Encoding-decoding of numbers explains biased judgments

Arthur Prat-Carrabin

2019 Sloan-Nomis Workshop on the Cognitive Foundations of Economic Behavior
02/23/2019

Encoding-decoding of numbers explains biased judgments

- Economic decisions often require the aggregation of multiple sources of information, a simple example of which is the computation of the average of some numbers.
- Human subjects averaging numbers in a comparison task, seem to *overweight* some numbers in comparison to others[1].

[1] *B. Spitzer, L. Waschke, and C. Summerfield, "Selective overweighting of larger magnitudes during noisy numerical comparison," Nature Human Behavior 1, art. 0145 (2017).*

Encoding-decoding of numbers explains biased judgments

- A possibility is that this selective weighting originates in the way the brain *encodes and decodes* presented stimuli.

- *Efficient* encoding: should be *adapted to the prior distribution of stimuli*[1].

➡ We design an average-comparison task, in which different prior distributions of numbers are used in different blocks of trials.

[1]*X.X. Wei and A.A. Stocker, "A Bayesian observer model constrained by efficient coding can explain 'anti-Bayesian' percepts," Nature Neuroscience 18: 1509 (2015).*

Outline

- Experimental design
- Behavioral data
  Models of noisy estimation
- Encoding-decoding models of estimation

Experimental design

- 10 numbers, alternating in color between red and green, presented in rapid succession (500ms).
- Each number is within the range [10.00, 99.99] and has two decimal points.
- Subjects choose whether the red numbers or green numbers have the larger average.

