A Bridge from Monty Hall to the Hot Hand: The Principle of Restricted Choice

Joshua B. Miller and Adam Sanjurjo

Suppose that you work in a restaurant where two regular customers, Ann and Bob, are equally likely to come in for a meal. Further, you know that Ann is indifferent among the 10 items on the menu, whereas Bob strictly prefers the hamburger. While in the kitchen, you receive an order for a hamburger. Who is more likely to be the customer: Ann or Bob?

One intuition is that we have learned nothing from the observation that a hamburger was ordered, as it does not rule out either Ann or Bob, so they must remain equally likely to be the customer. However, this intuition is wrong, as it fails to account for *how* Ann and Bob choose items from the menu. By contrast, once we do account for how they choose, then the correct intuition emerges right away: because ordering a hamburger is more consistent with Bob (who must order it) than with Ann (who may order it), the order is more likely to have been placed by Bob.

While it may be easy to resist the incorrect intuition when confronting this simple problem, doing so is not so straightforward once the way that choices are made becomes even slightly less transparent. Let us briefly consider two examples: the Monty Hall problem and the presumed debunking of the "hot hand" phenomenon.

■ *Joshua B. Miller is Associate Professor of Economics, University of Melbourne, Melbourne, Australia. Adam Sanjurjo is Associate Professor of Economics, University of Alicante, Spain. Both authors contributed equally, with names listed in alphabetical order. Their email addresses are Joshua.Benjamin.Miller@gmail.com and sanjurjo@ua.es.*

The Monty Hall problem is a probability puzzle known for its ability to confound the intuitions of both the layperson and the mathematically sophisticated. A standard version of the problem, taken from vos Savant (1990), is as follows:

> *Monty Hall problem: Suppose you're on a game show, and you're given the choice of three doors. Behind one door is a car, behind the others, goats. You pick a door, say #1, and the host, who knows what's behind the doors, opens another door, say #3, which has a goat. He says to you, "Do you want to pick door #2?" Is it to your advantage to switch your choice of doors?*

While the intuitively appealing answer is that either of the two remaining doors leads to the same chances of winning the car, the chances actually increase if the contestant switches from door #1 to door #2 (under natural assumptions that we discuss later). People typically get this problem wrong. For example, a robust finding in laboratory experiments is that roughly 80–90 percent of subjects incorrectly stay with the same door, rather than switch (for example, Friedman 1998). Further, even a number of mathematically inclined academics (including Paul Erdős) have expressed disbelief when told the correct answer (Vazsonyi 1999).[1]

The hot hand fallacy refers to people's tendency to believe that success breeds success, even when it does not. In the seminal study by Gilovich, Vallone, and Tversky (1985), the authors found that basketball players shoot no better after having just made several shots in a row, despite a near-unanimous belief reported by players, coaches, and fans that players shoot better in these situations. When confronted with the scientific evidence against their beliefs, even professional players and coaches were left unpersuaded, leading the hot hand to become known as a "massive and widespread cognitive illusion" (Kahneman 2011).[2]

However, with the recent discovery of a surprising statistical bias (Miller and Sanjurjo 2018), it appears that the basketball community may have been right all along. In particular, to estimate a player's probability of making a shot, conditional on having made several in a row, Gilovich, Vallone, and Tversky (1985) and subsequent studies (1) selected the shot attempts that immediately followed a streak of several made shots (for example, three) and then (2) calculated the player's shooting percentage on these shots. As discussed below, this procedure biases the researcher toward overselecting missed shots, which leads to an underestimate of the player's probability of success on these shots. Not only is this *streak selection bias* large enough to invalidate the conclusions of previous studies, but it masks significant evidence of substantial hot hand shooting in their data.

---

[1] Math puzzles of this sort have been noted for their importance in stimulating research ideas and illustrating principles from microeconomic theory (Friedman 1998; Kluger and Wyatt 2004; Fehr and Tyran 2005). The Monty Hall problem, in particular, has been studied extensively, including in the first issue of this journal (Nalebuff 1987). For more discussion, see Rosenhouse (2009) and the references therein.

[2] The hot hand fallacy has been offered as a candidate explanation for certain puzzles and anomalies in financial markets, sports wagering, casino gambling, and lotteries. See Benjamin (2018), Miller and Sanjurjo (2018), Rabin and Vayanos (2010), and the references therein.

While it may not appear that there is any connection between why people have difficulty understanding the Monty Hall problem and why researchers had long overlooked the bias in common measures of the hot hand, we show that the two are in fact intimately related. The first step in understanding the relation is to observe that both environments involve a procedure that selects an observation for analysis on the basis of the outcomes of other observations in the same dataset. In particular, just as Monty offers the contestant an opportunity to switch to another door, knowing that a goat is behind the door he just opened, the hot hand researcher selects a shot from a longer sequence of basketball shots, knowing that the previous several shots were made. The key step to connecting these two environments, and many others, is then to illuminate the information that is revealed by their respective selection procedures.

The tool that we use to draw out these connections is the *principle of restricted choice*, an inferential rule drawn from the card game contract bridge that makes clear the information revealed by the optimizing behavior of a constrained opponent. The principle's simple intuition is illustrated above in the opening example with Ann and Bob, where Bob is more restricted to choose the hamburger than Ann is, because while Ann might order the hamburger, Bob must. In the next section, we show that restricted choice is naturally quantified as the updating factor from the odds formulation of Bayes' rule. To illustrate how intuitive and general restricted choice thinking is, we apply it to a number of settings. First, we use it to solve several classic probability paradoxes, including the Monty Hall problem.[3] This exercise makes clear that restricted choice renders intuitive the typically difficult counterfactual (and hypothetical) reasoning that is inherent in Bayesian updating. By contrast, we describe how some commonly used heuristic approaches, while helpful for particular problems, can lead to mistakes when applied more generally. We also use the principle to solve a progression of novel coin-flip probability puzzles, and to make comparisons across puzzles. For example, we show that one of our coin-flip puzzles captures the essence of the hot hand selection bias and at the same time is virtually equivalent to the Monty Hall problem.

