

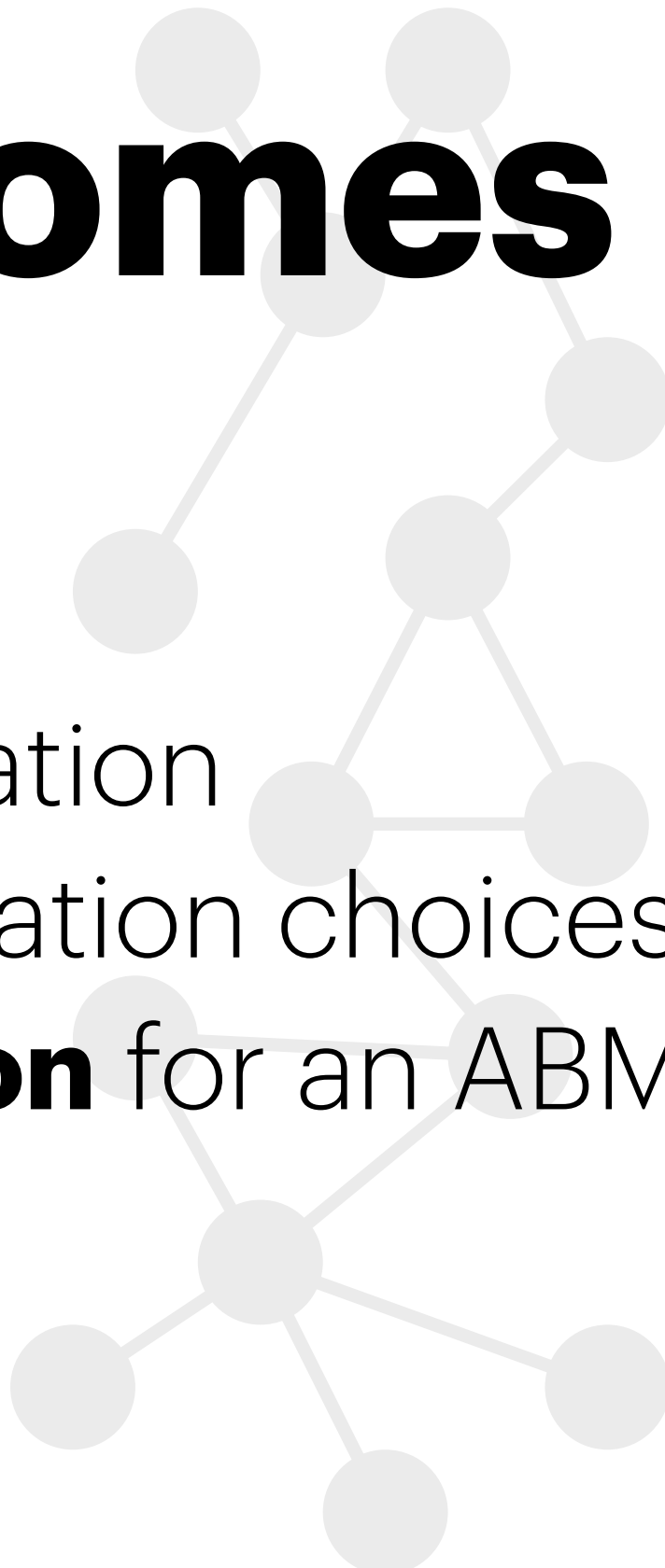
Validation

UGH...

