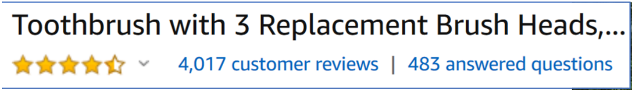
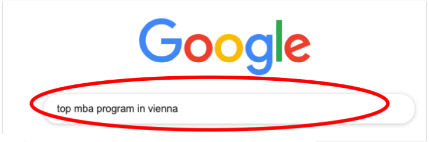
Bridging the Rational and Behavioral Worlds: Rationally Inattentive Decision Making and Implications on Business Operations

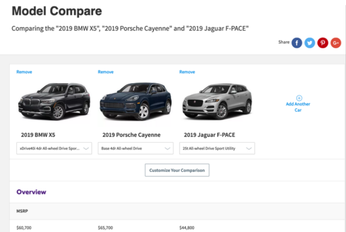
Tamer Boyacı
ESMT Berlin

April 19, 2023

Abundance of Information







Attention Economy

How do we make decisions?

- Rational decision making

How do we make decisions?

- Rational decision making

- Boundedly rational (behavioral) decision making

An Example...

- Selection of brands, models
- Prices, features (resolution, zoom etc.) and performance (color reproduction, etc.)
- Past experience with brands and knowledge of features (priors)
- Actively acquire and process information about choice options
- Make final decision (with incomplete information)

Some Things You Learn...

Fast and with ease

Slow and with difficulty

Which cam is more expensive?

Which cam makes better pictures?

$389

$409

$379

How Much You Learn...

What's at stake - relative gains and losses

$ 379

$ 55, 000

My Broad Research Agenda

My Broad Research Agenda

Model decision-making /choice under limited time and attention

- How much time and attention?
- What information to acquire?
- What choice to make?

My Broad Research Agenda

Model decision-making /choice under limited time and attention

- How much time and attention?
- What information to acquire?
- What choice to make?

Implications on business operations and firm decisions

- Pricing, assortment planning
- Information provisioning
- Service system design

My Broad Research Agenda

Model decision-making /choice under limited time and attention

- How much time and attention?
- What information to acquire?
- What choice to make?

Implications on business operations and firm decisions

- Pricing, assortment planning
- Information provisioning
- Service system design

Human - AI collaboration

- Judgement and decisions
- Accuracy, rate of errors
- Cognitive effort

My Broad Research Agenda

Model decision-making /choice under limited time and attention

- How much time and attention?
- What information to acquire?
- What choice to make?

Implications on business operations and firm decisions

- Pricing, assortment planning
- Information provisioning
- Service system design

Rational
Inattention
Theory

Human - AI collaboration

- Judgement and decisions
- Accuracy, rate of errors
- Cognitive effort

Some Related Publications

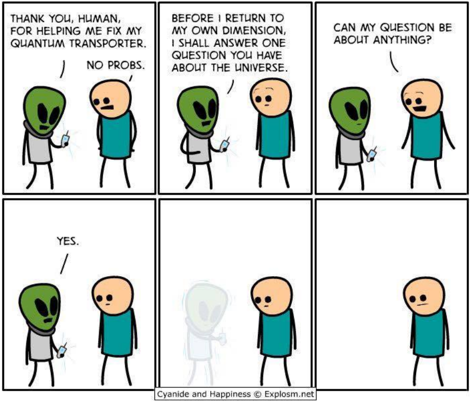
- T. Boyaci, Akcay, A. (2018) "Pricing When Customers Have Limited Attention". *Management Science* 64 (7): 2995–3014.
- Huettner, F.,T. Boyaci, Y. Akcay (2019) "Consumer Choice Under Limited Attention When Alternatives Have Different Information Costs". *Operations Research*, 67 (3), 671-699.
- Canyakmaz, C., T. Boyaci (2023) "Queuing Systems with Rationally Inattentive Customers". *Manufacturing & Service Operations Management*. 25(1), 266-287.
- T. Boyaci, C. Canyakmaz, F. de Vericourt "Human and Machine: The Impact of Machine Input on Human Decision Making Under Cognitive Limitations *Management Science*. Forthcoming.

Outline Going Forward

- A very quick intro to the theory of rational inattention (RI)
- Characterization of decisions under RI (with extensions)
- Impact on firm decisions and operations
  - Product choice and assortments
  - Pricing
- Human-AI collaboration
  - Impact of ML/AI input on decisions and accuracy
  - Cognitive effort

What is Rational Inattention?

