

Your Browsing History May Cost You

A Framework for Discovering Differential Pricing in Non-Transparent Markets

Aditya Karan (karan2@illinois.edu)

Naina Balepur (nainab2@illinos.edu)

Hari Sundaram (hs1@illinois.edu)

University of Illinois Urbana-Champaign



We collect **1.1M** flight records
and **4.2M** hotel records.

Some customers lose up to **\$6**
on average – **33x** what you
might lose by chance.



Shopping Alone vs. Together In Large Markets



In-person audits allow for easy consensus; online audits are difficult.



Prior Work Makes Assumptions about Identity

Datta (2014) – audit with sock puppets; use google settings to validate identity

Asplund (2020) – also audit with sock puppets; use correlation to validate identity

Mikians (2012) – train sock puppets with focus on personal and browser attributes



"On the Internet, nobody knows you're a dog."

Our Work is Grounded in Consensus on Behavior



Our Work is Grounded in Consensus on Behavior



We Collect Data from Kayak with Sock Puppets

- Use user profiles from existing literature

We Collect Data from Kayak with Sock Puppets

- Use user profiles from existing literature
- For every run, have agents visit websites based on profile and then simultaneously execute query

We Collect Data from Kayak with Sock Puppets

- Use user profiles from existing literature
- For every run, have agents visit websites based on profile and then simultaneously execute query
- Record good, seller, and pricing information

We Collect Data from Kayak with Sock Puppets

- Use user profiles from existing literature
- For every run, have agents visit websites based on profile and then simultaneously execute query
- Record good, seller, and pricing information

Seller ID	Profile ID	Flight #	Route	Good ID	Date	Price
S1	P1	AA2058	LGA-ORD	101	11Jun23	\$100
S1	P2	AA2058	LGA-ORD	101	11Jun23	\$105

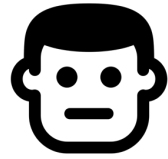
Average flight ticket price per profile is uninformative.



We Model Price Differences

Price Difference?

We Model Price Differences



Profile X

Query



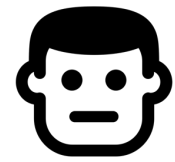
Profile Y



Seller

Price Difference?

We Model Price Differences



Profile X

Query



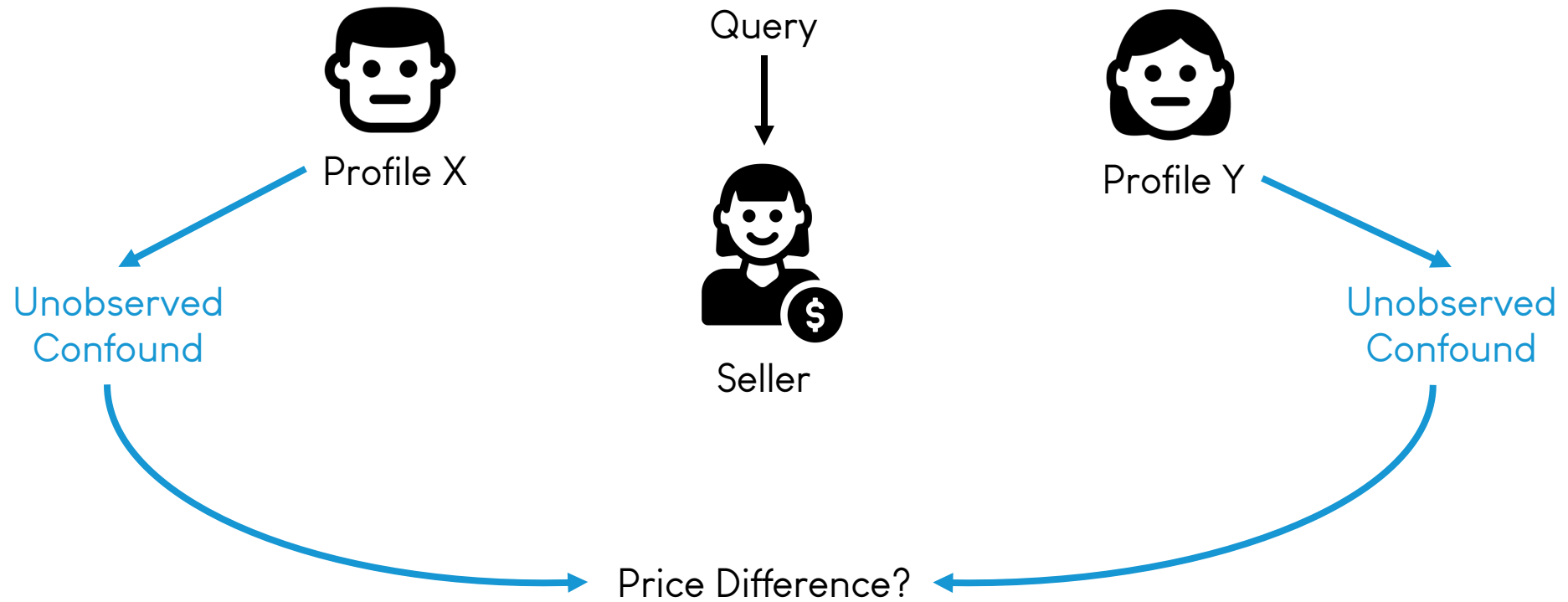
Seller



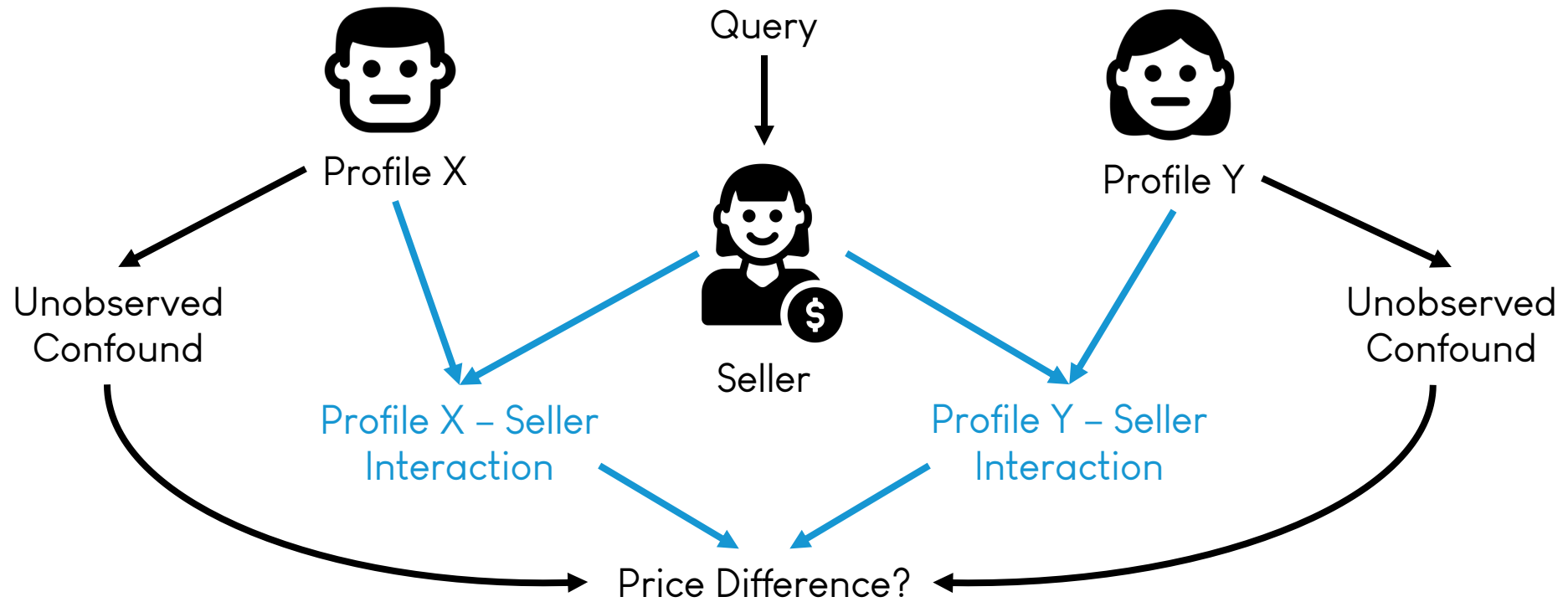
Profile Y

Price Difference?

