Earnings and hindsight bias

Chelley-steeley, Patricia L.; Kluger, Brian D.; Steeley, James M.

DOI:
10.1016/j.econlet.2015.07.005

Document Version
Peer reviewed version

Citation for published version (Harvard):
Chelley-steeley, PL, Kluger, BD & Steeley, JM 2015, 'Earnings and hindsight bias: an experimental study', *Economics Letters*, vol. 134, pp. 130-132. https://doi.org/10.1016/j.econlet.2015.07.005

Earnings and hindsight bias: An experimental study

Patricia Chelley-Steeley, Brian Kluger, Jim Steeley

PII: S0165-1765(15)00278-5
DOI: http://dx.doi.org/10.1016/j.econlet.2015.07.005
Reference: ECOLET 6819

To appear in: *Economics Letters*

Received date: 23 November 2014
Revised date: 5 July 2015
Accepted date: 6 July 2015

Please cite this article as: Chelley-Steeley, P., Kluger, B., Steeley, J., Earnings and hindsight bias: An experimental study. *Economics Letters* (2015), http://dx.doi.org/10.1016/j.econlet.2015.07.005

- Subjects estimate the probabilities for a series of upcoming events.
- Next, subjects value state claims based on the each event.
- After the events have taken place, subjects recall their probability estimates.
- Subjects' hindsight bias is calculated using initial and recalled probabilities.
- When subjects' earnings are higher, they exhibit more hindsight bias.

EARNINGS AND HINDSIGHT BIAS: AN EXPERIMENTAL STUDY

Patricia Chelley-Steeley
University of Birmingham

Brian Kluger*
University of Cincinnati

Jim Steeley
Aston Business School

Abstract

We conduct prediction experiments where subjects estimate, and later reconstruct probabilities of upcoming events. Subjects also value state-contingent claims on these events. We find that hindsight bias is greater for events where subjects earned more money.

JEL Code: G02

Keywords: Hindsight Bias, Experimental Finance

*Department of Finance
Lindner College of Business
University of Cincinnati Cincinnati, Ohio 45221-0195 USA
+1 513 556 1688
+1 513 556 0979 fax
brian.kluger@uc.edu

EARNINGS AND HINDSIGHT BIAS: AN EXPERIMENTAL STUDY

1. Introduction

Hindsight bias may be an important determinant of behavior in financial markets. Fischhoff and Beyth (1975) discovered hindsight bias, a propensity for subjects to remember or reconstruct probabilities in such a way as to make them seem to have "known it all along." Hindsight bias has been linked to overconfidence (Roese and Vohs (2012)), and overconfidence may be an especially important investment bias ((Shiller (2000), Hirshleifer (2001)). Our experiment extends the hindsight bias literature to a market-oriented environment, and our results show that subjects exhibit a greater degree of hindsight bias when they earn more, suggesting that superior investment performance leads to increased hindsight bias. This may in turn produce overconfidence.

We design a prediction experiment where subjects are asked to estimate probabilities and then to value a series of state-contingent claims that pay according to the outcomes of a series of future events. First, subjects are asked to estimate the probabilities for the events, each of which will be resolved over the next two weeks. Subjects then value state-contingent claims. After the events are resolved, subjects are invited back to collect their earnings, and are again asked to recall their earlier estimates. Differences between the initial and post-event probability estimates are used to calculate the existence and magnitude of the hindsight bias. We find that many participants exhibit hindsight bias, and we present evidence that the magnitude of the hindsight bias is greater when subjects earn very high profits.

The remainder of the paper is organized as follows. Section 2 presents the experimental design. Results are reported in Section 3. Discussion and concluding remarks follow in Section 4.

2. Experimental Design

Our experiments begin with a questionnaire directing subjects to estimate the probability of specified outcomes for a series of public events due to happen over the coming two-week period. These events are related to financial markets, sporting events, weather, and other events with high media exposure.

TABLE 1
An example of events used in the study.
By necessity, the events differ across sessions.

