Equilibrium Conditions for Catch-22 Situations

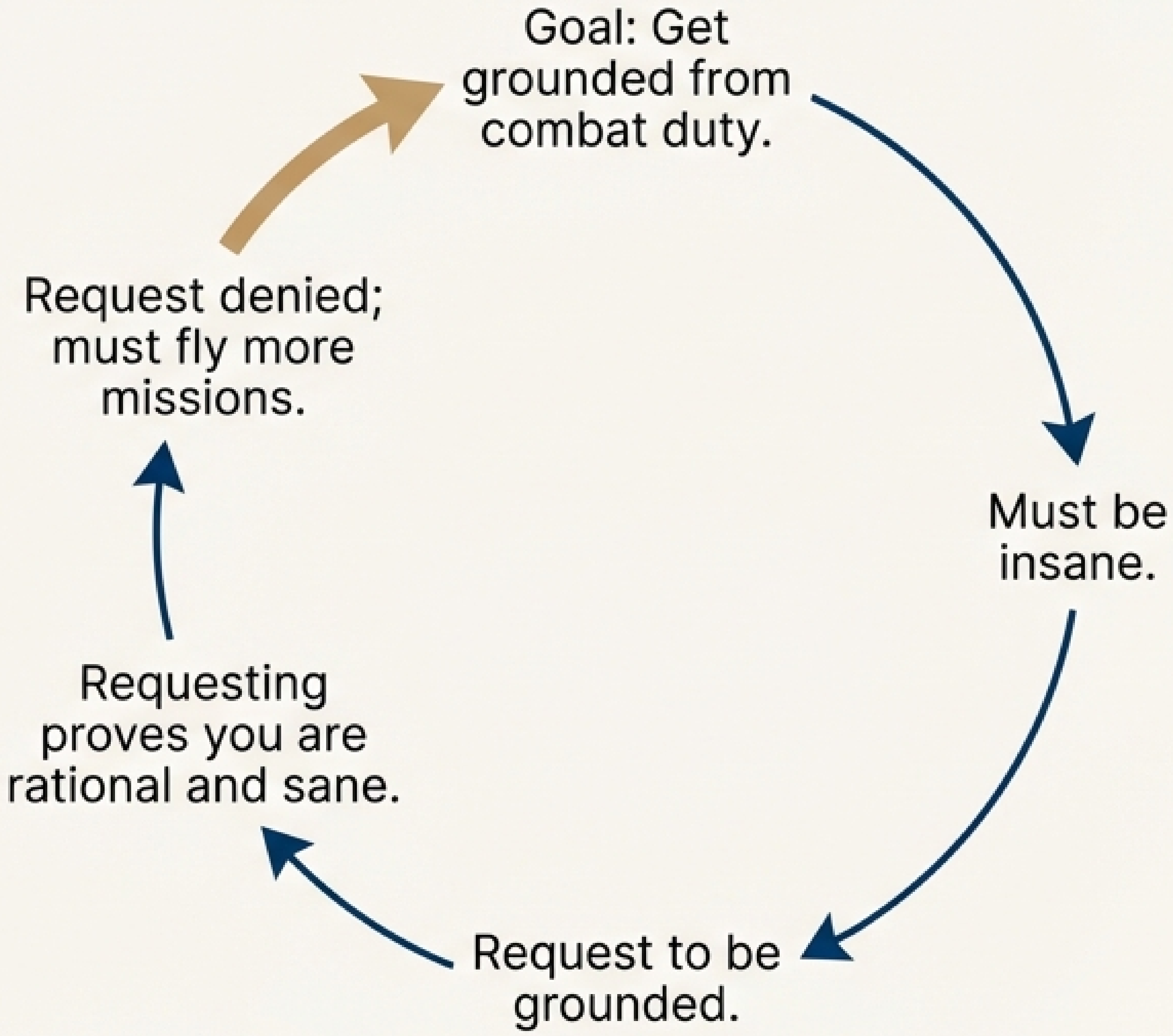
Joshua S. Gans
University of Toronto, Canada and NBER, USA
joshua.gans@utoronto.ca

DOI:10.22574/jmid.2025.12.002 | JEL Classification: C72, D82, J24, D86

The Canonical Paradox: Heller's Catch-22

"There was only one catch and that was Catch-22, which specified that a concern for one's own safety in the face of dangers that were real and immediate was the process of a rational mind. Orr was crazy and could be grounded. All he had to do was ask; and as soon as he did, he would no longer be crazy and would have to fly more missions."

