

ORIGINAL PAPER

Complexity and contingent reasoning

Hyoeun Park 

Division of Social Science and SSEL, Department of Economics, Abu Dhabi, NYU, UAE
Email: hyoeun.econ@gmail.com

(Received 7 April 2023; revised 5 February 2025; accepted 12 March 2025)

Abstract

Using a laboratory experiment, we investigate complexity in decision problems as a cause of failures in contingent reasoning. For this purpose, we introduce three dimensions of complexity to a decision problem: the number of contingencies, the dominance property of choices, and reducible states. Each decision problem is designed to reflect variations in complexity across the three dimensions. Experimental results show that the number of contingencies has the most significant effect on failures in contingent reasoning. The second dimension, the dominance property of choices, also has a statistically significant effect, though the effect size is smaller than in the existing literature. In contrast, the third complexity dimension has no impact; presenting the decision problem in a reduced or reducible form does not change subjects' performance on contingent reasoning. Additionally, we examine the *Power of Certainty* and show its existence. This effect is particularly pronounced when the number of contingencies is large.

Keywords: Complexity; Contingent reasoning; Laboratory experiment; decision-making; mistakes

JEL Codes: D91; C91; D89

1. Introduction

Deviations from optimal behavior have long been studied in economics using various approaches. In the experimental literature, many studies have examined such deviations, especially when contingent reasoning is engaged. Another strand of literature studies how complexity creates a cognitive burden and affects optimal choices.

In this paper, we bridge complexity in decision problems and contingent reasoning. Specifically, we investigate complexity in decision problems as a cause of failures in contingent reasoning. For this purpose, we introduce three dimensions of complexity: the number of contingencies, the existence of dominant or obviously dominant choices, and reducible states.

We do this in a decision problem that we develop, which has a similar structure to the decision problem studied by Martínez-Marquina et al. (2019). Specifically, we consider a labor market setting where there are two workers who vary in two dimensions: The revenue they bring and their willingness to accept. With this information about workers (revenue and willingness to accept) but without knowing which worker they will hire, a subject chooses *one* wage for a job for the potential hire. Calculating the optimal wage requires contingent reasoning because the subject needs to consider which workers would accept any given wage.

For complexity, we consider three dimensions: the number of contingencies, the existence of reducible states, and the dominance property of choices. We borrow the first two concepts from

Oprea (2020). He considers the number of states as a measure of complexity, whereas we consider the number of contingencies as a measure of complexity. In our setting, states represent the uncertainty resolved by nature, while contingencies are also affected by the worker's actions. For example, for each wage choice, both which worker they are assigned and whether the assigned worker would accept the wage matters. For the second dimension of complexity, a decision problem has reducible states if there exists an equivalent representation of the problem involving fewer states. We say this equivalent problem then involves "reduced" states. Finally, we also consider the dominant/obviously dominant strategies concept suggested by Li (2017), but adapt to individual choice problems.

In addition, we test the *Power of Certainty* (henceforth *PoC*) – suggested by Martínez-Marquina et al. (2019) – by changing the baseline setting to one without uncertainty while keeping the number of contingencies and the existence of dominant/obvious dominance the same.¹

Our results show that the number of contingencies has the largest effect on the failures in contingent reasoning. The existence of dominant and obviously dominant choices also has an effect, but the degree of obvious dominance is smaller than the existing literature suggests. Redundant states do not affect the failures in contingent reasoning. We also examine how often subjects choose un-dominated choices and suggest that people may exhibit contingent reasoning partially, even if they fail to implement it fully. Lastly, the *PoC* exists in our setting, too, and has the largest effect in problems with a higher number of contingencies.

1.1. Related literature

Existing papers in the literature that study contingent reasoning have done so using decision problems that differ from our labor market decision problem. For example, Charness & Levin (2009) and Martínez-Marquina et al. (2019) use an acquiring-a-company game to test for failures in contingent reasoning. Although these decision problems are different from ours, the problem studied by Martínez-Marquina et al. (2019) shares some similarities with ours. Specifically, both decision problems require subjects to eliminate dominated choices first. Then, non-dominated choices involve contingencies that depend on whether each firm (or worker) will reject a given offer or not. However, their work has a winner's curse feature that ours does not have. Other decision problems that have been studied include voting games (Ali et al., 2021; Esponda & Vespa, 2014), dynamic public goods games (Calford & Cason, 2024), investment decisions with endogenous selection (Esponda & Vespa, 2018), the sure-thing principle (Esponda & Vespa, 2021; Shafir & Tversky, 1992), and the Monty Hall problem (Friedman, 1998).²

Ours is not the first paper to use the number of states as a measure of complexity. As mentioned earlier, Oprea (2020) considers the number of states as one measure of complexity in a typing task.³ He shows that rules that require more states are more complex and they incur higher costs. We apply this concept of complexity to the number of contingencies in decision problems. Similarly, in her experiment, Puri (2020) defines the complexity of a lottery by its number of possible outcomes.⁴ For reducibility as a complexity measure, Oprea (2020) finds that reducible states are more costly than reduced states, which implies reducible states make a rule more complex.⁵

¹The third case, reducible states, always requires uncertainty in our design. Also, the results with uncertainty show that there are no effects of reducible states, so we omit this complexity when testing the *PoC*.

²Several studies have been conducted using theoretical frameworks in relation to the experimental works on contingent reasoning. Echenique et al. (2022) provides one theoretical explanation for a failure in contingent reasoning by weakening the dominance axiom. This relates to our focus on dominance choices.

³Oprea (2020) does not directly deal with rational decision-making. Rather, they focus on complexity in general decision rules themselves and costs incurred in decision rules. As we apply the complexity concepts to rational decision-making, we interpret cost as possible loss due to deviations from the optimal choices.

⁴Theory allows for a more general measure of complexity.

⁵Though the paper shows that the mistake rates of the tasks of the reducible state are just slightly higher than reduced states tasks, those mistakes are not perfectly compatible with our decision-problem. The mistake rates in the paper are failures to

In other related work, Halevy (2007) studies the Ellsberg paradox and shows that a substantial portion of subjects violate the axiom of reduction of compound lotteries, which relates to our reducibility result. Other than Li (2017), there are several papers that consider dominance in a decision problem. Kagel et al. (1987) and Kagel and Levin (1993) show failures in contingent reasoning in an auction setting, where optimal strategies in each auction have different dominance notions. Like ours, their settings also separate dominance and obvious dominance. Rabin & Weizsäcker (2009) also study dominated choices. They show that some subjects choose a first-order stochastically dominated option, which is consistent with our observation that a substantial portion of subjects fail to make an optimal choice, even when a dominant choice exists.

There are other papers on complexity that consider dimensions different from ours. Banovetz and Oprea (2022) and Mononen (2022) consider complexity as (mental) costs, which may explain why subjects fail in contingent reasoning with certain domains of complexity more than others. (For further work on complexity, see Araujo & Piermont, 2023; Bossaerts & Murawski, 2017; Campbell, 1988; Carvalho & Silverman, 2019; Echenique et al., 2011). In addition, there are papers in ambiguity literature that explain ambiguity preference with complexity. For example, Kovářík et al. (2016) explain the Ellsberg paradox via complexity aversion, and Aydogan et al. (2023) show that complexity and ambiguous preferences explain heterogeneity in Ellsberg's urn setting.

The rest of the paper is organized as follows: Section 2 presents the experimental design. Section 3 shows the results, and Section 4 concludes the paper with a discussion.

