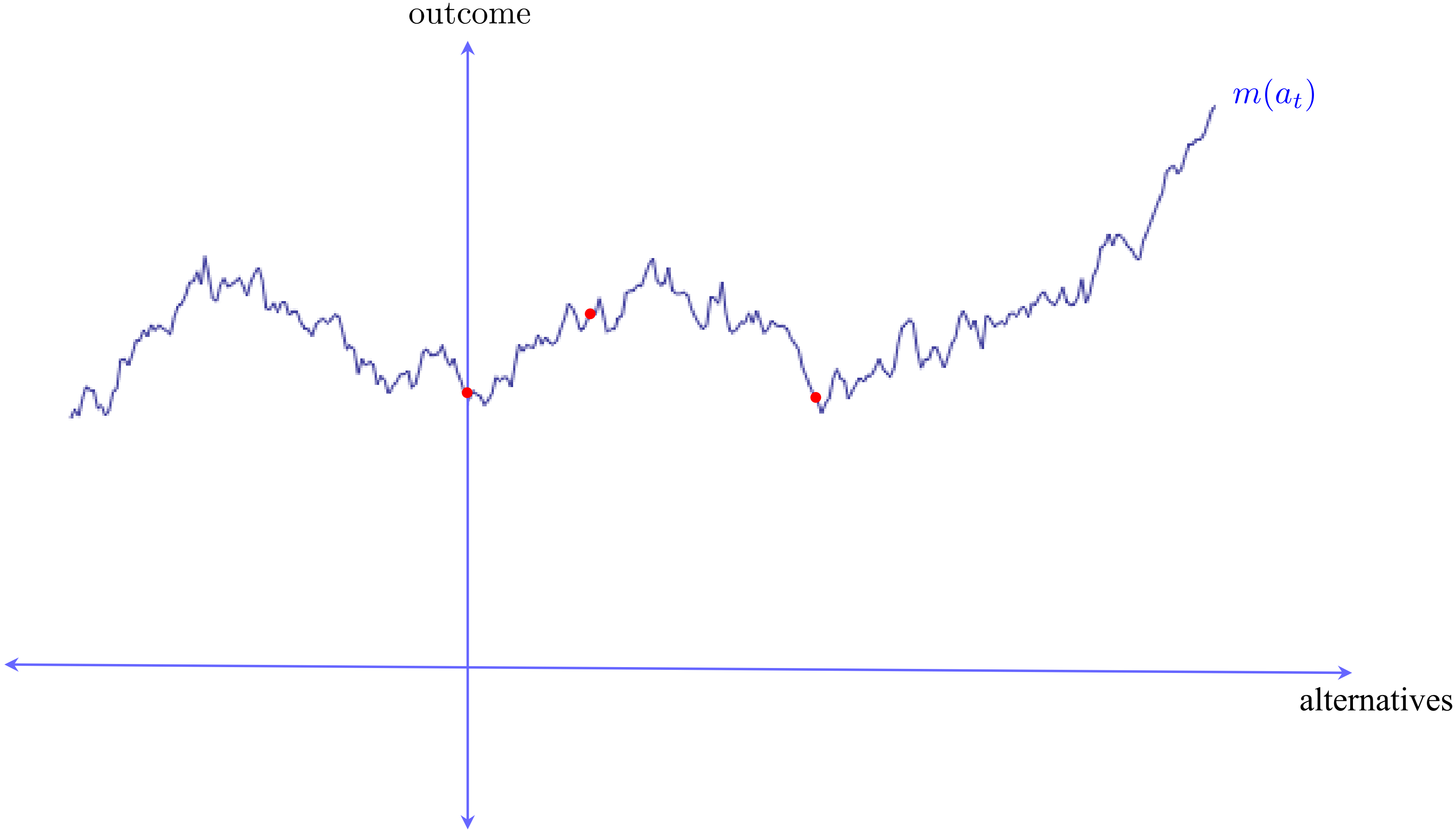
Beliefs on a Brownian bridge.

$$E[m(a)] = m_0 + \frac{a - a_0}{a_1 - a_0}(m(a_1) - m_0)$$

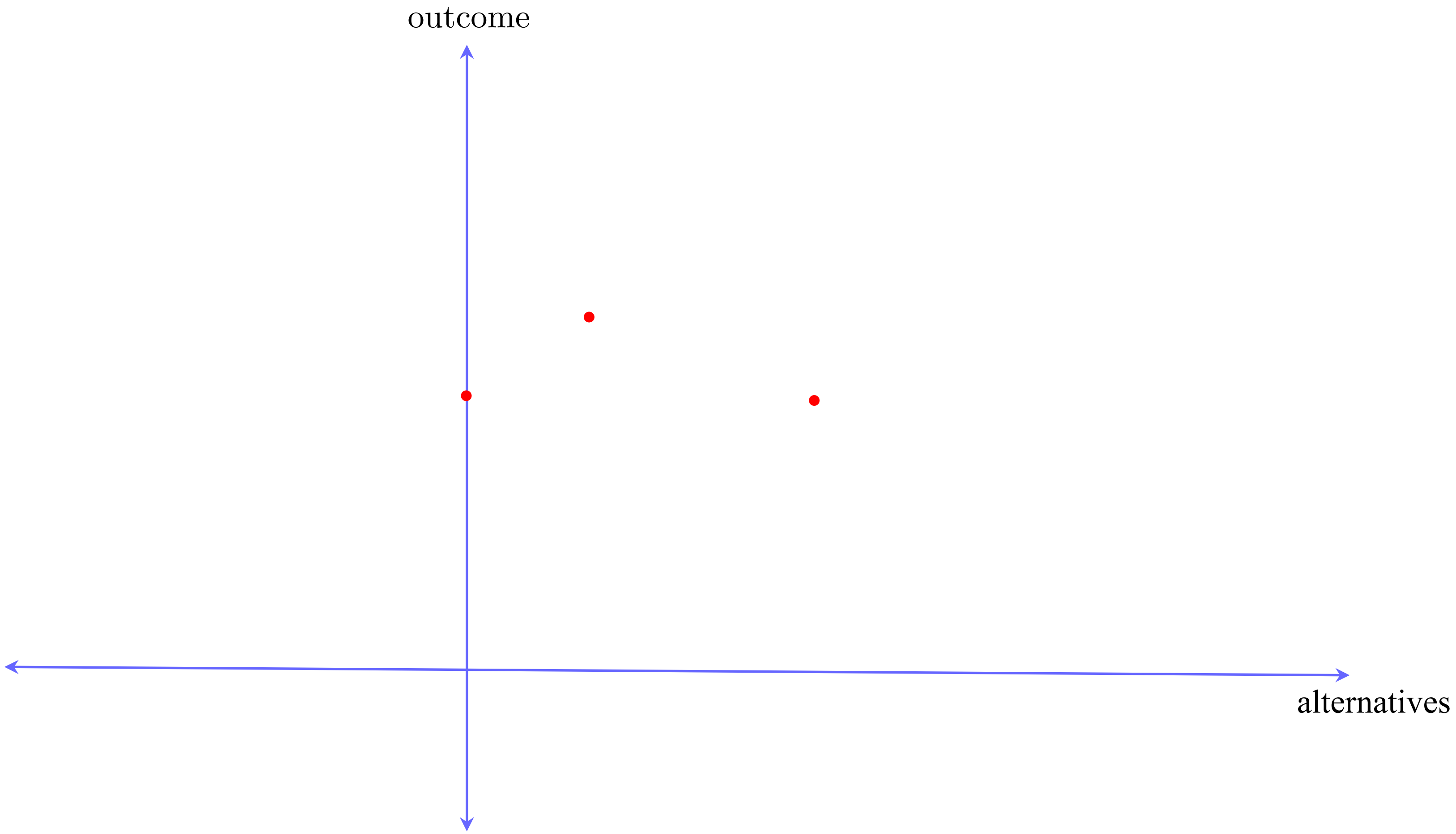
$$\text{Var}[m(a)] = \frac{(a_1 - a)(a - a_0)}{a_1 - a_0}\sigma^2$$

