BEHAVIORAL ECONOMICS · LECTURE 13

Overconfidence

Niels Boissonnet – Bielefeld University

Master Economics – Winter 2026/27

Before we start: a quick quiz

Instructions

For each question give a **lower bound** x and an **upper bound** y such that you are **90% sure** the true value is above x, and **90% sure** it is below y.

No phones!

Your intervals

1. The length of the river Nile (in km).
2. The height of Mount Everest (in m).
3. The year Bielefeld University was founded.
4. The average distance from the Earth to the Moon (in km).

And one more

5. In which **quarter** of this class will your exam grade be? (top 25%, second, third, bottom 25%)

Learning objectives

By the end of this lecture you should be able to:

- Define and measure **overprecision**, **overplacement** and **overestimation**.
- Derive, in a Bayesian model, when people **over-** or **under**estimate and over- or underplace.
- Explain why “more than 50% think they are better than average” is **not** by itself proof of irrationality, and how this can be tested.
- Analyse two economic consequences: **excess entry** and **three-part tariffs**.

Roadmap

Three faces of overconfidence

Where does overconfidence come from?

Economic consequences

Where we are in the course

The assumptions of the Robinson model

0. **Consistent preferences:** Robinson's choices can be described by a utility function u
1. **Costless optimization:** he always finds the option that maximizes u
2. **Correct beliefs & expected utility:** he knows p and maximizes expected utility
3. **Absolute, self-regarding consumption:** only his own (b, y) enters u
4. **Stable tastes:** the same u in both periods

Take-away

Lectures 10–12 relaxed the *evaluation* step (expected utility) and the knowledge of p. Today: Robinson maximizes expected utility, but with **wrong beliefs** about himself.

Incorrect beliefs

Two familiar views

- **Objective EU:** everyone agrees on the relevant probabilities (lotteries, Lectures 10–11).
- **Subjective EU:** people may hold *different* beliefs, but behave as if they had *some* prior (Lecture 12).

Incorrect beliefs

Two familiar views

- **Objective EU:** everyone agrees on the relevant probabilities (lotteries, Lectures 10–11).
- **Subjective EU:** people may hold *different* beliefs, but behave as if they had *some* prior (Lecture 12).

A third possibility

We know what the correct beliefs are, yet people make **systematic mistakes**.

- Errors in statistical reasoning: base-rate neglect, hot-hand fallacy, *gambler's fallacy* (Lecture 1: after 10 blacks, bet on red?).

Incorrect beliefs

Two familiar views

- **Objective EU:** everyone agrees on the relevant probabilities (lotteries, Lectures 10–11).
- **Subjective EU:** people may hold *different* beliefs, but behave as if they had *some* prior (Lecture 12).

A third possibility

We know what the correct beliefs are, yet people make **systematic mistakes**.

- Errors in statistical reasoning: base-rate neglect, hot-hand fallacy, *gambler's fallacy* (Lecture 1: after 10 blacks, bet on red?).

This lecture

Mistakes about **oneself**: **overconfidence**.

Take-away

"Overconfidence" covers three different biases (Moore and Healy, 2008). They are measured differently and need not go together.

Three kinds of overconfidence

Definitions

Let X_i be i's **actual** performance, X_j that of others, and $\mathbb{E}_i[\cdot]$ i's **beliefs**.

Bias	Definition	Measure
Overestimation	I think I perform better than I do	$\mathbb{E}_i[X_i] - X_i > 0$
Overplacement	I think I rank higher than I do	$(\mathbb{E}_i[X_i] - \mathbb{E}_i[X_j]) - (X_i - X_j) > 0$
Overprecision	I think I know more precisely than I do	hit rate of 80% intervals $< 80\%$

Take-away

Overestimation and overplacement are about the **level** of beliefs; overprecision is about their **spread**.

1

Three faces of overconfidence

Overprecision, overplacement and overestimation: three different biases, each with its own measure and evidence.

Overprecision

The belief that one's information is more precise than it actually is.

Overprecision

The belief that one's information is more precise than it actually is.

Measurement (our quiz)

Give x with $P(\text{value} > x) = 90\%$ and y with $P(\text{value} < y) = 90\%$.

- The **hit rate** is the share of intervals $[x, y]$ that contain the true value (across people or questions).
- **Benchmark:** with well-calibrated beliefs, the hit rate is $90\% - 10\% = 80\%$.

Overprecision

The belief that one's information is more precise than it actually is.

Measurement (our quiz)

Give x with $P(\text{value} > x) = 90\%$ and y with $P(\text{value} < y) = 90\%$.

- The **hit rate** is the share of intervals $[x, y]$ that contain the true value (across people or questions).
- **Benchmark:** with well-calibrated beliefs, the hit rate is $90\% - 10\% = 80\%$.

Our quiz

Nile $\approx 6{,}650$ km; Everest $8{,}849$ m; Bielefeld University 1969; Moon $\approx 384{,}400$ km. Hits per student should average 3.2 out of 4. **Our class:** _____.

Overprecision

The belief that one's information is more precise than it actually is.

Measurement (our quiz)

Give x with $P(\text{value} > x) = 90\%$ and y with $P(\text{value} < y) = 90\%$.

- The **hit rate** is the share of intervals $[x, y]$ that contain the true value (across people or questions).
- **Benchmark:** with well-calibrated beliefs, the hit rate is $90\% - 10\% = 80\%$.

Our quiz

Nile $\approx 6{,}650$ km; Everest $8{,}849$ m; Bielefeld University 1969; Moon $\approx 384{,}400$ km. Hits per student should average 3.2 out of 4. **Our class:** _____.

Evidence

Hit rates are typically **well below** 80%; Soll and Klayman (2004) report rates between 39% and 66%. Moore and Healy (2008): 90% intervals contain the truth only 73% of the time.

Overplacement

The belief that one's rank or relative performance is higher than it actually is.

