Discussion of:

Decision-Making under the Gambler's Fallacy

by:

Daniel Chen, Toby Moskowitz and Kelly Shue

Kent Daniel†

†Columbia Business School

RedRock Conference
11 September, 2015

The Gambler's Fallacy

- The gambler's fallacy (GF) is defined as an (incorrect) belief in negative autocorrelation of a non-autocorrelated random sequence.
  - Tversky and Kahneman (1971) and Kahneman and Tversky (1972) argue that the gambler's fallacy is consistent with "local representativeness." It is a belief in the law of small numbers.
  - For example, individuals who believe in the gambler's fallacy believe that after three red numbers appearing on the roulette wheel, a black number is "due," that is, is more likely to appear than a red number.

This Paper's Contribution

- There is a large literature documenting lab evidence consistent with GF.
- However, there is little evidence from real-world settings that GF affects outcomes.
- This paper documents that the decisions of "experts" in the real-world, with large and dramatic effects on outcomes, are are at least partly explained by the gambler's fallacy

The Law of Small Numbers

From Kahneman and Tversky (1972), p 435:

> *The gambler's fallacy, or the negative-recency effect, is a manifestation of the belief in local representativeness. ... if the proportions of the two outcomes are to be preserved in short segments, then a long sequence of one outcome must be followed by the other outcome in order to restore the balance.*

The Law of Small Numbers

From Kahneman and Tversky (1972), p 435:

> *The gambler's fallacy, or the negative-recency effect, is a manifestation of the belief in local representativeness. ... if the proportions of the two outcomes are to be preserved in short segments, then a long sequence of one outcome must be followed by the other outcome in order to restore the balance.*

- So, for example, if you ask someone to write down a long sequence of H/T realizations that would come from flipping an unbiased coin, their sequence will contain too many segments that look like:

  HTTHTHHT,

  and not enough that look like:

  HHHHHTTT.

The Bombing of London

From Kahneman and Tversky (1972):

> *In his* Introduction to Probability Theory*, Feller (1968, p. 160) describes an example which illustrates the erroneous belief in local representativeness. During the intensive bombing of London in the Second World War, it was generally believed that the bombing pattern could not be random, because a few sections of town were hit several times while many others were not hit at all. Thus, the pattern of hits violated local representativeness, and the randomness hypothesis seemed unacceptable.*

The Bombing of London

From Kahneman and Tversky (1972):

> *In his* Introduction to Probability Theory*, Feller (1968, p. 160) describes an example which illustrates the erroneous belief in local representativeness. During the intensive bombing of London in the Second World War, it was generally believed that the bombing pattern could not be random, because a few sections of town were hit several times while many others were not hit at all. Thus, the pattern of hits violated local representativeness, and the randomness hypothesis seemed unacceptable.*

Gambler's Fallacy – Model

The GF model comes from Rabin (2002):

- The case at time t is $y_t \in \{0, 1\}$; $P_t = P(y_t = 1) = \alpha$, and $y_t \perp y_{t-s} \forall s, t$.
  - That is, this is equivalent to drawing a ball from an urn, a fraction α of which are red, *with replacement.*
  - CMS go to a lot of effort to affirm that y_t itself is not negatively serially correlated.
- The agent also observes a noisy signal $S_t \in \{0, 1\}$, which is accurate and informative with probability μ.
- The agent then makes a "decision" D_t on whether the probability that the ball is red is above some threshold.

Gambler's Fallacy – Model

The GF model comes from Rabin (2002):

- The case at time t is $y_t \in \{0, 1\}$; $P_t = P(y_t = 1) = \alpha$, and $y_t \perp y_{t-s} \forall s, t$.
  - That is, this is equivalent to drawing a ball from an urn, a fraction α of which are red, *with replacement.*
  - CMS go to a lot of effort to affirm that y_t itself is not negatively serially correlated.
- The agent also observes a noisy signal $S_t \in \{0, 1\}$, which is accurate and informative with probability μ.
- The agent then makes a "decision" D_t on whether the probability that the ball is red is above some threshold.
- However, the GF agent incorrectly believes that the urn contains N balls, and that the balls are replaced only every 3rd round.