nb. drift term drops out

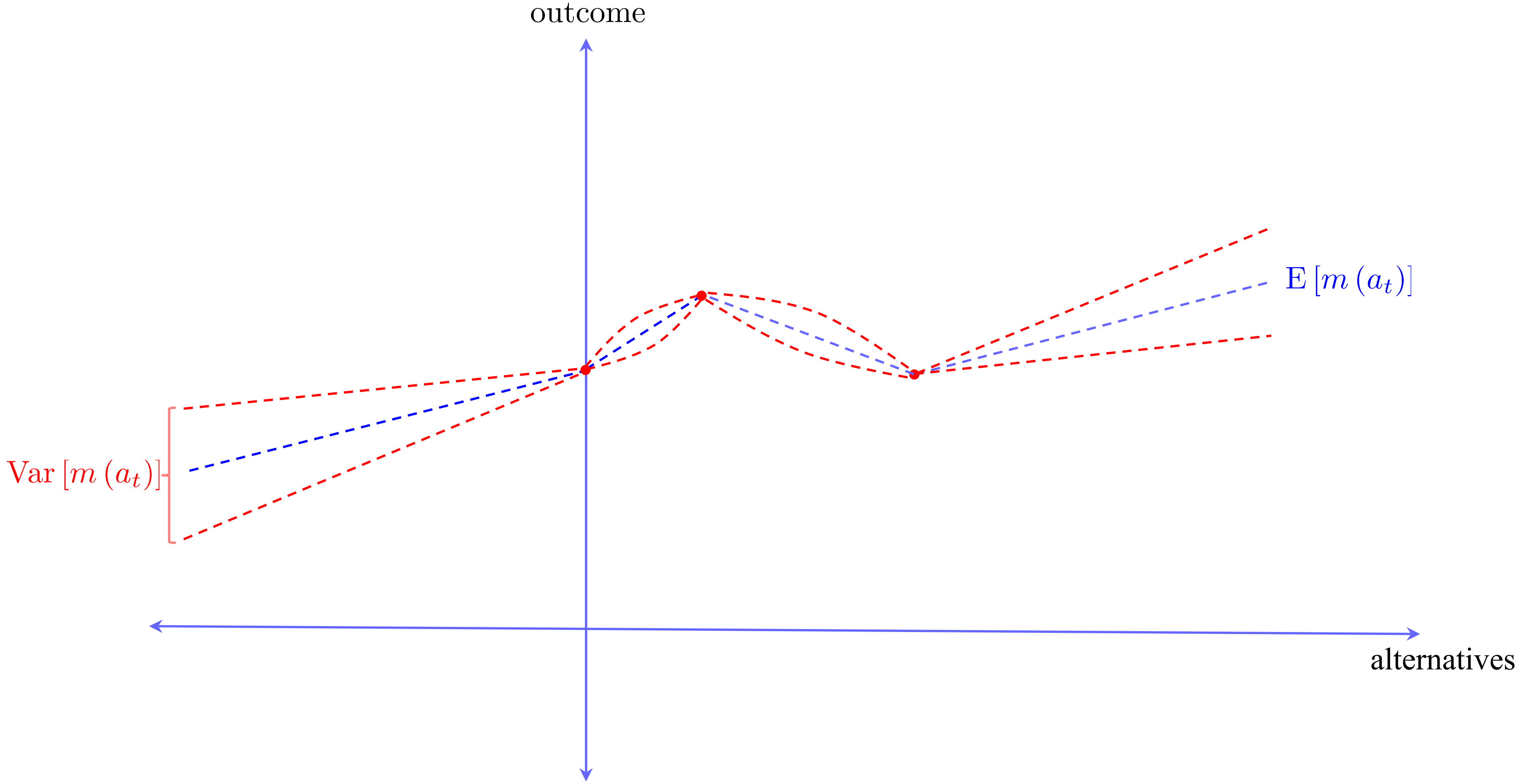
Belief Updating



Belief Updating



Belief Updating



- Maybe there should be a jump between finance and other industries.
- Sure! I won't do it today, but many ways to capture reality.
- Core idea the same: "nearby" choices produce "nearby" outcomes.

Embed in a Model of Search & Experimentation

- **Agents learn by trial and error.**
  - **Know "where they are" – status quo a_0 and $m_0 = m(a_0)$.**
  - **Possess theoretical knowledge $\mu > 0$ and $\sigma^2 > 0$**
  - **Obtain practical knowledge by experience.**

- Choice of alternative in each period t=1,2,3...
  - Sequential decisions by singly myopic agent or a sequence of short-lived agents.
  - (Relax partially—more later.)

- Related literature: Models of job search & bandit problems.
  - Far-sighted agents but no correlation.
  - Brownian motion = multi-armed bandit with correlated & deterministic arms.

Searching for What?

- What are agents trying to achieve?
  - Identify an ideal outcome?

    politics
  - Improve performance in unbounded way?

    firms, economic development, etc.
  - Identify all of the alternatives that satisfy a threshold?

    judicial decision making
  - Make sure every alternative can be sorted into buckets with high-enough confidence?
  - In all cases, agents must learn by trial-and-error (unless there is an expert to help!).

    communication – Session 2

Risk Distorts Decision Making in Orgs

- Arrow (1962):
  Risk averse managers engage in "*discrimination against risky enterprises as compared with the optimum*."

- Ed Catmull (2014), Founder of Pixar:

  "*While experimentation is scary to many, I would argue that we should be far more terrified of the opposite approach. Being too risk averse causes many companies to stop innovating and reject new ideas, which is the first stop on the path to irrelevance.*

- Question: How does risk affect firm performance?
  - When do firms experiment? How do they experiment?
  - When do they stop?
  - What are the implications for long run performance?

Preferences

- Utility function $u(m_t)$ satisfies
  - Non-satiation: $u'(m_t) > 0$
  - Risk aversion: $u''(m_t) < 0$
  - Standard Risk Aversion (Kimball 1993):
    - Logic: an undesirable risk cannot be made more desirable by an independent, loss-aggravating risk:

      $\mathrm{E}[u(m_t + x)] - u(m_t) \leq 0$ and $\mathrm{E}[u'(m_t + y)] - u'(m_t) \geq 0$ imply

      $$\mathrm{E}[u(m_t + x + y) - u(m_t + y)] \leq \mathrm{E}[u(m_t + x) - u(m_t)]$$

    - satisfied by most commonly used utility functions
    - implies decreasing absolute risk aversion

Preferences

- Utility function $u(m_t)$ satisfies
  - Non-satiation: $u'(m_t) > 0$
  - Risk aversion: $u''(m_t) < 0$
  - Standard Risk Aversion (Kimball 1993)
  - Crossing Condition:
    - coefficient of absolute risk aversion $r(m_t)$ crosses $\boxed{2\mu/\sigma^2}$
    - (this is a recurring threshold in my models.)

Preferences

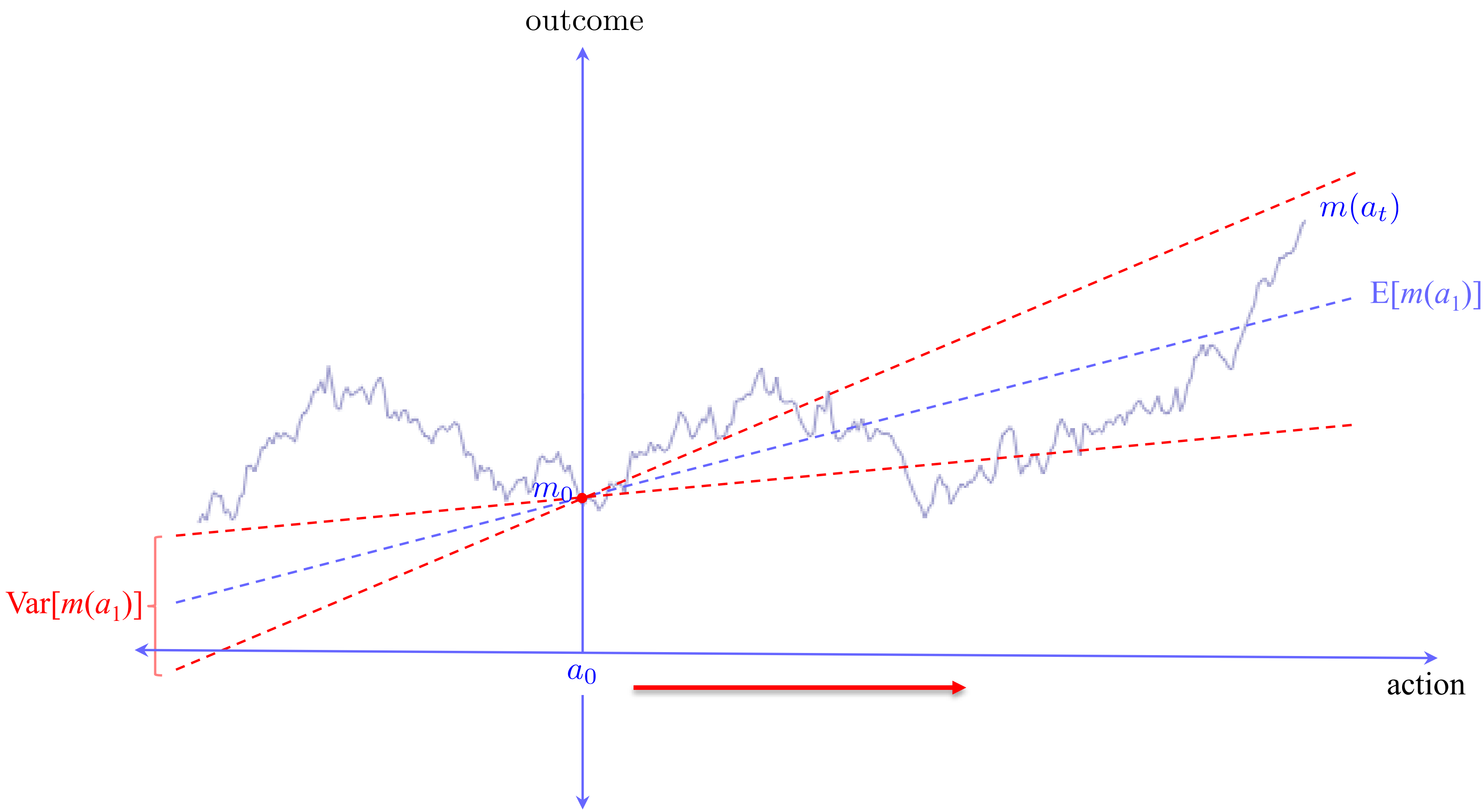
- Utility function $u(m_t)$ satisfies
  - Non-satiation: $u'(m_t) > 0$
  - Risk aversion: $u''(m_t) < 0$
  - Standard Risk Aversion (Kimball 1993)
  - Crossing Condition:
    - coefficient of absolute risk aversion $r(m_t)$ crosses $\boxed{2\mu/\sigma^2}$
    - (this is a recurring threshold in my models.)

