Inference by Believers in the Law of Small Numbers

Author(s): Matthew Rabin

Source: *The Quarterly Journal of Economics*, Aug., 2002, Vol. 117, No. 3 (Aug., 2002), pp. 775-816

Published by: Oxford University Press

Stable URL: https://www.jstor.org/stable/4132489

INFERENCE BY BELIEVERS IN THE LAW OF SMALL NUMBERS*

MATTHEW RABIN

People exaggerate the degree to which small samples resemble the population from which they are drawn. To model this belief in the "Law of Small Numbers," I assume that a person exaggerates the likelihood that a short sequence of i.i.d. signals resembles the long-run rate at which those signals are generated. Such a person believes in the "gambler's fallacy," thinking that early draws of one signal increase the odds of next drawing other signals. When uncertain about the rate, the person overinfers from short sequences of signals that the rate is more extreme than it is, and consequently infers that there is more variation in these rates among different sources than there is. Economic applications are discussed, such as how the model predicts that investors will believe in nonexistent variation in the quality of mutual-fund managers.

I. INTRODUCTION

Loosely put, the law of large numbers tells us that a large random sample from a population will have a distribution that closely resembles that of the overall population. But many people believe in what Tversky and Kahneman [1981] identified and labeled the "law of small numbers": they exaggerate how likely it is that a *small* sample resembles the parent population from which it is drawn. This paper develops a simple model reflecting this error, and studies how people making this error differ from Bayesians in their inferences. Most importantly, the model shows that the law of small numbers leads to the tendency to overinfer from short sequences and to believe in nonexistent variation. I illustrate some possible economic implications of the model, such as showing it can simultaneously explain short-run underreaction and medium-term overreaction by investors to recent corporate performance, as well as the tendency by investors to exaggerate the variation in skill among mutual-fund managers.

In Section II, I review some psychological evidence that peo-

* I thank Davis Beekman, Christopher Carpenter, Erik Eyster, David Huffman, Christopher Meissner, and Ellen Myerson for valuable research assistance and Colin Camerer, Erik Eyster, Sendhil Mullainathan, Lawrence Katz, Edward Glaeser, Dimitri Vayanos, an anonymous referee, and seminar participants at the University of California-Berkeley, Yale University, Johns Hopkins University, Harvard University-Massachusetts Institute of Technology, Cambridge University, Oxford University, and Tilberg University for helpful comments. I thank the Russell Sage, Alfred P. Sloan, John and Catherine MacArthur, and National Science Foundations (Award 9709485) for financial support.

The Quarterly Journal of Economics, August 2002

ple systematically depart from Bayesian reasoning with biases of the sort modeled in this paper, such as belief in the law of small numbers, the gambler's fallacy, and overinference. In Section III, I present the model. A person observes a sequence of binary signals of some underlying quality, such as a sequence of good or bad investments by a mutual-fund manager that signal her underlying competence, or a sequence of good or bad performances by a company that signals its long-run prospects. I assume that each value of the signal is generated randomly from a stationary probability that I shall refer to as the "rate." The person is a Bayesian and has correct probabilistic priors about this rate. *But,* whereas in reality these signals are generated by an i.i.d. process, the person believes that they are generated by random draws *without replacement* from an "urn" of $N < \infty$ signals, where the urn contains the proportion of the two values of the signal corresponding to the rate. This captures belief in the law of small numbers, since it means that the person believes that the proportion of signals must balance out to the population rate before N signals are observed. As N becomes infinitely large, the person becomes fully Bayesian; the smaller is N, the more he believes in the law of small numbers.

The model leads directly to the "gambler's fallacy": people expect the second draw of a signal to be negatively correlated with the first draw. Because we exaggerate how likely it is that a small set of coin flips yields very close to half heads and half tails, if early flips are disproportionately heads, the "law of averages" tells us that the next flips are more likely to be tails. And if an observer is sure that a particular fund manager invests successfully close to half the time even over short intervals, then he thinks that success in one year implies less than 1/2 chance of success in the following year.

In Section IV, I turn to the crux of the model, which examines how belief in the law of small numbers leads to overinference by a person who is uncertain about the rate by which signals are generated. Exaggerating the likelihood that a short sequence of signals will closely resemble the underlying rate leads to exaggerating the likelihood that the underlying rate resembles a short sequence of signals. If a person believes every pair of flips of a fair coin generates one head and one tail, then he believes that two heads in a row indicates a biased coin. If he believes that an average fund manager is successful once every two years, then he believes that a fund manager who is successful two years in a row must be unusually good. I formalize this overinference result by showing that, after two signals, a believer in the law of small numbers always has on average more extreme beliefs than he should.

In Section IV, I also discuss an immediate and important corollary of such overinference: when a person observes a small number of signals from each of a large population of different "sources," belief in the law of small numbers leads him to exaggerate the variation in rates among the sources. I believe that such "fictitious variation" is one of the economically most important implications of the law of small numbers. Consider an observer of financial analysts, and suppose that he observes two performances from a large number of analysts—as he might if he reads an article that lists the performances of a large number of mutual-fund managers over the last couple of years, or if he observed a series of them he has hired for brief durations. The model predicts that if, in truth, all analysts are average—and a Bayesian with any initial beliefs would eventually figure this out—the believer in the law of small numbers will infer that some analysts are good and some are bad. Because he underestimates how often average analysts will have consecutive successful or unsuccessful years, he interprets what he sees as evidence of the existence of good and bad analysts.

In Section V, I extend and apply the fictitious-variation result to the study of finance by presenting a stylized model of investors who observe a short series of performances by firms, and predict near-term performance from these observations. While different patterns may also be consistent with the law of small numbers, the model can predict that—due to the gambler's fallacy—investors underpredict repetition of short strings of performances while—due to overinference—they overpredict repetition of longer strings. This provides one plausible and conceptually parsimonious psychological account of a phenomenon in financial markets—short-term underreaction to announcements by firms but medium-term overreaction—that has recently been modeled using various rational and quasi-Bayesian models.

In Section VI, I investigate what the person infers when he observes many signals from a given source. Here a problem arises with the model, since the number of signals observed may be larger than the urn size the person believes in. To close the model and capture some of the psychological evidence reviewed in Section II, I assume that the person believes that the "urn" generating the signals is replaced every two periods. Such a deterministic and frequent replacement of the urn is of course highly artificial. But it serves to capture in a tractable way extensive evidence that people expect small subsequences of a long sequence to yield signals in approximately the same proportions as the overall sequence. That is, just as people do not expect the composition of small samples to differ dramatically from overall population proportions, they also do not expect to see "streaks" of signals that are not representative of the overall frequency in a sequence.

The person's theory of streaks implies that the inferences he makes depend not just on the proportions of signals he observes, but also on the precise sequence of those signals. Unlike when observing a small number of signals, after observing a long sequence of signals a believer in the law of small numbers is too likely to believe that the rate is *less* extreme than it is. This is because he struggles to explain why he is observing so many streaks of rare signals, which he thinks are *very* unlikely. In his attempt to explain such streaks, he may come to believe that the true rate is close to 50/50, even if such a moderate rate does not accord with the overall frequency of the signals. These results rely on the model's assumption that the person believes that the intrinsic rate he is observing does not vary over time. Rather than sticking with his belief in such a constant rate and trying to reconcile that belief with a surprising number of streaks, however, a person may explain such streaks by assuming that there is underlying variation even when there is none. Such a false belief, in fact, corresponds to the well-documented "hot-hand fallacy" whereby people tend to believe in nonexistent streakiness. I briefly discuss how this paper's model might be accommodated to incorporate such inference, and hence some intuition about how belief in the law of small numbers may help explain the hot-hand fallacy.

In Section VII, I examine inference by a person who decides what signals to observe based on his earlier observations, so that the sequence of signals a person observes is endogenous. Suppose that a person employs financial analysts one at a time, and decides when to switch analysts based on his beliefs about the talent of his current and other analysts. Assuming that he observes only the performance of his current analyst, I show that such a person will eventually become convinced that average talent is less than it is. The investor switches quickly from a fund manager who initially performs poorly—and when he does so he has overinferred that the analyst is bad. But he sticks with a fund manager who initially performs well—until he discovers (as he will) that she is average. Because he corrects his overly positive inference but not his overly negative inference, his beliefs are biased downward. Another interesting possibility is that there can be two different long-run steady-state beliefs that a person may converge on depending on his initial beliefs and the early signals he observes. If an investor initially believes (correctly) in relatively little quality dispersion among financial advisors, he will not switch advisors often, and hence will observe enough of each advisor to learn that she is average. But if he initially believes (falsely) in wide quality dispersion, he may frequently switch advisors after poor performance, and because he does not observe as much on average of each advisor, end up maintaining his belief in wide quality dispersion. More generally, different belief and behavior profiles may occur with high probability, even in single-person environments whose information structure is rich enough such that a Bayesian's steady-state beliefs and behavior would be deterministic and correct.

I conclude in Section VIII with a brief discussion of some of the limits of the model in this paper, and ideas for modifications to rectify these limits and facilitate further economic applications.

II. Evidence for the Law of Small Numbers

The term "the law of small numbers" was coined by Tversky and Kahneman [1971] to describe how people exaggerate the degree to which the probability distribution in a small group will closely resemble the probability distribution in the overall population. They and other researchers have emphasized the connection between the law of small numbers, the gambler's fallacy, regression errors, overinference from short sequences, and other mistakes.[1] In this section I review some of the broad array of evidence that supports the assumptions and results presented in the remainder of the paper about these various phenomena.

As an illustration of the basic belief in the law of small numbers, Tversky and Kahneman [1971] asked a group of mathematical psychologists to forecast the likelihood of replication of

1. Tversky and Kahneman relate the law of small numbers to another bias, *the representativeness heuristic.*

results in a variety of scenarios presented to them. Participants were, for instance, told that a pattern of behavior matching a theory had been identified as statistically significant in an experiment with twenty subjects, and asked to predict the likelihood that the pattern of behavior would reappear as statistically significant in a subsequent experiment with ten subjects. The respondents greatly exaggerated the likelihood of replication, apparently exaggerating the likelihood that true theories would show up as statistically significant even in small samples. Further evidence indicated that the source of the error was that people fundamentally expect that population proportions reliably show up even in small samples.[2]

Much of the psychological evidence (and some of the analysis below) concerns not people's beliefs about isolated small samples, but rather their beliefs about small subsequences embedded within long sequences of signals. Bar Hillel and Wagenaar [1991] review the extensive evidence for the existence of a "local representativeness" bias, in which people expect even short strings within a long sequence of signals to contain proportions similar to the long sequence. Most such evidence comes from two types of experiments, *production tasks* and *recognition tasks.* In production tasks, people are asked to produce "random sequences," typically instructed in a way that should evoke the desire to produce sequences resembling i.i.d. sequences with which they are familiar, such as being told to "generate a random sequence, such as one might expect from a long sequence of flips of an unbiased coin."

In a series of papers Rapoport and Budescu [1992, 1997] and Budescu and Rapoport [1994] confirm the general patterns reported in Bar Hillel and Wagenaar [1991]. Rapoport and Budescu

2. Here and elsewhere believers in the law of small numbers are too inattentive to sample size. This inattentiveness, however, manifests itself in a second way. Although people believe in the law of small numbers, they *do not* necessarily believe in the law of large numbers: while overestimating the resemblance of small samples to the overall population, people *underestimate* the resemblance large samples will have to the overall population. Kahneman and Tversky [1973], for instance, found that subjects on average thought that there was a greater than 1/10 chance that, of 1000 babies born on a given day, more than 750 would be male. The actual likelihood is (much) less than 1 percent. To overstate it a bit, people seem to have a universal probability distribution over sample means that is insensitive to the sample size. As Kahneman and Tversky [1973, p. 45] note, this has important implications for inference: people often infer a lot from statistics reported in percentage terms even from small sample sizes, but by the same token are not convinced when they should be by huge sample sizes. The results emphasized below pertain mostly to small-sample inference and prediction.