- Pioneered by 2011 Nobel Laureate Christopher A. Sims
- DM allocates scarce attention wisely
- DM is free to ask about anything



Rationally Inattentive Decision Making / Choice

- DM optimally chooses type and quantity of information, trading off the benefit of better information and its acquisition cost
- Information is quantified as reduction in Shannon Entropy (H):

$$H(X) = -\sum_i p_i \ln(p_i)$$

- Information costs are based on Shannon Mutual Information:
  - *Difference between entropy of X and entropy of X once Y is known:*

$$I(X, Y) = H(X) - H(X|Y)$$

- Information (Cognitive) Cost: $\lambda \cdot I$

Rationally Inattentive Decision Making / Choice

- $A = \{1, \ldots, n\}$ set of alternatives
- State $\Omega = (\Omega_1 \times \ldots \times \Omega_k \times \ldots \times \Omega_n)$ taking values $\omega \in \mathbb{R}^n$.
- Choosing i in state ω yields $u(i, \omega) \in \mathbb{R}$.
- DM prior belief distribution $g \in \Delta(\Omega)$
- DM can ask questions to sharpen beliefs at unit cost λ
- Information strategy: Joint dist. $f \in \Delta(\Omega \times S)$ of states and signals
  - For any signal, DM chooses option with highest payoff $\rightarrow R(f)$
  - Elicited signal reduces entropy $\rightarrow C(f)$
- Optimization problem of DM: Find f to maximize $R(f) - C(f)$

Rationally Inattentive Decision Making / Choice

- $A = \{1, \ldots, n\}$ set of alternatives
- State $\Omega = (\Omega_1 \times \ldots \times \Omega_k \times \ldots \times \Omega_n)$ taking values $\omega \in \mathbb{R}^n$.
- Choosing i in state ω yields $u(i, \omega) \in \mathbb{R}$.
- DM prior belief distribution $g \in \Delta(\Omega)$
- DM can ask questions to sharpen beliefs at unit cost λ
- Information strategy: Joint dist. $f \in \Delta(\Omega \times \mathrm{S})$ of states and signals
  - For any signal, DM chooses option with highest payoff $\rightarrow R(f)$
  - Elicited signal reduces entropy $\rightarrow C(f)$
- Optimization problem of DM: Find f to maximize $R(f) - C(f)$

Rationally Inattentive Decision Making / Choice

- When the cost of information is the same $\lambda > 0$ for all options, the conditional probability $p\,(i \mid \omega)$ of choosing i follows the *Generalized MNL* formula (Matejka & McKay, AER 2015):

$$\textbf{(GMNL)} \qquad p\,(i \mid \omega) = \frac{e^{\frac{u(i,\omega)}{\lambda}}\,p(i)}{\sum_{j \in A} e^{\frac{u(j,\omega)}{\lambda}}\,p\,(j)} \quad \text{almost surely,}$$

  where $p\,(i) := \int p\,(i \mid \omega)$ are unconditional probabilities that capture the effects of prior beliefs

  If $\lambda = 0$, highest payoff option is chosen with probability 1.

What does it tell us?

$$\textbf{(GMNL)} \quad p\left(i \mid \omega\right) = \frac{e^{\frac{u(i,\omega)}{\lambda}} p(i)}{\sum_{j \in A} e^{\frac{u(j,\omega)}{\lambda}} p\left(j\right)} \quad \text{almost surely,}$$

- The higher the pay-off $u(i,\omega)$, the more likely it will be selected
- An option that is a-priori attractive due to prior beliefs will be selected more (high $p(i)$)

  Such an option can be selected even if its true value is low

- The higher the information cost λ, the less information will be processed and the more choices will be driven by prior beliefs.

What does it tell us?

$$\textbf{(GMNL)} \quad p\left(i \mid \omega\right) = \frac{e^{\frac{u(i,\omega)}{\lambda}} p(i)}{\sum_{j \in A} e^{\frac{u(j,\omega)}{\lambda}} p\left(j\right)} \quad \text{almost surely,}$$

- The higher the pay-off $u(i, \omega)$, the more likely it will be selected
- An option that is a-priori attractive due to prior beliefs will be selected more (high $p(i)$)

  Such an option can be selected even if its true value is low

- The higher the information cost λ, the less information will be processed and the more choices will be driven by prior beliefs.

