ESTIMATING THE PREVALENCE OF DECEPTION IN ONLINE REVIEW COMMUNITIES

Myle Ott,[1] Claire Cardie,[1,2] Jeff Hancock[2,3]
Depts of Computer Science,[1] Information Science,[2] Communication[3]
Cornell University, Ithaca, New York

Consumers increasingly rate, review and research products and services online.

87%

"Positive information I've read online has **reinforced my decision** to purchase a product or service recommended to me."

Source: http://www.coneinc.com/negative-reviews-online-reverse-purchase-decisions

80%

"Negative information I've read online has made me **change my mind** about purchasing a product or service recommended to me."

Source: http://www.coneinc.com/negative-reviews-online-reverse-purchase-decisions

Perhaps unsurprisingly…

Is this an epidemic?

Fake reviews prompt Belkin apology

Amazon withdraws ebook explaining how to manipulate its sales rankings

Ebook claiming one can become a Kindle 'bestseller' simply by posting fake reviews temporarily removed from bookseller's listings

In a Race to Out-R

Historian Orlando F
damages for fake re

Orlando Figes posted reviews on Amazon praising his own work and rubbishing that of his rivals

Tripadvisor bribes: Hotel owners offer free rooms in return for glowing reviews

For $2 a Star, an On
Reviews

Author Claims To Manipulate Amazon Rankings By Buying Own Book Every Day

Company Settles Case of Reviews It Faked

Works Just as expected, May 14, 2007

By Laurie B. Cook - See all my reviews

REAL NAME

This review is from: **Belkin F5U301 CableFree 4-Port USB 2.0 Hub with Dongle (Electronics)**

Supplies good range and does provide true wireless USB. Software worked right out of the box. I have been recommending this nifty little device to all my friends. Very useful device.

How many
fake reviews
are out there?

How to tell if a review is fake

Which of these reviews is fake?

"I have stayed at many hotels traveling for both business and pleasure and I can honestly stay that The James is tops. The service at the hotel is first class. The rooms are modern and very comfortable. The location is perfect within walking distance to all of the great sights and restaurants. Highly recommend to both business travellers and couples."

"My husband and I stayed at the James Chicago Hotel for our anniversary. This place is fantastic! We knew as soon as we arrived we made the right choice! The rooms are BEAUTIFUL and the staff very attentive and wonderful!! The area of the hotel is great, since I love to shop I couldn't ask for more!! We will definatly be back to Chicago and we will for sure be back to the James Chicago."

Which of these reviews is fake?

Answer:

"My husband and I stayed at the James Chicago Hotel for our anniversary. This place is fantastic! We knew as soon as we arrived we made the right choice! The rooms are BEAUTIFUL and the staff very attentive and wonderful!! The area of the hotel is great, since I love to shop I couldn't ask for more!! We will definatly be back to Chicago and we will for sure be back to the James Chicago."

Ott et al. (2011)

Dataset

- Solicited 400 fake positive reviews of Chicago hotels
- Gathered 400 truthful positive reviews from TripAdvisor

Source: http://www.cs.cornell.edu/~myleott

Ott et al. (2011)

Identifying Deception

- 2 out of 3 undergraduates performed at-chance
- n-gram text categorization (SVM) is ≈ 90% accurate

Source: http://www.cs.cornell.edu/~myleott

Ott et al. (2011)

Features

Truthful reviews

- Tempered opinions
- More spatial details
- More nouns and adjectives
- More numbers and punctuation

Fake reviews

- Exaggerated opinions
- Greater focus on aspects external to the hotel
- More pronouns, verbs and adverbs
- More filler (blah, like)

Source: http://www.cs.cornell.edu/~myleott

How many
fake reviews
are out there?

Expectations

- Given that users increasingly rely on online reviews (Cone, 2011), rates of deception must be low
- On the other hand, rates of deception may vary across review communities and user groups

Expectations

Less deception

- Verified (high cost) review communities
- Low traffic (low benefit) review communities