Surveys

- 37% of professional engineers at a firm placed themselves in the top 5% of performers (Zenger, 1992).
- 93% of American and 69% of Swedish drivers rated themselves more skilful than the median driver in their country (Svenson, 1981).
- 68% of University of Nebraska professors place their teaching in the top 25% (Cross, 1977).

Overplacement

The belief that one's rank or relative performance is higher than it actually is.

Surveys

- 37% of professional engineers at a firm placed themselves in the top 5% of performers (Zenger, 1992).
- 93% of American and 69% of Swedish drivers rated themselves more skilful than the median driver in their country (Svenson, 1981).
- 68% of University of Nebraska professors place their teaching in the top 25% (Cross, 1977).

Tests

Dean and Ortoleva (2019) asked subjects to solve 17 Raven's Progressive Matrices. Predicted own score: 12; predicted average score: 11 ($p = 0.001$).

Overplacement

The belief that one's rank or relative performance is higher than it actually is.

Surveys

- 37% of professional engineers at a firm placed themselves in the top 5% of performers (Zenger, 1992).
- 93% of American and 69% of Swedish drivers rated themselves more skilful than the median driver in their country (Svenson, 1981).
- 68% of University of Nebraska professors place their teaching in the top 25% (Cross, 1977).

Tests

Dean and Ortoleva (2019) asked subjects to solve 17 Raven's Progressive Matrices. Predicted own score: 12; predicted average score: 11 ($p = 0.001$).

Our class (question 5)

_____% expect to be in the top quarter (calibrated: 25%).

Overestimation

The belief that you perform better than you actually do.

Exam grades

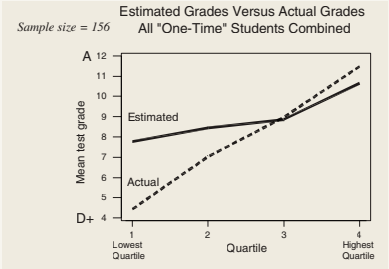
- Students estimate their exam grade right after the exam (Kennedy et al., 2002).
- Grouped by **actual** quartile:
  - bottom quartile: estimate about 8, actual about 4.5;
  - top quartile: slight **under**estimation.

Overestimation

The belief that you perform better than you actually do.

Exam grades

- Students estimate their exam grade right after the exam (Kennedy et al., 2002).
- Grouped by **actual** quartile:
  - bottom quartile: estimate about 8, actual about 4.5;
  - top quartile: slight **under**estimation.

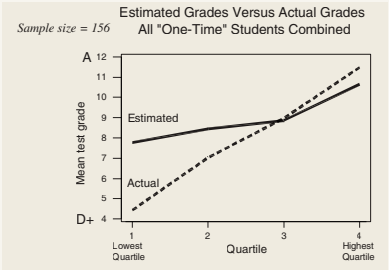


Overestimation

The belief that you perform better than you actually do.

Exam grades

- Students estimate their exam grade right after the exam (Kennedy et al., 2002).
- Grouped by **actual** quartile:
  - bottom quartile: estimate about 8, actual about 4.5;
  - top quartile: slight **under**estimation.



Take-away

The weakest students overestimate the most, the best ones underestimate: the same pattern as Kruger and Dunning (1999). Is it a bias, or **regression to the mean**?

Recap: three faces of overconfidence

Overprecision

Intervals too narrow; hit rates of 40–70% for 80–90% intervals.

Overplacement

Most people think they are above the median.

Overestimation

Low performers overestimate their own performance, high performers underestimate it.

Take-away

Three different biases, three different measures. The next question: are they **mistakes**, or what a Bayesian with noisy information would believe?

2

Where does overconfidence come from?

Bayesian agents with noisy signals show much apparent overconfidence; tests of the Bayesian story reveal a genuine bias.

Two classes of explanation

1. “Rational” overconfidence

- People are uncertain about their ability and update from noisy signals (Bayes’ rule).
- Apparent overconfidence can arise in the *aggregate* even if each belief is correct.
- Possibly combined with simple mistakes in information processing.

Two classes of explanation

1. "Rational" overconfidence

- People are uncertain about their ability and update from noisy signals (Bayes' rule).
- Apparent overconfidence can arise in the *aggregate* even if each belief is correct.
- Possibly combined with simple mistakes in information processing.

2. "Irrational" overconfidence

- **Motivated beliefs**: biases that protect the ego.
- We forget events that make us look bad.
- We discount signals telling us that we are bad.

Two classes of explanation

1. "Rational" overconfidence

- People are uncertain about their ability and update from noisy signals (Bayes' rule).
- Apparent overconfidence can arise in the *aggregate* even if each belief is correct.
- Possibly combined with simple mistakes in information processing.

2. "Irrational" overconfidence

- **Motivated beliefs**: biases that protect the ego.
- We forget events that make us look bad.
- We discount signals telling us that we are bad.

Strategy: work out what Bayesian agents would believe, then compare with the data.

A Bayesian model of confidence (Moore and Healy, 2008)

Assumptions

Person i takes a test. Her score is

$$X_i = S + L_i,$$

- $S \sim \mathcal{N}(m, v_S)$: the **ease** of the test, common to all (low S = hard test);
- $L_i \sim \mathcal{N}(0, v_L)$: i's **ability** relative to others, i.i.d. across people;
- S and the L_i are independent, and everybody knows m, v_S, v_L.

A Bayesian model of confidence (Moore and Healy, 2008)

Assumptions

Person i takes a test. Her score is

$$X_i = S + L_i,$$

- $S \sim \mathcal{N}(m, v_S)$: the **ease** of the test, common to all (low S = hard test);
- $L_i \sim \mathcal{N}(0, v_L)$: i's **ability** relative to others, i.i.d. across people;
- S and the L_i are independent, and everybody knows m, v_S, v_L.

Signal

After the test (before seeing her score), i observes a noisy **signal** about her own score:

$$Y_i = X_i + \varepsilon_i,$$

$\varepsilon_i \sim \mathcal{N}(0, v_\varepsilon)$ independent of S, L_i.