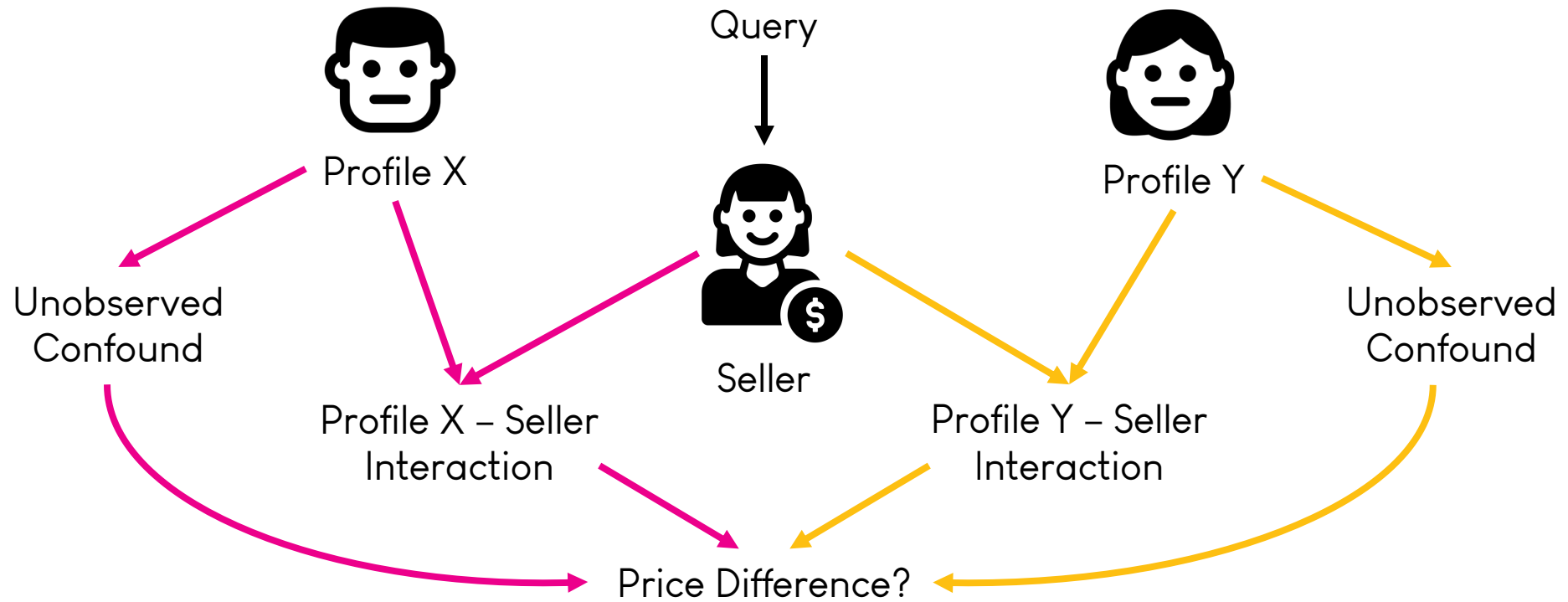
We Model Price Differences



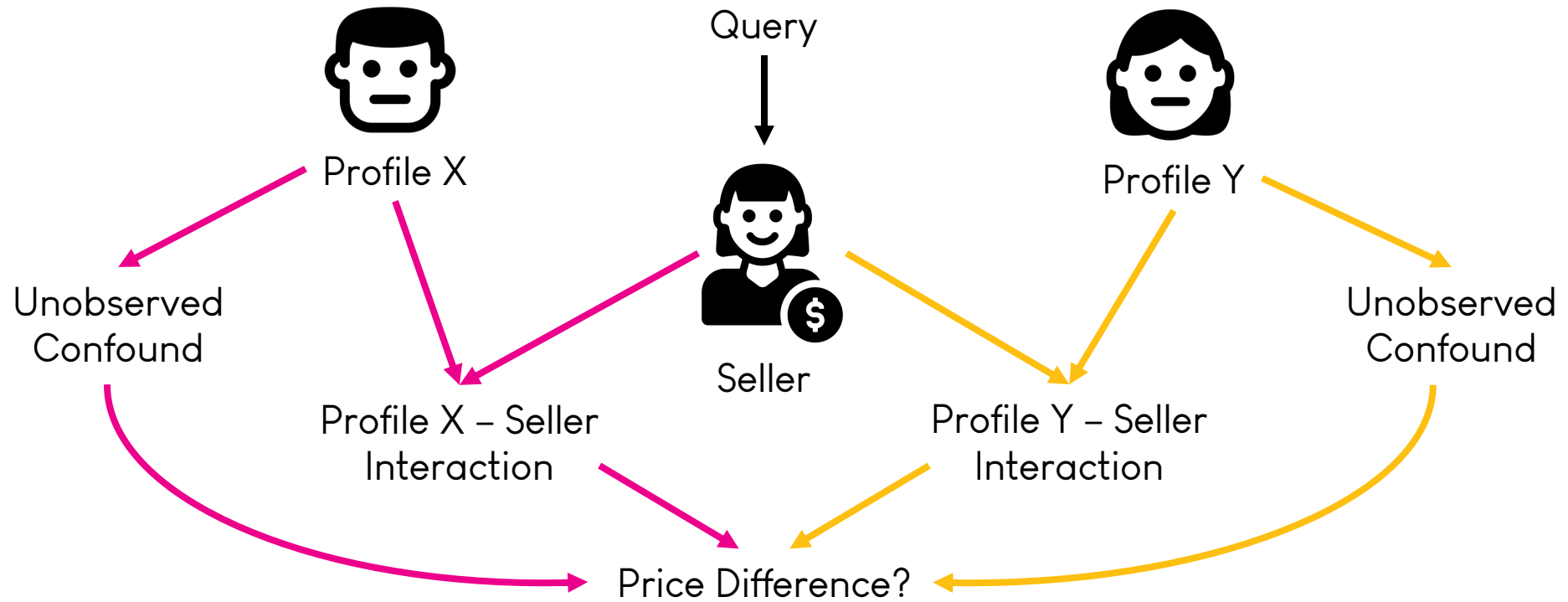
We Model Price Differences



We Model Price Differences



We Model Price Differences

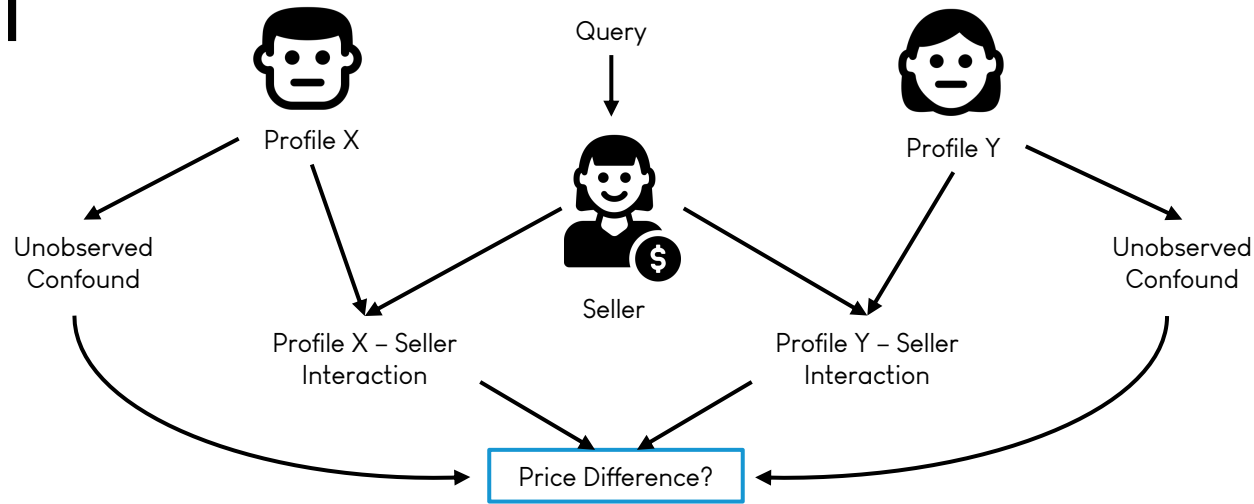


Good ID	Seller ID	Profile X ID	Price X	Profile Y ID	Price Y	Price Difference?
101	S1	P1	\$105	P2	\$100	True

Paired Difference Model

$$y_{i,j|k} \sim \text{Ber}(p_{i,j|k})$$

$y_{i,j|k}$ = indicator outcome variable;
buyers i and j see different prices from seller k