More deception

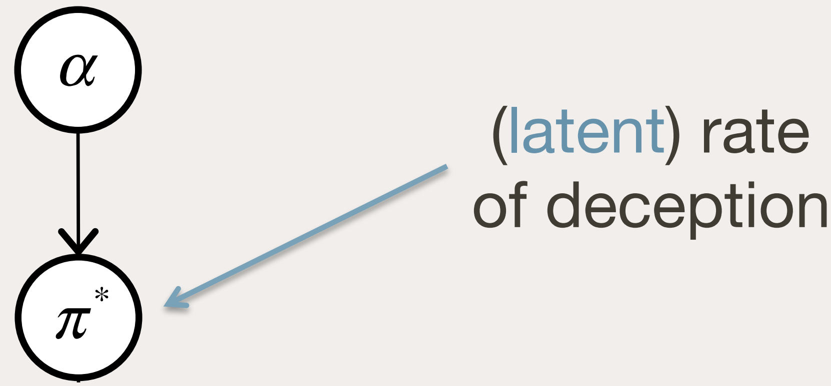
- Unverified (low cost) review communities
- High traffic (high benefit) review communities

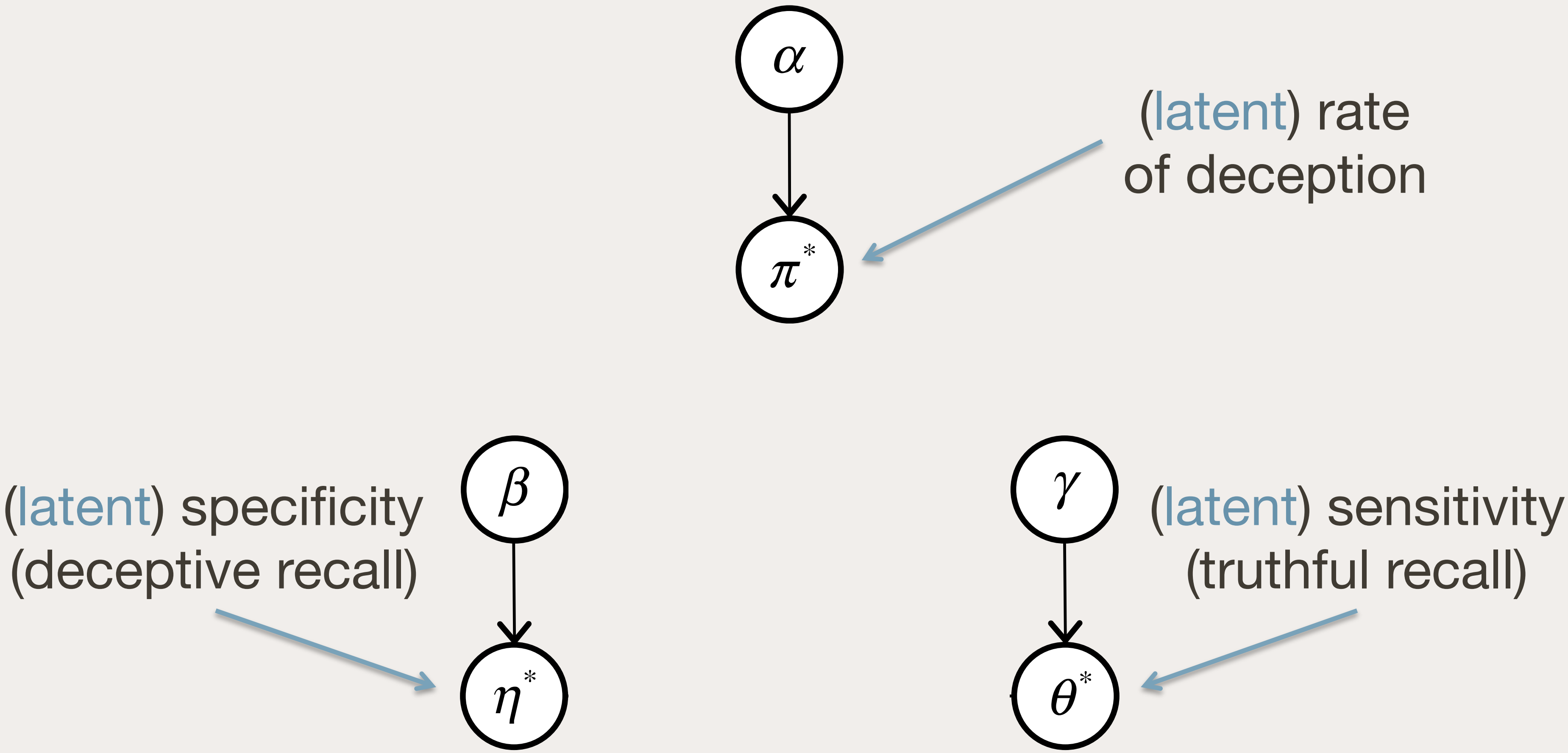
Approach

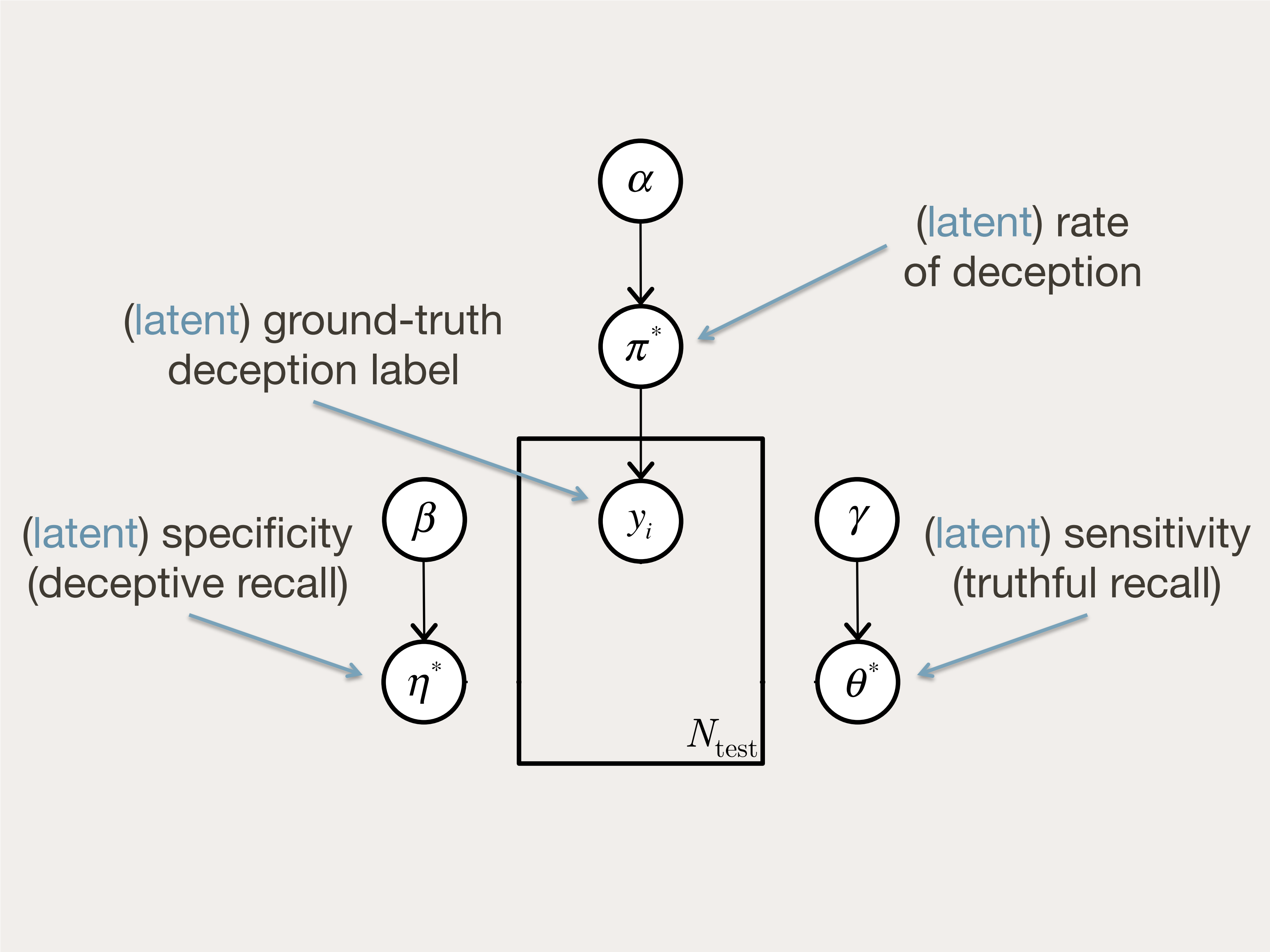
- Assume given a deception classifier
- Apply the classifier to some reviews
- Estimate the classifier's sensitivity and specificity, i.e., recall rates
- Estimate the rate of deception with a generative model