Lastly, we consider various empirical examples in which restricted choice thinking can help researchers become aware of (and avoid) the types of counterintuitive mistakes and biases that can arise when particular observations are selected for analysis on the basis of the outcomes of other observations in the same dataset. Our four examples include (1) a bias that arises in measures of dependence across time, illustrated with the hot hand literature; (2) a bias that arises in measures of dependence across space, illustrated with Schelling's (1971) well-known work on segregation; (3) an unexpected correlation known as Berkson's paradox, illustrated with the canonical case of two unrelated diseases that happen to be negatively correlated in the hospitalized population despite being uncorrelated in the general

---

[3]Reese (1960, p. 29) illustrates the principle of restricted choice with a problem nearly identical to the Monty Hall problem. Gillman (1992) appears to be the first to use the restricted choice principle to explain the intuition behind the Monty Hall problem.

population; and (4) a hypothetical example of ESP research gone wrong. These examples are chosen to illustrate some pitfalls that researchers can avoid by using restricted choice thinking.

The Principle of Restricted Choice

The principle of restricted choice was first introduced in the context of the card game contract bridge, to account for the information revealed by the actions of an agent with a known decision rule. Legendary bridge player Terence Reese succinctly illustrates the principle in *Master Play in Contract Bridge* (Reese 1960, p. 26): "Since East could have played either card indifferently from K–Q, the fact that he has played one affords an indication that he does not hold the other."[4] Another illustration, which requires no familiarity with card games, is provided in our Ann and Bob example from the beginning of this paper. To reiterate, Bob is more restricted to choose the hamburger than Ann is, because while Ann may order the hamburger, Bob must. As a result, once we find out that the customer ordered a hamburger, we should shift our beliefs toward the customer being Bob rather than Ann.

The principle of restricted choice provides an informal *intuition* for why beliefs should shift in a particular direction upon the arrival of new information and calls to mind the essential qualitative feature of Bayesian updating. Namely, Bayes' rule requires that the odds in favor of a proposition increase upon the arrival of information that is more likely in the case that the proposition is true, or conversely, that the odds in favor of a proposition decrease upon the arrival of information that is less likely in the case that the proposition is true.

From here on, we represent uncertainty with odds rather than probabilities, as this simplifies the reasoning in the types of problems we discuss. For example, a proposition with a 3/5 probability of being true has 3/5 "chances" in its favor for every 2/5 chances against. Given this, the odds in favor of the proposition can be written as 3/5:2/5, or equivalently as 3/2:1 (by dividing each term by 2/5, as odds are invariant to proportional scaling). In turn, the odds of 3/2:1 can be stated simply as the single number 3/2, taking as given that the chances against the proposition are 1. Of course, associated probabilities can be easily recovered from the odds; for example, a proposition with 3:2 odds in its favor has 3 chances in its favor out of $3 + 2 = 5$ total chances—or a probability of 3/5.

To see how restricted choice can be understood as Bayesian updating, let A ("Ann") and B ("Bob") represent the two hypothetical propositions (or models) that could have produced the observed outcome c ("hamburger") in the restaurant

---

[4] Reese (1960, chap. 3, p. 26) credits Alan Truscott, who wrote the daily bridge column for the *New York Times* from 1964 to 2005, for introducing restricted choice to the bridge community in the 1950s. Prior to that, Borel and Chéron (1940) use the concept, at least implicitly, by applying Bayes' rule to calculate probabilities in bridge problems.

example. Then, given the prior odds, which consist of the chances in favor of *B* (relative to the chance in favor of *A*), Bayes' rule gives the posterior odds in favor of *B* (relative to one chance in favor of *A*):

Posterior odds in favor of *B* = Likelihood ratio × Prior odds in favor of *B*.

The likelihood ratio, also known as the Bayes factor, represents the multiplicative factor by which the number of chances in favor of *B* increase, decrease, or stay the same upon observation of *c*.[5] For our purposes, it can be thought of as *B*'s restrictedness relative to *A*'s, that is, the degree to which *B* is more likely to produce outcome *c* than is *A*. The principle of restricted choice tells us that, upon observation of an outcome, the odds shift in the direction of the model that is more likely ("restricted") to produce that outcome.

To illustrate, in the Ann and Bob restaurant example, the prior odds in favor of Bob being the customer (relative to Ann) are 1:1. However, once a hamburger has been ordered, because Bob is more likely to order the hamburger than Ann is, the odds in favor of Bob must increase. In particular, if Ann is equally likely to order each of the 10 items, then because Bob orders the hamburger for sure, he is 10 times more restricted to choose the hamburger than Ann. Therefore, the odds in favor of the customer being Bob increase by a factor of 10 upon learning that the customer ordered a hamburger. Thus, the posterior odds in favor of Bob are 10:1. Finally, if we assume for simplicity that Ann and Bob are the only possible customers, then because there are 10 chances in favor of Bob for every 1 chance in favor of Ann, the probability that the hamburger order came from Bob is 10/11.

Restricted Choice in Some Classic Conditional Probability "Paradoxes"

We show how the simplicity and intuition of restricted choice reasoning extend to several related classic conditional probability puzzles that often tend to confound people's intuition. We start with Bertrand's box paradox (Bertrand 1889; Gorrochurn 2012; presentation below adapted from Rosenhouse 2009), then present two versions of the boy-or-girl paradox, and finally return to the Monty Hall problem.

> *Bertrand's box paradox: Three boxes are identical in external appearance. The first box contains two gold coins, the second two silver coins, and the third one gold coin and one silver coin. You choose a box at random and draw a coin. Suppose that you draw a gold coin. What is the probability that the other coin is also gold?*

---

[5]More formally, posterior odds satisfy $(\mathcal{R}_A^B(c) \times$ Prior chances in favor of *B*) : (Prior chances in favor of *A*), where $\mathcal{R}_A^B(c)$ is the likelihood ratio, or Bayes updating factor. The likelihood ratio is defined as the ratio of the probability of *c* conditional on *B* to its probability conditional on *A*, that is, $\mathcal{R}_A^B(c) = \dfrac{\Pr(c \,|\, B)}{\Pr(c \,|\, A)}$ (assuming $\Pr(c \,|\, A) > 0$). In the extreme case that $\Pr(c \,|\, A) = 0$, the odds in favor of *B* are 1:0 (assuming $\Pr(c \,|\, B) > 0$).