— Joseph Heller, *Catch-22* (1961)



An Economist's Question: Where Do These Rules Come From?

Central Question

Are Catch-22 rules merely "an expression of power dressed up in a rationale," or can they emerge as the logical, preference-driven outcome of a **strategic game**?

Paper's Hypothesis

We can model the paradox not as an arbitrary rule, but as a potential equilibrium outcome. The analysis must consider the payoffs to both sides to understand the rule's function and whether it is achieved efficiently.

Scope

This paper tackles the Catch-22 situation as spelled out by Heller head-on, exploring its plausibility through a formal game-theoretic model.

Formalizing the Paradox: The Game Setup

Players

- **Requester**: Knows their own type.
- **Gatekeeper**: Does not know the Requester's type.

Requester Types

- **Sane (S)**: Occurs with probability (1-p).
- **Insane (I)**: Occurs with probability p.

Sequence of Play

Nature

Chooses Requester's type (S or I).

Requester

Sends a message: Request (R) or Not Request ($\neg R$).

Gatekeeper

Observes the message and chooses an action: Grant (G) or Deny (D).

Outcomes

- Grant (G) leads to Outcome A.
- Deny (D) leads to Outcome B.

The Heart of the Game: Conflicting Preferences

Core Insight: The players' goals are in direct opposition, defining the strategic conflict. "Sanity" and "Insanity" are defined purely by these preferences.

Player	Preferred Outcome	Least Preferred Outcome
Sane Requester	**A** (Grounded)	**B** (Fly)
Insane Requester	**B** (Fly)	**A** (Grounded)
Gatekeeper	Sane gets **B** (Fly), Insane gets **A** (Grounded)	Sane gets **A**, Insane gets **B**

Summary: The Gatekeeper wants to achieve the exact opposite outcome for each type of Requester.

The Gatekeeper's Calculus

1. Equilibrium Concept

We solve for a **Perfect Bayesian Equilibrium** (PBE), where strategies are **best responses** and beliefs are updated via Bayes' rule.

2. The Gatekeeper's Decision

The Gatekeeper chooses to **Grant** (G) if their expected payoff from granting is higher than from denying. This decision hinges on their belief about the Requester's type after observing a message.

3. The Critical Threshold, q

Intuition

- q represents the Gatekeeper's "tipping point."
- If the posterior probability of the Requester being **Insane**, $P(I|message)$, is greater than q, the Gatekeeper's best response is to Grant.

Formal Definition

The Gatekeeper grants if $P(I|message) > q$, where:

$$q \equiv \frac{-\Delta S}{\Delta I - \Delta S} = \frac{g_S(B) - g_S(A)}{(g_I(A) - g_I(B)) + (g_S(B) - g_S(A))}$$

- $\Delta I = g_I(A) - g_I(B) > 0$ (Payoff gain from correctly grounding an Insane type)
- $\Delta S = g_S(A) - g_S(B) < 0$ (Payoff loss from incorrectly grounding a Sane type)

Proposition 1: The Equilibrium Conditions

Key Result: The equilibrium outcome depends entirely on the relationship between the prior probability of insanity (p) and the Gatekeeper's decision threshold (q).

If $p < q$ (Insanity is rare):

- **Outcome:** The unique equilibrium is for the Gatekeeper to **Always Deny**.

This is the **Catch-22** Equilibrium.

If $p > q$ (Insanity is common):

- **Outcome:** The unique equilibrium is for the Gatekeeper to **Always Grant**.

If $p = q$ (The tipping point):

- **Outcome:** A continuum of equilibria exists where both players are indifferent.