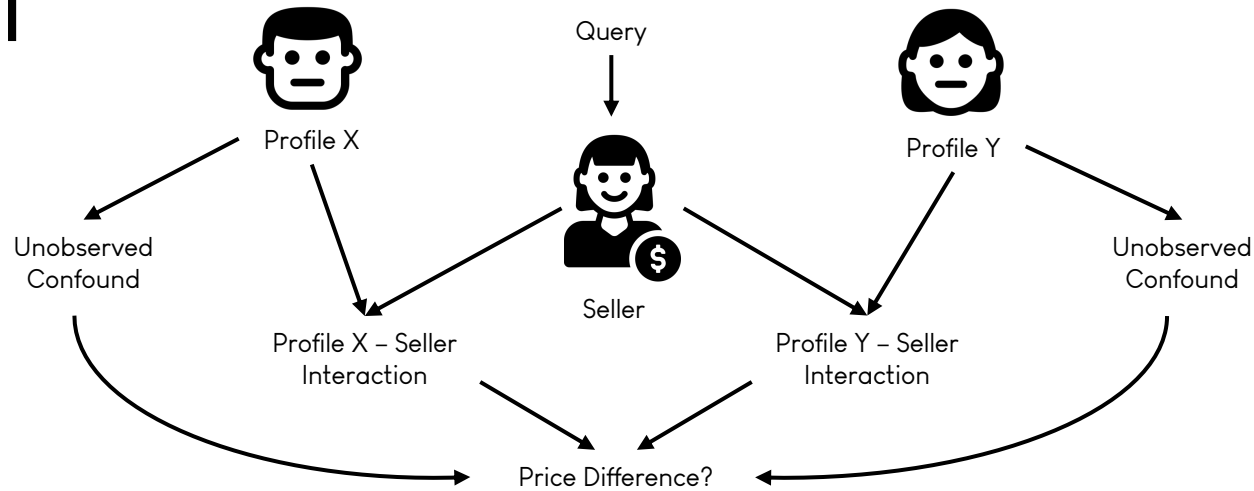


Paired Difference Model

$$y_{i,j|k} \sim \text{Ber}(p_{i,j|k})$$

$y_{i,j|k}$ = indicator outcome variable;
buyers i and j see different prices from seller k

$p_{i,j|k}$ = probability that buyers i and j
see different prices from seller k



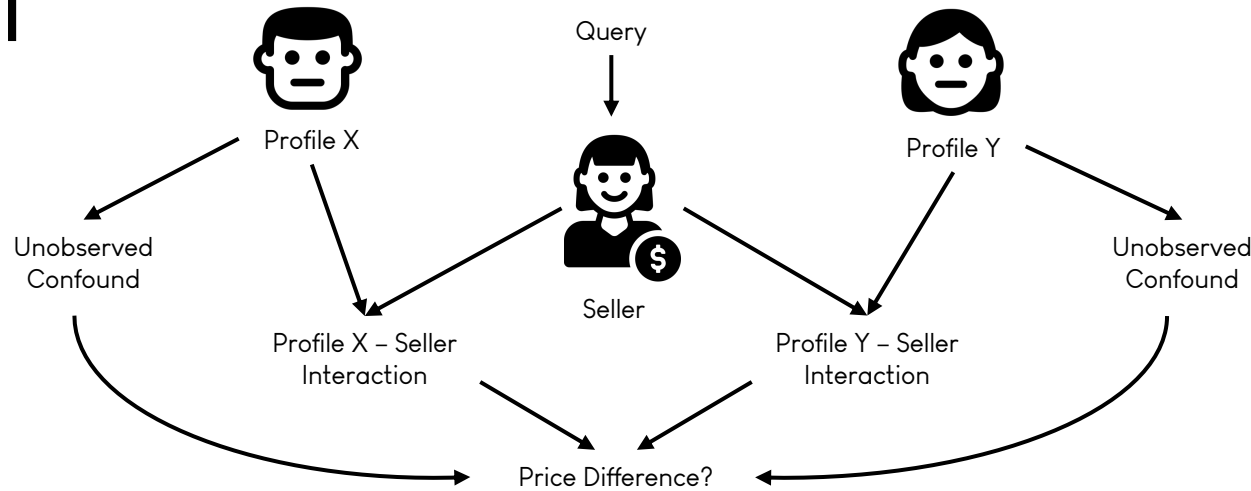
Paired Difference Model

$$y_{i,j|k} \sim \text{Ber}(p_{i,j|k})$$

$y_{i,j|k}$ = indicator outcome variable;
buyers i and j see different prices from seller k
 $p_{i,j|k}$ = probability that buyers i and j
see different prices from seller k

$$\text{Logit}(p_{i,j|k}) = \bar{p} + (\beta_i - \beta_j) + (\delta_{k,i} - \delta_{k,j})$$

\bar{p} = base probability that any buyers see different prices



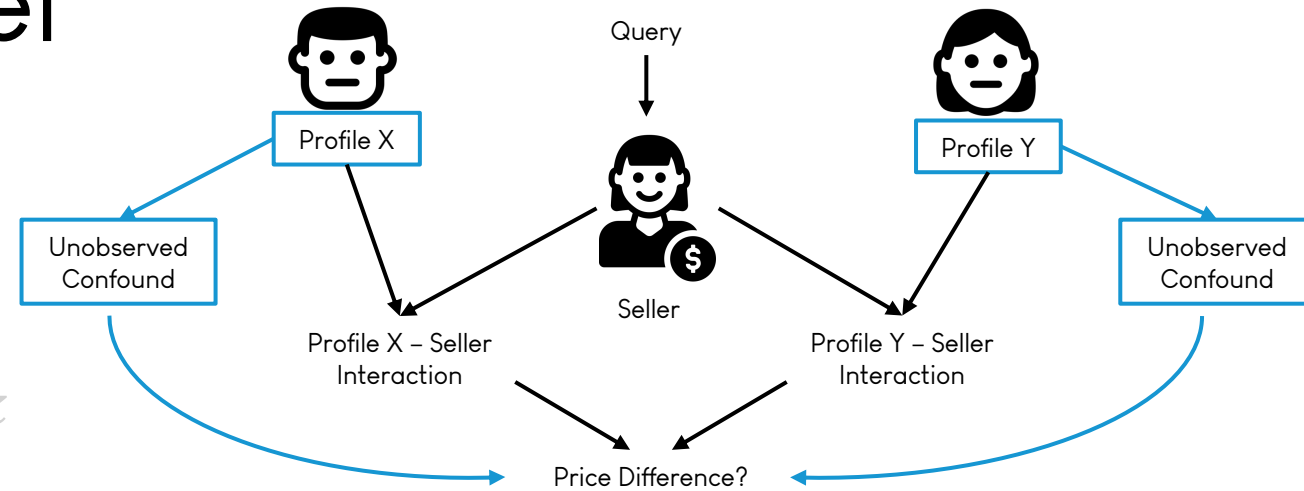
Paired Difference Model

$$y_{i,j|k} \sim \text{Ber}(p_{i,j|k})$$

$y_{i,j|k}$ = indicator outcome variable;
buyers i and j see different prices from seller k
 $p_{i,j|k}$ = probability that buyers i and j
see different prices from seller k

$$\text{Logit}(p_{i,j|k}) = \bar{p} + (\beta_i - \beta_j) + (\delta_{k,i} - \delta_{k,j})$$

\bar{p} = base probability that any buyers see different prices
 $\beta_i - \beta_j$ = difference in direct effect of profiles i and j on
probability of seeing different prices



Paired Difference Model

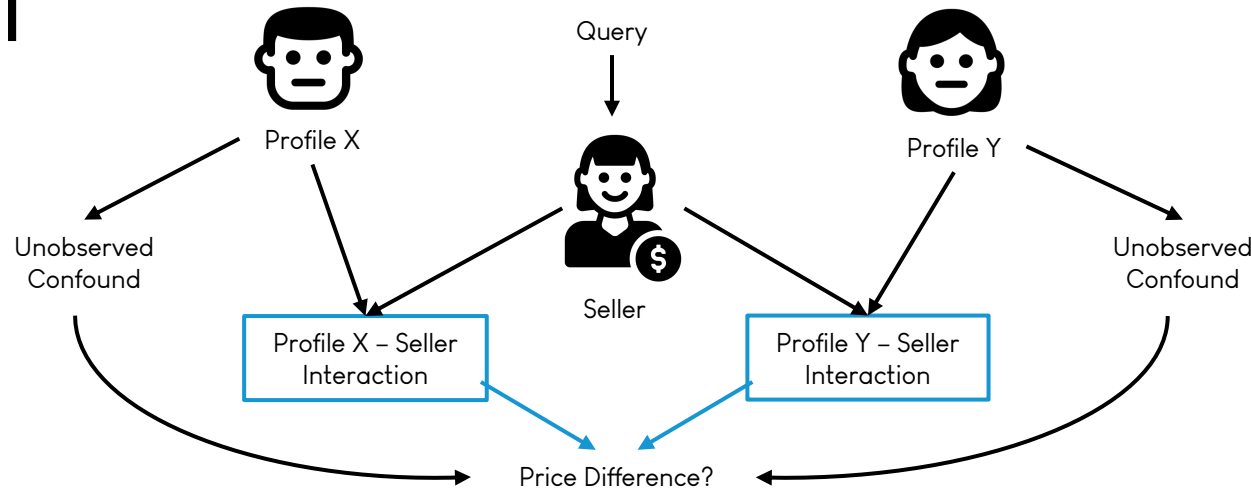
$$y_{i,j|k} \sim \text{Ber}(p_{i,j|k})$$

