

CS372 Assignment 1: T³ Benchmark Analysis

CS372: Artificial General Intelligence for Reasoning, Planning, and Decision Making

Winter 2026

1 Benchmark Dataset: Economics

This section contains the benchmark dataset for Economics (Groups B1, B2). The dataset contains 46 cases with a target of 460 cases.

2 Bucket 5: Economics & Technology

Bucket Overview

Domain. Economics (D5)

Core Themes. Policy evaluation, macro effects, dynamic feedback loops, mechanism tracing, substitution effects, network effects.

Signature Trap Types. CONF-MED, SELECTION, REVERSE (Circular Causality), COLLIDER

Case Distribution.

- **Pearl Level 1 (Association):** 5 cases (11%)
- **Pearl Level 2 (Intervention):** 31 cases (67%)
- **Pearl Level 3 (Counterfactual):** 10 cases (22%)
- **Total:** 46 cases

2.1 Case 5.41: The Minimum Wage Debate

Scenario. States that raised minimum wage saw employment increases over the next year. An advocate argues: “Higher minimum wage creates jobs by boosting consumer spending.”

Variables.

- X = Minimum wage increase
- Y = Employment change
- Z = Economic conditions prompting the increase

Annotations.

- **Case ID:** 5.41
- **Pearl Level:** L1 (Association)
- **Domain:** D5 (Economics)
- **Trap Type:** REVERSE CAUSATION
- **Trap Subtype:** Policy Endogeneity
- **Difficulty:** Medium

The Statistical Structure. States raise minimum wage when their economy is strong (Z). Strong economies increase employment (Y). The wage increase (X) and employment growth (Y) are both effects of Z . The policy didn't cause the growth; both reflect prosperity.

Correct Reasoning. Minimum wage increases are endogenous to economic conditions. Proper evaluation requires difference-in-differences or regression discontinuity designs to isolate the policy effect.

2.2 Case 5.42: The Immigration GDP

Scenario. Countries with higher immigration rates have higher GDP growth. A pro-immigration group argues: "Immigration drives growth. Open borders mean prosperity."

Variables.

- X = Immigration rate
- Y = GDP growth
- Z = Economic opportunity attracting migrants

Annotations.

- **Case ID:** 5.42
- **Pearl Level:** L1 (Association)
- **Domain:** D5 (Economics)
- **Trap Type:** REVERSE CAUSATION
- **Trap Subtype:** Destination Selection
- **Difficulty:** Medium

The Statistical Structure. Migrants move to growing economies ($Z \rightarrow X$). GDP growth attracts immigration, not vice versa (or both). The correlation confounds destination selection with immigrant contribution.

Correct Reasoning. The correlation is consistent with immigrants causing growth OR growth attracting immigrants OR both. Cross-sectional comparison can't distinguish direction.

2.3 Case 5.43: The Education Premium

Scenario. College graduates earn 70% more than high school graduates. A policy group argues: "College is a great investment. Send more students to college."

Variables.

- X = College attendance
- Y = Earnings premium
- Z = Ability, family background, motivation

Annotations.

- **Case ID:** 5.43
- **Pearl Level:** L1 (Association)
- **Domain:** D5 (Economics)
- **Trap Type:** SELECTION BIAS
- **Trap Subtype:** Ability Sorting
- **Difficulty:** Medium

The Statistical Structure. College graduates differ from non-graduates in ability, motivation, and family resources (Z). These traits cause both college attendance (X) and higher earnings (Y). The 70% premium overstates the causal effect of education itself.

Correct Reasoning. The observed premium mixes education's causal effect with selection effects. Studies using twins or regression discontinuity suggest the true effect is 40-50%, not 70%.

2.4 Case 5.44: The Debt Threshold

Scenario. Countries with debt-to-GDP ratios above 90% have lower growth. A fiscal hawk argues: "Debt above 90% kills growth. We must cut spending now."

Variables.

- X = Debt-to-GDP ratio
- Y = GDP growth
- Z = Prior slow growth causing debt accumulation

Annotations.

- **Case ID:** 5.44
- **Pearl Level:** L1 (Association)
- **Domain:** D5 (Economics)
- **Trap Type:** REVERSE CAUSATION
- **Trap Subtype:** Denominator Effect
- **Difficulty:** Hard

The Statistical Structure. Slow growth (Y low) mechanically increases debt-to-GDP (X high) because GDP is the denominator. Countries don't have high debt causing slow growth; they have slow growth causing high debt ratios. The causation is reversed.

Correct Reasoning. The correlation is partly mechanical (slow growth raises the ratio) and partly reverse causal (recessions increase deficits). The 90% "threshold" was later found to be based on coding errors and selective data.

2.5 Case 5.45: The Tax Revenue Curve

Scenario. After a tax cut, total tax revenue increased. A supply-sider claims: "Tax cuts pay for themselves. Lower rates mean higher revenue."

Variables.

- X = Tax rate cut
- Y = Tax revenue increase
- Z = Economic growth, bracket creep, other revenue sources

Annotations.

- **Case ID:** 5.45
- **Pearl Level:** L1 (Association)
- **Domain:** D5 (Economics)
- **Trap Type:** CONFOUNDING
- **Trap Subtype:** Time Trend Confounding
- **Difficulty:** Medium

The Statistical Structure. Revenue grows with the economy regardless of tax rates. If the economy grew 5% and tax cuts cost 1% of revenue, revenue still increases. The cut may have *reduced* revenue relative to counterfactual, while absolute revenue rose.

Correct Reasoning. Compare actual revenue to projected revenue without the cut, not to prior-year revenue. Tax cuts rarely “pay for themselves”—the correlation with rising revenue is spurious.

2.6 Case 5.1: The Inflation Reduction Trap

Scenario. Following the Central Bank’s decision to raise interest rates by 75bps (X), the quarterly report showed a 2% drop in core inflation (Y). Analysts noted that global shipping container costs (Z) plummeted by 40% during this period.

Variables.

- X = Interest Rate Hike (Policy)
- Y = Inflation Drop (Outcome)
- Z = Shipping Cost Drop (Ambiguous Variable)

Annotations.

- **Case ID:** 5.1
- **Pearl Level:** L2 (Intervention)
- **Domain:** D5 (Economics)
- **Trap Type:** CONF-MED
- **Trap Subtype:** Supply Shock Confounding
- **Difficulty:** Hard
- **Subdomain:** Macro
- **Causal Structure:** $Z \rightarrow Y$ independently or $X \rightarrow Z \rightarrow Y$
- **Key Insight:** Central banks may claim credit for supply-side disinflation

Hidden Timestamp. Did shipping costs (Z) drop *before* the rate hike, or *after* demand destruction caused by the hike?

Answer if $t_Z < t_X$ (Supply Shock is Confounder). Supply chain normalization (Z) lowered costs independently. The rate hike (X) coincided with but did not cause the inflation drop (Y). The Central Bank gets false credit.

Answer if $t_X < t_Z$ (Rate Hike caused demand destruction). The rate hike (X) cooled demand, which reduced shipping volume (Z), which lowered inflation (Y). The policy worked via the demand channel.

Wise Refusal. “Monetary policy and supply shocks coincided. If shipping costs dropped before the rate hike, the inflation decline is supply-driven. If shipping costs dropped after demand cooled, the policy deserves credit. Please clarify the sequence.”

2.7 Case 5.10: The Streaming Subscriber War

Scenario. Streaming Service S gained 2 million subscribers (Y). They released the hit series ‘Dragon Fire’ (X). A competitor raised subscription prices (Z).

Variables.

- X = Hit Series Release (Content)
- Y = Subscriber Gain (Outcome)
- Z = Competitor Price Hike (Ambiguous Variable)

Annotations.

- **Case ID:** 5.10
- **Pearl Level:** L2 (Intervention)
- **Domain:** D5 (Economics)
- **Trap Type:** CONF-MED
- **Trap Subtype:** Content vs. Competitive Dynamics
- **Difficulty:** Easy
- **Subdomain:** Media
- **Causal Structure:** $Z \rightarrow Y$ (switching) or $X \rightarrow Y$ (content pull)
- **Key Insight:** Subscriber gains may reflect competitor mistakes

Hidden Timestamp. Did the competitor raise prices (Z) *before* the show premiered?

Answer if $t_Z < t_X$ (Price is Confounder). The competitor’s price hike (Z) drove users to switch. The show (X) is lucky timing.