Event	Category
That the maximum temperature in Rome, Italy will exceed the maximum temperature in Madrid, Spain on Saturday May 25th.	Weather
That the film The Great Gatsby will be in the top 3 of the US box office chart for the weekend of 25th to 26th May.	Entertainment
That the New York Yankees baseball team will beat the New York Mets baseball team, at Yankee stadium on May 29th.	Sports

The next phase of our experiment consisted of twenty periods. At the start of each period an event was announced, selected from one of the events on the initial questionnaire. Subjects then value an asset based on that event using the Becker, Degroot and Marschak (1964) valuation-elicitation technique (BDM). The assets are state-contingent claims paying 100 trading dollars if the announced outcome occurs, or nothing otherwise. The BDM technique was implemented using a custom program written in z-Tree (Fischbacher, 2007). After subjects value the asset, the program selects a uniform random number from the interval [0,100]. If the random number is greater than the subject's value, then the subject is given trading dollars equal to the random number. But if the random number is less than or equal to the subject's value, then the subject is given

one share of the state-contingent claim. Therefore, each period, the subject will have either one asset certificate, or a cash balance corresponding to the random uniform number.

Approximately two weeks later, subjects received an email reminding them to collect their payment. Prior to payment, the follow-up questionnaire was administered. With the outcomes as well as their earnings for each of the events in hand, subjects were invited to recall the probabilities that they assigned earlier to the events, or to reconstruct their earlier estimates.

3. Results

The participants in the market experiments were students at Aston Business School, with most from either the MSc Finance and Investments program or the MBA program. The initial sessions took approximately ninety minutes including time for reading the instructions, questions, and some practice rounds. Forty-six subjects in total have participated, with average earnings at 22.55 GBP, and maximum earnings at 34.67 GBP.[1]

We now document the existence of hindsight bias in our sessions. Our measure of hindsight bias uses the Brier probability score (Brier (1950)), a measure of forecast accuracy, as a building block. The Brier score requires probability estimates for a series of N events. The Brier score (*PS)* is calculated using:

[1] One subject was excluded from the analysis because s/he put either 0% or 100% as the initial probabilities for almost all of the events. The subject may not have understood the concept of probabilities, but even if s/he did, it would be much easier for that subject to reconstruct the initial probabilities.

$$PS = \frac{1}{N} \sum_i (p_i - E_i)^2, \qquad (1)$$

where p is the subject's probability estimate for event i, and E is an indicator variable set to one if the event occurs, or set to zero otherwise. Lower values for the probability score signify greater forecast accuracy. A subject choosing the "correct" probability estimates will minimize the expected value of the score. The Brier Score is the same as the quadratic scoring rule, which is sometimes used in experimental economics to incentivize subjects to elicit probability estimates. Here, we use the Brier solely for measuring forecast accuracy.

Our design requires subjects to make two probability estimates per event. One estimate (IQ) is made via the initial questionnaire, and the other (2W) is made during the follow-up questionnaire. In addition, an implied estimate (BDM) is made indirectly. Since a subject's valuation of each state-contingent claim depends on his or her beliefs about the likelihood of the associated event, we can interpret the elicited value as a risk-neutral probability estimate.

TABLE 2
Comparing IQ, BDM and 2W Brier Scores

Paired t and Wilcoxon signed rank hypothesis tests comparing subjects' IQ, BDM and 2W Brier scores. The asterisk denotes a probability value less than 0.05 (two-tailed).

n = 45	Ho	t	Prob	Wilcoxon SR	Prob
	HS = 2W Brier - IQ Brier = 0	4.13	<.001*	332.5	<.001*
	BDM Brier = IQ Brier	2.81	.007*	223.5	.013*
	BDM Brier = 2W Brier	0.08	0.934	25.50	0.77

Differences between the IQ and 2W estimates provide a measure of hindsight bias (HS). The intuition is straightforward. Hindsight bias means that a subject will reconstruct probabilities to make them seem to be more accurate forecasters. Our results

show that the 2W Brier score is lower than the IQ Brier score, indicating that our subjects, at least in aggregate, are prone to hindsight bias.

We also reject the null that the IQ Brier equals the BDM Brier. There are two possible interpretations. First, recall that the IQ probabilities are not incentivized, but the BDM procedure is. Subjects may be answering more carefully. A second possibility is that the way we are framing the question in the BDM somehow leads to more accurate estimates. In the initial questionnaire, subjects are simply asked to state the probability that an event will occur. In the BDM procedure, subjects are asked to determine the value of a state-contingent claim based on the event.

The final comparison concerns the BDM and 2W Brier scores. It is tempting to interpret differences between the two scores as an implied hindsight bias. However, this interpretation is problematic. The BDM valuation requires subjects to provide a maximum willingness to pay, not a direct probability estimate. In most hindsight bias research, the focus is on probabilities, but it is certainly possible to imagine an analogous effect using values. For example, suppose a subject is asked to estimate next quarter's EPS for a company. After the EPS is released the subject is asked to recall their prior estimate. It may be an interesting topic for future research study whether hindsight bias would differ in magnitude when compared with the typical formulation in probabilities.