$y_{i,j|k}$ = indicator outcome variable;
buyers i and j see different prices from seller k
 $p_{i,j|k}$ = probability that buyers i and j
see different prices from seller k

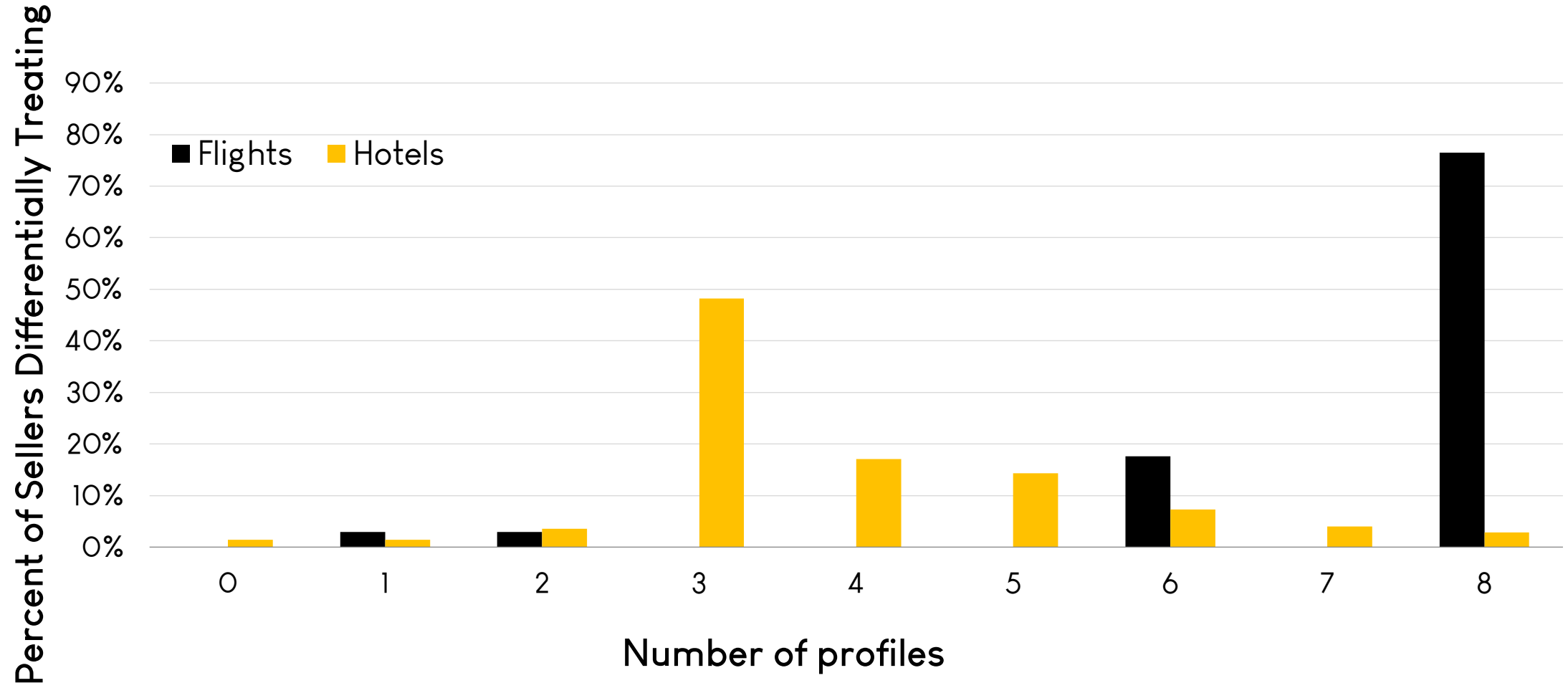
$$\text{Logit}(p_{i,j|k}) = \bar{p} + (\beta_i - \beta_j) + (\delta_{k,i} - \delta_{k,j})$$

\bar{p} = base probability that any buyers see different prices
 $\beta_i - \beta_j$ = difference in direct effect of profiles i and j on
probability of seeing different prices

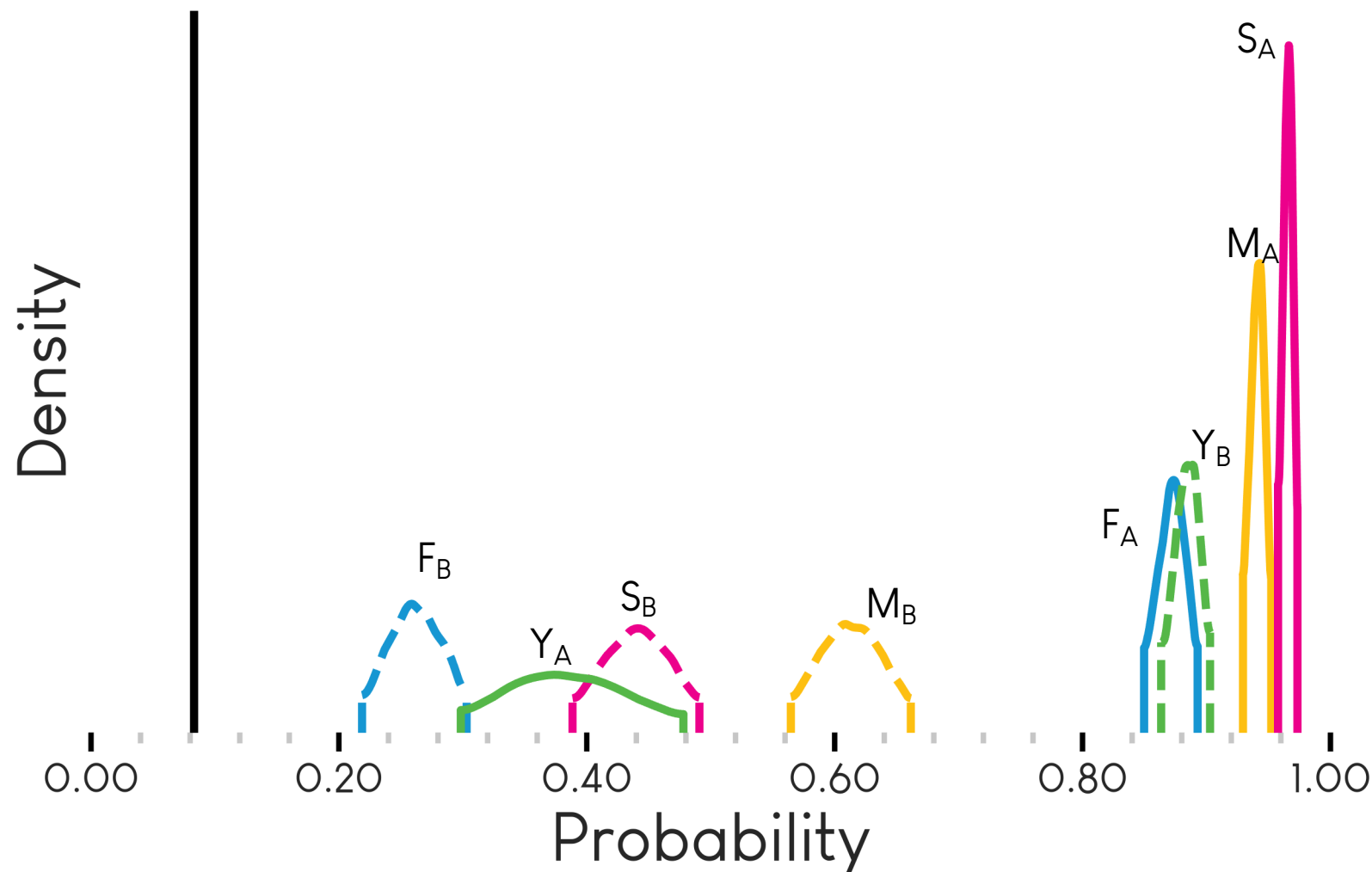
$\delta_{k,i} - \delta_{k,j}$ = difference in seller – profile interaction effect on
probability of seeing different prices



How Common is Differential Treatment Across Sellers?