2. Experimental design

2.1. Framework

We study a decision maker who faces wage decision problems under uncertainty with different dimensions of complexity. A decision-maker (henceforth DM) has to decide on one wage for a job when they do not know the exact quality of workers. In this section, we formalize the notions of complexities that we consider.

2.1.1. Baseline – one job problem

There is one available job, and the DM wants to hire a worker to maximize their profits. There are two candidate workers: Ann and Bob. Only one of Ann or Bob applies for the job, each with an equal chance. The DM has to decide *one* wage for the job before observing who applies. Ann and Bob have different willingnesses to accept (θ_A and θ_B) and bring different revenues (v_A and v_B). We assume Ann always brings higher revenue than Bob ($v_A > v_B$), Ann's willingness to accept is lower than that of Bob ($\theta_A < \theta_B$), and the revenue Ann brings is higher than Ann's willingness to accept ($v_A > \theta_A$). A worker will accept the job if and only if the wage the DM chooses is greater than or equal to their willingness to accept. Since Bob's willingness to accept is greater than Ann's, there is no wage that Ann rejects and Bob accepts. If the worker who applies rejects the wage, then the DM's payoff will be 0. In this problem, the set of states of the world is $\Omega = \{\mathbf{A}, \mathbf{B}\}$, where $\mathbf{A}$ represents the state where Ann applies for the job and $\mathbf{B}$ represents the other.

Niederle and Vespa (2023) describe the contingent reasoning process as thinking of all hypothetical states and computing payoffs for each action-state pair.⁶ We slightly modify this concept to define contingent reasoning. As a first step, we define the action-derived states as contingencies. Specifically, for any wage choice w , each state can be extended to be one where either a worker *Accepts* or *Rejects* (for simplicity, A and R , respectively). Thus, for the entire set of wage choices, $\mathcal{R}_+$, we can consider a Cartesian product of extended states $\{A, R\} \times \{A, R\} = \{(A, A), (A, R), (R, A), (R, R)\}$, where the

follow the commanded rules rather than rational decision-making. Thus, our results should be understood in relation to the cost-wise result.

⁶They focus on the acquiring-a-company problem. This problem has a similar structure to ours, which has two states of the world, a high value or low value, and a decision maker choosing one price.

first element of the pair represents Ann's decision and the second element represents Bob's decision.⁷ Each wage w will trigger exactly one of these extended states. We define each element of this Cartesian product as a contingency.

Next, for each of these contingencies, we call the wage that maximizes their utility in that contingency the *contingency-optimal* wage. Thus, there will be four contingency-optimal wages in our problem. We then define two levels of contingent reasoning: First, does the DM choose one of the contingency-optimal wages? Second, do they choose the utility-maximizing wage from that set?

In our baseline setting, we can simplify the set of contingencies. Since $\theta_A < \theta_B$, we can rule out (R, A) and consider three relevant contingencies for any wage w : (1) $w \geq \theta_B$, which triggers (A, A) , (2) $\theta_A \leq w < \theta_B$, which triggers (A, R) , and (3) $w < \theta_A$, which triggers (R, R) . Assuming expected utility, the DM evaluates each w according to

$$U(w) = \begin{cases} \frac{1}{2}u(v_A - w) + \frac{1}{2}u(v_B - w) & \text{if } w \geq \theta_B \\ \frac{1}{2}u(v_A - w) & \text{if } \theta_A \leq w < \theta_B \\ 0 & \text{if } w < \theta_A, \end{cases}$$

where $u(\cdot)$ is an increasing function with $u(0) = 0$. Since $u(\cdot)$ is an increasing function, U decreases in w . Thus, the contingency-optimal wage is the threshold value in each contingency: $w = \theta_B$, $w = \theta_A$, and $w \in [0, \theta_A)$ respectively. Our first question is whether the DM chooses one of these three contingency-optimal wages. Then, we ask if they choose optimally from that set. Since $v_A > \theta_A$, it is trivial that $w < \theta_A$ would give a lower utility than θ_A , so we can rule out $w \in [0, \theta_A)$. Thus, fully successful contingent reasoning requires choosing the wage that maximizes the DM's utility from $\{\theta_B, \theta_A\}$.

2.1.2. Various notions of dominance

As one dimension of complexity, we consider the dominance property of choices. Specifically, we introduce a modified concept of dominant and obviously dominant strategies, called dominant and obviously dominant choices. In addition to the well-known dominant strategies concept, Li (2017) suggests a concept called “obviously dominant strategies” and shows that obviously dominant strategies lead to rational choices more often than dominant strategies. We apply this notion, restricting choices first to those that are contingent-optimal.⁸

Specifically, let W^* be the set of contingency-optimal choices. By a slight abuse of notation, let u be a function of wage and state, (w, s) , in this subsection. We say that w^* is dominant (or a dominant choice) if $u(w^*, s) \geq u(w, s)$ for all $s \in \Omega$ and for all $w \in W^*$. We say that w^* is obviously dominant (or an obviously dominant choice) if $\min_{s \in \Omega} u(w^*, s) \geq \max_{s \in \Omega} u(w, s)$ for all $w \in W^*$. Note that any obviously dominant choice is also a dominant choice. For convenience, a “dominant choice” henceforth refers to a dominant but not obviously dominant choice. We regard a decision problem without a dominant choice to be more complex than one with a dominant choice, and a decision problem with a dominant choice is regarded as more complex than one with an obviously dominant choice.

In the baseline setting, if $v_B - \theta_B \leq 0 \leq v_A - \theta_B$, then θ_A is a dominant choice, as in Panel (b) of Figure 1. If $v_B - \theta_B \leq 0$, then θ_A is an obviously dominant choice (Panel (c) of Figure 1). When a decision problem does not have a dominant choice, the success in contingent reasoning is determined by the DM's risk preference. For example, as in Panel (a) of Figure 1, $w = \theta_A$ has a higher mean but also higher variance compared to $w = \theta_B$. In this case, the optimal choice depends on the DM's risk preference. How we elicit a risk preference in the experiment is explained in Subsection 2.2.

⁷The other way to think about this is that we can understand the action-derived states as the first-order perception of the payoff relevant states contingent on an action.

⁸Li (2017) uses “strategy” because the paper studies decision in game settings. In this paper, we use “choice” instead since we consider a one-person decision problem.

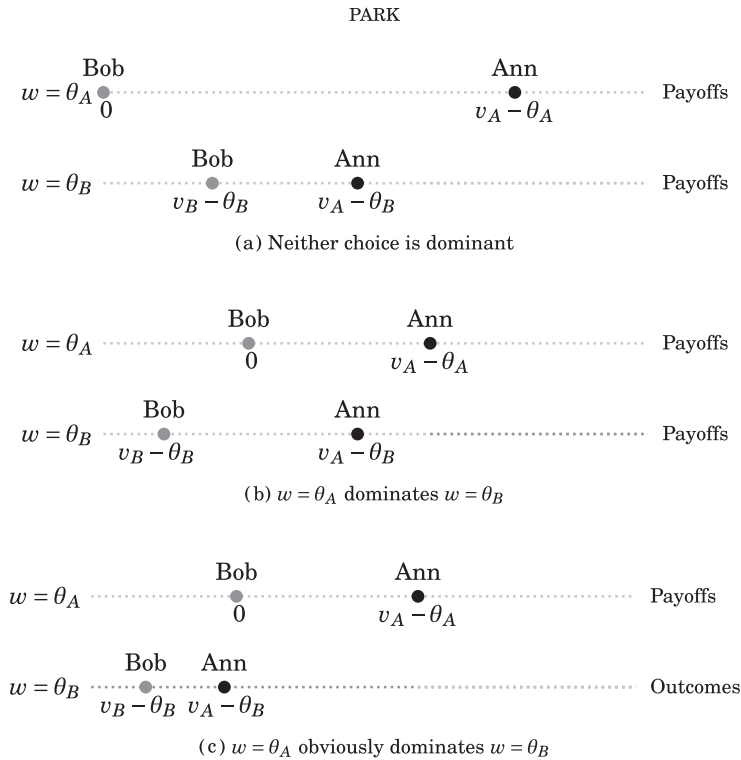


Fig. 1 Dominance property of choices and the payoffs given in each state when (a) neither choice is dominant, (b) $w = \theta_A$ dominates $w = \theta_B$, and (c) $w = \theta_A$ obviously dominates $w = \theta_B$.

