# Learning with misspecified causal models: Experimental tests

# PRELIMINARY AND INCOMPLETE

# DO NOT SHARE

Sandro Ambuehl and Jindi Huang *

July 8, 2025

**Abstract**

A recent theoretical literature studies learning through the lens of misspecified causal models and its implications for choice. We experimentally test the fundamental assumptions and key predictions of such models to answer three sets of questions. First, if subjects begin with a misspecified model, do they continue to interpret data through that lens as the literature assumes, or do they discover the misspecification and update the structure of their mental model? Second, in what way do individuals approach choice problems with unknown causal relationships? Do they learn causal networks, action-response associations, or heuristics? Third, do subjects exhibit the comparative statics and equilibrium behavior predicted in the theoretical literature? In our experiment, subjects interact with causal structures in many rounds about which they initially know nothing. In pilot experiments, our design reliably generates causal misperceptions which, for a large fraction of participants, persist until the final round, and cause costly decision mistakes. They indicate that subjects learn causal networks rather than mere action-outcome associations. Yet, some of the key theoretical predictions fail, largely due to subjects' lack of responsiveness to moderate changes in the magnitude of correlations in the data. We will complete the main data collection in time for the CESifo conference.

*JEL codes: C91, D01, D83*

*Ambuehl: Department of Economics and UBS Center for Economics in Society, University of Zurich, Blüemlisalpstrasse 10, 8006 Zürich, Switzerland, sandro.ambuehl@econ.uzh.ch. Huang: Department of Economics, University of Zurich, Blüemlisalpstrasse 10, 8006 Zürich, Switzerland, jindi.huang@econ.uzh.ch. This project started together with Heidi C. Thysen (then at the Norwegian School of Economics, now at Tide Buss AS) who made invaluable contributions to this project in its the initial stages. We are grateful to Rahul Bhui for helpful comments and discussions.

# 1 Introduction

A recent theoretical literature studies implications of agents learning through the lens of misspecified causal models (see, e.g. **??**, for reviews). An individual who views the world through the lens of a misspecified causal model will often make systematically biased choices, especially if the data from which she (mis)learns are generated by her own misinformed actions (as in **?**). For instance, such an individual may come to attribute causal power to an ineffective medication if she always takes it shortly before getting better, or to ineffective marketing spending that only occurs when a favorable environment yields ample funding for such campaigns. She may misperceive the causal relation between policy variables such as inflation and unemployment and thus vote for ineffective policies.

This paper experimentally tests the fundamental assumptions and key predictions of theories of learning through misspecified models. We address three sets of questions. First, if subjects have a misspecified model in mind, do they in fact interpret data through that lens (as in **?**), or do they update their models towards more faithful representations of causal structure, as a Bayesian would (as in **?**)? Second, in what way do individuals approach choice problems with unknown causal relationships? Do they learn causal networks through which they interpret the data (as in **?**), mere action-response associations (as in **?**), or decide based on heuristics such as win-stay-lose-shift? Third, if subjects interpret data through subjective causal networks that are misspecified, do subjects exhibit the equilibrium behavior and comparative statics predicted in the theoretical literature? Specifically, when subjects' beliefs feed back into themselves because ill-informed actions generate the data from which subjects update their beliefs, does behavior conform to the predictions of *personal equilibrium* of **?**?

We focus on settings in which decision makers repeatedly interact with stochastic causal systems, form beliefs about them, and choose actions accordingly. There are many examples, such as busines managers (e.g. the gas station managers in **?**), medical doctors who, over their career, observe a long sequence of patients presenting with similar symptoms, or central bank presidents who make monetary policy decisions several times each year. Each of these decisions is potentially informed by many variables (pricing, marketing effort, or even weather in the case of businesses' profits; patient demographics, behavior, medication choice and adherence in the case of medical doctors; inflation expectations, consumer sentiment, global tariff policies, exchange rates and many more in the case of central bank decisions). The relevant variables may be linked in various ways; they may be causes, mediators, symptoms, or non-causal predictors of other variables. A decision maker whose subjective causal model incorrectly assigns these causal roles to the variables she observes will often draw systematically biased inference and make systematically wrong decisions (**?**, reviews the theoretical implications of such causal misperceptions).

We answer our research questions using stylized experiments. These afford us the level of control and the extent of measurement we need to test the theoretical predictions. We consider decisions that

involve three binary variables: subjects' *action* $A$, a payoff-relevant *outcome* $Y$, and a third variable $X$ that is not directly payoff relevant but that may be a mediator, cause, predictor, or symptom of either or both other variables. This is the simplest setting that lets us study our research questions.