[1992, p. 355] asked subjects to "simulate the random outcome of tossing an unbiased coin 150 times in succession," while Rapoport and Budescu [1997, p. 612] asked subjects to "imagine a sequence of 150 draws with replacement from a well-shuffled deck, including five red and five black cards, and then call aloud the sequence of these binary draws." The switching rate between successive elements of the sequence people produced around 58 percent in each of two experiments. Moreover, subjects were less and less likely to choose a signal when more and more of the preceding choices were those same signals. The probability that a subject would produce a signal given that the previous 0, 1, 2, or 3+ signals chosen were that same signal were as follows:[3]

$$\begin{array}{rl} \Pr(A|B) & 58.5\% \\ \Pr(A|AB) & 46.0\% \\ \Pr(A|AAB) & 38.0\% \\ \Pr(A|AAA\ldots) & 29.8\%. \end{array}$$

Notice that these and many of the empirical findings discussed here show a belief in local representativeness, not merely a more generalized belief in negative autocorrelation in signals. Subjects viewed long streaks as more in need of balancing than short streaks, since they require more countersignals in order to make the local range of signals representative. This specific pattern of negative autocorrelation is the sequential analogue of the law of small numbers, and is the fundamental feature in how people interpret sequences that drives many of the model's results.

Rapoport and Budescu also develop interesting variations on the production tasks that more clearly than in previous studies provide subjects incentive to produce i.i.d. sequences. Experiments by O'Neill [1987], Rapoport and Budescu [1992], and Budescu and Rapoport [1994] studied the sequence of moves chosen by subjects in variants of a "matching pennies" game: two-player zero-sum games where players have a clear incentive to choose

3. I derived these numbers from Table 7 of Rapoport and Budescu [1997] as follows. $\Pr(A|B)$ is simply a percentage of two-tuples that were XY rather than XX; $\Pr(A|AB)$ is derived as the relative proportion of YXX sequences to XYX sequences, since this represents the percentage of time the subjects chose to repeat the second element rather than the first element in triplets where the first two elements differed. $\Pr(A|AAB)$ was derived as the relative frequency of $YXXX$ sequences to $XYYX$ sequences; and $\Pr(A|AAA\ldots)$ was derived as the relative frequency of $XXXX$ sequences to $XXXY$ sequences. For ease of presentation, the numbers I report are just the simple average of these numbers as derived from the "Observed" columns of Experiments 1 and 2.

unpredictably. Compared with the typical "produce-a-random-sequence" binary-choice experiment experiments, alternation probabilities were reduced but still statistically significantly greater than 50 percent.

Research on the local-representativeness bias employing recognition tasks, where participants are asked to identify which from a menu of sequences appear "random," yields results very similar to those from production-task research. Bar Hillel and Wagenaar [1991] report that, in both production and recognition tasks, the average "switching rate" is about 60 percent in representative binary studies. Finally, there is more limited evidence from *prediction tasks,* where subjects are asked to guess the next outcome in a series, such research also demonstrates that people exhibit the gambler's fallacy and local-representativeness bias.

This literature is not conclusive evidence for the gambler's fallacy for a variety of reasons.[4] In particular, when the underlying probability of a binary signal is 50/50, all predictions have an equally good chance of being right. Hence, any observed patterns in prediction or production tasks are not errors per se.[5] Some laboratory experiments, however, provide stronger evidence by identifying betting behavior that indicates that belief in the gambler's fallacy moves subjects away from 50/50 beliefs.

Not all the evidence for local representativeness come from the laboratory. There are, for instance, several studies demonstrating the existence of the gambler's fallacy in lottery play. Stronger evidence comes from pari-mutuel betting. Following similar research by Clotfelter and Cook [1993], Terrell [1994] studied the pari-mutuel pick-three lottery run by the State of New Jersey. Each day, bettors are allowed to buy tickets guessing the exact three-digit number the state would draw that day, where the state divides 52 percent of money bet on a given day's pick-three lottery evenly among those bettors who choose the winning number. Hence, the amount distributed to winners each

4. Moreover, while taken as a whole the experimental literature clearly supports the prevalence of the gambler's fallacy and local representativeness, the support is not universal. Experiments in Edwards [1961] and Lindman and Edwards [1961], for instance, provide harder-to-interpret evidence both for and against the gambler's fallacy.

5. And in many of the non-50/50 prediction-task experiments, it is difficult to interpret behavior because of *probability matching.* When making a long series of predictions, or a long series of choices whose payoffs depend on the signals, participants tend to pick particular outcomes in proportion to the frequency with which they expect the corresponding signal rather than always picking the highest-odds choice.

TABLE I

	Number	Mean
Winners repeating within 1 week	8	$349
Winners repeating between 1 and 2 weeks	8	$349
Winners repeating between 2 and 3 weeks	14	$308
Winners repeating between 3 and 8 weeks	59	$301
Winners not repeating within 8 weeks	1622	$260
All winners	1714	$262

Average payouts to winning numbers

day is an indicator of both the number of bettors on that number and the costliness of errors bettors are making. Tickets cost 50 cents each, so that if equal numbers of bettors choose all numbers, winnings on a given day should be $260. If the winnings are significantly higher than that, it indicates that too few bettors are betting that number. Terrell examines the data from 1785 daily drawings on New Jersey's pick-three numbers game from 1988 to 1992. He examines betting on numbers that have recently won. Table 1 from Terrell [1994, p. 311], reproduced here as Table I with some rounding and omitting standard deviations, reports the average winnings for numbers as a function of when that number last won.

The pattern clearly demonstrates the gambler's fallacy. From Table I it can be inferred, for instance, that 25 percent fewer lottery players in New Jersey bet on a number that won a week ago than if it won nine or more weeks ago.[6]

There is much less evidence on overinference based on small samples than on local representativeness and the gambler's fallacy. But there is some. A series of experiments by Grether [1980, 1992] and Camerer [1987] has subjects observe a series of draws coming from one of two underlying rates which they do not know, but which they make inferences from the draws. In these experiments, the subjects make gambles or investments whose attrac-

6. A study of racetrack betting by Metzger [1985] provides another verification of the gambler's fallacy in pari-mutuel betting; consider Metzger's study of betting at the race track. Among other biases, she finds that the odds given for long-shot horses late in a day are higher when a long-shot has won earlier in the day than when no long shots have won. Because the odds are continuously adjusted to reflect the amount of betting, this indicates that the bettors anticipated that a long shot winning early means one will not win later. Presumably, bettors take the view that lightning does not strike twice.

tiveness depends on their beliefs about an underlying rate. These experiments are careful to explain all aspects of the environment to the subjects, involve incentives, markets, and allow for learning. They all, to varying degrees, provide support for the overinference hypothesis.

Camerer [1987] studied a group of undergraduates at the Wharton School at the University of Pennsylvania, all of whom had taken both statistics and economics courses and participated in an asset-market experiment in which they had the incentive to correctly predict an underlying variable. Subjects observed three draws being drawn *with replacement* from one of two urns, urn X which contained exactly twice as many black balls as red balls, and urn Y which contained exactly twice as many red balls as black balls. The ex ante probability of urn X was .6. After observing the three draws, the subjects participated in an asset market for assets whose value depended on which urn was generating the draw of the balls. Hence, they were essentially betting on their posterior beliefs about the relative likelihood of X and Y.

Results demonstrated overinference by subjects: when either one out of three or two out of three balls were red, i.e., when the proportions of the three draws exactly reflected the proportions of either the X or Y urn, market behavior clearly indicated that subjects were exaggerating the likelihood that the urn was the one matching the proportion of red and black balls drawn. While significant overinference was not found when either zero or three of the balls were red, the strong results for the more moderate cases indicate overinference about the underlying rate based on small samples. Subjects persisted in making the mistake through 50 rounds of repetition.[7]

III. THE MODEL

Throughout the paper I consider a situation where there is a finite number of possible *rates,* $\theta \in [0,1]$, at which an infinite sequence of i.i.d. signals, s_t, $t = 1, 2, \ldots$, is generated. Each signal s_t takes on a value of either a or b, where for each t,

7. The results also clearly indicate that the overinference was due to the law of small numbers rather than the (related) phenomenon of base-rate neglect. Camerer notes that had subjects been neglecting the base rates of the two urns, the overinference following two-red/one-black draws should have been more severe than for one-red/two-black draws, since the 2:1 Y urn had lower likelihood. Grether [1980, 1992] found similar results, while emphasizing more base-rate neglect and the role of learning and incentives.

$\text{prob}(s_t = a) = \theta$. Let Θ denote the set of rates that occur with positive probability; the rate θ occurs with prior probability $\pi(\theta) > 0$, where $\Sigma_\Theta\ \pi(\theta) = 1$. Given θ, signals are generated by an i.i.d. process.

The model describes a person who begins with correct prior beliefs about the probability distribution π over possible rates Θ and is fully Bayesian. But to capture the belief in the law of small numbers, instead of understanding that the world is i.i.d., there is a positive integer N such that for each rate θ, the person believes signals are drawn without replacement from an "urn" of size N consisting of exactly θN a signals and $(1 - \theta)N$ b signals.[8] To reconcile his belief in finite urns and to model his belief in "local representativeness" as discussed in Section II, I assume that when he observes long sequences of signals the person believes that this urn is renewed after every two draws. That is, in every odd period he believes a first signal is drawn from an N-signal urn, and in every even period he believes a second signal is drawn *without replacement* from the same urn he drew from in the previous period. Assuming instead that the person thinks there is a constant 50 percent chance each period that the urn is renewed would make many aspects of the model less artificial, but I do not believe would change the qualitative results and would be far less tractable. As I develop the model, the advantage of the deterministic renewal of the urn will be clear: while the person does not recognize that the world is i.i.d., he believes that *pairs* of signals are generated by an i.i.d. process, which vastly simplifies analysis of his belief formation. Throughout the paper I assume that the person believes that the first signal he observes is the first signal of a new urn, even if he is aware that previous, unobserved sequences have been generated.

When N is large, the person perceives the signals to be close to uncorrelated, and his inference and predictions become that of a Bayesian as $N \rightarrow \infty$. But when N is small, the person is very biased. Suppose, for instance, that an observer is positive that a particular fund manager invests successfully with probability 1/2. If $N = 4$, the observer thinks the analyst has an "urn" of two good

8. While the psychological evidence shows the existence of a cognitive bias, the formal model also lends itself to a more literal Bayesian interpretation: people may completely understand the nature of i.i.d. stochastic processes, but merely underestimate how common such processes are. Little depends on interpreting anything in the paper as a cognitive error rather than merely assuming that people have an empirical misconception about what random processes prevail in the world.

years and two bad years. Then if the analyst is successful in her first year, the observer thinks that there is only a 1/3 chance that she will have an above-average year the following year.[9]

To avoid tedious repetition of the phrase "a person who believes in the law of small numbers," for the remainder of the paper I shall refer to a believer in the law of small numbers as "Freddy," named for a guy I once knew who believed in the law of small numbers.[10]

To make the model fully coherent, it must be that Freddy's prior beliefs always put positive weight on some rate whose urn contains at least two of both signals. This is necessary and sufficient to ensure that Freddy believes all sequences of a's and b's are possible. Formally, I require that for all $\theta \in \Theta$, θN is an integer, and that there exists $\theta \in \Theta$ such that min $[\theta N, (1-\theta)N] \geq 2$. With these restrictions, Freddy's "Bayesian" updating always uniquely determines his beliefs in all possible contingencies.[11]

For given priors π, let $\pi_t^N(h_t)$ represent an N-Freddy's posterior beliefs after history of signals h_t. I will use the notation $\pi_t^N(\cdot)$ throughout to represent beliefs by an N-Freddy following the tth signal, but depending on the context, I vary which variables are included as arguments in this function. Bayesian beliefs are $\pi_t^\infty(h_t) \equiv \lim_{N\to\infty} \pi_t^N(h_t)$.

The "gambler's fallacy" is a nearly tautological implication of the model: because an N-Freddy believes that there are only θN a signals and $(1-\theta)N$ b signals when the rate is θ, if in an odd

9. My model closely resembles one previously developed by Rapoport and Budescu [1997]. Their purpose is to explain "production tasks" of the sort discussed in Section II in which people are asked to generate sequences of numbers that look random. They assume that people do so as if they were choosing signals without replacement from an urn, but have memories of what they have done shorter than the size of the urn. Hence, their model of production of signals is a (stochastic and stationary) variant of my model. More broadly, this paper belongs to a small literature developing "quasi-Bayesian" models of biased information processing: a person is modeled as having a specific form of misreading of the world meant to correspond to a heuristic error, but then is assumed to operate as a Bayesian given this misreading. In this sense, it is related to the Barberis, Shleifer, and Vishny [1998] paper discussed in Section V as well as papers like Rabin and Schrag [1999] and Mullainathan [2002] which assume that people have the correct model of the world, but misread or misremember the signals they observe.

10. Actually, I have no recollection of Freddy believing in the law of small numbers. But since most people believe in it, probably Freddy did too.