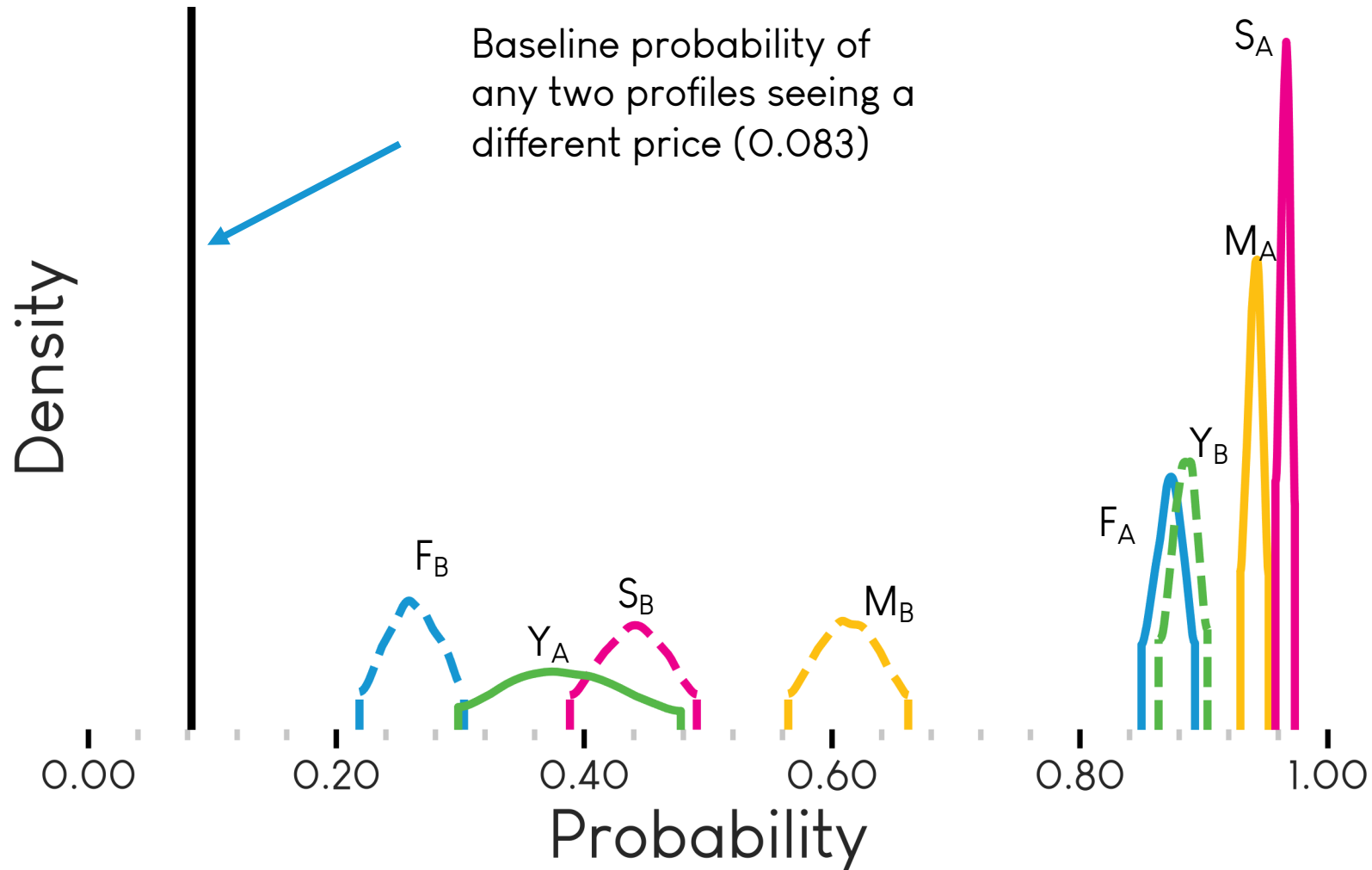


What is the Probability of Differential Treatment for a Specific Seller?