Model Predictions

The predictions of the GF model that CMS test are:

1. For a rational agent, decisions are not serially correlated.

Model Predictions

The predictions of the GF model that CMS test are:

1. For a rational agent, decisions are not serially correlated.
2. For a GF agent, decisions are negatively autocorrelated.

Model Predictions

The predictions of the GF model that CMS test are:

1. For a rational agent, decisions are not serially correlated.
2. For a GF agent, decisions are negatively autocorrelated.
3. A "Moderate" agent's ($\alpha \approx 0.5$) decisions are more serially correlated than an "extreme" agent's decisions.

Model Predictions

The predictions of the GF model that CMS test are:

1. For a rational agent, decisions are not serially correlated.
2. For a GF agent, decisions are negatively autocorrelated.
3. A "Moderate" agent's ($\alpha \approx 0.5$) decisions are more serially correlated than an "extreme" agent's decisions.
4. When S_t is noisier ($\mu \approx 0$), $\rho(D_t, D_{t-1})$ is more negative.

Testing the Model

CMS test the model using three amazing data sets:

- Immigration Judges from Asylum Courts
- Field experiment with Loan Application Reviews
- Major League Baseball Umpires

Magnitudes

The magnitudes of the GF effects are large:

- Immigration Judges:
  - "We estimate judges are up to 3.3 percentage points more likely to reject the current case if they approved the previous case."

Magnitudes

The magnitudes of the GF effects are large:

- Immigration Judges:
  - "We estimate judges are up to 3.3 percentage points more likely to reject the current case if they approved the previous case."
- Loan Officers:
  - "We find that up to nine percent of decisions are reversed due to negative autocorrelation in decisions under the flat incentive scheme among moderate decision-makers"
    - "the effect is significantly smaller under the stronger incentive schemes and among less moderate decision-makers."

Magnitudes

The magnitudes of the GF effects are large:

- Immigration Judges:
  - "We estimate judges are up to 3.3 percentage points more likely to reject the current case if they approved the previous case."
- Loan Officers:
  - "We find that up to nine percent of decisions are reversed due to negative autocorrelation in decisions under the flat incentive scheme among moderate decision-makers"
    - "the effect is significantly smaller under the stronger incentive schemes and among less moderate decision-makers."
- Major League Baseball Umpires:
  - "MLB umpires call the same pitches in the exact same location differently depending solely on the sequence of previous calls."
    - "... umpires are 1.5 percentage points less likely to call a pitch a strike if the previous pitch was called a strike."

GF or SCE?

- A sequential contrast effect (SCE) model could potentially explain the findings here.
  - However, under SCE, lagged quality should be more important than the lagged decision:

$$Y_{i,t} = \beta_0 + \beta_1 Y_{i,t-1} + \beta_2 Quality_{i,t-1} + Controls + \epsilon_{it}.$$

- For asylum cases, the authors find that $\beta_2 > 0$.
- In other words, when quality at $t-1$ was higher last time (controlling for the decision), the expert is more likely to accept in the next decision.

GF or SCE?

- A sequential contrast effect (SCE) model could potentially explain the findings here.
  - However, under SCE, lagged quality should be more important than the lagged decision:

$$Y_{i,t} = \beta_0 + \beta_1 Y_{i,t-1} + \beta_2 Quality_{i,t-1} + Controls + \epsilon_{it}.$$

- For asylum cases, the authors find that $\beta_2 > 0$.
- In other words, when quality at $t-1$ was higher last time (controlling for the decision), the expert is more likely to accept in the next decision.