A Bayesian model of confidence (Moore and Healy, 2008)

Assumptions

Person i takes a test. Her score is

$$X_i = S + L_i,$$

- $S \sim \mathcal{N}(m, v_S)$: the **ease** of the test, common to all (low S = hard test);
- $L_i \sim \mathcal{N}(0, v_L)$: i's **ability** relative to others, i.i.d. across people;
- S and the L_i are independent, and everybody knows m, v_S, v_L.

Signal

After the test (before seeing her score), i observes a noisy **signal** about her own score:

$$Y_i = X_i + \varepsilon_i,$$

$\varepsilon_i \sim \mathcal{N}(0, v_\varepsilon)$ independent of S, L_i.

Others' scores: $X_j = S + L_j$. Since L_j has mean 0 and is independent of Y_i, $\mathbb{E}[X_j \mid Y_i] = \mathbb{E}[S \mid Y_i]$.

Bayesian beliefs about oneself and about others

Posterior means

For jointly normal variables, $\mathbb{E}[Z \mid Y] = m_Z + \frac{\mathrm{Cov}(Z,Y)}{\mathrm{Var}(Y)}(Y - m_Y)$, with $\mathrm{Var}(Y_i) = v_S + v_L + v_\varepsilon$:

$$\mathbb{E}[X_i \mid Y_i] = \alpha_X\, m + (1 - \alpha_X)\, Y_i, \qquad \alpha_X = \frac{v_\varepsilon}{v_S + v_L + v_\varepsilon},$$

$$\mathbb{E}[X_j \mid Y_i] = \mathbb{E}[S \mid Y_i] = \alpha_S\, m + (1 - \alpha_S)\, Y_i, \qquad \alpha_S = \frac{v_L + v_\varepsilon}{v_S + v_L + v_\varepsilon}.$$

Bayesian beliefs about oneself and about others

Posterior means

For jointly normal variables, $\mathbb{E}[Z \mid Y] = m_Z + \frac{\mathrm{Cov}(Z,Y)}{\mathrm{Var}(Y)}(Y - m_Y)$, with $\mathrm{Var}(Y_i) = v_S + v_L + v_\varepsilon$:

$$\mathbb{E}[X_i \mid Y_i] = \alpha_X\, m + (1 - \alpha_X)\, Y_i, \qquad \alpha_X = \frac{v_\varepsilon}{v_S + v_L + v_\varepsilon},$$

$$\mathbb{E}[X_j \mid Y_i] = \mathbb{E}[S \mid Y_i] = \alpha_S\, m + (1 - \alpha_S)\, Y_i, \qquad \alpha_S = \frac{v_L + v_\varepsilon}{v_S + v_L + v_\varepsilon}.$$

Both beliefs are a weighted average of the **prior mean** m and the **signal** Y_i: they **regress** towards m.

$\alpha_S \geq \alpha_X$: beliefs about **others** are more regressive, since my signal tells me less about them.

Predictions: hard and easy tests

Average biases

Average over the people taking a test with ease S (signals unbiased, L_i average 0, so X_i and Y_i average S):

Predictions: hard and easy tests

Average biases

Average over the people taking a test with ease S (signals unbiased, L_i average 0, so X_i and Y_i average S):

Overestimation $=$ belief $-$ actual

$= \alpha_X m + (1 - \alpha_X)S - S = \alpha_X (m - S)$.

Predictions: hard and easy tests

Average biases

Average over the people taking a test with ease S (signals unbiased, L_i average 0, so X_i and Y_i average S):

Overestimation $=$ belief $-$ actual
$= \alpha_X m + (1-\alpha_X)S - S = \alpha_X\,(m-S)$.

Overplacement $=$
$\big(\mathbb{E}[X_i \mid Y_i] - \mathbb{E}[X_j \mid Y_i]\big) - (X_i - X_j)$, on average
$(\alpha_S - \alpha_X)(S-m) - 0$.

Predictions: hard and easy tests

Average biases

Average over the people taking a test with ease S (signals unbiased, L_i average 0, so X_i and Y_i average S):

Overestimation $=$ belief $-$ actual
$= \alpha_X m + (1-\alpha_X)S - S = \alpha_X\,(m-S)$.

Overplacement $=$
$\big(\mathbb{E}[X_i \mid Y_i] - \mathbb{E}[X_j \mid Y_i]\big) - (X_i - X_j)$, on average $(\alpha_S - \alpha_X)(S-m) - 0$.

Prediction

	Hard $(S < m)$	**Easy** $(S > m)$
Own score	overestimation	underestimation
Rank	underplacement	overplacement

Predictions: hard and easy tests

Average biases

Average over the people taking a test with ease S (signals unbiased, L_i average 0, so X_i and Y_i average S):

Overestimation $=$ belief $-$ actual $= \alpha_X m + (1-\alpha_X)S - S = \alpha_X\,(m - S)$.

Overplacement $=$ $\big(\mathbb{E}[X_i \mid Y_i] - \mathbb{E}[X_j \mid Y_i]\big) - (X_i - X_j)$, on average $(\alpha_S - \alpha_X)(S - m) - 0$.

Prediction

	Hard $(S < m)$	**Easy** $(S > m)$
Own score	overestimation	underestimation
Rank	underplacement	overplacement

Take-away

Across all tests ($\mathbb{E}[S] = m$) there is **no** bias on average. But on a given test, over- and underconfidence go in **opposite** directions depending on difficulty.

Worked example

Parameters

Let $m = 10$, $v_S = 4$, $v_L = 4$, $v_\varepsilon = 2$. Then $\alpha_X = \frac{2}{10} = 0.2$ and $\alpha_S = \frac{6}{10} = 0.6$. Consider an **average** person ($L_i = 0$) whose signal is exact ($\varepsilon_i = 0$), so $Y_i = X_i = S$.