11. Note that these restrictions require that $N \geq 4$.

period an a signal occurs, then if he is positive he is facing rate θ, he thinks the probability of an a signal next time is less than θ.[12]

LEMMA 1. Consider N and θ such that θN is an integer and $\pi(\theta) = 1$. For all even $t \geq 2$ and histories h_{t-2}, $\pi_t^N(s_t = a | s_{t-1} = b, h_{t-2}) = \theta N/(N-1) > \theta$ and $\pi_t^N(s_t = a | s_{t-1} = a, h_{t-2}) = (\theta N - 1)/(N-1) < \theta$. For all odd t, and histories h_{t-2}, $\pi_t^N(s_t = a | s_{t-1} = b, h_{t-2}) = \pi_t^N(s_t = a | s_{t-1} = a, h_{t-2}) = \theta$.

Lemma 1 shows how the Freddier the person is, the more severe is the gambler's fallacy. It also makes manifest the stark contrast between odd- and even-numbered signals. While this contrast will sometimes assist interpretation of the model by permitting a crisp separation of different effects, the distinction between even and odd periods is of course completely artificial, and will not be a focus in the analysis below.[13]

I believe that this simple model captures the important features of the belief in the law of small numbers, and that the error of thinking that signals are drawn "without replacement" from an urn is, although framed in the wording of an introductory probability course, somewhat close to the underlying psychology of the error people are making. I believe that all of the qualitative results of this paper would extend to more complicated models that capture the key features of belief in the law of small numbers.

A feature of the law of small numbers and of my model that plays a key role in the overconfidence results is that, while Freddy underestimates the likelihood of repetition of any signal given any underlying rate, his surprise relative to a Bayesian is greater the rarer the signal. This follows from the law of small numbers fairly directly because a group of all one signal is less representative of the rate the lower the rate. Most importantly for the formal results, the likelihood ratio of getting two a signals for the two rates θ, θ' is $(\theta N - 1)\theta/(\theta' N - 1)\theta'$, which is greater than the Bayesian likelihood ratio $\theta^2/(\theta')^2$ if and only if $\theta > \theta'$. This is important for the overinference results in the next section

12. All proofs are in the Appendix.

13. To verify that Freddy expects an average proportion of θ a signals if the true state is θ, note that he thinks the probability of getting two a's out of two signals in $\theta \cdot (\theta N - 1)/(N - 1)$, and the probability of getting one a is $2\theta \cdot (1 - \theta)N/(N - 1)$, yielding an average of 2θ a's after two signals.

because it is this predictable bias in the likelihood ratios that leads to the extreme beliefs.[14]

IV. Overinference

The most interesting implications of the law of small numbers come when Freddy is uncertain about the true rate, and makes inferences about the rate from the signals he observes. Suppose, for instance, that an observer believes that there is an equal chance a fund manager can be any of three types, bad, average, or good, who outperforms other mutual funds 1/4, 1/2, or 3/4 of the time, respectively. What does he infer from two successful years in a row by a particular fund? A Bayesian thinks such a sequence occurs with probability $1/4 \cdot 1/4 = 1/16$ for bad funds, $2/4 \cdot 2/4 = 4/16$ for average funds, and $3/4 \cdot 3/4 = 9/16$ for good funds. But an $N = 4$-Freddy believes that the probabilities are $1/4 \cdot 0/3 = 0/12$ if the fund is bad, $2/4 \cdot 1/3 = 2/12$ if the fund is average, and $3/4 \cdot 2/3 = 6/12$ if the fund is good. For each rate, Freddy assigns a lower probability to a streak of two a's than a Bayesian assigns—because he believes that no matter the rate, drawing the first a means there are fewer a's left for the second draw. But more importantly, Freddy's beliefs are too skewed toward believing that the fund is good, since making one less a available for the second draw has a proportionately greater impact when there are fewer a's to begin with. From his priors, Freddy forms probabilistic beliefs about the rate given an observed sequence of signals using a sort of warped Bayes' Law—applying Bayes' Law with his mistaken beliefs about how likely each sequence is given an underlying rate. While a Bayesian believes the probability that the analyst is good is 18/28, Freddy believes that the probability is $21/28 > 18/28$.

Freddy's beliefs following longer sequences can also be calculated. While Freddy wrongly believes that there is negative correlation within odd-even pairs of signals, he believes consecutive odd-even pairs of signals *are* distributed i.i.d. Hence, there is a reasonably simple formula determining Freddy's beliefs as a function of the number of aa, ab, and bb pairs of signals he observes. (Throughout the paper, the odd-even pair ab is meant to

14. This assumption corresponds to the language widely used in research on the law of small numbers, but because there are relatively few experiments for rates other than $\theta = 1/2$, to my knowledge this feature has relatively less empirical support than the other features.

be unordered, representing both ab and ba.) Suppose that, after either $2(q + r + s)$ or $2(q + r + s) + 1$ signals, Freddy observes q aa pairs, r ab pairs, and s bb pairs, in some fixed order, possibly followed by one unpaired signal, consisting of $y \in \{0,1\}$ a signals and $1 - y$ b signals. Freddy's beliefs about the likelihood of that particular sequence if the rate is θ are given by the following lemma.

LEMMA 2. Freddy believes that state θ generates the ordered sequence of q aa pairs, r ab ($=ba$) pairs, s bb pairs, followed by $y \in \{0,1\}$ isolated a signals or $z \in \{0,1\}$, $z \leq 1 - y$, isolated b signals with probability $\pi_x^N(q,r,s,y,z|\theta) = (\theta \cdot (\theta N - 1)/(N - 1))^q(2\theta \cdot (1 - \theta)N/(N - 1))^r((1 - \theta) \cdot ((1 - \theta)N - 1)/(N - 1))^s\theta^y(1 - \theta)^z$, where $x = 2(q + r + s) + y + z$.

The formula in Lemma 2 is derived directly from the formula in Lemma 1, and is a generalization of the above example. Implicit in Lemma 2 is the key intuition for how Freddiness affects belief formation and hence the core intuition in the paper: in each even period, Freddy thinks that one of whatever signal he observed in the previous period has been removed from the urn, making that signal less likely. If that signal *does* occur again, then Freddy exaggerates how strongly it indicates that the true rate is one that generates many such signals.

Given his theory of how signal sequences are generated, Freddy's beliefs are formed by plugging the formula in Lemma 2 into Bayes' Law.

LEMMA 3. Freddy's beliefs about the likelihood of rate θ^* following a (q,r,s,y,z) sequence of signals are

$$\pi_x^N(\theta^*|q,r,s,y,z) = \frac{\pi_x^N(q,r,s,y,z|\theta^*)\pi(\theta^*)}{\sum_{\theta\in\Theta} \pi_x^N(q,r,s,y,z|\theta)\pi(\theta)},$$

where $x = 2(q + r + s) + y + z$ and $\pi_x^N(q,r,s,y,z|\theta)$ is derived in Lemma 2.

A simple and uninteresting result is that Freddy's inferences from the first signal are undistorted.

LEMMA 4. For all N, and π, for all θ^*,

$$\pi_1^N(\theta^*|s_1 = a) = \frac{\theta^*\pi(\theta^*)}{\sum_{\theta\in\Theta} \theta\pi(\theta)}$$

and

$$\pi_1^N(\theta^*|s_1 = b) = \frac{(1-\theta^*)\pi(\theta^*)}{\sum_{\theta\in\Theta}(1-\theta)\pi(\theta)}.$$

That is, Freddy's beliefs after one signal are the same as a Bayesian's with the same priors. Combining this result with Lemma 1, Lemma 5 says the "gambler's fallacy" still holds for the second signal despite uncertainty about the rate.

LEMMA 5. For all N and π, for all θ^*, $\pi_1^N(s_2 = a|s_1 = a)$ and $\pi_1^N(s_2 = b|s_1 = b)$ are increasing in N.

Three comments help to interpret Lemma 5 and the other comparative statics on N presented in the paper. First, because as $N \to \infty$ Freddy becomes a Bayesian, a result on how the degree of Freddiness affects beliefs is also a comparison of Freddies to Bayesians. When the number is increasing in N, it means that Freddy has a lower value for that number than does a Bayesian. Second, the wording in all these results is loose, since changing N in arbitrary ways typically affects the coherence of the model by rendering θN to be nonintegers for some N. But noting that if (π,N_0) is coherent, so is (π,kN_0) for all positive integers k, all comparative statics on N are proved by changing the integer value k for a fixed N_0. Third, most results in this paper can be stated in terms of precise formulas whose qualitative features are of interest. While I state some of the results in terms of these precise formulas (as in Lemmas 2, 3, and 4), I present others, such as Lemma 5, solely in terms of their qualitative features. The proofs contain the precise formulas.

The first "overinference" result is that Freddy infers too much about the likely extremity of the rate from an extreme sequence of signals. Proposition 1 formalizes the result by showing that when all of the signals are of one type, Freddy exaggerates the relative likelihood of rates that are most likely to generate that signal. Let h_t^a be the sequence of t a signals, and h_t^b be the sequence of t b signals. Then:

PROPOSITION 1. For all $t > 1$, and θ, $\hat{\theta} \in \Theta$ such that $\theta > \hat{\theta}$, $\pi_t^N(\theta|h_t^a)/\pi_t^N(\hat{\theta}|h_t^a)$ and $\pi_t^N(\hat{\theta}|h_t^b)/\pi_t^N(\theta|h_t^b)$ are both strictly decreasing in N.

That is, following an extreme sequence of signals, Freddy's beliefs are skewed toward those rates where the signals are more likely. Note that this implies that Freddy's predictions of signals in odd periods $\pi_t^N(s_{t+1} = a|h_t^a)$ and $\pi_t^N(s_{t+1} = b|h_t^b)$, when the gambler's fallacy does not kick in, are both decreasing in N.

Proposition 1 says that following an extreme sequence of signals Freddy infers too strongly that he is facing an extreme rate. Proposition 2 shows that a similar bias holds after any sequence of signals where exactly half are a's and half b's: in such cases, Freddy exaggerates the likelihood that the true rate is close to 1/2. Let $H_t^{1/2}$ be the set of all t-sequences with exactly the same number of a's and b's. Then Proposition 2 holds.

PROPOSITION 2. For all even t and all $h_t \in H_t^{1/2}$, and for all θ, $\hat{\theta} \in \Theta$ such that either $\theta > \hat{\theta} \geq 1/2$ or $\theta \leq \hat{\theta} < 1/2$, $\pi_t^N(\theta|h_t)/\pi_t^N(\hat{\theta}|h_t)$ is weakly increasing in N.

It turns out that $H_t^{1/2}$, h_t^a, and h_t^b are special types of sequences: it is not true generally that Freddy necessarily exaggerates the likelihood that the true rate resembles the proportion of signals he has received when those signals are mixed. For many long sequences of signals, in fact, the opposite is true. I return to that issue in Section VI. But one form of overinference holds even after observing a long sequence of signals: Freddy is prone to be overconfident that the true rate is consistent with the overall direction of the signals rather than the opposite. Formally, Proposition 3 states that given any symmetric prior distribution π, whenever the majority of signals in the sequence h_t are a's, Freddy will, for any $\theta > 1/2$, exaggerate the relative likelihood that the rate is θ rather than $1 - \theta$.

PROPOSITION 3. For all symmetric π, rates $\theta > 1/2$ such that $\pi(\theta) > 0$, and histories h_t yielding more a signals than b signals, $\pi_t^N(\theta|h_t)/\pi_t^N(1 - \theta|h_t)$ is decreasing in N.

Propositions 1, 2, and 3 characterize Freddy's beliefs following given sequences of signals. But what sequence Freddy observes is itself a (stochastic) function of the true rate. I now turn to characterizing Freddy's possible beliefs as a function of the true rate. This will allow us to compare the mean and variance of Freddy's beliefs with a Bayesian's. Let $E_t^N(h_t) \equiv \sum_{\theta \in \Theta} \pi_t^N(\theta|h_t) \cdot \theta$ be the mean value of Freddy's probabilistic beliefs about the rates following sequence of signals h_t. That is, $E_t^N(h_t) \in [0,1]$ is Freddy's perception given that he has observed h_t of the expected value of the rate, which is also his estimated probability that the signal in the next odd period will be an a. Let $f_{t,\pi}^N$ be the probability distribution over the values of $E_t^N(h_t)$ given the probability distribution π over the rates. While somewhat cumbersome notationally and conceptually, the distributions $f_{t,\pi}^N$ play an important role in intuiting the implications of Freddiness. They represent the distribution of Freddy's expected beliefs about the underlying rate and the actual distribution of rates. If $f_{t,\pi}^N$ has a higher variance than the corresponding Bayesian distribution, $f_{t,\pi}^{\infty}$, then Freddy's beliefs are too dispersed due to overinference.