Given that a gold coin was drawn, it is impossible that the all-silver box was chosen. Thus, two possible boxes remain: all gold and mixed. Given this, it becomes intuitively appealing to conclude that the probability that the other coin is gold is $1/2$. However, what this reasoning misses is that a draw that occurs from the all-gold box is more restricted to "choose" (draw) a gold coin, because with the gold box one *must* draw a gold coin, whereas with the mixed box one can draw either a gold or a silver coin. In this case, one is twice as restricted to choose the gold coin from the all-gold box relative to the mixed box. Therefore, by the principle of restricted choice, the updated odds in favor of the draw having come from the all-gold box are 2:1, double the prior odds of 1:1. This implies that the probability that the other coin is also gold is equal not to $1/2$ but rather to $2/3$.

Next, we consider the boy-or-girl paradox (as presented in problem 1 of Bar-Hillel and Falk 1982; see also Gardner 1961):

> *Boy-or-girl paradox: Mr. Smith is a father of two. We meet him walking along the street with a young boy whom he proudly introduces as his son. What is the probability that Mr. Smith's other child is also a boy?*

The intuitive answer to this problem is $1/2$, and under usual assumptions, this answer is correct—but for reasons that differ from the intuition many people bring to the problem. Let us assume that Mr. Smith chooses his walking companion at random from among his two children (without discriminating). With this, the problem becomes close to Bertrand's box paradox: Mr. Smith's children are drawn, without replacement, from either an all-boy "box," an all-girl box, or a mixed-gender box. The key difference, however, is that the types of boxes are not all equally likely. In particular, the equivalent of the mixed box—one boy and one girl—has 2:1 prior odds in its favor, relative to any single-gender box, because there are two birth order possibilities in the mixed-gender box (boy–girl and girl–boy). Analogous to Bertrand's box paradox, learning that the randomly chosen walking companion is a boy makes the posterior odds in favor of both children being boys (relative to mixed gender) double the prior odds, because the choice of a boy is twice as restricted in the all-boys case. Thus, because the prior odds were 1:2 "in favor" of all boys (relative to mixed gender), or $1/2{:}1$, the posterior odds are 1:1 in favor of all boys. Finally, because there remain only two possible compositions of children—all boys or mixed gender—the probability that Mr. Smith has all boys is $1/2$. As a result, the probability that his other child is a boy is $1/2$.[6]

---

[6]Another common version of the boy-or-girl paradox is as follows: "Mr. Smith says: 'I have two children and at least one of them is a boy.' Given this information, what is the probability that the other child is a boy?" (Fox and Levav 2004, p. 631). If one assumes that Mr. Smith would say nothing (or its equivalent) in the case that he were to have two girls, then in this version of the problem, Mr. Smith is equally restricted to report "boy" in the cases of boy–girl, girl–boy, and boy–boy, so prior and posterior odds are identical. As a result, the correct probabilities can be computed simply by an enumeration of the sample space and elimination of the impossible girl–girl combination. Therefore, failure to see the correct answer in this version of the boy-or-girl paradox can arise not because of a failure to incorporate

Now consider another version of the boy-or-girl paradox (equivalent to problem 2 in Bar-Hillel and Falk 1982, with slightly adapted language):

Younger boy-or-girl paradox: Mr. Smith is a father of two. We meet him walking along the street with a boy whom he proudly introduces as his eldest child. What is the probability that Mr. Smith's younger child is also a boy?

Because the younger child must be either a boy or a girl, the intuitively appealing response to this question is, again, 1/2. This response is correct if we assume that Mr. Smith chooses his walking companion at random between his two children, regardless of gender, as in the basic boy-or-girl paradox.

But while gender neutrality is a natural assumption, another possibility is that Mr. Smith has the unfortunate attitude of being willing to walk only with sons. Assume that this is so, but that if he has two boys, then he is indifferent between walking companions and chooses one of the boys at random. Under these assumptions, observing the gender of Mr. Smith's walking companion yields redundant information. That is, if we had merely observed Mr. Smith walking with a child, without any further information, we would already know that the child must be a boy, and that the possible birth order combinations are thus boy–girl, girl–boy, and boy–boy.

However, in the current problem, we additionally discover that Mr. Smith's walking companion is his eldest child, which eliminates the possibility of boy–girl, reducing the possibilities to girl–boy and boy–boy. With this, the intuitive response is again 1/2, but now this response is wrong. The reason why is that it fails to take into account that the degree of restrictedness in Mr. Smith's choice varies across these hypothetical birth orders. In particular, if the younger child is a girl (girl–boy), then Mr. Smith's choice of walking partner will be the older boy for sure. On the other hand, if the younger child is also a boy (boy–boy), then Mr. Smith is equally likely to choose each boy. This means that when the younger child is a girl, Mr. Smith's choice is twice as restricted. Therefore, the posterior (relative) odds in favor of girl–boy are double the prior odds of 1:1. Thus, the probability of girl–boy is 2/3. As a result, the probability of boy–boy is 1/3—that is, there is a 1/3 chance that the younger child is a boy.

We now return to the Monty Hall problem, which is essentially identical to the younger boy-or-girl paradox just discussed, in which Mr. Smith is willing to walk only with sons. With respect to the statement of the problem in vos Savant (1990; see also Selvin 1975), we change the door numbers (without loss of generality) in order to facilitate comparison with the coin-flip problems presented below:

---

the subtleties of Bayesian reasoning, but simply because of a failure to appreciate the subtleties of the sample space. A classic example of this type of mistake is Leibniz's error, which is believing that 11 and 12 are equally probable when rolling a pair of fair dice, because there is just one way for each sum to be partitioned into two numbers less than (or equal to) 6 (Gorroochurn 2012).

> *Monty Hall problem: Suppose you're on a game show, and you're given the choice of three doors. Behind one door is a car, behind the others, goats. You pick a door, say #3, and the host, who knows what's behind the doors, opens another door, say #1, which has a goat. He says to you, "Do you want to pick door #2?" Is it to your advantage to switch your choice of doors?*

As with the boy-or-girl paradox, the correct answer depends on conditions that have not yet been specified. One possibility is that Monty, the host, follows a rule that he must always reveal a goat from behind one of the two doors that the contestant does not choose. Further, in the case that Monty has two goats to choose from, he chooses a door (uniformly) at random.

Under these conditions, because one goat and one car will always remain covered once Monty reveals a goat, an intuitively appealing conclusion is that the odds in favor of the car being behind door #2 are 1:1 (relative to door #3), meaning that the contestant should be indifferent about switching.