	Hard test, $S = 6$	**Easy test**, $S = 14$
Actual own score X_i	6	14
Belief about own score $0.2 \cdot 10 + 0.8\,Y_i$	6.8 $(+0.8)$	13.2 (-0.8)
Belief about others $0.6 \cdot 10 + 0.4\,Y_i$	8.4 (actual 6)	11.6 (actual 14)
Believed rank: own − others	-1.6 (underplacement)	$+1.6$ (overplacement)

Worked example

Parameters

Let $m = 10$, $v_S = 4$, $v_L = 4$, $v_\varepsilon = 2$. Then $\alpha_X = \frac{2}{10} = 0.2$ and $\alpha_S = \frac{6}{10} = 0.6$. Consider an **average** person ($L_i = 0$) whose signal is exact ($\varepsilon_i = 0$), so $Y_i = X_i = S$.

	Hard test, $S = 6$	**Easy test**, $S = 14$
Actual own score X_i	6	14
Belief about own score $0.2 \cdot 10 + 0.8\, Y_i$	6.8 $(+0.8)$	13.2 (-0.8)
Belief about others $0.6 \cdot 10 + 0.4\, Y_i$	8.4 (actual 6)	11.6 (actual 14)
Believed rank: own − others	-1.6 (underplacement)	$+1.6$ (overplacement)

A perfectly Bayesian, average student overestimates her score on the hard test **and** believes she is below average.

The predicted pattern and earlier evidence

Prediction

Beliefs about oneself are regressive; beliefs about others even more so (figure: 10-item quiz, prior mean 5).

Earlier evidence

- Earlier studies found both over- and **under**confidence (Stankov and Crawford, 1997), and both over- and **under**placement (Kruger and Dunning, 1999).
- Is this explained by **task difficulty**? Moore and Healy (2008) test it with easy, medium and hard trivia quizzes.

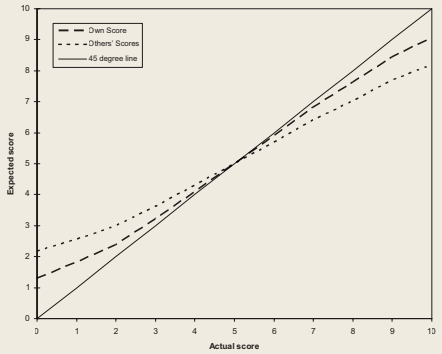


Figure 1. An example of the theory's prediction of beliefs about performance by self and others on a 10-item trivia quiz as a function of the actual score of the person doing the predicting, assuming the person expected a score of 5 prior to taking the quiz.

Moore and Healy (2008): results

Average bias after the quiz

Average bias measured after the quiz (interim phase), overall over 18 rounds; standard deviations in parentheses.

Quiz difficulty	Overestimation	Overplacement	Model predicts
Easy	-0.22 (0.93)	0.48 (2.59)	underestimate, overplace
Medium	0.01 (1.27)	0.04 (3.91)	no bias
Hard	0.79 (1.50)	-1.36 (2.39)	overestimate, underplace

Source: Moore and Healy (2008), Tables 1 and 2 (measured in quiz points).

Moore and Healy (2008): results

Average bias after the quiz

Average bias measured after the quiz (interim phase), overall over 18 rounds; standard deviations in parentheses.

Quiz difficulty	Overestimation	Overplacement	Model predicts
Easy	-0.22 (0.93)	0.48 (2.59)	underestimate, overplace
Medium	0.01 (1.27)	0.04 (3.91)	no bias
Hard	0.79 (1.50)	-1.36 (2.39)	overestimate, underplace

Source: Moore and Healy (2008), Tables 1 and 2 (measured in quiz points).

Findings

- Exactly the predicted sign pattern: overestimation and overplacement move in **opposite** directions.
- But **overprecision** is found everywhere: 90% intervals contain the truth only 73% of the time.

Take-away

Much over- and underestimation and placement is consistent with Bayesian updating; overprecision is not.

Can “93% above the median” be rational? (Benoît and Dubra, 2011)

Three types of drivers

Three skill levels, equally likely (one third each). Each driver knows only whether **she** had an accident.

Skill	High	Medium	Low
$P(\text{accident} \mid \text{skill})$	$1/20$	$9/16$	$47/80$
$P(\text{skill} \mid \text{no accident})$	$19/36 \approx 0.53$	$35/144 \approx 0.24$	$11/48 \approx 0.23$

Can “93% above the median” be rational? (Benoît and Dubra, 2011)

Three types of drivers

Three skill levels, equally likely (one third each). Each driver knows only whether **she** had an accident.

Skill	High	Medium	Low
$P(\text{accident} \mid \text{skill})$	$1/20$	$9/16$	$47/80$
$P(\text{skill} \mid \text{no accident})$	$19/36 \approx 0.53$	$35/144 \approx 0.24$	$11/48 \approx 0.23$

Computation

- Overall accident rate: $\frac{1}{3}\left(\frac{1}{20} + \frac{9}{16} + \frac{47}{80}\right) = \frac{1}{3} \cdot \frac{96}{80} = 0.4.$
- So **60%** of drivers have had no accident, and each of them rationally believes she is in the **top third** with probability $\frac{(1/3)(19/20)}{0.6} = \frac{19}{36} > \frac{1}{2}$.

Can “93% above the median” be rational? (Benoît and Dubra, 2011)

Three types of drivers

Three skill levels, equally likely (one third each). Each driver knows only whether **she** had an accident.

Skill	High	Medium	Low
$P(\text{accident} \mid \text{skill})$	$1/20$	$9/16$	$47/80$
$P(\text{skill} \mid \text{no accident})$	$19/36 \approx 0.53$	$35/144 \approx 0.24$	$11/48 \approx 0.23$

Computation

- Overall accident rate: $\frac{1}{3}\left(\frac{1}{20} + \frac{9}{16} + \frac{47}{80}\right) = \frac{1}{3} \cdot \frac{96}{80} = 0.4.$
- So **60%** of drivers have had no accident, and each of them rationally believes she is in the **top third** with probability $\frac{(1/3)(19/20)}{0.6} = \frac{19}{36} > \frac{1}{2}.$

Take-away

60% of perfectly Bayesian drivers say “I am most likely among the best third”. Survey percentages alone cannot prove overconfidence.