79.60

44.92

95.12

66.29

35.54

87.07

38.93

69.19

51.07

12.07

R

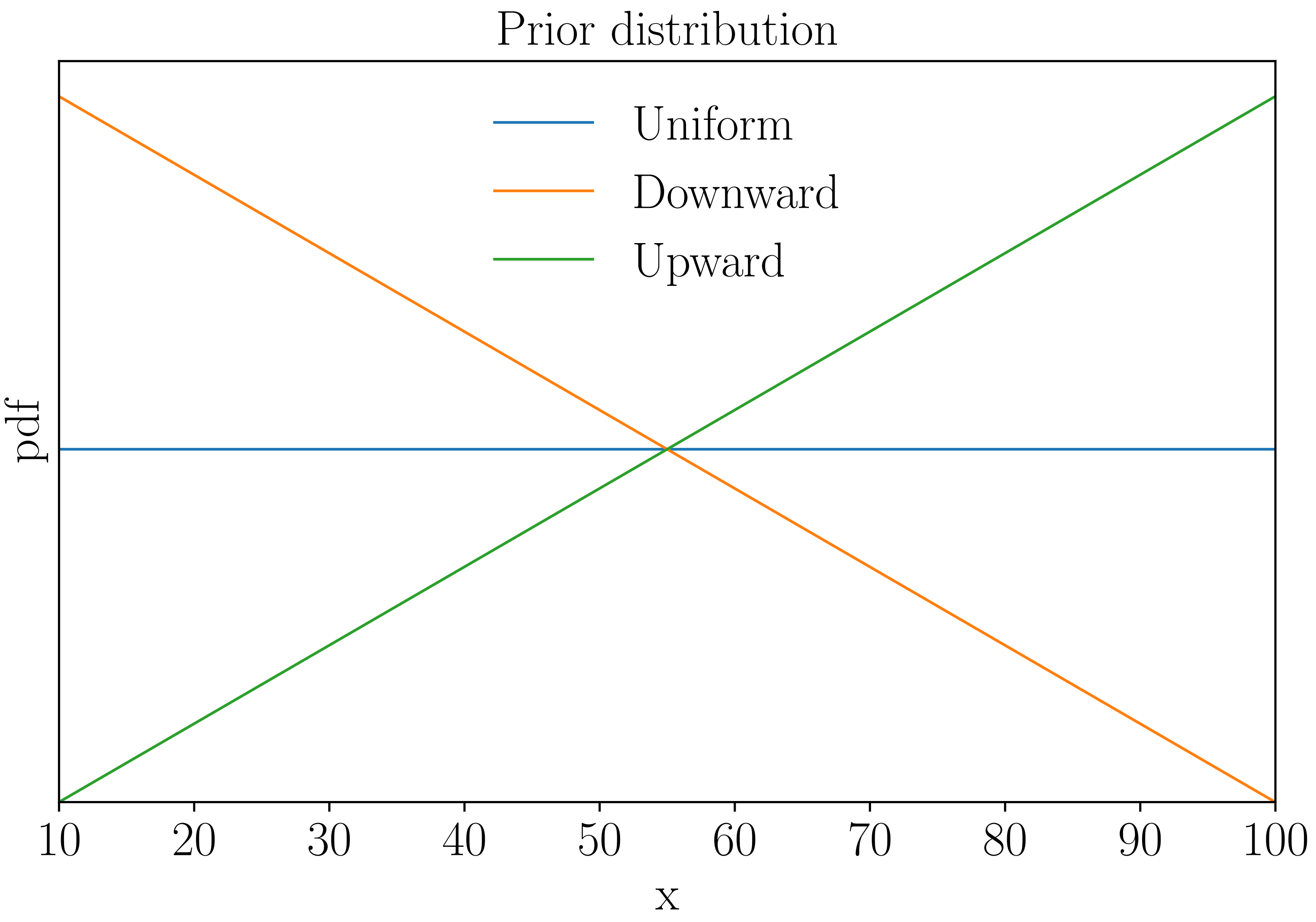
G

You gained 60.05

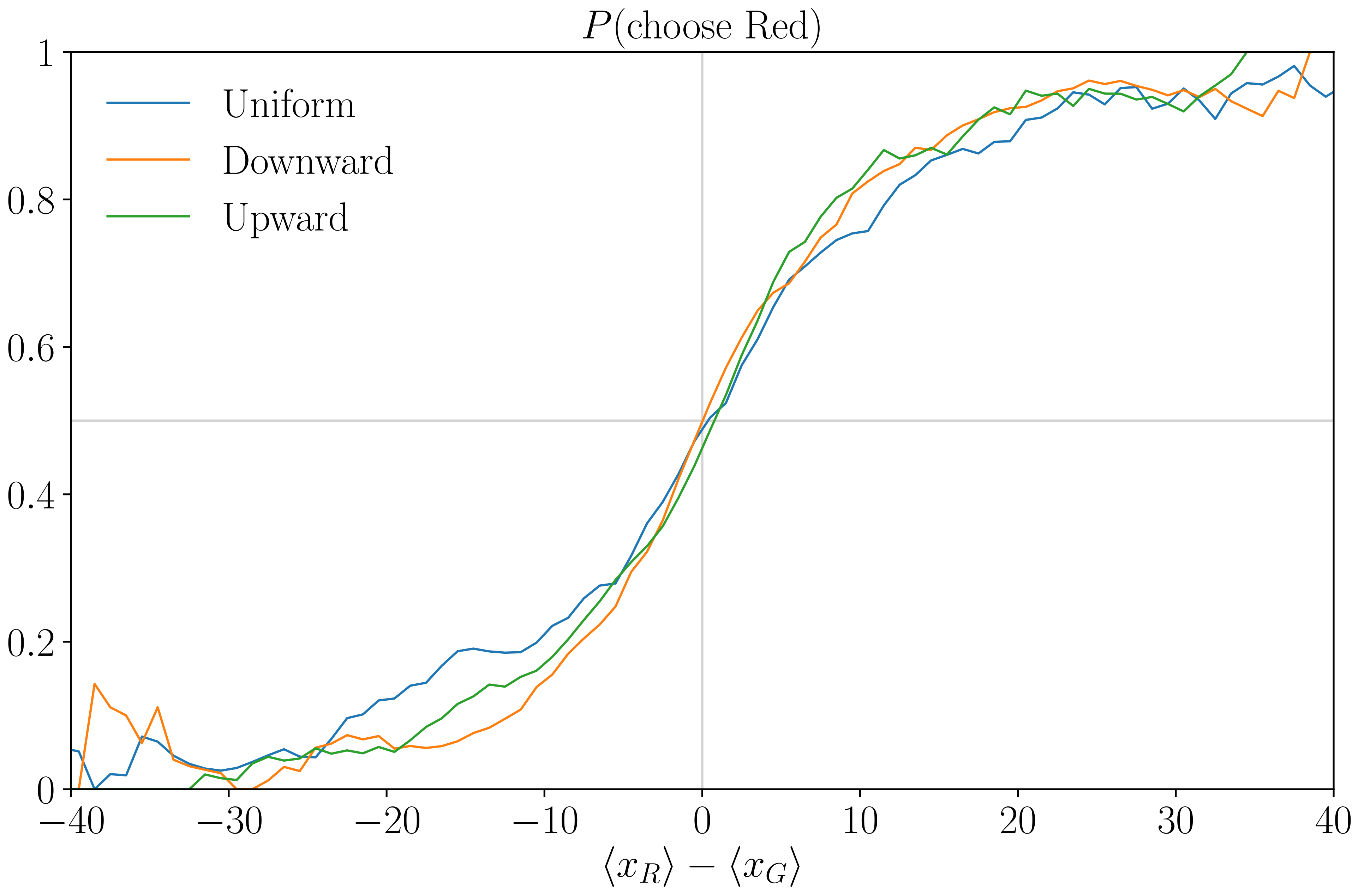
You would have gained 55.91

Experimental design

- In different blocks of trials, numbers are sampled from different prior distributions.