Nevertheless, as in the previous problems, this simple reasoning is incorrect. To see why, notice first that before Monty opens door #1, the contents behind doors #1 and #2, respectively, are one of the following, each with equal probability: car–goat, goat–car, or goat–goat. However, once Monty reveals a goat behind door #1, the remaining possible arrangements behind doors #1 and #2 become goat–car and goat–goat. Because Monty must open door #1 in the case of goat–car, whereas he opens it only half of the time in the case of goat–goat, he is twice as restricted to open it in the case of goat–car. Therefore, given that the prior (relative) odds in favor of goat–car were 1:1, the posterior odds must double—that is, the odds in favor of the car being behind door #2 are now 2:1 (relative to door #3). As a result, it is in the contestant's interests to switch doors, as the probability of winning the car by doing so is 2/3.

Restricted Choice as a General-Purpose Approach

Throughout this paper we illustrate how the restricted choice approach is intuitive and straightforward to apply to a range of conditional probability problems. By contrast, while other approaches can do an excellent job of shaking people out of incorrect initial intuitions, they tend to employ either ad hoc explanations that do not readily generalize across problems or formal explanations that do, but at the expense of being less intuitive.

For example, in the *Parade Magazine* article in which she discussed the Monty Hall problem, vos Savant (1990) offered a modification of the problem to make more salient the benefit of switching after Monty opens a door to reveal a goat. She wrote, "Here's a good way to visualize what happened. Suppose there are a million doors, and you pick door #1. Then the host, who knows what's behind the doors and will always avoid the one with the prize, opens them all except door #777,777. You'd switch to that door pretty fast, wouldn't you?"

This modification effectively conveys the restricted choice intuition in a way that helps make the correct answer—to switch doors—more transparent. In the terms we have been using, because Monty must leave door #777,777 closed when the car is behind it, whereas he has a 1/999,999 probability of leaving it closed when the car is behind door #1, he is 999,999 times more restricted to leave door #777,777 closed when the car is behind door #777,777 (versus door #1). Because the prior odds between the two doors are 1:1, the posterior odds become 999,999:1 in favor of door #777,777. Indeed, when experimental subjects face a many-door version of the Monty Hall problem, they correctly decide to switch doors at a rate of approximately 85 percent, compared with only 15 percent when facing the standard version (Page 1998).

While the many-doors modification of the Monty Hall problem does lead to an immediate improvement in the rate of correct responses, it also has some important limitations. For one, when experimental subjects who face the manipulation then go back to the standard version of the Monty Hall problem, they proceed to make the wrong choice at rates similar to subjects that never faced the many-door version (Page 1998). Second, it does not indicate how to compute the posterior odds, which is necessary if one wishes to ascertain the value of switching. Third, the modification seems unlikely to be useful as a general problem-solving tool, as it is difficult to adapt to the other problems we have discussed so far.[7]

Another common approach to solving the Monty Hall problem—and the highest-voted answer on the question-and-answer website Mathematics Stack Exchange (https://math.stackexchange.com/q/96832)—involves answering as if the contestant decides whether to commit to switching *before* Monty chooses which of the two remaining doors to open (see also Krauss and Wang 2003). While this heuristic approach answers a slightly different problem, it appears to help people see that always switching yields the best of what the two remaining doors have to offer, and thus yields the car 2/3 of the time.

While reasoning through the Monty Hall problem without conditioning on which door Monty opens may help people shake off certain incorrect intuitions, this best-of-two-doors approach also has some important limitations. For one, it is not clear how to generalize it to address the other conditional probability problems discussed in the previous section. More importantly, because the best-of-two-doors approach ignores which of the two doors was opened, as well as Monty's rule for choosing between them in the case of two goats, the resulting probability—while correct numerically—is not the *conditional* probability that the problem implicitly requests. To see why this matters, assume that in the case that Monty has two goats to choose between, he always reveals the goat behind the lower-numbered door (rather than randomizing between the two doors, as implicitly assumed above). While the best-of-two-doors intuition still indicates that it is always strictly beneficial

---

[7]In the boy-or-girl problems, the analogous modification is for Mr. Smith to walk with all but one of his 999,999 children, and to meet him walking with only boys. In Bertrand's box paradox there would be 999,999 coins in each box, 999,998 coins would be drawn, and they would all need to be gold.

for the contestant to switch, this is no longer true in the event that Monty opens the lower-numbered door! Instead, the contestant should be indifferent between switching and not switching because Monty is now equally restricted to open the lower-numbered door, regardless of whether his two options are goat–goat or goat–car. Finally, while one could claim that this argument is unnatural, because Monty should be expected to randomize uniformly in the case of two goats, in the next section we provide an example of a coin-flip version of the Monty Hall problem in which the best-of-two-doors intuition fails to provide the correct answer even when Monty does randomize uniformly.

Yet another approach to solving conditional probability problems is to describe the sample space in detail and calculate the conditional probability directly. In the Monty Hall problem, for example, given the contestant's initial choice, one can generate all four (prize-placement, door-opened) combinations, and their probabilities, by laying out Monty's two-stage decision tree in which he first places the car behind one of the three doors (at random) and then chooses which door to open (according to his rule). One can then grind out the correct answer using the definition of conditional probability, rather than Bayes' rule.[8] While certainly correct, the relative disadvantage of sample space arguments is that they are typically more complex, and the intuition is less transparent.

When it comes to conditional probability problems, ad hoc intuitive explanations—as well as more complicated formal explanations—may be correct as far as they go. However, they are limited relative to restricted choice in terms of building a broader intuition for how the probability of interest in these kinds of problems can be altered by seemingly small changes in the selection procedure.

Restricted Choice in Coin-Flip Puzzles

In this section, we introduce a progression of coin-flip puzzles ("paradoxes") and solve them using restricted choice reasoning. The *next flip paradox* is nearly identical to the Monty Hall problem. When combined with the *alternation paradox*, it provides an explanation of why the earlier studies that purported to demonstrate a hot hand fallacy were actually biased. We then extend the alternation paradox into the *streak-reversal paradox*, which illustrates how these statistical puzzles can be related to selection bias in slightly richer settings.

> *Next flip paradox: Jack flips a coin three times, then tells you that the first flip is a heads. What is the probability that the second flip is also a heads?*

---

[8] More broadly, one can use a natural frequency intuition to arrive at the correct conditional probabilities for the Monty Hall problem. Gigerenzer and Hoffrage (1995) adapt the natural sampling approach to reframe conditional probability problems so that subjects can apply the definition directly, rather than updating priors with Bayes' rule.