We focus on two settings. In the first, which we call the *Reverse Causality* problem, subjects take action before observing the realization of any other variable, but may mistake a symptom of the outcome as a cause. This setting implements the main example of **?** and provides the strongest and most detailed tests of the theoretical predictions. Yet, in many real-world cases, decision makers choose only after observing any information about the current state. Managers, for instance, may observe changes in business conditions before deciding about marketing spending, medical doctors observe patient histories before deciding on treatment, and patients observe internal states before deciding whether to consume over-the-counter medications. Hence, our second setting presents the decision maker with information before she acts. Because this setting lets the decision maker condition her action on another variable, inference about the causal effect of her action may be confounded by that other variable. We thus call this second setting the *Confounded Choice* problem. The two settings differ not only in terms of which variables are observable at the time of taking the action, but also in terms of the information needed to infer the correct causal model. In the Reverse Causality problem, the action is ancestral, so observational data suffices to infer the true causal effect of the action on the outcome. In the Confounded Choice problem, by contrast, correlational information alone does not reveal the causal effect of the action on the outcome. Instead, explicit experimentation must disconnect the action choice from the confounding variable to draw correct causal inferences.

In the first setting, the richest implications among all qualifying three-node causal networks emerge if the true causal structure is described by the DAG $A \rightarrow X \leftarrow Y$, but the subject believes that the relation is $A \rightarrow X \rightarrow Y$. In short, the decision maker mistakes $X$ for a cause of $Y$ even though it is, in fact, a symptom of $Y$. This is the structure of the Dieter's Dilemma of **?**. Assuming that the action $A$ masks the symptom $X$, that the action is costly, and that the decision maker prefers the outcome $Y$ to be off rather than on, a Bayesian quickly learns that the action is ineffective and cease to take it. Yet, as **?** shows, a misspecified model learner will come to take the action with an intermediate probability that depends on the noisiness of the link between $X$ and $Y$, and on the payoff and cost parameters. We test the key best-response dynamic that undergirds this equilibrium through treatments in which we exogenously manipulate the action probability the decision maker observes: when the action is taken rarely, the symptom is not masked, so the decision maker observes a strong correlation between symptom and outcome. She misinterprets this correlation causally, thus believes that the action is effective in improving the outcome, and hence takes it often. When the action is taken often, by contrast, the symptom is often masked, the observed correlation between symptom and outcome low, and the decision maker's assessment of the power of the action to affect the outcome is more pessimistic. So she takes the action rarely. In addition to best responses to exogenously given

action probabilities, we consider treatments in which the decision maker's actions fully determine the data she observes. These let us test equilibrium comparative statics.

In our second setting, the Confounded Choice problem, the decision maker observes a variable $X$, and then decides whether to take action $A$ to potentially influence the outcome $Y$ she is interested in. We consider the case in which $X$ predicts the outcome $Y$, and the costly action $A$ does not affect anything, so that the true DGP is described by $A \leftarrow X \rightarrow Y$. A decision maker whose action choice responds to $X$, however, will observe a correlation between her action $A$ and the outcome $Y$, and may thus come to believe that she can causally influence the outcome via the subjective model $X \rightarrow A \rightarrow Y$. For instance, a manager may have excess funds to spend on marketing whenever the economy is strong, and may misattribute the correlation between marketing spending and strong sales (both driven by the state of the economy) to a causal effect of marketing expenditure on spending. Here, we test whether the causal misperception arises and persists, as well as the comparative statics of changing the strength of the relation between $Y$ and $X$.

Implementing these settings experimentally requires addressing three challenges. First, subjects need to interpret the data through a misspecified model at some point. To induce such interpretations, we present the realizations of the variables to subjects in a temporal order that may or may not match the causal order.[1] Specifically, we inform subjects that they will see the realization of $X$ before the realization of $Y$ in each round, no matter whether $X$ causes $Y$ or $Y$ causes $X$. When the causal and temporal orders differ, subjects who infer causal order from temporal order will interpret the data through the lens of a mistaken causal model. Second, we seek to prevent contamination from prior beliefs about causal structure from outside the laboratory. Hence, we frame the experimental task as a game. Next to the subject, there are two computer players (representing $X$ and $Y$) who may respond to the subject, to each other, or act independently. The subject and the computer players each have a button they can decide to press. The subjects' own decision whether to press the button may directly affect her payoff. In addition, one of the computer players' decisions also affects her payoff. Third, decision-making in a multi-round setting with unknown causal structures presents decision makers with an explore / exploit dilemma, yet our interest centers on steady-state behavior. To study best-response behavior, we experimentally suppress the exploration motive. In some rounds, subjects are exogenously forced to take a particular action and observe the consequences of that choice. In other rounds, subjects can freely choose their action, but do not observe the consequences of their choice. To study equilibrium predictions, subjects must be able to both freely choose the action and observe its consequences, preventing the suppression of exploration motives. Instead, we focus on the last 20 rounds when exporation motives have presumably vanished.

We have collected data from two pilot experiments on prolific.com. We will run the main sessions before the CESifo conference date in October. One pilot considers the Confounded Choice problem

[1] We do not simply tell subjects under what causal assumptions they should interpret the data. The reason is that subjects may not internalize such information, in which case they would not interpret data through the lens of that model, and, if the model is misspecified, the information would constitute deception.

with the goal of testing whether causal misperceptions arise; it does not include treatment variations. The second pilot considers the Dieter's Dilemma with variation in the key parameters to test the key theoretical predictions: exogenous variation in the action frequency to test best-response dynamics, and the extent of noise in the link between outcome $Y$ and symptom $X$ to test the key equilibrium prediction.