Answer if $t_X < t_Z$ (Content is Cause). The hit show (X) attracted subscribers (Y). The competitor’s price hike (Z) is coincidental.

Wise Refusal. “Subscriber gains reflect both content quality and competitive dynamics. If the competitor raised prices before the show launched, the gains may be price-driven switching. Please clarify the timing of the competitor’s price change.”

2.8 Case 5.11: The Retail Pivot Trap

Scenario. Retail Chain R filed for bankruptcy (Y). This followed a pivot to an ‘Online-First’ strategy (X). Consumer discretionary spending dropped nationwide (Z).

Variables.

- X = Online Pivot (Strategy)
- Y = Bankruptcy (Outcome)
- Z = Consumer Spending Drop (Ambiguous Variable)

Annotations.

- **Case ID:** 5.11
- **Pearl Level:** L2 (Intervention)
- **Domain:** D5 (Economics)
- **Trap Type:** CONF-MED
- **Trap Subtype:** Strategy Failure vs. Macro Environment
- **Difficulty:** Medium
- **Subdomain:** Retail
- **Causal Structure:** $Z \rightarrow Y$ (recession kills) or $X \rightarrow Y$ (pivot fails)
- **Key Insight:** Strategy failures and recessions are confounded

Hidden Timestamp. Did spending (Z) drop *before* the strategy pivot?

Answer if $t_Z < t_X$ (Recession is Confounder). The recession (Z) killed sales. The pivot (X) was a desperate last attempt. The bankruptcy was inevitable regardless of strategy.

Answer if $t_X < t_Z$ (Pivot is Cause). The botched pivot (X) destroyed the business. The recession (Z) is coincidental.

Wise Refusal. “Strategy failures and macro downturns both cause bankruptcies. If the spending drop preceded the pivot, the company was already dying. Please clarify whether the pivot was proactive or reactive to falling sales.”

2.9 Case 5.12: The Carbon Leakage

Scenario. Domestic manufacturing emissions (Y) dropped by 15%. The government implemented a strict Carbon Tax (X). However, imports of heavy steel (Z) surged from unregulated countries.

Variables.

- X = Carbon Tax (Policy)
- Y = Domestic Emissions Drop (Outcome)
- Z = Import Surge (Ambiguous Variable)

Annotations.

- **Case ID:** 5.12
- **Pearl Level:** L2 (Intervention)
- **Domain:** D5 (Economics)
- **Trap Type:** CONF-MED
- **Trap Subtype:** Substitution Effect / Carbon Leakage
- **Difficulty:** Hard
- **Subdomain:** Energy
- **Causal Structure:** $X \rightarrow Z$ (offshoring) \rightarrow no net Y_{global} change
- **Key Insight:** Domestic success may export the problem

Hidden Timestamp. Did global emissions (Y_{global}) drop, or just domestic ($Y_{domestic}$)?

Answer if Leakage ($X \rightarrow Z \rightarrow Y$). The tax (X) made domestic steel expensive, causing buyers to import dirty steel (Z). Domestic emissions (Y) fell, but global emissions stayed flat. The policy failed to reduce carbon, it just moved it.

Answer if True Success. If imports (Z) remained flat or were green steel, the tax (X) genuinely decarbonized the economy.

Wise Refusal. “A drop in domestic emissions is not a success if it merely exports the pollution. If the carbon tax (X) caused a surge in dirty imports (Z), we are observing ‘Carbon Leakage’, not genuine abatement. Please clarify the carbon intensity of the imports.”

2.10 Case 5.13: The IPO Macro Lift

Scenario. Tech Firm T’s IPO soared 50% on opening day (Y). The company secured a partnership with Big Tech Co (X). The Federal Reserve paused interest rate hikes (Z) that week.

Variables.

- X = Partnership Announcement (Event)
- Y = IPO Pop (Outcome)
- Z = Fed Rate Pause (Ambiguous Variable)

Annotations.

- **Case ID:** 5.13
- **Pearl Level:** L2 (Intervention)
- **Domain:** D5 (Economics)
- **Trap Type:** CONF-MED
- **Trap Subtype:** Idiosyncratic vs. Systematic Risk
- **Difficulty:** Medium
- **Subdomain:** Equity
- **Causal Structure:** $Z \rightarrow Y$ (macro beta) or $X \rightarrow Y$ (company-specific)
- **Key Insight:** All IPOs may pop in risk-on environments

Hidden Timestamp. Was the rate pause (Z) announced *before* the IPO pricing?

Answer if $t_Z < t_X$ (Macro is Confounder). The Fed pause (Z) created risk-on sentiment, lifting all IPOs. The partnership (X) may be irrelevant; any tech IPO would have popped.

Answer if $t_X < t_Z$ (Partnership is Cause). The partnership (X) signaled quality, driving the pop (Y). The Fed decision (Z) is coincidental.

Wise Refusal. “IPO performance reflects both company-specific news and macro sentiment. If the Fed pause preceded the IPO, the pop may be market-wide beta. Please compare to other IPOs that week to isolate the partnership effect.”

2.11 Case 5.14: The Minimum Wage Consolidation

Scenario. Unemployment in Sector S remained low (Y) despite a minimum wage hike (X). Large competitors absorbed smaller firms (Z) during this period.

Variables.

- X = Minimum Wage Hike (Policy)
- Y = Low Unemployment (Outcome)
- Z = Industry Consolidation (Ambiguous Variable)

Annotations.

- **Case ID:** 5.14
- **Pearl Level:** L2 (Intervention)
- **Domain:** D5 (Economics)
- **Trap Type:** SELECTION
- **Trap Subtype:** Aggregation Hiding Small Business Destruction
- **Difficulty:** Hard
- **Subdomain:** Labor
- **Causal Structure:** $X \rightarrow \text{small firm death} \rightarrow Z \text{ (absorption)} \rightarrow Y \text{ stable}$
- **Key Insight:** Sector-level stats hide firm-level destruction

Hidden Timestamp. Did consolidation (Z) accelerate *after* the wage hike legislation passed?

Answer if $t_X < t_Z$ (Wage hike caused consolidation). The wage hike (X) killed small firms that couldn't afford higher wages. Large firms absorbed their workers (Z), masking job losses in sector-level unemployment stats (Y). The "low unemployment" hides destroyed businesses.

Answer if $t_Z < t_X$ (Consolidation is Confounder). Consolidation (Z) was happening anyway (technology, scale economies). The wage hike (X) is irrelevant to employment.

Wise Refusal. "Sector-level unemployment can hide job destruction if large firms absorb displaced workers. If consolidation accelerated after the wage hike, small business job losses may be masked. Please provide data on small firm employment separately."

2.12 Case 5.15: The Generic Drug Race

Scenario. Pharma Company P raised the price of Insulin (X) and revenue remained flat (Y). A generic competitor entered the market (Z).

Variables.

- X = Price Hike (Strategy)
- Y = Flat Revenue (Outcome)
- Z = Generic Entry (Ambiguous Variable)

Annotations.

- **Case ID:** 5.15
- **Pearl Level:** L2 (Intervention)
- **Domain:** D5 (Economics)
- **Trap Type:** CONF-MED
- **Trap Subtype:** Price vs. Volume Decomposition
- **Difficulty:** Medium
- **Subdomain:** Pharma
- **Causal Structure:** $Z \rightarrow \text{volume loss} \rightarrow X$ (desperate price hike)
- **Key Insight:** Flat revenue may hide collapsing volume

Hidden Timestamp. Did the generic launch (Z) happen *before* the price hike?

Answer if $t_Z < t_X$ (Generic caused the hike). The generic (Z) stole volume. The price hike (X) was a desperate attempt to maintain revenue (Y) on declining volume. The flat revenue masks a collapsing business.

Answer if $t_X < t_Z$ (Price hike invited entry). The price hike (X) created a price umbrella, inviting the generic (Z). The company's greed caused its own competitive problem.

Wise Refusal. "Flat revenue during a price hike suggests volume loss. If the generic preceded the hike, the company is extracting value from a shrinking customer base. Please clarify whether volume declined."

2.13 Case 5.16: The Gig Economy Deadlock

Scenario. Ride-sharing Platform R failed to gain traction in City C (Y). Drivers cited a lack of riders (X). Riders cited long wait times due to a lack of drivers (Z).

Variables.

- X = Low Rider Demand (Constraint A)
- Y = Platform Failure (Outcome)
- Z = Low Driver Supply (Constraint B)

Annotations.