The answer to this question depends on conditions that have not yet been specified. In particular, if Jack had decided to select which of the first two flip outcomes to reveal at random, or had simply planned on always revealing the outcome of the first flip, then the correct answer will be 1/2. This is precisely as in the basic boy-or-girl paradox, in which Mr. Smith chooses a child at random, regardless of gender. But instead, say that Jack was interested only in the respondent's beliefs about the probability that heads follows heads. Thus, assume that Jack had selected one of the first two flip outcomes at random according to the criterion that it be a heads (so that with two tails he could not have asked the question). In this case, the answer changes.

Under this selection criterion, the next flip paradox is nearly identical to the standard Monty Hall problem. In particular, just as Monty is able to look behind each door before opening one, which in turn reveals information regarding the location of the car, Jack looks at the outcome of each coin flip before selecting one, which in turn reveals information about the location of heads. To see the parallel more clearly, let Jack now be the game show host instead of Monty. In this game, Jacks flips three coins, leaving each behind a separate door. He then asks the contestant to choose one of the three doors, informing her that she will receive a prize if the door she chooses conceals a tails flip. Once the contestant has chosen one of the doors, Jack opens one of the other two doors at random, according to the criterion that he must reveal a heads flip (in the case of two tails flips, he cannot open either door). Finally, Jack offers the contestant the opportunity to switch. Assume that the contestant's initial choice is the third door, and that Jack opens the first door, revealing the first flip to be a heads. In this case, the first two flip outcomes must be either heads–tails or heads–heads. Then, by the same restricted choice reasoning as in the Monty Hall problem, Jack is twice as restricted to open the first door in the case of heads–tails as he is in the case of heads–heads. As a result, the probability that the second flip is a heads is $1/3$.[9]

Although the contestant can extract information about the second coin flip from the knowledge that the first flip is a heads, this does not imply that coins have memory. Instead, the contestant exploits the fact that Jack has inspected the outcome of the first two flips before choosing, which means that Jack's choice (probabilistically) reflects his knowledge. More subtly, this also implies that time's arrow is irrelevant—that is, if Jack were to instead reveal that the second flip was a heads, then the probability of heads on the previous (first) flip would similarly be 1/3.

Another coin-flip problem, the alternation paradox, brings us one step closer to illustrating the streak selection bias; indeed, this problem happens to be the exact probabilistic representation of the simple three-flip example of the bias given in table 1 of Miller and Sanjurjo (2018).

---

[9]A slight modification makes the next flip paradox identical to the younger boy-or-girl paradox: in this version, Jack flips the coin twice, then chooses one of the heads flips at random (final flip included) and tells you that it is the first flip.

> *Alternation paradox: Jack will flip a coin three times, then select a flip that is immediately preceded by a heads, at random. Assuming that Jack has a flip to select, what is the probability that the selected flip is a heads?*

In order for Jack to select a flip, he must inspect the outcomes of the first two flips. Given that at least one of the two has come up heads, it is clearly impossible that the sequence could have started with two tails. Let H__ be the event that Jack selects the second flip, which is preceded by a heads on the first flip; let _H_ be the event that Jack selects the third flip, which is preceded by a heads on the second flip. Conditional on Jack having chosen a flip, these events are equally likely. In the case that Jack selects the second flip, the probability that it is a heads is simply the solution to the next flip paradox, namely, $\Pr(HH\_|H\_\_) = 1/3$. On the other hand, if Jack selects the third flip, then $\Pr(\_HH|\_H\_) = \Pr(\_\_H) = 1/2$, as the outcome of the last flip cannot restrict Jack's choice of which immediate heads successor to select. It then immediately follows that

$$\Pr(\text{Heads}|\text{Flip preceded by a heads})$$

$$= \Pr(HH\_|H\_\_) \times \Pr(H\_\_) + \Pr(\_HH|\_H\_) \times \Pr(\_H\_)$$

$$= \left(\frac{1}{3} \times \frac{1}{2}\right) + \left(\frac{1}{2} \times \frac{1}{2}\right)$$

$$= \frac{5}{12}.$$

The next problem extends the alternation paradox to 100 flips and streak lengths of 3.

> *Streak reversal paradox: Jack, now a researcher, observes the outcome of 100 flips of a fair coin. He selects all of the flips that are immediately preceded by three consecutive heads and calculates the proportion of heads on these flips. He expects this proportion to be 0.5. Is he correct?*

While Jack's expectation is intuitively appealing, it turns out to be incorrect. In particular, the expected value of this proportion is not 0.50 but 0.46 (for the formula, see Miller and Sanjurjo 2018).

To see how the principle of restricted choice provides intuition for the streak reversal paradox, first observe that the expected proportion can be represented as a probability. In particular, the proportion of heads among the flips that Jack has selected is equal to the probability of heads on a flip chosen at random from among these flips. Next, imagine Jack choosing a flip at random from among the flips that he selected (those immediately preceded by three consecutive heads). If Jack were to choose, say, flip number 42, then intuition suggests that the odds of heads on that flip are 1:1. As with the alternation paradox, this intuition would be correct if Jack were to have chosen the flip *before* having examined the sequence. However, because

Jack instead examined the sequence first, then chose flip 42 based on information that he had regarding the outcomes of other flips in the sequence (including flip 42), this intuition is incorrect.

To see why the odds of heads on flip 42 are not 1:1, first observe that if flip 42 were a heads, then flips 39–42 would be HHHH, making flip 43 also immediately follow (at least) three consecutive heads. In this case Jack could have chosen flip 43 instead of flip 42. On the other hand, if flip 42 were instead a tails, then flips 39–42 would be HHHT, making it impossible for Jack to choose flip 43 (or 44, or 45). This implies that with a tails on flip 42 Jack would be relatively more restricted (likely) to choose flip 42, as there would be comparatively fewer eligible flips (on average) in the sequence from which to choose. Finally, the fact that Jack is more restricted to choose flip 42 in the case that it is a tails makes the likelihood that the flip he chose was a tails greater than the unconditional (prior) probability of flipping a tails, which in turn implies that the (posterior) probability that flip 42 is a heads is less than 0.5.[10]

This reasoning holds for any flip that Jack may choose, unless it happens to be the final flip of the sequence. For that flip, the posterior odds of a heads versus a tails are the same as the prior odds for the same reason given in the explanation of the alternation paradox.