	Grant Asylum Dummy	
	(1)	(2)
Lag grant	-0.0356***	-0.0352***
	(0.00788)	(0.00785)
Lag case quality	0.00691*	0.00520
	(0.00385)	(0.00360)
p-value lag case quality < 0	0.0367	0.0751
Quality Measure	1	2
N	23981	23973
R^2	0.228	0.228

Out of Sample Data: *The QE Musical Competition*

From Flôres and Ginsburgh (1996), Ginsburgh and Ours (2003), and Ginsburgh (2003).

- "... the best know competition for violin and piano."
- Run using the same format since 1951.

> *Among those who have won the first prize, we single out a few well-known names: Leonid Kogan (1951), Leon Fleisher(1952), Vladimir Askenazy (1956). ... Lazar Berman(1956), Guidon Kremer(1967), and Emmanuel Ax (1972) [were also highly ranked.] (p. 97)*

- The 12 finalists of the competition perform a set of works, including a concerto newly composed for the competition
  - The order of the final appearance is drawn at random.

Queen Elizabeth Musical Competition

Piano competition results – 1952-1991

Competition	*Order of appearance for the following final ranks:*											
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
Piano												
1952	3	8	10	12	6	2	9	11	5	4	1	7
1956	9	6	7	10	4	3	5	2	1	12	11	8
1960	5	10	2	4	7	6	12	8	9	3	1	11
1964	4	10	2	12	8	6	7	11	1	3	5	9
1968	9	2	11	8	6	12	1	5	4	3	7	10
1972	7	10	12	8	1	6	2	11	5	9	3	4
1975	12	11	4	8	7	10	3	6	5	1	9	2
1978	6	3	4	8	9	7	5	2	10	11	12	1
1983	10	3	9	2	4	6	12	5	7	1	8	11
1987	8	7	11	5	2	12	6	3	10	9	4	1
1991	10	11	9	6	3	5	8	2	12	1	4	7

Queen Elizabeth Musical Competition

Piano competition results – 1952-1991

Competition	*Order of appearance for the following final ranks:*											
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
Piano												
1952	3	8	10	12	6	2	9	11	5	4	1	7
1956	9	6	7	10	4	3	5	2	1	12	11	8
1960	5	10	2	4	7	6	12	8	9	3	1	11
1964	4	10	2	12	8	6	7	11	1	3	5	9
1968	9	2	11	8	6	12	1	5	4	3	7	10
1972	7	10	12	8	1	6	2	11	5	9	3	4
1975	12	11	4	8	7	10	3	6	5	1	9	2
1978	6	3	4	8	9	7	5	2	10	11	12	1
1983	10	3	9	2	4	6	12	5	7	1	8	11
1987	8	7	11	5	2	12	6	3	10	9	4	1
1991	10	11	9	6	3	5	8	2	12	1	4	7

- In 9/11 of the competitions, the winner is from the final group of 4.

Queen Elizabeth Musical Competition

Piano competition results – 1952-1991

Competition	*Order of appearance for the following final ranks:*											
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
Piano												
1952	3	8	10	12	6	2	9	11	5	4	1	7
1956	9	6	7	10	4	3	5	2	1	12	11	8
1960	5	10	2	4	7	6	12	8	9	3	1	11
1964	4	10	2	12	8	6	7	11	1	3	5	9
1968	9	2	11	8	6	12	1	5	4	3	7	10
1972	7	10	12	8	1	6	2	11	5	9	3	4
1975	12	11	4	8	7	10	3	6	5	1	9	2
1978	6	3	4	8	9	7	5	2	10	11	12	1
1983	10	3	9	2	4	6	12	5	7	1	8	11
1987	8	7	11	5	2	12	6	3	10	9	4	1
1991	10	11	9	6	3	5	8	2	12	1	4	7

- In 9/11 of the competitions, the winner is from the final group of 4.
- The competition result determines future success:
  - Ginsburgh and Ours (2003) shows that the (random) ordering determines success.
  - The selection by the “expert” judges appears to play a major role in future career success.