In both pilot experiments, we find pronounced causal misperceptions and the corresponding sub-optimal behavior. In the Confounded Choice problem, subjects systematically correlate their action with the predictor variable they observe, but then attribute independent causal power to their action. As a result, they frequently pay for taking the ineffective action. In the Dieter's Dilemma, subjects consistently mistake $X$, which is a the symptom $Y$, as a cause of $Y$, and thus invest in the ineffective costly action $A$ which, in fact, affects $X$.

With respect to our first question (do people stay with misspecified models or do they converge to more faithful representations of the DGP) we find that the majority of subjects do not converge to the correct model across 100 rounds, even though a Bayesian with uniform priors would place most probability mass on the correct model already within 10 rounds. This finding is consistent with the hypothesis that subjects continue to interpret data through the lens of their misspecified causal model rather than converge to a more faithful representation of causal structure.

With respect to our second question (how subjects approach choice problems with unknown causal relationships), we find that the average subject in the Dieter's Dilemma does not merely learn an action-outcome association, but instead forms beliefs about the variable $X$ which are predictive of her choices. The average subject does exhibit hallmarks of heuristic decision making, especially win-stay / lose shift. In each of the last 20 rounds, subjects who have have chosen $A = 1$ in the previous round are 15 p.p. more likely to keep choosing $A = 1$ if the outcome was good than if it was bad, in line with win-stay. The outcome has a much lower effect of 3 p.p. among subjects who have chosen $A = 0$ in the previous round, presenting weaker evidence for lose-shift.[2] At the same time, we observe a significant fraction of subjects who begin to play $A = 0$ throughout, consistent with the use of procedures such as Bayesian learning or model-free reinforcement learning that converge to the optimal action. Overall, the presence of indications of all four decision modes we consider suggests a considerable amount of individual heterogeneity.

Regarding our third question (predicted comparative statics), our pilot sample of 59 subjects does not reveal clear effects. Subjects' chosen action frequency does not vary depending on whether the exogenous action probability in the data is 20% or 80%; we thus do not detect the best-response behavior that would undergird the equilibrium predictions of personal equilibrium in **?**. Neither do we observe the predicted equilibrium comparative statics that strengthening the link between symptom and outcome in the Dieter's Dilemma increases the equilibrium action frequency.

[2] A possible reason is that subjects who have learned that there is no economic value from playing $A = 1$ will choose $A = 0$ regardless.

The pilot results rest on across-subject averages. It is unlikely that all subjects will approach causal inference problems in the same way. The main experiments will allow us to classify individuals into four types. The Misspecified Model Learner follows the predictions of **?**. The Bayesian learner updates her beliefs about the causal relations in accordance with Bayes' law. The Model-Free Reinforcement Learner (henceforth: MFRL) learns cause-effect associations without interpreting the data through the lens of a causal network . Finally, the Win-Stay/Lose-Shift type (henceforth: WSLS) implements a win-stay/lose-shift strategy or a generalization thereof without learning causal effects. We use exogenous manipulation in the action frequency to infer subjects' types. A Misspecified Model Learner who believes that $X$ is a cause of the action will take the action more often the lower the observed exogenous action frequency in the data. A Bayesian learner, by contrast, as well as a MFRL, will learn not to take the action. Finally, a WSLS type will take the action with a fixed probability that does not depend on the exogenous action frequency in the data. We present subjects with two treatments, one in which the exogenous action frequency is low, and one in which it is high. We classify people according to the action frequency they choose in each case in response to the exogenous action frequency. (To distinguish Bayesian and MFRL learners, we include measurements that let subjects manipulate $X$ rather than $A$ )

Our work is related to a sizeable theoretical literature that studies the implications of misspecified learning. The framework of **?** that models decision makers as fitting (misspecified) causal DAGs to data is closely related to other concepts, which our tests thus also address. These include analogy-based expectations (**??**), and cursed equilibrium (**?**). They are a special case of Berk-Nash equilibrium (**?**). While those frameworks concern games, our tests, which focus on individual decision problems, address assumptions those frameworks place on best response functions.

Our work is also related to a vast literature in cognitive science that shows how the causal DAGs approach can explain large swaths of human cognition **??**, for recent book-length reviews. That literature focuses on how individuals acquire correct causal models rather than on the implications of misspecified causal models. Neither does it consider equilibrium concepts such as the personal equilibrium concept of **?** that is a key focus of our empirical tests.

Within experimental economics, our work is related to **?** who show, in the case of the canonical Bayesian updating problem, that subjects learn better from case-by-case feedback about stochastic realizations when they are not given an explanation of the problem, since that explanation causes the typical base-rate neglect issue. Our work differs from theirs in two key ways. First, we study *causal* inference. Second, we consider the equilibrium implications when subjects generate the data from which they subsequently learn. Our work is also related to **?** who study how narratives affect the interpretation of exogenously provided data, and of the narratives subjects invent themselves. Unlike our paper, they do not let subjects interact with the causal systems, and hence cannot test equilibrium concepts.