A second issue is that in our follow-up questionnaire, subjects are specifically asked to recall their initial estimates, not the values submitted as a part of the BDM procedure. Differences between the BDM and 2W Brier scores do not necessarily indicate hindsight bias.

To study high earnings and hindsight bias, we divide each subject's responses into two subgroups. For each event, earnings can only be between zero and one hundred.[2] For each subject, we calculate our hindsight measure twice, once using those events where earnings were greater than a cutoff, and again using only events where the subject's earnings were less than a cutoff. We perform this exercise for cutoff values of 50, 60, 70, and 80.

TABLE 3
Hindsight Bias and Earnings

Paired t and Wilcoxon signed rank hypothesis tests comparing subjects' hindsight measures, with high and low earnings. The hindsight measure is calculated using the only the events where earnings are greater than the cutoff value, and again using events where earnings are less than the cutoff value. Difference is $HS_{\text{Subject i} \mid \text{events where i's earning} \geq \text{Cutoff}} - HS_{\text{Subject i} \mid \text{events where i's earning} < \text{Cutoff}}$. The null hypothesis is: Difference = 0. All tests are two-tailed.

n=45	Cutoff	Difference	t	Prob	Wilcoxon SR	Prob
	80	.038	2.30	.03*	222.0	0.01*
	70	.025	1.44	0.15	204.0	0.02*
	60	.012	0.63	0.54	28.0	0.76
	50	.015	0.84	0.40	65.5	0.47

We conduct paired tests with the null hypothesis that a subject's hindsight bias is the same with high and low earnings. The null is rejected with the higher cutoff values (70 and 80), providing support for our hypothesis that successful performance is associated with increased hindsight bias. Our subjects are prone to a greater amount of hindsight bias for events where they are very successful, where they earn close to the maximum possible amount.

[2] Aggregating across subjects, earnings were 0 for 29% of the questions answered, and 100 for 11% of the questions answered. Earnings were above 50 for 53%, above 60 for 46%, above 70 for 38% and above 80 for 30% of the questions.

4. Summary

Psychologists have discovered that individuals are often subject to hindsight bias, a tendency to remember or reconstruct probabilities in such a way as to make them seem to have "known it all along." We measure the hindsight bias as the "improvement" in forecast accuracy between the initial probability estimates and the reconstructed estimates after the event outcomes are known. Our first results pertain to forecast accuracy. We detect hindsight bias in many subjects. Further, we find that valuation of the state contingent claim helps subjects to improve forecast accuracy over estimates based on initial questionnaires.

Our next results concern hindsight bias and performance. Subjects' hindsight bias is greater for events where they earned greater earnings. Perhaps high earnings confirm the notion that a subject has known it all along. Our experiments are a first step in understanding the dynamic properties of behavioral biases, and support the existence of a link between hindsight bias and performance. Better performance is associated with a greater degree of hindsight bias.

5. Acknowledgements

We are very grateful to Lucy Ackert, Erik Angner, Richard Deaves, Steven Jordan, Ulrich Schmidt and David Stolin, as well as seminar participants at both McMaster University and University of Virginia for helpful comments and conversations.

6. References

Becker, Gordon M., Morris H. DeGroot, and Jacob Marschak: 1964, "Measuring Utility by a Single- response Sequential Method." *Behavioral Science* 9, 226– 232.

Brier, Glenn: 1950, Verification of forecasts expressed in terms of probabilities, *Monthly Weather Review*, 78, 1–3.

Fischbacher, Urs: 2007, Z-Tree: Zurich Toolbox for Ready-made Economic Experiments, *Experimental Economics* 10(2), 171-178.

Fischhoff, Baruch and Ruth Beyth: 1975, I Knew it Would Happen: Remembered Probabilities of Once-Future Things, *Organizational Behavior and Human Performance*, 13, 1-16.

Hirshleifer, David: 2001, Investor Psychology and Asset Pricing, *Journal of Finance* 56(4), 1533-1597.

Roese, Neal J. and Kathleen D. Vohs: 2012, Hindsight Bias, *Perspectives on Psychological Science* 7(5), 411–426.

Shiller, Robert J.: 2000, *Irrational Exuberance*. Broadway Books, New York, New York.