Simulating Language
4: Iterated Learning

Simon Kirby
simon.kirby@ed.ac.uk

Lab 3, Question 1

"Can you produce a result like the Hudson Kam & Newport (2005) results *for adults*, i.e. that adult learners fairly accurately track the frequency of a linguistic variant in their input? What kinds of priors and what kinds of data does this work for?"

- What would "tracking the frequency of a linguistic variant in the input" look like in our model?

- Under what conditions does this occur?

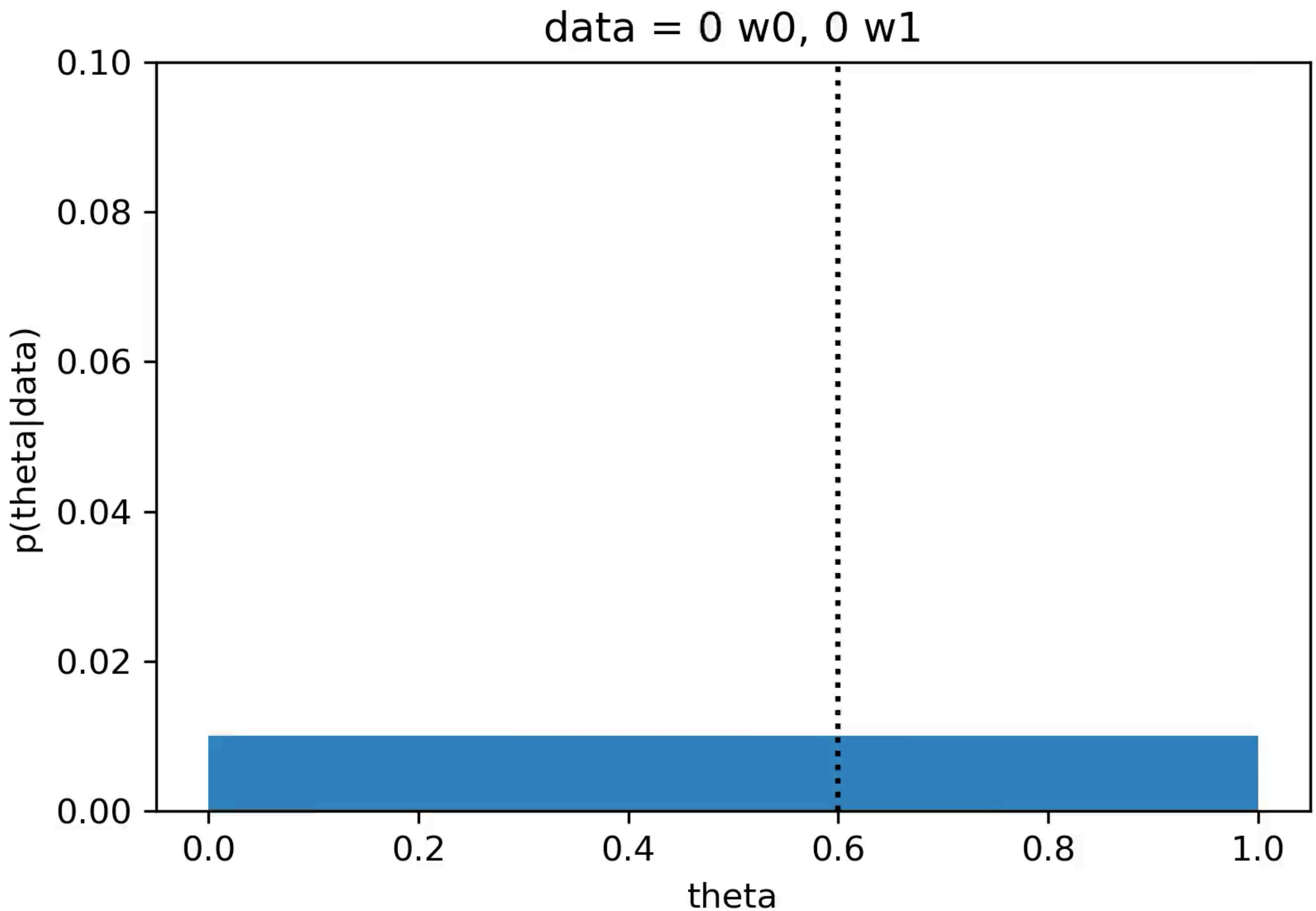
Lab 3, Question 2

"Can you produce a result like the Hudson Kam & Newport (2005) results *for children*, i.e. that children tend to regularise, sometimes producing only one variant even when their data contains variation? Again, what kinds of priors and what kinds of data does this work for?"

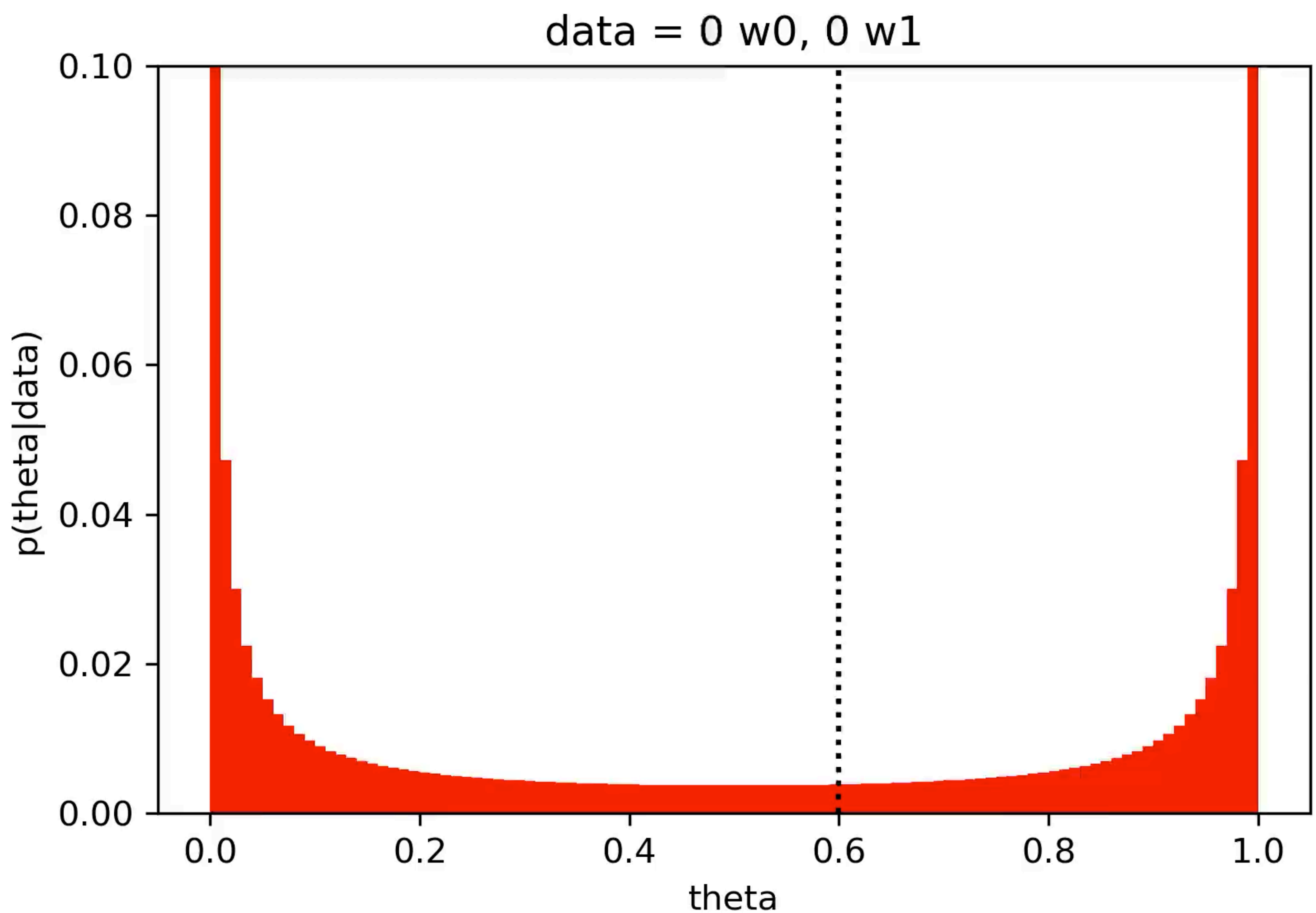
- What would regularisation look like in our model?

- Under what conditions does this occur?

Uniform prior (alpha=1)



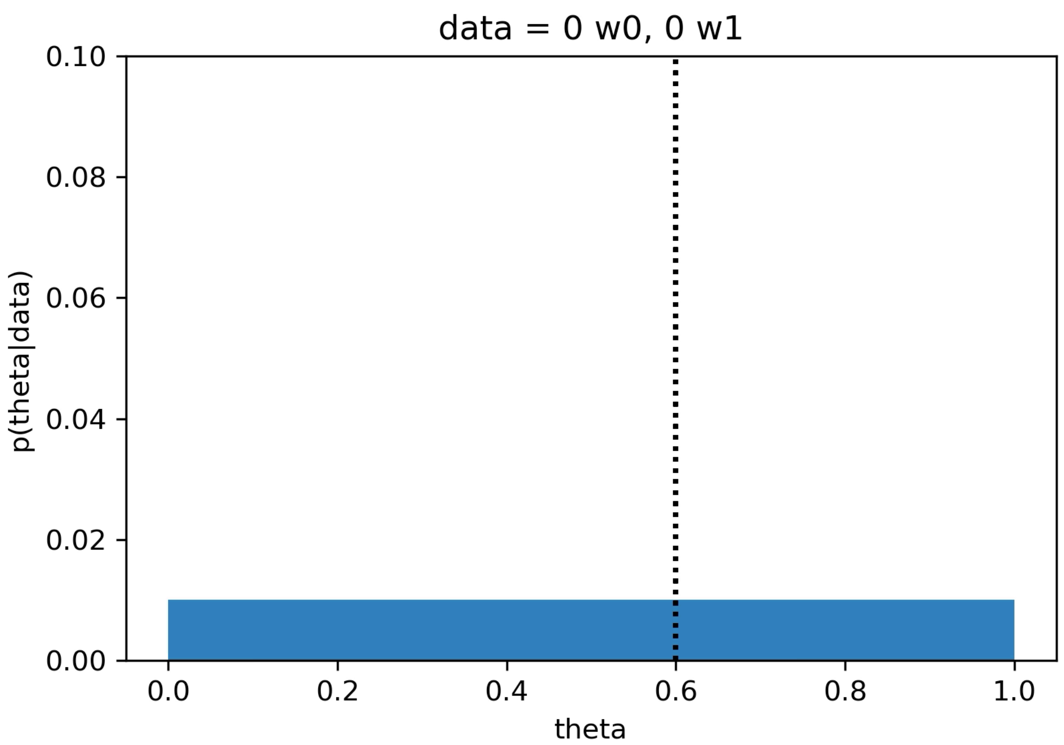
Regularity prior (alpha=0.1)



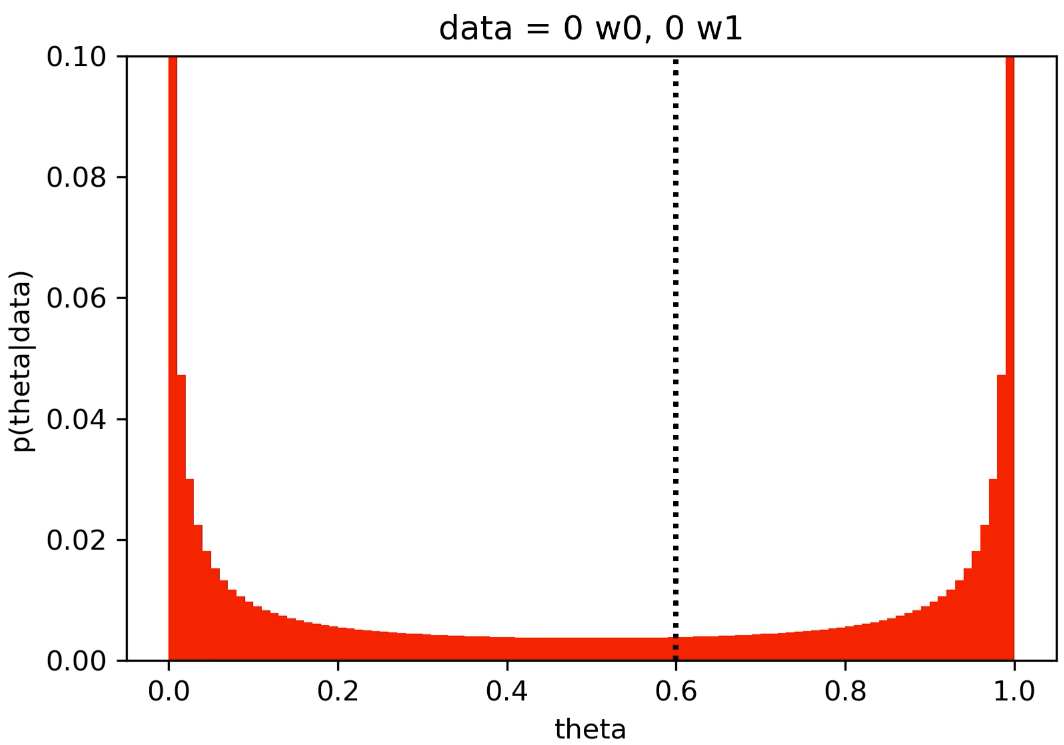
Data obscures the prior

$$P(\theta|d) \propto P(d|\theta)P(\theta)$$

Unbiased learner

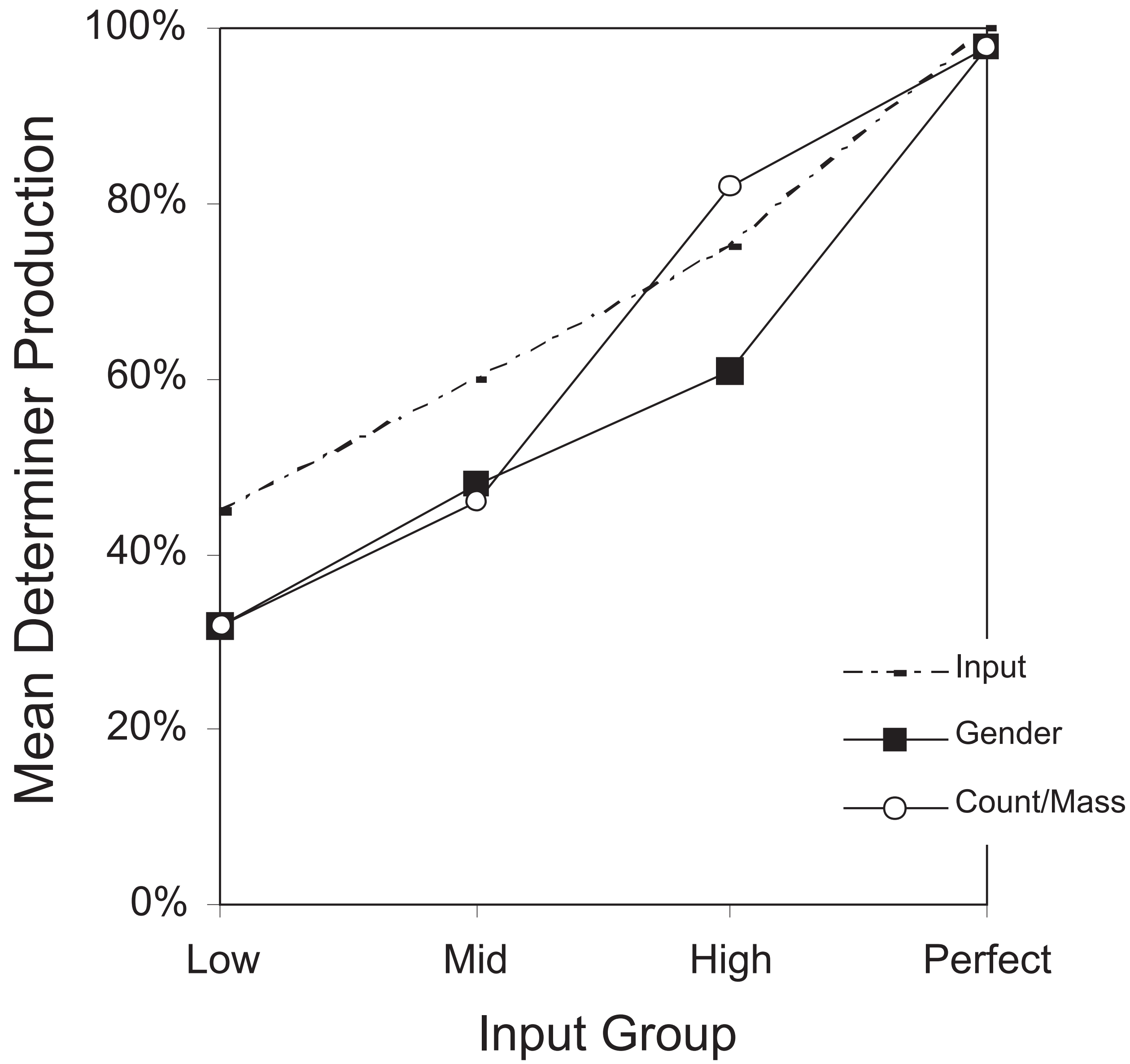


Biased learner



Data obscures the prior $P(\theta|d) \propto P(d|\theta)P(\theta)$

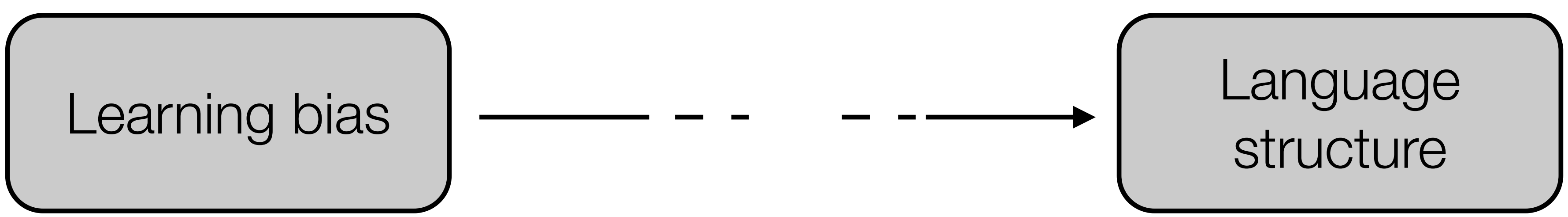
Unbiased learner? Biased learner?