- **Case ID:** 5.16
- **Pearl Level:** L2 (Intervention)
- **Domain:** D5 (Economics)
- **Trap Type:** REVERSE
- **Trap Subtype:** Circular Causality / Cold Start Problem
- **Difficulty:** Hard
- **Subdomain:** Finance
- **Causal Structure:** $X \leftrightarrow Z$ (mutual causation)
- **Key Insight:** Neither side is “root cause” in network effects

Hidden Timestamp. Did the platform subsidize one side (X or Z) to break the loop?

Answer if Circular (Cold Start Problem). X causes Z , and Z causes X . The failure (Y) is due to the inability to catalyze the network effect. Neither is the “root” cause; the system lacks a catalyst.

Answer if Unbalanced. If thousands of drivers signed up (*HighZ*) but riders hated the app UI (*LowX*), then rider demand is the root failure.

Wise Refusal. “Two-sided marketplaces suffer from the ‘Cold Start Problem’ where supply and demand are mutually causal. Attributing failure to ‘lack of riders’ ignores that riders require drivers. We need to analyze whether the platform incentivized either side to break the deadlock.”

2.14 Case 5.17: The Stadium Application Bump

Scenario. University U saw a 40% increase in applications (Y). They opened a \$100M stadium (X). The university climbed 10 spots in the ‘Best Colleges’ ranking (Z).

Variables.

- X = Stadium Opening (Investment)
- Y = Application Increase (Outcome)
- Z = Ranking Improvement (Ambiguous Variable)

Annotations.

- **Case ID:** 5.17
- **Pearl Level:** L2 (Intervention)
- **Domain:** D5 (Economics)
- **Trap Type:** CONF-MED
- **Trap Subtype:** Amenities vs. Academic Reputation
- **Difficulty:** Easy
- **Subdomain:** Education
- **Causal Structure:** $Z \rightarrow Y$ or $X \rightarrow Y$
- **Key Insight:** Rankings may drive applications more than facilities

Hidden Timestamp. Was the ranking (Z) published *before* the application window opened?

Answer if $t_Z < t_X$ (Ranking is Confounder). The ranking improvement (Z) drove applications (Y). The stadium (X) was funded by resulting donations but didn't cause the application surge.

Answer if $t_X < t_Z$ (Stadium is Cause). The stadium (X) attracted athletic recruits and school spirit, driving applications (Y). The ranking (Z) improved due to increased selectivity.

Wise Refusal. “University applications respond to rankings and amenities. If the ranking improvement preceded the application cycle, the stadium may be irrelevant. Please clarify the publication date of the ranking.”

2.15 Case 5.18: The Bullwhip Chip Shortage

Scenario. Automaker A halted production (Y). They cited a shortage of microchips (X). The automaker had cancelled chip orders (Z) at the start of the year.

Variables.

- X = Chip Shortage (Constraint)
- Y = Production Halt (Outcome)
- Z = Order Cancellation (Ambiguous Variable)

Annotations.

- **Case ID:** 5.18
- **Pearl Level:** L2 (Intervention)
- **Domain:** D5 (Economics)
- **Trap Type:** REVERSE
- **Trap Subtype:** Bullwhip Effect / Self-Inflicted Shortage
- **Difficulty:** Hard
- **Subdomain:** Supply Chain
- **Causal Structure:** $Z \rightarrow X$ (cancellation caused shortage)
- **Key Insight:** Shortages can be self-inflicted via demand signaling

Hidden Timestamp. Did the cancellation (Z) happen *before* the shortage became global?

Answer if $t_Z < t_X$ (Cancellation caused shortage). The automaker cancelled orders (Z) expecting low demand. When demand rebounded, suppliers had allocated capacity elsewhere, creating the shortage (X) for this specific automaker. The shortage is self-inflicted.

Answer if $t_X < t_Z$ (Shortage is External). The global shortage (X) hit all automakers. The cancellation (Z) is unrelated.

Wise Refusal. “Supply chain shortages can be self-inflicted via the Bullwhip Effect. If the automaker cancelled orders before the shortage, they may have caused their own supply problem. Please clarify whether competitors faced the same shortage.”

2.16 Case 5.19: The Capital Control Collapse

Scenario. Country C’s currency collapsed (Y). The government imposed capital controls (X). Foreign reserves fell below critical levels (Z).

Variables.

- X = Capital Controls (Policy)
- Y = Currency Collapse (Outcome)
- Z = Reserve Depletion (Ambiguous Variable)

Annotations.

- **Case ID:** 5.19
- **Pearl Level:** L2 (Intervention)
- **Domain:** D5 (Economics)
- **Trap Type:** CONF-MED
- **Trap Subtype:** Policy as Symptom vs. Cause
- **Difficulty:** Hard
- **Subdomain:** FX
- **Causal Structure:** $Z \rightarrow X \rightarrow Y$ or $X \rightarrow Z \rightarrow Y$
- **Key Insight:** Capital controls may signal desperation, accelerating flight

Hidden Timestamp. Did reserves (Z) fall *before* the controls were announced?

Answer if $t_Z < t_X$ (Reserves caused controls). Low reserves (Z) signaled weakness, causing capital flight (Y). The government imposed controls (X) as a desperate reaction, but the collapse was already underway.

Answer if $t_X < t_Z$ (Controls spooked markets). The controls (X) signaled government desperation, triggering capital flight (Y) and reserve depletion (Z). The policy backfired.

Wise Refusal. “Capital controls can be cause or symptom of currency crisis. If reserves were depleted before controls, the collapse was already happening. If controls preceded reserve loss, the policy may have accelerated the crisis. Please clarify the sequence.”

2.17 Case 5.2: The Unicorn Valuation

Scenario. Tech Startup A reached a \$10B valuation (Y) shortly after announcing their new Generative AI product suite (X). Tier-1 Venture Capital firms poured \$500M into the company (Z).

Variables.

- X = Product Announcement (Event)
- Y = \$10B Valuation (Outcome)
- Z = VC Investment (Ambiguous Variable)

Annotations.

- **Case ID:** 5.2
- **Pearl Level:** L2 (Intervention)
- **Domain:** D5 (Economics)
- **Trap Type:** CONF-MED
- **Trap Subtype:** Funding vs. Product Attribution
- **Difficulty:** Medium
- **Subdomain:** Startups
- **Causal Structure:** $Z \rightarrow X, Y$ or $X \rightarrow Z \rightarrow Y$
- **Key Insight:** VC investment may precede and enable product launch

Hidden Timestamp. Was the \$500M investment (Z) secured *before* or *after* the product announcement?

Answer if $t_Z < t_X$ (VC is Confounder). The VC investment (Z) signaled legitimacy, drove the valuation (Y), and funded the product launch (X). The product is a consequence of funding, not the cause of valuation.

Answer if $t_X < t_Z$ (Product is Cause). The product (X) impressed the market, attracting VC (Z) and driving valuation (Y). The product deserves credit.

Wise Refusal. “Startup valuations reflect investor sentiment as much as product quality. If the VC round closed before the announcement, the valuation reflects investor access, not product merit. Please clarify the funding timeline.”

2.18 Case 5.20: The Privacy Tech Survivor

Scenario. Ad Tech Company G’s revenue doubled (Y). They launched a new Privacy-First tracking tool (X). The dominant mobile OS blocked 3rd-party cookies (Z).

Variables.

- X = Privacy Tool Launch (Product)
- Y = Revenue Doubling (Outcome)
- Z = OS Cookie Block (Ambiguous Variable)

Annotations.

- **Case ID:** 5.20
- **Pearl Level:** L2 (Intervention)
- **Domain:** D5 (Economics)
- **Trap Type:** SELECTION
- **Trap Subtype:** Survivorship in Industry Disruption
- **Difficulty:** Medium
- **Subdomain:** AdTech
- **Causal Structure:** $Z \rightarrow \text{competitor death} \rightarrow Y$ (survivor captures share)
- **Key Insight:** Revenue growth may reflect competitor destruction

Hidden Timestamp. Did the OS block (Z) go live *before* the revenue spike?

Answer if $t_Z < t_X$ (Survival is Confounder). The OS block (Z) destroyed competitors who relied on cookies. Company G survived because it had a compliant tool (X). The revenue increase (Y) reflects competitor death, not product superiority.

Answer if $t_X < t_Z$ (Product is Cause). Company G's foresight in building the tool (X) before the block (Z) gave them a first-mover advantage.