Some Empirical Implications

We provide four empirical examples of how applying the principle of restricted choice can in some instances help us as researchers to avoid making critical mistakes in our design of experiments, analysis of data, and interpretation of results.

The Presumed Debunking of the Hot Hand

Having gone through the solutions to the coin-flip puzzles, it is now straightforward to explain the bias built into the seminal study of the hot hand fallacy by Gilovich, Vallone, and Tversky (1985) and similar studies that followed.

The original study conducted a controlled shooting experiment in which collegiate basketball players attempted 100 shots, from locations on the court at which they are expected to make half of them. To test for a hot hand, the authors compared each player's shooting percentage immediately following a streak of successes (makes) with his/her percentage immediately following a streak of failures (misses). Under their null hypothesis of no hot hand shooting, these two percentages are expected to be the same, and under the alternative hypothesis of hot hand shooting, the percentage following successes is expected to be larger than the percentage following failures.

---

[10]This explanation omits some details; see appendix A of Miller and Sanjurjo (2018) for a complete proof.

While this null hypothesis may seem correct, the streak reversal paradox makes clear that, perhaps counterintuitively, it is not. Indeed, if a robot player's shot outcomes were to be determined by repeated tosses of a fair coin (no hot hand), the expected shooting percentage following streaks of success would not be 0.50, but 0.46. By symmetry, the expected percentage following streaks of failures would be 0.54. Taking the difference, the total bias is 8 percentage points.[11] This means that if a researcher were to observe no difference in a player's shooting percentages, it would actually constitute (sizeable) evidence of the hot hand!

Upon correction for this bias in the shooting percentages of each of the players in the original study, the positive 3 percentage point average hot hand effect reported there (not statistically significant) becomes a statistically significant 13 percentage point effect (Miller and Sanjurjo 2018). This is a large effect and is roughly equal to the difference between a median and a top three-point shooter in the 2015–2016 NBA season.

Similarly biased measures were also used in the replications of the original hot hand study: a close replication with Olympic basketball players (Avugos, Bar-Eli, Ritov, and Sher 2013) and another using elite shooters from the annual NBA three-point "shootout" (Koehler and Conley 2003). As with the original study, a bias-corrected reanalysis reveals substantial evidence of hot hand shooting in both datasets (Miller and Sanjurjo 2018).

Clustering and Segregation

While the coin-flip puzzles and the streak selection bias pertain to measures of sequential dependence in time-series data, it turns out that the time dimension itself is not central to the bias. Instead, the key is that the selection of the data to be analyzed is determined by the outcomes of other (adjacent) flips in the same dataset. This in turn suggests the possibility of a more general selection bias that applies to measures of dependence across space just as easily as is does to measures of dependence across time.

Consider an $n \times n$ grid of cells, each colored red or blue according to the outcome of a fair coin flip. Now, suppose that we are interested in the probability that a cell is a red, given that all of its neighbors are red. An intuitive way to estimate this probability would be to select the subset of all cells that are surrounded by red and then calculate the proportion of red among these cells. However, this estimate will be biased downward due to a mechanism that is essentially identical to the bias that emerges in the one-dimensional setting of the streak reversal paradox. In particular, if one were to choose a cell from among those surrounded by red, the probability that this cell is blue would be greater than 50 percent. This is because if it were blue, then none of its neighbors could be surrounded by red, which would

---

[11] In fact, the bias is actually a bit more severe than this, due to an additional selection effect that is driven by the exclusion of sequences that do not have both of the following: (1) at least one shot that immediately follows a streak of made shots and (2) at least one shot that immediately follows a streak of missed shots.

lead to fewer such cells, making the probability of choosing any such cell, including itself, more likely.

The bias in this measure of the similarity between a cell and its neighbors suggests the possibility of such a bias appearing in measures of clustering in location preferences, as in studies of racial segregation. Indeed, this description of a grid of cells with two possible values is reminiscent of the classic work by Schelling (1971) on patterns of segregation. While it has no bearing on Schelling's main results, a bias happens to exist in one of his measures of clustering—"the average proportion of neighbors of like or opposite color." The reason for the bias is similar to that described in the previous paragraph. In particular, imagine choosing a cell at random from among the red cells. If more of that cell's neighbors are blue, then fewer red cells are available to be drawn. This in turn makes the chosen cell more likely to have been chosen to begin with. By consequence, a representative red cell is expected to have a higher proportion of blue neighbors than red neighbors.[12]

This bias extends to any spatial arrangement of outcomes, including lattices and networks. It is closely related to a bias in a well-known measure of spatial association, Moran's *I* (Moran 1950). The extent to which the magnitude of these biases is empirically relevant depends on the definition of a cluster and the size of the grid under consideration.[13]

Berkson's Paradox

Berkson's paradox (sometimes called Berkson's bias) is a form of selection bias. The original example involved a hypothetical case of two diseases that, while not associated in the general population, become negatively associated in the population of hospitalized patients (Berkson 1946). It is sometimes referred to as the "admission rate bias" (Sackett 1979), or as an instance of "collider bias" (for example, Westreich 2012), and it can be illustrated with the following example (adapted from Pearl 2009):

> *Berkson's paradox: Suppose that a randomly selected high school student has a 50 percent chance of having good SAT scores, along with a 50 percent chance of having good grades, and that the attributes are independent. Further, suppose that every student*

[12] Schelling (1971, p. 156) briefly considers this biased measure of segregation. Specifically, Schelling writes, "If we count neighbors of like color and opposite color for each of the 138 randomly distributed stars and zeros in [Schelling's figure 7], we find that zeros on the average have 53 percent of their neighbors of the same color, stars 46 percent. (The percentages can differ because stars and zeros can have different numbers of blank neighboring spaces.)" Of course, Schelling's main result was not to measure segregation but rather to show that a relatively weak preference for being near one's own type, together with the possibility of movement, would often lead to much stronger patterns of segregation.

[13] The bias in Moran's *I* measure of spatial autocorrelation is typically small, with an expected value of $-1/(n-1)$, where n is the total number of cells. For the cluster-related measures of association discussed above, the bias is stronger, but still weaker than the streak-related measures in time-series data. For example, in a 50×50 grid, the probability that one of the cells surrounded by 8 reds is itself red is approximately 48 percent, whereas in a 2,500-cell linear grid, the probability that one of the cells with 8 consecutive red cells to its left is itself red is approximately 44 percent.

with at least one good attribute applies to university and highlights his/her single best attribute in the application. If an applicant highlights good grades, then what is the probability that the applicant has good SAT scores?