- An example of preferences that satisfy these conditions:

$$u(m_t) = \alpha m_t - e^{-\beta m_t},$$

where $\alpha > 0$ and $\beta > 2\mu/\sigma^2$.

Directed Search



- Agent knows which direction to search is more likely to yield better outcomes
  - Directed search (note: not in bandit models).
- How far should she move? Should she move? Faces a risk-return trade-off.

The First Agent's Problem

random variable
drawn from
standard normal

payoff from status quo

$$\max_{\Delta_1 \geq 0} \mathrm{E}\Big[u\Big(m_0 + \mu\Delta_1 + \sigma\sqrt{\Delta_1}\,z\Big)\Big]$$

$\Delta_1 = a_1 - a_0 \geq 0$

expected gain in payoff

The First Agent's Problem

- Expected utility

$$\mathrm{E}\Big[u\Big(m_0 + \mu\Delta_1 + \sigma\sqrt{\Delta_1}\,z\Big)\Big]$$

  is concave in Δ_1.

  $\leftrightarrow$ Standard Risk Aversion

- A non-trivial solution to the first agent's problem exists.

  $\leftrightarrow$ Crossing Condition

The First Agent's Problem

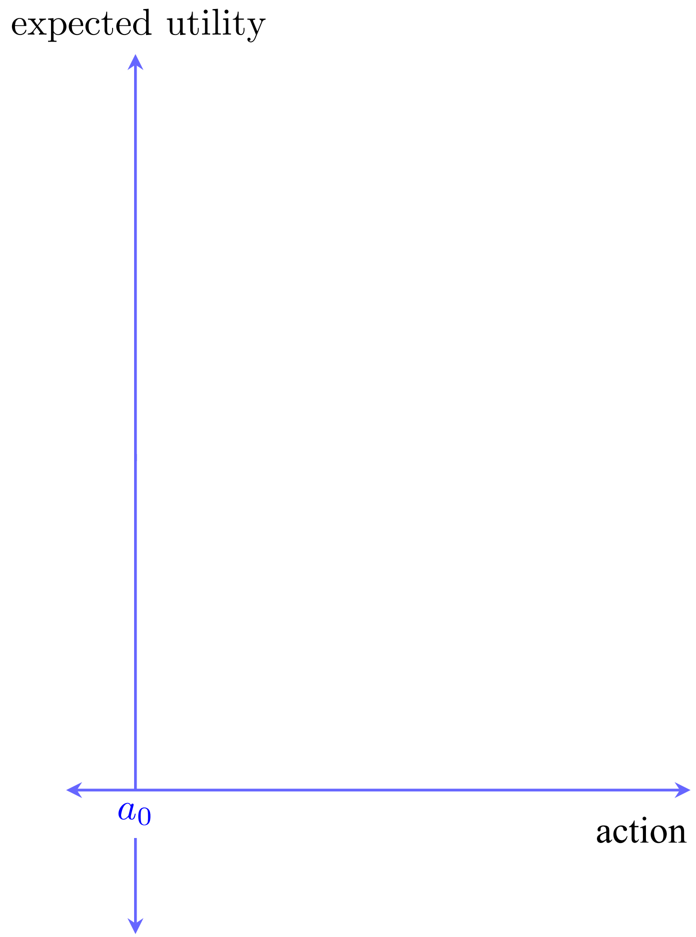
- Marginal expected utility is given by

$$\frac{\mathbf{dE}[u(m_1)]}{\mathbf{d}\boldsymbol{\Delta}_1} = \mathbf{E}\left[u'(m_1)\right] \frac{\sigma^2}{2} \left(\boxed{\frac{2\mu}{\sigma^2} - \left(- \frac{\mathbf{E}\left[u''(m_1)\right]}{\mathbf{E}[u'(m_1)]} \right)} \right),$$

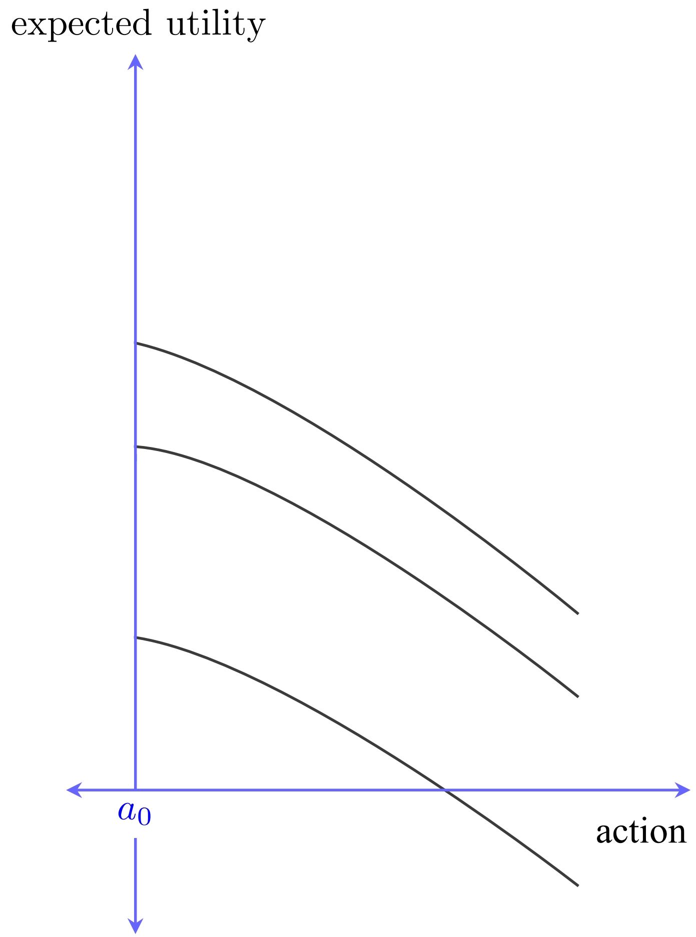
  where

$$m_1 = m_0 + \mu \Delta_1 + \sigma \sqrt{\Delta_1} z$$

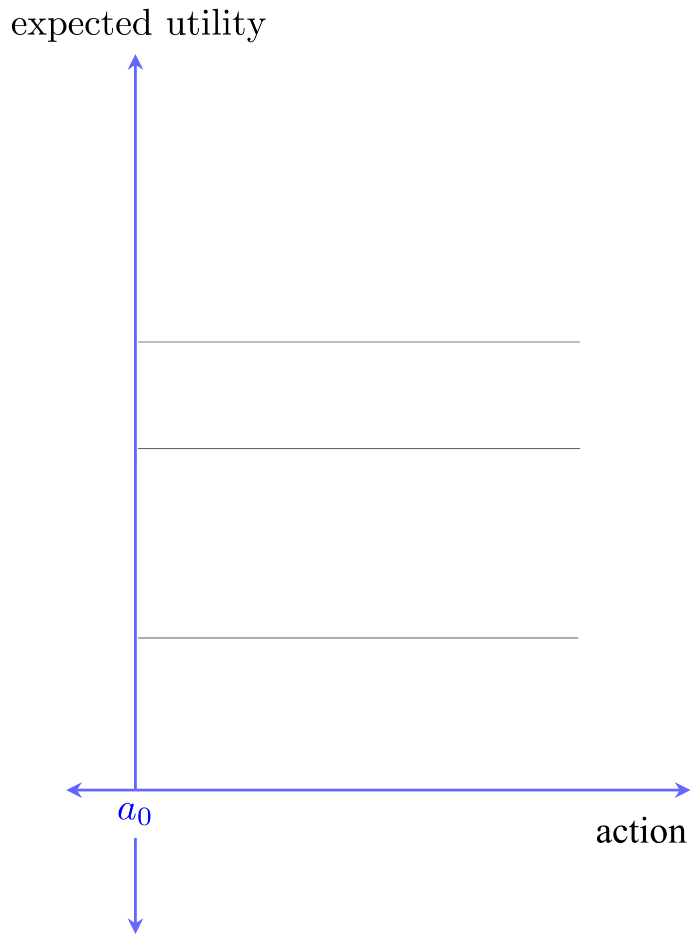
EXPECTED UTILITY IF $r(m) > 2\mu/\sigma^2$ FOR ALL m