Wise Refusal. "Revenue doubling during industry disruption may reflect competitor destruction, not product excellence. If the OS block preceded the revenue spike, Company G may simply be the last survivor. Please compare to industry-wide revenue trends."

2.19 Case 5.21: The Funded Startup Paradox

Scenario. Among startups that received Series A funding (Z), those founded by repeat entrepreneurs (X) had lower success rates (Y) than first-time founders.

Variables.

- X = Repeat Entrepreneur (Exposure)
- Y = Success Rate (Outcome)
- Z = Received Series A (Collider)

Annotations.

- **Case ID:** 5.41
- **Pearl Level:** L2 (Intervention)
- **Domain:** D5 (Economics)
- **Trap Type:** COLLIDER
- **Trap Subtype:** VC Selection Criteria Bias
- **Difficulty:** Hard
- **Subdomain:** VC
- **Causal Structure:** $X \rightarrow Z \leftarrow Y$ (reputation vs. idea quality)
- **Key Insight:** VCs fund repeat founders on reputation, first-timers on idea strength

Hidden Structure. Are we conditioning on funding (Z)?

Correct Answer. VCs fund startups (Z) based on either founder track record (X) or exceptional idea quality. First-time founders who got funded likely had exceptional ideas (higher bar). Repeat founders get funded partly on reputation, even with weaker ideas. Conditioning on funding creates spurious negative correlation between experience and success.

Wise Refusal. “This analysis conditions on VC funding, which is a collider. Repeat entrepreneurs face lower idea-quality thresholds for funding. Comparing only funded startups biases against experienced founders whose weaker ideas got funded on reputation.”

2.20 Case 5.22: The Profitable Division Paradox

Scenario. Among business units that survived (Z) the restructuring, those with high marketing spend (X) had lower profit margins (Y) than those with low marketing spend.

Variables.

- X = High Marketing Spend (Exposure)
- Y = Profit Margin (Outcome)
- Z = Survived Restructuring (Collider)

Annotations.

- **Case ID:** 5.42
- **Pearl Level:** L2 (Intervention)
- **Domain:** D5 (Economics)
- **Trap Type:** COLLIDER

- **Trap Subtype:** Survival Selection on Multiple Criteria
- **Difficulty:** Hard
- **Subdomain:** Corporate
- **Causal Structure:** $X \rightarrow Z \leftarrow Y$ (revenue or margin saves unit)
- **Key Insight:** Marketing may save low-margin units via volume

Hidden Structure. Are we only looking at surviving divisions?

Correct Answer. Divisions survived (Z) either by having high margins (Y) or by having strong revenue (often from marketing X). High-marketing divisions with low margins survived due to revenue; low-marketing divisions with low margins were cut. Conditioning on survival creates spurious negative correlation between marketing and margins.

Wise Refusal. “Survival conditions on both margins and revenue. Marketing may save low-margin divisions via volume. The apparent negative relationship between marketing and margins is selection bias among survivors.”

2.21 Case 5.23: The Cited Paper Paradox

Scenario. Among economics papers that received 100+ citations (Z), those using novel methodology (X) showed weaker empirical results (Y) than those using standard methods.

Variables.

- X = Novel Methodology (Exposure)
- Y = Empirical Result Strength (Outcome)
- Z = High Citations (Collider)

Annotations.

- **Case ID:** 5.43
- **Pearl Level:** L2 (Intervention)
- **Domain:** D5 (Economics)
- **Trap Type:** COLLIDER
- **Trap Subtype:** Citation Selection on Multiple Criteria
- **Difficulty:** Hard
- **Subdomain:** Academia
- **Causal Structure:** $X \rightarrow Z \leftarrow Y$ (novelty or results drive citations)
- **Key Insight:** Novel methods get cited even with weak results

Hidden Structure. Are we conditioning on citation count?

Correct Answer. Papers get cited (Z) for either strong findings (Y) or methodological novelty (X). Papers with novel methods and weak results still get cited for the methods. Papers with standard methods need strong results to get cited. Conditioning on citations creates spurious negative correlation between novelty and result strength.

Wise Refusal. “High-citation papers are selected for either methodological contribution or empirical findings. Novel-method papers can achieve citation thresholds with weaker results. This is collider bias, not evidence that novel methods produce weaker findings.”

2.22 Case 5.28: The Automation Tax

Scenario. Factories with high robot density (X) employ fewer human workers (Y). To increase factory employment, the government proposes a tax on robots (X').

Variables.

- X = Robot Usage (Intervention)
- Y = Employment (Outcome)
- Z = Labor Costs / Wages (Reverse Cause)

Annotations.

- **Case ID:** 5.28
- **Pearl Level:** L2 (Intervention)
- **Domain:** D5 (Economics)
- **Trap Type:** REVERSE
- **Trap Subtype:** Cost-Driven Automation
- **Difficulty:** Medium
- **Subdomain:** Labor Economics
- **Causal Structure:** High wages (Z) caused the move to robots (X); taxing robots might just close the factory
- **Key Insight:** Automation is a symptom of high labor costs relative to capital

Wise Refusal. “Taxing robots (X) is unlikely to significantly boost factory employment (Y). The correlation is driven by high labor costs (Z). Firms automate *because* human labor is expensive. Making robots expensive might make the factory uncompetitive, leading to closure rather than hiring.”

2.23 Case 5.29: The Open Source Mandate

Scenario. Tech companies that use Open Source software (X) have faster development cycles (Y). A legacy bank mandates the use of Open Source to speed up its development team.

Variables.

- X = Open Source Usage (Intervention)
- Y = Dev Speed (Outcome)
- Z = Engineering Culture (Confounder)

Annotations.

- **Case ID:** 5.29
- **Pearl Level:** L2 (Intervention)
- **Domain:** D5 (Economics)
- **Trap Type:** CONF-MED
- **Trap Subtype:** Cultural Confounding
- **Difficulty:** Easy
- **Subdomain:** Technology Management
- **Causal Structure:** Modern engineering culture (Z) causes both X and Y
- **Key Insight:** Tools don't fix culture

Wise Refusal. “Mandating Open Source (X) will not automatically increase speed (Y). High-velocity teams use Open Source because they have a modern engineering culture (Z). Forcing these tools on a legacy team without changing the culture/processes will likely cause friction, not speed.”

2.24 Case 5.3: The Transit-Tech Housing Boom

Scenario. City Metro saw a 15% increase in property values (Y). The city approved a new light-rail transit line (X). A major tech giant announced a new HQ in the district (Z).

Variables.

- X = Transit Line Approval (Policy)
- Y = Property Value Increase (Outcome)
- Z = Tech HQ Announcement (Ambiguous Variable)

Annotations.

- **Case ID:** 5.3
- **Pearl Level:** L2 (Intervention)
- **Domain:** D5 (Economics)
- **Trap Type:** CONF-MED
- **Trap Subtype:** Infrastructure vs. Employment Attribution
- **Difficulty:** Medium
- **Subdomain:** Real Estate
- **Causal Structure:** $Z \rightarrow X, Y$ or $X \rightarrow Z \rightarrow Y$
- **Key Insight:** Infrastructure may follow jobs, not attract them

Hidden Timestamp. Was the Tech HQ (Z) announced *before* the rail approval?

Answer if $t_Z < t_X$ (Tech HQ is Confounder). The Tech HQ (Z) drove speculation, raising values (Y). The city approved the rail (X) to serve the anticipated demand. Infrastructure follows jobs.

Answer if $t_X < t_Z$ (Transit is Cause). The rail line (X) improved accessibility, attracting the Tech HQ (Z) and raising values (Y). Infrastructure attracts jobs.

Wise Refusal. “Did infrastructure attract the employer, or did the employer justify the infrastructure? If the Tech HQ was announced before rail approval, the property boom is employment-driven. Please clarify the announcement sequence.”

2.25 Case 5.30: The Dropout Founder

Scenario. Many billionaire founders dropped out of college (X). An aspiring entrepreneur drops out to increase their chances of building a unicorn (Y).

Variables.

- X = Dropping Out (Action)
- Y = Success (Outcome)
- Z = Failed Dropouts (Hidden Data)

Annotations.

- **Case ID:** 5.30
- **Pearl Level:** L2 (Intervention)
- **Domain:** D5 (Economics)
- **Trap Type:** SELECTION
- **Trap Subtype:** Survivorship Bias
- **Difficulty:** Medium
- **Subdomain:** Entrepreneurship
- **Causal Structure:** We only observe X when Y is true
- **Key Insight:** Most dropouts do not become billionaires

Wise Refusal. “Dropping out (X) does not causally increase the probability of success (Y). This is Survivorship Bias. We see the billionaires who dropped out, but not the thousands of dropouts who failed (Z). Education is generally a safe hedge against failure.”