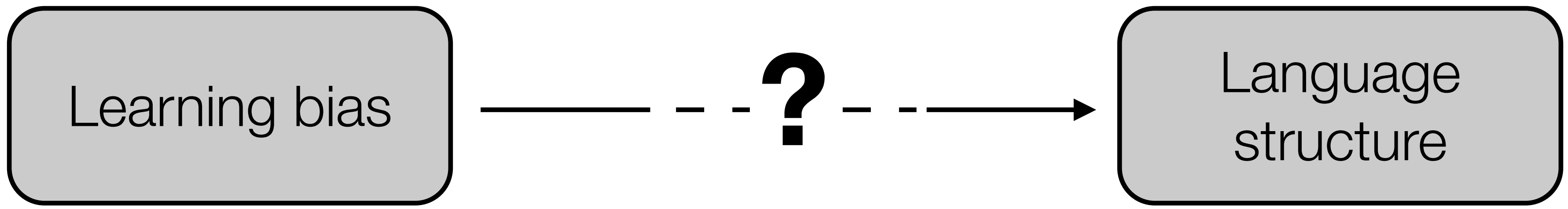
What makes languages regular?

- We're interested in explaining why languages are the way they are (e.g. regular)

- We're arguing it's due to something about our learning bias (e.g. learners prefer regular languages)



The problem of linkage



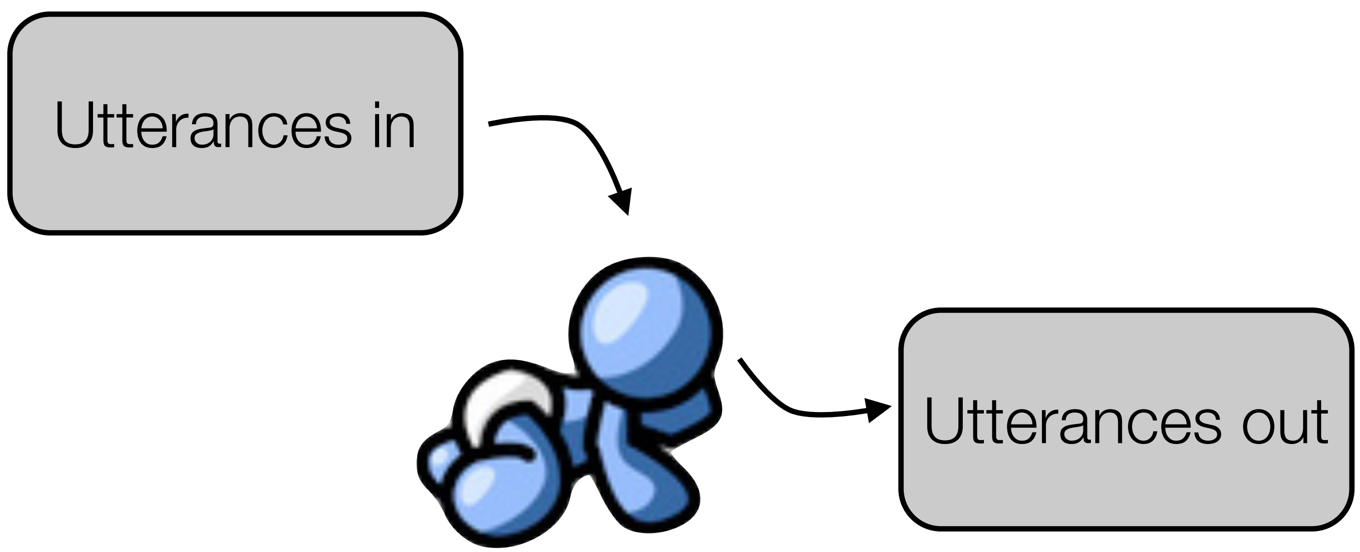
THE PROBLEM OF LINKAGE

- But there's something wrong here. Given enough data, the different learning biases seem to lead to the same outcome.

- Two problems:
  - Where does the data come from in the first place?
  - And **how exactly** does learning bias (a property of an individual's cognition) lead to language structure (a universal property of population behaviour)?

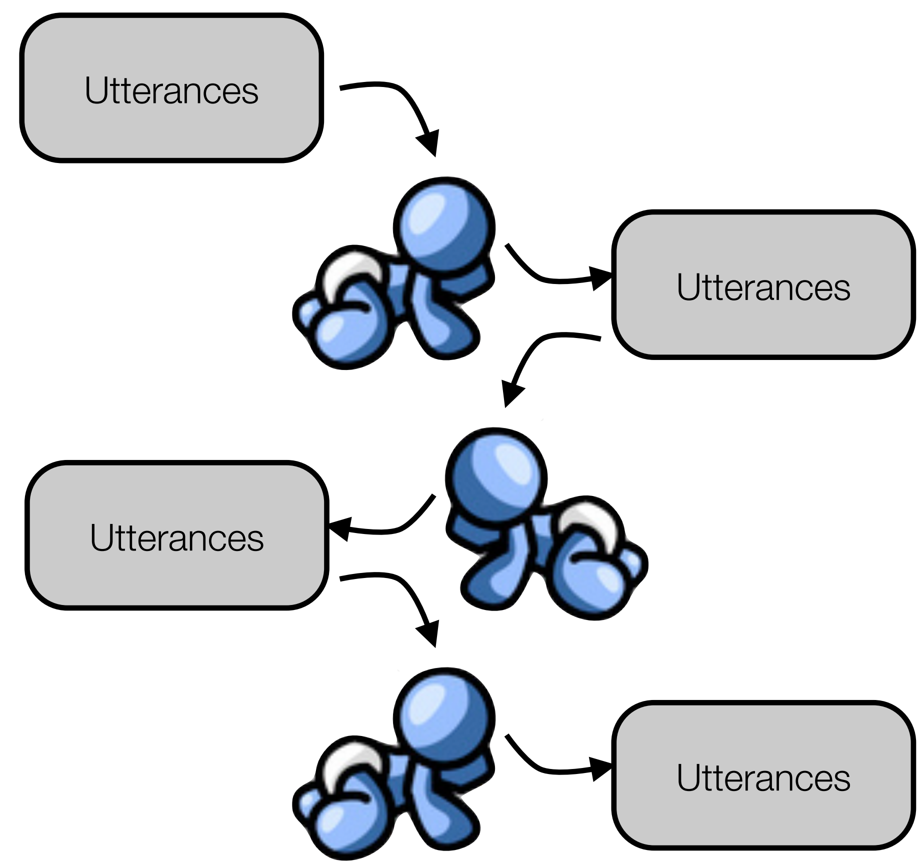
Solving the problem of linkage

- Where does the language data come from that our learners have to acquire?



Solving the problem of linkage

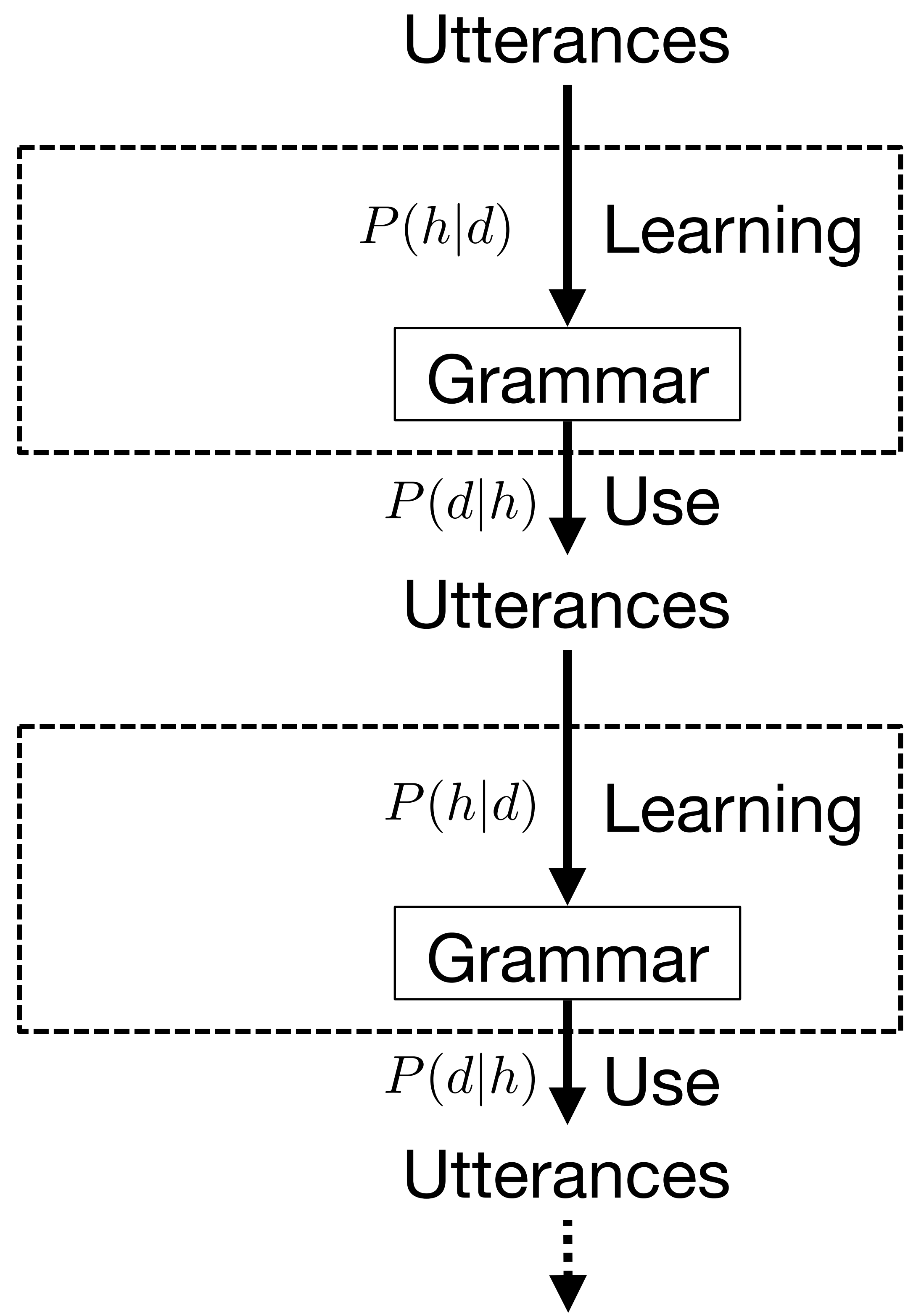
- Where does the language data come from that our learners have to acquire?
- From other learners!
- Language persists over time by repeatedly being learned and used by multiple individuals in a population
- It is out of this continual process of *iterated learning* that the structure of language emerges
- Note, this is *cultural* rather than biological evolution



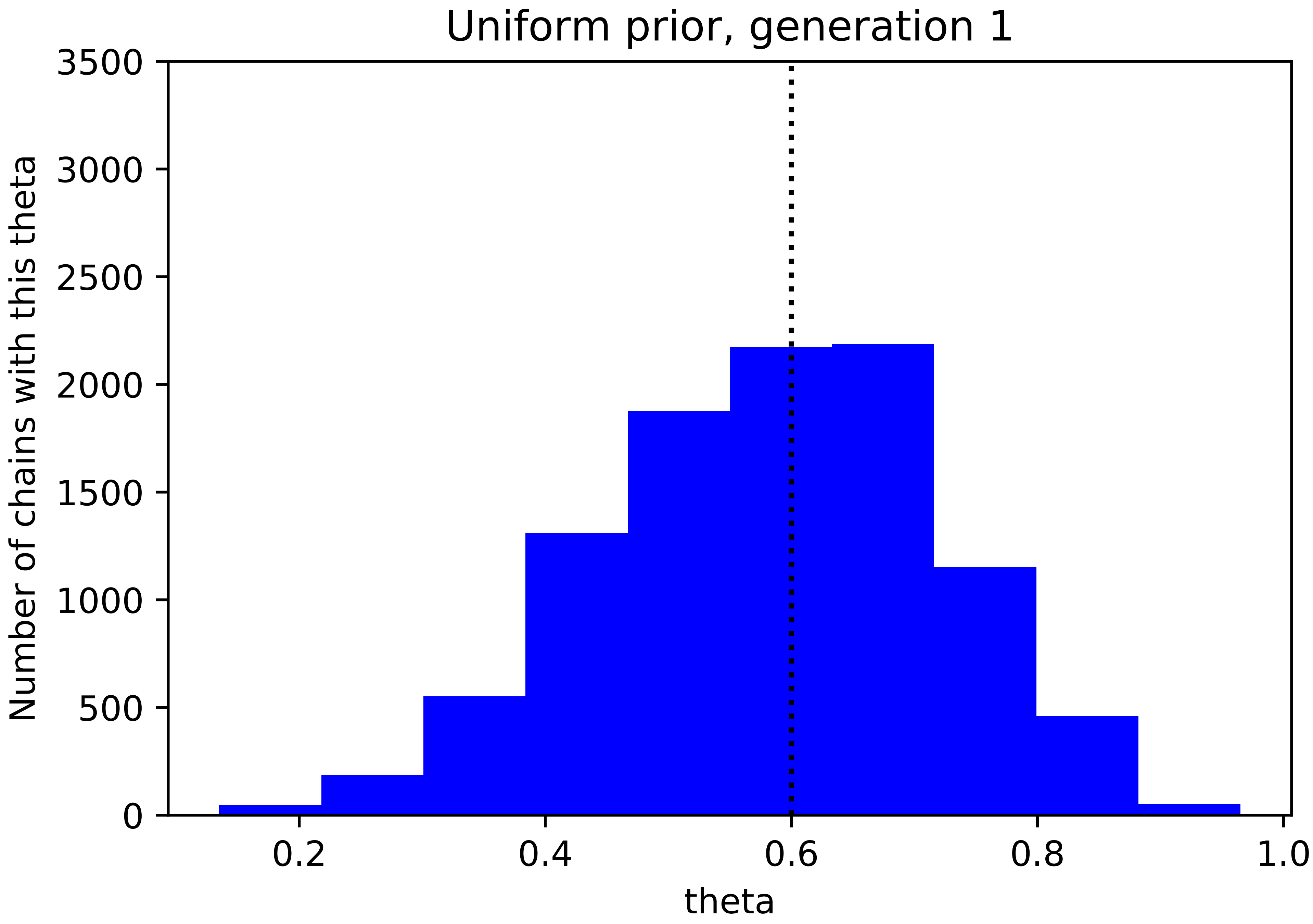
Modelling **iterated learning**

Simulate language transmission from learner to learner.