It follows from Proposition 1 that for all t, $E_t^N(h_t^a)$ is decreasing in N: Freddy has too high an estimate of θ following a string of a signals. Similarly, $E_t^N(h_t^b)$ is increasing in N. Hence, it is trivial that the range of possible beliefs that Freddy might have is decreasing in N: for all π and t, the size of the support of $f_{t,\pi}^N$ is decreasing in N. In this sense the variation in Freddy's beliefs is too great.

Beyond this, little general can be said about the distribution of beliefs. One barrier to saying more reflects a feature of the model that is important in its own right: for many prior probabilities of rates, there may be predictable drift in Freddy's average beliefs as he gets information.[15] To disentangle overinference from such drift in beliefs, many results will be easier to formulate and interpret when limiting analysis to *symmetric* prior beliefs, distributions π such that for all $d \in [0,1/2]$, $\pi(1/2 + d) = \pi(1/2 - d)$. It is easy to verify in such cases that $f_{t,\pi}^N$ will be symmetric with mean 1/2. Indeed, focusing on symmetric distributions, a strong result can be stated for two or three signals. Namely, the distribution of Freddy's expected beliefs after either two or three signals is too dispersed.

PROPOSITION 4. For all symmetric distributions π and all $\hat{N} > N$, $f_{2,\pi}^N$ is a mean-preserving spread of $f_{2,\pi}^{\hat{N}}$, and $f_{3,\pi}^N$ is a mean-preserving spread of $f_{3,\pi}^{\hat{N}}$.

Propositions 1–4 together constitute the main "overinference" results of the paper, which say that Freddy infers too much from a *short* sequence of signals. Propositions 1 and 2 indicate that for all possible combinations of two signals, Freddy believes too strongly that the underlying rate is that which most resembles the observed signals. Propositions 3 and 4 say that the

15. Freddy, of course, does not understand this drift: because he is a Bayesian with the wrong model of the world, he obeys the law of iterated expectations in the sense that his own expectation of future beliefs is his current beliefs.

distribution of the mean of Freddy's beliefs has too high a variance after two or three observations of the signals. Analyzing Freddy's beliefs after longer sequences of signals is considerably more complicated, and I return to it in Section VI.

I turn now to exploring Freddy's beliefs about the distribution of rates when he observes a small number of signals from a large number of *different* rates. For instance, Freddy may form beliefs about the distribution of the quality of funds based on observing a small number of performances from each of many funds.

Let p be Freddy's prior beliefs over the set of probability distributions, π, that might prevail, where $p(\pi)$ is the probability p assigns to π. Let M be the number of signals Freddy observes for each draw of a rate, and assume that Freddy observes infinitely many different draws. I shall refer to each draw of a rate as a *source*. Let $\pi^N_{p,M}(h)$ be Freddy's beliefs about the possible distribution of rates after observing an infinite sequence h of M signals from each source.

If M is very large, results are complicated by the issues I discuss in the next section. Of greater interest is what happens when Freddy observes a large number of sources, but only observes a small number of signals from each source. In fact, when Freddy observes only two signals from each of a large number of sources, it follows from the overinference result in Proposition 3 that Freddy comes to believe that there is more variation in rates than there really is.

PROPOSITION 5. For all $N < \infty$ and symmetric π such that $\pi(\theta = 1) < 1/2$:

1) There exists a strict mean-preserving spread of π, $\hat{\pi}$, such that if $p(\hat{\pi}) > 0$ and $p(\pi) = 1 - p(\hat{\pi})$, then for all h whose proportions of signal pairs match those generated by π, $\pi^N_{p,2}(h)$ will assign probability 1 to $\hat{\pi}$.
2) There does not exist a $\tilde{\pi}$, where π is a strict mean-preserving spread of $\tilde{\pi}$, where $p(\pi) > 0$, such that for any h whose proportion of signal pairs match those generated by π $\pi^N_{p,2}(h)$ assigns positive probability to $\tilde{\pi}$.

Proposition 5 says that when observing two signals per source, Freddy may come to believe that there is more dispersion in rates than there really is, but will never come to believe that there is less dispersion. Hence, when Freddy does not observe many predictions by each fund manager, he will believe in more variation in expertise than really exists.

To illustrate more concretely the intuition of some of the above results, and to facilitate analysis of some economic applications, consider the simple class of symmetric distributions of the sort used in all of the above illustrations: Freddy has symmetric beliefs over probability distribution of $\Theta = \{1/2 - d, 1/2, 1/2 + d\}$, where $d \in (0,1/2)$, and beliefs $\pi(\theta = 1/2 - d) = \pi(\theta = 1/2 + d) = q$ and $\pi(\theta = 1/2) = 1 - 2q$. Let q^* be the true distribution, and let Freddy's prior beliefs over the possible probability distributions over types be $p(q)$.

Of special interest is Freddy's perceived distribution of rates, $\tilde{q}$, as a function of the actual distribution of rates, q, following the observation of a pair of signals from an infinite number of different sources. A stark example will prove useful as a template for many further examples. Suppose that there is a distribution over three possible rates, $\Theta = \{0,1/2,1\}$, and $\pi(\theta = 0) = \pi(\theta = 1) = q$, $\pi(\theta = 1/2) = 1 - 2q$ is the true distribution, for $q \in [0,1/2]$. Consider a 4-Freddy. He understands that rate $\theta = 1$ always generates a's and rate $\theta = 0$ always generates b's. But Freddy differs from a Bayesian in his beliefs about the probability of sequences generated by $\theta = 1/2$. Whereas $\theta = 1/2$ actually generates pairs aa, ab, and bb in proportions 1/4, 1/2, and 1/4, 4-Freddy thinks it generates them in proportions 1/6, 2/3, and 1/6. Hence, Freddy thinks that when the distribution is $\pi(\theta = 0) = \pi(\theta = 1) = \tilde{q}$, he will see aa or bb pairs $\tilde{q} * 0 + (1 - 2\tilde{q}) * 1/6 + \tilde{q} * 0 = 1/6 + 2/3\tilde{q}$ of the time, and ab pairs $2/3 - 4/3\tilde{q}$ of the time. If the distribution is q, then he will actually observe aa pairs $q * 1 + (1 - 2q) * 1/4 + q * 0 = 1/4 + 1/2q$ of the time. Hence, setting $1/6 + 2/3\tilde{q} = 1/4 + 1/2q$, we see that in the limit as Freddy observes two signals each from an infinite number of rates, he will come to believe $\tilde{q} = 1/8 + 3/4q$. Because $1/8 + 3/4q > q$ for all $q \in [0,1/2)$, Freddy exaggerates how common the extreme rates are.

Now suppose that, in the same population of sources, Freddy observes four signals each from an infinite number of sources.[16] Freddy's beliefs about the likelihood of all combinations of four signals are shown in Table II.

To derive what Freddy's eventual beliefs will be from the

16. Studying the case where $\theta = 1$ or $\theta = 0$ obscures the more general possibility that Freddy underestimates the frequency of extreme rates, since in this case he will never observe any surprising countersignals if the true rate is extreme, and hence here he will never come to underestimate the variance in distribution.

TABLE II

	Probability of the sequence if $\theta = 1/2$		
	# Permutations	Actual likelihood	4-Freddy's perceived likelihood
aa aa	1	1/16	$1 \times 1/6 \times 1/6 = 1/36$
bb bb	1	1/16	$1 \times 1/6 \times 1/6 = 1/36$
aa ab	4	4/16	$4 \times 1/6 \times 1/3 = 8/36$
bb ab	4	4/16	$4 \times 1/6 \times 1/3 = 8/36$
ab ab	4	4/16	$4 \times 1/3 \times 1/3 = 16/36$
aa bb	2	2/16	$2 \times 1/6 \times 1/6 = 2/36$

above table, we must confront a possibility not seen in the earlier examples: that Freddy cannot form *any* beliefs to explain the frequency of signal combinations he observes. Freddy expects, for instance, to see twice as many $\{ab,ab\}$ foursomes as $\{aa,ab\}$, irrespective of q, so he is perplexed to see roughly equal numbers. Hence, Freddy must choose among an array of what he thinks are very implausible explanations for the pattern he observes. To see what this inference process involves, imagine that Freddy receives $16 \cdot X$ quadruplets of signals in exactly the expected proportions according to q. It can be shown that Freddy would update his beliefs toward the $\tilde{q}$ that maximizes the likelihood function,

$$L(\tilde{q}) \equiv \left[\left(\frac{1}{36} + \frac{17}{18}\,\tilde{q} \right)^{2(1/16+7/8q)} \cdot Z \cdot (1 - 2\tilde{q})^{7/8-7/4q} \right]^{X},$$

where Z is a term that does not depend on $\tilde{q}$. When he observes 1600 or 16,000 quadruplets in roughly these proportions—which he will—then he comes to believe firmly in $\tilde{q}$. This yields the solution that after a very large number of observations, Freddy believes

$$\tilde{q} = \frac{5}{136} + \frac{63}{68}\,q.$$

Notice that if $q = 0$, then $\tilde{q} = 5/136$: if there is no variation, Freddy comes to believe that there is. Indeed, for all $q < 1/2$ Freddy exaggerates variance. To use an example I will return to, if $q = 1/7$, Freddy will believe $\tilde{q} = 23/136 > 1/7$. This result captures the intuition that Freddy is likely to exaggerate the prevalence of extreme rates.

At the beginning of this section I stressed that the model assumes that Freddy starts with correct priors and has mistaken beliefs due solely to erroneous updating. But the fact that Freddy tends to exaggerate the variance in rates suggests that Freddy's "priors" may not be correct after all: when he has previously inferred the distribution of rates from small numbers of observations of each source, he will have overly dispersed prior beliefs about the rates for each new source he is facing. Hence, Freddy not only overinfers because of a bad updating, but because of bad priors. He'll infer too much about the extreme talent of a new fund manager not only because he'll infer too much from small samples, but because when he hires her he exaggerates how likely it is that she is talented. Indeed, while Lemma 4 indicates that Freddy's inferences after one signal are the same as a Bayesian's, in this setting his beliefs will be overdispersed even after one signal.

To illustrate this point, and to build the foundation for the application in the next section, consider the situation just discussed, where 4-Freddy has observed four signals each from a large number of sources, where $\pi(\theta = 1/2) = 5/7$, $\pi(\theta = 0) = \pi(\theta = 1) = 1/7$, and had original priors with support $\Theta \in \{0,1/2,1\}$. Freddy will believe after observing one a from a new source that the probability he is facing rate $\theta = 1$ is

$$\frac{23/136 \cdot 1}{23/136 \cdot 1 + 90/136 \cdot 1/2} = 23/68 \approx 33.8 \text{ percent},$$

whereas a Bayesian with correct priors of $q = 1/7$, believes she is facing rate $\theta = 1$ with probability

$$\frac{1/7 \cdot 1}{1/7 \cdot 1 + 5/7 \cdot 1/2} = 2/7 \approx 28.6 \text{ percent}.$$

I use this example to discuss the implications of inference about the underlying rate for what is often of greater interest: what Freddy predicts about future signals. Freddy's mispredictions about coming signals implicate not only overinference but also the gambler's fallacy, and the two have opposite implications. The numbers 33.8 percent and 28.6 percent represent beliefs about the rate, not prediction of the next signal. Freddy predicts that the second signal following one a will be an a with probability $23/68 \cdot 1 + 45/68 \cdot 3 \approx 55.9$ percent. A Bayesian, by contrast, predicts that the next signal will be a with probability $2/7 \cdot 1 +$

$5/7 \cdot 1/2 \approx 64.3$ percent. The reason that Freddy underestimates the probability of an ensuing a is because of his belief in the gambler's fallacy; while he underestimates the probability that $\theta = 1/2$, he believes that *if* $\theta = 1/2$, then the next signal will be b with probability 2/3.