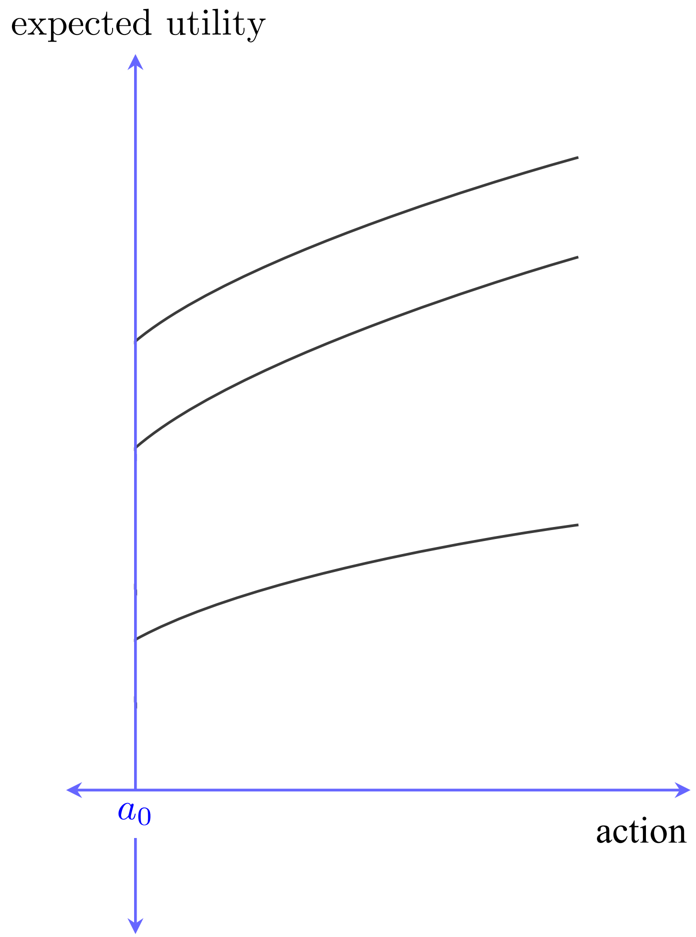
EXPECTED UTILITY IF $r(m) > 2\mu/\sigma^2$ FOR ALL m



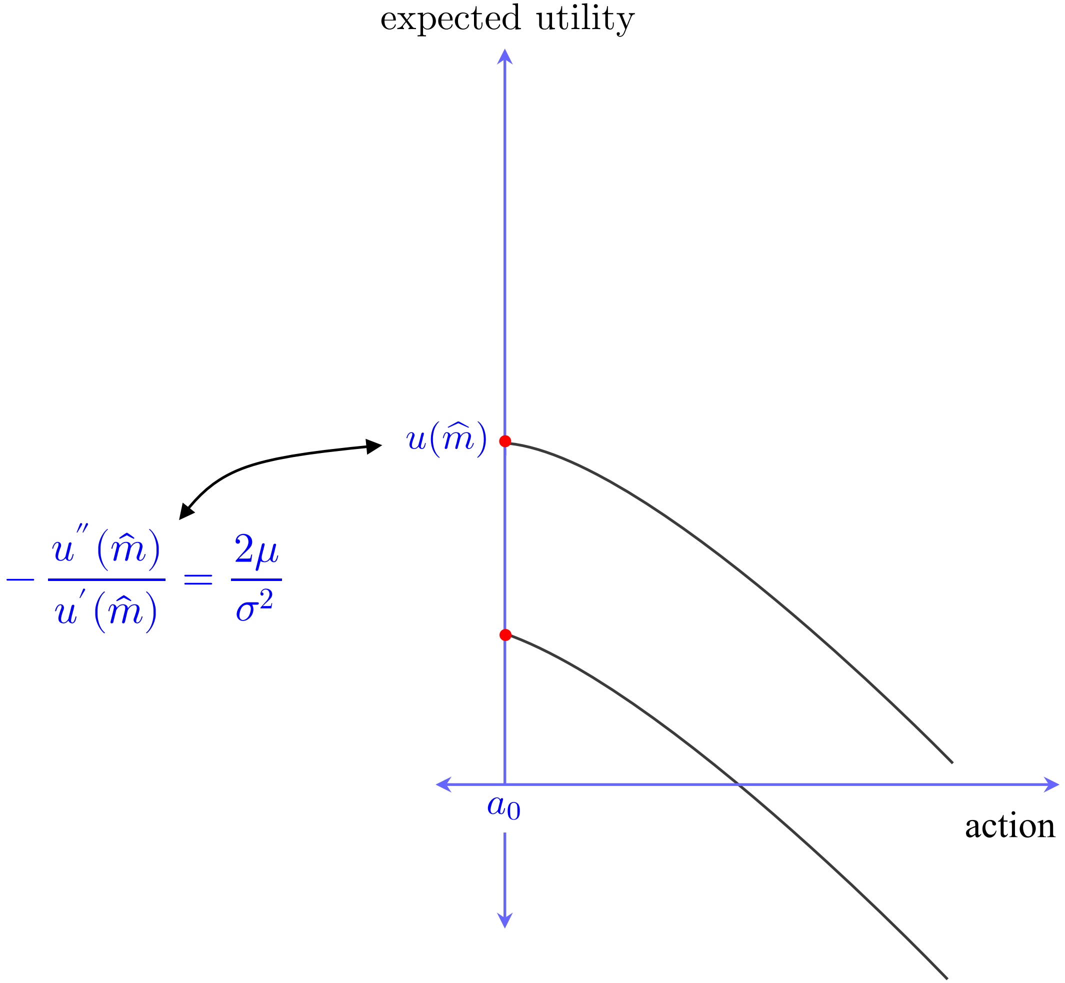
Expected Utility if $r(m) = 2\mu/\sigma^2$ For All m



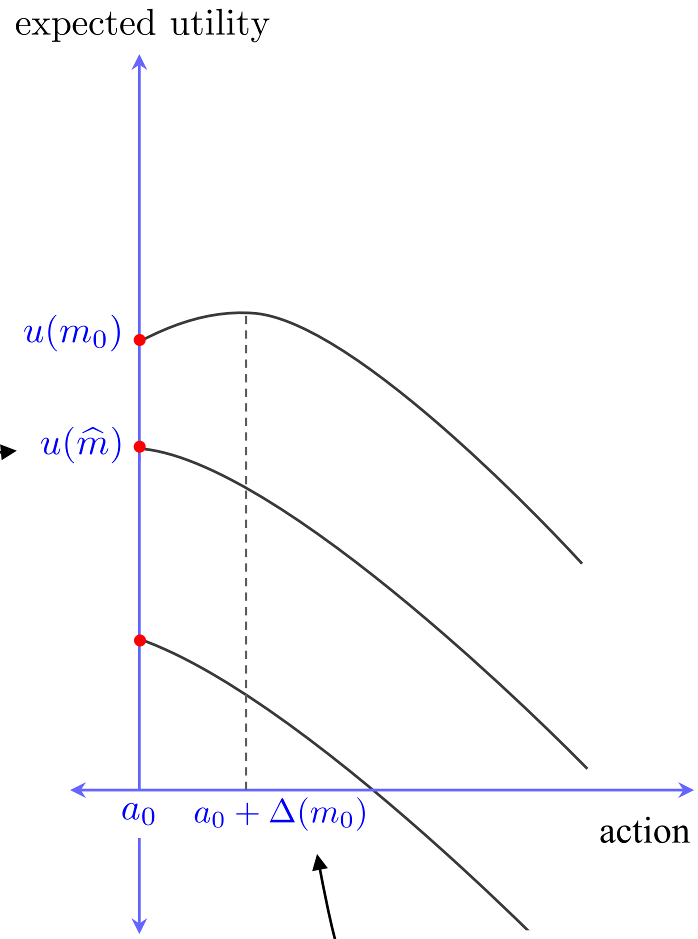
Expected Utility if $r(m) < 2\mu/\sigma^2$ For All m



EXPECTED UTILITY IF $r(m)$ CROSSES $2\mu/\sigma^2$



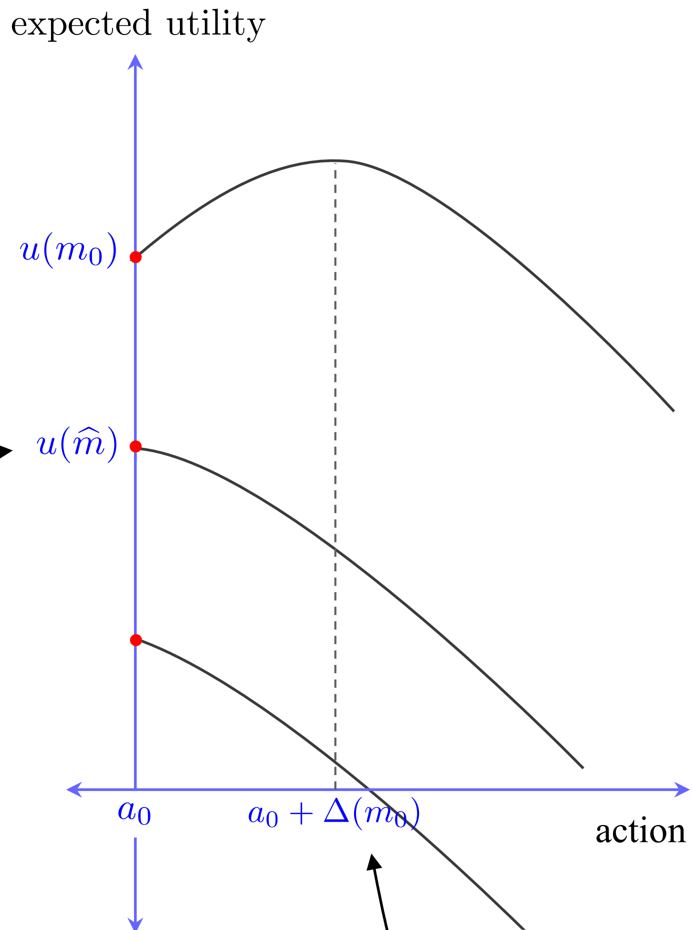
Expected Utility if $r(m)$ Crosses $2\mu/\sigma^2$



$$-\frac{u''(\hat{m})}{u'(\hat{m})} = \frac{2\mu}{\sigma^2}$$

$$-\frac{\mathbf{E}\left[u''\left(m_0 + \mu\Delta + \sqrt{\Delta}\sigma z\right)\right]}{\mathbf{E}\left[u'\left(m_0 + \mu\Delta + \sqrt{\Delta}\sigma z\right)\right]} = \frac{2\mu}{\sigma^2}$$