This paper proceeds as follows. Section **??** formally defines the types we distinguish. Section **??** details our strategy to identify these types. Section **??** outlines key design elements. Section **??** presents our pilot results.

# 2 Theory

The DGP in the Dieter's Dilemma is consistent with the following DAG: $A \rightarrow X \leftarrow Y$, where $A, X, Y \in \{0, 1\}$. We parametrize it as $P(Y) = \gamma$, and $P(X|A, Y)$ is the noisy-OR. That is, $P(X = 1|A, Y) = 1 - (1 - \epsilon)^A(1 - \epsilon)^Y$, where $\epsilon \in [0, 0.5]$. Hence, $X$ is on with a high probability if $A$ or $Y$ is on; otherwise it is off. In one treatment, we instead use the noisy-AND function, $P(X = 1|A, Y) = (1 - \epsilon)^A(1 - \epsilon)^Y$, so that $X$ is on with a high probability only if both $A$ and $Y$ are on.

The subjects' payoff function is $u(A, X, Y) = Y - cA$ where $c > 0$ is a cost parameter. Because the action does not causally affect $Y$ but it cost is positive, optimal behavior entails never taking the action.

## 2.1 Type definitions

We consider four main types: (i) Bayesian (Subsection **??**), (ii) Misspecified model learner (Subsection **??**) (iii) Win-stay / lose-shift (Subsection **??**) (iv) Model-free reinforcement learner (Subsection **??**). Throughout, our predictions consider long-run equilibrium behavior and abstract from short-term adjustments.

## 2.2 Bayesian learner

We assume the Bayesian's priors include the true DGP. Therefore, following standard results, the Bayesian's beliefs will converge to the true DGP.

In our analysis of the Dieter's Dilemma, we assume that the agent chooses actions through a logit choice function with precision parameter $\beta \geq 0$. Specifically,

$$P(A_{t+1} = 1|H_t) = \frac{\exp\left(\beta Eu(A_{t+1} = 1|H_t)\right)}{\sum_{a\in\{0,1\}} \exp\left(\beta Eu(A_{t+1} = a|H_t)\right)} \tag{1}$$

For simplicity, we impose risk-neutrality ($\forall x : u(x) = x$).

We assume the Bayesian considers the set of DAGs $\mathcal{D}$ consisting of all three-node DAGs with the action as an ancestral node. The Bayesian updates about both the causal structure and effect strengths. We assume the Bayesian starts from a uniform prior over all DAGs in $\mathcal{D}$. Regarding updating about DAG parameters, we use the standard Dirichlet-Multinomial formulation with the BDe prior (see **?**, for details). Given sufficient variation of the action in the data, the Baysian's beliefs converge to the true data-generating process as the number of rounds tends to infinity.

## 2.3 Misspecified model learner

The misspecified model learnier is a special case of the Baysian learner that puts full prior weight on a single DAG that is inconsistent with the DGP. All other elements are the same. We assume that the MML interprets the data through the DAG $A \to X \to Y$. Accordingly, the MML believes that $P(Y=1|A) = P(Y=1|X=1)P(X=1|A) + P(Y=1|X=0)P(X=0|A)$.[3]

If the DGP is given by $A \to X \leftarrow Y$ with $P(A=1) = \alpha$, and $P(Y=1) = \gamma$, and $P(X=1|A=a, Y=y) = \epsilon$ for $a = y = 0$ and $P(X=1|A=a, Y=y) = 1-\epsilon$ otherwise (noisy OR), the expected causal effect of selecting $A=1$ is given by

$$\hat{P}(Y=1|A=1,\alpha) = \frac{(1-\epsilon)^2\gamma}{\epsilon + (1-2\epsilon)\gamma + \alpha(1-2\epsilon)(1-\gamma)} + \frac{\epsilon^2\gamma}{1-\epsilon-(1-2\epsilon)\gamma - \alpha(1-2\epsilon)(1-\gamma)},$$

$$\hat{P}(Y=1|A=0,\alpha) = \frac{(1-\epsilon)\gamma\big[\epsilon + (1-2\epsilon)\gamma\big]}{\epsilon + (1-2\epsilon)\gamma + \alpha(1-2\epsilon)(1-\gamma)} + \frac{\epsilon\gamma\big[1-\epsilon-(1-2\epsilon)\gamma\big]}{1-\epsilon-(1-2\epsilon)\gamma - \alpha(1-2\epsilon)(1-\gamma)}.$$

An interior equilibrium exists if the equation $\hat{P}(Y=1|A=1,\alpha) - k - \hat{P}(Y=1|A=0,\alpha) = 0$ has a solution $\alpha^*$ strictly within the unit interval. In this case, the foregoing expression is positive if $\alpha \leq \alpha^*$ and negative otherwise. Accordingly, for the Misspecified Model Learner, $P(A=1)$ is higher for $\alpha \leq \alpha^*$ than for $\alpha \geq \alpha^*$

## 2.4 Model-free reinforcement learner

We specify our model-free reinforcement-learner as in **?**. Specifically, the agent has propensity $q_1(t)$ to play $A=1$ and $q_0(t)$ to play $A=0$ in period $t$. Propensities map into choice probabilities through the Luce function, as in equation (**??**) where the expected utility terms are replaced with the propensities.