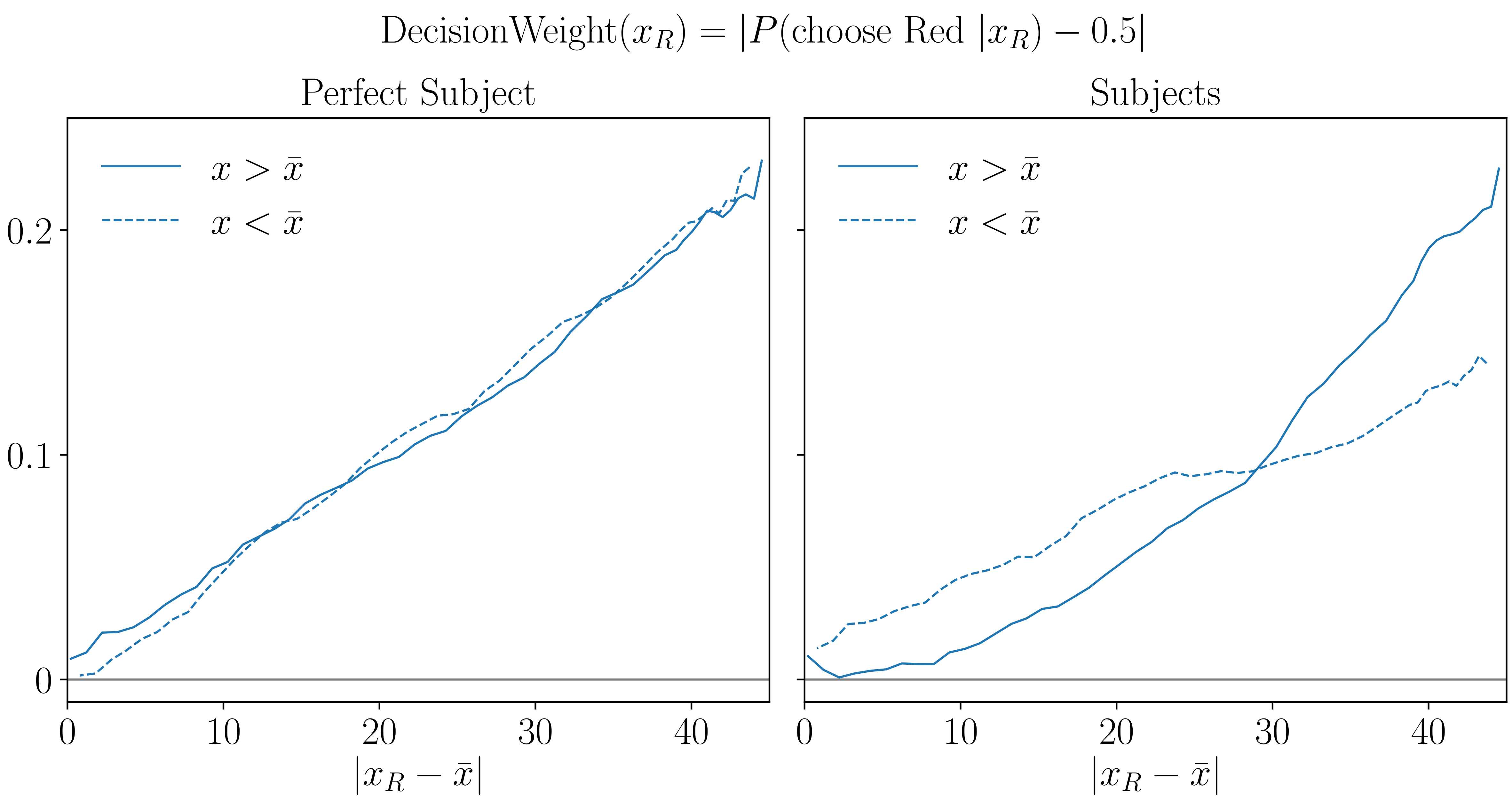


Results



- This suggests there is noise in the decision process.

Results



- Different numbers seem to be weighted differently in the decision process.

Models of subjective estimates

- Errors in decision suggest a model of noisy estimation :

- We consider :

Models of subjective estimates

- Errors in decision suggest a model of noisy estimation :

- We consider :

$$\hat{x} \mid x \sim N(x,\ s^2)$$

The number is perceived with noise.

Models of subjective estimates

- Errors in decision suggest a model of noisy estimation :

- We consider :

$\hat{x} \mid x \sim N(x,\ s^2)$	The number is perceived with noise.
$N(m(x),\ s^2)$	A transformation of the number is observed with noise. It should capture unequal weighting, and improve accuracy[1].