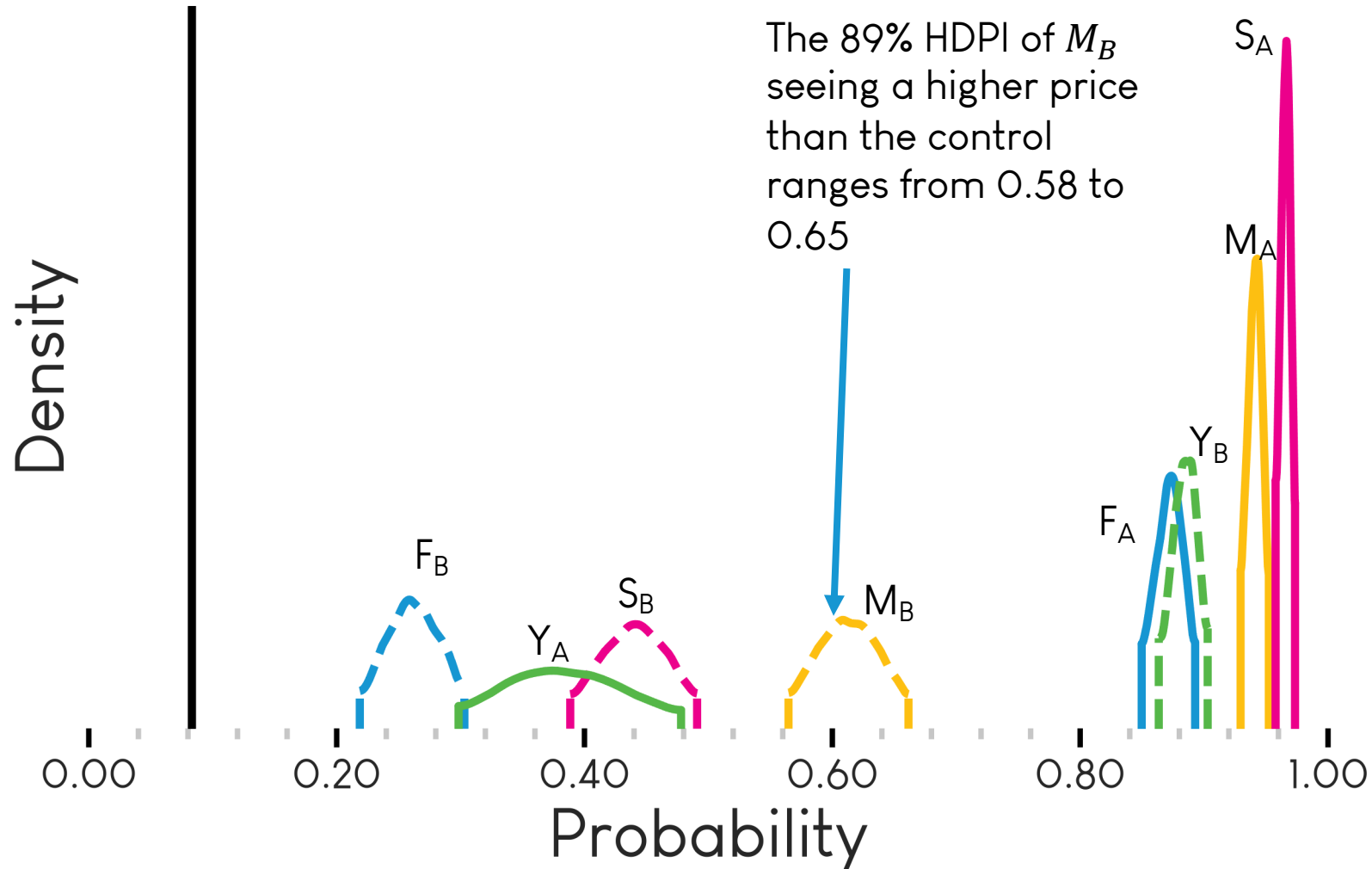


$F_A M_A S_A Y_A$ from Agarwal 2020; $F_B M_B S_B Y_B$ from Asplund 2020

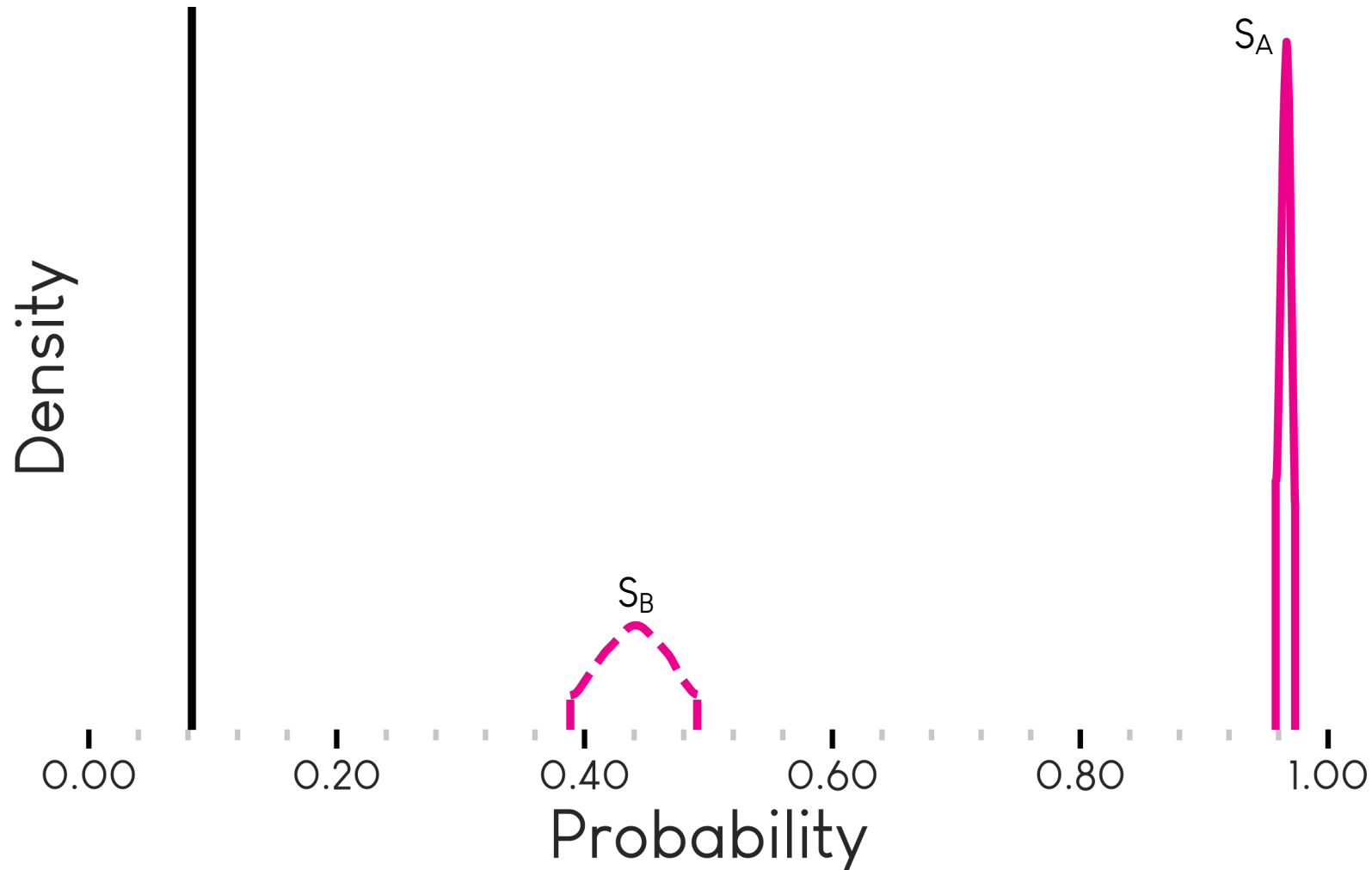
What is the Probability of Differential Treatment for a Specific Seller?



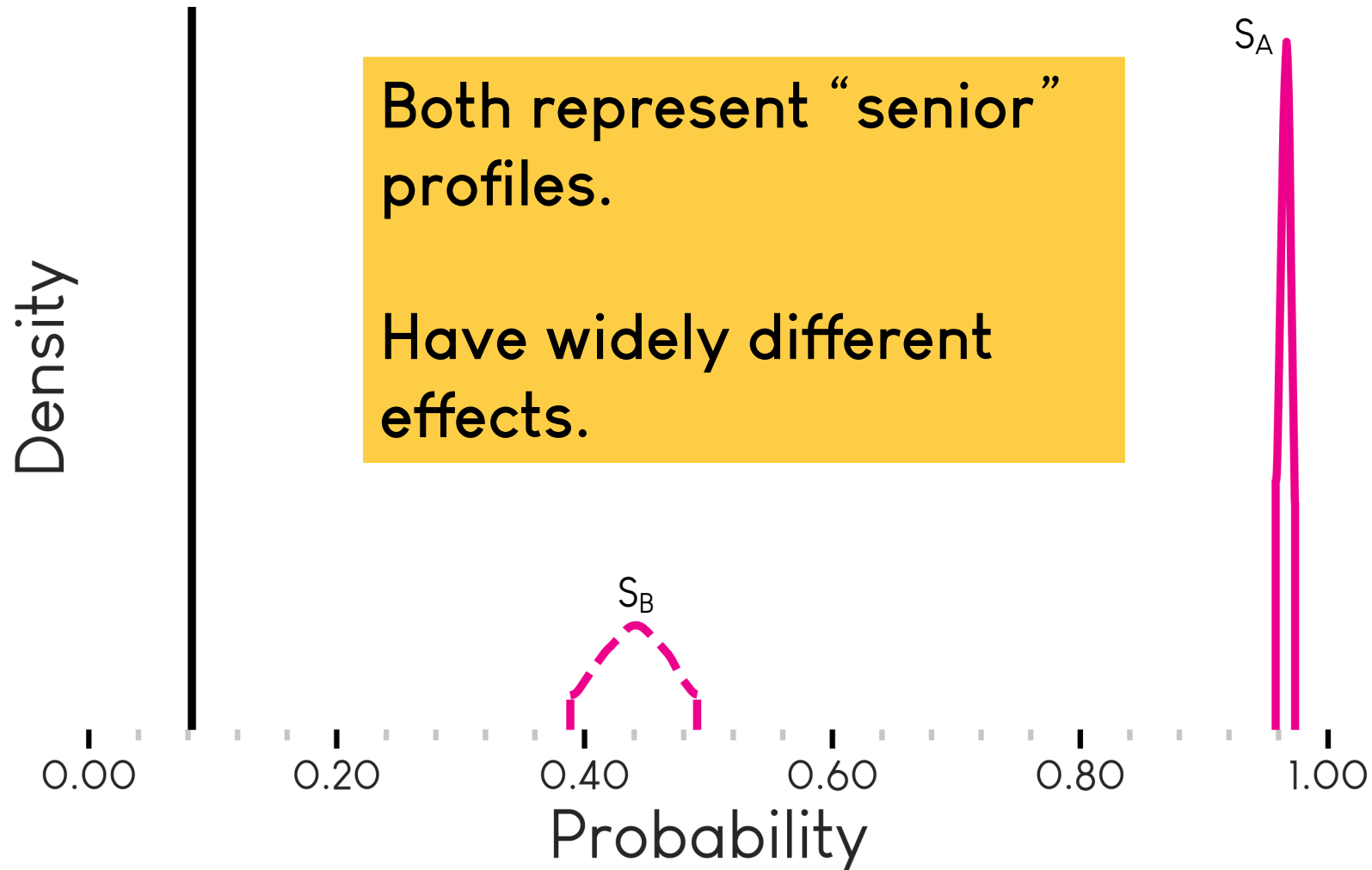
What is the Probability of Differential Treatment for a Specific Seller?



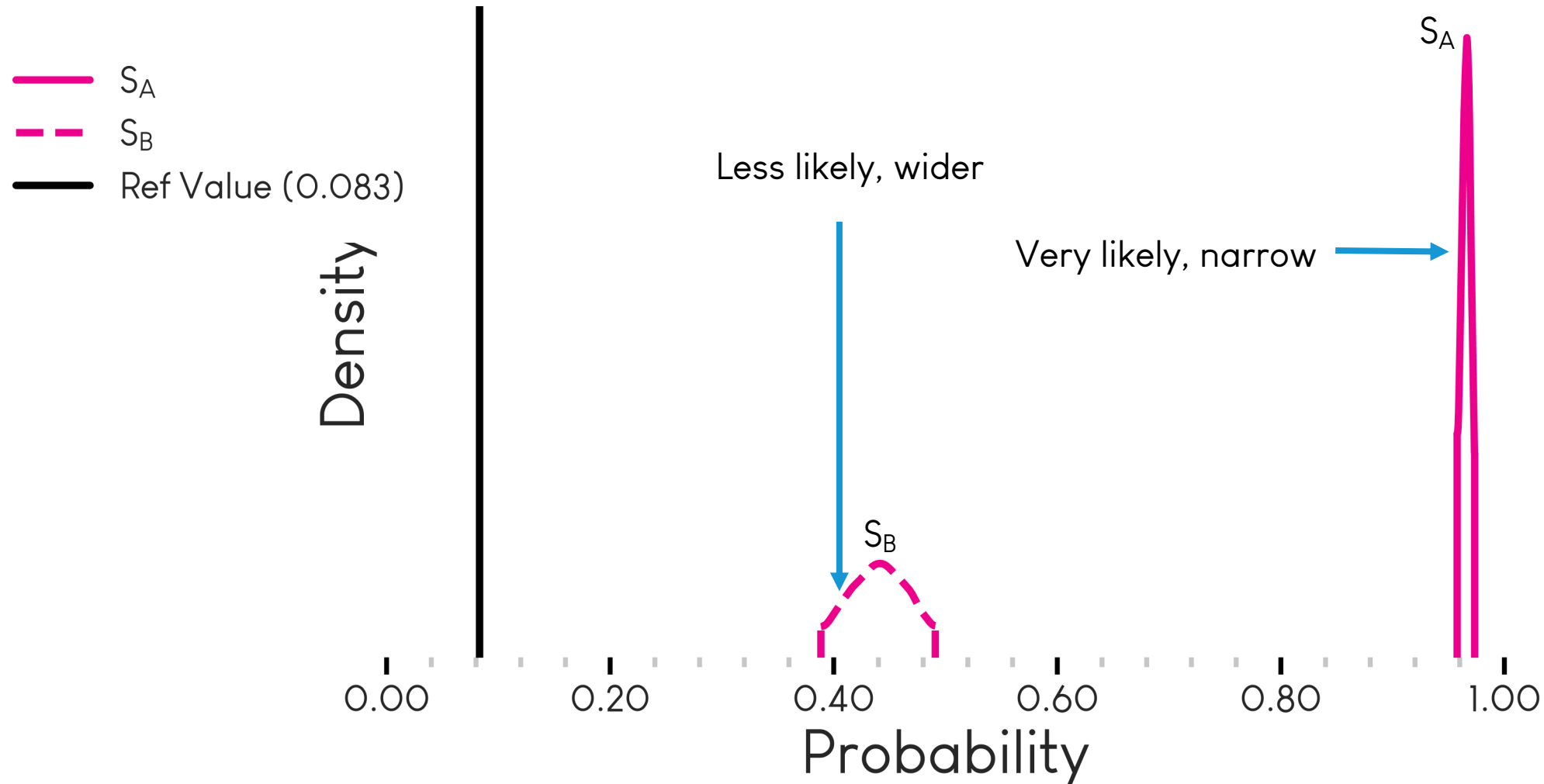
What is the Probability of Differential Treatment for a Specific Seller?



