

# Deception: Theoretical Considerations

Joel Sobel

November 7, 2019

# Disclaimer

I'm not sure what I'm doing here

- I'm just a game theorist

# Disclaimer

I'm not sure what I'm doing here

- I'm just a game theorist
- so when you talk

# Disclaimer

I'm not sure what I'm doing here

- I'm just a game theorist
- so when you talk
- I'm Ginger



# Disclaimer

I'm not sure what I'm doing here

- I'm just a game theorist
- so when you talk
- I'm Ginger
- and when I talk



# Disclaimer

I'm not sure what I'm doing here

- I'm just a game theorist
- so when you talk
- I'm Ginger
- and when I talk
- you may be bored.



# Organization

## 1 Brief Motivation

# Organization

- 1 Brief Motivation
- 2 A Formal Model



# Organization

- 1 Brief Motivation
- 2 A Formal Model
- 3 Definition(s) of Deception

# Organization

- 1 Brief Motivation
- 2 A Formal Model
- 3 Definition(s) of Deception
- 4 Deception versus Lying

# Organization

- 1 Brief Motivation
- 2 A Formal Model
- 3 Definition(s) of Deception
- 4 Deception versus Lying
- 5 Two Properties

# Organization

- 1 Brief Motivation
- 2 A Formal Model
- 3 Definition(s) of Deception
- 4 Deception versus Lying
- 5 Two Properties
- 6 Omission and Deception

# Organization

- 1 Brief Motivation
- 2 A Formal Model
- 3 Definition(s) of Deception
- 4 Deception versus Lying
- 5 Two Properties
- 6 Omission and Deception
- 7 Deception and Damage

# Organization

- 1 Brief Motivation
- 2 A Formal Model
- 3 Definition(s) of Deception
- 4 Deception versus Lying
- 5 Two Properties
- 6 Omission and Deception
- 7 Deception and Damage
- 8 Theory versus Practice

# Motivation

- For me: Curiosity, Coherence, Cost of Lying.

# Motivation

- For me: Curiosity, Coherence, Cost of Lying.
- For today: Deception as public policy issue.



# Motivation

- For me: Curiosity, Coherence, Cost of Lying.
- For today: Deception as public policy issue.

# Motivation

- For me: Curiosity, Coherence, Cost of Lying.
- For today: Deception as public policy issue.

In fact:

Lip service to public policy. I hope my curiosity is enough to maintain your interest.

# Basic Model

- 1 Two players: Sender and Receiver.

# Basic Model

- 1 Two players: Sender and Receiver.
- 2 Sender observes  $\theta \in \Theta$ .

# Basic Model

- 1 Two players: Sender and Receiver.
- 2 Sender observes  $\theta \in \Theta$ .
- 3 Sender sends message  $m \in M$ .

# Basic Model

- 1 Two players: Sender and Receiver.
- 2 Sender observes  $\theta \in \Theta$ .
- 3 Sender sends message  $m \in M$ .
- 4 In some applications  $S$  also takes action  $x \in X$ .

# Basic Model

- 1 Two players: Sender and Receiver.
- 2 Sender observes  $\theta \in \Theta$ .
- 3 Sender sends message  $m \in M$ .
- 4 In some applications  $S$  also takes action  $x \in X$ .
- 5  $R$  observes  $m$ , not  $x$ ;  $x$  may influence  $R$ 's preferences directly, not  $m$ .

# Basic Model

- 1 Two players: Sender and Receiver.
- 2 Sender observes  $\theta \in \Theta$ .
- 3 Sender sends message  $m \in M$ .
- 4 In some applications  $S$  also takes action  $x \in X$ .
- 5  $R$  observes  $m$ , not  $x$ ;  $x$  may influence  $R$ 's preferences directly, not  $m$ .
- 6 Receiver takes action  $y \in Y$ .



# Basic Model

- 1 Two players: Sender and Receiver.
- 2 Sender observes  $\theta \in \Theta$ .
- 3 Sender sends message  $m \in M$ .
- 4 In some applications  $S$  also takes action  $x \in X$ .
- 5  $R$  observes  $m$ , not  $x$ ;  $x$  may influence  $R$ 's preferences directly, not  $m$ .
- 6 Receiver takes action  $y \in Y$ .
- 7 Preferences  $U^i(\theta, x, y, m)$ ;  $U^R(\cdot)$  independent of  $m$ .