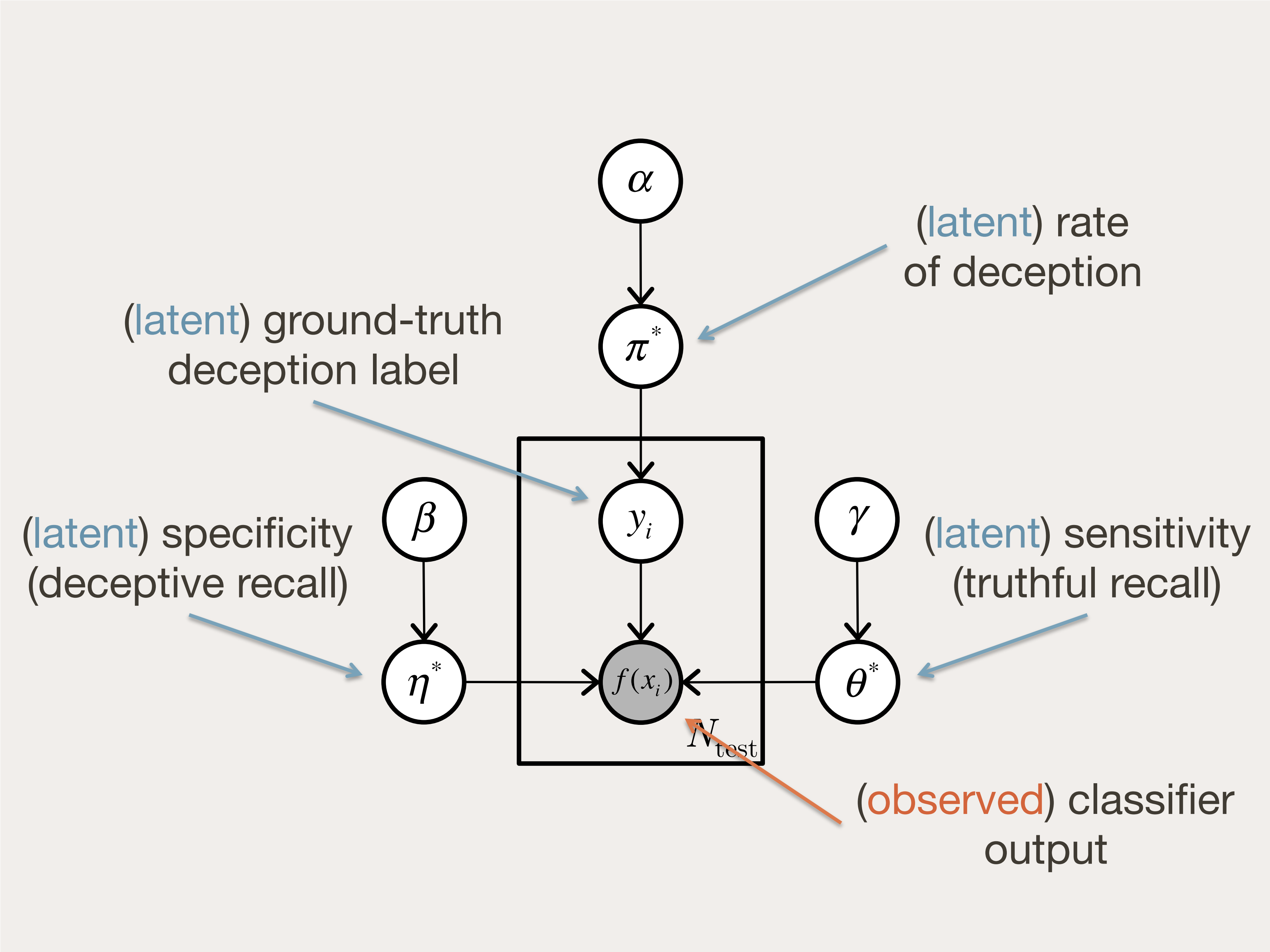
Generative Storyline

- Sample (latent) rate of deception
- Sample (latent) sensitivity
- Sample (latent) specificity
- For each review:
  - Sample (latent) ground-truth deception label
  - Sample (observed) classifier output


α
π^*
(latent) rate
of deception


α
π^*
(latent) rate of deception
β
η^*
(latent) specificity (deceptive recall)
γ
θ^*
(latent) sensitivity (truthful recall)

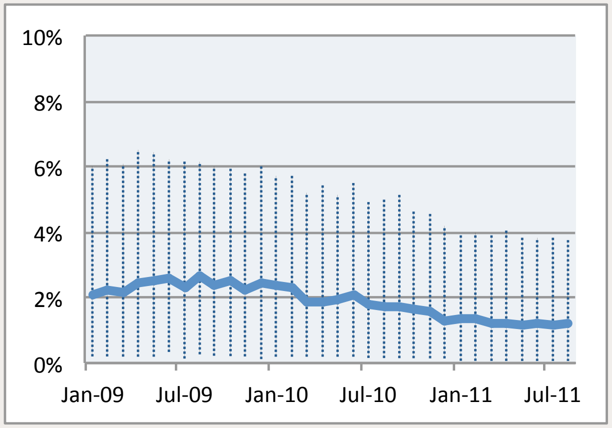

(latent) rate of deception
(latent) ground-truth deception label
(latent) specificity (deceptive recall)
(latent) sensitivity (truthful recall)
α
π^*
y_i
β
η^*
γ
θ^*
N_{test}