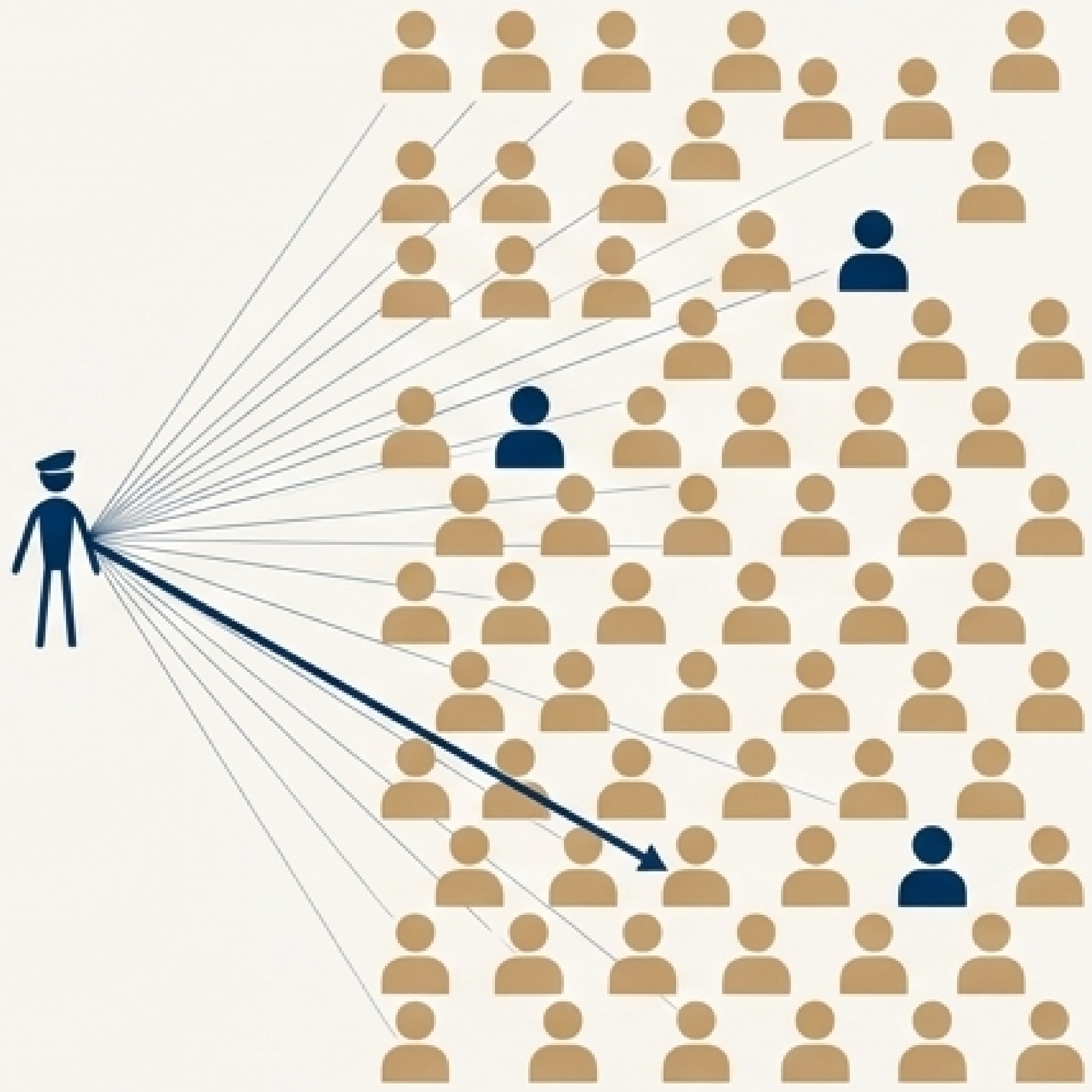
The Logic of the Catch: Why 'Always Deny'?

Focus on the $p < q$ case: When the probability of encountering an Insane type is low, why does the Gatekeeper ignore the Requester's message?

The 'Babbling Equilibrium': The Requester's message is 'cheap talk.' Since it costs nothing to send, it is not a credible signal of type.

Gatekeeper's Rationale: The Gatekeeper's optimal strategy is to ignore the uninformative message and act based on the prior probability p.

An Efficient Response: Because p is low, the Gatekeeper minimizes the expected error by assuming any given Requester is Sane. The "Always Deny" rule is an efficient response to population-level statistics, designed to avoid the costly mistake of grounding a Sane pilot. It is not an arbitrary rule, but a reflection of the Gatekeeper's preferences under uncertainty.