# Basic Model

- 1 Two players: Sender and Receiver.
- 2 Sender observes  $\theta \in \Theta$ .
- 3 Sender sends message  $m \in M$ .
- 4 In some applications  $S$  also takes action  $x \in X$ .
- 5  $R$  observes  $m$ , not  $x$ ;  $x$  may influence  $R$ 's preferences directly, not  $m$ .
- 6 Receiver takes action  $y \in Y$ .
- 7 Preferences  $U^i(\theta, x, y, m)$ ;  $U^R(\cdot)$  independent of  $m$ .
- 8 Prior  $P(\theta)$ .

## Basic Model

- 1 Two players: Sender and Receiver.
- 2 Sender observes  $\theta \in \Theta$ .
- 3 Sender sends message  $m \in M$ .
- 4 In some applications  $S$  also takes action  $x \in X$ .
- 5  $R$  observes  $m$ , not  $x$ ;  $x$  may influence  $R$ 's preferences directly, not  $m$ .
- 6 Receiver takes action  $y \in Y$ .
- 7 Preferences  $U^i(\theta, x, y, m)$ ;  $U^R(\cdot)$  independent of  $m$ .
- 8 Prior  $P(\theta)$ .
- 9 All sets are finite (unless I want them to be infinite).

## Model Includes . . .

- 1 Standard (Spence) signaling (when education isn't productive).

## Model Includes . . .

- 1 Standard (Spence) signaling (when education isn't productive).
- 2 Cheap talk.

## Model Includes . . .

- 1 Standard (Spence) signaling (when education isn't productive).
- 2 Cheap talk.
- 3 Verifiable information (disclosure) games.

## Model Includes . . .

- 1 Standard (Spence) signaling (when education isn't productive).
- 2 Cheap talk.
- 3 Verifiable information (disclosure) games.
- 4 Some sequential games with incomplete information.

# Deception

## Definition (Deception)

Let  $\mu$  be a probability distribution on  $\Theta$ .



## Definition (Deception)

Let  $\mu$  be a probability distribution on  $\Theta$ .

- 1 The message  $m$  is **deceptive** given  $\theta$  and  $\mu$  if there exists a message  $n$  such that  $\mu(\theta | n) > 0$  and a number  $p \in (0, 1)$ , and a distribution  $\rho$  satisfying  $\rho(\theta) = 0$  such that

$$\mu(\cdot | m) = p\mu(\cdot | n) + (1 - p)\rho \quad (1)$$

## Definition (Deception)

Let  $\mu$  be a probability distribution on  $\Theta$ .

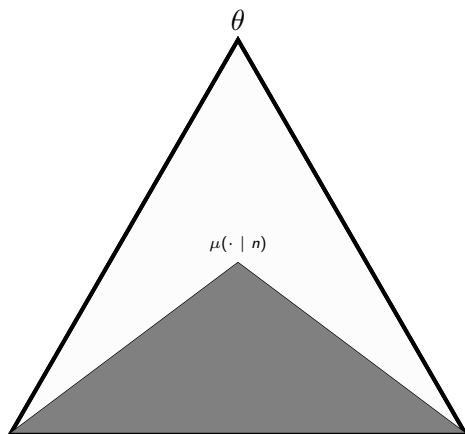
- 1 The message  $m$  is **deceptive** given  $\theta$  and  $\mu$  if there exists a message  $n$  such that  $\mu(\theta | n) > 0$  and a number  $p \in (0, 1)$ , and a distribution  $\rho$  satisfying  $\rho(\theta) = 0$  such that

$$\mu(\cdot | m) = p\mu(\cdot | n) + (1 - p)\rho \quad (1)$$

- 2 The message  $m$  is **strongly deceptive** given  $\theta$  and  $\mu$  if there exists  $n$  such that  $\mu(\cdot | m) \neq \mu(\cdot | n)$  and  $p \in [0, 1)$  such that

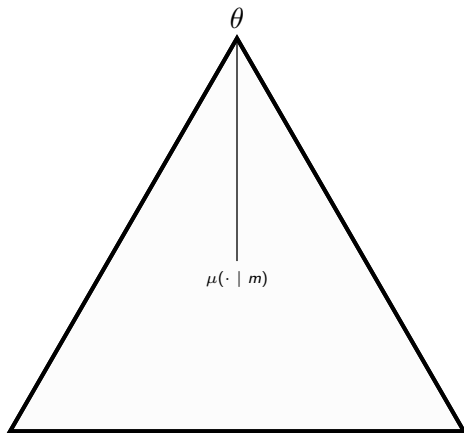
$$\mu(\cdot | n) = p\mu(\cdot | m) + (1 - p)I(\cdot | \theta). \quad (2)$$

## Deception Illustrated



Any beliefs in shaded region are deceptive.