How does the bias affect the end result of iterated learning?

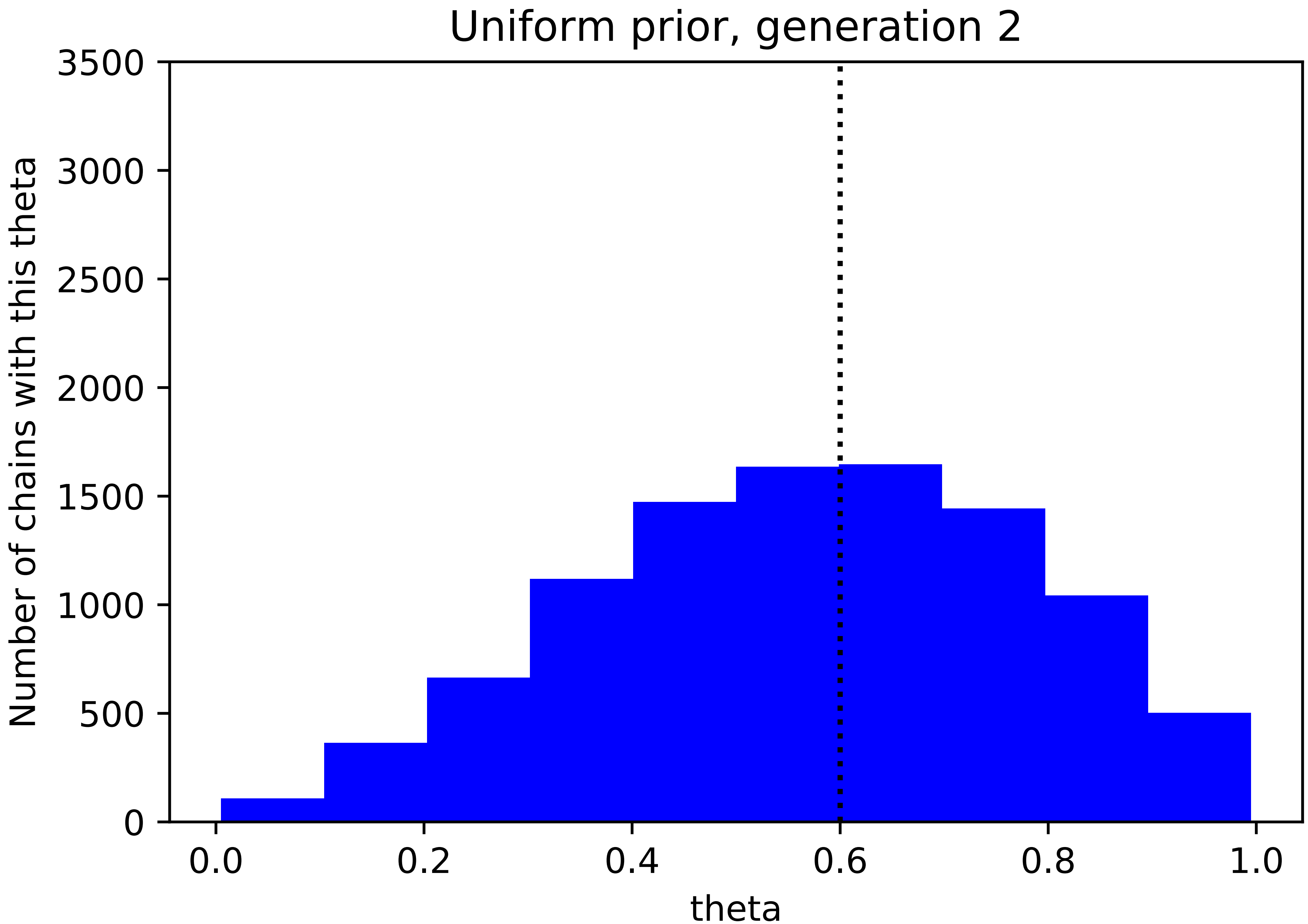


Watching the prior reveal itself

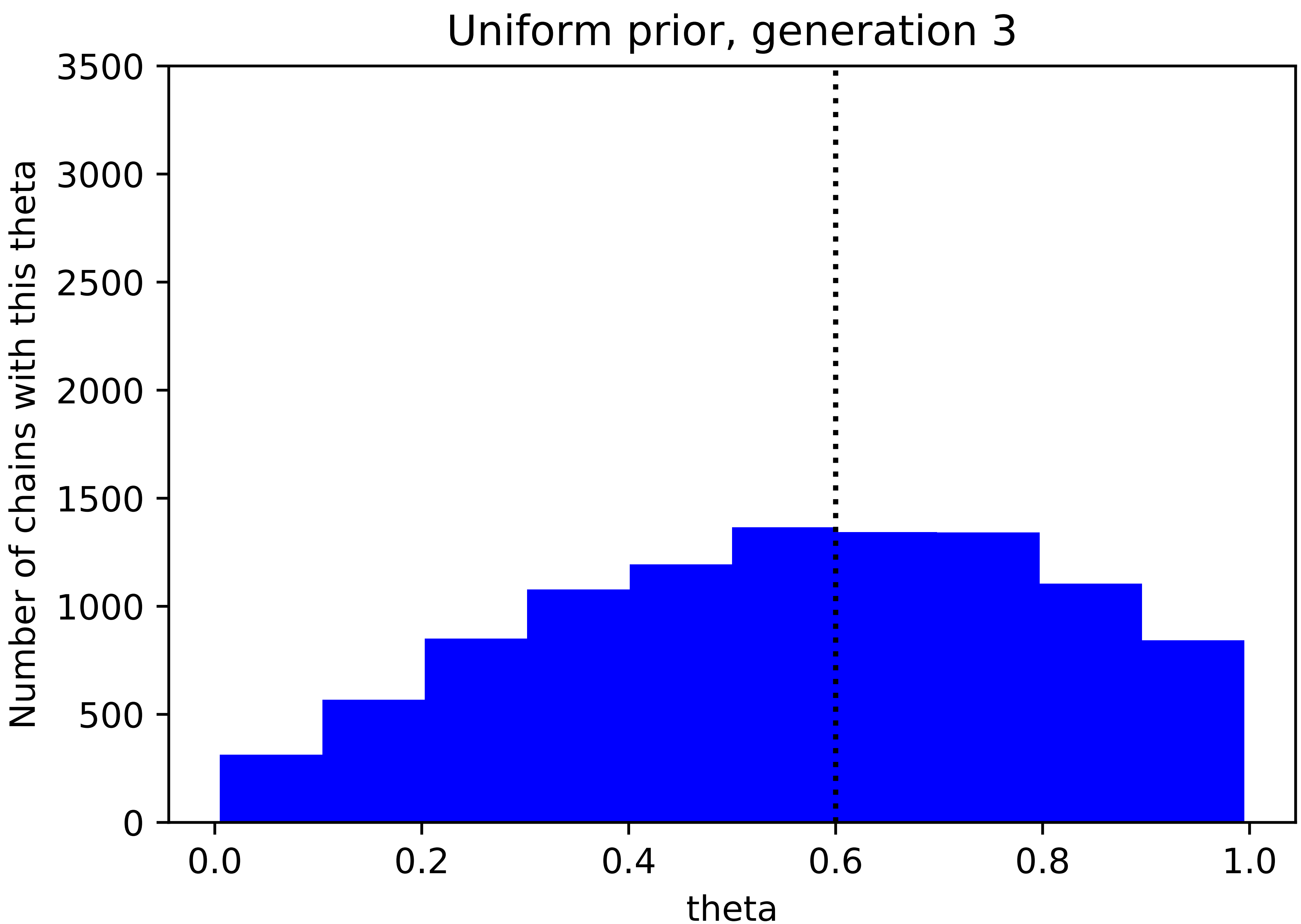


10,000 runs of the simulation

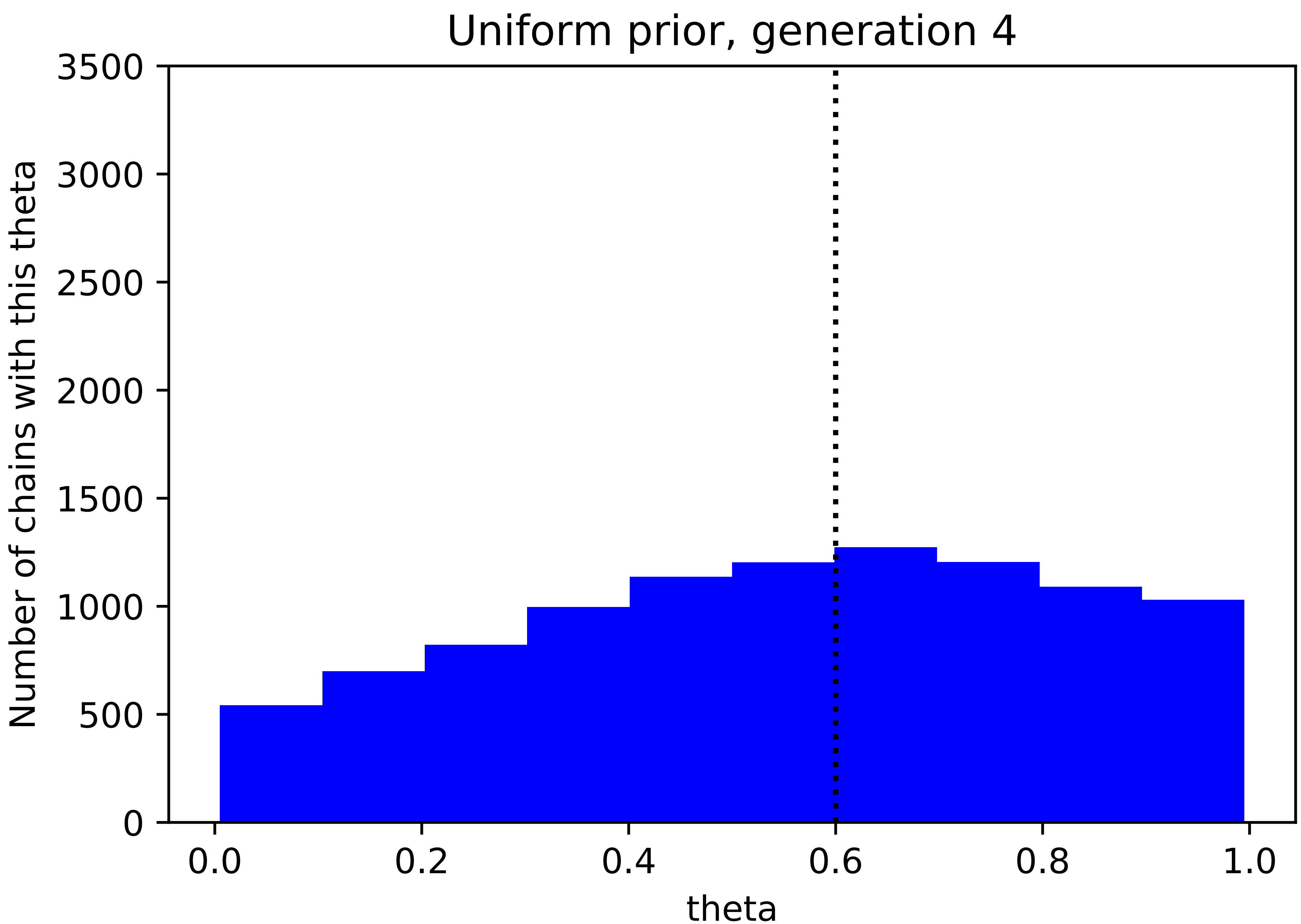
Watching the prior reveal itself



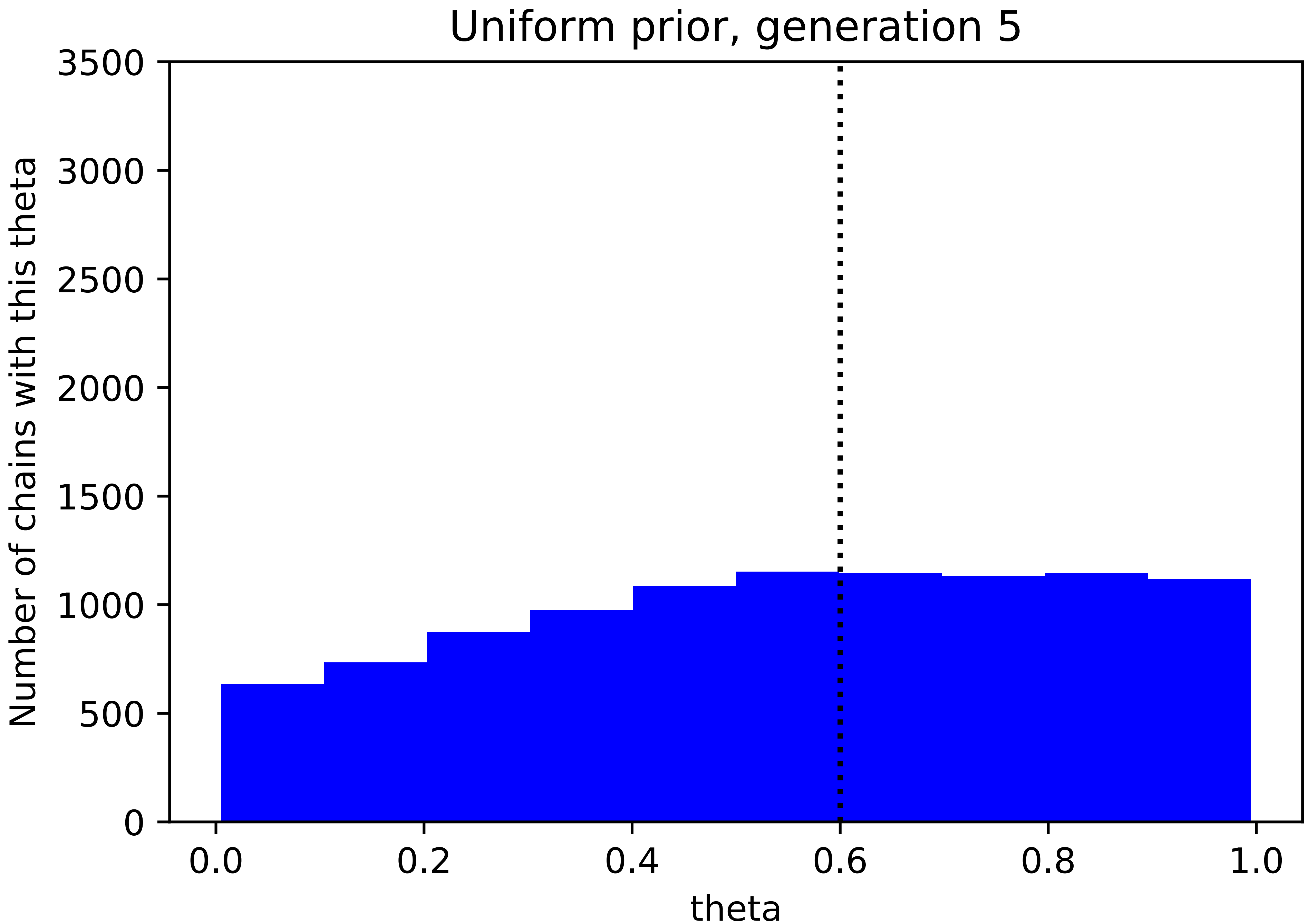
Watching the prior reveal itself



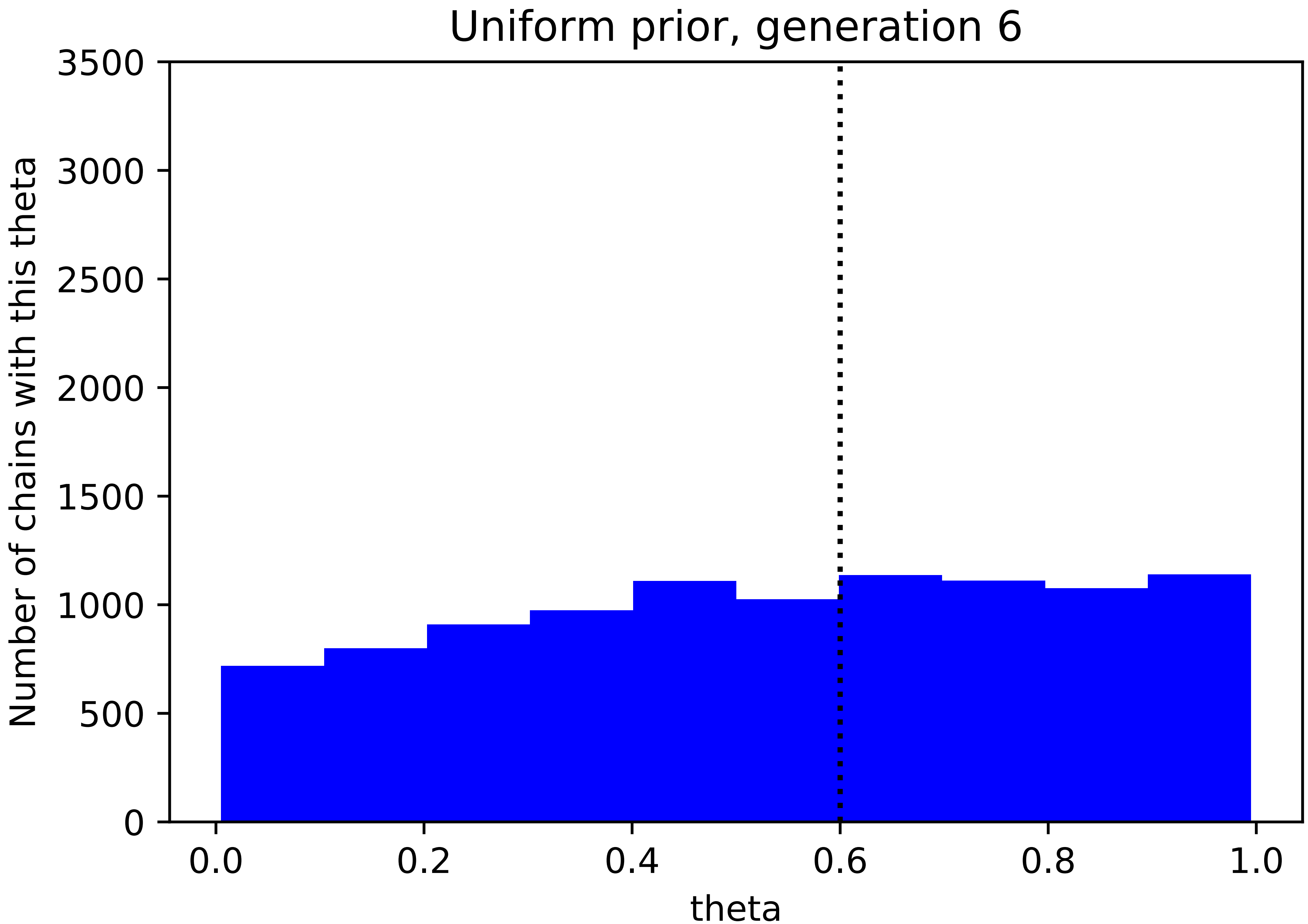