2.1.3. Reducible problem

Oprea (2020) proposes several notions of complexity in a decision problem based on the automata literature. “Reducibility” is one such notion. If a decision problem can be reduced to an equivalent problem with fewer states, then the non-reduced (we call it reducible) problem is more complex than the reduced problem. In line with this measure of complexity, we construct a decision problem with a reducible state that can be reduced to the *one job* problem.⁹

Consider the following decision that corresponds to this complexity. Similar to the *one job* problem, suppose that there is only one available job position. In the *one job* problem, there is always only one applicant, either Ann or Bob. In our reducible problem, there are now three states: Ann applies (A), Bob applies (B), and both apply (AB). If the true state is that both apply for the job (AB), then a randomization device decides who will get the job. The DM has to decide the wage w before knowing which is the true state.

Let a probability that the randomization device assigns to Ann in state AB be r and let the probability of a state ω being realized be p_ω . If we assume that the DM satisfies the reduction of compound lotteries axiom then this decision problem is equivalent to a decision problem with $\Omega = \{A, B\}$ with probabilities $(p_A + p_{AB}r, p_B + p_{AB}(1 - r))$. If $p_A + p_{AB}r = p_B + p_{AB}(1 - r) = \frac{1}{2}$, then the problem can be reduced to the *one job* problem. We regard this *reducible* problem as more complex than the corresponding (reduced) *one job* problem.

⁹More precisely, Oprea (2020) refers to this notion of complexity/simplicity as “the lowest-state equivalent rule that delivers an identical sequence of actions over all possible sequence of events.” The concept “state” used in the paper will be described with more details in the following subsection.

Within the reducible problems, we can also create additional complexity due to the dominance property of choices (without dominant choices, with dominant choices, and with obviously dominant choices) by using the same parametrizations as the *one job* problems.

2.1.4. Two jobs problem: Increasing the number of contingencies

Suppose now that there are two jobs, job x and job y . There is only one available position for each job. There are two workers, Ann and Bob, who differ in their willingness to accept for each job (θ_{ij} , i is Ann or Bob, and j is job x or job y), and the revenue they bring for each job is (v_{ij}). We impose the parameter conditions that Bob's willingness to accept is higher than Ann's for both jobs ($\theta_{Bj} > \theta_{Aj}$, $j = x, y$), Ann brings more revenue than Bob for both jobs ($v_{Aj} > v_{Bj}$, $j = x, y$), both workers have a higher willingness to accept for job x ($\theta_{ix} > \theta_{iy}$, $i = \text{Ann, Bob}$), and Ann brings more revenue than her willingness to accept for both jobs ($v_{Aj} > \theta_{Aj}$, $j = x, y$).

With equal chances, (1) *Ann applies for job x and Bob applies for job y* , or (2) *Ann applies for job y and Bob applies for job x* . That means the states of the world are in $\Omega = \{(\mathbf{AB}), (\mathbf{BA})\}$ where first letter indicates which worker applies to job x . Under this uncertainty, the DM must decide one wage for each job, namely, w_x and w_y . Like in the *one job* case, a worker will accept a wage offer as long as it is greater than or equal to their willingness to accept.¹⁰ The payoff of the DM is the sum of their earnings from the two jobs.

Even if the number of states is the same as the *one job* case, the number of contingencies is different. After ruling out non-feasible contingencies, there are nine contingencies: $\{\text{Both accept job } x, \text{ Only Ann accepts job } x, \text{ Both rejects job } x\} \times \{\text{Both accept job } y, \text{ Only Ann accepts job } y, \text{ Both rejects job } y\}$.¹¹ Assuming an expected utility representation, we have the following form:

$$U(w) = \begin{cases} \frac{1}{2}u(v_{Ax} - w_x + v_{By} - w_y) + \frac{1}{2}u(v_{Bx} - w_x + v_{Ay} - w_y) & \text{if } w_x \geq \theta_{Bx}, w_y \geq \theta_{By} \\ \frac{1}{2}u(v_{Ax} - w_x) + \frac{1}{2}u(v_{Bx} - w_x + v_{Ay} - w_y) & \text{if } w_x \geq \theta_{Bx}, \theta_{Ay} \leq w_y < \theta_{By} \\ \frac{1}{2}u(v_{Ax} - w_x) + \frac{1}{2}u(v_{Bx} - w_x) & \text{if } w_x \geq \theta_{Bx}, w_y < \theta_{Ay} \\ \frac{1}{2}u(v_{Ax} - w_x + v_{By} - w_y) + \frac{1}{2}u(v_{Ay} - w_y) & \text{if } \theta_{Ax} \leq w_x < \theta_{Bx}, w_y \geq \theta_{By} \\ \frac{1}{2}u(v_{Ax} - w_x) + \frac{1}{2}u(v_{Ay} - w_y) & \text{if } \theta_{Ax} \leq w_x < \theta_{Bx}, \theta_{Ay} \leq w_y < \theta_{By} \\ \frac{1}{2}u(v_{Ax} - w_x) & \text{if } \theta_{Ax} \leq w_x < \theta_{Bx}, w_y < \theta_{Ay} \\ \frac{1}{2}u(v_{By} - w_y) + \frac{1}{2}u(v_{Ay} - w_y) & \text{if } w_x < \theta_{Ax}, w_y \geq \theta_{By} \\ \frac{1}{2}u(v_{Ay} - w_y) & \text{if } w_x < \theta_{Ax}, \theta_{Ay} \leq w_y < \theta_{By} \\ 0 & \text{if } w_x < \theta_{Ax}, w_y < \theta_{Ay} \end{cases}$$

By a similar logic as the *one job* case, we can rule out wages $w_j < \theta_{Aj}$ for any job j as they are dominated by $w = \theta_{Aj}$. Then the only possible optimal wages are either θ_{Aj} or θ_{Bj} , for $j \in \{x, y\}$. Thus, the contingency-optimal wage profiles (w_x, w_y) are given by $W^* = \{(\theta_{Ax}, \theta_{Ay}), (\theta_{Ax}, \theta_{By}), (\theta_{Bx}, \theta_{Ay}), (\theta_{Bx}, \theta_{By})\}$. In other words, after eliminating trivially dominated cases, the number of relevant contingencies is four.