What does it tell us?

$$\textbf{(GMNL)} \quad p\,(i \mid \omega) = \frac{e^{\frac{u(i,\omega)}{\lambda}}\,p(i)}{\sum_{j \in A} e^{\frac{u(j,\omega)}{\lambda}}\,p\,(j)} \quad \text{almost surely,}$$

- The higher the pay-off $u(i, \omega)$, the more likely it will be selected
- An option that is a-priori attractive due to prior beliefs will be selected more (high $p(i)$)

Such an option can be selected even if its true value is low

- The higher the information cost λ, the less information will be processed and the more choices will be driven by prior beliefs.

What does it tell us?

$$\textbf{(GMNL)} \quad p\left(i \mid \omega\right) = \frac{e^{\frac{u(i,\omega)}{\lambda}} p(i)}{\sum_{j \in A} e^{\frac{u(j,\omega)}{\lambda}} p\left(j\right)} \quad \text{almost surely,}$$

- The higher the pay-off $u(i, \omega)$, the more likely it will be selected
- An option that is a-priori attractive due to prior beliefs will be selected more (high $p(i)$)

  Such an option can be selected even if its true value is low

- The higher the information cost λ, the less information will be processed and the more choices will be driven by prior beliefs.

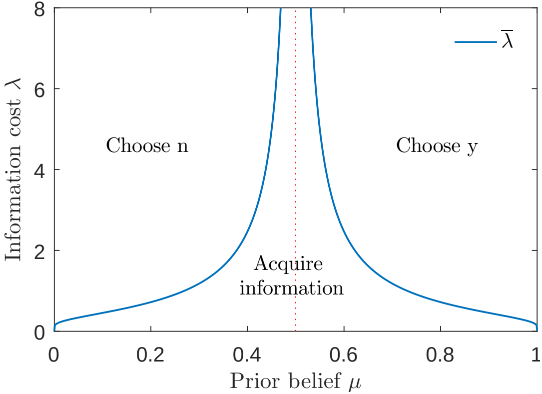
Example: Making The Right Decision (Accuracy)

- State of the world is $\omega \in \Omega = \{g, b\}$ representing "Good" and "Bad"
- DM's prior belief that state is good is μ
- DM needs to choose one of two actions $a \in A = \{y, n\}$
- Immediate payoffs $\Leftrightarrow$ decision accuracy

	$\omega = \mathbf{g}$	$\omega = \mathbf{b}$
a=**y**	1	0
a=**n**	0	1

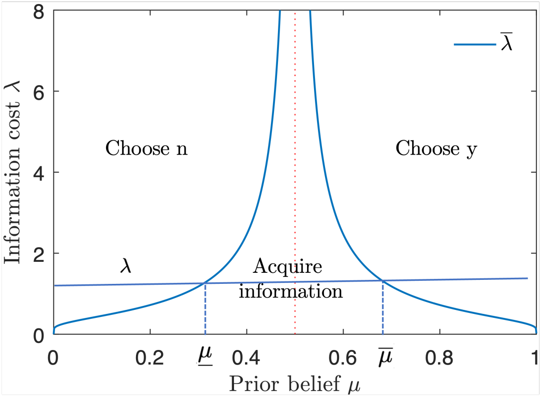
- DM may exert cognitive effort (λ) to refine her belief

Optimal Choice p^* - Cognitive Cost Threshold



- As the DM is more uncertain a-priori (μ is close to 1/2), she is ready to tolerate high cognitive costs to learn more

Optimal Choice p^* - Belief Threshold



- DM processes information only in the range $(\underline{\mu}, \overline{\mu})$
- As the cognitive cost increases, DM relies more on her prior (μ)

Rationally Inattentive Choice When Information Costs Differ

- Let the alternatives be are ordered such that $\lambda_1 \leq \lambda_2 ... \leq \lambda_N$
- How would the DM allocate attention? What is the cost $C(f)$?
  - Distinguish *inferential (implied)* and *direct* information
  - Be efficient (prioritize cheaper channels)
- The information cost $C(f)$ is based on *conditional mutual information*
- **Theorem:** For any information cost $0 < \lambda_1 \leq \lambda_2 .. \leq \lambda_n < \infty$, the optimal conditional choice probabilities satisfy

$$p(i \mid \omega) = \frac{e^{\frac{u(i,\omega)}{\lambda_n}} p(i)^{\frac{\lambda_1}{\lambda_n}} \prod_{k=1}^{n-1} p(i \mid \omega_{1..k})^{\frac{\lambda_{k+1}-\lambda_k}{\lambda_n}}}{\sum_{j \in A} e^{\frac{u(j,\omega)}{\lambda_n}} p(j)^{\frac{\lambda_1}{\lambda_n}} \prod_{k=1}^{n-1} p(j \mid \omega_{1..k})^{\frac{\lambda_{k+1}-\lambda_k}{\lambda_n}}}$$