Watching the prior reveal itself



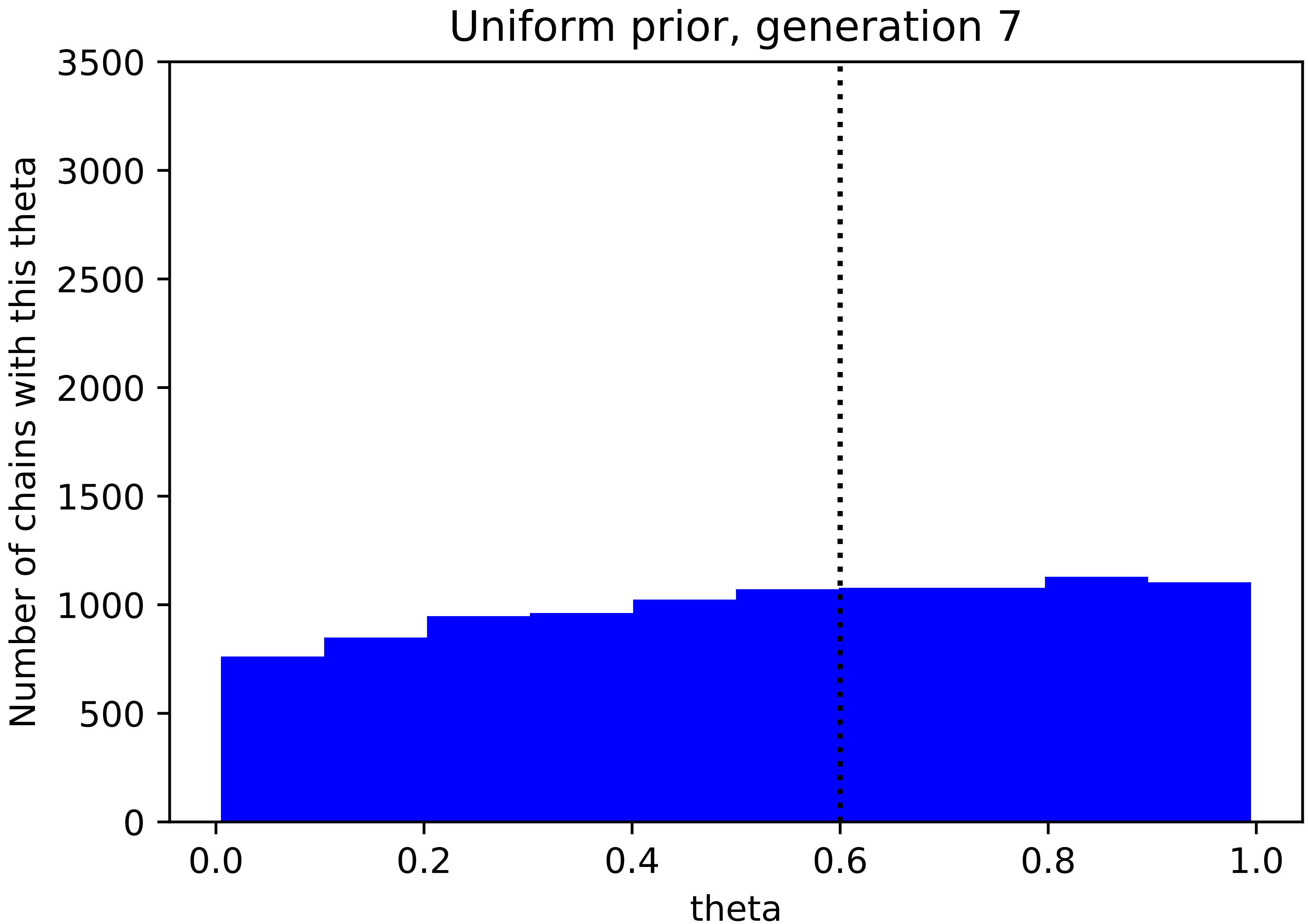
Watching the prior reveal itself



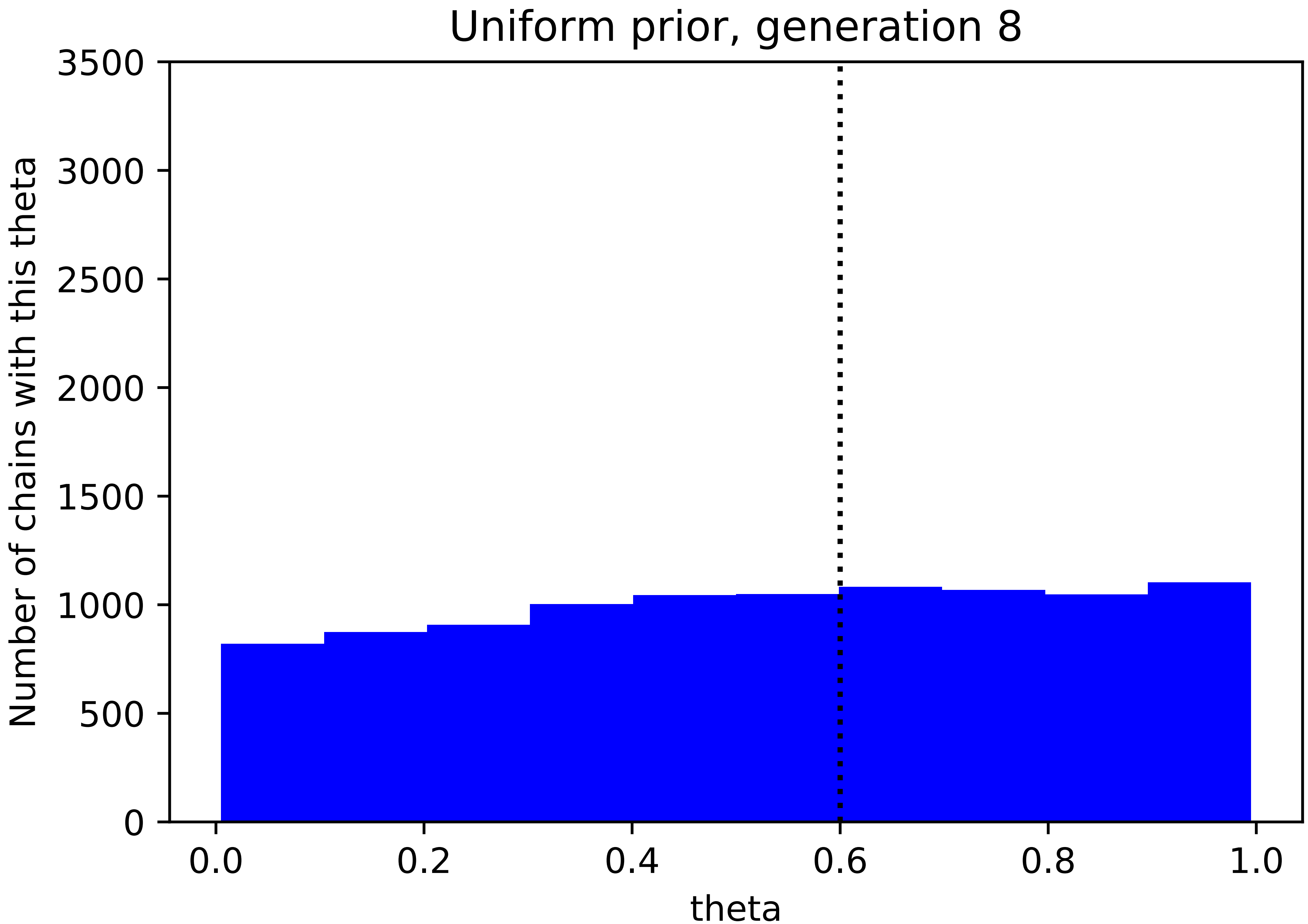
Watching the prior reveal itself



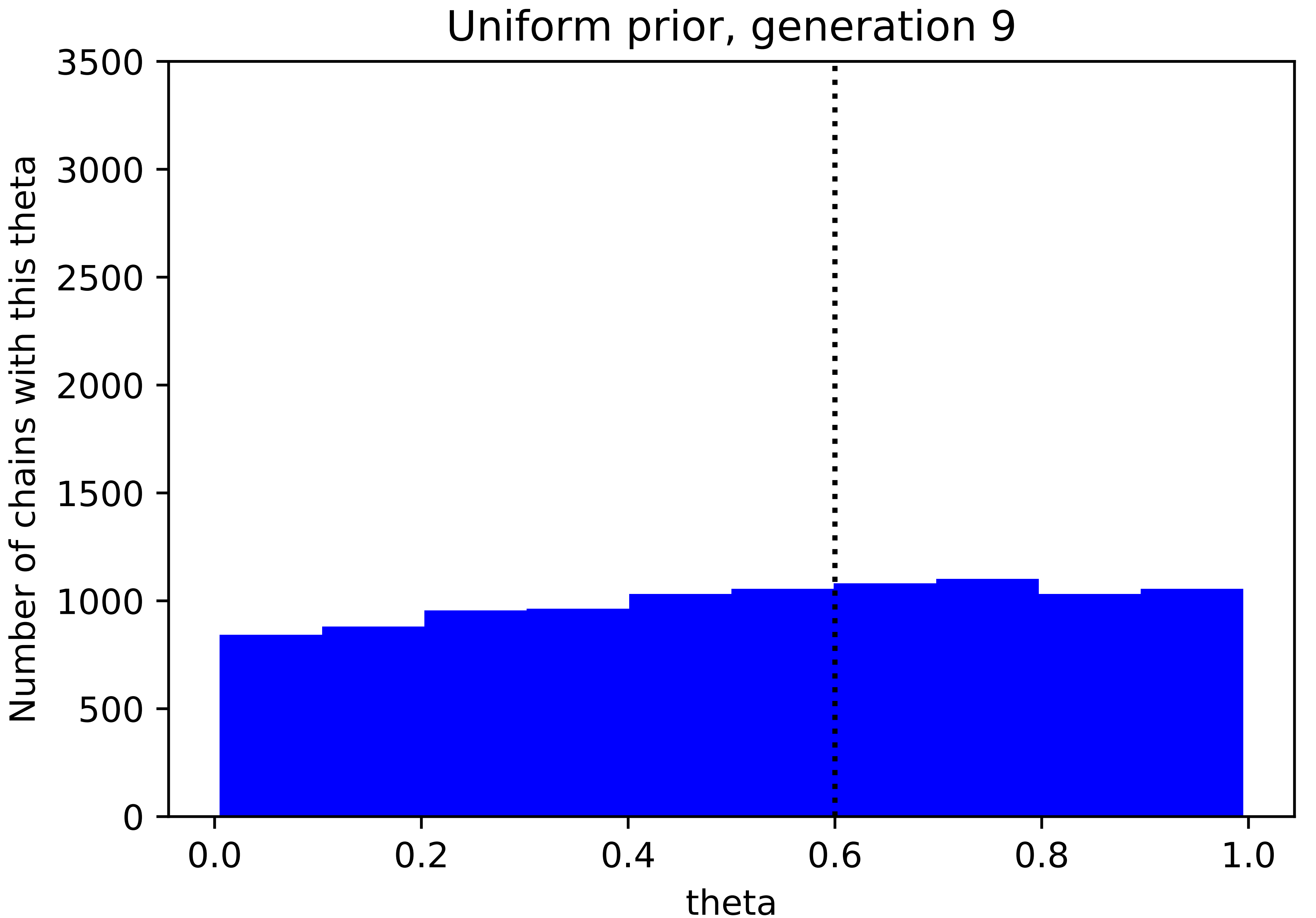
Watching the prior reveal itself



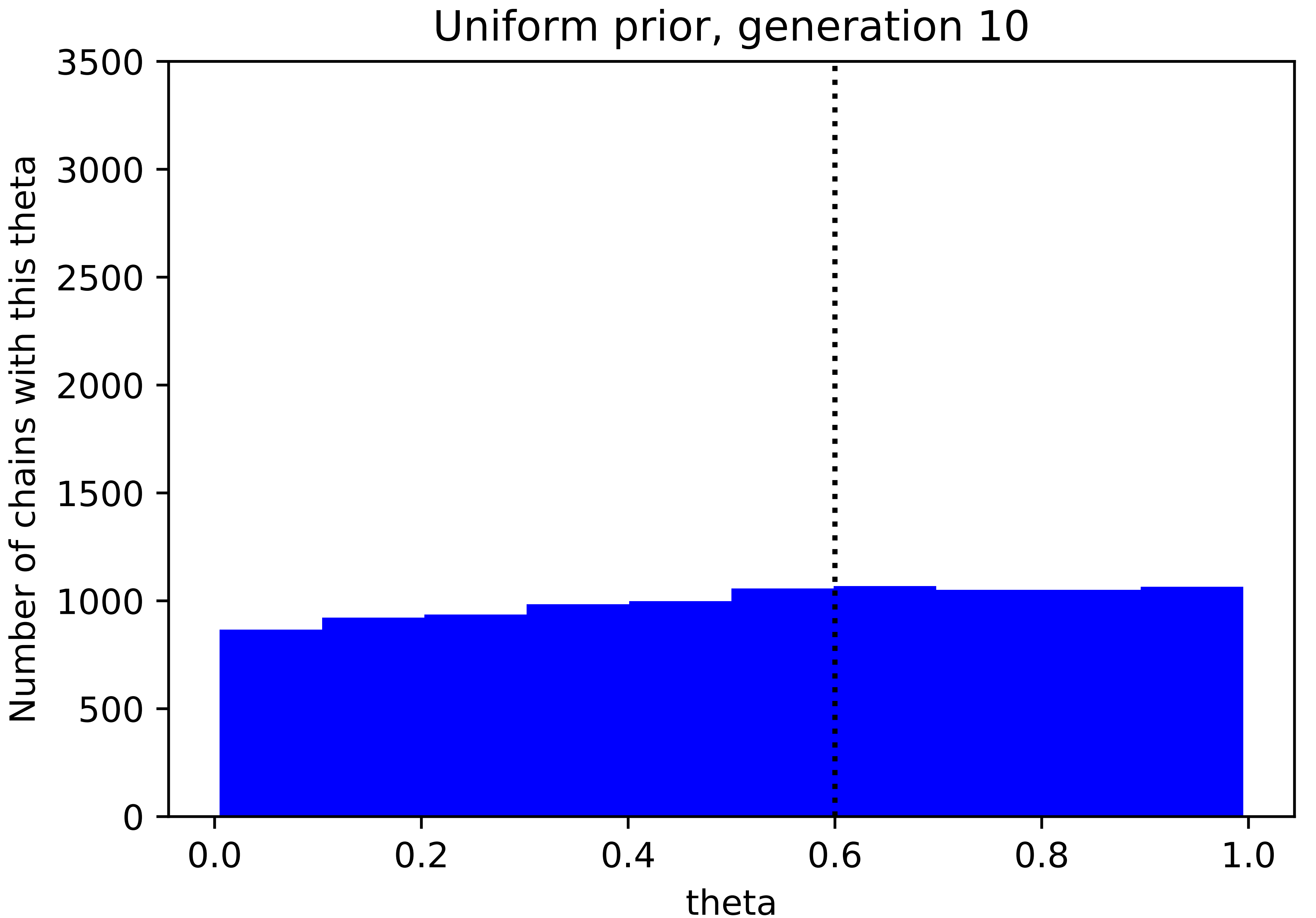
Watching the prior reveal itself



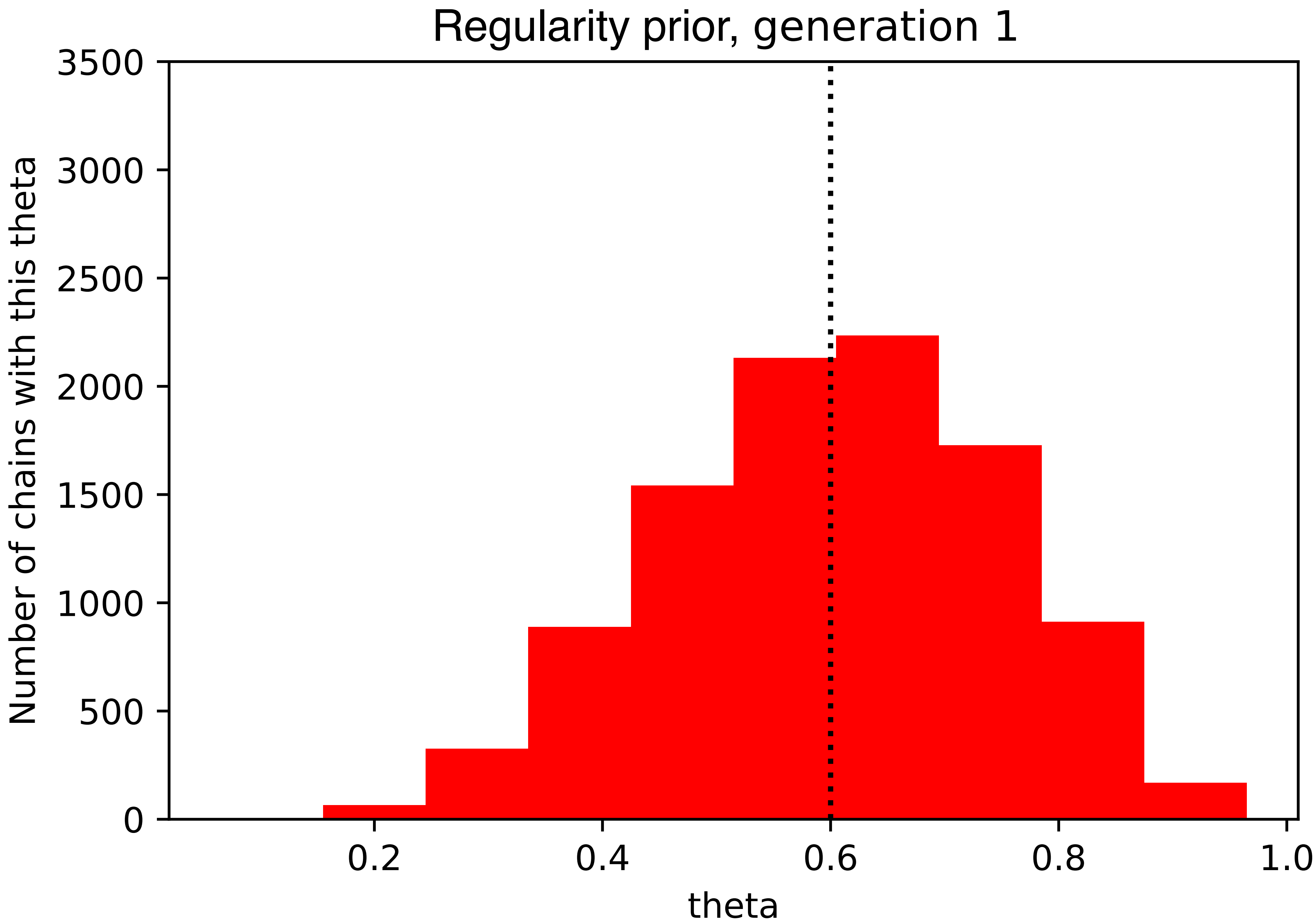