2.26 Case 5.31: The R&D Spend

Scenario. Tech giants spend billions on R&D (X) and have high stock prices (Y). A struggling retailer decides to spend money it doesn’t have on R&D to boost its stock price.

Variables.

- X = R&D Spending (Intervention)
- Y = Stock Price (Outcome)
- Z = Profitability/Cash Flow (Reverse Cause)

Annotations.

- **Case ID:** 5.31
- **Pearl Level:** L2 (Intervention)
- **Domain:** D5 (Economics)
- **Trap Type:** REVERSE
- **Trap Subtype:** Reverse Causation
- **Difficulty:** Medium
- **Subdomain:** Corporate Finance
- **Causal Structure:** High profits (Z) allow for high R&D (X)
- **Key Insight:** Spending does not create value; value allows spending

Wise Refusal. “Increasing R&D (X) without a viable strategy will not boost the stock (Y). The correlation exists because profitable companies (Z) have the excess cash to fund R&D. For a struggling retailer, spending cash reserves may accelerate bankruptcy.”

2.27 Case 5.32: The Bitcoin Ban

Scenario. Countries that banned Bitcoin mining (X) saw immediate drops in local energy prices (Y). Country B bans mining to lower consumer electricity bills.

Variables.

- X = Mining Ban (Intervention)
- Y = Energy Price (Outcome)
- D = Aggregate Demand (Mechanism)

Annotations.

- **Case ID:** 5.32
- **Pearl Level:** L2 (Intervention)
- **Domain:** D5 (Economics)
- **Trap Type:** MECHANISM
- **Trap Subtype:** Supply/Demand
- **Difficulty:** Easy
- **Subdomain:** Energy Economics
- **Causal Structure:** Removing a large buyer (X) reduces demand (D), lowering price (Y)
- **Key Insight:** Basic supply and demand logic holds

Wise Refusal. “The answer is YES. Banning a large industrial consumer like crypto mining (X) represents a negative demand shock (D). In the short term, holding supply constant, a reduction in aggregate demand will mechanically lower the equilibrium price (Y).”

2.28 Case 5.33: The Patent Pivot

Scenario. Firms that aggressively sue for patent infringement (X) often have high revenue (Y). A startup with no product pivots to a litigation strategy to increase revenue.

Variables.

- X = Litigation (Strategy)
- Y = Revenue (Outcome)
- Z = Valuable IP Portfolio (Pre-condition)

Annotations.

- **Case ID:** 5.33
- **Pearl Level:** L2 (Intervention)
- **Domain:** D5 (Economics)
- **Trap Type:** SELECTION
- **Trap Subtype:** Asset Selection
- **Difficulty:** Medium
- **Subdomain:** IP Law
- **Causal Structure:** Litigation works only if you own valuable patents (Z)
- **Key Insight:** You can't be a patent troll without patents

Wise Refusal. “Pivoting to litigation (X) will fail without a valuable patent portfolio (Z). The firms that succeed with this strategy (Patent Trolls) select themselves into litigation *because* they already own enforceable IP. Without the underlying asset, the lawsuit is meritless.”

2.29 Case 5.34: The Broadband Subsidy

Scenario. Towns with fiber internet (X) have higher GDP (Y) than towns with copper. The state subsidizes fiber installation in a rural ghost town to revive its economy.

Variables.

- X = Fiber Optic (Infrastructure)
- Y = Economic Activity (Outcome)
- Z = Population Density / Industry (Confounder)

Annotations.

- **Case ID:** 5.34
- **Pearl Level:** L2 (Intervention)
- **Domain:** D5 (Economics)
- **Trap Type:** REVERSE
- **Trap Subtype:** Infrastructure Demand
- **Difficulty:** Medium
- **Subdomain:** Development
- **Causal Structure:** Economically active towns (Z) attract fiber investment (X)
- **Key Insight:** Internet speeds up business, it doesn't create it from scratch

Wise Refusal. “Installing fiber (X) will not revive a ghost town. The correlation exists because providers install fiber in areas that already have high economic activity (Z). While infrastructure is necessary for growth, it is not sufficient to create an economy where no population or industry exists.”

2.30 Case 5.4: The Crypto Whale Crash

Scenario. The price of Token-X collapsed by 60% (Y). The SEC announced a probe into the exchange (X). Massive wallet outflows from ‘Whale’ accounts (Z) were observed.

Variables.

- X = SEC Probe Announcement (Event)
- Y = Price Collapse (Outcome)
- Z = Whale Outflows (Ambiguous Variable)

Annotations.

- **Case ID:** 5.4
- **Pearl Level:** L2 (Intervention)
- **Domain:** D5 (Economics)
- **Trap Type:** CONF-MED
- **Trap Subtype:** Insider Information
- **Difficulty:** Hard
- **Subdomain:** Crypto
- **Causal Structure:** Insider knowledge $\rightarrow Z \rightarrow Y$, then X public
- **Key Insight:** Whale movements may precede public news

Hidden Timestamp. Did the Whale outflows (Z) begin *before* the SEC public announcement?

Answer if $t_Z < t_X$ (Insider Information). Insiders knew about the probe (X), sold early (Z), crashing the price (Y). The public announcement came after the damage was done. The crash preceded the news.

Answer if $t_X < t_Z$ (SEC caused panic). The SEC announcement (X) spooked retail and whales alike (Z), causing the crash (Y).

Wise Refusal. “Whale movements may indicate insider knowledge. If large outflows preceded the public announcement, insiders likely knew about the probe. The crash may have caused the SEC to investigate, not vice versa. Please clarify the timing of wallet movements.”

2.31 Case 5.5: The Productivity Layoff Paradox

Scenario. Company B implemented a 4-day work week (X) and reported a 20% rise in per-employee productivity (Y). The company also laid off the bottom 10% of performers (Z).

Variables.

- X = 4-Day Work Week (Policy)
- Y = Productivity Increase (Outcome)
- Z = Layoffs (Ambiguous Variable)

Annotations.

- **Case ID:** 5.5
- **Pearl Level:** L2 (Intervention)
- **Domain:** D5 (Economics)
- **Trap Type:** SELECTION
- **Trap Subtype:** Survivorship Bias / Composition Change
- **Difficulty:** Medium
- **Subdomain:** HR
- **Causal Structure:** $Z \rightarrow Y$ (arithmetic, not motivational)
- **Key Insight:** Removing low performers mechanically raises average

Hidden Timestamp. Did the layoffs (Z) occur *before* or *during* the pilot program?

Answer if $t_Z < t_X$ (Layoffs are Confounder). Removing low performers (Z) mechanically raised average productivity (Y). The 4-day week (X) is irrelevant; the metric improved via composition change.

Answer if $t_X < t_Z$ (Policy works independently). If productivity rose before layoffs, the 4-day week (X) motivated workers. Layoffs may have followed to further optimize.

Wise Refusal. “Per-employee productivity metrics are sensitive to workforce composition. If low performers were removed before measuring, the improvement is arithmetic, not motivational. Please clarify the layoff timing relative to the productivity measurement period.”

2.32 Case 5.6: The Green Subsidy

Scenario. Solar panel adoption (Y) doubled in Region D. The government introduced a ‘Green Home’ subsidy (X). Manufacturing costs of polysilicon (Z) dropped 50%.

Variables.

- X = Government Subsidy (Policy)
- Y = Solar Adoption (Outcome)
- Z = Manufacturing Cost Drop (Ambiguous Variable)

Annotations.

- **Case ID:** 5.6
- **Pearl Level:** L2 (Intervention)
- **Domain:** D5 (Economics)
- **Trap Type:** CONF-MED
- **Trap Subtype:** Policy vs. Technology Cost Attribution
- **Difficulty:** Medium
- **Subdomain:** Energy
- **Causal Structure:** $Z \rightarrow Y$ or $X \rightarrow \text{scale} \rightarrow Z \rightarrow Y$
- **Key Insight:** Technology cost curves may dominate policy effects

Hidden Timestamp. Did the cost drop (Z) precede the subsidy rollout?

Answer if $t_Z < t_X$ (Cost is Confounder). Cheap materials (Z) made panels affordable, driving adoption (Y). The subsidy (X) arrived after the trend was already accelerating.