- Choice probabilities are further adjusted based on what the DM learns (more) about the options with lower cost of information

Rationally Inattentive Choice When Information Costs Differ

- Let the alternatives be are ordered such that $\lambda_1 \leq \lambda_2 ... \leq \lambda_N$
- How would the DM allocate attention? What is the cost $C(f)$?
  - Distinguish *inferential (implied)* and *direct* information
  - Be efficient (prioritize cheaper channels)
- The information cost $C(f)$ is based on *conditional mutual information*
- **Theorem:** For any information cost $0 < \lambda_1 \leq \lambda_2 .. \leq \lambda_n < \infty$, the optimal conditional choice probabilities satisfy

$$p(i \mid \omega) = \frac{e^{\frac{u(i,\omega)}{\lambda_n}} p(i)^{\frac{\lambda_1}{\lambda_n}} \prod_{k=1}^{n-1} p\left(i \mid \omega_{1..k}\right)^{\frac{\lambda_{k+1}-\lambda_k}{\lambda_n}}}{\sum_{j \in A} e^{\frac{u(j,\omega)}{\lambda_n}} p(j)^{\frac{\lambda_1}{\lambda_n}} \prod_{k=1}^{n-1} p\left(j \mid \omega_{1..k}\right)^{\frac{\lambda_{k+1}-\lambda_k}{\lambda_n}}}$$

- Choice probabilities are further adjusted based on what the DM learns (more) about the options with lower cost of information

Example: Impact of Products With Lower Information Cost

Example: Impact of Products With Lower Information Cost

You believe that Sony cameras are of high quality in general ($p(Sony)$ will be high)

Example: Impact of Products With Lower Information Cost

You believe that Sony cameras are of high quality in general ($p(Sony)$ will be high)

It is hard to gain info about the Sony camera, but there is a very similar Canon camera ($\lambda_{Canon} < \lambda_{Sony}$)

Example: Impact of Products With Lower Information Cost

You believe that Sony cameras are of high quality in general ($p(Sony)$ will be high)

It is hard to gain info about the Sony camera, but there is a very similar Canon camera ($\lambda_{Canon} < \lambda_{Sony}$)

You study Canon camera and decide you don't like the digital zoom. You believe Sony camera is similar so ($p(Sony|\omega_{Canon})$ will be low)

Example: Impact of Products With Lower Information Cost

You believe that Sony cameras are of high quality in general ($p(Sony)$ will be high)

It is hard to gain info about the Sony camera, but there is a very similar Canon camera ($\lambda_{Canon} < \lambda_{Sony}$)

You study Canon camera and decide you don't like the digital zoom. You believe Sony camera is similar so ($p(Sony|\omega_{Canon})$ will be low)

In reality, Sony camera's digital zoom is very good (ω_{Sony} is high)

Example: Strong Failure of Regularity

You are well informed about Nikon Sony can be better or worse

	State 1	State 2
Nikon	1	1
Sony	0.6	1.2

Example: Strong Failure of Regularity

You are well informed about Nikon Sony can be better or worse

	State 1	State 2
Nikon	1	1
Sony	0.6	1.2

Hard to gain info on Sony ($\lambda = 1$) Sony is never selected 0%

Example: Strong Failure of Regularity

You are well informed about Nikon Sony can be better or worse

	State 1	State 2
Nikon	1	1
Sony	0.6	1.2

Hard to gain info on Sony ($\lambda = 1$) Sony is never selected 0%

An inferior Sony (e.g., no flash) is included in choice set ($\lambda' = 0.2$)

	State 1	State 2
Sony Inferior	0.5	0.9

Example: Strong Failure of Regularity

You are well informed about Nikon Sony can be better or worse

	State 1	State 2
Nikon	1	1
Sony	0.6	1.2

Hard to gain info on Sony ($\lambda = 1$) Sony is never selected 0%

An inferior Sony (e.g., no flash) is included in choice set ($\lambda' = 0.2$)