Following 2 a's, Freddy will believe he is facing rate $\theta = 1$ with probability

$$\frac{23/136 \cdot 1}{23/136 \cdot 1 + 90/136 \cdot 1/6} = 23/38 \approx 60.5 \text{ percent,}$$

and hence predict a third a with probability $23/38(1) + 15/38(1/2) \approx 80.3$ percent. A Bayesian thinks he is facing $\theta = 1$ with probability

$$\frac{1/7 \cdot 1}{1/7 \cdot 1 + 5/7 \cdot 1/4} = 4/9 \approx 44.4 \text{ percent,}$$

and hence predicts a third a with probability $4/9(1) + 5/9(1/2) \approx 72.2$ percent. That is, while Freddy underestimates the probability of a second a in a row, he *exaggerates* the probability of a third a in a row. This is because in predicting the signal following renewal of the urn, the gambler's fallacy does not kick in. But to see that overinference can in fact overpower the gambler's fallacy even when they are both at play, consider Freddy's prediction after three initial a's. Now he thinks he is facing $\theta = 1$ with probability

$$\frac{23/136 \cdot 1}{23/136 \cdot 1 + 90/136 \cdot 1/12} = 46/61 \approx 75.4 \text{ percent,}$$

and hence predicts the next signal to be an a with probability $46/61(1) + 15/61(1/3) \approx 83.6$ percent. The Bayesian thinks three a's implies that he is facing $\theta = 1$ with probability

$$\frac{1/7 \cdot 1}{1/7 \cdot 1 + 5/7 \cdot 1/8} = 8/13 \approx 61.5 \text{ percent,}$$

and predicts a fourth a with probability $8/13(1) + 5/13(1/2) \approx 80.8$ percent. Here, overinference about the rate overwhelms the gambler's fallacy and leads to exaggerated prediction of repeated recent signals. This result does not say that Freddy will persistently exaggerate repetition. For long sequences representing the true underlying rate, Freddy will eventually figure out the true rate or (as shown in Section VI) underestimate its extremity, and hence the gambler's fallacy will reemerge. But the example indicates that Freddy is prone to go from underpredicting repetition of streaks in the short run to overpredicting repetition in the medium run. This pattern may help explain a widely discussed anomaly in financial markets, to which I turn in the next section.

V. A Financial Application

The previous example indicates that the law of small numbers might provide a plausible, intuitive, and parsimonious explanation for a financial anomaly that has recently received attention: De Bondt and Thaler [1990], Barberis, Shleifer, and Vishny [1998], and others show that investors in stock markets and other financial markets seem to 1) underreact in the short term to good and bad news about a firm's financial prospects, but 2) *over*react to in the medium or longer term to such news. The underreaction is evidenced by the fact that, as found in Cutler, Poterba, and Summers [1991], the return in many financial markets is positively autocorrelated from one period to the next, across periods of a month, a quarter, or a year. More studies show that the returns to stocks in the short period following better-than-average unexpected earnings by a firm are higher than the earnings following worse-than-average unexpected earnings, suggesting that the prices of these stocks set by investors are not immediately taking into account good or bad news. But there is an opposite pattern when firms or markets perform consistently well or poorly over a longer horizon. Cutler, Poterba, and Summers [1991] show a slight negative correlation in the returns in markets over horizons of 3–5 years, and Campbell and Shiller [1988] also find that returns are negatively correlated over time. A series of articles beginning with De Bondt and Thaler [1987] have shown that the returns to specific portfolios of stocks that have very poor returns over a period tend to significantly outperform portfolios with good returns, indicating that investors are too pessimistic about the future prospects of portfolios that have performed poorly recently.[17]

17. There is also some suggestive evidence gathered from financial markets constructed in the laboratory. Andreassen and Kraus [1990] test the investment patterns of undergraduates in an artificial laboratory experiment with stock series derived from real world stock performances. Investors in these markets tend to sell after stocks go up and buy when prices fall, consistent with the gambler's fallacy. This evidence is inconclusive for many reasons—such behavior

Several other models have been developed to explain these underreaction/overreaction phenomena. Barberis, Shleifer, and Vishny [1998] construct a quasi-Bayesian model where performance is really a random walk, but where investors believe in either of two false models of the world—that returns are negatively autocorrelated, or that they are positively autocorrelated. They show that, given this false model, investor behavior will track the observed pattern of underreaction and overreaction. Daniel, Hirshleifer, and Subrahmanyam [1998] build a model combining the well-established finding that people tend to be overconfident in the precision of the signals they get with the assumption that people overestimate their own insight.[18] Hong and Stein [1999] explain these patterns by combining that assumption of slow diffusion of firm-specific information with a relaxation of the rational-expectations assumption, and Hong, Lim, and Stein [1998] test and confirm some predictions of this model. Berk, Green, and Naik [1999] and Barberis and Huang [2000] each provide (very different) rational-choice explanations of the observed phenomena.

Belief in the law of small numbers may provide an alternative account of the underreaction/overreaction phenomenon. There are three clear advantages of this account: first, while I do not claim that the observed patterns follow inevitably from belief in the law of small numbers, it does seem to be quite a natural outcome, and hence tightly connects the financial anomalies to a far more general psychological phenomenon. Second, while the existing quasi-Bayesian models all employ two separate psychological phenomena, this account can provide a relatively simple, unified model derived from one psychological bias. Finally, the model connects the underreaction/overreaction phenomenon with the more general and potentially more important fictitious-variation bias in financial markets. In a sense, the overreaction can be seen as the time-series analogy of the "cross-sectional" belief

is consistent, for instance, with rational expectations combined with loss aversion and a preference for grabbing gains. De Bondt [1993] finds evidence in a similarly constructed financial market that investors tend to extrapolate recent performance of the market, and also collects survey evidence indicating more directly that investors believe too strongly that trends will continue.

18. Overconfidence could be generated by belief in the law of small numbers. Daniel, Hirshleifer, and Subrahmanyam [1998] invoke the overconfidence in a different way than I do below, by assuming that expert investors overinfer from private signals they receive, rather than from the recent, public performance of the stock.

TABLE III

	$\pi_t^\infty(\theta = 1 \mid \cdot)$	$\pi_t^4(\theta = 1 \mid \cdot)$	$\pi_t^\infty(s_t = a \mid \cdot)$	$\pi_t^4(s_t = a \mid \cdot)$
θa	$2/7 \approx 28.6\%$	$23/68 \approx 33.8\%$	$9/14 \approx 64.3\%$	$19/34 \approx 55.9\%$
θaa	$4/9 \approx 44.4\%$	$23/38 \approx 60.5\%$	$13/18 \approx 72.2\%$	$61/76 \approx 80.3\%$
θba	0%	0%	$1/2 = 50\%$	$1/2 = 50\%$
θaaa	$8/13 \approx 61.5\%$	$46/61 \approx 75.4\%$	$21/26 \approx 80.8\%$	$51/61 \approx 83.6\%$
θbaa	0%	0%	$1/2 = 50\%$	$1/3 \approx 33.3\%$
θaba	0%	0%	$1/2 = 50\%$	$1/3 \approx 33.3\%$
θbba	0%	0%	$1/2 = 50\%$	$1/3 \approx 33.3\%$

that there is more variation in intrinsic corporate performance than there is. Overreaction, in this light, is a manifestation of belief that there is more to learn than there really is from sequences of observations.

To see how the model may lead to the under- and overreaction phenomenon, consider the following environment. All investors live infinitely, and they invest at random in one stock for four months, and then move on to another stock, etc., never reinvesting in earlier stocks. Their eventual beliefs about the distribution of underlying quality of stocks is determined by the model, where an a signal is a positive shock to a firm's value and b is a negative shock, where in actuality these shocks do not predict more positive or negative shocks. A given company lives infinitely, with the same number of potential investors each month observing them, with a turnover rate of 1/4 of investors each period. The performance of each stock is i.i.d., with 5/7 of stocks having underlying quality $\theta = 1/2$, and 1/7 each are $\theta = 1$ and $\theta = 0$.

I now examine average beliefs by investors observing a company as a function of the company's recent history. Consider all the possible histories of the company that next year's investors can observe in which the most recent performance has been an a: $\{\theta a, \theta aa, \theta ba, \theta aaa, \theta baa, \theta aba, \theta bba\}$, where θ represents the fact that the investor did not observe the previous sequence. Table III summarizes the relevant numbers derived from the results calculated above.

From Table III, and from the assumption that equal numbers of new investors regenerate every fourth period, we can compare the beliefs of Bayesians versus 4-Freddies averaged among the four relevant cohorts of observers, about the next signal. Table IV derives the average likelihood that the next signal is an a among

TABLE IV

	0	1	2	3	Bayesian *Avg*	0	1	2	3	4-Freddy *Avg*
bba	1/2	9/14	1/2	1/2	.536	1/2	19/34	1/2	1/3	.473
aba	1/2	9/14	1/2	1/2	.536	1/2	19/34	1/2	1/3	.473
baa	1/2	9/14	13/18	1/2	.591	1/2	19/34	61/76	1/3	.549
aaa	1/2	9/14	13/18	21/26	.668	1/2	19/34	61/76	51/61	.674

investors who have observed 0, 1, 2, or 3 of the most recent performances, as a function of what those performances have been.

While not numerically dramatic, Table IV reveals the pattern discussed: for "short" sequences of recent performance—a streak of one or two a's—an investment pool of Freddies will underreact to the string. But for a "long" sequence—three or more a's in a row—an investment pool of Freddys will overreact: exaggerating the likelihood that the observed firm is good.

I do not claim that this pattern inheres in the logic of the law of small numbers; it depends on the many parameters of the model.[19] But the logic of the model seems to lead naturally to this pattern if the true variance in performance is small and investors do not make too many observations of each firm.

VI. Inference from Long Sequences

Previous sections have for the most part emphasized inference and predictions following only a few signals from one or more sources. I turn now to an exploration of predictions and inference following long sequences. Propositions 1 and 2 show that Freddy overinfers that the rate is extreme from an extreme sequence of signals, and overinfers that the rate is close to 1/2 from a 50/50 sequence of signals. When Freddy has observed just two signals, these are the only types of sequences he can observe. But longer sequences of signals typically do not fall into either of these categories. For many other types of sequences, in fact, Freddy may underestimate the likelihood that the rate closely corresponds to the proportion of signals in the sequence.

19. On the other hand, neither do we know whether empirically observed investor behavior inheres in the logic of the market—it too may exist only because of the prevailing constellation of market parameters.

To illustrate this, consider again an observer who thinks a fund manager might be any of three types, bad, average, or good, outperforming other funds 1/4, 1/2, or 3/4 of the time. Suppose that the manager being observed outperforms others 3/4 of the time. If Freddy knew only about this frequency statistic, then he would reach the obvious and correct conclusion—that this analyst is good. Suppose, however, that Freddy observes the sequence of six successful followed by two unsuccessful performances, $aaaaaabb$. Despite the fact that $6/8 = 3/4$ of the signals are a's, a 4-Freddy perceives this sequence as surely coming from $\theta = 1/2$ rather than $\theta = 3/4$. This is because he thinks an odd-even streak of 2 straight b's is *impossible* when $\theta = 3/4$, when these two signals are drawn from an urn containing 3 a's and 1 b. In a 4-Freddy's mind, good analysts simply are not unsuccessful two years in a row. After observing a good analyst for a long time, Freddy will almost surely observe all possible pairs of signals—two successful years of investing, one successful and one bad, and (less often) two unsuccessful years in a row. Hence, he will almost surely infer that the analyst is average—since the only type of analyst who can have both two unsuccessful years and two successful years in a row are average ones. He believes this despite his surprise that this supposedly average analyst is successful 3/4 of the time.

In fact, quite generally Freddy is prone to believe that the rate is less extreme than it is after observing a very large number of signals. To see this, note that the proportion of aa, ab, and bb pairs given the true rate θ^* and an infinite number of observations equals almost exactly $(\theta^*)^2$ aa's, $2\theta^*(1 - \theta^*)$ ab's, and $(1 - \theta^*)^2 bb$'s. Having received these proportions, therefore, Freddy thinks that this distribution was generated by the rate θ that is most likely to generate such proportions. Lemma 6 derives Freddy's limit beliefs from his maximum-likelihood estimate of the rate.

LEMMA 6. Suppose that the true rate is θ^*. Then $\lim_{t\to\infty}\pi_t^N(\hat{\theta}|\pi) = 1$ for all π such that $\pi(\hat{\theta}) > 0$, where

$$\hat{\theta} = \arg\max_{\theta\in\Theta}\left[\left(\theta\cdot\frac{\theta N - 1}{N-1}\right)^{(\theta^*)^2}\left(2\theta\cdot\frac{(1-\theta)N}{N-1}\right)^{2(1-\theta^*)\theta^*}\times\left((1-\theta)\cdot\frac{(1-\theta)N-1}{N-1}\right)^{(1-\theta^*)^2}\right].$$

Lemma 6 says that as the number of signals observed becomes arbitrarily large, Freddy's beliefs converge to certainty about the rate.[20] More interestingly, it can be shown that $\hat{\theta}$ is never farther away from 1/2 than is θ^*. Combined with the example above, this shows that Freddy never thinks that the rate may be more extreme than it is, but sometimes thinks that it may be strictly less extreme. I state this important corollary to Lemma 6 as Proposition 6.