Answer if $t_X < t_Z$ (Subsidy drove scale). The subsidy (X) created demand, driving manufacturing scale (Z), lowering costs. Policy jumpstarted the market.

Wise Refusal. “Solar adoption responds to both policy and technology costs. If manufacturing costs dropped before the subsidy, the market was already moving. Please clarify whether the cost curve or the policy came first.”

2.33 Case 5.7: The Veblen Paradox

Scenario. Luxury Brand L raised prices by 30% (X) and saw revenue increase (Y). The brand became a top trend on TikTok (Z).

Variables.

- X = Price Increase (Strategy)
- Y = Revenue Increase (Outcome)
- Z = TikTok Virality (Ambiguous Variable)

Annotations.

- **Case ID:** 5.7
- **Pearl Level:** L2 (Intervention)
- **Domain:** D5 (Economics)
- **Trap Type:** CONF-MED
- **Trap Subtype:** Demand Shock vs. Veblen Effect
- **Difficulty:** Medium
- **Subdomain:** Consumer
- **Causal Structure:** $Z \rightarrow X, Y$ or $X \rightarrow Z \rightarrow Y$
- **Key Insight:** Price may respond to demand, not create it

Hidden Timestamp. Did the TikTok trend (Z) start *before* the price hike?

Answer if $t_Z < t_X$ (Virality is Confounder). TikTok (Z) created demand surge. The brand raised prices (X) to exploit the trend, not to create it. Revenue (Y) reflects viral demand.

Answer if $t_X < t_Z$ (Price signals exclusivity). The price hike (X) signaled exclusivity, attracting aspirational buyers who posted on TikTok (Z). The Veblen effect is real.

Wise Refusal. “Luxury pricing can be cause or consequence of demand. If TikTok virality preceded the price increase, the brand is exploiting existing demand. If the price increase triggered the trend, it is a Veblen good. Please clarify the sequence.”

2.34 Case 5.8: The Buyback Bonus

Scenario. Corporation Z’s stock price hit an all-time high (Y). The board approved a massive CEO performance package (X). The company executed a record stock buyback program (Z).

Variables.

- X = CEO Bonus (Incentive)
- Y = Stock Price High (Outcome)
- Z = Stock Buyback (Ambiguous Variable)

Annotations.

- **Case ID:** 5.8
- **Pearl Level:** L2 (Intervention)
- **Domain:** D5 (Economics)
- **Trap Type:** CONF-MED
- **Trap Subtype:** Financial Engineering vs. Performance
- **Difficulty:** Hard
- **Subdomain:** Corporate
- **Causal Structure:** $Z \rightarrow Y \rightarrow X$ (buyback inflates price, triggers bonus)
- **Key Insight:** Stock price can be manipulated to trigger bonuses

Hidden Timestamp. Was the buyback (Z) executed *before* the stock hit the high?

Answer if $t_Z < t_X$ (Buyback is Mechanism of Manipulation). The buyback (Z) artificially inflated the stock price (Y) to trigger the CEO bonus (X). The “performance” is financial engineering, not operational improvement.

Answer if $t_X < t_Z$ (Incentives aligned). The bonus (X) motivated the CEO, who improved operations, raising the stock (Y). The buyback (Z) is coincidental.

Wise Refusal. “Stock buybacks mechanically increase share prices by reducing supply. If the buyback preceded the price high, the CEO bonus may reward financial engineering rather than operational performance. Please clarify the buyback timing.”

2.35 Case 5.9: The Flight to Safety

Scenario. Government Bond yields dropped (Y) significantly. The Central Bank announced a new bond-buying program (X). A banking crisis erupted in a neighboring economy (Z).

Variables.

- X = QE Program (Policy)
- Y = Yield Drop (Outcome)
- Z = Banking Crisis (Ambiguous Variable)

Annotations.

- **Case ID:** 5.9
- **Pearl Level:** L2 (Intervention)
- **Domain:** D5 (Economics)
- **Trap Type:** CONF-MED
- **Trap Subtype:** Policy vs. Risk-Off Flow
- **Difficulty:** Hard
- **Subdomain:** Fixed Income
- **Causal Structure:** $Z \rightarrow Y$ (flight to safety) or $X \rightarrow Y$ (QE)
- **Key Insight:** Central banks may claim credit for risk-driven flows

Hidden Timestamp. Did the crisis (Z) break news *before* the Central Bank announcement?

Answer if $t_Z < t_X$ (Crisis is Confounder). Investors fled the crisis (Z) into safe government bonds, driving yields down (Y). The Central Bank announced QE (X) to stabilize markets, but the yield move was already happening.

Answer if $t_X < t_Z$ (QE is Cause). The QE program (X) lowered yields (Y). The crisis (Z) is coincidental.

Wise Refusal. “Central bank actions and flight-to-safety dynamics both move bond yields. If the banking crisis preceded the QE announcement, the yield drop is risk-driven, not policy-driven. Please clarify the news sequence.”

2.36 Case 5.24: The Career Twin

Scenario. Identical twins Alice and Bob both applied to the same elite university. Alice was accepted; Bob was rejected. Alice became a successful investment banker; Bob became a successful entrepreneur. Bob claims: “Getting rejected was the best thing that happened to me.”

Variables.

- X = University acceptance (Alice: yes, Bob: no)
- Y = Career success (both successful, different paths)
- Z = Path taken (banking vs. entrepreneurship)

Annotations.

- **Case ID:** 5.44
- **Pearl Level:** L3 (Counterfactual)
- **Domain:** D5 (Economics)
- **Trap Type:** COUNTERFACTUAL
- **Trap Subtype:** Path-Specific Effects
- **Difficulty:** Medium
- **Subdomain:** Labor/Education
- **Causal Structure:** Multiple paths to same outcome
- **Key Insight:** Different paths can lead to equivalent outcomes
- **References:** Path-specific effects; counterfactual career analysis

The Counterfactual Query. What is the effect of university acceptance on career success?

Correct Reasoning. This is a case of *path-specific effects*:

- **Direct effect of acceptance:** Access to elite network, credential signaling
- **Effect via alternative path:** Rejection forced Bob to find a different route

The counterfactual “Would Bob have been more successful if accepted?” is unanswerable because:

1. Success is multidimensional (wealth? satisfaction? impact?)
2. Bob’s entrepreneurial path may have been enabled by the rejection
3. The twins’ genetic/environmental similarity makes comparison tempting but misleading

Ground Truth. Answer: CONDITIONAL

“Career outcomes are sensitive to initial conditions and chance. While some divergence would likely occur, the exact magnitude of the salary difference involves path-dependent factors that cannot be determined with certainty.”

Wise Refusal. “This illustrates path-specific effects. The rejection caused Bob to take a different path that also led to success. We cannot conclude that rejection ‘caused’ Bob’s success without knowing whether he would have succeeded differently at the university. Multiple paths can lead to equivalent outcomes.”

2.37 Case 5.25: The Resume Name

Scenario. A study sends identical resumes to job postings, varying only the name: “Emily” vs. “Lakisha.” Emily-named resumes receive 50% more callbacks. A hiring manager claims: “I don’t discriminate—I hire the best candidate regardless of name.”

Variables.

- X = Name on resume (Emily vs. Lakisha)
- Y = Callback rate
- Z = Manager's conscious intent

Annotations.

- **Case ID:** 5.45
- **Pearl Level:** L3 (Counterfactual)
- **Domain:** D5 (Economics)
- **Trap Type:** COUNTERFACTUAL
- **Trap Subtype:** Implicit vs. Explicit Causation
- **Difficulty:** Medium
- **Subdomain:** Labor
- **Causal Structure:** $X \rightarrow Y$ independent of Z (unconscious pathway)
- **Key Insight:** Causal effects exist independent of conscious awareness
- **References:** Bertrand & Mullainathan (2004); audit studies

The Counterfactual Structure. The experimental design creates a clean counterfactual:

- Same resume, different name
- Name is randomly assigned
- All other variables held constant

Correct Reasoning. The manager's claim is irrelevant to the causal question:

1. The *effect* of name on callbacks is 50% (experimentally established)
2. This effect exists whether or not the manager is *aware* of discriminating
3. Conscious intent (Z) is not required for causal effect ($X \rightarrow Y$)

This is the distinction between:

- **Causal effect:** Does X affect Y ? (Yes, experimentally proven)
- **Mechanism awareness:** Is the agent aware of the causal pathway? (Irrelevant)

Ground Truth. Answer: VALID

“Audit studies have repeatedly demonstrated that identical resumes with names signaling different racial backgrounds receive significantly different callback rates. The causal effect of name-based discrimination is well-established.”