[1] *B. Spitzer, L. Waschke, and C. Summerfield, "Selective overweighting of larger magnitudes during noisy numerical comparison," Nature Human Behavior 1, art. 0145 (2017).*

Models of subjective estimates

- Errors in decision suggest a model of noisy estimation :

- We consider :

$\hat{x} \mid x \sim N(x,\ s^2)$	The number is perceived with noise.
$N(m(x),\ s^2)$	A transformation of the number is observed with noise. It should capture unequal weighting, and improve accuracy[1].
$N(x,\ s^2(x))$	Different numbers are perceived with different amounts of noise.

[1] *B. Spitzer, L. Waschke, and C. Summerfield, "Selective overweighting of larger magnitudes during noisy numerical comparison," Nature Human Behavior 1, art. 0145 (2017).*

Models of subjective estimates

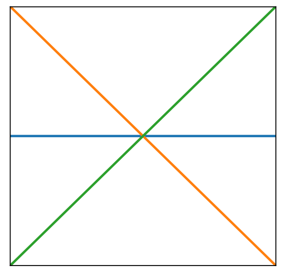
- Errors in decision suggest a model of noisy estimation :

- We consider :

$\hat{x}\,\lvert x \sim N(x,\ s^2)$	The number is perceived with noise.
$N(m(x),\ s^2)$	A transformation of the number is observed with noise. It should capture unequal weighting, and improve accuracy[1].
$N(x,\ s^2(x))$	Different numbers are perceived with different amounts of noise.
$N(m(x),\ s^2(x))$	A transformation of the number is observed, with varying noise.

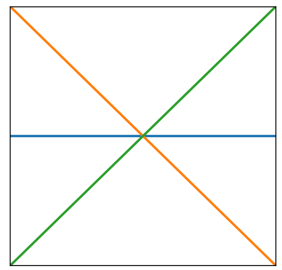
[1] *B. Spitzer, L. Waschke, and C. Summerfield, "Selective overweighting of larger magnitudes during noisy numerical comparison," Nature Human Behavior 1, art. 0145 (2017).*

Models of subjective estimates



	Same parameters	Prior-specific parameters
$\hat{x} \mid x \sim N(x,\ s^2)$		
$N(m(x),\ s^2)$		
$N(x,\ s^2(x))$		
$N(m(x),\ s^2(x))$		

Models of subjective estimates



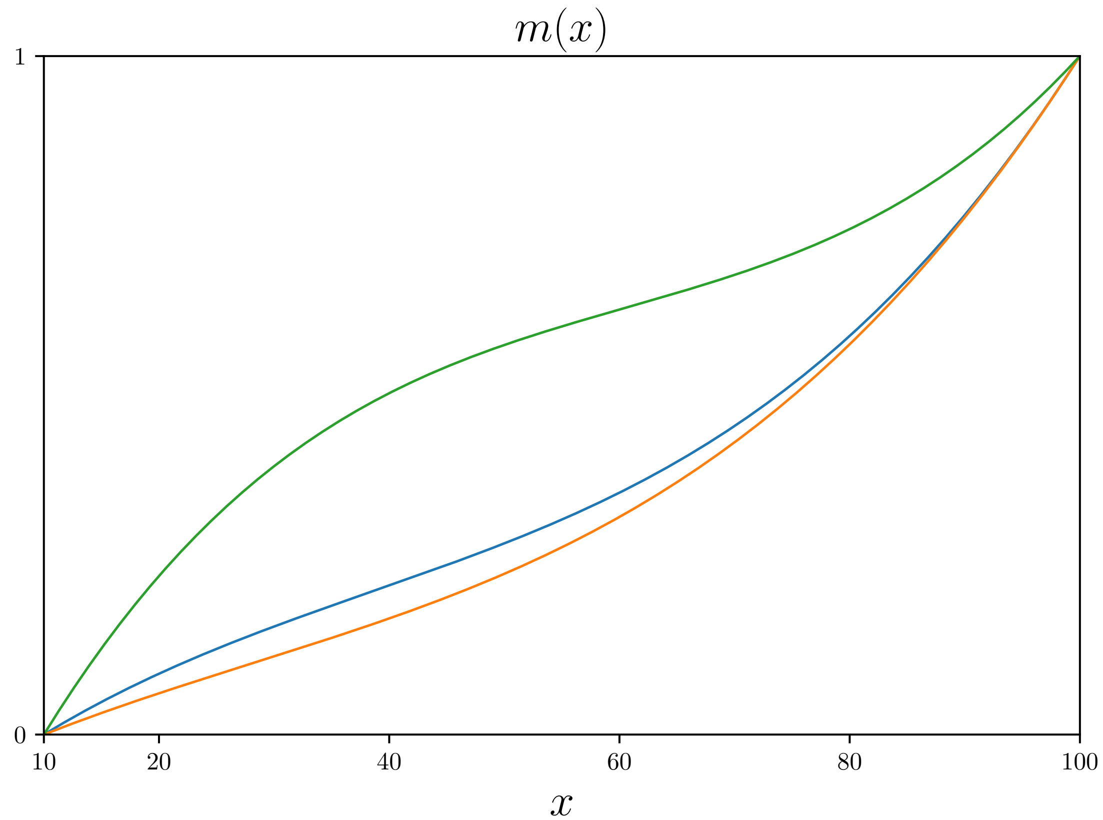
ΔBIC	Same parameters	Prior-specific parameters
$\hat{x} \mid x \sim N(x,\ s^2)$	262	212
$N(m(x),\ s^2)$	140	75
$N(x,\ s^2(x))$	262	203
$N(m(x),\ s^2(x))$	66	**0**