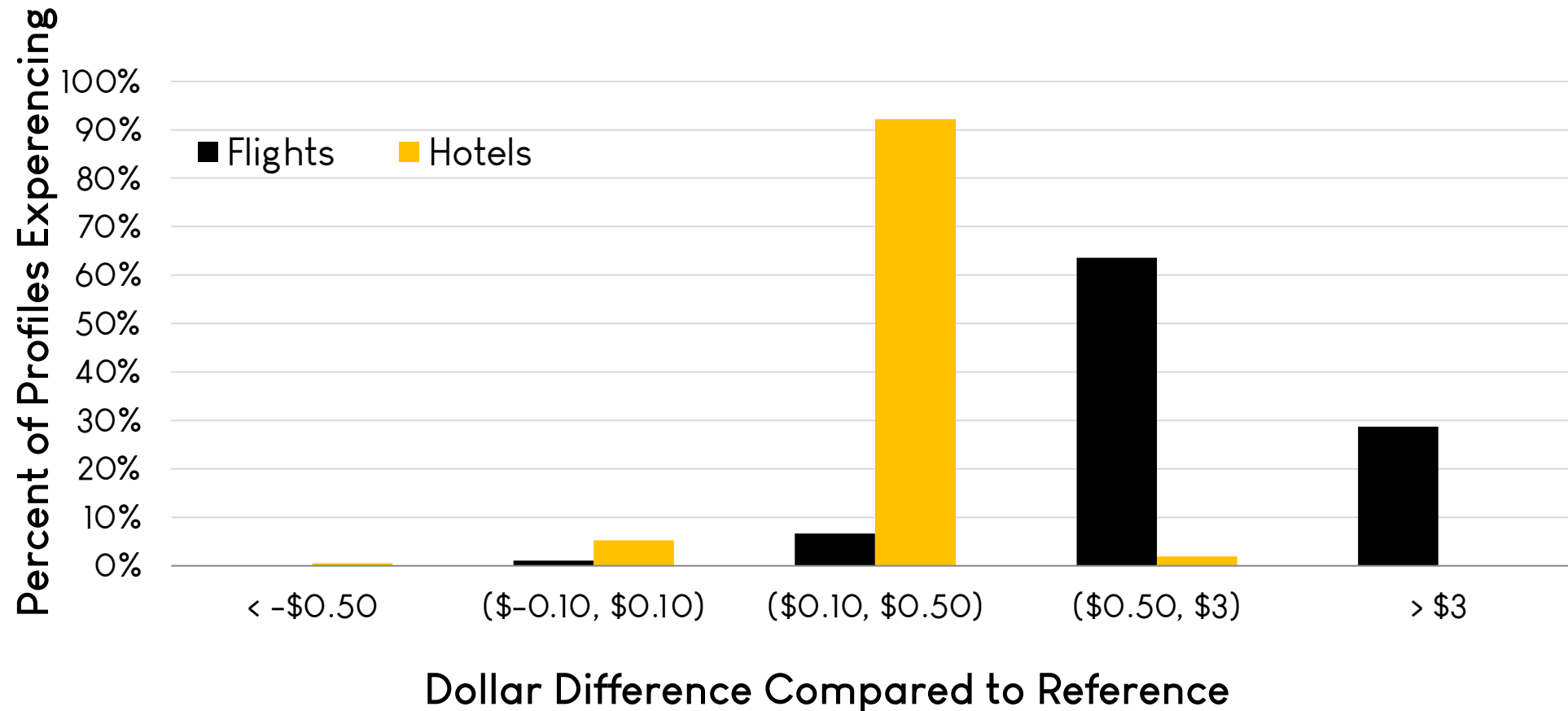
What is the Probability of Differential Treatment for a Specific Seller?



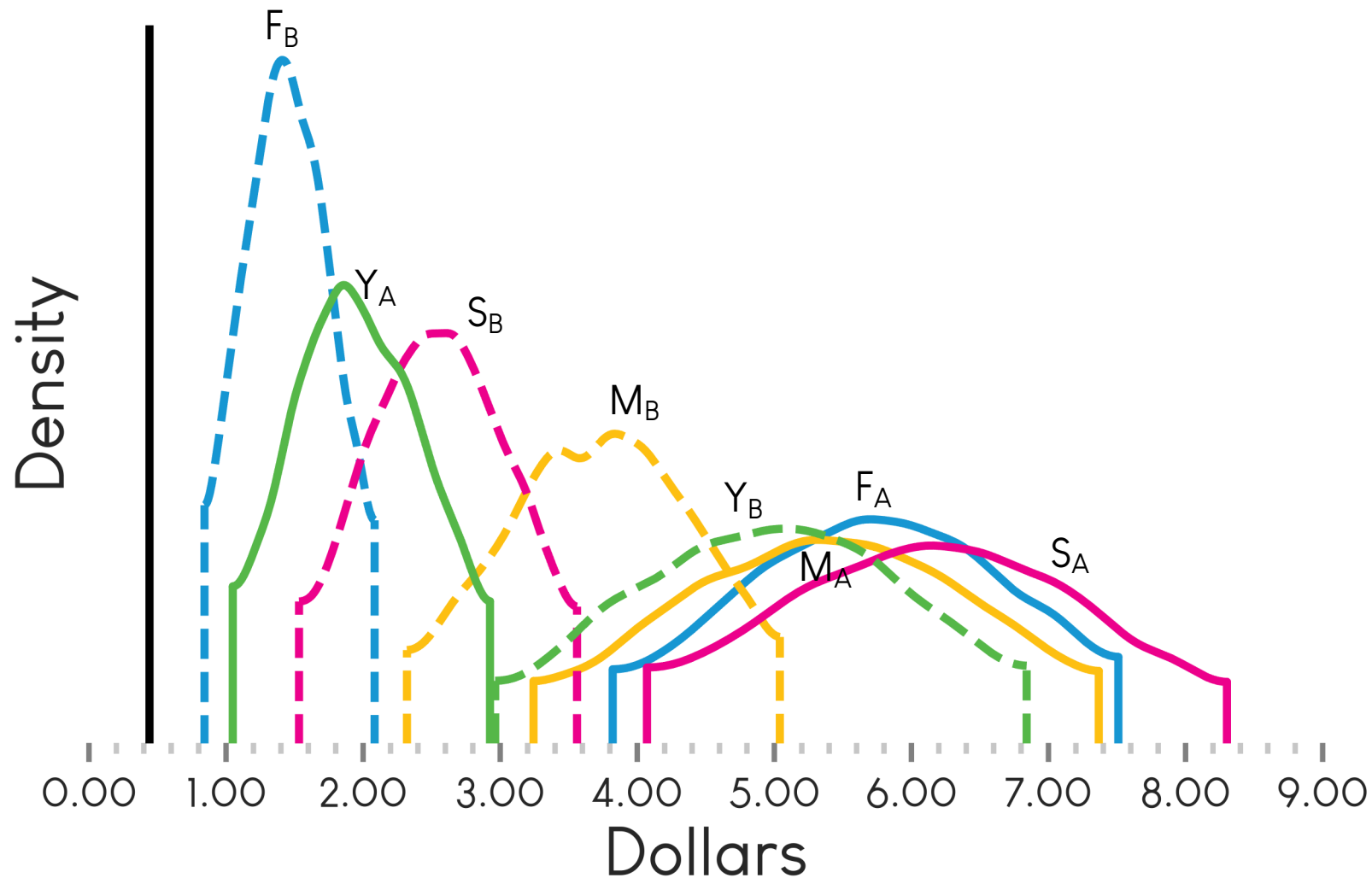
What is the Probability of Differential Treatment for a Specific Seller?