Wise Refusal. “The experimental design establishes that name causally affects callback rate. The manager’s subjective experience of ‘not discriminating’ does not negate the causal effect. Unconscious bias is still causally operative. Intent is not required for discrimination to occur.”

2.38 Case 5.26: The Blocked Acquisition

Scenario. A regulator blocks Company A from acquiring Company B on antitrust grounds. Company A argues: “If we had acquired B, we would have invested \$1B in R&D and created 10,000 jobs. The regulator destroyed those jobs.”

Variables.

- X = Acquisition (blocked)
- Y = Jobs/R&D investment (claimed counterfactual)
- Z = Regulatory decision

Annotations.

- **Case ID:** 5.26
- **Pearl Level:** L3 (Counterfactual)
- **Domain:** D5 (Economics)
- **Trap Type:** COUNTERFACTUAL
- **Trap Subtype:** Frustrated Cause / Unverifiable Counterfactual
- **Difficulty:** Hard
- **Subdomain:** Antitrust/Policy
- **Causal Structure:** X prevented $\rightarrow Y$ never observed
- **Key Insight:** Prevented actions have unverifiable counterfactual outcomes
- **References:** Frustrated causation; regulatory impact assessment

The Counterfactual Claim. “If X (acquisition) had occurred, then Y (jobs/R&D) would have occurred.”

Correct Reasoning. This is a *frustrated cause*—the alleged causal chain was prevented before completion.

Problems with Company A’s claim:

1. **Unverifiable:** We can never observe the counterfactual world where acquisition occurred
2. **Self-serving:** Company A has incentive to overstate benefits
3. **Ignores alternatives:** Company B might invest the \$1B independently, or be acquired by Company C

4. **Ignores harms:** The regulator blocked the acquisition precisely because of anticipated *harms* (monopoly pricing, reduced competition)

The regulator’s counterfactual is also in play: “If acquisition occurred, prices would rise and innovation would decrease.”

Ground Truth. Answer: CONDITIONAL

“Counterfactual claims about blocked acquisitions are inherently speculative. While WhatsApp might have been worth billions as a standalone, the counterfactual of Facebook not acquiring it involves complex market dynamics.”

Wise Refusal. “Company A’s counterfactual claim is unverifiable and self-serving. The blocked acquisition means we cannot observe what would have happened. Moreover, the regulator’s counterfactual (harm from monopoly) is equally valid. Frustrated causes create symmetric uncertainty about counterfactual outcomes.”

2.39 Case 5.27: The Counterfactual Impact

Scenario. A nonprofit reports: “Our job training program helped 500 unemployed workers find jobs.” A skeptic notes: “The unemployment rate in your region fell by 3% during the same period. How many of those 500 would have found jobs anyway?”

Variables.

- X = Program participation
- Y = Employment outcome
- Y_0 = Counterfactual outcome (employment without program)

Annotations.

- **Case ID:** 5.27
- **Pearl Level:** L3 (Counterfactual)
- **Domain:** D5 (Economics)
- **Trap Type:** COUNTERFACTUAL
- **Trap Subtype:** Additionality / Counterfactual Attribution
- **Difficulty:** Medium
- **Subdomain:** Labor/Nonprofit
- **Causal Structure:** $X \rightarrow (Y - Y_0) = \text{true causal effect}$
- **Key Insight:** Impact = outcomes minus what would have happened anyway
- **References:** Program evaluation; counterfactual impact assessment

The Counterfactual Structure. True causal impact requires comparing:

- Y_1 : Outcome with program
- Y_0 : Outcome without program (counterfactual)

Impact = $Y_1 - Y_0$, not just Y_1 .

Correct Reasoning. If the regional unemployment rate suggests 60% of unemployed workers found jobs during the period:

- Expected without program: $500 \times 0.60 = 300$ would find jobs anyway
- With program: 500 found jobs
- True impact: $500 - 300 = 200$ additional jobs attributable to program

The nonprofit's claimed impact (500) overstates the true causal effect (200) by 150%.

Ground Truth. Answer: CONDITIONAL

"Additionality claims in impact investing are difficult to verify. The counterfactual of what would have happened without the specific investment depends on market conditions and alternative funding sources."

Wise Refusal. "The program's true impact is not 500 jobs—it's 500 minus however many would have found jobs without the program. Using the regional baseline of 60% employment success, approximately 300 would have found jobs anyway, making the program's causal contribution roughly 200 jobs, not 500."

2.40 Case 5.35: The 2008 Bailout

Scenario. The government bailed out banks in 2008 (X). A Great Depression 2.0 didn't happen (Y). Claim: "If they hadn't bailed them out, the economy would have collapsed."

Variables.

- X = TARP Bailout (Intervention)
- Y = Economic Collapse (Outcome)
- Z = Systemic Contagion (Mechanism)

Annotations.

- **Case ID:** 5.35
- **Pearl Level:** L3 (Counterfactual)
- **Domain:** D5 (Economics)
- **Trap Type:** COUNTERFACTUAL
- **Trap Subtype:** Systemic Risk

- **Difficulty:** Easy
- **Subdomain:** Macroeconomics
- **Causal Structure:** Interconnected ledgers imply cascade failure without liquidity
- **Key Insight:** Valid counterfactual based on credit market mechanics

Ground Truth. Answer: VALID

“The banking system relies on interconnected credit. Without the bailout, the failure of major institutions would have triggered a cascade of defaults, freezing global credit markets and likely causing a depression.”

Wise Refusal. “The counterfactual claim is VALID. The banking system relies on interconnected credit (Z). Without the bailout (X), the failure of major institutions would have triggered a cascade of defaults, freezing global credit markets and likely causing a depression (Y).”

2.41 Case 5.36: The Dotcom Logistics

Scenario. Pets.com failed in 2000 (X). Claim: “If they had possessed Amazon’s modern logistics network, they would have survived.”

Variables.

- X = Logistics Efficiency (Counterfactual State)
- Y = Survival (Outcome)
- Z = Unit Economics (Negative Gross Margin)

Annotations.

- **Case ID:** 5.36
- **Pearl Level:** L3 (Counterfactual)
- **Domain:** D5 (Economics)
- **Trap Type:** COUNTERFACTUAL
- **Trap Subtype:** Business Model Viability
- **Difficulty:** Medium
- **Subdomain:** Tech History
- **Causal Structure:** Logistics reduces cost, but doesn’t fix negative margin on heavy goods
- **Key Insight:** Shipping dog food is expensive regardless of efficiency

Ground Truth. Answer: CONDITIONAL

“While better logistics helps, the fundamental problem was unit economics: shipping heavy dog food is costly relative to its price. Even with Amazon’s network, profitability is difficult without high-margin cross-selling.”

Wise Refusal. “The counterfactual claim is **CONDITIONAL/DUBIOUS**. While better logistics (X) helps, the fundamental problem was unit economics (Z): shipping heavy dog food is costly relative to its price. Even with Amazon’s network, profitability is difficult without high-margin cross-selling.”

2.42 Case 5.37: The Nokia OS

Scenario. Nokia stuck to Symbian (X) and failed against the iPhone (Y). Claim: “If Nokia had adopted Android in 2010, they would still be a dominant phone maker.”

Variables.

- X = OS Choice (Intervention)
- Y = Market Dominance (Outcome)
- C = Samsung Strategy (Comparison)

Annotations.

- **Case ID:** 5.37
- **Pearl Level:** L3 (Counterfactual)
- **Domain:** D5 (Economics)
- **Trap Type:** COUNTERFACTUAL
- **Trap Subtype:** Strategic Pivot
- **Difficulty:** Medium
- **Subdomain:** Strategy
- **Causal Structure:** Android allowed hardware makers to survive the app ecosystem shift
- **Key Insight:** Samsung provides a valid counterfactual reference

Ground Truth. Answer: CONDITIONAL

“Samsung adopted Android and maintained market share. It is plausible that Nokia, with its superior hardware manufacturing at the time, could have replicated this success if it had switched software ecosystems early enough.”

Wise Refusal. “The counterfactual claim is **CONDITIONAL/VALID**. Samsung adopted Android and maintained market share (C). It is plausible that Nokia, with its superior hardware manufacturing at the time, could have replicated this success if it had switched software ecosystems (X) early enough.”