A Robust Finding

Testing an Alternative Timing: What if the Gatekeeper commits to a rule *before* the Requester sends their message?

1. **Gatekeeper**: Publicly commits to a policy $(\alpha_R, \alpha_{\neg R})$.
2. **Requester**: Observes the policy and chooses their message (R or $\neg R$) to maximize their payoff.

The equilibrium outcomes remain **identical**.

- If $p < q$, the Gatekeeper's optimal committed policy is to **Always Deny**.
- If $p > q$, the optimal policy is to **Always Grant**.

Conclusion: The emergence of the Catch-22 equilibrium is not an artifact of the game's timing. The finding is robust to whether the Gatekeeper acts reactively or pre-commits to a rule.

Application: The Labor Market "Experience Trap"

The Paradox

"You can't get a job without experience, but you can't get experience without a job."

Mapping the Model

Heller Model	Labor Market Analogy
Requester	Worker
Gatekeeper	Firm
Sane / Insane	Low-Potential / High-Potential
Grounded / Fly	Hired / Not Hired
Request	Job Application

Key Difference

Unlike the Heller model, this Catch-22 is typically a rule without an "inference component." We explore the conditions under which the *rule* emerges in equilibrium.

The Labor Market Model

Setup

- A fixed pool of N workers.
- A proportion p are High-Potential (H-type); $(1-p)$ are Low-Potential (L-type).
- All workers start as inexperienced. Experience is only gained through employment.

Productivity

- Firms face a downward-sloping labor demand.
- Marginal product of an experienced H-type is $H(E)$; for an L-type it is $L(E)$.
- $H'(E) < 0$ and $L'(E) < 0$ due to diminishing marginal returns.

The Hiring Decision

- A firm hiring an inexperienced worker pays a wage equal to their expected productivity:

$$w_{\text{inexp}}(E) = pH(E) + (1-p)L(E)$$

- Hiring occurs only if this expected contribution is positive.

Proposition 2: The Equilibrium Experience Trap

Initial Condition

Assume hiring is initially profitable, i.e., $pH(0) + (1-p)L(0) > 0$.