Watching the prior reveal itself



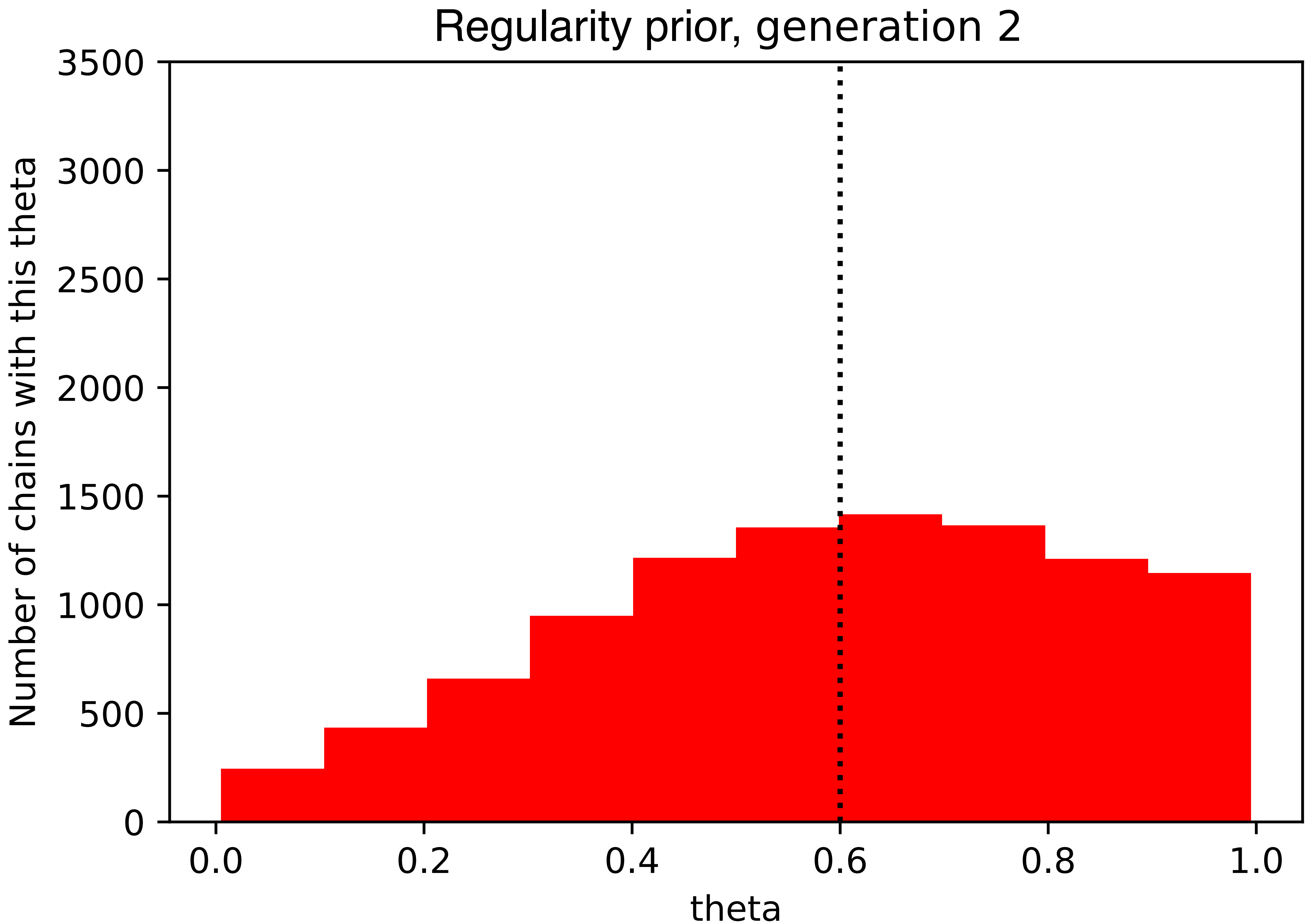
Watching the prior reveal itself



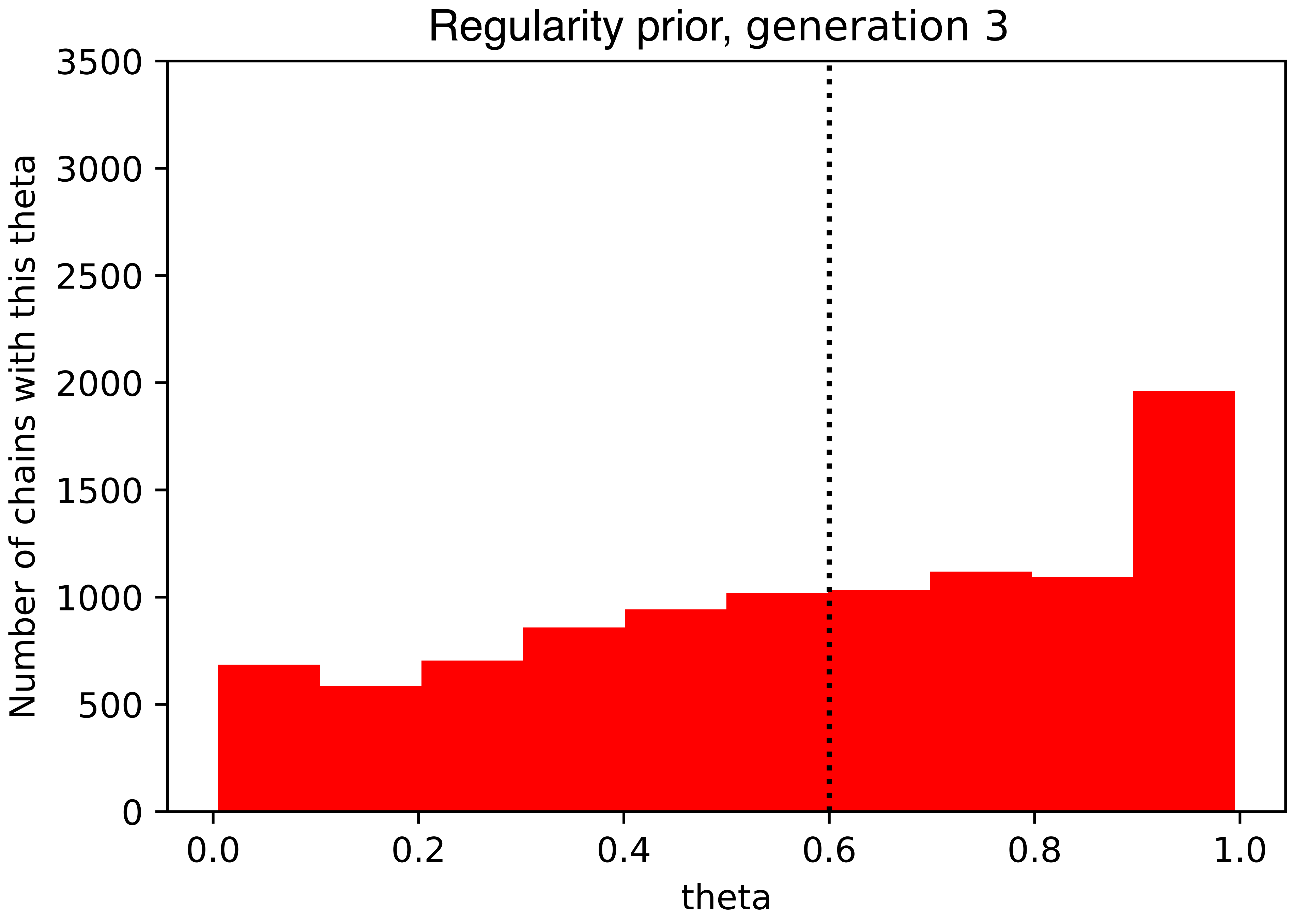
Watching the prior reveal itself



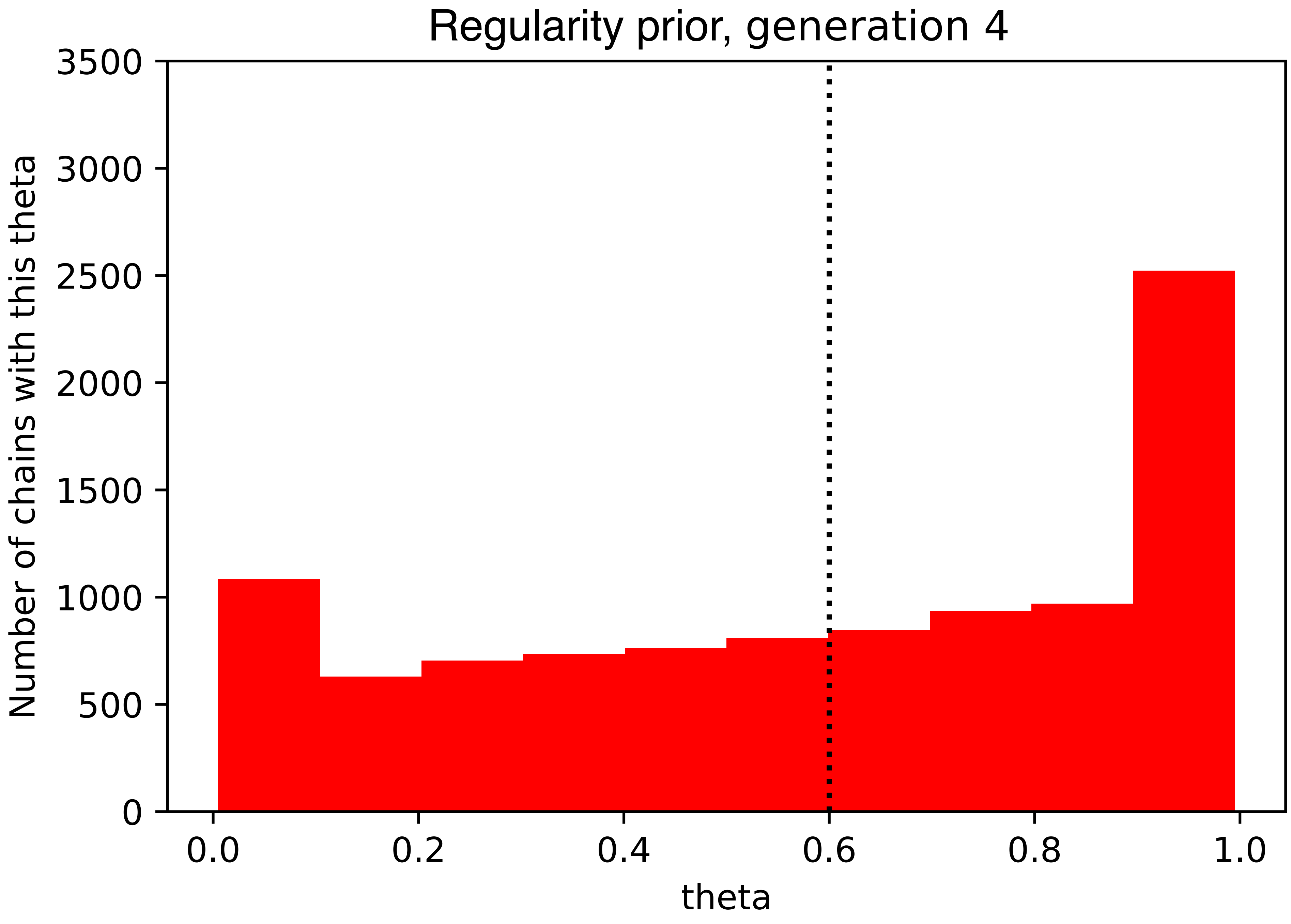
Watching the prior reveal itself



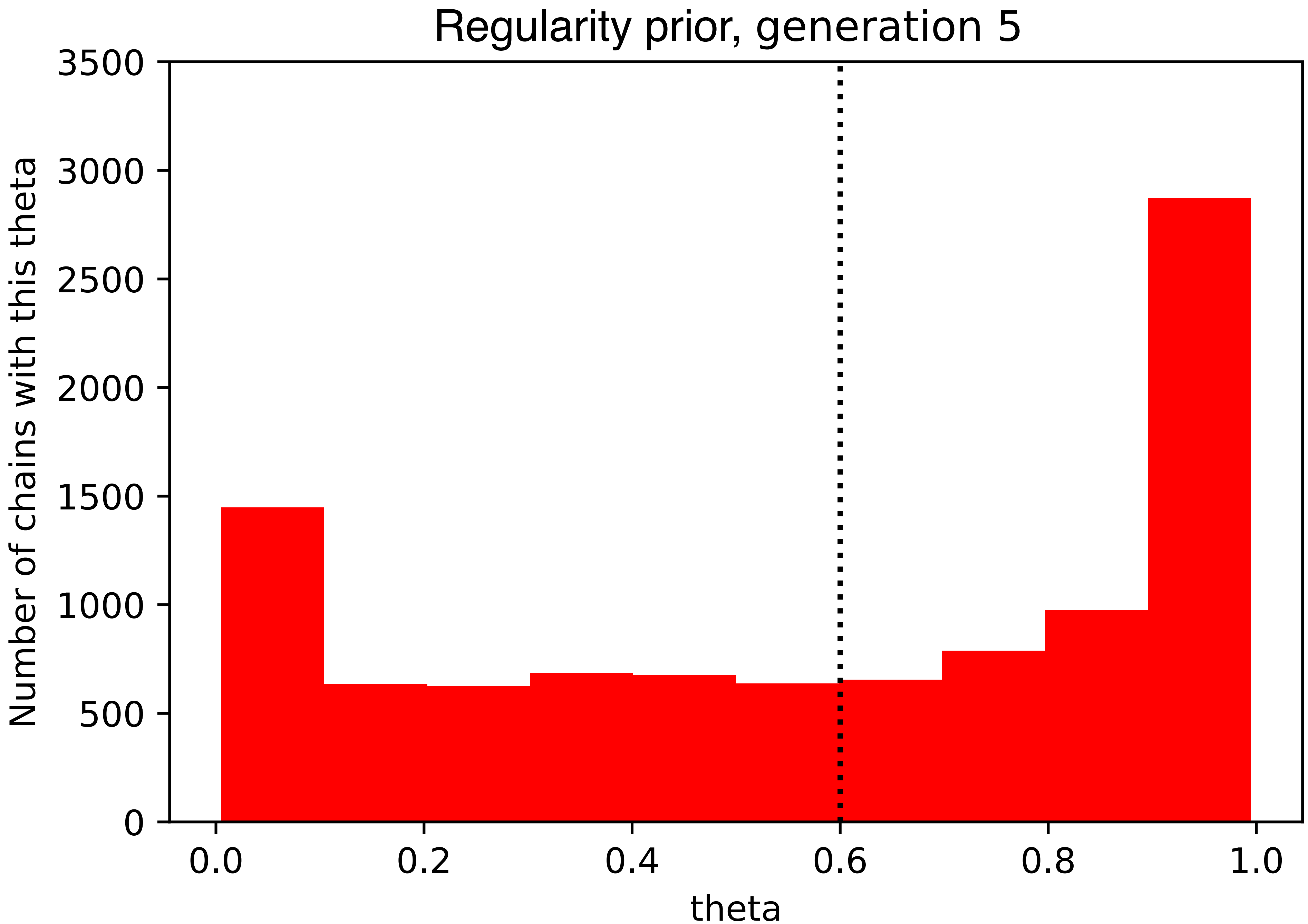
Watching the prior reveal itself



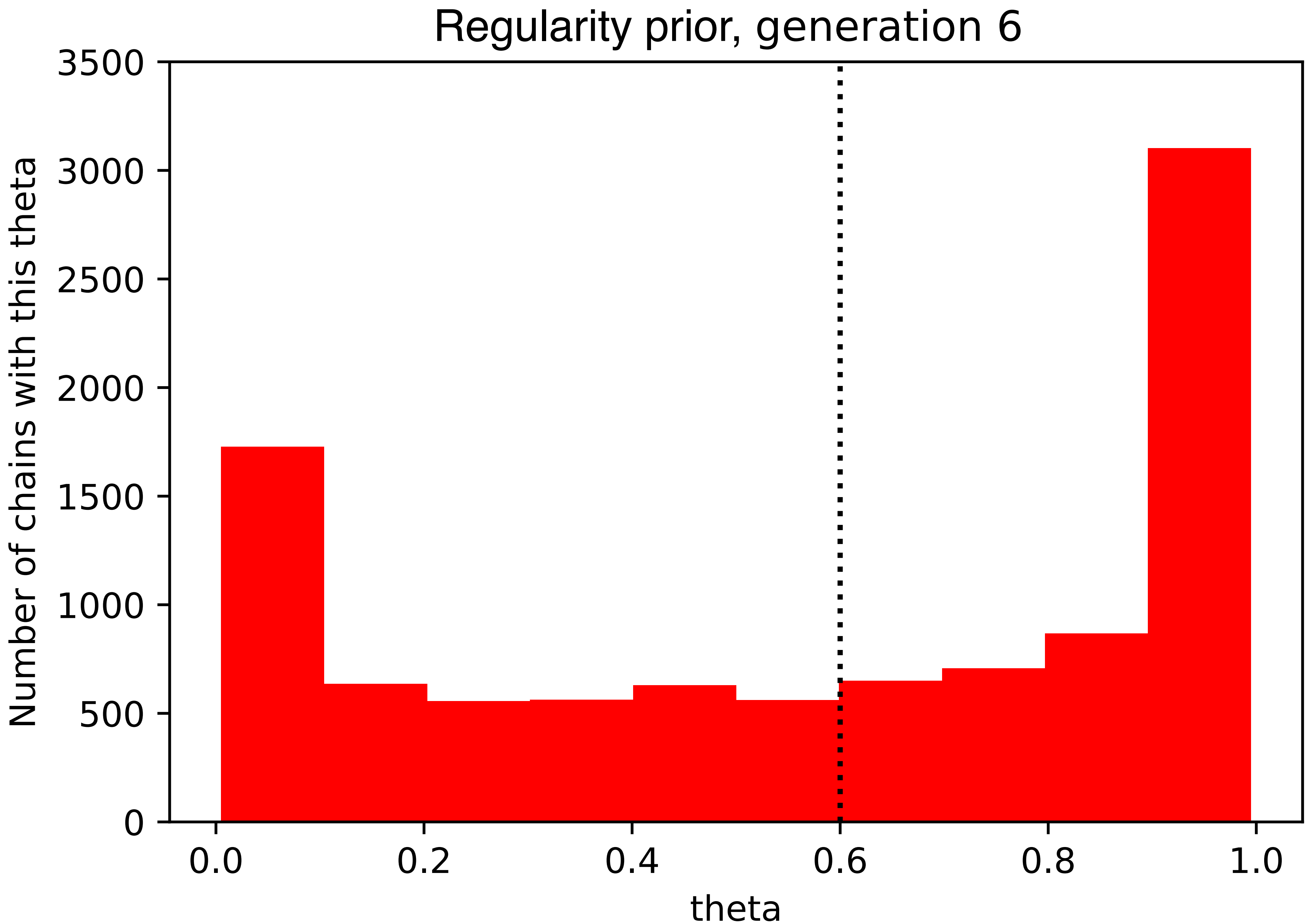
Watching the prior reveal itself