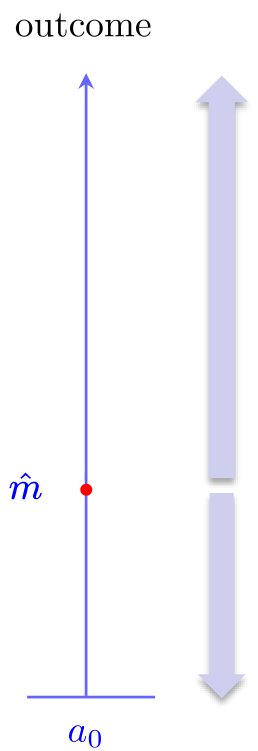
Expected Utility If $r(m)$ Crosses $2\mu/\sigma^2$



$$-\frac{u''(\hat{m})}{u'(\hat{m})} = \frac{2\mu}{\sigma^2}$$

$$-\frac{\mathbf{E}\left[u''\left(m_0 + \mu\Delta + \sqrt{\Delta}\sigma z\right)\right]}{\mathbf{E}\left[u'\left(m_0 + \mu\Delta + \sqrt{\Delta}\sigma z\right)\right]} = \frac{2\mu}{\sigma^2}$$

Optimal Learning—First Agent



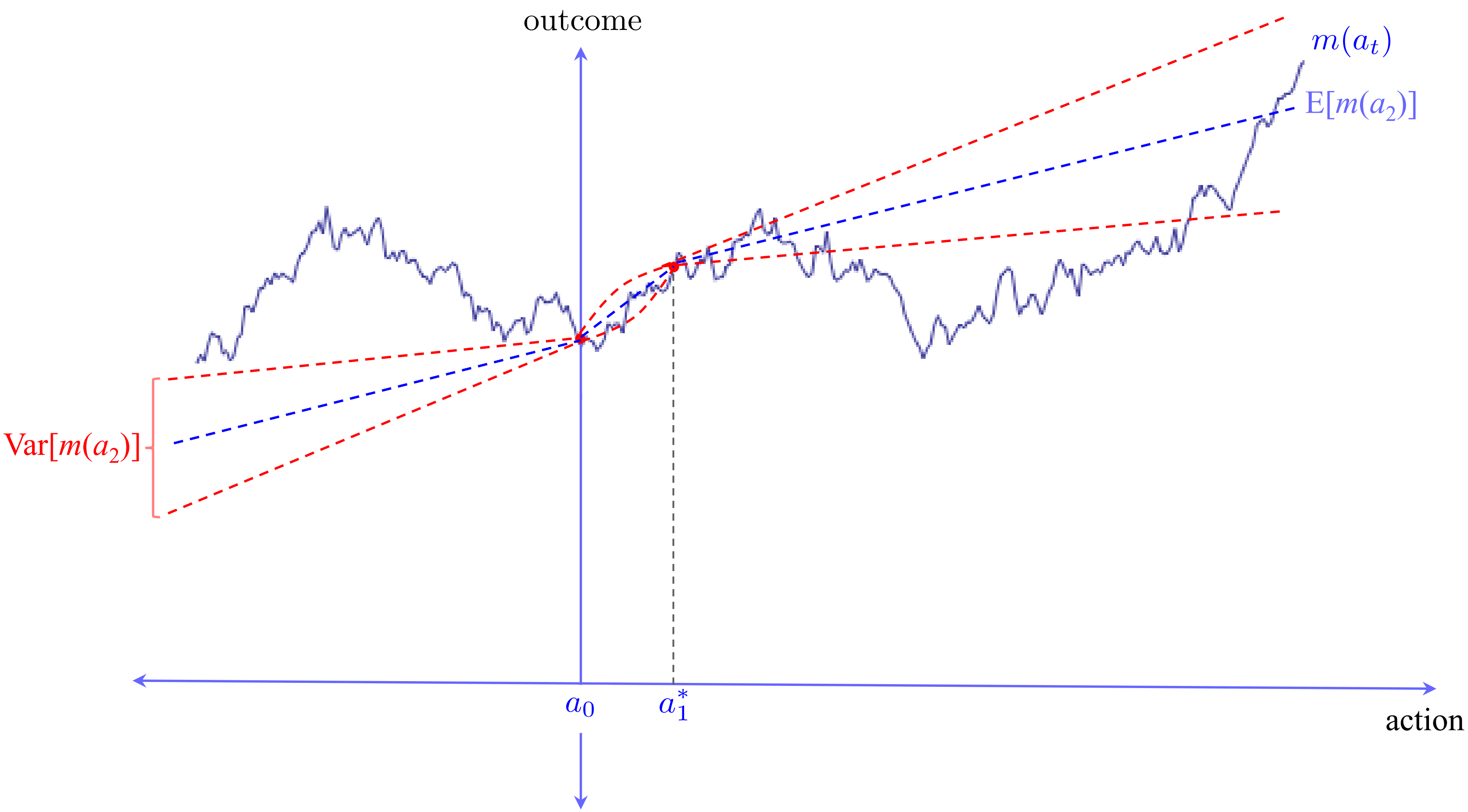
take unique action $a_0 + \Delta(m_0)$

increasing in m_0 and μ
decreasing in σ^2 and
risk aversion

decreasing in μ
increasing in σ^2 and
risk aversion

stay at a_0

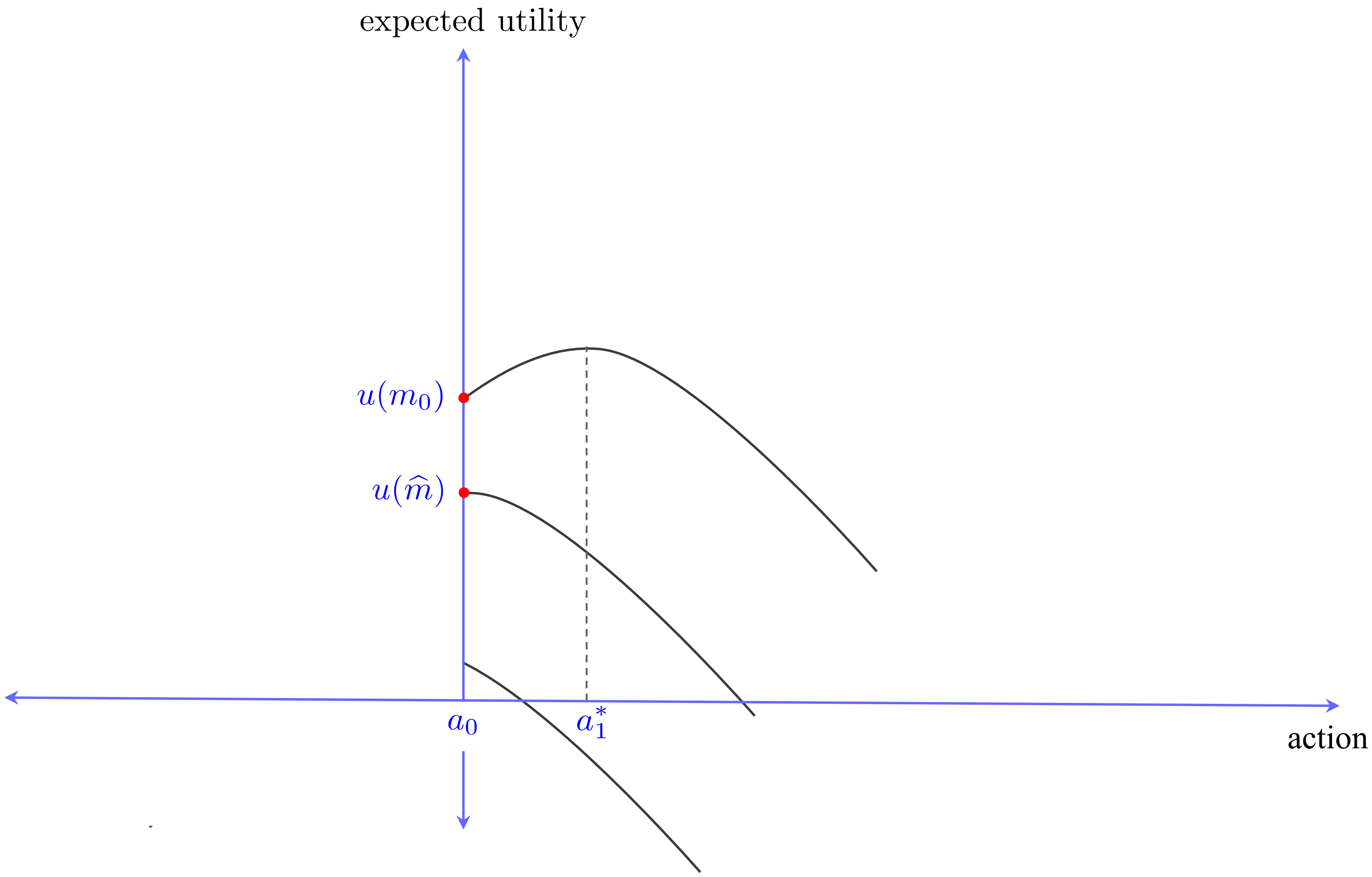
Second Agent—Beliefs



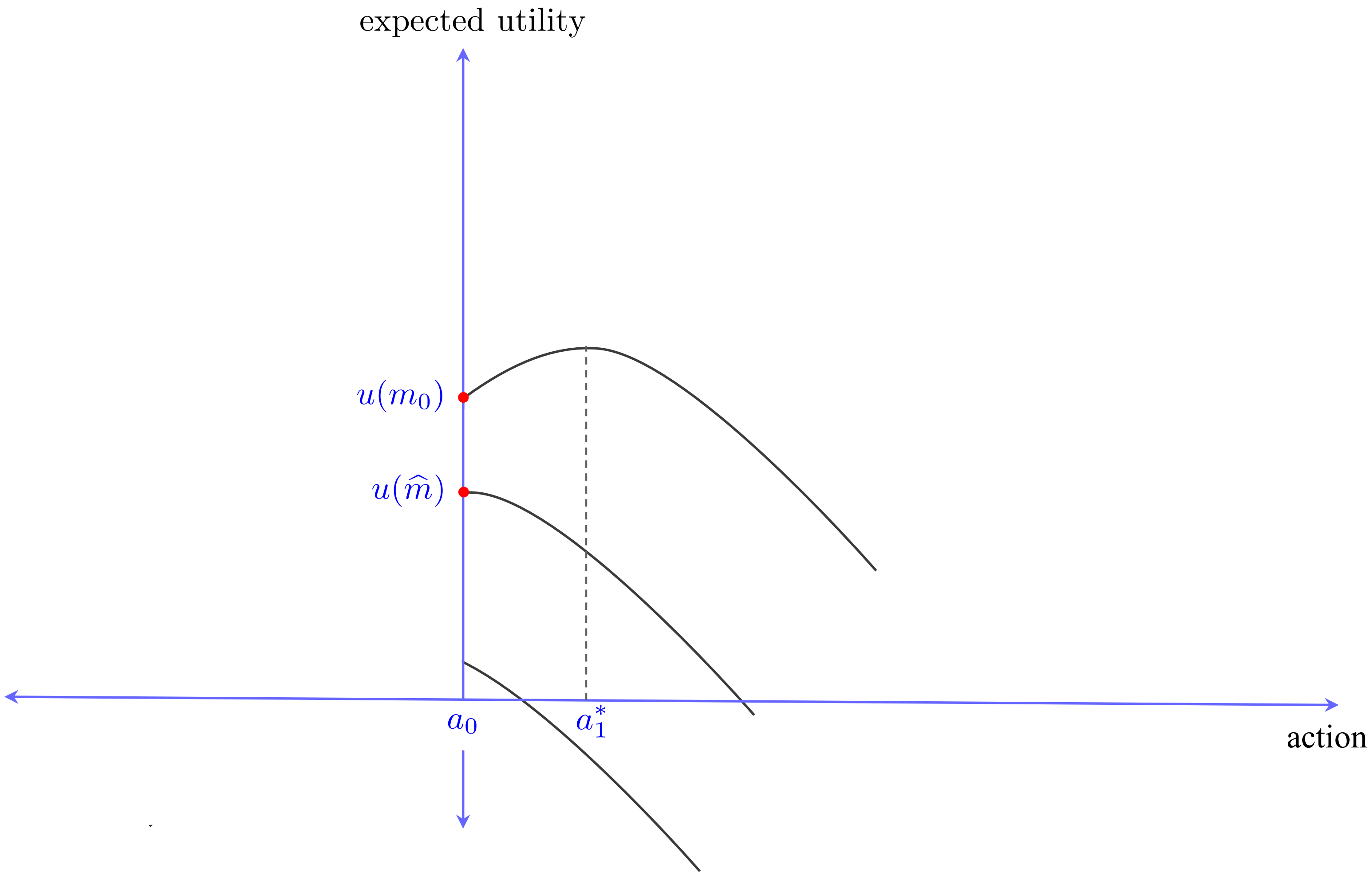
Second Agent—The Constrained Problem

$$\max_{\Delta_2 \geq 0} \mathrm{E}\left[u\left(m_1^* + \mu\Delta_2 + \sigma\sqrt{\Delta_2}\,z\right)\right]$$