In each period $t$, the agent updates propensities. The propensity for the action not taken remains unchanged. If the agent played $A=a$, then

$$q_a(t+1) = q_a(t) + R(x)$$

Here, $R(x) = x - \gamma$ where $x$ is the payoff obtained in period $t$ and $\gamma = P(Y=1)$ is the expected success rate.[4] Under the above assumption, the MFRL converges to $P(A=1) = 0$. The expected difference in propensities is not only positive, but diverges as $t \to \infty$, so the choice is not noisy.

[3] If the variables were Gaussian, this formula would simply say that the MML believes that the causal effect of $A$ on $Y$ is $\mathrm{cov}(A,X)\,\mathrm{cov}(X,Y)$.

[4] **?** start with the assumption that $R(x) = x - x_{min}$ but then argue that that assumption is problematic and argue for more general reference points. In our case, that formulation implies the implausible result that in case of exogenous data in which $P(A=1) = \alpha$, the learner will play $A=1$ for sure if $\alpha \geq \frac{c+\gamma}{c+2\gamma}$.

Let $n$ denote the normalizing factor in $R(x) = x - n$. Then, the reinforcements in each round are given as follows:

$$\begin{aligned} R(A=1,Y=1) &= 1-c-n \\ R(A=1,Y=0) &= -c-n \\ R(A=0,Y=1) &= 1-n \\ R(A=0,Y=0) &= 0-n \end{aligned}$$

Therefore, the expected attractions of action $A = a$ in period $t$, written $E\big[q_a(t)\big]$, are given by

$$\begin{aligned} E\big[q_1(t)\big] &= t\alpha\,(\gamma - c - n), \\ E\big[q_0(t)\big] &= t(1-\alpha)\,(\gamma - n), \end{aligned}$$

The difference in these terms, given by $E\big[q_0(t) - q_1(t)\big] = t\Big[(1-2\alpha)(\gamma - n) + \alpha c\Big]$, diverges to infinity if and only if $h(\alpha) := (\gamma - n) + \alpha\big(c - 2(\gamma - n)\big) \geq 0$. This condition holds for all $\alpha \in [0,1]$ if $n \in [\gamma - c, \gamma]$.[5]

## 2.5 Win-stay / lose-shift

Win-stay/lose-shift is a popular heuristic in game theory when there are two possible actions (as in our case). The agent picks an action, and stays with the action as long as the outcome is good. If a bad outcome arises, he agent switches to the other action.

We allow for noisy implementations of the heuristic. Formally, letting $\epsilon \in [0,1]$ denote the noise parameter, the agent's strategy is given by $P(A_{t+1} = 1 | A_t = a, Y_t = y) = a_{ay}$ where $a_{11} = a_{00} = 1 - \epsilon$ and $a_{01} = a_{10} = \epsilon$.

As we show in Appendix **??**, letting $\pi_1$ denote the probability with which the agent takes action 1, the ergodic action distribution is given by

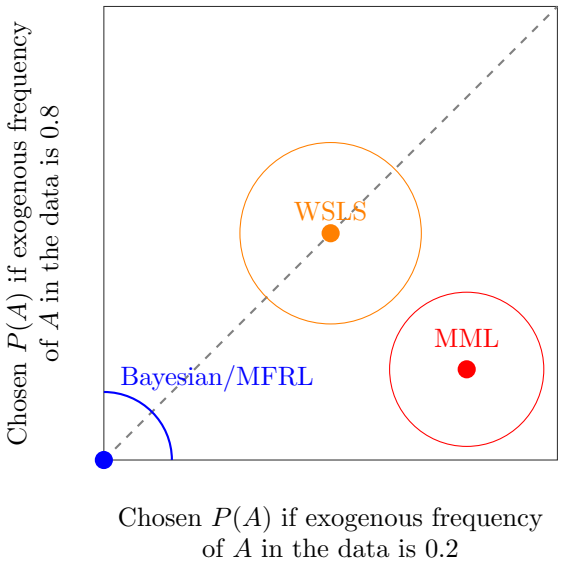
$$\frac{\pi_1}{1-\pi_1} = \frac{1 - P(Y=1|A=1)}{1 - P(Y=0|A=0)}$$

In any DGP with $A \perp\!\!\!\perp Y$, if $P(A=1) = \alpha$ and $P(Y=1) = \gamma$, then

$$\begin{aligned} E\big[P(A_{t+1}=1)\big] &= \alpha\gamma a_{11} + \alpha(1-\gamma)a_{10} + (1-\alpha)\gamma a_{01} + (1-\alpha)(1-\gamma)a_{00} \\ &= (1-\gamma) + \epsilon(2\gamma - 1) + \alpha(2\gamma - 1)(1 - 2\epsilon) \end{aligned}$$

[5] Proof: $h(0) \geq 0$ holds iff $n \leq \gamma$. $h(1) \geq 0$ holds iff $n \geq \gamma - c$. If $h(0) \geq 0$ and $h(1) \geq 0$, then linearity of $h$ implies $h(\alpha) \geq 0 \forall \alpha \in [0,1]$.