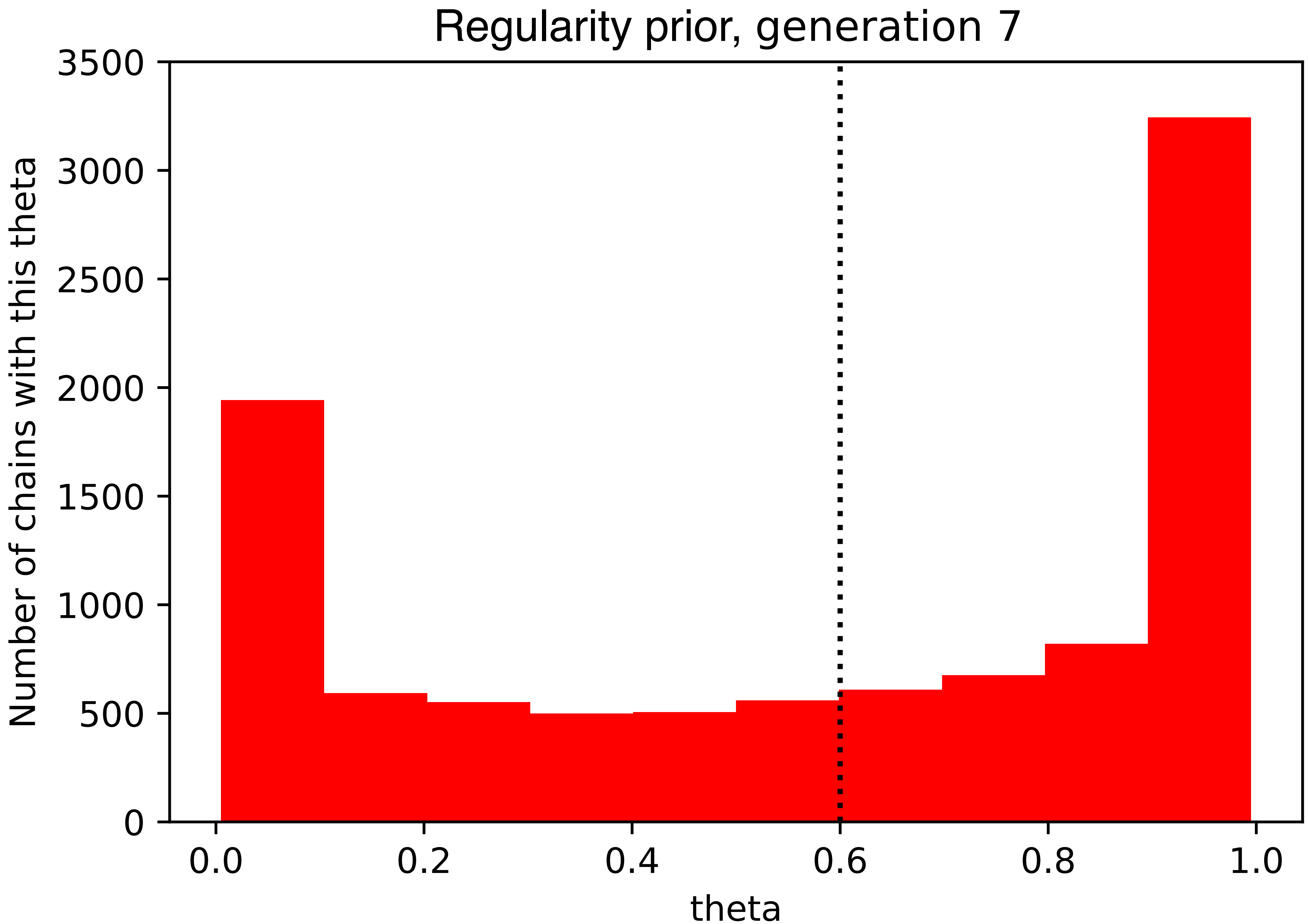
Watching the prior reveal itself



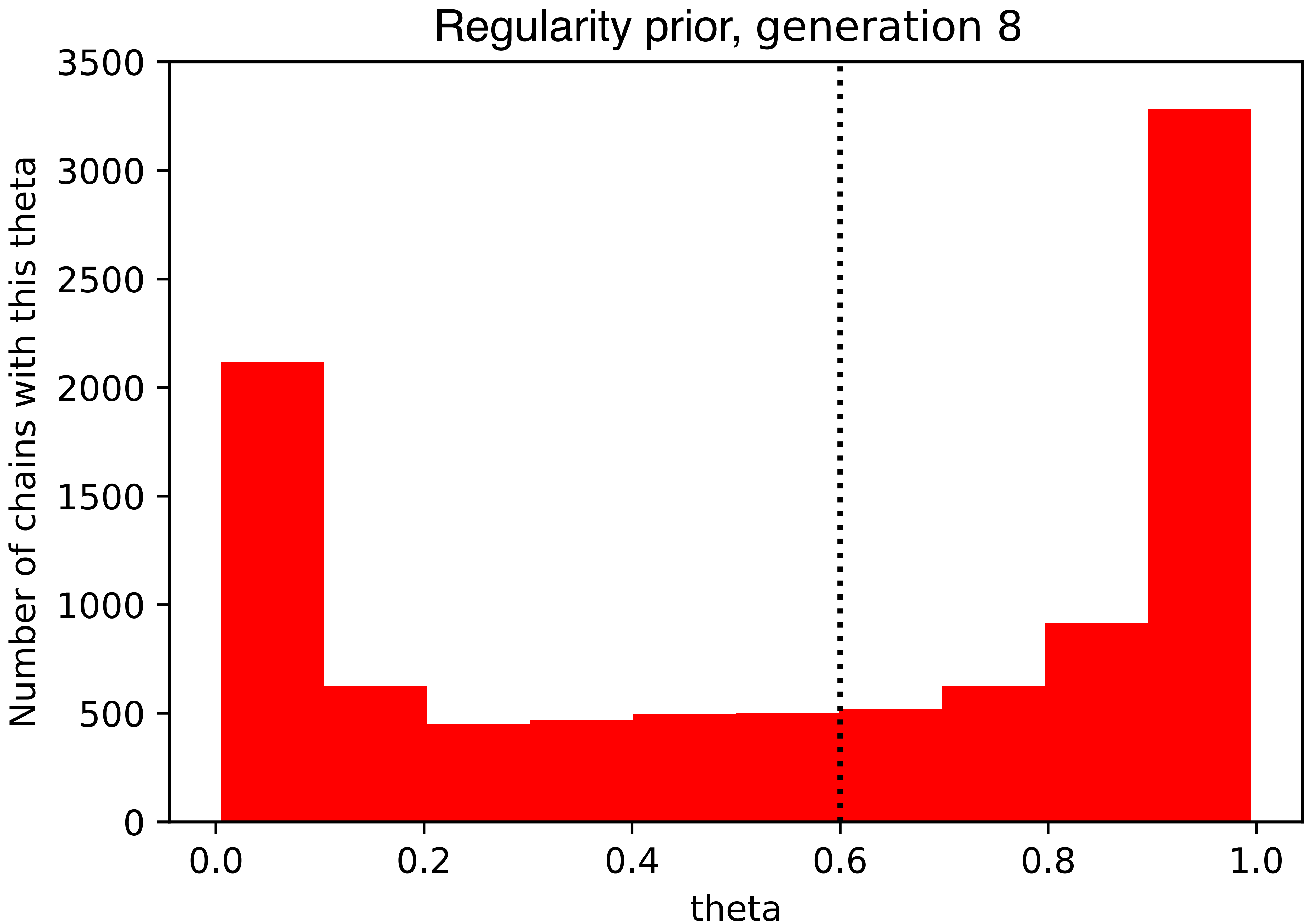
Watching the prior reveal itself



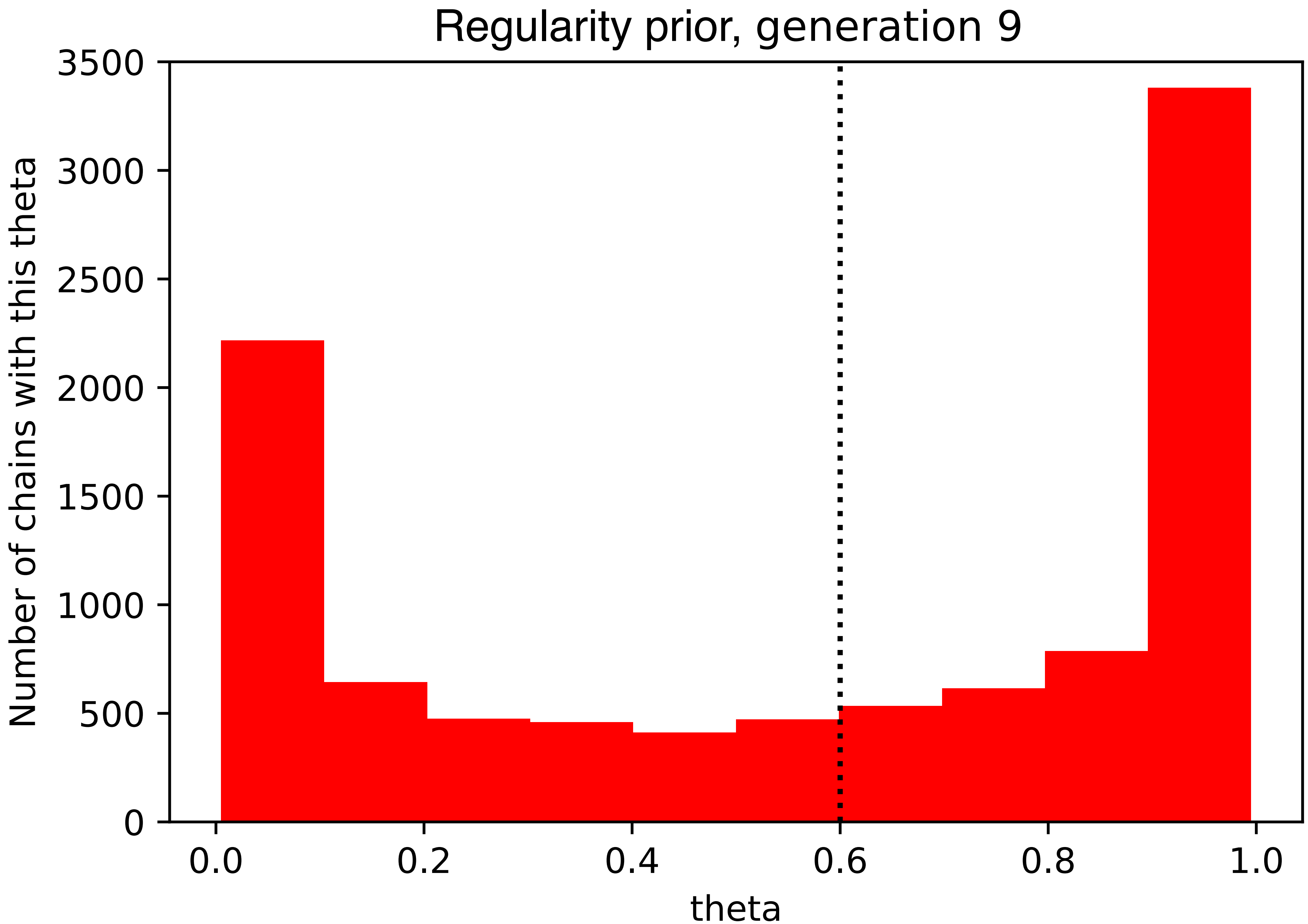
Watching the prior reveal itself



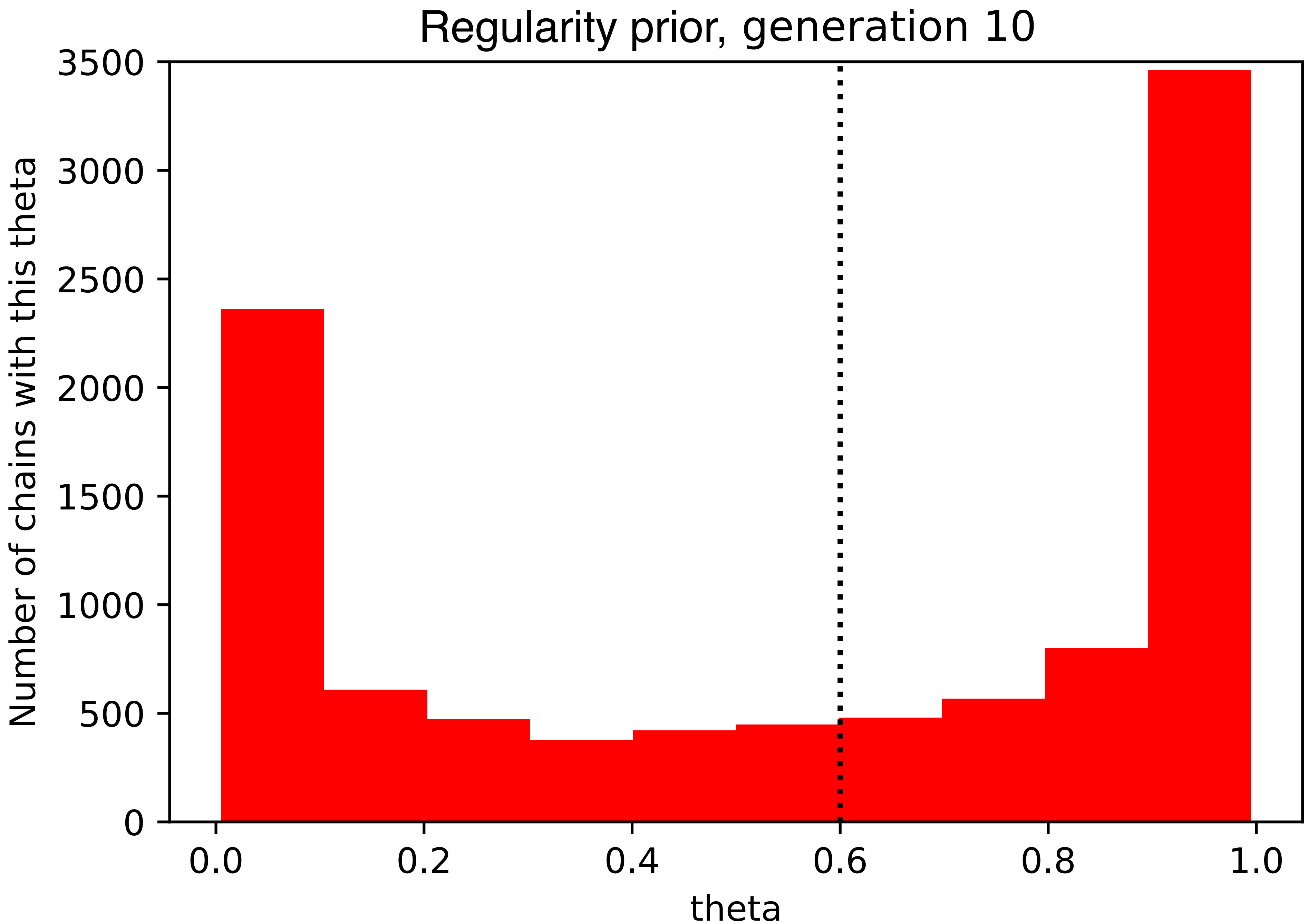
Watching the prior reveal itself



Watching the prior reveal itself



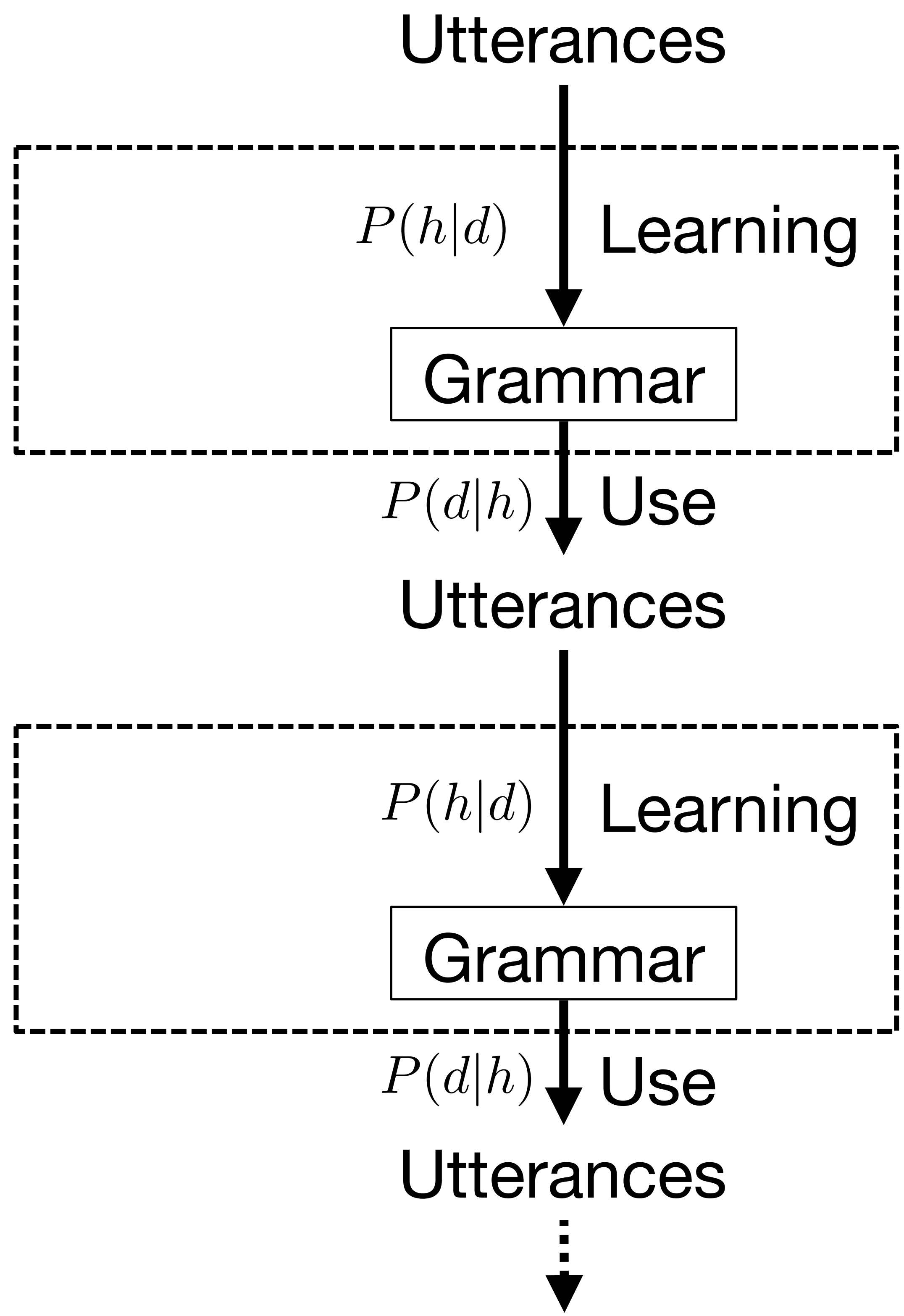
Watching the prior reveal itself