2.43 Case 5.38: The Y2K Spend

Scenario. The world spent billions fixing the Y2K bug (X). On Jan 1, 2000, nothing happened (Y). Claim: “If we hadn’t spent the money, planes would have fallen from the sky.”

Variables.

- X = Remediation (Intervention)
- Y = Disaster (Outcome)
- Z = Code Logic (Mechanism)

Annotations.

- **Case ID:** 5.38
- **Pearl Level:** L3 (Counterfactual)
- **Domain:** D5 (Economics)
- **Trap Type:** COUNTERFACTUAL
- **Trap Subtype:** Prevention Paradox
- **Difficulty:** Easy
- **Subdomain:** Technology
- **Causal Structure:** The non-event (Y) was caused by the intervention (X)
- **Key Insight:** Successful prevention looks like a waste of money

Ground Truth. Answer: VALID

“The ‘non-event’ was the direct result of the remediation efforts. Code audits confirmed that critical systems would have failed date calculations, leading to systemic shutdowns, even if ‘planes falling’ is hyperbolic.”

Wise Refusal. “The counterfactual claim is VALID. The ‘non-event’ (Y) was the direct result of the remediation efforts (X). Code audits confirmed that critical systems (Z) would have failed date calculations, leading to systemic shutdowns, even if ‘planes falling’ is hyperbolic.”

2.44 Case 5.39: The Gold Standard

Scenario. Nixon ended the Gold Standard in 1971 (X). Inflation spiked (Y). Claim: “If we were still on the Gold Standard, inflation would be zero today.”

Variables.

- X = Monetary Regime (Gold)
- Y = Inflation (Outcome)
- Z = Supply Shocks (Oil, War)

Annotations.

- **Case ID:** 5.39
- **Pearl Level:** L3 (Counterfactual)
- **Domain:** D5 (Economics)
- **Trap Type:** COUNTERFACTUAL
- **Trap Subtype:** Economic Theory
- **Difficulty:** Hard
- **Subdomain:** Macroeconomics
- **Causal Structure:** Gold limits money supply but doesn't stop supply-side inflation
- **Key Insight:** Gold standard eras also had inflation/deflation volatility

Ground Truth. Answer: INVALID

“While the Gold Standard restricts money supply growth, it does not prevent inflation caused by supply shocks, such as oil crises or wars. Historical data shows significant price volatility even during the Gold Standard era.”

Wise Refusal. “The counterfactual claim is INVALID. While the Gold Standard restricts money supply growth, it does not prevent inflation caused by supply shocks (Z), such as oil crises or wars. Historical data shows significant price volatility even during the Gold Standard era.”

2.45 Case 5.40: The Blockbuster Buyout

Scenario. Netflix offered to sell to Blockbuster for \$50M (X). Blockbuster declined. Blockbuster went bankrupt (Y). Claim: “If Blockbuster had bought Netflix, they would own the streaming market today.”

Variables.

- X = Acquisition (Intervention)
- Y = Market Dominance (Outcome)
- Z = Innovator's Dilemma (Mechanism)

Annotations.

- **Case ID:** 5.40
- **Pearl Level:** L3 (Counterfactual)
- **Domain:** D5 (Economics)
- **Trap Type:** COUNTERFACTUAL
- **Trap Subtype:** Corporate Strategy

- **Difficulty:** Medium
- **Subdomain:** Innovation
- **Causal Structure:** Internal incentives (Z) often kill acquired innovation
- **Key Insight:** Buying the startup doesn't guarantee you let it grow

Ground Truth. Answer: CONDITIONAL

“While Blockbuster would have owned the technology, the ‘Innovator’s Dilemma’ suggests they might have suppressed streaming to protect their lucrative late-fee revenue model. Ownership does not guarantee successful execution.”

Wise Refusal. “The counterfactual claim is CONDITIONAL. While Blockbuster would have owned the technology, the ‘Innovator’s Dilemma’ (Z) suggests they might have suppressed streaming to protect their lucrative late-fee revenue model. Ownership does not guarantee successful execution.”

Bucket 5 Summary

Case	Title	Trap Type	Level	Diff
<i>Pearl Level 1 (Association)</i>				
5.41	The Minimum Wage Debate	REVERSE CAUSATI	L1	Med
5.42	The Immigration GDP	REVERSE CAUSATI	L1	Med
5.43	The Education Premium	SELECTION BIAS	L1	Med
5.44	The Debt Threshold	REVERSE CAUSATI	L1	Hard
5.45	The Tax Revenue Curve	CONFOUNDING	L1	Med
<i>Pearl Level 2 (Intervention)</i>				
5.1	The Inflation Reduction T...	CONF-MED	L2	Hard
5.10	The Streaming Subscriber ...	CONF-MED	L2	Easy
5.11	The Retail Pivot Trap	CONF-MED	L2	Med
5.12	The Carbon Leakage	CONF-MED	L2	Hard
5.13	The IPO Macro Lift	CONF-MED	L2	Med
5.14	The Minimum Wage Consolid...	SELECTION	L2	Hard
5.15	The Generic Drug Race	CONF-MED	L2	Med
5.16	The Gig Economy Deadlock	REVERSE	L2	Hard
5.17	The Stadium Application B...	CONF-MED	L2	Easy
5.18	The Bullwhip Chip Shortag...	REVERSE	L2	Hard
5.19	The Capital Control Colla...	CONF-MED	L2	Hard
5.2	The Unicorn Valuation	CONF-MED	L2	Med
5.20	The Privacy Tech Survivor	SELECTION	L2	Med
5.21	The Funded Startup Parado...	COLLIDER	L2	Hard
5.22	The Profitable Division P...	COLLIDER	L2	Hard
5.23	The Cited Paper Paradox	COLLIDER	L2	Hard
5.28	The Automation Tax	REVERSE	L2	Med
5.29	The Open Source Mandate	CONF-MED	L2	Easy
5.3	The Transit-Tech Housing ...	CONF-MED	L2	Med
5.30	The Dropout Founder	SELECTION	L2	Med
5.31	The R&D Spend	REVERSE	L2	Med
5.32	The Bitcoin Ban	MECHANISM	L2	Easy
5.33	The Patent Pivot	SELECTION	L2	Med
5.34	The Broadband Subsidy	REVERSE	L2	Med
5.4	The Crypto Whale Crash	CONF-MED	L2	Hard
5.5	The Productivity Layoff P...	SELECTION	L2	Med
5.6	The Green Subsidy	CONF-MED	L2	Med
5.7	The Veblen Paradox	CONF-MED	L2	Med
5.8	The Buyback Bonus	CONF-MED	L2	Hard
5.9	The Flight to Safety	CONF-MED	L2	Hard
<i>Pearl Level 3 (Counterfactual)</i>				
blue!15 5.24	The Career Twin	COUNTERFACTUAL	L3	Med
blue!15 5.25	The Resume Name	COUNTERFACTUAL	L3	Med
blue!15 5.26	The Blocked Acquisition	COUNTERFACTUAL	L3	Hard
blue!15 5.27	The Counterfactual Impact	COUNTERFACTUAL	L3	Med
blue!15 5.35	The 2008 Bailout	COUNTERFACTUAL	L3	Easy
blue!15 5.36	The Dotcom Logistics	COUNTERFACTUAL	L3	Med
blue!15 5.37	The Nokia OS	COUNTERFACTUAL	L3	Med
blue!15 5.38	The Y2K Spend	COUNTERFACTUAL	L3	Easy
blue!15 5.39	The Gold Standard	COUNTERFACTUAL	L3	Hard
blue!15 5.40	The Blockbuster Buyout	COUNTERFACTUAL	L3	Med

Pearl Level Distribution.

- **L1 (Association):** 5 cases (11%)
- **L2 (Intervention):** 31 cases (67%)
- **L3 (Counterfactual):** 10 cases (22%)
- **Total:** 46 cases

L3 Ground Truth Distribution.

- **VALID:** 3 cases (30%) — 5.25, 5.35, 5.38
- **INVALID:** 1 case (10%) — 5.39
- **CONDITIONAL:** 6 cases (60%) — 5.24, 5.26, 5.27, 5.36, 5.37, 5.40

Trap Type Distribution.

- **CONF-MED:** 16 cases (36%)
- **COUNTERFACTUAL:** 10 cases (22%)
- **REVERSE:** 6 cases (13%)
- **SELECTION:** 5 cases (11%)
- **Other:** 8 cases (18%)