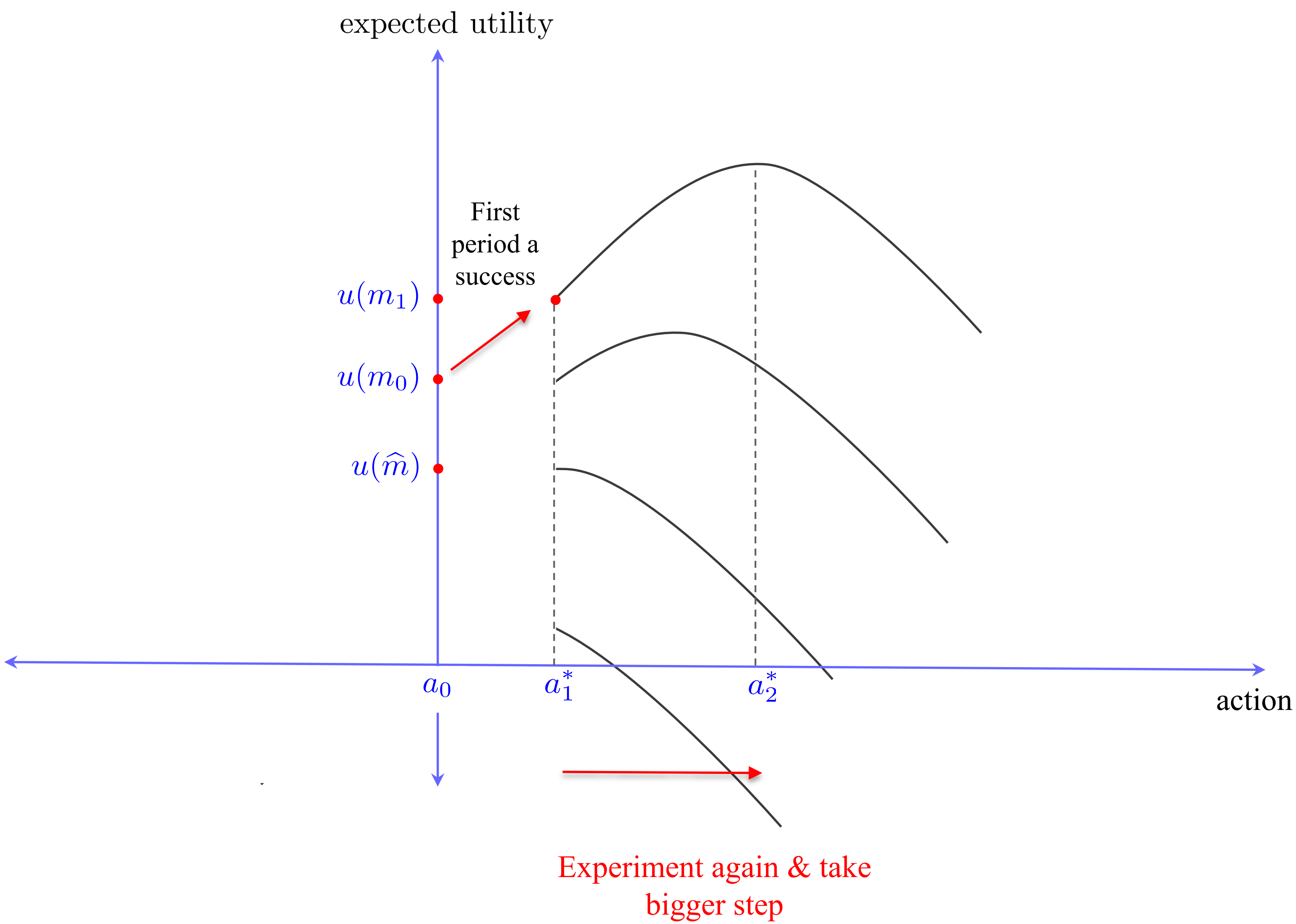
First Agent—The Problem



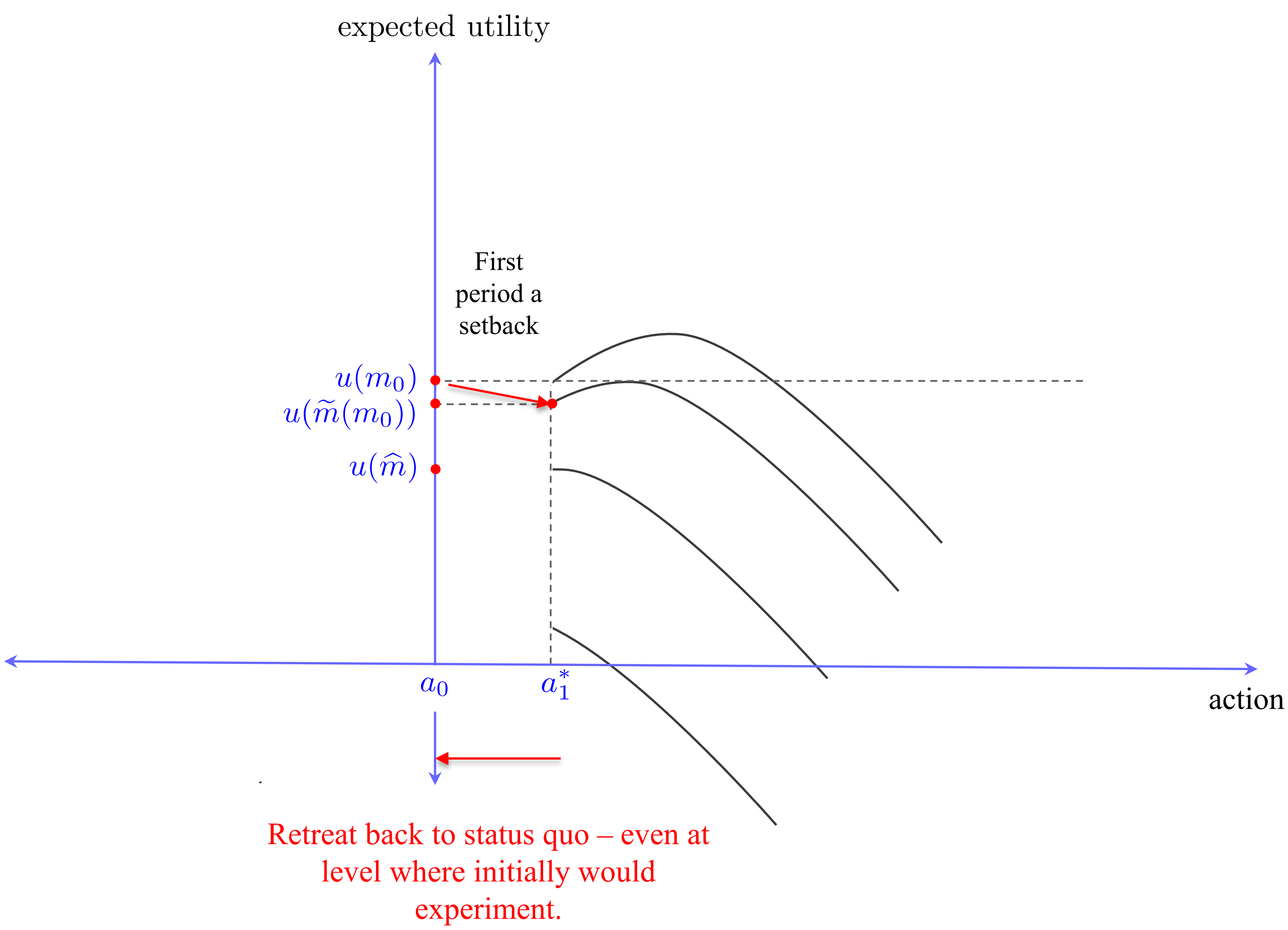
Second Agent—The Constrained Problem



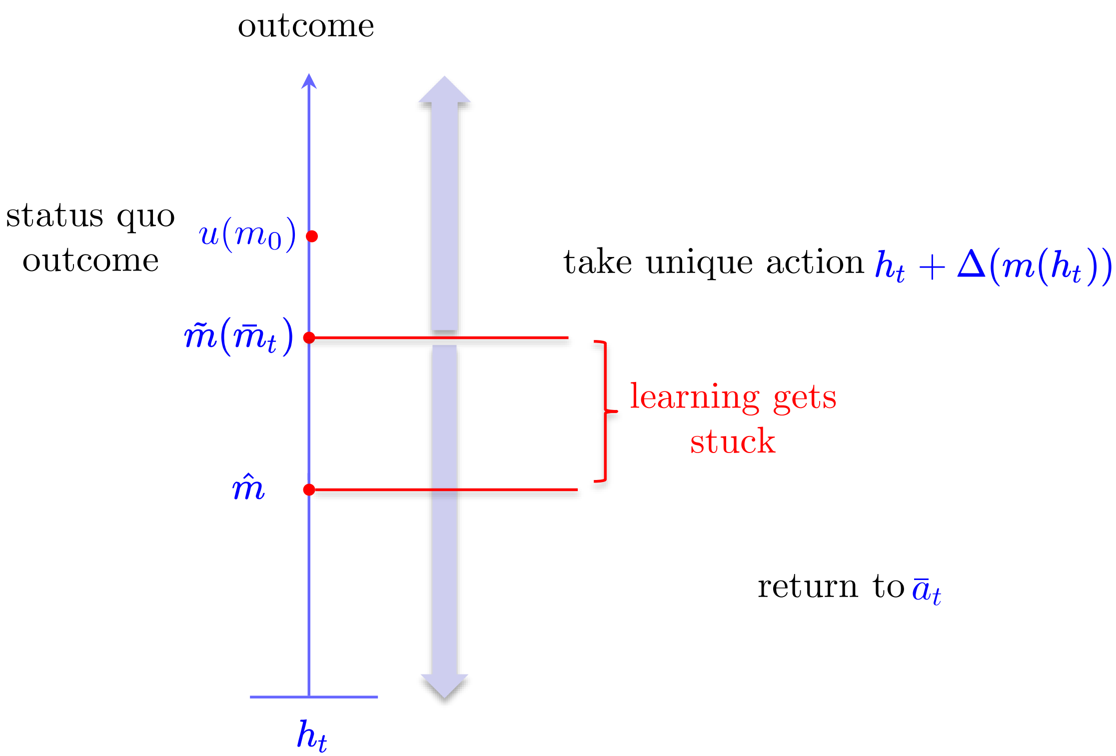
Second Agent—The Full Problem



Second Agent—The Full Problem



Optimal Learning—Period $t > 1$



Search Behavior: Summary

- If status quo outcome is sufficiently low, experimentation never starts.
  - a risk-aversion "poverty trap."
- Learning can get "stuck" following a sub-standard but not terrible performance.
  - Definition of "sub-standard" is increasing in performance.
- Optimal action depends on only two factors: frontier & peak performance.
  - Search continues unless frontier drops sufficiently below peak performance.
  - Lower performance makes agents more "tentative" in experimenting.
  - Not because pessimistic, but because less risk tolerant.
  - Reverse holds for success at the frontier.