PROPOSITION 6. Suppose that N is even. If the true rate is $\theta^* > 1/2$, $\lim_{t\to\infty}\pi_t^N(\hat{\theta}) = 1$ for some $\hat{\theta} \in [1/2,\theta^*]$. If the true rate is $\theta^* < 1/2$, $\lim_{t\to\infty}\pi_t^N(\hat{\theta}) = 1$ for some $\hat{\theta} \in [1/2,\theta^*]$. If the true rate is $\theta^* = 1/2$, $\lim_{t\to\infty}\pi_t^N(\hat{\theta} = 1/2) = 1$.

The logic behind these results is that Freddy observes more streaks than he expects, and the unexpected streakiness of the signals can outweigh the mean frequency of signals in determining Freddy's beliefs.[21]

Despite its intuitive basis (once we have retrained our intuitions), my guess is that the "conservatism" identified by Proposition 6 is not that important in pragmatic terms. More generally, there are reasons to be cautious about interpreting the relevance of limit results, and in fact this may be a good juncture to point out an important feature of the model that applies to all the limit results in this paper. When Freddy's limit beliefs are different from a Bayesian's, the difference depends on the fact that Freddy places *exactly* zero probability on the stochastic structure of the world being as it actually is. If Freddy placed any positive probability on the world being i.i.d., then he would eventually come to believe it is i.i.d. This is because the pattern Freddy observes surprises him immensely in almost every case, and all limit results involve Freddy choosing the least implausible of two unlikely explanations, rather than providing an explanation he finds plausible.

One might argue that it would be more realistic to assume that Freddy will eventually figure out that his theory of negative

20. Lemma 6 is of course well posed only if there is a unique maximand, which I shall assume there is.

21. In the model of this paper, with the urn renewal after every two periods, I have not found examples where Freddy's beliefs do not converge to the true rate θ^* when $\min\,[\theta^* N,(1 - \theta^*)N] \geq 1$. I do not know whether there are such examples. But in the considerably more complicated model that has three-period renewal and $N = 100$, $\theta^* = .97$ will eventually be rejected in favor of beliefs that $\theta = .96$, even though Freddy will never observe a sequence that he considers impossible if $\theta = .97$.

autocorrelation is wrong. There are three related reasons why, in my view, this does not render the model irrelevant. First, although obscured by the modeling techniques, the quasi-Bayesian approach is meant as a model of a boundedly rational person. The reasoning needed to correct the error is often as difficult as, or more difficult than, that needed to avoid the error in the first place. Second, the model itself is simplified to keep it tractable for the analysts; in a realistically complicated model identifying the mistake is likely to be much more complicated. Third, empirically people do not correct this error. Existing evidence—typically from either smarter-than-average, more-educated-than-average twenty year olds in reasonably naturalistic experimental settings, or from "the real world"—shows not that these people make these errors before they know better, but rather that they make these mistakes given their past experience. The hypothesis that reasoning will be ubiquitously Bayesian given realistic experience levels *is* what is tested and rejected by these data.[22]

Because the limit results in this paper clearly require Freddy to correctly understand the pattern he is seeing, they are suspect in light of evidence that people are poor at judging such patterns; the model's prediction about inferences people make based on the patterns they observe in long sequences may be misleading because the patterns they "observe" are not the patterns that are there. But one such pattern-recognition bias stands out as having special significance in the context of this paper: the hot-hand fallacy. This is the tendency for people to perceive a "hot hand" (positive autocorrelation) in what are actually i.i.d. sequences of

22. The pattern of signals can only affect Freddy's beliefs, of course, if he observes this pattern. When Freddy does not observe, or does not attend to, the precise sequence in which his signals arrive. In this case, Freddy can only infer from the frequency of the two signals, and hence it can be proved he will eventually discover the true rate. Beyond this, I have little to say about what Freddy believes when he does not observe the sequence of signals. I do not know, for instance, whether Freddy necessarily overinfers the extremity of the rate if he does not observe the sequence of signals. I conjecture but have not proved that the following is true. Consider θ, $\hat{\theta} \in \Theta$ such that $\theta > \hat{\theta}$. For all $\{x_a, x_b\}$ such that

$$\frac{x_a}{x_a + x_b} > \theta, \; \frac{\pi^N_{x_a+x_b}(\theta \mid \{x_a, x_b\})}{\pi^N_{x_a+x_b}(\hat{\theta} \mid \{x_a, x_b\})}$$

is strictly decreasing in N. For all $\{x_a, x_b\}$ such that

$$\frac{x_a}{x_a + x_b} < \hat{\theta}, \; \frac{\pi^N_{x_a+x_b}(\hat{\theta} \mid \{x_a, x_b\})}{\pi^N_{x_a+x_b}(\theta \mid \{x_a, x_b\})}$$

is strictly decreasing in N.

signals. Variants of this misperception sometimes show up in experiments. In "prediction-task" experiments—in which participants predict coming signals as a function of recent signals—there are some cases where the predominant pattern is for people to overpredict continuation of recent signals rather than to commit the gambler's fallacy. The clear majority of experiments, however, are more consistent with the gambler's fallacy.[23]

But belief in the hot hand has been documented much more widely in the field. Gilovich, Vallone, and Tversky [1985] and Tversky and Gilovich [1989a, 1989b] have demonstrated that, while basketball fans believe that basketball players are streak shooters whose "on" and "off" nights cannot be explained by randomness, such a hot hand does not in fact exist (or at least not nearly to the degree that people believe in it). Camerer [1989] shows that organized gambling on basketball games exhibits a small hot-hand bias, insofar as betting indicates a belief that winning streaks and losing streaks are more likely to continue than they actually are.

At first blush, the hot-hand fallacy may seem in contradiction to the gambler's fallacy, since it suggests that people expect to see too many long strings of the same signal rather than too much alternation. Some cases where the hot-hand fallacy prevails do indeed represent an important caveat to the findings of this paper.[24] As many researchers have intuited, however, the hot-hand fallacy may in fact *derive from* the law of small numbers rather than contradict it. The hot-hand fallacy is, in fact, generally interpreted as coming from people's perception that observed streaks are too long to be due to chance. That is, it is precisely because people expect to see more switching among signals than they actually will that they mistake true i.i.d. randomness for streakiness.

A model of how people develop a belief in nonexistent hot hands could be developed building from the model in this paper.

23. Moreover, in experiments with which I am familiar, because they do not control subjects' prior beliefs about the underlying signal probability, predictions by subjects of repetition of signals may simply be inference about the generic likelihood of those signals.

24. But for many questions of economic significance, including those emphasized in this paper, the "long-wave" positive autocorrelation of the hot-hand bias does not undermine the predictions of the gambler's fallacy. For instance, it may not matter much whether somebody who observes an analyst who has recently done well overinfers her intrinsic talent, or merely infers that she is on a hot or cold streak; as long as she thinks the streak is likely to continue for a while, she may treat hot or cold average analysts as if they were good or bad analysts.

In this paper I have assumed that Freddy believes firmly that the string of urns that leads to local representativeness all have the same rate of signal generation. Suppose, however, that he thinks it is possible that the underlying rate of the urn might stochastically change, but do so more rarely than the urn is renewed. This belief in "long-wave" positive autocorrelation may be quite reasonable. There are many plausible explanations, for instance, for why a basketball player might get hot or cold. Maybe when a player is shooting well, he becomes confident, rather than tentative, in taking shots, and this improves his game. When he is doing poorly, he is nervous, and forces bad shots, etc.[25] Indeed, there surely *is* a hot hand in some sports phenomena. But the law of small numbers provides a natural intuition for why somebody who begins with the belief that a stochastic process might or might not involve long-wave positive autocorrelation will over time come to believe in such autocorrelation even when none exists. Faced with actual independence of signals, people develop a bogus belief in a form of positive autocorrelation in signal generation that to them explains the missing negative autocorrelation they expected due to the gambler's fallacy.[26] Such a model would predict the gradual development of belief in the hot hand in those settings—such as basketball shooting—where people find it a priori plausible, but continue believing solely in the gambler's fallacy in contexts where they do not find streakiness plausible.[27] Besides being an important phenomenon in its own right, such a model of the hot-hand fallacy would also counteract the logic discussed above of overly moderate beliefs after long

25. This suggests an intriguing possibility if confidence and underconfidence *do* influence performance: there may be a hot hand if and only if people *believe* in the hot hand. If a person accepts that his performance is i.i.d., he'll never gain nor lose confidence even in the face of streaks, and so his performance will be i.i.d. But if he believes in the hot hand, his fluctuations in confidence will extend what would otherwise be random streaks.

26. For this type of model to work, it would be crucial that Freddy not believe it possible that the urns' rate changes as often as the urns are renewed, since then the countervailing positive and negative autocorrelation he comes to believe in would generate a de facto i.i.d. signal process.

27. The findings in Edwards [1961], discussed in Section II, lend support to this interpretation of the hot-hand fallacy. He finds that in the first 200 trials of a flip of a coin people's predictions correspond to the gambler's fallacy (as he defines it), but that in the last 800 trials their error switches to the hot-hand fallacy. This could be because participants observe less negative autocorrelation than they had predicted given their belief in the law of small numbers, and hence over time came to believe that they were observing more long streaks of signals than they really were.

sequences, so that in such a model Proposition 6 would no longer hold.

VII. Inference and Endogenous Observations

Thus far in the paper I have assumed that which signals Freddy observes is independent of the realization of the signals he has observed earlier. But people often choose what to observe, and do so in part based on the beliefs they have formed from earlier observations. Because belief in the law of small numbers influences how people interpret signals, therefore, it can also influence which signals they observe. I illustrate some potential implications of such endogenous observations in this section.

To illustrate how Freddy's choice of behavior may influence his belief formation, suppose that he believes that there is some variance in the usefulness of interacting with certain people; he thinks, for instance, that some financial analysts provide profitable advice while others do not. Now suppose that Freddy quits employing analysts when he thinks the expected benefit of searching for a better one exceeds the benefit (net of transactions costs) of sticking with the current one.

In such contexts, it is likely that Freddy will come to believe that analysts are worse than they truly are. The intuition is straightforward: Freddy is likely to switch analysts after a short sequence of negative signals and stay with his current analyst after a short sequence of good signals. Because Freddy switches after overinferring that a fund manager is bad, but sticks around to correct his beliefs when he overinfers that a fund manager is good, he will end up exaggerating the prevalence of bad analysts.

Consider again a world where all financial analysts are successful one-half of the time, but Freddy believes it is possible that there are also some analysts who are always successful and others who are never successful. Suppose that Freddy employs an infinite sequence of analysts to help him invest, where he has the opportunity to switch analysts after two signals, and *must* switch after four signals. Assume—crucially—that Freddy observes the performance of only those analysts whom he hires.

Let the signal $s_t = a$ correspond to successful investment by the analyst, and $s_t = b$ correspond to unsuccessful investment. Assume that Freddy wishes to maximize $\sum_{t=1}^{\infty} \delta^t i(s_t)$, where $i(a) = 1$ and $i(b) = 0$. That is, Freddy wishes to maximize the present discounted sum of his money, where he earns more if the analyst he has hired is successful than if she is unsuccessful. Assume that δ is very close to 1, so that Freddy wants to maximize average per-period payoff.[28]

Before investigating which switching behavior Freddy chooses, I first address the question of what he will come to believe given different possible switching behaviors. I denote Freddy's beliefs by $(\tilde{q}_a, \tilde{q}_b)$, where he believes that proportion $\tilde{q}_a$ of the analysts are good ($\theta = 1$), $\tilde{q}_b$ are bad ($\theta = 0$), and $1 - \tilde{q}_a - \tilde{q}_b$ are average ($\theta = 1/2$).