Like in the *one job* problem, the range of parameters can be divided into three cases: (1) $0 \leq v_{Bj} - \theta_{Bj} \leq v_{Aj} - \theta_{Bj} \leq v_{Aj} - \theta_{Aj}$ for all $j \in \{x, y\}$, (2) $v_{Aj} - \theta_{Aj} \geq v_{Aj} - \theta_{Bj}$ for all $j \in \{x, y\}$, and (3) $\min\{v_{Aj} - \theta_{Aj}, 0\} \geq \max\{v_{Aj} - \theta_{Bj}, v_{Bj} - \theta_{Bj}\}$ for all $j \in \{x, y\}$. The first case does not

¹⁰Each worker's acceptance decision is not affected by the other's decision. That means hiring only one of the workers is possible, depending on wage choices. For example, suppose that Ann applies for job x and Bob applies for job y . If $w_x > \theta_{Ax}$ and $w_y < \theta_{By}$, then Ann will accept, but Bob will reject, and the DM will get a payoff only from Ann at job x .

¹¹Since $\theta_{Bj} > \theta_{Aj}$, $j = x, y$, it is not feasible for any job j that Ann rejects and Bob accepts.

have a dominant choice, thus an optimal choice depends on the DM's risk preference. Again, how we elicit a risk preference will be explained in Subsection 2.2. The second case is one with a dominant choice, and the third case is one with an obviously dominant choice. In these two cases, $(\theta_{Ax}, \theta_{Ay})$ is the optimal wage profile.

2.1.5. Power of certainty

Martínez-Marquina et al. (2019) show that if a decision problem is represented in two different ways, one in a probabilistic version and the other one in a deterministic version, then failures in contingent reasoning increase in the probabilistic version. They call the difference in failure rates due to the existence of uncertainty the *Power of Certainty*. Even though the *PoC* might not be directly related to complexity, we also test whether the *PoC* exists in our setting with a simple modification.

The *PoC* environment offers the deterministic version of the wage decision problem we have discussed thus far. The *PoC* environment is equivalent to the main model in terms of the contingency and (expected) outcomes but has no randomness in the states of the world. We implement this by having two jobs, one for Ann and one for Bob. A worker will accept a wage as long as it is greater than or equal to their willingness to accept. The DM chooses only *one* wage that is used for both positions. In other words, the DM does not choose different wages for Ann and Bob. They choose only one wage, and Ann and Bob decide whether to accept the offer independently. Thus, they can hire two workers, one worker, or no worker depending on the chosen wage w . If w is greater than or equal to a worker's willingness to accept, the worker accepts the offer. The DM is paid the sum of their earnings from both jobs. To have consistency in payoffs with the *one job* problem, we divide total profits by two. We refer to this as the *one job PoC* problem since it mimics the original *one job* problem but features deterministic problems.

Contingent reasoning requires choosing a wage from the set $\{\theta_A, \theta_B\}$. With an expected utility representation, we have

$$U(w) = \begin{cases} u(\frac{1}{2}(v_A - w) + \frac{1}{2}(v_B - w)) & \text{if } w \geq \theta_B \\ u(\frac{1}{2}(v_A - w)) & \text{if } \theta_A \leq w < \theta_B \\ 0 & \text{if } w < \theta_A \end{cases}$$

This is equivalent to the original *one job* problem not only in a contingency sense but also in a utility sense if the DM is risk-neutral.

The *two jobs PoC* problem is also modified in a similar way. There are two kinds of jobs, job x and job y , and there are two available positions for each job. Both Ann and Bob apply for both jobs, and they can work for both jobs if they accept both jobs. Each worker's willingness to accept and the revenue they bring are known to the DM. The DM submits *one wage* for each job, *i.e.*, one wage for job x , and one wage for job y . Depending on the wage chosen, the DM can hire two workers, one worker, or no workers for each job. For consistency, we again divide total profits by two. The contingency-optimal wages are $W^* = \{(\theta_{Ax}, \theta_{Ay}), (\theta_{Ax}, \theta_{By}), (\theta_{Bx}, \theta_{Ay}), (\theta_{Bx}, \theta_{By})\}$. Thus, the relevant contingencies and preferences are the same as the original problem if the DM is risk-neutral. We can derive the expected utility from similar arguments. Note that since there is no uncertainty, risk preferences do not affect optimal choices in either *PoC* treatment.

The *PoC* treatment cannot include reducible problems because certainty is a key feature of that design. Thus, we exclude reducible problems.

Table 1. Complexity and hypotheses

		H3			
		obviously dominant choice	dominant choice	no dominant choice	
H1	one job	1J-Obv	1J-Dom	1J-None	H2
	two jobs	2J-Obv	2J-Dom	2J-None	
	reducible	H4 Red-Obv	Red-Dom	Red-None	
	one job PoC	1J-Obv-PoC	1J-Dom-PoC	1J-None-PoC	
	two jobs PoC	2J-Obv-PoC	2J-Dom-PoC	2J-None-PoC	

2.2. Experimental implementation and hypotheses

2.2.1. Main treatment

Table 1 summarizes our design and hypotheses. The main treatment of the experiment employs a within-subjects design consisting of two parts: (1) wage choice problems and (2) lottery choice problems. Each subject experiences both. The first part of the experiment contains the labor market setting where a subject chooses a wage(s) for a job(s) as complexity in decision problems varies. In total, there are nine types of questions: {one job, two jobs, reducible} $\times$ {no dominant choice, dominant choice, obviously dominant choice}. Each question type has three questions with different parameters; thus, there is a total of 27 questions. That means each cell of **Table 1** has three questions. After the 27 wage decision problems, subjects have 27 lottery choice problems that are equivalent to the wage decision problems. The purpose of the lottery choice problems is to measure subjects' risk preferences. By doing so, we can identify subjects' optimal wage choices for the problems where there is no dominant choice in the equivalent wage choice problem.

From this structure, we say that a subject fails in contingent reasoning if they choose (1) any wage other than workers' minimum willingness to accept (for example, in the *one job* case, any wage other than θ_A and θ_B) or (2) a wage in the original problem that does not correspond to their chosen lottery in the lottery choice problem. With this concept of failures in contingent reasoning, we have the following three hypotheses.

Hypothesis 1. Failures in contingent reasoning increase in the number of contingencies.

The first hypothesis can be tested by comparing the first (*one job* problems) and second row (*two jobs* problems) results in **Table 1**. That means if a subject is more successful in the nine questions from the first row compared to the nine questions from the second row, we can argue that failures in contingent reasoning increase in the number of contingencies.

Hypothesis 2. Failures in contingent reasoning are more common in the reducible than in the equivalent reduced problem.

This hypothesis is also related to the reducibility argument in Oprea (2020). The questions in the third row (*reducible* problems) in **Table 1** can be further reduced to the first row (*one job (reduced)* problems). Thus, if subjects perform better in questions corresponding to the first row compared to the third row, then this hypothesis is supported.

Hypothesis 3. Failures in contingent reasoning decrease when a dominant choice exists. It further decreases when an obviously dominant choice exists.

This hypothesis follows from Li (2017) and can be tested by comparing frequencies of choosing optimal choices in the first (obviously dominant choice), second (dominant choice), and third (no dominant choice) columns of Table 1.

The second part of the experiment, lottery choices, is devised to identify optimal choices in problems without a dominant choice. As stated in the previous section, when there is no dominant choice, there is no objectively optimal wage from the set of contingency-optimal choices. Instead, the utility-maximizing choice depends on the DM's risk preferences.