	State 1	State 2
Sony Inferior	0.5	0.9

Sony camera is chosen 29%
(Inferior Sony never selected)

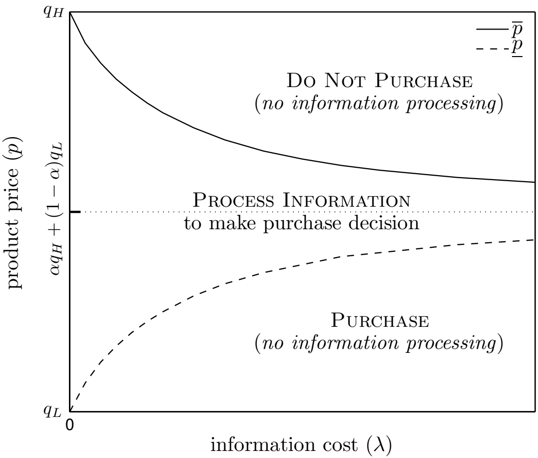
Pricing for Rationally Inattentive Customers

Buy or Not?

- Product price p (firm's decision; fully observed by customers)
- Customers not fully aware of the true quality (q) of the product
  - ➠ $v = q - p$
  - ➠ q : (High quality – q_H) or (Low quality – q_L)
- Customers' prior beliefs:
  - ➠ q_H with probability μ and q_L with $1 - \mu$
- Rationally inattentive customers with cost of information λ
- No-purchase option $v = 0$

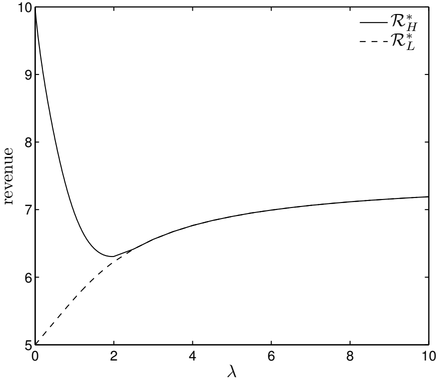
Pricing for Rationally Inattentive Customers

- Purchasing and information processing strategy



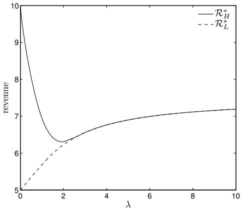
Impact of Costly Information (Too Little Time)

- Example: $q_H = 10$, $q_L = 5$, $\mu = 0.5$



- Information provision: search vs experience vs credence goods

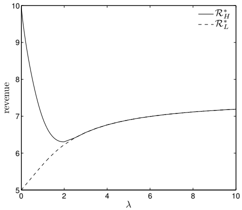
Impact of Costly Information (Too Little Time)



What are the implications?

- Prices of highly experiential and credence goods should converge regardless of the quality
  - ➠ Your mechanic/doctor should overcharge for simple procedures and undercharge for complicated procedures
  - ➠ They should obscure, conceal, blur information
- Seller of search goods
  - ➠ Low quality:Obscure, conceal, blur information
  - ➠ High quality:Proactively reveal information

Impact of Costly Information (Too Little Time)

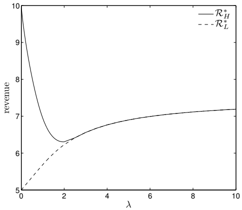


What are the implications?

- Prices of highly experiential and credence goods should converge regardless of the quality
  - ➠ Your mechanic/doctor should overcharge for simple procedures and undercharge for complicated procedures
  - ➠ They should obscure, conceal, blur information

- Seller of search goods
  - ➠ Low quality:Obscure, conceal, blur information
  - ➠ High quality:Proactively reveal information

Impact of Costly Information (Too Little Time)



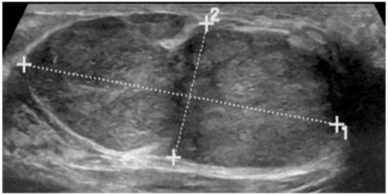
What are the implications?

- Prices of highly experiential and credence goods should converge regardless of the quality
  - Your mechanic/doctor should overcharge for simple procedures and undercharge for complicated procedures
  - They should obscure, conceal, blur information

- Seller of search goods
  - Low quality:Obscure, conceal, blur information
  - High quality:Proactively reveal information

Human - Machine/AI Collaboration

- Breast ultrasound - Detecting malignant tumors

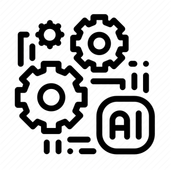


Example of a true negative.