If Freddy never switches after two signals, he will always observe four signals, and Section V shows that his eventual beliefs will be $\tilde{q}_a = \tilde{q}_b = 5/136$. To see what happens when Freddy switches after a bb pair, but not otherwise, notice that out of every sixteen analysts Freddy employs, he observes on average one (aa,aa), one (aa,bb), two (ab,bb), four (aa,ab), four (ab,ab), and four (bb) combinations. The fact that Freddy abandons a fund manager after bb means that one-fourth of the time he will observe just bb from a fund manager, rather than four signals. Eleven of these sixteen sequences involve mixes of a's and b's and hence can only be generated by $\theta = 1/2$. The sequence (aa,aa) can be generated by either $\theta = 1/2$ or $\theta = 1$, and bb can be generated by $\theta = 1/2$ or $\theta = 0$. If 4-Freddy believes that the distribution is (q_a, q_b), he believes that the frequency of the eleven mixed sequences is $29/36(1 - q_a - q_b)$, the frequency of (aa,aa) combinations is $1/36(1 - q_a - q_b) + 1q_a = 1/36(1 + 35q_a - q_b)$, and the frequency of bb pairs is $1/6(1 - q_a - q_b) + q_b = 1/6(1 - q_a + 5q_b)$. From this it can be shown that Freddy will come to believe $\tilde{q}_a = 9/232$ and $\tilde{q}_b = 25/232$.[29] Freddy's beliefs will be the same if he instead switched after both bb and ab, because Freddy will see the same proportions of (aa,aa) and

28. The assumption that Freddy can or must switch after even periods is important, and perhaps leads to misleading conclusions. Suppose that Freddy could switch after one signal, and had to switch after two signals. Then the gambler's fallacy is likely to dominate his behavior. As such, depending on his beliefs, he is likely to voluntarily switch after observing an a signal, but not after observing a b, since he thinks his current analyst is likely to revert to mean. If Freddy were observing many more signals per analyst, and (more importantly) were making decisions about whether to stay with an analyst for a while, then even if the exact timing of Freddy's switch is determined by his belief in the gambler's fallacy, his bigger-scale decision about how long to remain with an analyst will likely be dominated by his belief about the analyst's general merits. Fleshing out this logic in a more complicated model would be difficult, and hence focusing on the two-period/four-period model serves as a useful way to capture these issues.

29. These are the beliefs that maximize $(1 - q_a - q_b)^{11}(1 - q_a + 4q_b)^4(1 + 35q_a - q_b)$.

(bb) combinations as when he switches only on bb; he will see different mixed combinations of signals, but all mixed combinations mean the same thing to Freddy—that he is observing rate $\theta = 1/2$ for sure.

Notice here Freddy forms biased beliefs, $\tilde{q}_a \neq \tilde{q}_b$, because he is switching from "bad" analysts before he has found out that they are not really that bad, but sticking with "good" analysts—long enough to discover that they are not that good. Notice also that now Freddy exaggerates the variance even more than when he exogenously observed four signals: both $\tilde{q}_a$ and $\tilde{q}_b$ have gone up. Intuitively, Freddy is now observing fewer signals from each analyst, which generally raises his perception of variance.

To figure out Freddy's expected average payoffs from different behaviors is somewhat complicated, and has a somewhat complicated connection to his actual payoffs.[30] Letting C be the cost of premature switches, and assuming that forced switches are free, then it can be shown that Freddy's perceptions of payoffs are as follows:

$$\text{If never switch:} \quad 1 + \tilde{q}_a - \tilde{q}_b - \frac{C}{2}$$

$$\text{If switch after } bb\text{:} \quad \frac{11 + 13\tilde{q}_a - 11\tilde{q}_b - 6C}{11 + \tilde{q}_a - 5\tilde{q}_b}$$

$$\text{If switch after } bb \text{ or } ab\text{:} \quad \frac{7 + 17\tilde{q}_a - 7\tilde{q}_b - 6C}{7 + 5\tilde{q}_a - \tilde{q}_b}.$$

30. First, if Freddy switches after the first pair proportion x of the time, then the proportion of pairs, y, that are first pairs will be given by $y = yx + (1 - y)1$, since there is a y chance his pair was new last time, yet he switched, and $1 - y$ chance he was old last time, and forced to switch. Hence, $y = 1/(2 - x)$. If Freddy plans to switch on bb only, then he expects to switch proportion $1/6(1 - \tilde{q}_a - \tilde{q}_b) + \tilde{q}_b$ of the time, so that his perceived proportion of new pairs will be $6/(11 + \tilde{q}_a - 5q_b)$. Freddy perceives the average payoff from a new pair to be $\tilde{q}_a \cdot 2 + (1 - \tilde{q}_a - \tilde{q}_b) \cdot 1 = 1 + \tilde{q}_a - \tilde{q}_b$. He perceives the payoff from an old pair that followed an ab initial pair to be 1. He perceives the expected payoff of an old pair following an initial aa pair to be $(1 + 11\tilde{q}_a - \tilde{q}_b)/(1 + 5\tilde{q}_a - \tilde{q}_b)$. Freddy thinks the proportions of overall pairs that will be 2^{nd} pairs following initial aa and ab pairs will be $(1 + 5\tilde{q}_a - \tilde{q}_b)/(11 + \tilde{q}_a - 5\tilde{q}_b)$ and $(4 - 4\tilde{q}_a - 4\tilde{q}_b)/(11 + \tilde{q}_a - 5\tilde{q}_b)$, respectively. So Freddy's expected average payoff from switching following bb pairs can be calculated to be $(11 + 13\tilde{q}_a - 11\tilde{q}_b - 6C)/(11 + \tilde{q}_a - 5\tilde{q}_b)$. Now suppose that Freddy's strategy is to switch except after aa. Then Freddy believes that proportion $6/(7 + 5\tilde{q}_a - \tilde{q}_b)$ of pairs will be new pairs, with expected payoff $1 + \tilde{q}_a - \tilde{q}_b$, and proportion $(1 + 5\tilde{q}_a - \tilde{q}_b)/(7 + 5\tilde{q}_a - \tilde{q}_b)$ of signals will be 2^{nd} pairs following aa initial pairs, with expected payoffs $(1 + 11\tilde{q}_a - \tilde{q}_b)/(1 + 5\tilde{q}_a - \tilde{q}_b)$. Hence, Freddy's perception of payoffs from this switching strategy will be $(7 + 17\tilde{q}_a - 7\tilde{q}_b - 6C)/(7 + 5\tilde{q}_a - \tilde{q}_b)$. Freddy's expected average payoff from never switching voluntarily will be $1 + \tilde{q}_a - \tilde{q}_b - C/2$.

Finally, these payoffs can be calculated when $\tilde{q}_a = \tilde{q}_b = 5/136$, and for $\tilde{q}_a = 9/232$, $\tilde{q}_b = 25/232$:

	$\tilde{q}_a = \tilde{q}_b = \frac{5}{136}$	$\tilde{q}_a = \frac{9}{232}, \tilde{q}_b = \frac{25}{232}$
Never switch	$\frac{2 - C}{2}$	$\frac{99 - 58C}{116}$
Switch after bb	$\frac{1506 - 816C}{1476}$	$\frac{399 - 232C}{406}$
Switch after ab or bb	$\frac{1002 - 816C}{972}$	$\frac{267 - 232C}{274}$.

From this last table, Freddy's potential switching strategies can be determined. The table can be used to make three relevant statements. When Freddy has beliefs $\tilde{q}_a = \tilde{q}_b = 5/136$, he refrains from voluntary switching if and only if $C > 25/78 \approx .32$. When Freddy has beliefs $\tilde{q}_a = 9/232$ and $\tilde{q}_b = 25/232$, he refrains from switching if and only if $C > 1923/5510 \approx .35$. Finally, if Freddy prefers to switch, then he does so if and only if he sees bb; he never switches following ab or aa. Taking these statements together, when $C < .32$, Freddy switches when the analyst he hires performs at bb, and when $C > .35$, Freddy never voluntarily switches. But when $C \in (25/78, 1923/5510) \approx (.32, .35)$, Freddy strictly prefers never switching *if* he has not been switching, but prefers to switch at bb if he has been switching. This shows that there are two steady-state belief-behavior combinations for the same parameters—one where Freddy switches a lot because he thinks there is variance in expertise that merits shopping around, and one where he does not. This is driven by the endogeneity of beliefs, which would not arise for a Bayesian. Because in one of these steady states Freddy is incurring more-than-necessary search costs, in addition to showing how errors in belief-formation can lead to multiple steady-state belief-behavior combinations, this example illustrates how belief in the law of small numbers may lead to inefficient expenditures by people in pursuit of entirely illusory expert opinions.[31]

31. This, in turn, raises the possibility that seemingly harmful interventions that interfere with choices people make can in fact be beneficial; in the examples here, for instance, raising the switching cost can make Freddy better off.

VIII. Discussion and Conclusion

The model in this paper helps see how several different phenomena logically derive from the same underlying judgmental bias. In doing so, it also ties together the *scale* of these phenomena; the strength of the gambler's fallacy determines the degree of overinference, the scope of the false-variation bias, etc. This tight structure makes the model precise and refutable, and it would be of some interest to see how well the model does in simultaneously explaining the scope of these phenomena in relevant economic circumstances. But I suspect that the simple model of this paper will not calibrate well. As it stands, there is only one parameter of the model—how big an "urn" the person believes in—that provides a degree of freedom in specifying the nature of a person's belief in the law of small numbers. Allowing a more general (and more complicated) model that allows more parameters while preserving the qualitative features of the law of small numbers will likely be needed to allow greater ability to delink the precise scale of the phenomena associated with the law of small numbers.

Other modifications to the model are needed to make the model more realistic. The most obvious is modifying the artificial distinction between even and odd periods. The best way to fix this artificial feature may sometimes be the one I used in the applications above—simply taking an average over the odd and even periods. But for other applications this may not be adequate, and a more stationary model would need to be developed. To apply the law of small numbers more widely, it will of course be important to be able to apply it to continuous-variable models; mutual funds do not merely do better than average or worse than average—but a lot better than average, a little better than average, etc. It is less obvious how to apply the law of small numbers to such environments, however, since little systematic evidence has been gathered on the gambler's fallacy and related biases outside the context of binary variables. It is nonetheless natural to suppose that people believe a very good performance by an average fund is more due for a bad performance than a modestly good performance. Verifying the behavioral validity and developing a tractable model of such a hypothesis would be crucial to developing further applications of the law of small numbers.

Finally, developing a model of the hot hand along the lines discussed in Section VI will be especially important; while I and others have argued that the hot-hand fallacy, just like the gambler's fallacy, is a manifestation of the law of small numbers, it does nonetheless sometimes make opposite predictions as the gambler's fallacy. Developing a careful model of the relationship among these phenomena will therefore be important in many economic applications.

Appendix: Proofs

Proof of Lemma 1. Algebra.

Proof of Lemma 2. Algebra.

Proof of Lemma 3. Bayes Rule.

Proof of Lemma 4. Bayes Rule.

Proof of Lemma 5.

$$\pi_1^N(s_2 = a|s_1 = a) = \sum_{\theta\in\Theta} \pi_1^N(\theta|s_1 = a)\cdot\pi_1^N(s_2 = a|\theta,s_1 = a)$$

$$= \sum_{\theta\in\Theta} \pi_1^N(\theta|s_1 = a)\cdot\frac{\theta N - 1}{N - 1}.$$

Since Lemma 4 says that $\pi_1^N(\theta|s_1 = a)$ is independent of N and for all $\theta \in \Theta$, $\theta N/(N - 1)$ is increasing in N, the lemma is established. Mutatis mutandi for $\pi_1^N(s_2 = b|s_1 = a)$.

Proof of Proposition 1. For even t,

$$\frac{\pi_t^N(\theta|h_t^a)}{\pi_t^N(\hat\theta|h_t^a)} = \frac{(\theta(\theta N - 1)/(N - 1))^{t/2}}{(\hat\theta(\hat\theta N - 1)/(N - 1))^{t/2}} = \left(\frac{\theta(\theta N - 1)}{\hat\theta(\hat\theta N - 1)}\right)^{t/2},$$

which is decreasing in N iff $\theta(\theta N - 1)/(\hat\theta(\hat\theta N - 1))$ is decreasing in N. This is true iff $\theta > \hat\theta$. The argument easily extends for odd t, and for $\pi_t^N(\hat\theta|h_t^b)/\pi_t^N(\theta|h_t^b)$.