ΔBIC with best model

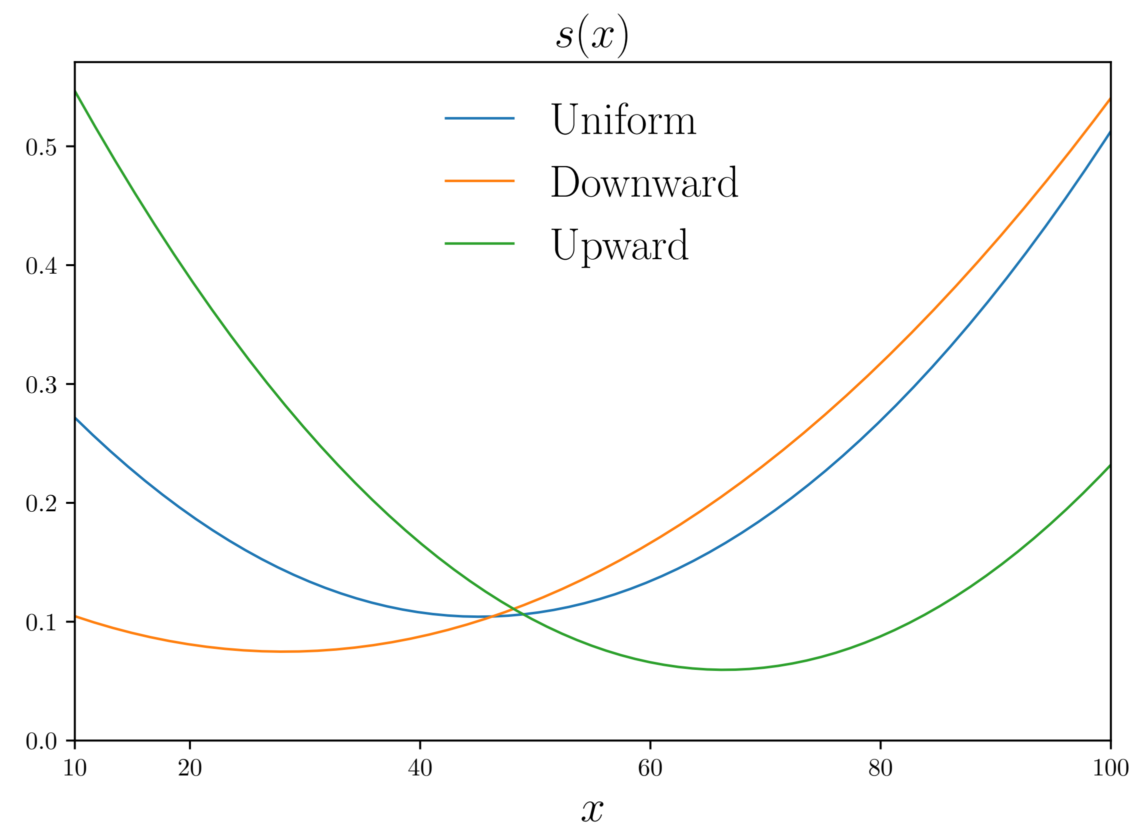
Transformation + varying noise

$$\hat{x} \mid x \sim N(m(x), s^2(x))$$

Transformation of the number



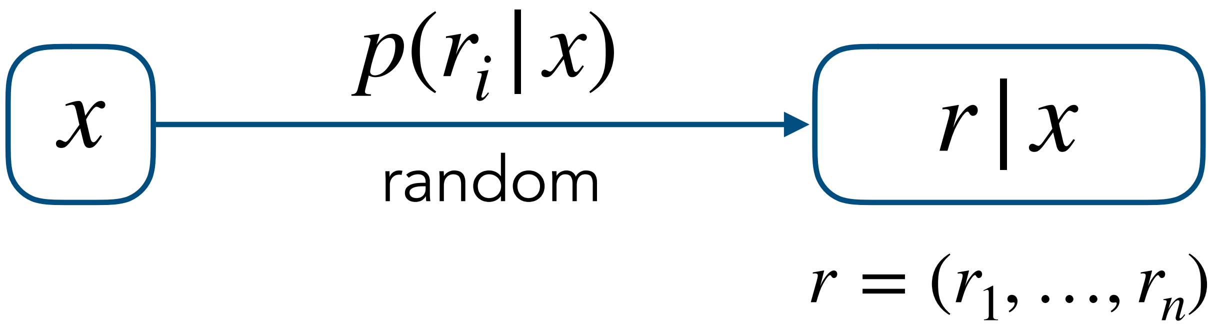
Varying noise



- What determines the shapes of these curves?