Rather than assuming subjects are risk averse, we introduce the lottery choice problems that are induced by the contingency-optimal wages from part 1 and check whether choices are consistent between the two. That means we create lotteries using only $\{\theta_A, \theta_B\}$ or $\{(\theta_{Ax}, \theta_{Ay}), (\theta_{Ax}, \theta_{By}), (\theta_{Bx}, \theta_{Ay}), (\theta_{Bx}, \theta_{By})\}$. Let $\mathcal{L}(p_1, z_1; \dots; p_k, z_k)$ be a simple lottery that gives each z_j with probability p_j . Similarly, let $\mathcal{L}(p_1, \mathcal{L}_1; \dots; p_k, \mathcal{L}_k)$ be a compound lottery that gives each simple lottery $\mathcal{L}_j$ with probability p_j . Then, each lottery choice problem has one of the following three types:

- (1) A choice between $\mathcal{L}(\frac{1}{2}, v_A - \theta_A; \frac{1}{2}, 0)$ and $\mathcal{L}(\frac{1}{2}, v_A - \theta_B; \frac{1}{2}, v_B - \theta_B)$, for the *one job* problems, which corresponds to $w = \theta_A$ and $w = \theta_B$ respectively,
- (2) A choice between $\bar{\mathcal{L}}(\frac{1}{3}, \mathcal{L}_1(\frac{1}{2}, v_A - \theta_A; \frac{1}{2}, 0); \frac{1}{3}, \mathcal{L}_2(1, v_A - \theta_A; 0, 0); \frac{1}{3}, \mathcal{L}_3(0, v_A - \theta_A; 1, 0))$ and $\bar{\mathcal{L}}(\frac{1}{3}, \mathcal{L}_1(\frac{1}{2}, v_A - \theta_B; \frac{1}{2}, v_B - \theta_B); \frac{1}{3}, \mathcal{L}_2(1, v_A - \theta_B; 0, v_B - \theta_B); \frac{1}{3}, \mathcal{L}_3(0, v_A - \theta_B; 1, v_B - \theta_B))$ for the *reducible* problems, corresponding to $w = \theta_A$ and $w = \theta_B$ respectively,
- (3) A choice among $\mathcal{L}(\frac{1}{2}, v_{Ax} - \theta_{Ax}; \frac{1}{2}, v_{Ay} - \theta_{Ay})$, $\mathcal{L}(\frac{1}{2}, v_{Ax} - \theta_{Ax} + v_{By} - \theta_{By}; \frac{1}{2}, v_{Ay} - \theta_{By})$, $\mathcal{L}(\frac{1}{2}, v_{Ax} - \theta_{Ay}; \frac{1}{2}, v_{Ay} - \theta_{Ay} + v_{Bx} - \theta_{Bx})$, and $\mathcal{L}(\frac{1}{2}, v_{Ax} - \theta_{Bx} + v_{By} - \theta_{By}; \frac{1}{2}, v_{Ay} - \theta_{By} + v_{Bx} - \theta_{Bx})$ for the *two jobs* problems, corresponding to $(w_x, w_y) = (\theta_{Ax}, \theta_{Ay})$, $(w_x, w_y) = (\theta_{Ax}, \theta_{By})$, $(w_x, w_y) = (\theta_{Bx}, \theta_{Ay})$, and $(w_x, w_y) = (\theta_{Bx}, \theta_{By})$ respectively.

In problems without a dominant choice, the success of the contingent reasoning behavior is determined by the consistency between the wage decision problem and the lottery choice problem. If a subject chooses a wage other than one of the contingency-optimal wages, then regardless of their lottery choice, the subject is said to fail in contingent reasoning. If the subject chooses one of the contingent-optimal wages but chooses a lottery that is induced by a different wage in that set, then it is still a failure in contingent reasoning since the subject does not have a consistent preference. The subject succeeds in contingent reasoning only if they choose one of the contingency-optimal wages in the wage choice problem to accept and the equivalent lottery in the correspondent lottery choice problem.

Even though problems with a dominant choice and an obviously dominant choice have optimal solutions regardless of the DM's risk preference, we still ask subjects to choose between two/four lotteries to see whether they behave consistently.

2.2.2. Power of certainty treatment

For the *Power of Certainty* treatment, each subject has 18 questions presented from the first two rows (*one job* and *two jobs* problems) of Table 1 with a deterministic representation (*one job PoC* and *two jobs PoC* problems). For consistency, we present payoffs as $\frac{1}{2}(v_i - w)$ if a worker i accepts wage offer w . Since optimal wages are determined regardless of the DM's risk preference, subjects in this treatment do not have lottery choice problems. Following Martínez-Marquina et al. (2019), we hypothesize that the *PoC* treatment decreases failures in contingent reasoning.

Hypothesis 4. Failures in contingent reasoning decrease in the deterministic problems.

2.2.3. Procedure

At the start of the experiment, subjects were given instructions and comprehension check problems. Next, they faced 27 wage decision problems in random order and then 27 lottery choice problems in random order. In the *PoC* treatment, subjects instead answered 18 wage decision problems in random order, but were not given lottery choice problems.

In total, 162 Ohio State undergraduate students participated, 113 subjects for the main treatment and 49 subjects for the *Power of Certainty* treatment.¹² Subjects were recruited using ORSEE (Greiner, 2015). The participation fee was \$6 for the main treatment and \$3 for the *PoC* treatment. Subjects also received an additional payment from one randomly selected question. Each point in the selected question is converted to \$0.30. The experiment was programmed by oTree (Chen et al., 2016).

3. Results

This section shows the results of the experiment. There was no learning effect, so we include all decision problems for the analysis. Appendix A provides a graph for the test of the learning effect.

3.1. The number of contingencies

The questions corresponding to the first row (*one job* problems) and second row (*two jobs* problems) of Table 1 are different only in the number of contingencies. Each row has three questions for each dominance property of choices (without dominance, dominance, and obvious dominance). Thus, comparing the results from the nine questions in the first row and the nine questions in the second row allows for testing the effect of the number of contingencies (Hypothesis 1) on failures in contingent reasoning.

We graphically illustrate our results. Figure 2 illustrates the results for the number of contingencies. It shows the cumulative distribution functions (CDFs) of the number of questions answered incorrectly, either with wages that correspond to inconsistent preferences or wages outside of the contingency-optimal set. As the CDFs show, subjects submit incorrect answers more often in the *two jobs* problems. This implies that subjects fail more in contingent reasoning with a higher number of contingencies. The one-sided Wilcoxon signed-rank test confirms that the difference in two distributions is statistically significant with more failures in the *two jobs* problems (p -value < 0.0001). Also, the Kolmogorov-Smirnov test confirms that the entire distributions are different (p -value < 0.0001).

Our regression results further confirm that the number of contingencies affects the failures in contingent reasoning. For each class of question, if a subject submits correct choices for all nine questions, then they are defined as a *rational* type. For *one job* problems, 26.5% of subjects are classified as a *rational* type (p -value < 0.01). On the other hand, for *two jobs* problems, only 3.5% of subjects are classified as a *rational* type (p -value $= 0.27$). In addition, we conduct Fisher's exact test, and the test result is consistent, with p -values less than 0.01. Detailed tables are in Appendix A.

We, therefore, confirm that a larger number of contingencies increases the failure in contingent reasoning.

Result 1. Increasing the number of contingencies increases the failures in contingent reasoning.