A test of rational overplacement

Prediction (Benoît and Dubra, 2011)

Suppose people state the quantile they are **most likely** in, using Bayes' rule. Then, for every quantile k:

*among those who say "I am in quantile k", the most common **true** quantile is k.*

A test of rational overplacement

Prediction (Benoît and Dubra, 2011)

Suppose people state the quantile they are **most likely** in, using Bayes' rule. Then, for every quantile k:

*among those who say "I am in quantile k", the most common **true** quantile is k.*

Why?

Each person who says k has $P(\text{in } k) \geq P(\text{in } j)$ for all j, given her information. Averaging over all who say k, quantile k must be the most frequent among them.

A test of rational overplacement

Prediction (Benoît and Dubra, 2011)

Suppose people state the quantile they are **most likely** in, using Bayes' rule. Then, for every quantile k:

*among those who say "I am in quantile k", the most common **true** quantile is k.*

Why?

Each person who says k has $P(\text{in } k) \geq P(\text{in } j)$ for all j, given her information. Averaging over all who say k, quantile k must be the most frequent among them.

Data

Burks et al. (2013): about 1,000 trainee truck drivers, IQ and numeracy tests, incentivized beliefs about their own quintile.

Burks et al. (2013): results

The empirical allocation functions $\hat{q}_k(s_j)$

	Numeracy test					IQ test				
	s_1	s_2	s_3	s_4	s_5	s_1	s_2	s_3	s_4	s_5
t_5	0.0	0.0	0.1	0.27	**0.62**	0.004	0.016	0.121	0.271	**0.579**
t_4	0.004	0.009	0.091	**0.298**	0.59	0.0	0.014	0.168	**0.355**	0.461
t_3	0.0	0.0125	**0.181**	0.362	0.443	0.006	0.031	**0.262**	0.375	0.325
t_2	0.004	**0.0**	0.272	0.377	0.345	0.0	**0.04**	0.39	0.363	0.204
t_1	**0.02**	0.02	0.401	0.376	0.175	**0.033**	0.11	0.42	0.322	0.104

Notes: The empirical allocation function indicates for each ability quintile k, what fraction of individual put themselves in ability quintile j.

Rows: true quintile t_k (5 = top). Columns: stated quintile s_j. Entries: share of quintile t_k stating s_j.

Numeracy

- Column s_3: 40% of the **bottom** quintile say "middle", vs. 18% of the true middle quintile.
- So the modal true type of those who say "middle" is **not** middle: the prediction fails.

Burks et al. (2013): results

The empirical allocation functions $\hat{q}_k(s_j)$

	Numeracy test					IQ test				
	s_1	s_2	s_3	s_4	s_5	s_1	s_2	s_3	s_4	s_5
t_5	0.0	0.0	0.1	0.27	**0.62**	0.004	0.016	0.121	0.271	**0.579**
t_4	0.004	0.009	0.091	**0.298**	0.59	0.0	0.014	0.168	**0.355**	0.461
t_3	0.0	0.0125	**0.181**	0.362	0.443	0.006	0.031	**0.262**	0.375	0.325
t_2	0.004	**0.0**	0.272	0.377	0.345	0.0	**0.04**	0.39	0.363	0.204
t_1	**0.02**	0.02	0.401	0.376	0.175	**0.033**	0.11	0.42	0.322	0.104

Notes: The empirical allocation function indicates for each ability quintile k, what fraction of individual put themselves in ability quintile j.

Rows: true quintile t_k (5 = top). Columns: stated quintile s_j. Entries: share of quintile t_k stating s_j.

Numeracy

- Column s_3: 40% of the **bottom** quintile say "middle", vs. 18% of the true middle quintile.
- So the modal true type of those who say "middle" is **not** middle: the prediction fails.

Personality

- Below the median in *social dominance*, 33% think they are in the top 20%; above the median, 55%. In both groups, 20% are.

Bayesian updating alone cannot explain this: **ego and social signalling** matter.

Asymmetric updating: Möbius et al. (2014)

Design

- Subjects take an IQ test. Belief elicited: probability μ of being in the **top half**.

  (Incentivized: choose μ such that you are indifferent between winning \$10 with probability μ and winning \$10 if you are in the top half.)

Asymmetric updating: Möbius et al. (2014)

Design

- Subjects take an IQ test. Belief elicited: probability μ of being in the **top half**.

  (Incentivized: choose μ such that you are indifferent between winning \$10 with probability μ and winning \$10 if you are in the top half.)

- Then 4 signals, each **75% accurate**: if you are in the top half, the signal says "top" with prob. 0.75 and "bottom" with prob. 0.25 (and conversely). Beliefs are elicited after each signal.

Asymmetric updating: Möbius et al. (2014)

Design

- Subjects take an IQ test. Belief elicited: probability μ of being in the **top half**.

  (Incentivized: choose μ such that you are indifferent between winning \$10 with probability μ and winning \$10 if you are in the top half.)

- Then 4 signals, each **75% accurate**: if you are in the top half, the signal says "top" with prob. 0.75 and "bottom" with prob. 0.25 (and conversely). Beliefs are elicited after each signal.

Bayesian benchmark (in odds)

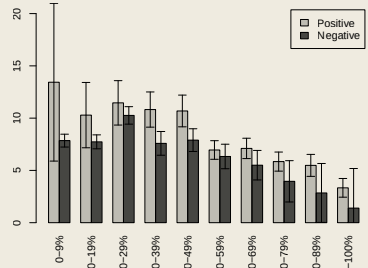
$$\frac{\mu'}{1-\mu'} = \frac{\mu}{1-\mu} \times \begin{cases} 3 & \text{after a "top" signal,} \\ \frac{1}{3} & \text{after a "bottom" signal.} \end{cases}$$

Example

From $\mu = 0.5$, one "top" signal $\Rightarrow \mu' = 0.75$; "top" then "bottom" $\Rightarrow$ back to 0.5.