Queen Elizabeth Musical Competition

Piano competition results – 1952-1991

Competition	*Order of appearance for the following final ranks:*											
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
Piano												
1952	3	8	10	12	6	2	9	11	5	4	1	7
1956	9	6	7	10	4	3	5	2	1	12	11	8
1960	5	10	2	4	7	6	12	8	9	3	1	11
1964	4	10	2	12	8	6	7	11	1	3	5	9
1968	9	2	11	8	6	12	1	5	4	3	7	10
1972	7	10	12	8	1	6	2	11	5	9	3	4
1975	12	11	4	8	7	10	3	6	5	1	9	2
1978	6	3	4	8	9	7	5	2	10	11	12	1
1983	10	3	9	2	4	6	12	5	7	1	8	11
1987	8	7	11	5	2	12	6	3	10	9	4	1
1991	10	11	9	6	3	5	8	2	12	1	4	7

- The average ranking of the contestants grouped with the winner should, given the random ordering, be 7.

Queen Elizabeth Musical Competition

Piano competition results – 1952-1991

Competition	*Order of appearance for the following final ranks:*											
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
Piano												
1952	3	8	10	12	6	2	9	11	5	4	1	7
1956	9	6	7	10	4	3	5	2	1	12	11	8
1960	5	10	2	4	7	6	12	8	9	3	1	11
1964	4	10	2	12	8	6	7	11	1	3	5	9
1968	9	2	11	8	6	12	1	5	4	3	7	10
1972	7	10	12	8	1	6	2	11	5	9	3	4
1975	12	11	4	8	7	10	3	6	5	1	9	2
1978	6	3	4	8	9	7	5	2	10	11	12	1
1983	10	3	9	2	4	6	12	5	7	1	8	11
1987	8	7	11	5	2	12	6	3	10	9	4	1
1991	10	11	9	6	3	5	8	2	12	1	4	7

- The average ranking of the contestants grouped with the winner should, given the random ordering, be 7.
  - In the data, it is almost 8, with a p-value of 12.5%.
    - Not hugely significant, but consistent with the hypothesis here.

Finance Applications

- The results here are fascinating, in that they show that, in settings with substantial uncertainty, experts rely on heuristics like local representativeness.
- Some work on behavioral asset pricing models incorporate GF:
  - In Barberis, Shleifer, and Vishny (1998), local representativeness (in the "mean reverting" regime) produces momentum.
    - In the "trending" regime, investors believe the firm has "hot hands."
- Social psychology literature shows that authorities/experts can have a large effect on individual decision making.
  - e.g., Milgram (1963) electric shock experiments.
- In the recent financial crisis, markets did perhaps not pay sufficient attention to the incentives of rating agencies.
- If experts make mistakes like GF, and markets don't take this into account, the economic effects could be large.
  - Does GF show up in analyst behavior? rating agencies? FRB chairs? Treasury Secretaries?

References I

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny, 1998, A model of investor sentiment, *Journal of Financial Economics* 49, 307–343.

Flôres, Jr., Renato G, and Victor A Ginsburgh, 1996, The queen elisabeth musical competition: how fair is the final ranking?, *The Statistician* 45, 97–104.

Ginsburgh, Victor, 2003, Awards, success and aesthetic quality in the arts, *The Journal of Economic Perspectives* 17, 99–111.

Ginsburgh, Victor A., and Jan C. van Ours, 2003, Expert opinion and compensation: Evidence from a musical competition, *The American Economic Review* 93, 289–296.

Kahneman, Daniel, and Amos Tversky, 1972, Subjective probability: A judgement of representativeness, *Cognitive Psychology* 3, 430–454.

Milgram, Stanley, 1963, Behavioral study of obedience, *Journal of Abnormal and Social Psychology* 67, 371–378.

Rabin, Matthew, 2002, Inference by believers in the law of small numbers., *Quarterly Journal of Economics* 117, 775–816.

Tversky, Amos, and Daniel Kahneman, 1971, Belief in the law of small numbers, *Psychological Bulletin* 76, 105–110.