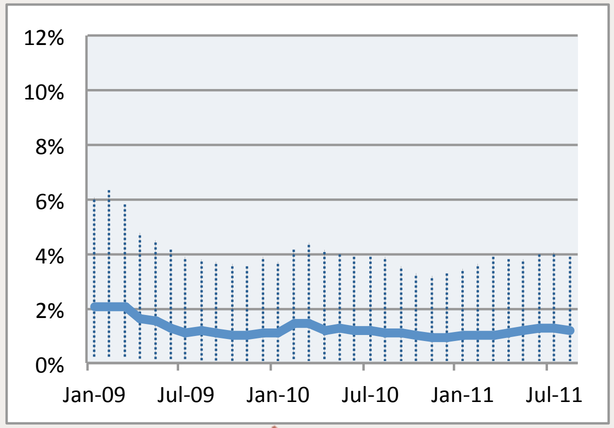


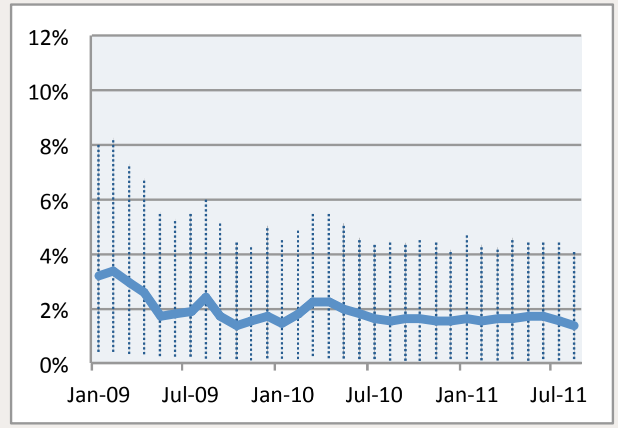
α
(latent) rate
of deception
π^*
(latent) ground-truth
deception label
y_i
β
γ
(latent) specificity
(deceptive recall)
(latent) sensitivity
(truthful recall)
η^*
$f(x_i)$
θ^*
N_{test}
(observed) classifier
output

Approach

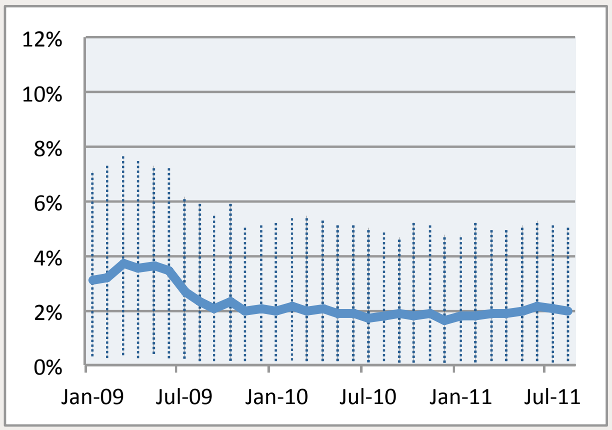
- Gibbs sampling
- Apply model to reviews from six hotel review communities:
  - High cost: Expedia, Hotels.com, Orbitz, Priceline
  - Low cost: TripAdvisor and Yelp