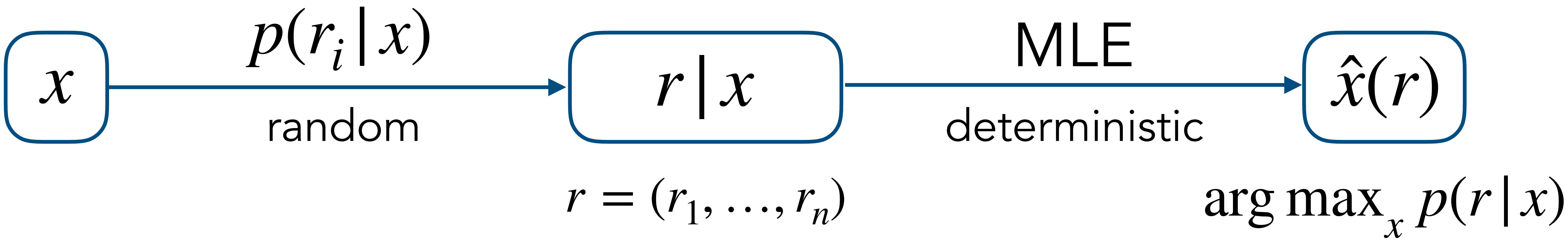
Inference as a Constraint

- We now present an approach in which assumptions are made on how the brain computes the estimates $\hat{x}$.



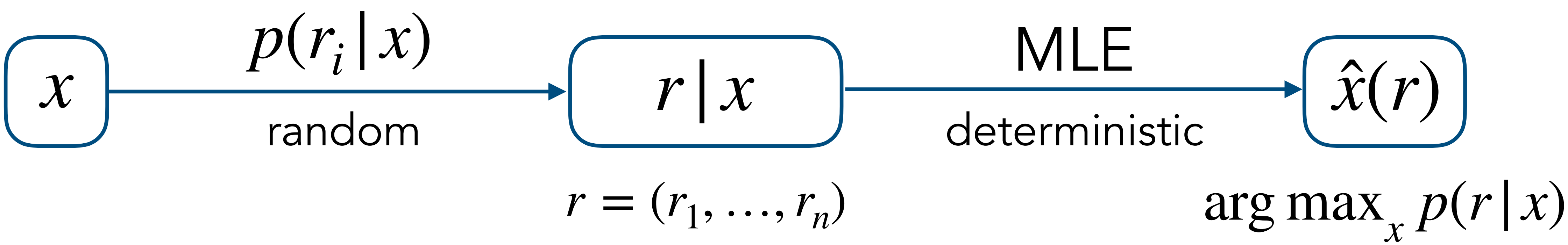
Inference as a Constraint

- We now present an approach in which assumptions are made on how the brain computes the estimates $\hat{x}$.



Inference as a Constraint

- We now present an approach in which assumptions are made on how the brain computes the estimates $\hat{x}$.



- What constraint does that impose on $p(\hat{x} \mid x)$?

Properties of the MLE

- The MLE is, up to the order $\frac{1}{\sqrt{n}}$, *unbiased* and *efficient*.

- Approximately:

$$\hat{x}^{MLE} \,|\, x \sim N\left(x, \frac{1}{nI(x)}\right).$$

where $I(x)$ is the Fisher information of $p(r_i \,|\, x)$.

- This corresponds exactly to our "varying-noise" model

$$\hat{x} \,|\, x \sim N\left(x, s^2(x)\right).$$

Properties of the MLE

- The MLE has a bias of order $\frac{1}{n}$.