Modelling **iterated learning**

Simulate language transmission from learner to learner.

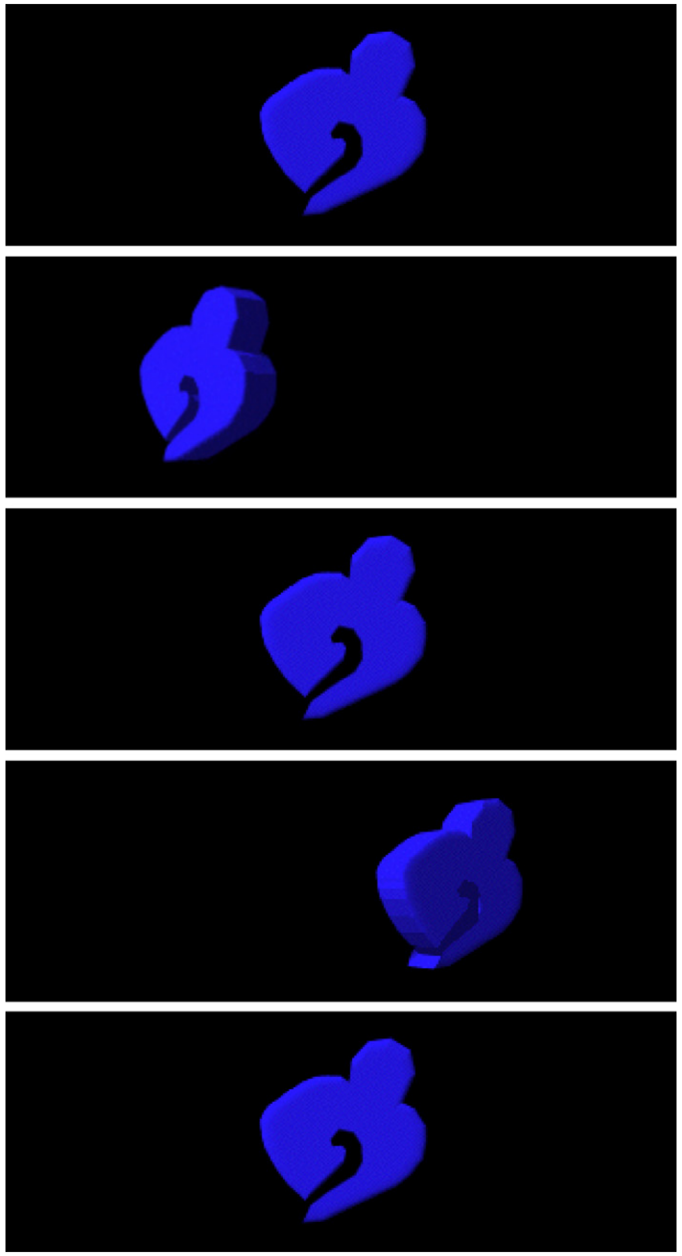
Over time, the bias reveals itself

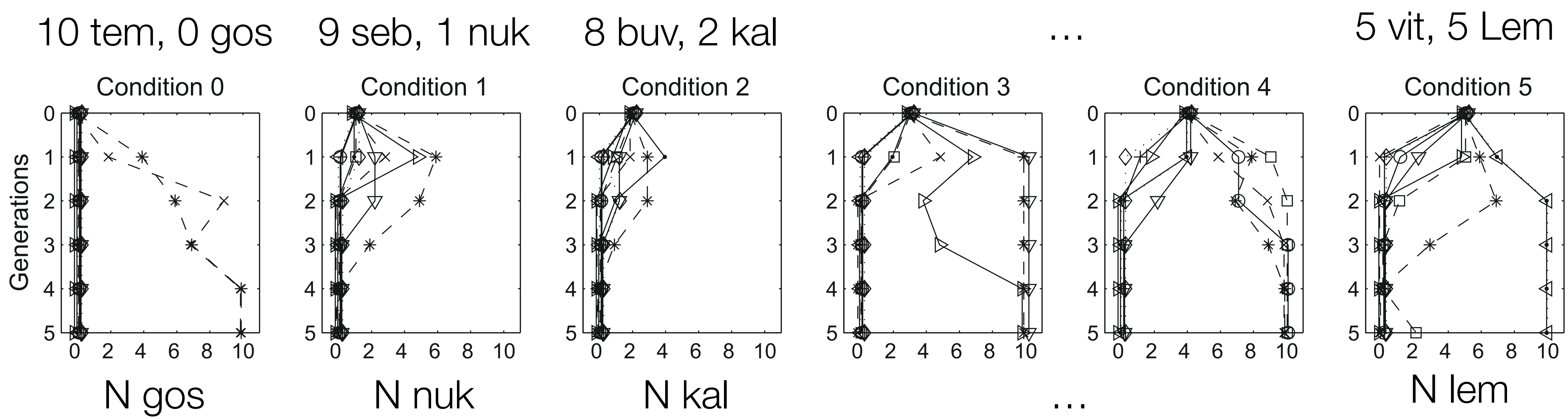


Reali, F., & Griffiths, T. L. (2009). The evolution of frequency distributions: Relating regularization to inductive biases through iterated learning. Cognition, 111, 317–328.