¹² At first, we recruited 114 subjects, but one subject's data was missing. We first ran the main treatment and then decided to run the *PoC* treatment. Power calculation gives 46 subjects to replicate the results from Martínez-Marquina et al. (2019), with 113 subjects for the main treatment.

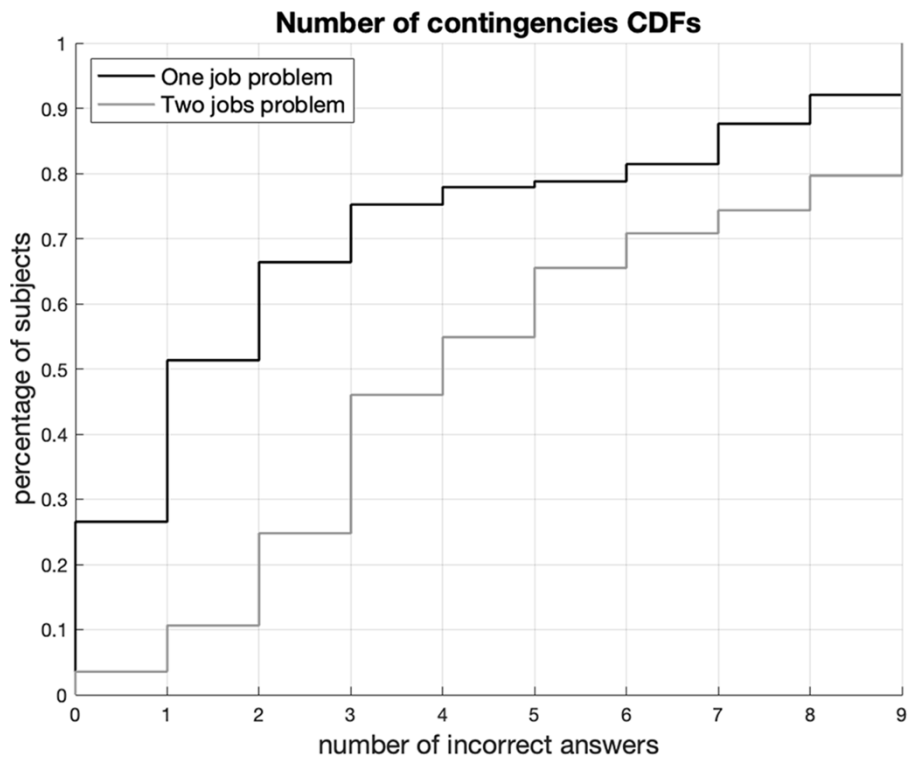


Fig. 2 CDFs of the number of incorrect answers

3.2. Reducibility

By the same logic applied for testing the effect of the number of contingencies, we compare the first row (*one job (reduced) problems*) and the third row (*reducible problems*) of Table 1 to test the effect of reducibility.

As the CDFs in Figure 3 depict, the two distributions are very close. Specifically, the one-sided and two-sided Wilcoxon-signed rank tests show that there is no statistically significant difference (p -value = 0.7174 and p -value = 0.7174, respectively). Also, the Kolmogorov-Smirnov test confirms that there is no difference in the shape of distributions (p -value = 0.7189). This suggests that reducibility does not increase the failures in contingent reasoning.

The regressions corroborate the result. In particular, the contrast p -value is 0.767, meaning there is no statistical difference between the two. Fisher’s exact test result also shows that the difference between reduced and reducible problems is insignificant. Detailed result tables for these tests are in Appendix A.

Result 2. There is no difference in the failures in contingent reasoning between the reducible problem and the equivalent reduced problem.

3.3. Various notions of dominance

The following result shows the effect of complexity due to the dominance property of choices: without a dominant choice, with a dominant choice, and with an obviously dominant choice. How this dimension of complexity affects a decision can be tested by comparing the three columns in Table 1.

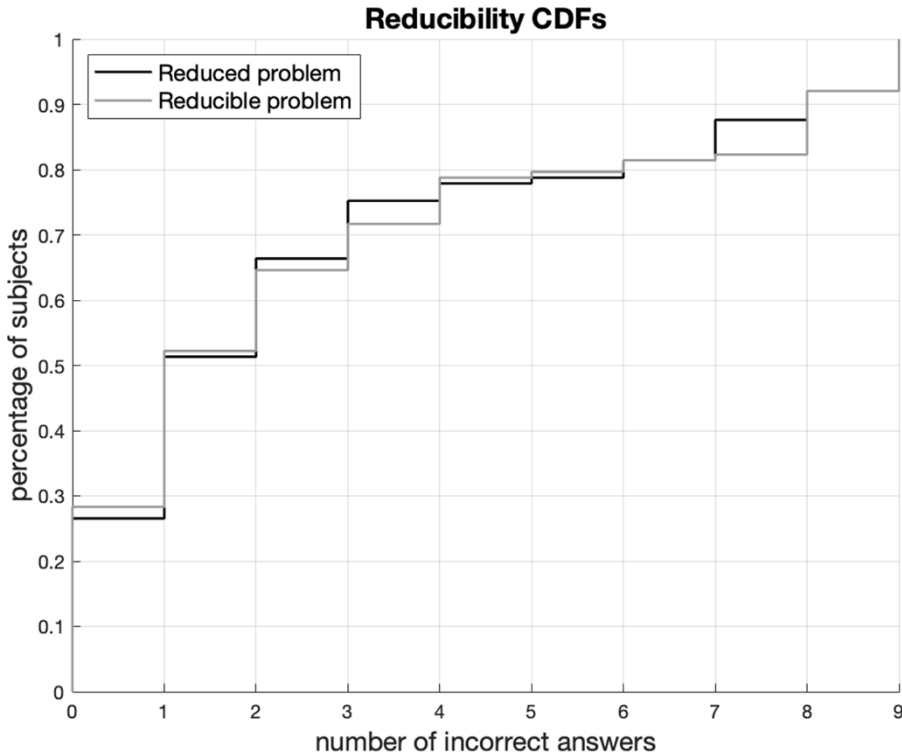


Fig. 3 CDFs of the number of incorrect answers

Figure 4 shows the distributions of the number of incorrect answers across dominance property of choices. As we hypothesized, subjects perform the worst without dominant choices. For example, the percentage of subjects who submit incorrect answers for more than five questions (over half of the total) is 23% for obviously dominant (Obv) choices and 28% for dominant (Dom) choices. However, this percentage increases drastically to 48% when no dominant choices are present (NoDom). Moreover, a difference across the three distributions is statistically significant (p -value < 0.001 with a Kruskal-Wallis test). For the Kolmogorov-Smirnov test, the obviously dominant choices and dominant choices are different at the marginal (p -value = 0.0513), and without dominant choices are significantly different (p -value < 0.001). Thus, the CDFs verify that the complexity due to the dominance property of choices affects the rate of failures in contingent reasoning, as we hypothesized.

Regression results confirm these conclusions. Subjects fail in contingent reasoning the most when a dominant choice does not exist (0.9% *rational* type with p -value = 0.82). The difference between a dominant choice (34.5% *rational* type with p -value < 0.01) and an obviously dominant choice (50.4% *rational* type with p -value < 0.01) is statistically significant, but the magnitude of difference is smaller than the difference between without dominant choice and dominant choice. The regression table and the Fisher's exact test result, which is consistent with all other results, are in Appendix A.

Result 3. Failures in contingent reasoning increase when there is no dominant choice. The improvement due to a dominant choice compared to an obviously dominant choice exists but is smaller.