- For a Gaussian likelihood, we have, approximately:

$$\hat{x}^{MLE} \,|\, x \sim N\left(x + \frac{1}{4}\frac{d}{dx}\left(\frac{1}{nI(x)}\right), \frac{1}{nI(x)}\right).$$

- This looks like our best-fitting model

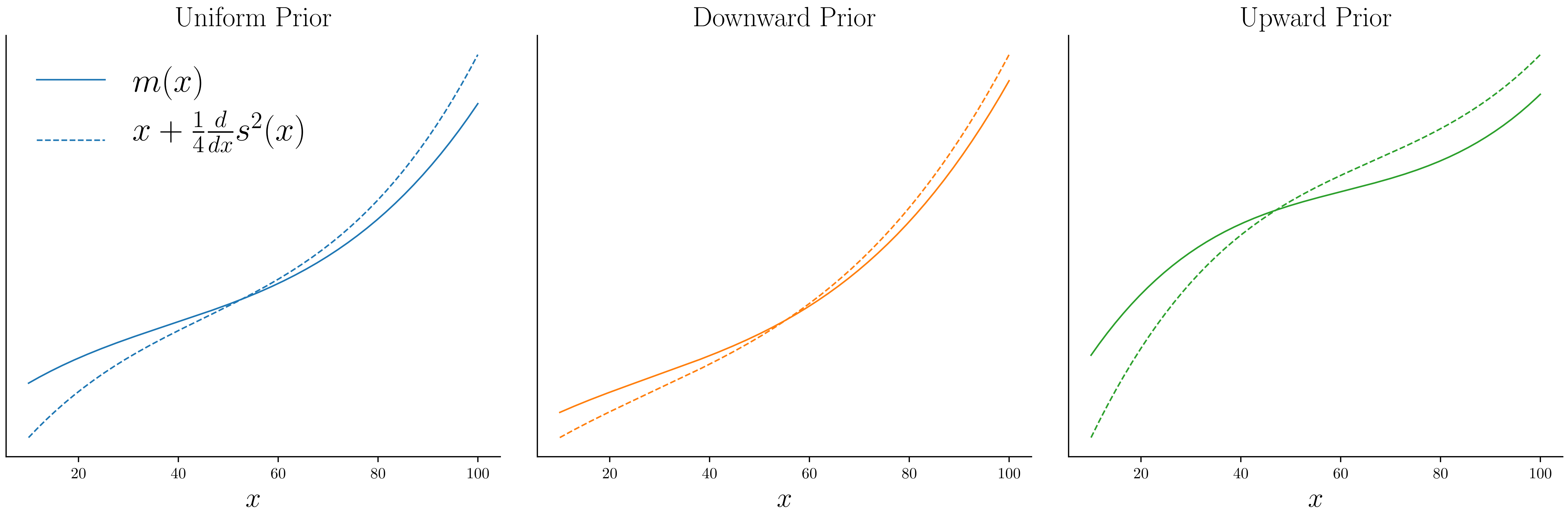
$$\hat{x} \,|\, x \sim N\left(m(x), s^2(x)\right),$$

but our two functions, $m(x)$ and $s(x)$, are now *constrained* by a *single* function, $I(x)$.

Properties of the MLE

- This predicts

$$m(x) = x + \frac{1}{4}\frac{d}{dx}\Big(s^2(x)\Big).$$



MLE-constrained model

$$\hat{x} \mid x \sim N\left(x + \frac{1}{4}\frac{d}{dx}\left(\frac{1}{nI(x)}\right), \frac{1}{nI(x)}\right)$$

MLE-constrained model

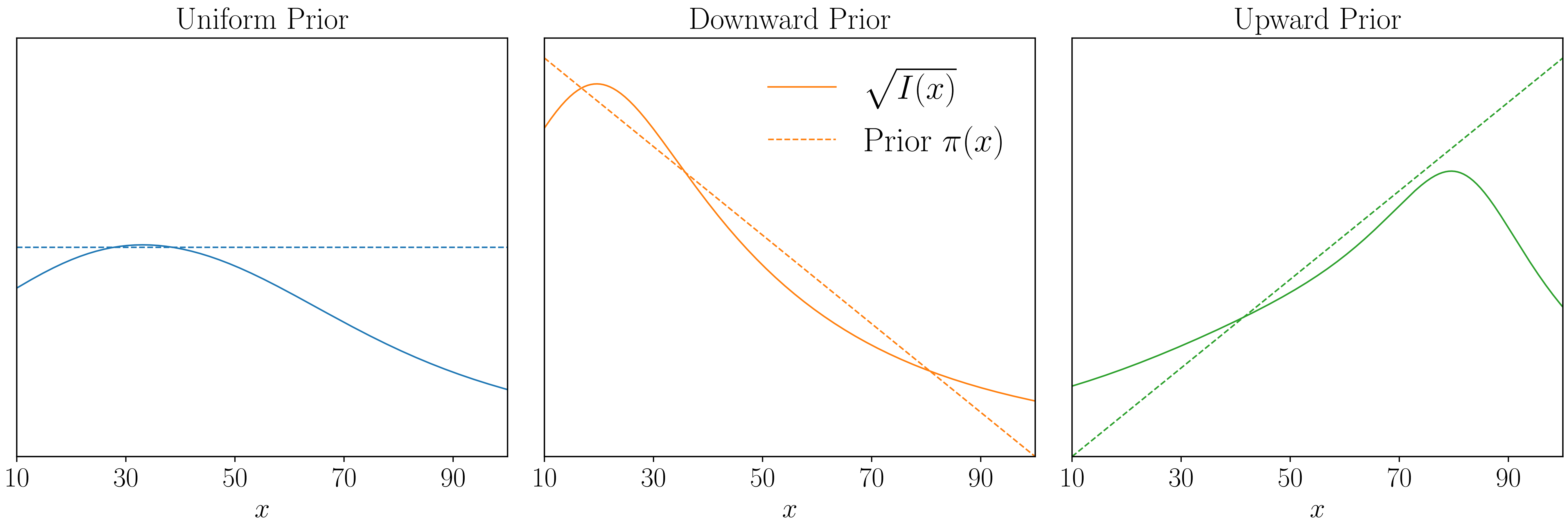
$$\hat{x} \,|\, x \sim N\left(x + \frac{1}{4}\frac{d}{dx}\left(\frac{1}{nI(x)}\right), \frac{1}{nI(x)}\right)$$