**ANOTHER VALIDATION
MEME**

makeameme.org

Learning outcomes



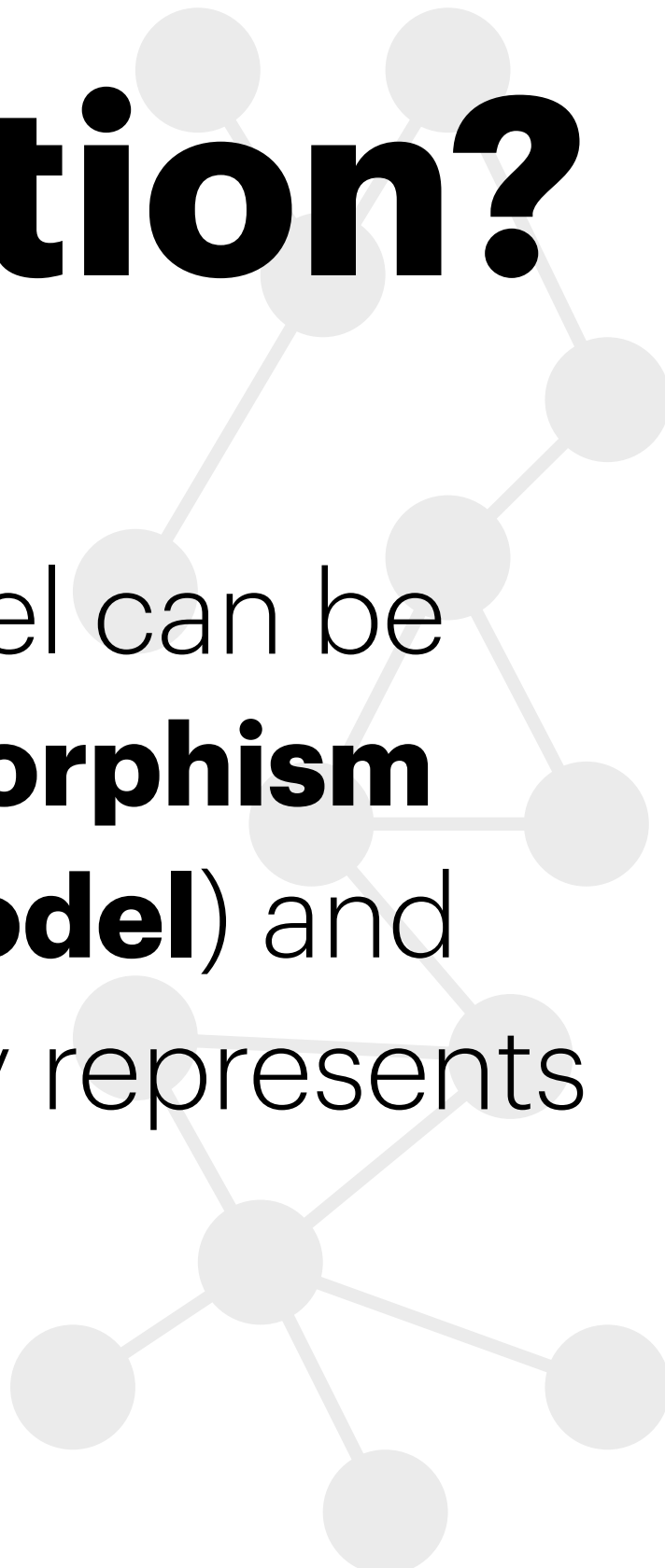
Understand the **concept** of validation

Discuss the **implications** of validation choices

Learn how to **approach validation** for an ABM

What is validation?

Definition: the **validity** of a model can be defined as the degree of **homomorphism** between a certain system (**the model**) and another system that it purportedly represents (**the real-world system**)



What is validation?

Definition: the **validity** of a model can be defined as the degree of

Is the model good at reproducing the real system?
Is there a better alternative?

that it purportedly represents (**the real-world system**)

Dimensions of validation



Concept validation - is it consistent with theory?

Empirical validation - is it consistent with data?