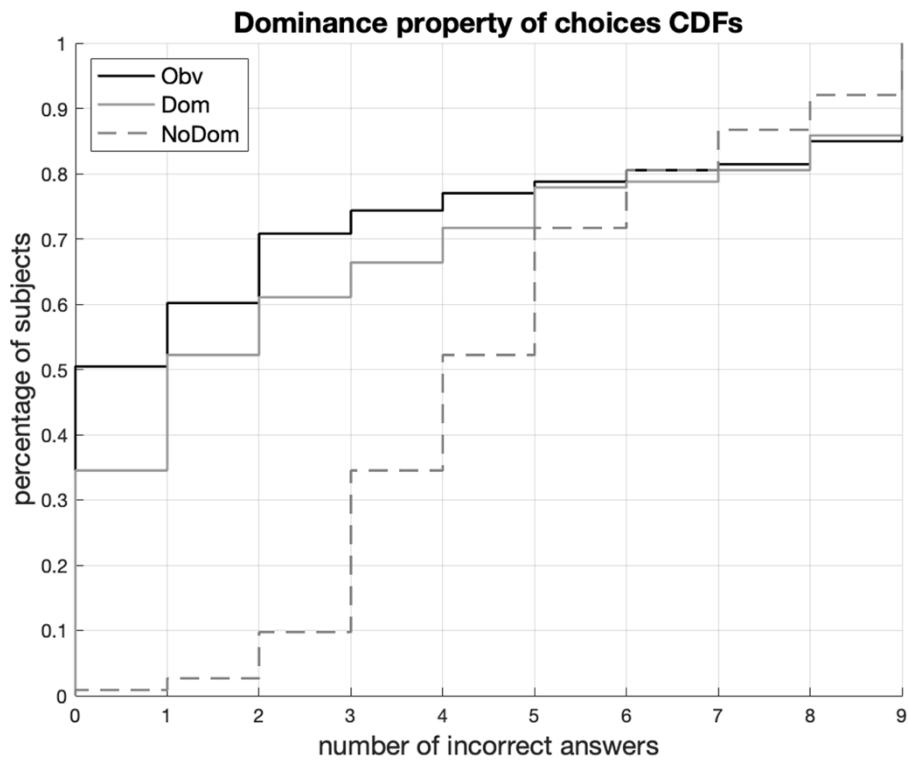


Fig. 4 CDFs of the number of incorrect answers

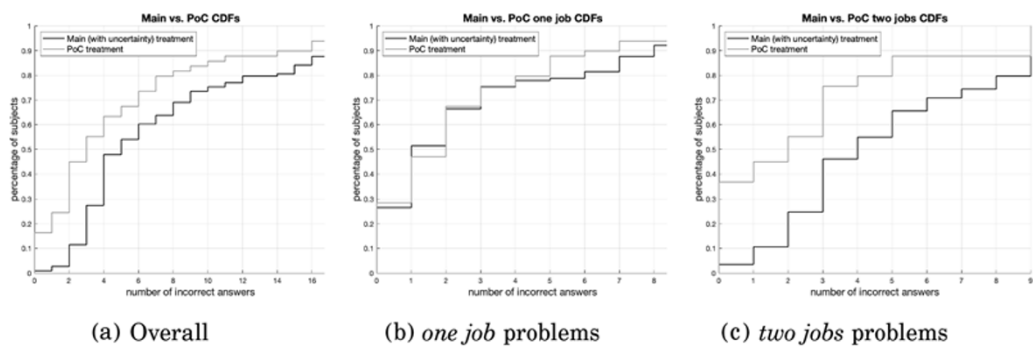


Fig. 5 PoC CDFs. a) Overall b) One job problems c) Two jobs problems

3.4. Power of certainty

We test whether the *PoC* persists in our setting. For consistency, we exclude reducible problems from the main treatment and compare the remaining 18 questions with the *PoC* treatment, which represents the same problems in a deterministic way. We ran a between-subject design for testing the *PoC*.

Figure 5a illustrates the difference in distributions between the main treatment and the *PoC* treatment. The difference is statistically significant (Wilcoxon rank sum test, p -value < 0.001 ; Kolmogorov-Smirnov test, p -value < 0.001) with subjects submitting incorrect answers more often in the main treatment. Thus, this shows the existence of the *PoC* in our setting.

Table 2. Inconsistent and consistent wages

	Inconsistent	Consistent	Abs. failure
One Job	27.43%	26.55%	46.02%
Two Jobs	30.97 %	3.54%	65.49%
Reducible	28.32%	28.32%	43.36%
NoDom	57.52%	0.89 %	41.19%
Dom	–	34.51%	65.49%
Obv	–	50.44%	49.56%

We further investigate the *PoC* by separating the *one job* and *two jobs* problems. [Figure 5b](#) and [Figure 5c](#) show *one job* and *two jobs* problem distributions, respectively. There is no difference between the two treatments when we restrict our attention to the *one job* problems (Wilcoxon rank sum test, p -value = 0.87 and Kolmogorov-Smirnov test, p -value = 0.5602). On the other hand, in the *two jobs* problems, subjects in the main treatment perform worse than those in the *PoC* treatment. Thus, the *PoC* is driven by the *two jobs* problems (p -value < 0.001 for both Wilcoxon rank sum test and Kolmogorov-Smirnov test). Regressions uphold the results as well. The regression tables are in Appendix A.

Result 4. The *Power of Certainty* exists but is only driven by the *two jobs* cases.

3.5. Contingency-optimal but inconsistent wages

One concern that can arise would be a distinction between contingent reasoning and consistency in preference: If a subject chooses one of the contingency-optimal wages that is not consistent with their lottery choice (we call this inconsistent wage from now), then it is unclear whether this is due to failure in contingent reasoning, inconsistency in preference, or both. Alternatively, such inconsistency could be due to a preference for randomization over lotteries (Agranov & Ortoleva, 2017; Agranov et al., 2023). Even though we cannot distinguish between these explanations, we can still look at frequency of subjects choosing contingency-optimal wages as a minimal measure of contingent reasoning.¹³

The first column of [Table 2](#) shows the percentage of subjects who always choose a contingency-optimal wage but choose an inconsistent wage at least once. These numbers suggest that substantial subjects can at least rule out wages that fall outside of the set of contingency-optimal wages. The second column is a reiteration of the previous sections and indicates the percentage of subjects who always choose a contingency-optimal and consistent wage. The third column is the absolute failure, representing the percentage of subjects who do not choose contingency-optimal wages at least once.

From [Table 2](#), we can see that there is a substantial portion of subjects who unquestionably fail in contingent reasoning. As for the effect of complexity, the result is ambiguous: The 65.49% who fail in dominant choices represent all possible failures of contingent thinking, while without dominant choices, the failures of contingent reasoning are spread between the “inconsistent” and “absolute failure” categories. Whether without dominant choices (the more complex treatment) has a higher rate of failures in contingent reasoning thus depends on what share of “inconsistent” subjects are, in fact, failing in contingent reasoning.

¹³Since the *PoC* problems do not have uncertainty, it does not engage any kind of problem regarding inconsistent choices. Thus, we focus only on the main treatment here.

4. Discussion

In this paper, we study how the complexity of a decision problem can affect failures in contingent reasoning. For the three dimensions of complexity we consider, we find that reducible states do not affect the failures in contingent reasoning, while the other two have an effect.