  - Complementarity between risk tolerance and experimentation.

Long-Run Performance

- Within period:

  less risk aversion → more experimentation

- Across periods:

  more experimentation → less risk aversion

- Complementarity between risk tolerance and experimentation.
- Optimal learning propagates initial performance differences.

Long-Run Performance

- Suppose there are two "organizations": H and L.
- The production function for organization $k = H, L$ is characterized by drift μ, variance σ^2, and status quo income ${m_0}^k$, where

$$m_0^H > m_0^L > \hat{m}.$$

- Proposition: We then have

$$\mathbf{E}_1\left[m_t^*(m_0^H) - m_t^*(m_0^L)\right] > m_0^H - m_0^L$$

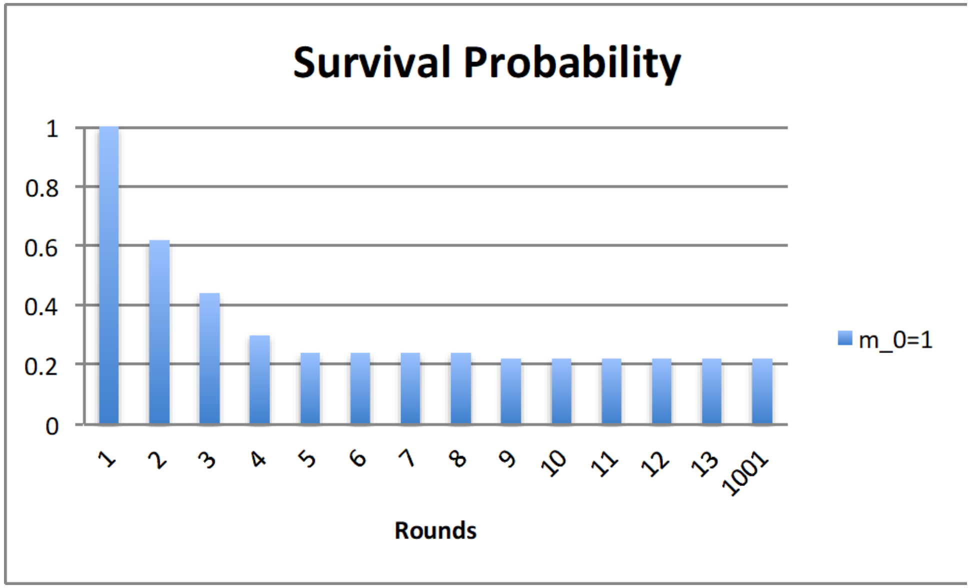
for $t = 1, 2, \ldots$, where $E_1[\ ]$ are the expectations taken in the first period.

- Long-run performance depends on initial conditions.
  - Differences grow in expectation over time.