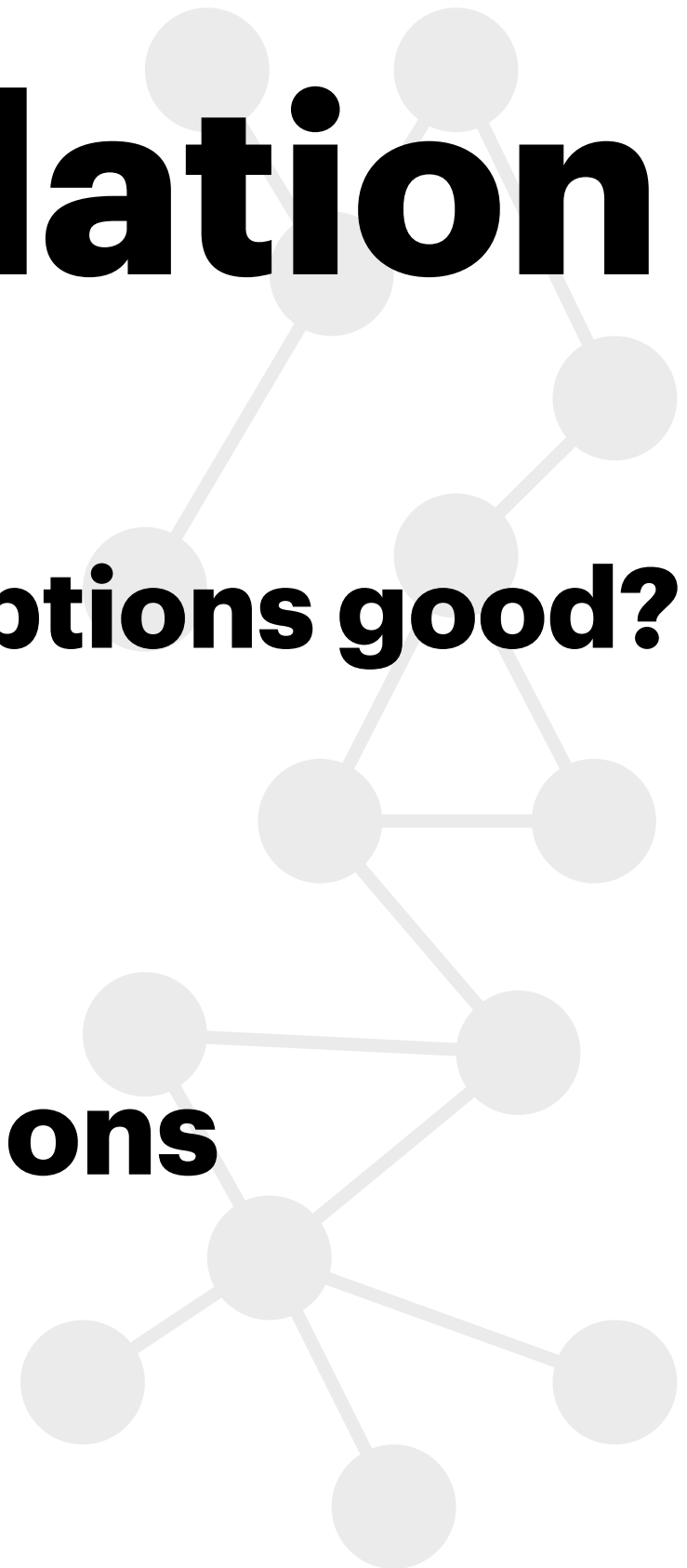
Empirical validation

Input validation - are the assumptions good?

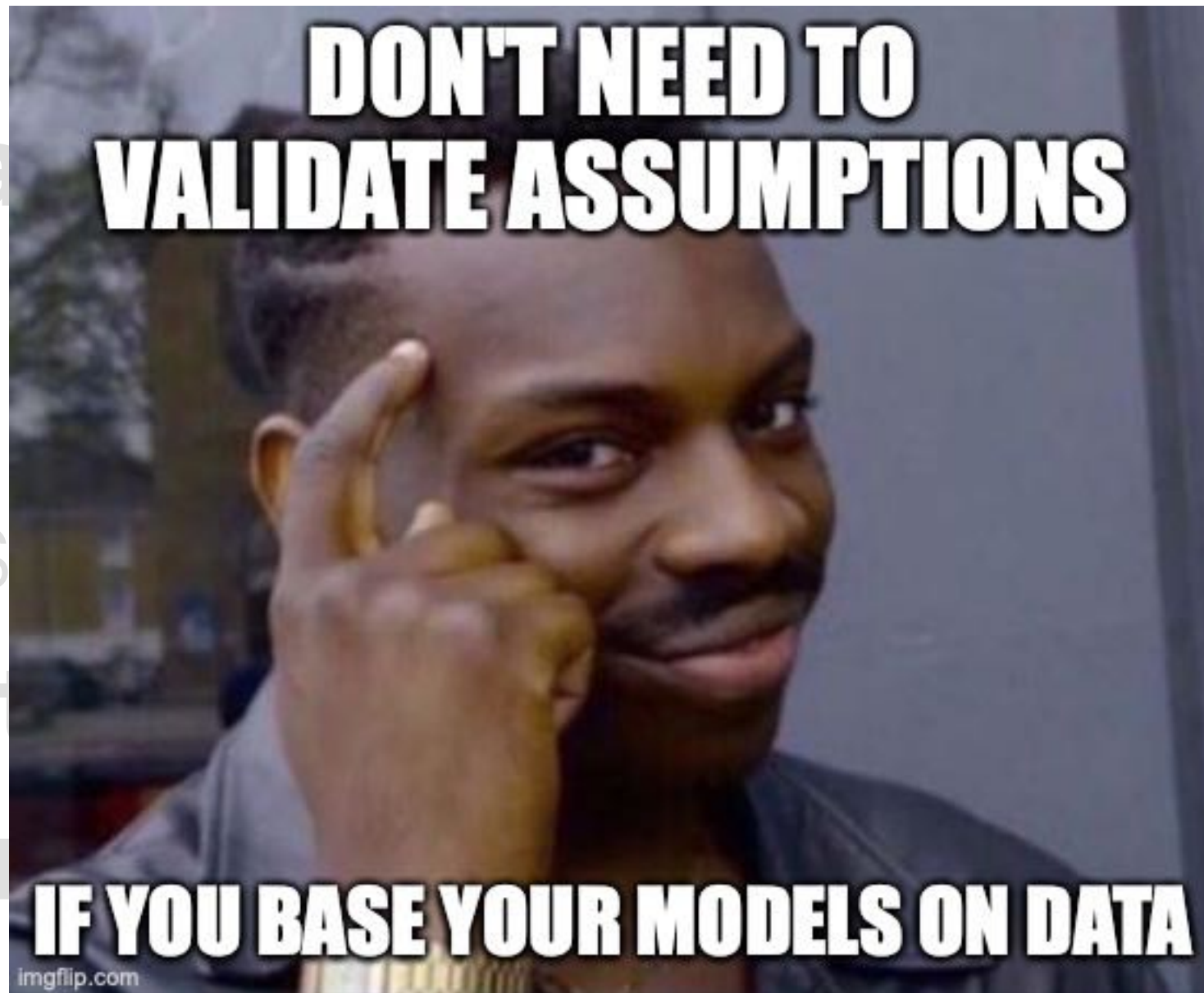
Two classes of input validation:

Parameters and initial conditions

Structural assumptions



Empirical validation



Input variables
good?

Two classes

Parameters

Structure

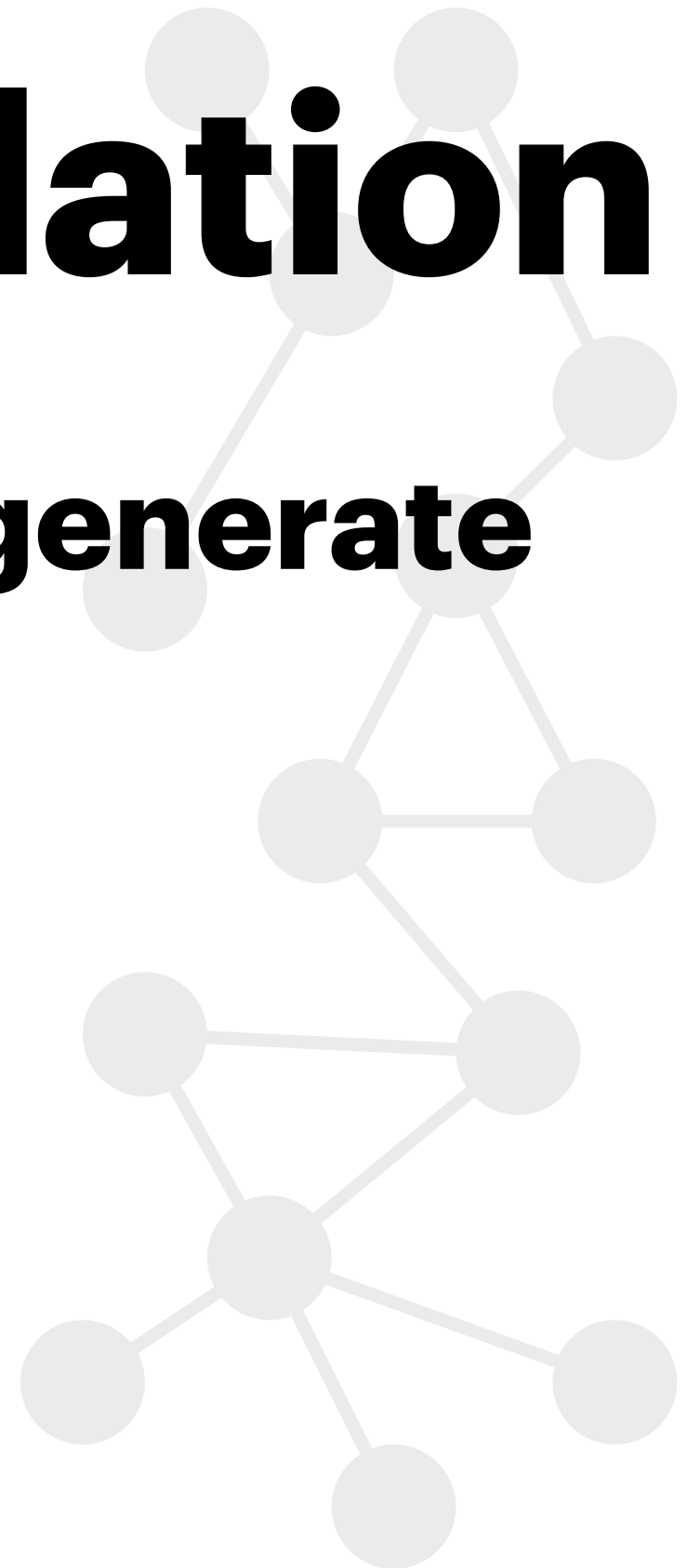
assumptions

assumptions

Empirical validation

Output validation - does it generate plausible implications?