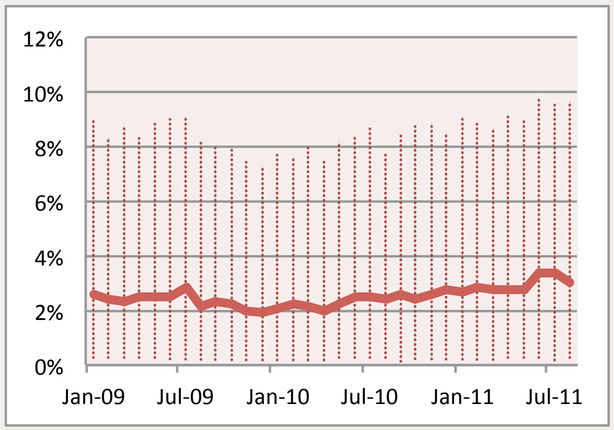


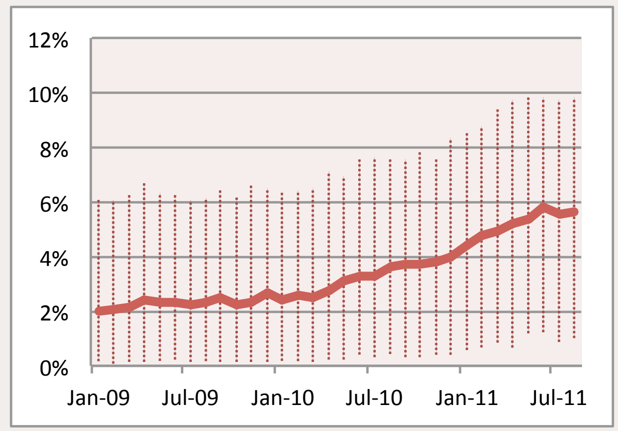




ORBITZ



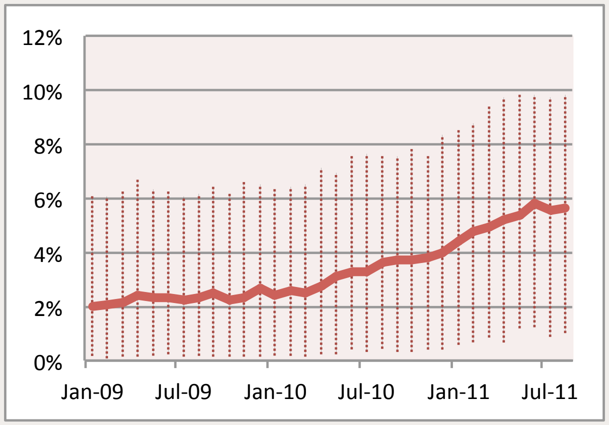




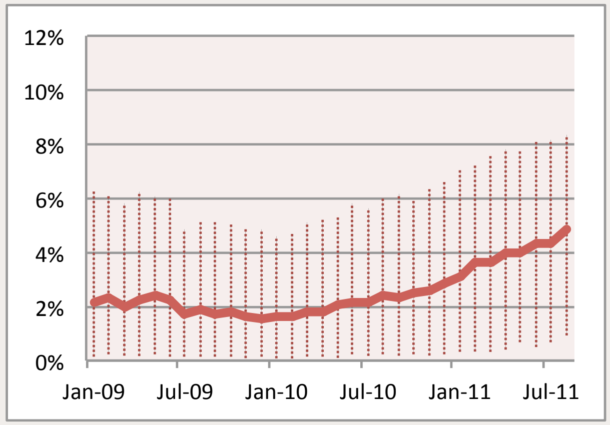
hotels.com

The rate of deception varies according to the **costs and benefits** of posting fake reviews.

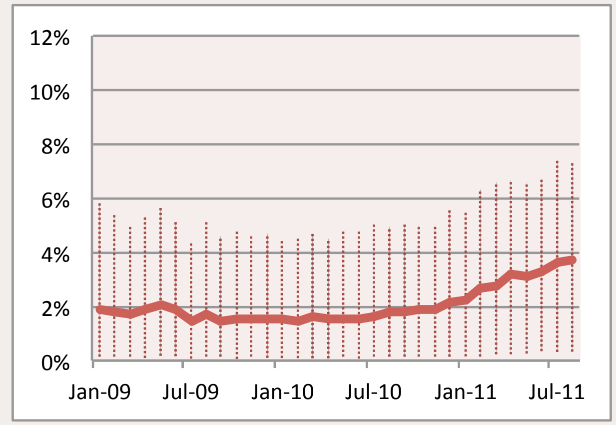
Increasing the cost should reduce deception.



Users with
≥ 1 review



Users with
≥ 2 reviews



Users with
≥ 3 reviews

Conclusion

- Presented a framework for estimating the rate of deception in online review communities using a noisy classifier.
- Explored the rates of deception in six popular review communities.
- Showed how review posting costs can be manipulated to reduce deception.
- Demo at ReviewSkeptic.com.

Thank you. Questions?

- Presented a framework for estimating the rate of deception in online review communities using a noisy classifier.
- Explored the rates of deception in six popular review communities.
- Showed how review posting costs can be manipulated to reduce deception.
- Demo at ReviewSkeptic.com.

ReviewSkeptic.com