ΔBIC	Same parameters	Prior-specific parameters
$\hat{x} \,\vert x \sim N(x, s^2)$	304	254
$N(m(x), s^2)$	182	117
$N(x, s^2(x))$	304	245
$N(m(x), s^2(x))$	108	42
$I(x)$-based	76	**0**

ΔBIC with best model

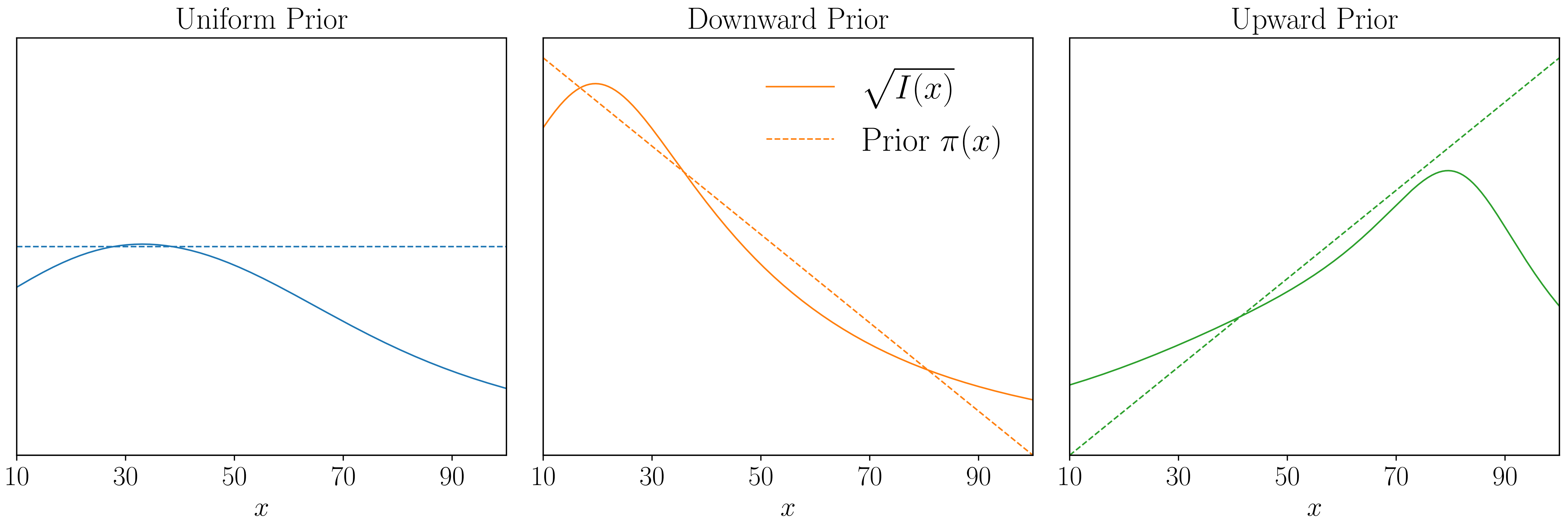
MLE-constrained model

- The fitted Fisher information in comparison with the prior:



MLE-constrained model

- The fitted Fisher information in comparison with the prior:



- This suggests an *efficient coding* of the numbers.

Summary

- In our average-comparison task, subjects seem to unequally weight different numbers in their decisions.
- We introduce a MLE-based model, in which

  (i) an *encoding* of the number, characterized by $I(x)$,

  (ii) is followed by a *maximum-likelihood estimation* of the number, based on the encoded evidence.
- This model makes a specific prediction relating the bias and the variance of the estimates,
- And it best accounts for the behavioral data.
- Lastly, the encoding Fisher information $I(x)$ seems *efficiently* adapted to the prior.

Thank you!