Möbius et al.: results

Figure 3: Asymmetry



Mean absolute belief revisions by decile of prior belief in being of type equal to the signal received. For example, a subject with prior belief $\hat{\mu} = 0.8$ of being in the top half who received a signal T and a subject with prior belief $\hat{\mu} = 0.2$ who received a signal B are both plotted at $x = 80\%$. T-bars indicate 95% confidence intervals.

Belief revisions

- Mean belief revision after a signal, by decile of prior belief: **light** = positive signal, **dark** = negative.
- People update **more** after good news than after bad news (**asymmetry**), and less than Bayes overall (**conservatism**).
- After 2 positive and 2 negative signals (Bayes: no change), beliefs rise by 4.8 points on average.

Take-away

Good news about ourselves is weighted more than bad news: a **motivated** source of overconfidence.

Recap: where does overconfidence come from?

Statistics

- **Regression to the mean** (Moore and Healy): Bayesian agents overestimate on hard tasks and overplace on easy ones; no bias on average.
- **Rational overplacement** (Benoît and Dubra): most people can rationally believe they are above average.

But the data reject a purely Bayesian story

- overprecision is everywhere;
- the Benoît–Dubra test fails (Burks et al., 2013);
- updating is **asymmetric** (Möbius et al., 2014).

Take-away

Some apparent overconfidence is statistics; a robust core is a genuine bias.

3

Economic consequences

Overconfidence leads to excess entry into markets, and firms design contracts that exploit it.

Excess entry

Business failure

- Many new businesses fail: of US manufacturing firms entering between 1963 and 1982, 62% exited within 5 years and 80% within 10 years (Dunne, Roberts and Samuelson, 1988).
- Is there **excess entry**: too many new firms joining markets?

Excess entry

Business failure

- Many new businesses fail: of US manufacturing firms entering between 1963 and 1982, 62% exited within 5 years and 80% within 10 years (Dunne, Roberts and Samuelson, 1988).
- Is there **excess entry**: too many new firms joining markets?

Overconfidence could cause it

- **overestimation:** “my product is great”;
- **overplacement:** “I am better than the competition”.

Excess entry

Business failure

- Many new businesses fail: of US manufacturing firms entering between 1963 and 1982, 62% exited within 5 years and 80% within 10 years (Dunne, Roberts and Samuelson, 1988).
- Is there **excess entry**: too many new firms joining markets?

Overconfidence could cause it

- **overestimation:** “my product is great”;
- **overplacement:** “I am better than the competition”.

Camerer and Lovallo (1999) test this in the lab, where the “right” amount of entry is known.

The entry game (Camerer and Lovallo, 1999)

Rules

- N players decide simultaneously whether to enter a market with **capacity** c (announced each round).
- Staying out pays 0. Entrants are **ranked**:
  - the top c share \$50 (table);
  - all other entrants **lose \$10**.
- So the total profit of entrants is $50 - 10\,(E - c)$ if $E > c$ entrants: zero when $E = c + 5$.
- Each subject starts with a \$10 stake.

Payoffs

TABLE 1—RANK-BASED PAYOFFS

	Payoff for successful entrants as a function of "c"			
Rank	2	4	6	8
1	33	20	14	11
2	17	15	12	10
3		10	10	8
4		5	7	7
5			5	6
6			2	4
7				3
8				2

Payoff by rank, for capacity $c = 2, 4, 6, 8$.

Design: chance vs. skill

Two conditions

- **Rank by chance:** a random draw. **Rank by skill:** number of correct answers in a logic-puzzle or trivia quiz.
- Each subject plays 12 rounds of each condition. **Ranks are revealed only at the end**, after all entry decisions.

Design: chance vs. skill

Two conditions

- **Rank by chance:** a random draw. **Rank by skill:** number of correct answers in a logic-puzzle or trivia quiz.
- Each subject plays 12 rounds of each condition. **Ranks are revealed only at the end**, after all entry decisions.

Two subject pools

- **standard** recruitment (sessions 1–4);
- **self-selected**: told in advance that earnings depend on skill at sports/current-events trivia (sessions 5–8).

Design: chance vs. skill

Two conditions

- **Rank by chance:** a random draw. **Rank by skill:** number of correct answers in a logic-puzzle or trivia quiz.
- Each subject plays 12 rounds of each condition. **Ranks are revealed only at the end**, after all entry decisions.

Two subject pools

- **standard** recruitment (sessions 1–4);
- **self-selected**: told in advance that earnings depend on skill at sports/current-events trivia (sessions 5–8).

Prediction

If people **overplace** themselves, there is more entry, and lower industry profit, in the skill condition.

Results

TABLE 5—AVERAGE DIFFERENCE IN EXPECTED PROFITS PER ENTRANT BETWEEN RANDOM AND SKILL CONDITIONS

Measure	Experiment 1	Experiment 2	Experiment 3	Experiment 4	Experiment 5	Experiment 6	Experiment 7	Experiment 8	Total
$\Pi_r - \Pi_s$	1.635	0.477	−1.19	0.24	1.62	2.49	3.16	1.80	1.31
	(1.98)	(1.41)	(1.72)	(2.41)	(1.32)	(1.27)	(1.61)	(1.20)	(2.04)
# of S's with	10/12	10/13	3/11	7/14	12/13	12/13	13/13	11/12	78/101
$\Pi_r - \Pi_s < 0$ (percent)	(83)	(77)	(27)	(50)	(92)	(92)	(100)	(92)	(77)
# of S's with	0/12	0/13	0/12	2/15	12/15	15/16	12/14	11/14	52/111
$\Pi_s < 0$ (percent)	(0)	(0)	(0)	(13)	(80)	(94)	(86)	(79)	(47)

Findings

- More entry in the **skill** condition than in the random condition.
- Subjects' own forecasts imply $1.31 lower expected profit per entrant in skill rounds.
- Industry profit lower in skill rounds by $9.14 (standard pool) and by **$27.10** (self-selected pool).