**Figure 1:** Assignment of types.




***Notes***: Subjects outside the acceptance regions are classified as 'other'.

Accordingly, when $\gamma = \frac{1}{2}$, the (noisy) WSLS type will choose $A_t = 1$ with probability 1/2 regardless of the exogenous action probability in the data, $\alpha$. If good outcomes are common ($\gamma > \frac{1}{2}$), higher rates of $\alpha$ increase the WSLS type's propensity to choose $A_t = 1$ as the increased coincidence of good outcomes with taking the action causes the agent to stay at taking the action more often. Likewise, if good outcomes are rare ($\gamma < \frac{1}{2}$)) the increased coincidence of bad outcomes with taking the action causes the agent to switch to not taking the action more often.

# 3 Identification

## 3.1 Estimating types

To derive predictions about behavior using exogenous variation in the action probability, we consider long-run equilibrium behavior. Specifically, we assume that the Bayesian and the Misspecified Model Learner have observed the true distribution of the variables. The Misspecified Model Learner fits his model to that distribution. The WSLS type is mechanical.

The key characteristic of **?**'s personal equilibrium is the way in which the misspecified model learner will adjust his probability of taking the costly action as a function of the correlations observed in the data. In the context of the Dieter's Dilemma with the noisy-OR link, the misspecified model learner will infer a strong causal effect of $A$ on $Y$ when $P(A)$ in the data is low, and a weak effect if $P(A)$ in the data is high.

**Figure 2:** Experimental interface.


***Notes***: At the beginning of a round, the second and third tile are blank. Once the subject chooses an action, the second tile is filled. Once the subject presses the 'c' key, the third tile appears. Pressing the space bar initiates the next round.

**Mapping choices to types** We obtain the best-response function for each subject. Each subject is characterized by a pair $(\alpha_{0.2}, \alpha_{0.8})$ which denotes the subject's action probability when the exogenous action probabilities in the data the subject observes are 0.2 and 0.8, respectively. We set $\gamma = \frac{1}{2}$. We assume that the logit noise parameter (equation **??**) is small enough such that $P(A|\alpha = 0.2) < 0.2$. Then the following hold,

(i) For the Misspecified Model learner, $\alpha_{0.2} > \alpha_{0.8}$. For all other types, $\alpha_{0.2} = \alpha_{0.8}$

(ii) For WSLS, $\alpha_{0.2} = 0.5$

(iii) For Bayes and MFRL, $\alpha_{0.2} = \alpha_{0.8} = 0$

We assign subjects to types using the following method. We exogenously select a critical $p$-value $p$. For each subject, and for each type, we conduct a Wald test of the joint hypothesis $\hat{\alpha}_{0.2} = \alpha^t_{0.2}$ and $\hat{\alpha}_{0.8} = \alpha^t_{0.8}$. If we reject the hypothesis for each type, the subject remains unassigned. If we accept the hypothesis for one or multiple types, we assign the subject to the type that is closest to

the subjects' choice. To measure that distance, we use the Wald statistic, given by

$$W(a,t) = \frac{n(a_1 - t_1)^2}{t_1(1-t_1)} + \frac{n(a_2 - t_2)^2}{t_2(1-t_2)}$$

That is, we assign the subject to the type $t$ for which $W(a,t)$ is smallest. (The critical value for $W(a,t)$ is $\chi^2_{2,1-p}/n$) See Appendix **??** for the derivation of the statistic.

# 4 Design structure

Our design has two parts: exogenous and endogenous data, which we detail in subsections **??** and **??**, respectively. The part using exogenous data lets us place subjects in broad type categories depending on how their choice probabilities vary with the observed action frequency in the data.

The part using endogenous data tests the key substantive predictions of the equilibrium models: observations affect inference affect actions affect observations.

The endogenous data part lets us test more fine-grained predictions than the exogenous-data part. With exogenous data, we test whether subjects exhibit the best-response dynamics that drive them towards an interior equilibrium, that is, whether an equilibrium exists. With endogenous data, we test whether the equilibrium exhibits the theoretically predicted comparative statics. Because these comparative statics are predicted only for subjects of certain types, we will study the comparative statics of equilibria for the subset of subjects classified as that type.

## 4.1 Exogenous data: Main Treatments

Our exogenous data treatments have dual goals. First, we seek to test whether subjects exhibit the comparative statics that will lead to interior or multiple equilibria (both are consistent with misspecified model learners but inconsistent with Bayesian or reinforcement learning). Second, we seek to use subjects' behavior to classify them into types, so we can conduct type-specific equilibrium analysis in the endogenous data treatments.

We employ the following three settings:

(i) Dieter's dilemma, noisy-OR, low noise

(ii) Dieter's dilemma, noisy-OR, high noise

(iii) Dieter's dilemma, noisy-AND

We run each of these twice, one in a condition in which the exogenous action probability is 0.2 and one with exogenous action probability 0.8.