An iterated vocabulary learning experiment

- 6 objects, each object has two labels
- Training: see objects labelled 10 times each
- Testing: label each object 10 times
- Initial language:
    - Object 1: “tef” 10 times, “gos” 0 times
    - Object 2: “seb” 9 times, “nuk” 1 time
    - Object 3: “buv” 8 times, “kal” 2 times
    - …
    - Object 6: “vit” 5 times, “lem” 5 times



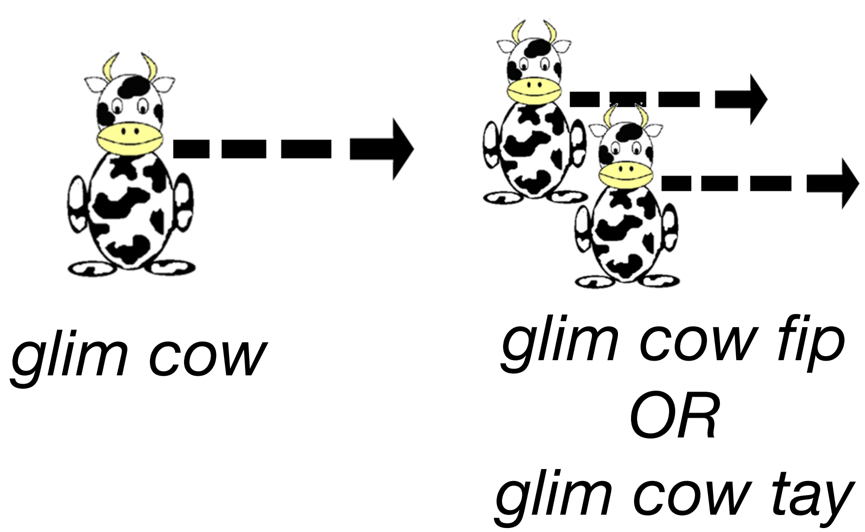


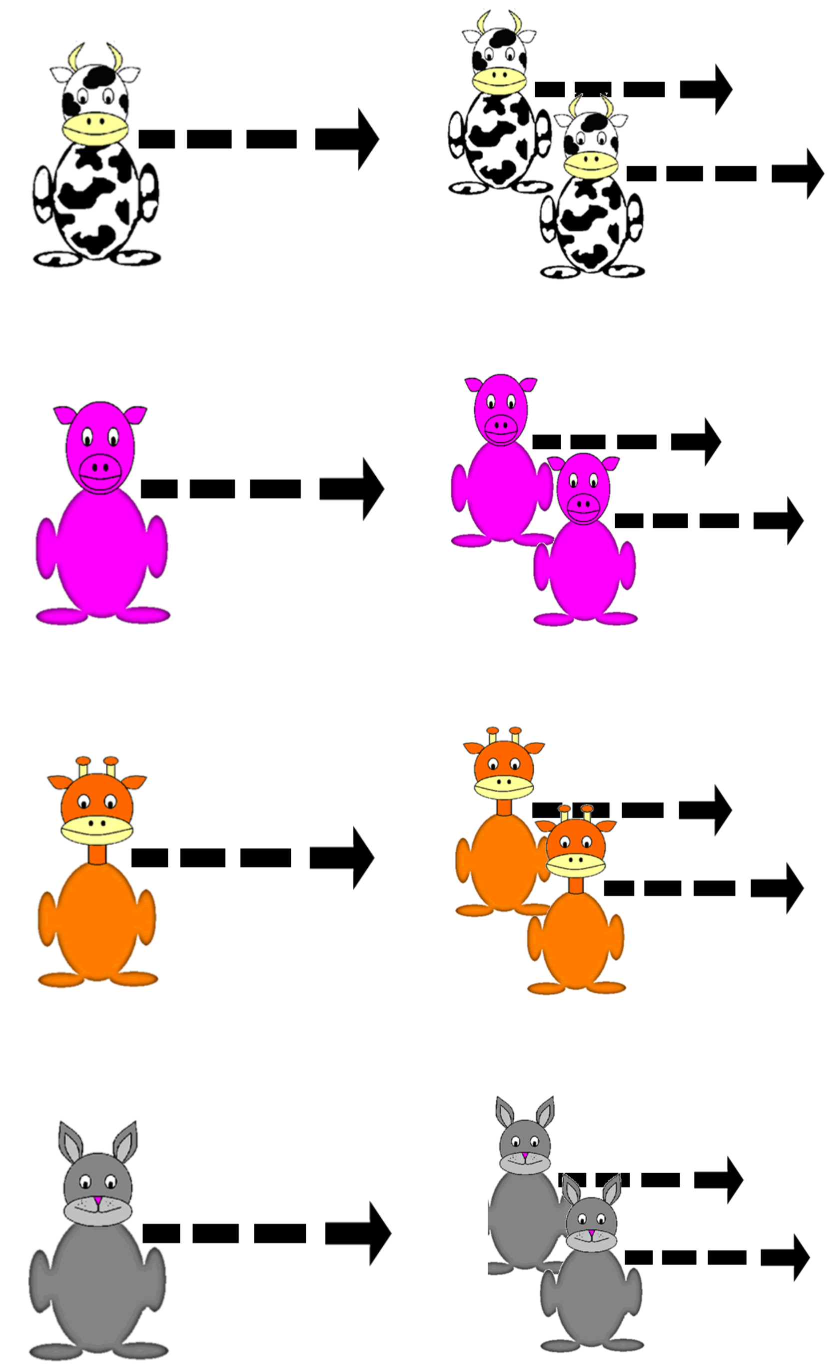
10 tem, 0 gos
9 seb, 1 nuk
8 buv, 2 kal
…
5 vit, 5 Lem
Condition 0
Condition 1
Condition 2
Condition 3
Condition 4
Condition 5
Generations
0
1
2
3
4
5
0 2 4 6 8 10
N gos
N nuk
N kal
…
N lem

Smith, K., & Wonnacott, E. (2010). Eliminating unpredictable variation through iterated learning. Cognition, 116, 444–449.

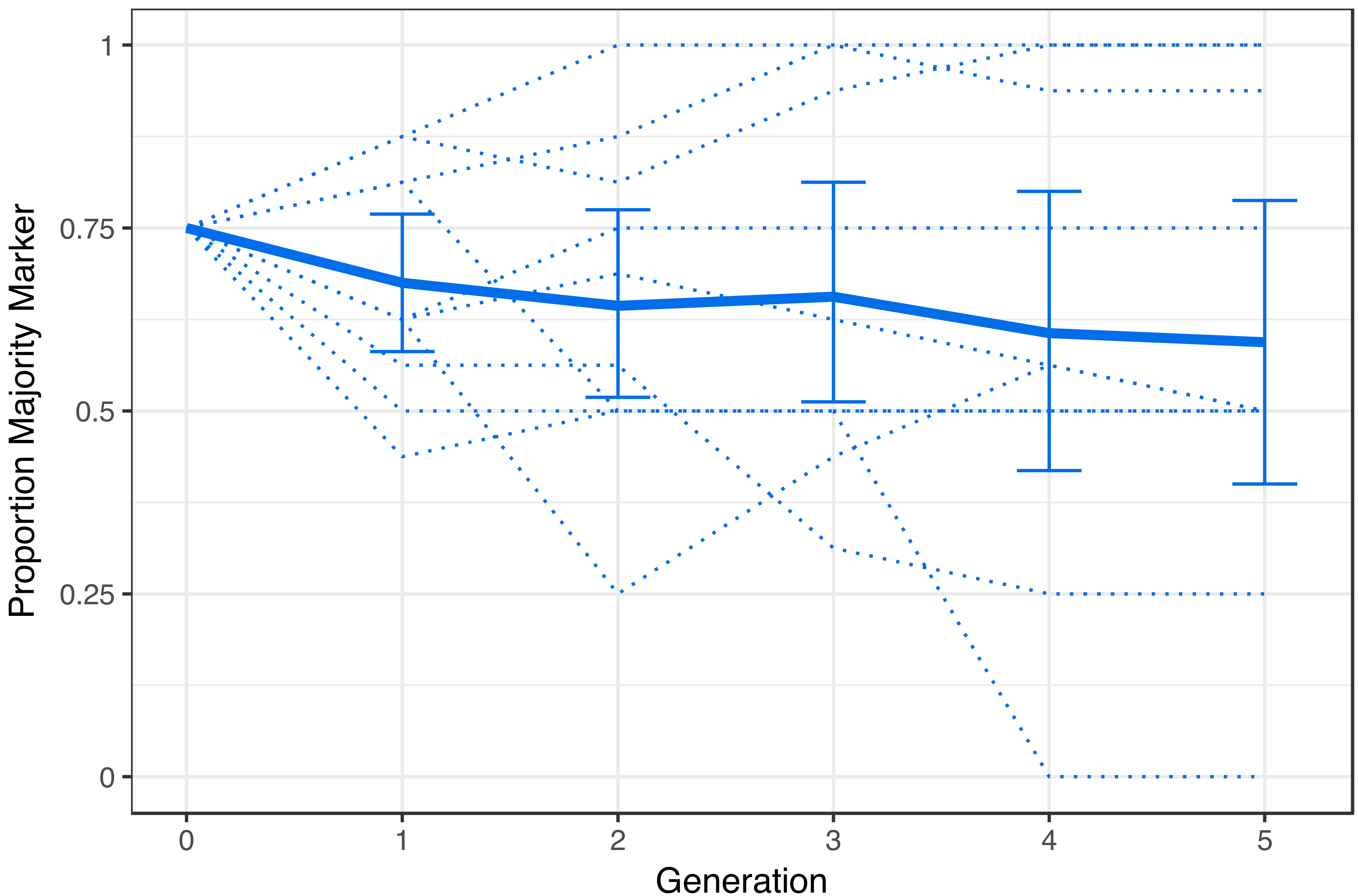
An iterated artificial language learning experiment

- 4 animals, presented in singular or plural
- Training: see scenes plus descriptions
- Testing: produce descriptions
- Initial language:

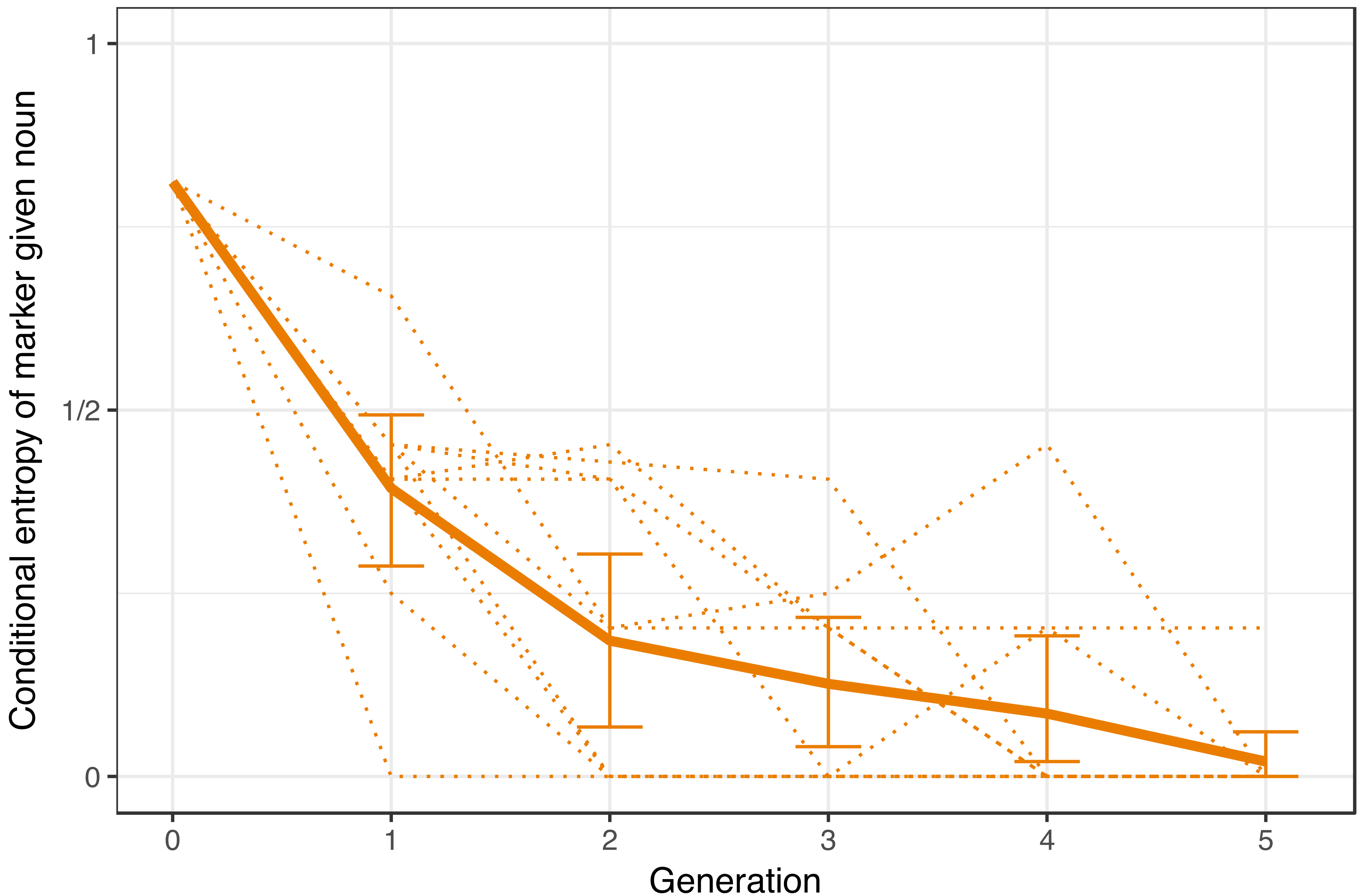




Variation usually maintained…



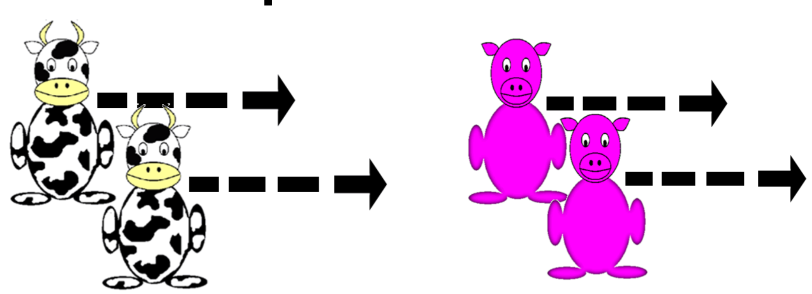
… but becomes predictable



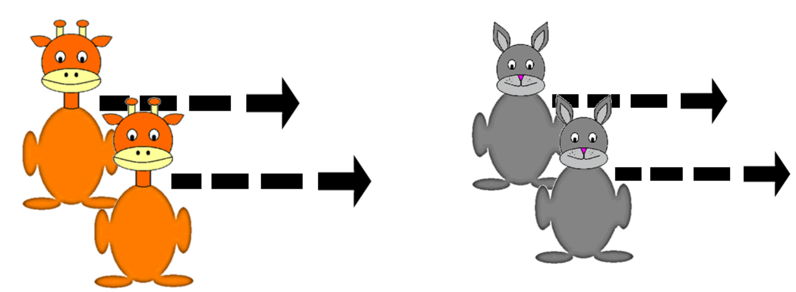
Simple system of conditioned variation

Predictable variation, rather than zero variability, gradually develops: a simple **noun class** system

“fip” nouns



“tay” nouns



Summary and next up

- **Beta-binomial model** allows us to model how learners respond to variability
- Two important insights:
  - If you study learning in individuals, data can obscure the prior
  - The prior can reveal itself over iterated learning
- Lab experiments show this same cumulative, gradual regularization can produce patterns of conditioned variation, like we see in natural languages
- Lab: iterated Bayesian learning
- Next lecture: Communication and the Rational Speech Act model