## Strong Deception Illustrated



$\mu(\cdot | m)$  strongly deceptive only if it is possible to induce beliefs on line segment.

# Lying is not Deception

*Red Bull gives you wings.*

# Lying is not Deception

*Red Bull gives you wings.*

# Lying is not Deception

*Red Bull gives you wings.*

If no one believes your lies, then they are not deceptive.

## Example

- 1  $S$  learns  $\theta$  (an element of  $\{1, \dots, 10\}$ ) or nothing.



## Example

- 1  $S$  learns  $\theta$  (an element of  $\{1, \dots, 10\}$ ) or nothing.
- 2  $S$  can state type or “I don’t know.”

## Example

- 1  $S$  learns  $\theta$  (an element of  $\{1, \dots, 10\}$ ) or nothing.
- 2  $S$  can state type or “I don’t know.”
- 3  $R$  pays expected value of type;  $S$  prefers higher payment.

## Example

- 1  $S$  learns  $\theta$  (an element of  $\{1, \dots, 10\}$ ) or nothing.
- 2  $S$  can state type or “I don’t know.”
- 3  $R$  pays expected value of type;  $S$  prefers higher payment.
- 4 In equilibrium:  $S$  says “I don’t know” if she doesn’t know or if type is small.

## Example

- 1  $S$  learns  $\theta$  (an element of  $\{1, \dots, 10\}$ ) or nothing.
- 2  $S$  can state type or “I don’t know.”
- 3  $R$  pays expected value of type;  $S$  prefers higher payment.
- 4 In equilibrium:  $S$  says “I don’t know” if she doesn’t know or if type is small.
- 5 Saying “I don’t know” when  $S$  knows is deceptive.

# Omission/Silence and Deception

- Formally, omission is not defined in abstract model.

## Omission/Silence and Deception

- Formally, omission is not defined in abstract model.
- It is not hard to add structure so that formulation includes message labeled “silence” or incomplete disclosure.

## Omission/Silence and Deception

- Formally, omission is not defined in abstract model.
- It is not hard to add structure so that formulation includes message labeled “silence” or incomplete disclosure.
- We did that in the previous example.

## Omission/Silence and Deception

- Formally, omission is not defined in abstract model.
- It is not hard to add structure so that formulation includes message labeled “silence” or incomplete disclosure.
- We did that in the previous example.
- Receiver is free to interpret the distinguished message arbitrarily.



## Omission/Silence and Deception

- Formally, omission is not defined in abstract model.
- It is not hard to add structure so that formulation includes message labeled “silence” or incomplete disclosure.
- We did that in the previous example.
- Receiver is free to interpret the distinguished message arbitrarily.
- These beliefs determine whether distinguished message is deceptive.

## Example: Zyllyz

Fictional drug marketed as a cognitive aid.

Possible Advertisements:

Interpreted using my formulation:

## Example: Zyllyz

Fictional drug marketed as a cognitive aid.

Possible Advertisements:

- 1 *Buy Zyllyz.*

Interpreted using my formulation:

## Example: Zyllyz

Fictional drug marketed as a cognitive aid.

Possible Advertisements:

- 1 *Buy Zyllyz.*

Interpreted using my formulation:

## Example: Zyllyz

Fictional drug marketed as a cognitive aid.

Possible Advertisements:

- 1 *Buy Zyllyz.*
- 2 *Zyllyz helps you think forward and backward.*

Interpreted using my formulation:

## Example: Zyllyz

Fictional drug marketed as a cognitive aid.

Possible Advertisements:

- 1 *Buy Zyllyz.*
- 2 *Zyllyz helps you think forward and backward.*

Interpreted using my formulation:

## Example: Zyllyz

Fictional drug marketed as a cognitive aid.

Possible Advertisements:

- 1 *Buy Zyllyz.*
- 2 *Zyllyz helps you think forward and backward.*
- 3 *Eight out of ten economic theorists can become  
10 economists with regular treatments of Zyllyz.*

Interpreted using my formulation:

## Example: Zyllyz

Fictional drug marketed as a cognitive aid.