Assuming that an applicant highlights an attribute at random (uniformly) in the case that both attributes are good, this problem is identical to the younger boy-or-girl paradox, as well as a two-coin version of the next flip paradox. In this problem, each attribute is good or not good with a 50–50 chance, just as each child is a boy or not a boy with a 50–50 chance. As a result, an applicant who has good grades and poor SAT scores is twice as restricted to highlight good grades compared with an applicant with both good grades and good SAT scores. Thus, as prior odds are even, the principle of restricted choice leads to posterior odds of 2:1 in favor of the applicant having good grades and poor SAT scores; that is, given an emphasis on good grades, the probability that the applicant has poor SAT scores is 2/3.

While it remains true that, among the applicants with good grades, half of them also have good SAT scores, this subgroup constitutes just 1/3 of the applicant pool. The remaining 2/3 of the applicants, on the other hand, have just one good attribute. Thus, there will be a negative correlation between attributes in the applicant pool, despite the correlation in the general population being zero.

This phenomenon could easily lead a casual observer to fallacious beliefs. For example, a student (or professor) who spends enough time in a university environment may come to believe (incorrectly) that certain attributes that are associated with good grades (like diligence) are in general inversely related to those attributes associated with good SAT scores (like brilliance). This mistake is analogous to a gambler holding the belief that streaks are more likely to end rather than continue, because in his personal experience this is, in fact, representative of a typical night at the casino (as conveyed in the streak reversal paradox, presented above, and in gambler's verity, presented below).

It is not difficult to imagine that a similar bias may be present in experiments in which performance on behavioral tasks that involve cognitive ability is correlated with a personality measure such as conscientiousness. Indeed, because experimental subjects in research studies may be further selected on attributes such as budget constraints and intellectual curiosity, one can similarly imagine the discovery of appealing new correlations that are nevertheless spurious—such as a hypothetical negative correlation between measures of intellectual curiosity and greedy or selfish behavior in experimental tasks. As one example, Murray, Johnson, McGue, and Iacono (2014) proposed that empirical work documenting an (internally valid) negative correlation between conscientiousness and cognitive ability may instead merely be reporting a statistical artifact that is driven by a selection bias identical to Berkson's paradox.

A Hypothetical Case: Gambler's Verity and Psi Research

The same bias that underlies the alternation paradox can be used to generate a puzzle in which a strategy for predicting randomly generated outcomes can appear

to outperform what would be expected by chance. In particular, this can happen if a researcher is unaware of the implicit selection bias that the strategy generates.

> *Gambler's verity: Imagine a roulette wheel in which half of the slots are red and half are black (for simplicity). Jill will observe exactly three spins of the wheel and has committed to the following betting strategy: whenever observing a red (R), bet black (B) on the next spin; otherwise, do not bet. Do you expect Jill to win half of her bets?*

Jill's betting strategy will restrict her to betting on the second spin, the third spin, or both. Thus, there are three possible outcomes: she will win on none of her bets, half of them, or all of them. While intuition may suggest that she is expected to win on half of her bets, this is incorrect, as it overlooks the fact that the three outcomes are not equally likely. To see this, we can enumerate the sample space, as follows: if the sequence is BBB or BBR, Jill will not bet; otherwise, for the remaining six equally likely sequences, she will bet. Given that Jill bets, she has a $3/6 = 1/2$ probability of winning all of her bets (RBR, RBB, BRB), a $1/6$ probability of winning half of them (RRB), and a $2/6 = 1/3$ probability of losing all of them (BRR, RRR). As a result, Jill is expected to win more bets than she loses, with an expected win rate of $(1/2 \times 1) + (1/6 \times 1/2) + (1/3 \times 0) = 0.58$. In fact, her high success rate immediately follows from the solution to the alternation paradox, which we solved using the same restricted choice thinking as in our solution to the Monty Hall problem. That is, Jill's expected win rate is equivalent to the statement that for a randomly selected flip that is immediately preceded by a heads, the probability of a tails (an alternation) is $1 - 5/12 = 0.58$.

While it appears that Jill has discovered a strategy with which she can expect to win money, this is not true. In particular, relative to the high-probability sequences in which she walks away ahead, in the low-probability sequences in which she walks away behind she wagers 50 percent more and her absolute (negative) profit is 50 percent greater. The key to this asymmetry is that in some of these sequences Jill is betting only once, but in others she is betting twice. Specifically, conditional on walking away ahead, the sequences RBR, RBB, and BRB are equally likely, and in each sequence Jill wagers once and wins once. On the other hand, conditional on walking away behind, while the sequences BRR and RRR are also equally likely, for the sequence BRR Jill wagers once and loses, but for the sequence RRR she wagers twice and loses twice. As a result, when Jill walks away behind ($1/3$ probability), she is expected to wager 1.5 times with a net payoff of -1.5, whereas when she walks away ahead ($1/2$ probability), she is expected to wager 1 time with a net payoff of 1. As a result, given fair odds, Jill is expected to break even. This (sad) state of affairs brings to mind the old Las Vegas proverb: the probability of winning is inversely proportional to the amount of the wager.

To see how the gambler's verity problem could have implications for social science research, consider the hypothetical case of the amazing Zener, an ESP master who claims to have a scientifically validated method to train people in precognition. In order to validate his method, he devises a test to prove that his

students can do better than chance at predicting the outcomes of coin flips. For each student, a group of objective third-party researchers will flip a coin 100 times, and the student will predict only on flips for which he/she "senses" the ensuing outcome. According to Zener, not all of his trainees have learned how to predict, so he requests that the researchers merely count how many of his students predict at better than chance rates.

Following these instructions, the researchers find that of the 1,000 students tested, 490 predict at a rate better than chance, 395 at a rate worse than chance, and 115 at the rate of chance. Thus, the odds are found to be substantially in favor of a student predicting at better-than-chance rates, relative to worse-than-chance rates. Furthermore, the average student is observed to have a 54 percent success rate on his/her predictions. Mystified by these statistically significant results, the researchers are left to conclude that Zener must indeed have amazing abilities.