- The Human: MD
- The Machine: Deep learning image analysis

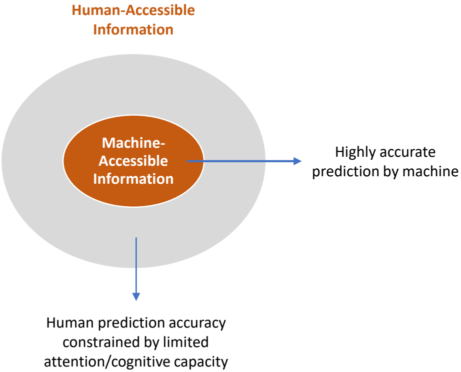
Humans and Machines Are Complementary

- Flexible - can assess any information
- Limited cognitive capacity



- Rigid - extract a limited subset of information
- Immense computing power

Fundamental Questions



- What is the impact of machine on human decisions?
- What is the impact of machine on accuracy & nature of errors?
- What is the impact of machine on cognitive effort spent?

The Task: Making The Right Decision (Accuracy)

- State of the world is $\omega \in \Omega = \{g, b\}$ representing "Good" and "Bad"
- DM's prior belief that state is good is μ
- DM needs to choose one of two actions $a \in A = \{y, n\}$
- Baseline model: Immediate payoffs $\Leftrightarrow$ decision accuracy

	$\omega =$ **g**	$\omega =$ **b**
a=**y**	1	0
a=**n**	0	1

- DM is rationally inattentive and may exert cognitive effort (λ) to refine her belief

Human Machine Collaboration

- Information sources are partitioned into two distinct subsets, $X_1, X_2 \in \{+, -\}$, one of which only human can assess

$$\underbrace{\underbrace{X_1}_{\substack{\text{machine can} \\ \text{assess } \textbf{perfectly}}} \qquad \underbrace{X_2}_{\substack{\text{machine cannot} \\ \text{assess}}}}_{\substack{\text{human can assess both} \\ \text{under cognitive limitations}}}$$

- True state is "good" only if both sources indicate a positive outcome $\rightarrow$ DM's prior belief that state is good is $\mu = Prob(x_1 = +, x_2 = +)$

- Given the machine's evaluation $x_1 \in \{-, +\}$, DM updates belief to μ^x:

$$\mu^- = 0 \text{ and } \mu^+ = \frac{\mu}{\mu + \pi(+, -)} > \mu.$$

Human Machine Collaboration

- Information sources are partitioned into two distinct subsets, $X_1, X_2 \in \{+, -\}$, one of which only human can assess

$$\underbrace{\overset{X_1}{\overbrace{\text{machine can assess \textbf{perfectly}}}} \quad \overset{X_2}{\overbrace{\text{machine cannot assess}}}}_{\text{human can assess both under cognitive limitations}}$$

- True state is "good" only if both sources indicate a positive outcome $\rightarrow$ DM's prior belief that state is good is $\mu = Prob(x_1 = +, x_2 = +)$

- Given the machine's evaluation $x_1 \in \{-, +\}$, DM updates belief to μ^x:

$$\mu^- = 0 \text{ and } \mu^+ = \frac{\mu}{\mu + \pi(+,-)} > \mu.$$

Impact of Machine on Human Decision

- $A^\star$, $V^\star$, $C^\star$, $p^\star$, $\alpha^\star$, $\beta^\star$
- $A^\star_m$, $V^\star_m$, $C^\star_m$, $p^\star_m$, $\alpha^\star_m$, $\beta^\star_m$

A: Accuracy

V: Objective Value

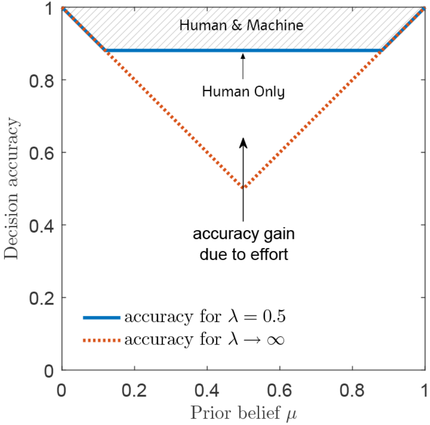
C: Cognitive Cost

p: Decision (Probability of choosing $a = y$)