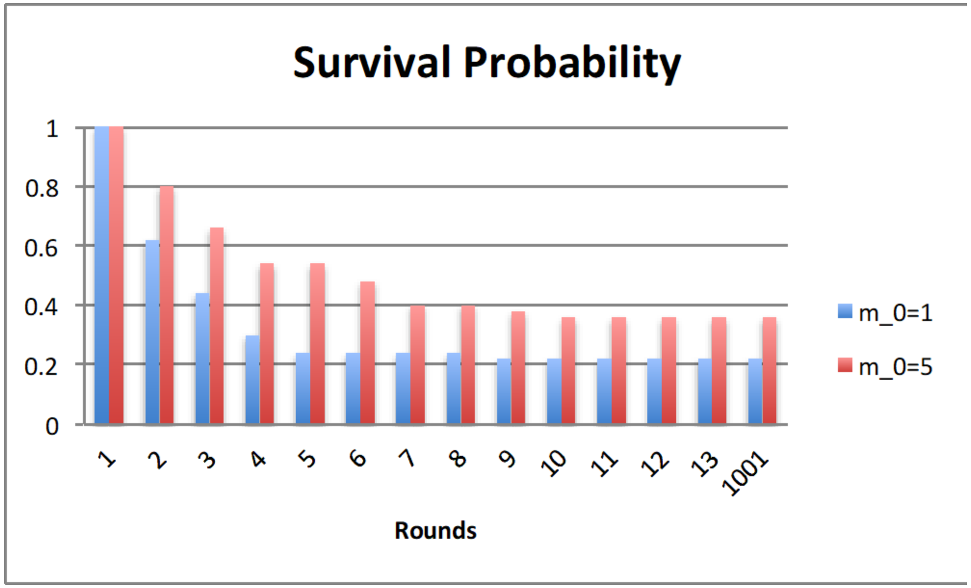
Numerical Example

- Suppose the utility function is given by: $u(m_t) = \alpha m_t - e^{-\beta m_t}$,

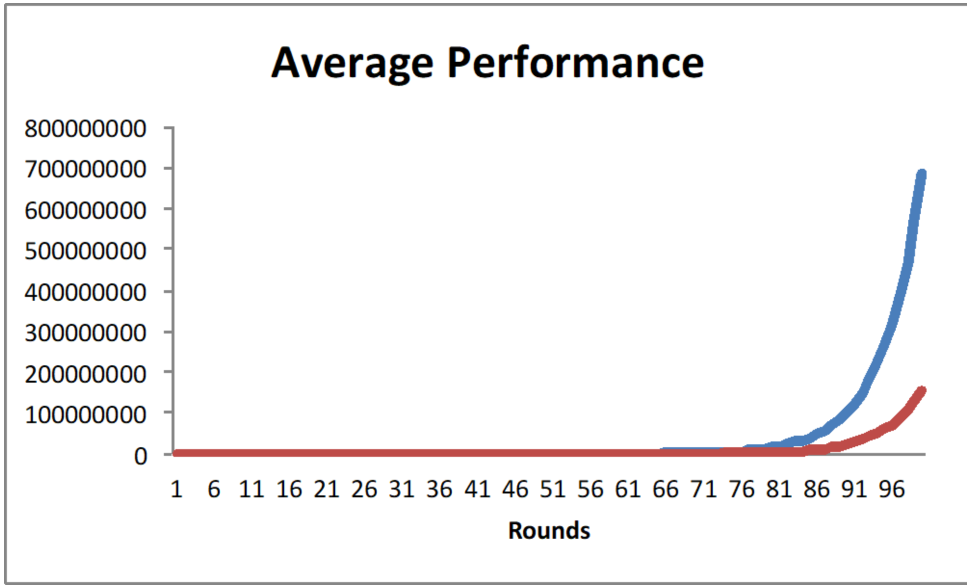
  where $\alpha > 0$ and $\beta > 2\mu/\sigma^2$.



Results from 50 runs with $\alpha = \sigma^2 = 1$, $\mu = ¼$, $\beta = 1.1$



Results from 50 runs with $\alpha = \sigma^2 = 1$, $\mu = ¼$, $\beta = 1.1$

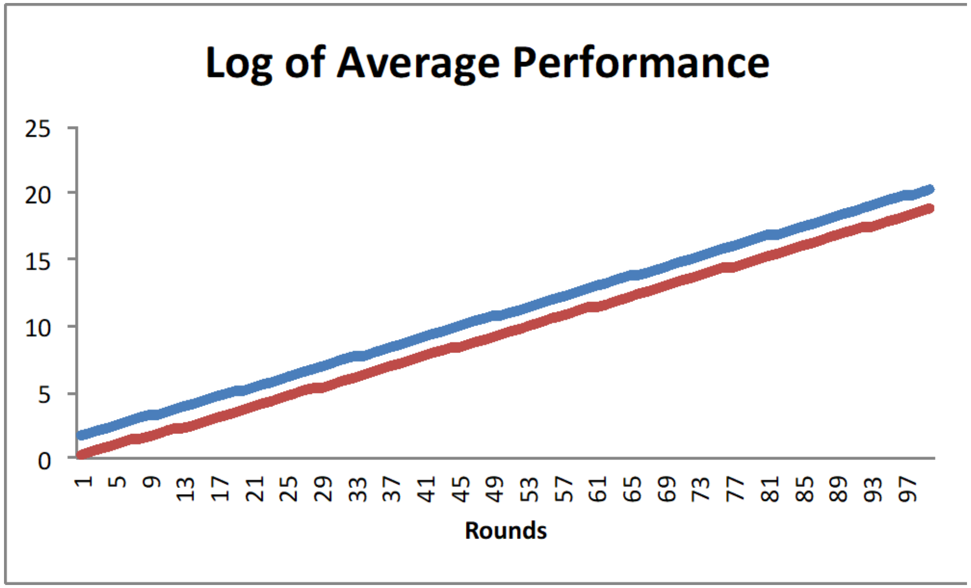


Results from 50 runs with $\alpha = \sigma^2 = 1$, $\mu = \frac{1}{4}$, $\beta = 1.1$

Gibrat's Law

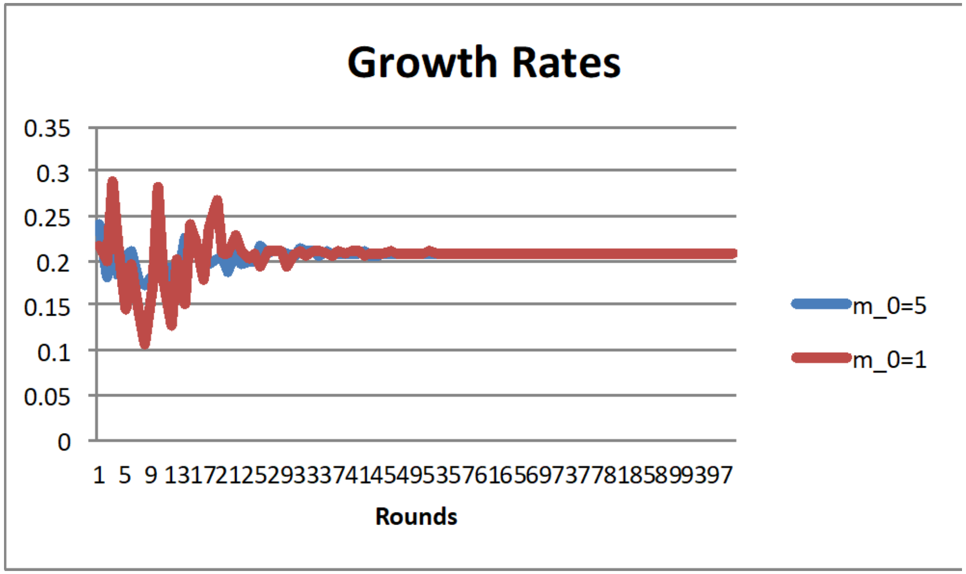
"The proportional rate of growth of a firm is independent of its absolute size."

- Gives rise to log-normal firm size distribution.
- Applied to city sizes, country economic development, …



Results from 50 runs with $\alpha = \sigma^2 = 1$, $\mu = ¼$, $\beta = 1.1$

Gibrat's Law



Results from 50 runs with $\alpha = \sigma^2 = 1$, $\mu = ¼$, $\beta = 1.1$

Summary

- Organizations either get bogged down in the first few periods ...

  or they escape the difficulties of their early histories.

- → Two-pronged growth:
  - Shake out phase: some get caught in poverty trap.
  - Growth phase: survivors grow indefinitely. Differences increase exponentially.

- Initial differences therefore grow exponentially over time.
  - Persistent Performance Differences among Seemingly Similar Enterprises (PPD of SSE).

- Progress is path dependent.

Far Sighted Agents – Some Progress

- This paper: Callander & Matouschek (2019, AEJ).
  - Search dynamics follow same structure.
  - Experimentation more valuable so higher cut-offs to stop.

- Garfagnini and Strulovici (2016, RES).
  - Two period horizon, zero drift.
  - Waves in search & experimentation—exploit vs. explore.

- Jovanovic and Rob (1990, Ecma).
  - Each dimension a Brownian motion. Choose only once on each.

- Callander and Hummel (2014, Ecma).
  - Two politicians take turns in office.
  - Strategic "preemptive experimentation".
  - Shape future policy by changing policy knowledge.

- Wong (2020, wp). Continuous search.
- Urgun and Yariv (2020, wp) ... Later today!

Applications

- Product innovation – Callander (2011, AER)
  - Phases of the Product Life Cycle: monotonic phase & triangulating phase.
- Policymaking.
  - Repeated elections among two policy-motivated parties – Callander (2011, APSR).
  - *"The country needs and, unless I mistake its temper, the country demands bold, persistent experimentation. It is common sense to take a method and try it; If it fails, admit it frankly and try another. But above all, try something."*
    Franklin D. Roosevelt, 1932.
  - Dynamic agenda setting – w/ McCarty (in progress).
- Organizations. w/ Matouschek (in progress).
  - Imitation across firms. One common BM + one idiosyncratic BM = performance.
  - Coordination within firms.
  - Pilot projects & experimentation.
- Multi-dimensional search – w/ Umberto Garfagnini (stagnating…).