Possible Advertisements:

- 1 *Buy Zyllyz.*
- 2 *Zyllyz helps you think forward and backward.*
- 3 *Eight out of ten economic theorists can become IO economists with regular treatments of Zyllyz.*

Interpreted using my formulation:

- 1 Basically silence. Not dishonest. Deceptive if (for example): Zyllyz is worthless;  $R$  interprets statement as uninformative relative to prior;  $S$  could credibly state "Zyllyz is worthless."



## Example: Zyllyz

Fictional drug marketed as a cognitive aid.

Possible Advertisements:

- 1 *Buy Zyllyz.*
- 2 *Zyllyz helps you think forward and backward.*
- 3 *Eight out of ten economic theorists can become IO economists with regular treatments of Zyllyz.*

Interpreted using my formulation:

- 1 Basically silence. Not dishonest. Deceptive if (for example): Zyllyz is worthless; *R* interprets statement as uninformative relative to prior; *S* could credibly state “Zyllyz is worthless.”
- 2 No operational meaning. Not dishonest. Deception as above.

## Example: Zyllyz

Fictional drug marketed as a cognitive aid.

Possible Advertisements:

- 1 *Buy Zyllyz.*
- 2 *Zyllyz helps you think forward and backward.*
- 3 *Eight out of ten economic theorists can become IO economists with regular treatments of Zyllyz.*

Interpreted using my formulation:

- 1 Basically silence. Not dishonest. Deceptive if (for example): Zyllyz is worthless;  $R$  interprets statement as uninformative relative to prior;  $S$  could credibly state "Zyllyz is worthless."
- 2 No operational meaning. Not dishonest. Deception as above.
- 3 Could be false. May or may not be deceptive (independent of the truth of the statement) depending on the interpretation.

# Damage

Let  $\bar{u}^R(\theta, x, m)$  be the Receiver's expected utility when  $S$  takes action  $x$ , sends the message  $m$ , and the true state is  $\theta$ .

Definition (Damaging Behavior)

# Damage

Let  $\bar{u}^R(\theta, x, m)$  be the Receiver's expected utility when  $S$  takes action  $x$ , sends the message  $m$ , and the true state is  $\theta$ .

## Definition (Damaging Behavior)

- 1 The pair  $(m, x)$  is a **damaging action** given  $\theta$  and  $y(\cdot)$  if there exists a message  $n$  such that  $\bar{u}^R(\theta, x, y(m)) < \bar{u}^R(\theta, x, y(n))$ .

# Damage

Let  $\bar{u}^R(\theta, x, m)$  be the Receiver's expected utility when  $S$  takes action  $x$ , sends the message  $m$ , and the true state is  $\theta$ .

## Definition (Damaging Behavior)

- 1 The pair  $(m, x)$  is a **damaging action** given  $\theta$  and  $y(\cdot)$  if there exists a message  $n$  such that  $\bar{u}^R(\theta, x, y(m)) < \bar{u}^R(\theta, x, y(n))$ .
- 2 The strategy  $(m^*, x^*)$  is a **damaging strategy** given  $y(\cdot)$  if there exists a  $\theta$  such that  $(m^*(\theta), x^*(\theta))$  is a damaging action given  $\theta$  and  $y(\cdot)$ .

# Properties of Damaging Messages

## Proposition

*If the Sender's strategy set is finite, then for each  $\theta$  and  $y(\cdot)$ , the Sender will have a strategy that is not damaging.*

# Properties of Damaging Messages

## Proposition

*If the Sender's strategy set is finite, then for each  $\theta$  and  $y(\cdot)$ , the Sender will have a strategy that is not damaging.*

## Proposition

*If  $y(\cdot)$  is constant, then no Sender strategy is damaging given  $y(\cdot)$ .*

## Result: Informal

“Deception causes damage. Damage requires deception.”

if  $m$  is deceptive given  $\theta$ , then any  $R$  does better after  $n$ .

If  $m$  is not deceptive given  $\theta$ , then there exists a specification preferences for which the message  $m$  is not damaging given  $\theta$  (assuming  $R$  responds optimally).



# Difference Between Deception and Strong Deception

- 1 Deception is easier.

# Difference Between Deception and Strong Deception

- 1 Deception is easier.
- 2 The “specification of preferences” can come from a smaller class.

## Deception: Other Definitions

- 1 Different notions of “further from the truth.” Research question that pairs definition of deception with the set of preferences for which deception is damaging.