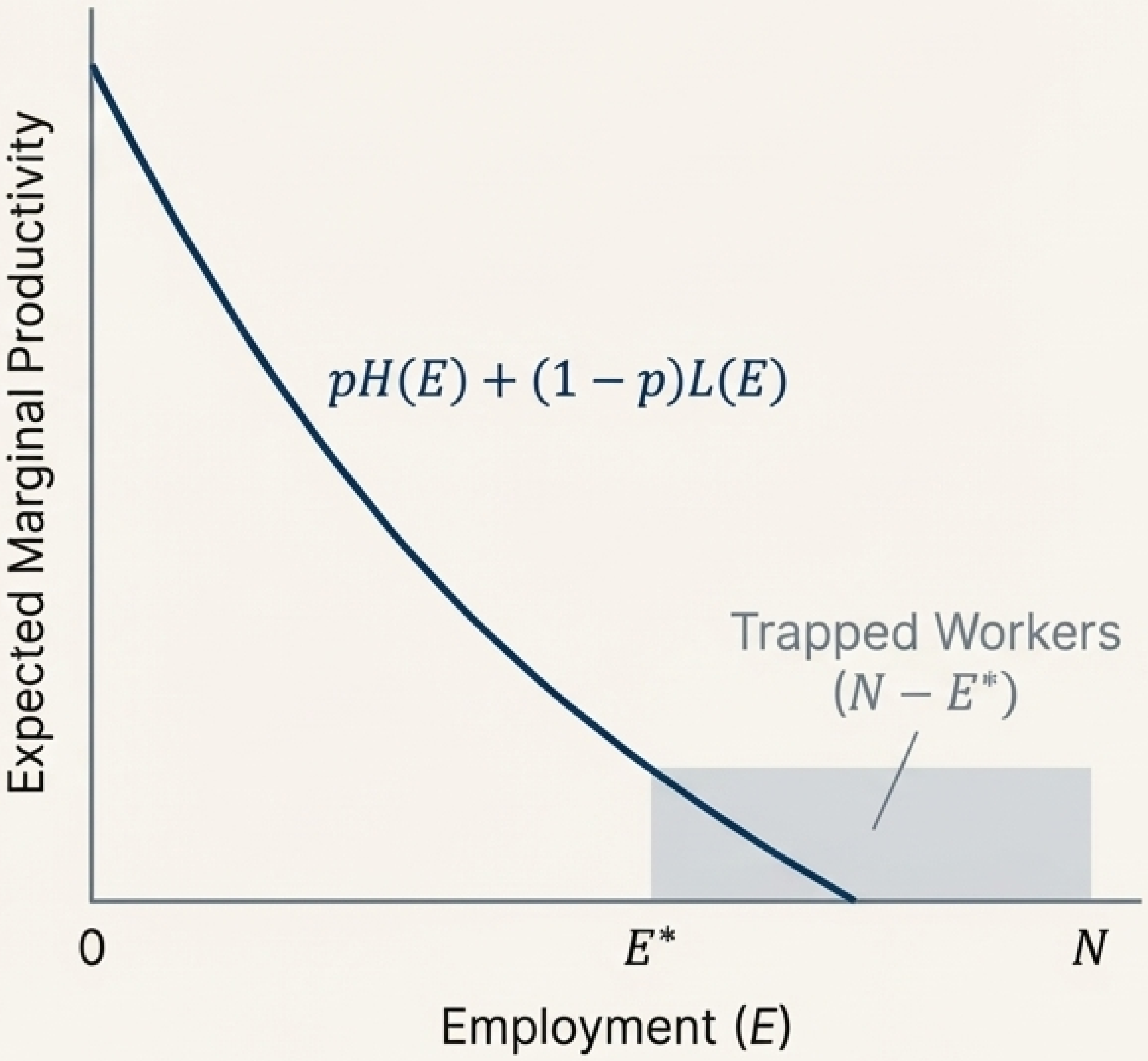
The Equilibrium

The firm hires workers until the employment level E^* is reached, where the expected marginal product of the next inexperienced hire is zero:

$$pH(E^*) + (1-p)L(E^*) = 0$$

Outcome: A Partial Catch-22

- E^* workers are hired and gain experience.
- The remaining $N - E^*$ workers are permanently unemployed and inexperienced. They are trapped: they cannot get experience without a job, and the firm will not hire them at zero expected marginal productivity.



The General Structure of a Catch-22 Situation

Common Elements

1. **Players**
    - **Seekers**: Desire a beneficial outcome they cannot achieve autonomously.
    - **Gatekeepers**: Control access to that outcome.

2. **Circular Prerequisite**
   The Gatekeeper requires the Seeker to possess an attribute that can only be acquired ***after*** obtaining the outcome they seek.

3. **Asymmetric Information**
   The Gatekeeper cannot directly observe the Seeker's true state or potential, leading to screening rules that inadvertently create the trap.

4. **Equilibrium Conditions**
   The paradox arises when parameters (probabilities, costs, marginal products) push the equilibrium to a state where there is no profitable deviation for any player that could break the cycle.

A Re-framing of the Paradox

Key Takeaway: Catch-22 situations are not necessarily products of irrationality, arbitrary power, or pure oppression.

The Economic Explanation: They can be understood as **stable, rational equilibria** that arise from strategic interactions under uncertainty and asymmetric information. The paradox has an internal logic.

> "The paradoxical state is stable precisely because it aligns with all players' equilibrium incentives and the structural constraints of the game."

From Theory to Intervention

- **Insight**

  Understanding the equilibrium logic behind a no-win scenario can inspire interventions. By identifying the key parameters that create the trap, we can propose changes to break the cycle.

Examples

Labor Market

Subsidies for hiring inexperienced workers or improved screening technologies could shift parameters ($p_H(E) + (1-p)L(E) > 0$) and eliminate the trap.

General

Interventions can relax prerequisites, provide credible signals, or alter payoffs to shift from a Catch-22 equilibrium to one where Seekers can achieve their desired outcomes.

Final Statement

By casting Catch-22s into a strategic and equilibrium-centric framework, we gain a deeper understanding of how such paradoxes arise and persist. **They are not mere narrative curiosities; they represent stable outcomes embedded in the logic of incentives, information, and constraints.**

Joshua S. Gans | joshua.gans@utoronto.ca