However, the researchers' conclusion is premature, as the observed results can easily occur in the absence of precognition. In fact, this outcome is close to what would be *expected* if Zener had instructed his students to simply predict a tails whenever the previous three flips are heads—the equivalent of predicting a streak reversal (tails) in the setting of the streak reversal paradox.

Conclusion

We have shown that the usefulness of the principle of restricted choice as an inferential tool extends well beyond the settings of contract bridge and the Monty Hall problem. When naturally quantified as the updating factor in the odds form of Bayes' rule, restricted choice provides a simple, intuitive, and general approach to thinking through and solving classic conditional probability puzzles. Moreover, it can be used to identify novel biases in important empirical settings. Thus, the principle is capable of helping researchers avoid certain intuitively appealing but critical errors when designing experiments, analyzing data, and interpreting results.

References

Avugos, Simcha, Michael Bar-Eli, Ilana Ritov, and Eran Sher. 2013. "The Elusive Reality of Efficacy–Performance Cycles in Basketball Shooting: Analysis of Players' Performance under Invariant Conditions." *International Journal of Sport and Exercise Psychology* 11(2): 184–202.

Bar-Hillel, Maya, and Ruma Falk. 1982. "Some Teasers concerning Conditional Probability." *Cognition* 11(2): 109–22.

Benjamin, Daniel J. 2019. "Errors in Probabilistic Reasoning and Judgment Biases." Chap. 2 in *Handbook of Behavioral Economics*, vol. 2, edited by B. Douglas Bernheim, Stefano DellaVigna, and David Laibson. Amsterdam: Elsevier.

Berkson, Joseph. 1946. "Limitations of the Application of Fourfold Table Analysis to Hospital Data." *Biometrics Bulletin* 2(3): 47–53.

Bertrand, Joseph. 1889. *Calcul des probabilités.* Paris: Gautier-Villars et Fils.

Borel, Émile, and André Chéron. 1940. *Théorie mathématique du bridge à la portée de tous.* Paris: Gauthier-Villars.

Fehr, Ernst, and Jean-Robert Tyran. 2005. "Individual Irrationality and Aggregate Outcomes." *Journal of Economic Perspectives* 19(4): 43–66.

Fox, Craig R., and Jonathan Levav. 2004. "Partition-Edit-Count: Naive Extensional Reasoning in Judgment of Conditional Probability." *Journal of Experimental Psychology: General* 133(4): 626–42.

Friedman, Daniel. 1998. "Monty Hall's Three Doors: Construction and Deconstruction of a Choice Anomaly." *American Economic Review* 88(4): 933–46.

Gardner, Martin. 1961. *The Second Scientific American Book of Mathematical Puzzles and Diversions.* New York: Simon and Schuster.

Gigerenzer, Gerd, and Ulrich Hoffrage. 1995. "How to Improve Bayesian Reasoning without Instruction: Frequency Formats." *Psychological Review* 102(4): 684–704.

Gillman, Leonard. 1992. "The Car and the Goats." *American Mathematical Monthly* 99(1): 3–7.

Gilovich, Thomas, Robert Vallone, and Amos Tversky. 1985. "The Hot Hand in Basketball: On the Misperception of Random Sequences." *Cognitive Psychology* 17(3): 295–314.

Gorroochurn, Prakash. 2012. *Classic Problems of Probability.* New Jersey: John Wiley and Sons.

Kahneman, Daniel. 2011. *Thinking, Fast and Slow.* New York: Farrar, Straus and Giroux.

Kluger, Brian D., and Steve B. Wyatt. 2004. "Are Judgment Errors Reflected in Market Prices and Allocations? Experimental Evidence Based on the Monty Hall Problem." *Journal of Finance* 59(3): 969–98.

Koehler, Jonathan J., and Caryn A. Conley. 2003. "The 'Hot Hand' Myth in Professional Basketball." *Journal of Sport and Exercise Psychology* 25(2): 253–59.

Krauss, Stefan, and X. T. Wang. 2003. "The Psychology of the Monty Hall Problem: Discovering Psychological Mechanisms for Solving a Tenacious Brain Teaser." *Journal of Experimental Psychology: General* 132(1): 3–22.

Miller, Joshua B., and Adam Sanjurjo. 2018. "Surprised by the Hot Hand Fallacy? A Truth in the Law of Small Numbers." *Econometrica* 86(6): 2019–47.

Moran, P. A. P. 1950. "Notes on Continuous Stochastic Phenomena." *Biometrika* 37(1/2): 17–23.

Murray, Aja L., Wendy Johnson, Matt McGue, and William G. Iacono. 2014. "How Are Conscientiousness and Cognitive Ability Related to One Another? A Re-Examination of the Intelligence Compensation Hypothesis." *Personality and Individual Differences* 70: 17–22.

Nalebuff, Barry. 1987. "Puzzles: Noisy Prisoners, Manhattan Locations, and More." *Journal of Economic Perspectives* 1(1): 185–91.

Page, Scott E. 1998. "Let's Make a Deal." *Economics Letters* 61(2): 175–80.

Pearl, Judea. 2009. "Causal Inference in Statistics: An Overview." *Statistical Surveys* 3: 96–146.

Rabin, Matthew, and Dimitri Vayanos. 2010. "The Gambler's and Hot-Hand Fallacies: Theory and Applications." *Review of Economic Studies* 77(2): 730–78.

Reese, Terence. 1960. *Master Play in Contract Bridge.* New York: Dover Publications.

Rosenhouse, Jason. 2009. *The Monty Hall Problem: The Remarkable Story of Math's Most Contentious Brain Teaser.* New York: Oxford University Press.

Sackett, David L. 1979. "Bias in Analytic Research." *Journal of Chronic Diseases* 32(1/2): 51–63.

Schelling, Thomas C. 1971. "Dynamic Models of Segregation." *Journal of Mathematical Sociology* 1(2): 143–86.

Selvin, Steve. 1975. "Letters to the Editor: A Problem in Probability." *American Statistician* 29(1): 67.

Vazsonyi, Andrew. 1999. "Which Door Has the Cadillac?" *Decision Line* 30(1): 17–19.

vos Savant, Marilyn. 1990. "Ask Marilyn." *Parade Magazine*, September 1990, p. 15.

Westreich, Daniel. 2012. "Berkson's Bias, Selection Bias, and Missing Data." *Epidemiology* 23(1): 159.