Results

TABLE 5—AVERAGE DIFFERENCE IN EXPECTED PROFITS PER ENTRANT BETWEEN RANDOM AND SKILL CONDITIONS

Measure	Experiment 1	Experiment 2	Experiment 3	Experiment 4	Experiment 5	Experiment 6	Experiment 7	Experiment 8	Total
$\Pi_r - \Pi_s$	1.635	0.477	−1.19	0.24	1.62	2.49	3.16	1.80	1.31
	(1.98)	(1.41)	(1.72)	(2.41)	(1.32)	(1.27)	(1.61)	(1.20)	(2.04)
# of S's with	10/12	10/13	3/11	7/14	12/13	12/13	13/13	11/12	78/101
$\Pi_r - \Pi_s < 0$ (percent)	(83)	(77)	(27)	(50)	(92)	(92)	(100)	(92)	(77)
# of S's with	0/12	0/13	0/12	2/15	12/15	15/16	12/14	11/14	52/111
$\Pi_r < 0$ (percent)	(0)	(0)	(0)	(13)	(80)	(94)	(86)	(79)	(47)

Findings

- More entry in the **skill** condition than in the random condition.
- Subjects' own forecasts imply $1.31 lower expected profit per entrant in skill rounds.
- Industry profit lower in skill rounds by $9.14 (standard pool) and by **$27.10** (self-selected pool).

Take-away

Reference group neglect: entrants think about their own skill, not about the skill of those they compete against. Self-selection makes it worse.

Selling to overconfident consumers (Grubb, 2009)

Setup

A stylized phone-contract example. You are the (monopoly) phone company:

- fixed cost **$50** per consumer, marginal cost **5¢** per minute;
- the consumer values minutes at **45¢** each, up to a satiation point θ, and at 0 beyond;
- **period 1:** contract signed; **period 2:** θ is realized and minutes are used.

Selling to overconfident consumers (Grubb, 2009)

Setup

A stylized phone-contract example. You are the (monopoly) phone company:

- fixed cost **\$50** per consumer, marginal cost **5¢** per minute;
- the consumer values minutes at **45¢** each, up to a satiation point θ, and at 0 beyond;
- **period 1:** contract signed; **period 2:** θ is realized and minutes are used.

Uncertain usage

At signing, θ is unknown:

$$\theta = 100,\ 400 \text{ or } 700 \text{ minutes},$$

each with probability $\frac{1}{3}$ $(\mathbb{E}[\theta] = 400)$.

Selling to overconfident consumers (Grubb, 2009)

Setup

A stylized phone-contract example. You are the (monopoly) phone company:

- fixed cost **\$50** per consumer, marginal cost **5¢** per minute;
- the consumer values minutes at **45¢** each, up to a satiation point θ, and at 0 beyond;
- **period 1:** contract signed; **period 2:** θ is realized and minutes are used.

Uncertain usage

At signing, θ is unknown:

$$\theta = 100,\ 400 \text{ or } 700 \text{ minutes},$$

each with probability $\frac{1}{3}$ $(\mathbb{E}[\theta] = 400)$.

Question: which tariff (fixed fee + price per minute) maximizes profit?

Rational consumer: a two-part tariff

Optimal tariff

- Price at **marginal cost** (5¢): the consumer uses exactly θ minutes, so all efficient trades take place.
- Extract the **expected surplus** with the up-front fee:

Numbers

$$\begin{aligned} F &= \mathbb{E}\big[(45-5)\text{¢} \times \theta\big] \\ &= 0.40 \times \tfrac{1}{3}(100+400+700) \\ &= 0.40 \times 400 = \$160. \end{aligned}$$

Rational consumer: a two-part tariff

Optimal tariff

- Price at **marginal cost** (5¢): the consumer uses exactly θ minutes, so all efficient trades take place.
- Extract the **expected surplus** with the up-front fee:

Numbers

$$\begin{aligned} F &= \mathbb{E}\big[(45-5)\text{¢} \times \theta\big] \\ &= 0.40 \times \tfrac{1}{3}(100+400+700) \\ &= 0.40 \times 400 = \$160. \end{aligned}$$

Profit $= F - \$50 = \$110$ per consumer (the per-minute price exactly covers the per-minute cost).

Rational consumer: a two-part tariff

Optimal tariff

- Price at **marginal cost** (5¢): the consumer uses exactly θ minutes, so all efficient trades take place.
- Extract the **expected surplus** with the up-front fee:

Numbers

$$\begin{aligned} F &= \mathbb{E}\big[(45-5)\text{¢} \times \theta\big] \\ &= 0.40 \times \tfrac{1}{3}(100+400+700) \\ &= 0.40 \times 400 = \$160. \end{aligned}$$

Profit $= F - \$50 = \$110$ per consumer (the per-minute price exactly covers the per-minute cost).

Take-away

With a rational consumer, the monopolist offers a **two-part tariff** with marginal-cost pricing: efficient usage, all surplus extracted up front.

Overconfident consumer: a three-part tariff

Real-world contracts

- Real-world contracts are often **three-part tariffs**: a fixed fee; an allowance of free minutes; a high price beyond.
- Can **overconfidence** explain them?

Overconfident consumer: a three-part tariff

Real-world contracts

- Real-world contracts are often **three-part tariffs**: a fixed fee; an allowance of free minutes; a high price beyond.
- Can **overconfidence** explain them?

Optimal contract under overprecision

- Suppose the consumer is sure that $\theta = 400$ (right mean, no variance): **overprecision**.
- Optimal contract:
  - fee **\$180**,
  - **0¢** for the first 400 minutes,
  - **45¢** per minute beyond 400.

Overconfident consumer: a three-part tariff

Real-world contracts

- Real-world contracts are often **three-part tariffs**: a fixed fee; an allowance of free minutes; a high price beyond.
- Can **overconfidence** explain them?