### 4.2 Endogenous data: Treatments

Early rounds present subjects with an explore/exploit dilemma: subjects may intentionally take actions they believe do not maximize their current-round payoff because doing so provides information about the DGP that lets them take better actions in the future. These exploration incentives grow weaker the shorter the subject's horizon in the remainder of the game. Our types capture both learning from a given observed sequence of data and how beliefs myopically map into choices, but they do not consider exploration motives. We will assume that in the last 10 rounds, subjects no longer consider how their actions affect their future learning so that behavior in the last 10 rounds is myopic – that is, the exploration motive will have declined to negligible levels.[6]

We consider four main treatments in the Reverse Causality problem. They vary the nature of the link-function (noisy-OR or noisy-AND), the amount of noise in the link function (high or low). We consider two main treatments in the Confounded Choice problem that vary the amount of noise.

## 5 Pilot results

We ran a pilot with 59 subjects on prolific.com to assess whether our experiment can generate misperceptions about causal relations. We ran the Dieter's Dilemma in a single parametrization. The noise parameter for the noisy-OR function is $\epsilon = 0.05$. Subjects were paid for one of the 100 rounds, selected at random. They received \$2 if $Y = 1$ in that round and \$0 otherwise (in addition to a completion payment). Choosing $A = 1$ cost \$0.25, choosing $A = 0$ was free.

Figure **??** shows the key results. Panel A the CDF of subjects' responses to the question of how much the chance of $Y = 1$ changes if they choose $A = 0$ rather than $A = 1$. Half of the subjects perceive a causal relation, even though none is present in the DGP. To compare the results against the theoretical benchmark, note that the correlation that a subject with misspecified DAG $A \to X \to Y$ perceives depends on the action frequency she chose. At the mean action frequency $P(A) = 0.5$, given the parameters used in the experiment, the MML perceives an effect of approximately 25 percentage points. About 40% of subjects perceive an effect of around this magnitude or greater. While these beliefs are not incentivized, our incentivized belief elicitation shows similar results. Table **??** lists subjects' willingness to pay to change the value of a given variable from off to on. While the small sample prevents statistical significance, we generally observe positive WTP amounts, consistent with the hypothesis that subjects perceived causal relations. We also find that incentivized beliefs about the causal effects in the system are strongly related to the frequency with which subjects take the action (though unincentivized beliefs are not).

[6] The issue also arises, in principle, in the empirical literature on repeated games where subjects have an incentive to explore to learn about the opponent's strategy or type. That literature ignores exploration motives. (see, e.g. **?**)

Panel B considers the comparative statics predicted by **?**. It shows the CDF of action frequencies in the last 20 rounds of each game as a function of the DGP. Contrary to the predictions, increasing the noise that links outcome $Y$ and symptom $Y$ has no effect on mean action frequencies.

Panel C tests the best-response behavior. The empirical analog to Figure **??**, it plots each subject as a two-dimensional vector $(\hat{p}_{0.2}, \hat{p}_{0.8})$, where $p_q$ denotes the frequency with which the subject choose to take the action when given a choice after having observed data with exogenous action probability $q$. The best-response dynamics that undergird the **?** equilibrium require $\hat{p}_{0.2} > \hat{p}_{0.8}$. As the figure shows, very few subjects satisfy this condition. Instead, many subjects appear to take the action with a relatively constant probability no matter the frequency of the action in the observed data, as evident by the fact that many choices cluster around the diagonal. If that probability is low, such behavior is consistent with rational learning; if it is high, it is consistent with win-stay / lose-shift behavior.

The main sessions will include additional treatments and their analysis will include far more detailed statistics to draw reliable conclusions.

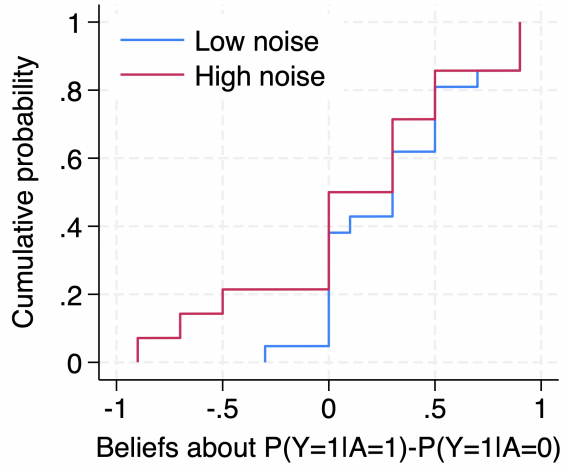
**Table 1:** Incentivized beliefs about causal effects

| | Mean | S.E. | $p$-value |
|---|---|---|---|
| Effect of $A$ on $X$ | | | |
| $P(X=1\|A=1;Y=0) - P(X=1\|A=0;Y=0)$ | 0.100 | 0.248 | 0.691 |
| $P(X=1\|A=1;Y=1) - P(X=1\|A=0;Y=1)$ | 0.357 | 0.218 | 0.113 |
| Direct effect of $A$ on $Y$ | | | |
| $P(Y=1\|A=1;X=0) - P(Y=1\|A=0;X=0)$ | -0.260 | 0.224 | 0.257 |
| $P(Y=1\|A=1;X=1) - P(Y=1\|A=0;X=1)$ | 0.135 | 0.212 | 0.532 |
| Effect of $X$ on $Y$ | | | |
| $P(Y=1\|A=0;X=1) - P(Y=1\|A=0;X=0)$ | 0.288 | 0.155 | 0.074 |
| $P(Y=1\|A=1;X=1) - P(Y=1\|A=1;X=0)$ | 0.188 | 0.186 | 0.322 |