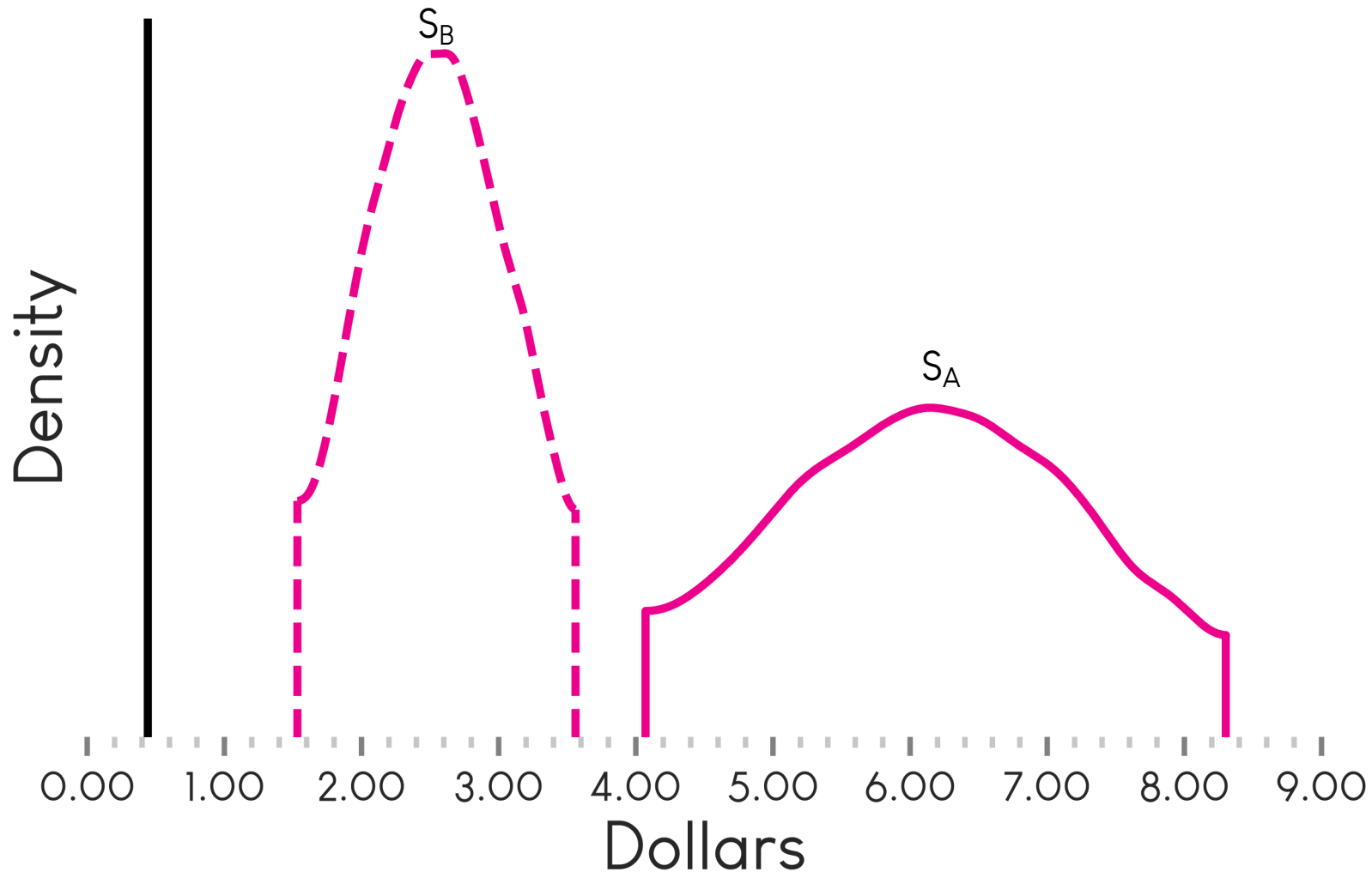
How Much More Do Consumers Pay?



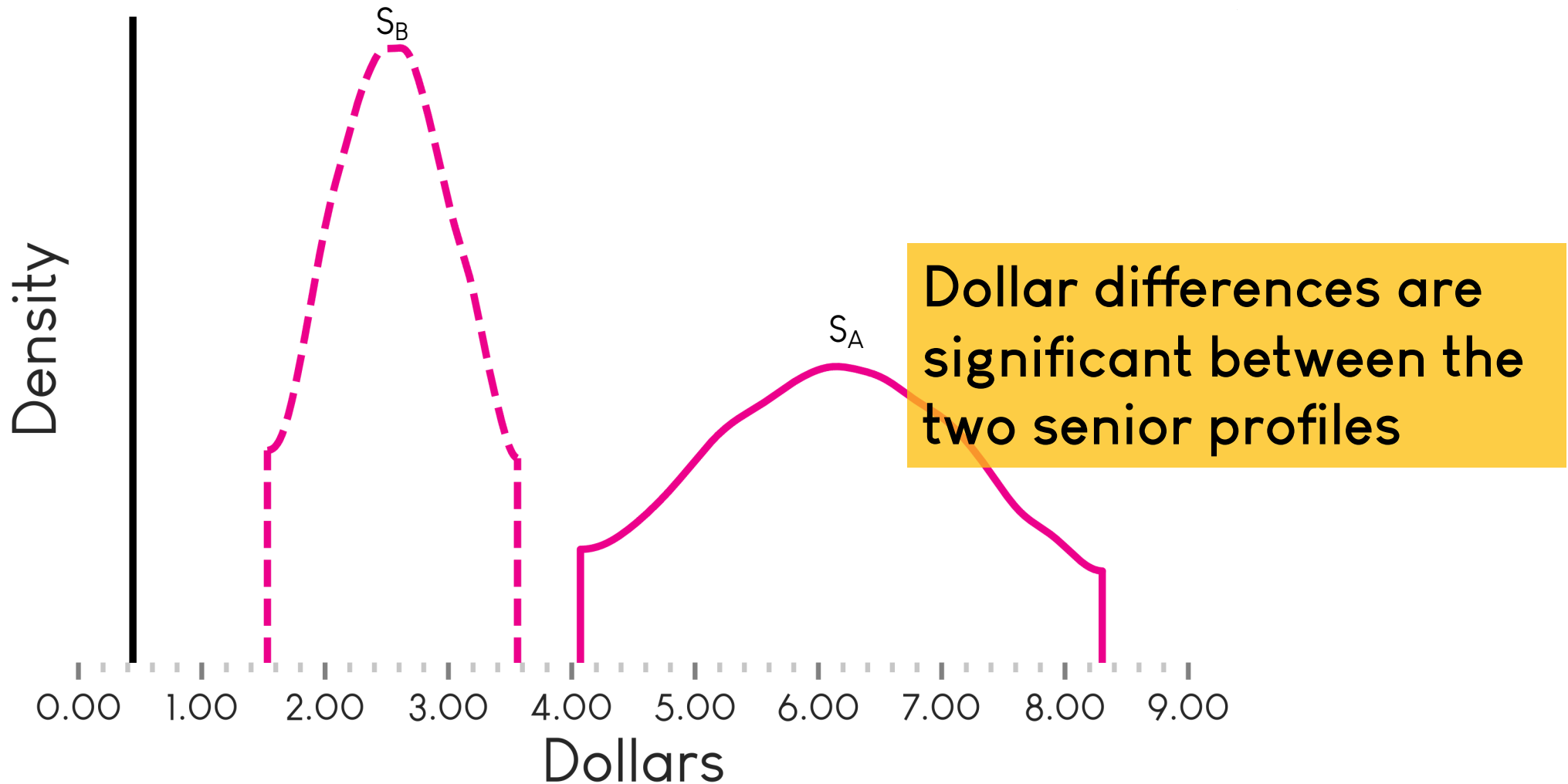
What is the Dollar Impact for a Specific Seller?



What is the Dollar Impact for a Specific Seller?



What is the Dollar Impact for a Specific Seller?



Future Work Can Mitigate Consumer Loss



Laws requiring disclosures



Limit tracking



Consumer cooperation

You may have paid too much

Why?

- 1) Lack of price transparency
- 2) Computational advertising infrastructure