Specifically, we find that an increase in the number of contingencies increases the failures in contingent reasoning. We also find that dominance affects the failures in contingent reasoning. A dominant choice reduces the failures compared to when a dominant choice does not exist, and obviously dominant choices further reduce the failures. However, the difference between obviously dominant and dominant choices is not as significant as the existing literature in strategic settings suggests. For example, Kagel et al. (1987) show that in an auction setting, subjects barely deviate from equilibrium strategies when there is an obviously dominant strategy, but persistently deviate when an auction has a dominant but not obviously dominant strategy. Li (2017) also shows that subjects' performance is improved substantially with an obviously dominant strategy compared to a dominant strategy. The difference in degree between this paper and the existing literature suggests that even if an (obvious) dominant choice is a parallel concept to an (obviously) dominant strategy, the impact on rational decision-making can differ depending on the context.

Lastly, we show that the *PoC* exists but is only significant in the *two jobs* problems. The magnitude of the *PoC* in our paper is smaller than that found by Martínez-Marquina et al. (2019). The difference could be due to several reasons. One possibility is a population difference: OSU undergraduates and MTurk subjects might have different sophistication levels. It is possible that as the level of sophistication of the population increases, the *PoC* is more likely to be relevant for problems with a higher number of contingencies. Thus, there appear to be interesting interactions between sophistication and the number of contingencies. The difference in results could also be driven by other design factors, such as how risk preferences are measured or a difference in the structure of the decision problems.

In future studies, we can extend the experimental design to capture the different degrees of complexities. One possible direction is to test *s*-complexity aversion – suggested by Oprea (2020) – in a contingent reasoning environment by asking whether subjects are willing to bear a cost to avoid a larger number of contingencies. This would test whether failures in contingent reasoning are due to cognitive limitations or preferences when the number of contingencies vary. Also, we can dive deeper into reproducibility by changing the distribution in reducible problems. Our current experiment uses only uniform distributions. We could change the distribution to one that requires more calculations but is still equivalent to a reduced problem. If it requires significantly more calculation and the results change, then that would isolate a different level of complexity in the same family of problems. These directions would help further understand how complexity and failures in contingent reasoning are related.

Supplementary material. The supplementary material for this article can be found at <https://doi.org/10.1017/eec.2025.7>.

Replication packages. To obtain replication material for this article, <https://doi.org/10.17605/OSF.IO/QEMG4>.

Acknowledgements. The author thanks Paul J. Healy, John Rehbeck, Yaron Azrieli, Leeat Yariv, participants at the ESA Global Meetings (2022), FUR (2022), MEA 86th Annual Meetings, members of the OSU Theory/Experimental Reading Group, two anonymous referees, and the editor, for valuable comments.

Funding statement. Funding for this project was provided by the JMCB Small Research Grant at OSU. This project was determined exempt from IRB review under protocol 2021E0330.

References

- Agranov, M., Healy, P. J., & Nielsen, K. (2023). Stable randomisation. *The Economic Journal*, 133(655), 2553–2579.
- Agranov, M., & Ortoleva, P. (2017). Stochastic choice and preferences for randomization. *Journal of Political Economy*, 125(1), 40–68.

- Ali, S. N., Mihm, M., Siga, L., & Tergiman, C. (2021). Adverse and advantageous selection in the laboratory. *American Economic Review*, 111(7), 2152–2178.
- Araujo, F. A., & Piermont, E. (2023). Unawareness and risk taking: The role of context. *Journal of Economic Behavior & Organization*, 208, 61–79.
- Aydogan, I., Berger, L., & Th  roude, V. (2023). More ambiguous or more complex? An investigation of individual preferences under uncertainty.
- Banovetz, J., & Oprea, R. (2022). Complexity and procedural choice. *American Economic Journal: Microeconomics*, 15(2), 384–413.
- Bossaerts, P., & Murawski, C. (2017). Computational complexity and human decision-making. *Trends in Cognitive Sciences*, 21(12), 917–929.
- Calford, E. M., & Cason, T. N. (2024). Contingent reasoning and dynamic public goods provision. *American Economic Journal: Microeconomics*, 16(2), 236–266.
- Campbell, D. J. (1988). Task complexity: A review and analysis. *The Academy of Management Review*, 13(1), 40–52.
- Carvalho, L., Silverman, D. (2019). Complexity and sophistication. Working Paper 26036, National Bureau of Economic Research.
- Charness, G., Levin, D. (2009). The origin of the winner's curse: A laboratory study. *American Economic Journal: Microeconomics*, 1(1), 207–236.
- Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9, 88–97.
- Echenique, F., Golovin, D., & Wierman, A. (2011). Complexity and economics: Computational constraints may not matter empirically. *ACM SIGecom Exchanges*, 10(1), 2–5.
- Echenique, F., Miyashita, M., Nakamura, Y., Pomatto, L., & Vinson, J. (2022). Twofold multiprior preferences and failures of contingent reasoning. *Journal of Economic Theory*, 202, 105448.
- Esponda, I., & Vespa, E. (2014). Hypothetical thinking and information extraction in the laboratory. *American Economic Journal: Microeconomics*, 6(4), 180–202.
- Esponda, I., & Vespa, E. (2018). Endogenous sample selection: A laboratory study. *Quantitative Economics*, 9(1), 183–216.
- Esponda, I., & Vespa, E. (2021). Contingent thinking and the sure-thing principle: Revisiting classic anomalies in the laboratory 91(5), 2806–2831.
- Friedman, D. (1998). Monty hall's three doors: Construction and deconstruction of a choice anomaly. *American Economic Review*, 88(4), 933–46.
- Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with ORSEE. *Journal of the Economic Science Association*, 1(1), 114–125.
- Halevy, Y. (2007). Ellsberg revisited: An experimental study. *Econometrica*, 75(2), 503–536.
- Kagel, J. H., Harstad, R. M., & Levin, D. (1987). Information impact and allocation rules in auctions with affiliated private values: A laboratory study. *Econometrica*, 55(6), 1275–1304.
- Kagel, J. H., & Levin, D. (1993). Independent private value auctions: Bidder behaviour in first-, second- and third-price auctions with varying numbers of bidders. *Economic Journal*, 103(419), 868–879.
- Kov  r  k, J., Levin, D., & Wang, T. (2016). Ellsberg paradox: Ambiguity and complexity aversions compared. *Journal of Risk and Uncertainty*, 52, 47–64.
- Li, S. (2017). Obviously strategy-proof mechanisms. *American Economic Review*, 107(11), 3257–87.
- Mart  nez-Marquina, A., Niederle, M., & Vespa, E. (2019). Failures in contingent reasoning: The role of uncertainty. *American Economic Review*, 109(10), 3437–3474.
- Mononen, L., 2022. On preference for simplicity and probability weighting. Working paper.
- Niederle, M., & Vespa, E. (2023). Cognitive limitations: Failures of contingent thinking. *Annual Review of Economics*, 15, 307–328.
- Oprea, R. (2020). What makes a rule complex?. *American Economic Review*, 110(12), 3913–3251.
- Puri, I. (2020). Preference for simplicity, 1–48. Available at SSRN 3253494.
- Rabin, M., & Weizs  cker, G. (2009). Narrow bracketing and dominated choices. *The American Economic Review*, 99(4), 1508–1543.
- Shafir, E., & Tversky, A. (1992). Thinking through uncertainty: Nonconsequential reasoning and choice. *Cognitive Psychology*, 24(4), 449–474.