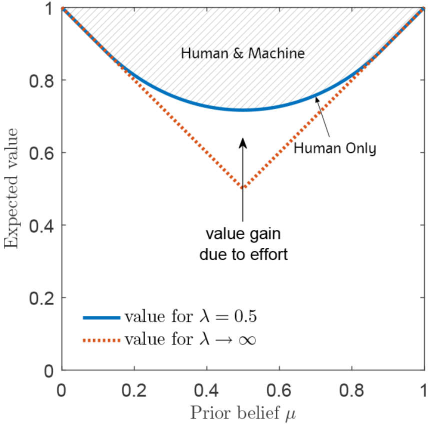
α: False positive rate

β: False negative rate

Impact on Accuracy and Value



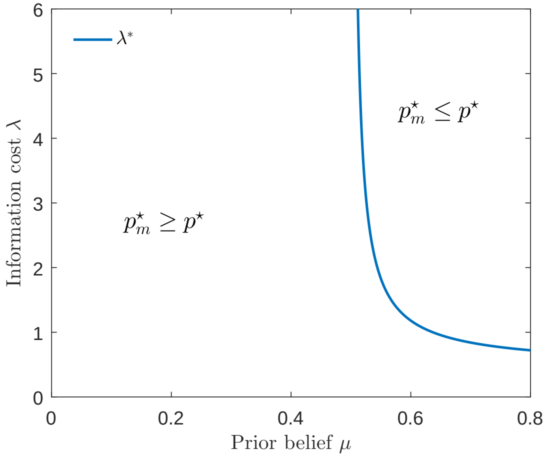
(a) Decision Accuracy



(b) Expected Value

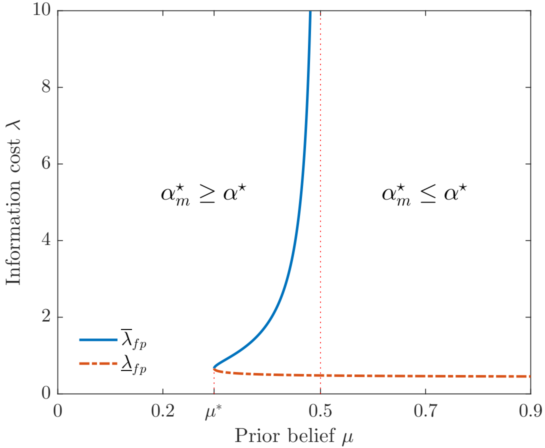
- DM's decision accuracy and expected utility **always increase** with machine

Impact on Choice Probability



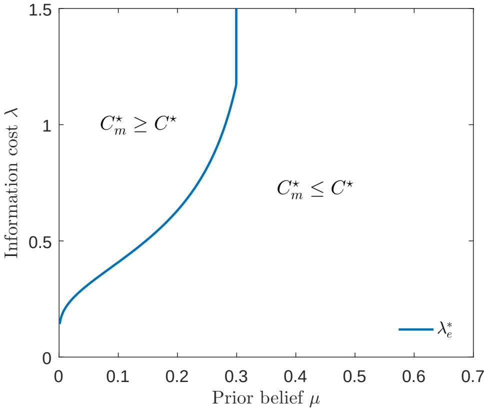
- The machine may increase the variability of the DM's decision

Impact on False Positive Errors



- When the DM sufficiently favors the bad state, a positive assessment may make her more uncertain or favor the good state

Impact on Cognitive Effort



- The machine can increase cognitive effort, especially when
  - The good state is less likely (low μ)
  - DM is cognitively constrained (high λ)

Human - Machine / AI Collaboration

- Overall accuracy is improved due to collaboration
- Collaboration most beneficial for identifying a relatively likely state
  - Errors are reduced
  - "Efficiency" of the DM is improved
- Collaboration less beneficial for identifying a relatively unlikely state, especially when the DM is cognitively constrained
  - False positive conclusions increase
  - "Efficiency" of the DM is reduced
- Results are robust
  - Generalized pay-off structures
  - Mistrust against the machine or the machine is imprecise

In Conclusion

- Rational inattention is a powerful theory for decision-making under
  - Limited time and attention
  - Limited information processing capacity
- Analytical characterizations of optimal decisions can be derived
- Empirical estimation and validation are developing
- Many academically and practically relevant applications
  - Assortment optimization
  - Pricing
  - Services ...
- Basis for modeling Human - AI collaboration
  - Impact on decisions and error rates
  - Impact on cognitive effort