Optimal contract under overprecision

- Suppose the consumer is sure that $\theta = 400$ (right mean, no variance): **overprecision**.
- Optimal contract:
  - fee **\$180**,
  - **0¢** for the first 400 minutes,
  - **45¢** per minute beyond 400.

The consumer *expects* to use 400 minutes worth $0.45 \times 400 = \$180$ and pay nothing more: she accepts.

Why is this profitable?

One block of minutes at a time

Start from the two-part tariff and change one block of minutes at a time:

- **Minutes 0–100:** price $5 \to 0$¢. Consumer values it at \$5 (+\$5 on the fee); firm loses \$5 for every type. Neutral.
- **Minutes 100–400:** price $5 \to 0$¢. Consumer values it at \$15 (she is sure to use them): fee $+\$15$. Firm loses \$15 only if $\theta \geq 400$: \$10 in expectation. **Gain \$5.**
- **Minutes 400–700:** price $5 \to 45$¢. Consumer thinks she never uses them: fee unchanged. Firm earns $300 \times 0.40 = \$120$ if $\theta = 700$: \$40 in expectation. **Gain \$40.**

Why is this profitable?

One block of minutes at a time

Start from the two-part tariff and change one block of minutes at a time:

- **Minutes 0–100:** price $5 \to 0¢$. Consumer values it at \$5 (+\$5 on the fee); firm loses \$5 for every type. Neutral.
- **Minutes 100–400:** price $5 \to 0¢$. Consumer values it at \$15 (she is sure to use them): fee $+\$15$. Firm loses \$15 only if $\theta \geq 400$: \$10 in expectation. **Gain \$5.**
- **Minutes 400–700:** price $5 \to 45¢$. Consumer thinks she never uses them: fee unchanged. Firm earns $300 \times 0.40 = \$120$ if $\theta = 700$: \$40 in expectation. **Gain \$40.**

Comparison

	Two-part	Three-part
Fee	\$160	\$180
Expected profit	\$110	\$155
Consumer's *actual* expected surplus	\$0	−\$45

Take-away

The firm sells the minutes the consumer **overvalues** cheaply and charges a lot for the ones she **undervalues**.

Recap: economic consequences

Excess entry

Overplacement and reference group neglect lead to too many entrants; self-selection makes it worse.

Contract design

A monopolist facing overprecise consumers offers **three-part tariffs**: free allowance + steep overage charges.

Welfare

Overconfident consumers **lose** from such contracts, and the firm gains.

Take-away

Biases do not wash out in markets: firms can **design** contracts around them.

Exercises

1. (Easy)

A student gives 80% intervals for 10 questions and 5 contain the truth. What hit rate would a calibrated student have on average? How likely is 5 or fewer hits if she is calibrated?

2. (Intermediate)

In the Moore–Healy model take $m = 10$, $v_S = 6$, $v_L = 2$, $v_\varepsilon = 2$. Case (a): hard test $S = 5$, person with $L_i = 1$, $\varepsilon_i = 0$. Case (b): easy test $S = 15$, $L_i = -1$, $\varepsilon_i = 0$. Compute her beliefs about her own and others' scores. Does she over- or underestimate, over- or underplace?

3. (Intermediate)

In Möbius et al. (2014), start from $\mu = 0.5$ and receive signals "top", "top", "bottom". Find (a) the Bayesian posterior; (b) the posterior of someone who treats "bottom" signals as only 60% accurate; (c) both, after 2 "top" and 2 "bottom".

4. (Advanced)

In the phone example, compute the consumer's **actual** expected surplus under the two-part and under the three-part tariff, for the overconfident consumer. Who gains and who loses from overconfidence? Why does the firm not raise the price of minutes 0–100?

Final take-aways

1. Overconfidence comes in three forms: **overprecision**, **overplacement**, **overestimation**.
2. Much over- and underestimation and placement is what **Bayesian** agents with noisy signals would show (regression to the mean; hard vs. easy tasks).
3. A robust core is a genuine bias: **overprecision**, failure of the Benoît–Dubra test, **asymmetric** updating.
4. Consequences: **excess entry** (reference group neglect) and contracts that exploit consumers, e.g. **three-part tariffs**.

References I

Benoît, J.-P. and Dubra, J. (2011). Apparent overconfidence. *Econometrica*, 79(5):1591–1625.

Burks, S. V., Carpenter, J. P., Goette, L., and Rustichini, A. (2013). Overconfidence and social signalling. *Review of Economic Studies*, 80(3):949–983.

Camerer, C. and Lovallo, D. (1999). Overconfidence and excess entry: An experimental approach. *American Economic Review*, 89(1):306–318.

Dean, M. and Ortoleva, P. (2019). The empirical relationship between nonstandard economic behaviors. *Proceedings of the National Academy of Sciences*, 116(33):16262–16267.

Grubb, M. D. (2009). Selling to overconfident consumers. *American Economic Review*, 99(5):1770–1807.

Kennedy, E. J., Lawton, L., and Plumlee, E. L. (2002). Blissful ignorance: The problem of unrecognized incompetence and academic performance. *Journal of Marketing Education*, 24(3):243–252.

Kruger, J. and Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6):1121.

Möbius, M. M., Niederle, M., Niehaus, P., and Rosenblat, T. S. (2014). Managing self-confidence. *NBER Working paper*.

Moore, D. A. and Healy, P. J. (2008). The trouble with overconfidence. *Psychological Review*, 115(2):502.

Soll, J. B. and Klayman, J. (2004). Overconfidence in interval estimates. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(2):299.

Stankov, L. and Crawford, J. D. (1997). Self-confidence and performance on tests of cognitive abilities. *Intelligence*, 25(2):93–109.

Svenson, O. (1981). Are we all less risky and more skillful than our fellow drivers? *Acta Psychologica*, 47(2):143–148.

Zenger, T. R. (1992). Why do employers only reward extreme performance? Examining the relationships among performance, pay, and turnover. *Administrative Science Quarterly*, 37(2):198–219.