Proof of Proposition 2. All $h_t \in H_t^{1/2}$ correspond to some sequence of ab $(=ba)$ pairs and an equal number of aa and bb pairs. Given this and Lemma 2, it is sufficient to show that both $\pi^N(ab|\theta)/\pi^N(ab|\hat\theta)$ and $\pi^N(aa,bb|\theta)/\pi^N(aa,bb|\hat\theta)$ are nondecreasing in N. The first expression equals

$$\frac{\theta(\theta(1 - \theta)N)/(N - 1)}{\hat\theta(1 - \hat\theta)N/(N - 1)} = \frac{\theta(1 - \theta)}{\hat\theta(1 - \hat\theta)}$$

which does not depend on N. The second expression equals

$$\frac{\theta(\theta N-1)/(N-1)(1-\theta)((1-\theta)(N-1)/(N-1))}{\hat{\theta}(\hat{\theta}N-1)/(N-1)(1-\hat{\theta})((1-\hat{\theta})(N-1)/(N-1))}.$$

This is equal to $(\theta(1-\theta)/(\hat{\theta}(1-\hat{\theta})))((\theta(1-\theta)N^2-N+1)/(\hat{\theta}(1-\hat{\theta})N^2-N+1$, which can be shown to be increasing in N iff $(\hat{\theta}(1-\hat{\theta})-\theta(1-\theta))(N-2)>0$. If $N>2$—which it must be to make the model coherent—then this inequality holds iff $\hat{\theta}(1-\hat{\theta})>\theta(1-\theta)$. This is always true when either $\theta>\hat{\theta}>1/2$ or $\theta<\hat{\theta}<1/2$.

Proof of Proposition 3. Since π is symmetric, $\pi_t^N(\theta|h_t)/\pi_t^N(1-\theta|h_t)=\pi^N(h_t|\theta)/\pi^N(h_t|1-\theta)$. If h_t involves a total of r aa pairs, s bb pairs, and t ab pairs, then

$$\begin{aligned}\frac{\pi^N(h_t|\theta)}{\pi^N(h_t|1-\theta)}&=\frac{(\theta\cdot(\theta N-1/N-1))^r((1-\theta)\times((1-\theta)N-1/N-1))^s(\theta((1-\theta)N/N-1))^t}{((1-\theta)((1-\theta)N-1/N-1))^r\times(\theta\cdot(\theta N-1/N-1))^s((1-\theta)(\theta N/N-1))^t}\\&=\left[\frac{\theta(\theta N-1)}{(1-\theta)((1-\theta)N-1)}\right]^r\left[\frac{(1-\theta)((1-\theta)N-1)}{\theta(\theta N-1)}\right]^s\\&=\left[\frac{\theta(\theta N-1)}{(1-\theta)((1-\theta)N-1)}\right]^{r-s}\\&=\left(\frac{\theta}{1-\theta}\right)^{r-s}\left(\frac{\theta N-1}{(1-\theta)N-1}\right)^{r-s}.\end{aligned}$$

We know that if $x_a>x_b$ then $r>s$. Hence, the result is established by the fact that $(\theta N-1)/((1-\theta)N-1)$ is decreasing in N when $\theta>1/2$.

Proof of Proposition 4. It is trivial to show that for all π: $E_2^N(aa)>E_2^N(ab)>E_2^N(bb)$ for all N, $E_2^N(ab)=1/2$ for all N, $E_2^N(aa)$ is decreasing in N, and $E_2^N(bb)$ is increasing in N. If π is symmetric, then for all N $E_2^N(\pi)=1/2$. Since $\pi(aa)=\pi(bb)$, and these values are independent of N, this establishes the proposition. Irrespective of N, the third signal will be used appropriately for updating beliefs, so it will not affect the result. (Once four or more signals occur, it is possible to get a mixture of both aa and bb odd-even pairs.)

Proof of Proposition 5. Let $d(\pi,N)$ be an N-Freddy's beliefs about the proportion each of aa and bb pairs that the symmetric distribution π will generate. First, notice that if π' is a mean-preserving spread of π, then $d(\pi',N) > d(\pi,N)$ for all N (including $N = \infty$). Second, notice that for all π, $d(\pi,N)$ is increasing in N. Notice further yet that if $\hat{\pi} = (1 - k)\pi + k\pi'$ for $k \in (0,1)$, then $d(\hat{\pi},N) = (1 - k)d(\pi,N) + kd(\pi',N)$.

Now choose π generating real distribution $(s,1 - 2s,s)$ of aa, ab, and bb pairs. Since $S = d(\pi,\infty)$, $d(\pi,N) < S$ for all $N < \infty$. Then by choosing any mean-preserving spread, π', of π such that $d(\pi',N) > S$ (π' such that $\pi'(\theta = 1) = \pi'(\theta = 0) = 1/2$ always works), we can choose the $k \in (0,1)$ such that $\hat{\pi} = k\pi' + (1 - k)\pi$ generates $S = kd(\pi',N) + (1 - k)d(\pi,N)$ aa pairs. Since $\hat{\pi}$ is a mean-preserving spread of π whenever π' is, this proves the first part of the proposition. Since any $\tilde{\pi}$ where π is a mean-preserving spread of $\tilde{\pi}$ yields $d(\tilde{\pi},N) < d(\pi,N)$, this can be used to show the second part of the proposition.

Proof of Lemma 6. By the law of large numbers, after an infinite sequence of signals, Freddy will get proportions very close to $(\theta^*)^2$ aa odd-even pairs, $2(1 - \theta^*)\theta^*$ ab pairs, and $(1 - \theta^*)^2$ bb pairs. Hence, Freddy's beliefs will converge to putting full weight on the beliefs $\hat{\theta}$ that maximize the likelihood of observing such proportions. Applying Lemma 2, therefore, we get the maximization stated in this lemma.

Proof of Proposition 6. We must show that the $\hat{\theta}$ that maximizes the likelihood function in Lemma 6 has the specified properties for all θ^* and N. Taking the derivative of the logarithm of the likelihood function in Lemma 6 with respect to θ, we get

$$\frac{\partial L}{\partial \theta} = \frac{(\theta^*)^2}{\theta} + \frac{(\theta^*)^2 N}{\theta N - 1} + \frac{2(1 - \theta^*)\theta^*}{\theta} - \frac{2(1 - \theta^*)\theta^*}{1 - \theta} - \frac{(1 - \theta^*)^2}{1 - \theta} - \frac{(1 - \theta^*)^2 N}{(1 - \theta)N - 1}.$$

It can be shown that $\partial^2 L/\partial\theta^2 < 0$ for all θ^*, N, and θ, by observing that all terms are strictly decreasing in θ. By observing further that the likelihood function in Lemma 6 is zero in the limit as either $\theta \searrow 0$ or $\theta \nearrow 1$, we know that the unique $\hat{\theta}$ is an interior solution satisfying the first-order condition $\partial L/\partial\theta = 0$. Moreover, because $\partial L/\partial\theta$ is strictly decreasing, if we can show that, for $\theta^* > 1/2$, $\partial L/\partial\theta|_{\theta=\theta^*} < 0$, and $\partial L/\partial\theta|_{\theta=1/2} > 0$, we will have established that $\hat{\theta} \in [1/2,\theta^*]$ for even N and $\theta^* > 1/2$.

Algebra shows that $\partial L/\partial\theta|_{\theta=\theta^*} = 1 - 2\theta^* + (\theta^*)^2N/(\theta^*N - 1) - (1 - \theta^*)^2N/((1 - \theta^*)N - 1)$, which can be shown to be negative when $\theta^* > 1/2$. Algebra shows that

$$\left.\frac{\partial L}{\partial\theta}\right|_{\theta=1/2} = 2(\theta^*)^2 - 2(1 - \theta^*)^2 + (\theta^{*2} - (1 - \theta^*)^2)\frac{N}{1/2N - 1},$$

which is positive for all N, when $\theta^* > 1/2$.

To show that $\hat{\theta} \in [\theta^*,1/2]$ when $\theta^* < 1/2$ is the same. And it is easy to establish that when $\theta = 1/2$, $\partial L/\partial\theta = 0$ is solved by $\hat{\theta} = 1/2$.

Department of Economics, University of California at Berkeley

References

Andreassen, P. B., and S. J. Kraus, "Judgmental Extrapolation and the Salience of Change," *Journal of Forecasting,* IX (1990), 347–372.

Barberis, Nicholas, and Ming Huang, "Mental Accounting, Loss Aversion, and Individual Stock Returns," mimeo, University of Chicago, 2000.

Barberis, Nicholas, Andrei Shleifer, and Robert W. Vishny, "A Model of Investor Sentiment," *Journal of Financial Economics,* XLIX (1998), 307–343.

Bar-Hillel, Maya, and Willem A. Wagenaar, "Perceptions of Randomness," *Advances of Applied Mathematics,* XII (1991), 428–454.

Berk, Jonathan, Richard Green, and Vasant Naik, "Optimal Investment, Growth Options, and Security Returns," *Journal of Finance,* LIV (1999), 1553–1607.

Budescu, David V., and Amnon Rapoport, "Subjective Randomization in One- and Two-Person Games," *Journal of Behavioral Decision Making,* VII (1994), 261–278.

Camerer, Colin F., "Do Biases in Probability Judgments Matter in Markets? Experimental Evidence," *American Economic Review,* LXXVII (1987), 981–997.

——, "Does the Basketball Market Believe in the 'Hot Hand'?" *American Economic Review,* LXXIX (1989), 1257–1261.

Campbell, John Y., and Robert J. Shiller, "Stock Prices, Earnings, and Expected Dividends," *Journal of Finance,* XLIII (1988), 661–676.

Clotfelter, Charles, and Philip Cook, "The 'Gambler's Fallacy' in Lottery Play," *Management Science,* XXXIX (1993), 1521–1525.

Cutler, David M., James M. Poterba, and Lawrence H. Summers, "Speculative Dynamics," *Review of Economic Studies,* LVIII (1991), 529–546.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, "Investor Psychology and Security market Under- and Overreactions," *Journal of Finance,* LIII (1998), 1839–1885.

De Bondt, Werner F. M., "Betting on Trends—Intuitive Forecasts of Financial Risk and Return," *International Journal of Forecasting,* IX (1993), 355–371.

De Bondt, Werner F. M., and Richard H. Thaler, "Further Evidence on Investor Overreaction and Stock Market Seasonality," *Journal of Finance,* XLII (1987), 557–581.

De Bondt, Werner F. M., and Richard H. Thaler, "Do Security Analysts Overreact?" *American Economic Review,* LXXX (1990), 52–57.

Edwards, Ward, "Probability Learning in 1000 Trials," *Journal of Experimental Psychology,* LXII (1961), 385–394.

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

Grether, David M., "Bayes' Rule as a Descriptive Model: The Representativeness Heuristic," *Quarterly Journal of Economics,* XCV (1980), 537–557.
——, "Testing Bayes' Rule and the Representativeness Heuristic," *Journal of Economic Behavior and Organization,* XVII (1992), 31–57.
Hong, Harrison, Terrence Lim, and Jeremy Stein, "Bad News Travels Slowly: Size, Analyst Coverage and the Profitability of Momentum Strategies," mimeo, MIT Sloan School, 1998.
Hong, Harrison, and Jeremy Stein, "A Unified Theory of Underreaction, Momentum Trading and Overreaction in Asset Markets," *Journal of Finance,* LIV (1999), 2143–2184.
Kahneman, Daniel, and Amos Tversky, "On the Psychology of Prediction," *Psychological Review,* LXXX (1973), 237–251.
Lindman, Harold, and Ward Edwards, "Supplementary Report: Unlearning the Gambler's Fallacy," *Journal of Experimental Psychology,* LXII (1961), 630.
Metzger, Mary A., "Biases in Betting: An Application of Laboratory Findings," *Psychological Reports,* LVI (1985), 883–888.
Mullainathan, Sendhil, "A Memory Based Model of Bounded Rationality," *Quarterly Journal of Economics,* CXVII (2002), 735–774.
O'Neill, B., "Nonmetric Test of the Minimax Theory of Two-Person Zero-Sum Games," *Proceedings of the National Academy of Sciences,* LXXXIV (1987), 2106–2109.
Rabin, Matthew, and J. Schrag, "First Impressions Matter: A Model of Confirmatory Bias," *Quarterly Journal of Economics,* CXIV (1999), 37–82.
Rapoport, Amnon, and David V. Budescu, "Generation of Random Series in Two-Person Strictly Competitive Games," *Journal of Experimental Psychology: General,* CXXI (1992), 352–363.
Rapoport, Amnon, and David V. Budescu, "Randomization in Individual Choice Behavior," *Psychology: Review,* CIV (1997), 603–617.
Terrell, Dek, "A Test of the Gambler's Fallacy—Evidence from Pari-Mutuel Games," *Journal of Risk and Uncertainty,* VIII (1994), 309–317.
Tversky, Amos, and Daniel Kahneman, "Belief in the Law of Small Numbers," *Psychological Bulletin,* LXXVI (1971), 105–110.
Tversky, Amos, and Thomas Gilovich, "The Cold Facts about the 'Hot Hand' in Basketball," *Chance,* II (1989a), 16–21.
Tversky, Amos, and Thomas Gilovich, "The Hot Hand: Statistical Reality or Cognitive Illusion?" *Chance,* II (1989b), 31–34.