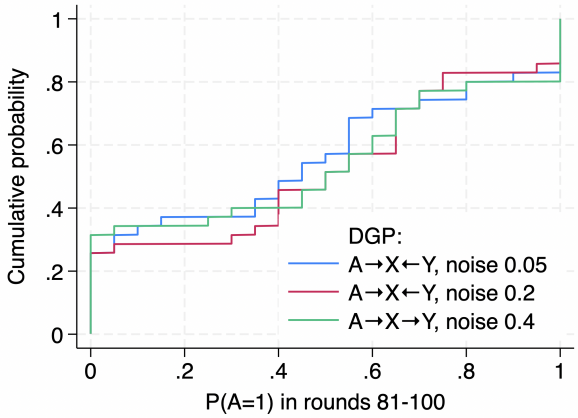
***Notes:*** Means show the WTP to switch the controlled variable from off to on.

**Figure 3:** Pilot results

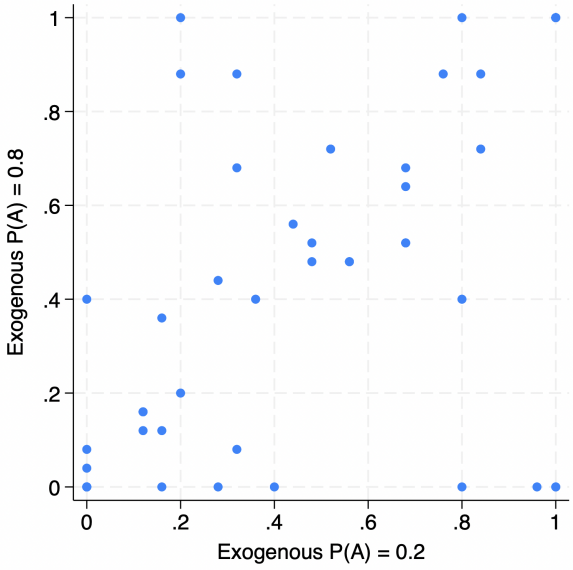
**A. Beliefs about the causal effect of $A$ on $Y$**




**B. Distribution of action frequencies by treatment in endogenous rounds**




**C. Distribution of action frequencies in exogenous rounds**



# References


**Aoyagi, Masaki, Guillaume R Fréchette, and Sevgi Yuksel**, "Beliefs in repeated games: An experiment," *American Economic Review*, 2024, *114* (12), 3944–3975.

**Bohren, J Aislinn and Daniel N Hauser**, "Misspecified Models in Learning and Games," *Annual Review of Economics*, 2024, *17*.

**Erev, Ido and Alvin E Roth**, "Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria," *American economic review*, 1998, pp. 848–881.

**Esponda, Ignacio and Demian Pouzo**, "Berk–Nash equilibrium: A framework for modeling agents with misspecified models," *Econometrica*, 2016, *84* (3), 1093–1130.

**_ , Emanuel Vespa, and Sevgi Yuksel**, "Mental models and learning: The case of base-rate neglect," *American Economic Review*, 2024, *114* (3), 752–782.

**Eyster, Erik and Matthew Rabin**, "Cursed equilibrium," *Econometrica*, 2005, *73* (5), 1623–1672.

**Griffiths, Thomas, Charles Kemp, and Joshua Tenenbaum**, "Bayesian models of cognition," *Unpublished*, 2008.

**Griffiths, Thomas L, Nick Chater, and Joshua B Tenenbaum**, *Bayesian models of cognition: reverse engineering the mind*, MIT Press, 2024.

**Han, Yi, David Huffman, and Yiming Liu**, "Minds, Models and Markets: How Managerial Cognition Affects Pricing Strategies," June 2025. Working paper.

**Jehiel, Philippe**, "Analogy-based expectation equilibrium," *Journal of Economic theory*, 2005, *123* (2), 81–104.

**_ and Frédéric Koessler**, "Revisiting games of incomplete information with analogy-based expectations," *Games and Economic Behavior*, 2008, *62* (2), 533–557.

**Kendall, Chad W and Constantin Charles**, "Causal Narratives," *Unpublished*, 2022.

**Koller, Daphne and Nir Friedman**, *Probabilistic graphical models: principles and techniques*, MIT press, 2009.

**Spiegler, Ran**, "Bayesian networks and boundedly rational expectations," *The Quarterly Journal of Economics*, 2016, *131* (3), 1243–1290.

**_** , "Behavioral implications of causal misperceptions," *Annual Review of Economics*, 2020, *12*, 81–106.

**Waldmann, Michael, ed.**, *The Oxford handbook of causal reasoning*, Oxford University Press, 2017.