## Deception: Other Definitions

- 1 Different notions of “further from the truth.” Research question that pairs definition of deception with the set of preferences for which deception is damaging.
- 2 Prior beliefs and prior optimal actions as a benchmark: A message would be deceptive if it induces beliefs that are less accurate (in some sense) than the prior; damaging if the message lowers the Receiver’s payoff relative to the prior optimal action.  
Problems: sometimes message will be deceptive even if it does not influence beliefs; deception can’t be avoided.

## Deception: Other Definitions

- 1 Different notions of “further from the truth.” Research question that pairs definition of deception with the set of preferences for which deception is damaging.
- 2 Prior beliefs and prior optimal actions as a benchmark: A message would be deceptive if it induces beliefs that are less accurate (in some sense) than the prior; damaging if the message lowers the Receiver’s payoff relative to the prior optimal action.  
Problems: sometimes message will be deceptive even if it does not influence beliefs; deception can’t be avoided.
- 3 FTC: deception is misleading, benchmark is “reasonable consumer;” deception must be damaging. Reasonable: credulous or equilibrium.

## Deception: Other Definitions

- 1 Different notions of “further from the truth.” Research question that pairs definition of deception with the set of preferences for which deception is damaging.
- 2 Prior beliefs and prior optimal actions as a benchmark: A message would be deceptive if it induces beliefs that are less accurate (in some sense) than the prior; damaging if the message lowers the Receiver’s payoff relative to the prior optimal action.  
Problems: sometimes message will be deceptive even if it does not influence beliefs; deception can’t be avoided.
- 3 FTC: deception is misleading, benchmark is “reasonable consumer;” deception must be damaging. Reasonable: credulous or equilibrium.
- 4 Philosophers: manipulation of beliefs, intentional, and not about consequences.

# Theory Versus Practice: Beliefs

- Where do beliefs come from?

# Theory Versus Practice: Beliefs

- Where do beliefs come from?
- Theory assumes that beliefs are “in the head of the Sender.”



# Theory Versus Practice: Beliefs

- Where do beliefs come from?
- Theory assumes that beliefs are “in the head of the Sender.”
  - 1 Right choice for theory (because  $S$ 's beliefs determine  $S$ 's behavior).

# Theory Versus Practice: Beliefs

- Where do beliefs come from?
- Theory assumes that beliefs are “in the head of the Sender.”
  - 1 Right choice for theory (because  $S$ 's beliefs determine  $S$ 's behavior).
  - 2 Not useful for practice (cannot reliably identify  $S$ 's beliefs).

# Theory Versus Practice: Beliefs

- Where do beliefs come from?
- Theory assumes that beliefs are “in the head of the Sender.”
  - 1 Right choice for theory (because  $S$ 's beliefs determine  $S$ 's behavior).
  - 2 Not useful for practice (cannot reliably identify  $S$ 's beliefs).
- Practice: Beliefs come from “representative consumer.” If so, lies become deceptive.

For example: All claims – but especially numerical ones – are interpreted literally. (Miniwheat.)

# Theory Versus Practice: Silence

- Theory does treat silence differently.

# Theory Versus Practice: Silence

- Theory does treat silence differently.
- Practice probably does, but omission may be viewed as deceptive. Examples:

# Theory Versus Practice: Silence

- Theory does treat silence differently.
- Practice probably does, but omission may be viewed as deceptive. Examples:
  - 1 Pt Fusion: incomplete description of confidentiality policy.

# Theory Versus Practice: Silence

- Theory does treat silence differently.
- Practice probably does, but omission may be viewed as deceptive. Examples:
  - 1 Pt Fusion: incomplete description of confidentiality policy.
  - 2 Machinima: incomplete information about the source of reviews.

# Theory Versus Practice: Silence

- Theory does treat silence differently.
- Practice probably does, but omission may be viewed as deceptive. Examples:
  - 1 Pt Fusion: incomplete description of confidentiality policy.
  - 2 Machinima: incomplete information about the source of reviews.
  - 3 POW Wonderful: failure to provide most recent evidence.



# Theory Versus Practice: Non Rational Consumers

- Theory permits behavioral agents, so there is no logical conflict.

# Theory Versus Practice: Non Rational Consumers

- Theory permits behavioral agents, so there is no logical conflict.
- Practical problems about the extent to which public policy should identify and correct lack of sophistication. (For example, failure to be skeptical in disclosure games.)

## Theory Versus Practice: Glib Summary

In practice, easy cases involve factually incorrect statements.

Hard cases involve “loose” statements that lead reasonable people to draw incorrect inferences.