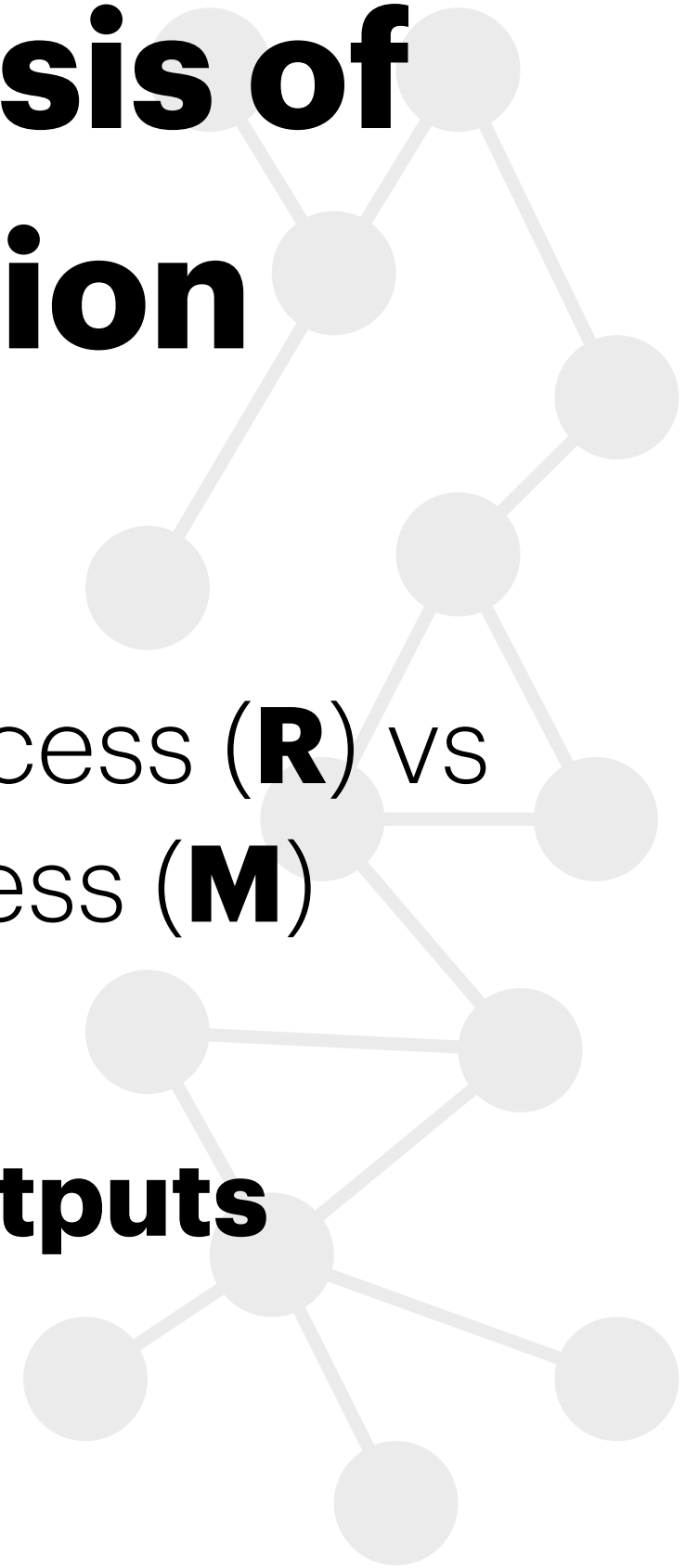
Calibration/estimation



Methodological basis of empirical validation

Real-world data-generating process (**R**) vs
model data-generating process (**M**)

We need to compare the **outputs**
of **R** and **M**



Assessment

A model is said to be:

Useful if exhibit **at least some** of the observed historical patterns

Accurate if it exhibits **only** patterns observed historically

Complete if it exhibits **all** historical observed patterns



Assessment

$R \cap M = \emptyset$ **Useless**

$M \subset R$ **Incomplete**

$R \subset M$ **Inaccurate/redundant**

$M \iff R$ **Complete and accurate**



Tractability vs accuracy

Simplification

Stylised facts

Trade-off between tractability and accuracy



Instrumentalism vs realism

Realism says theory is real

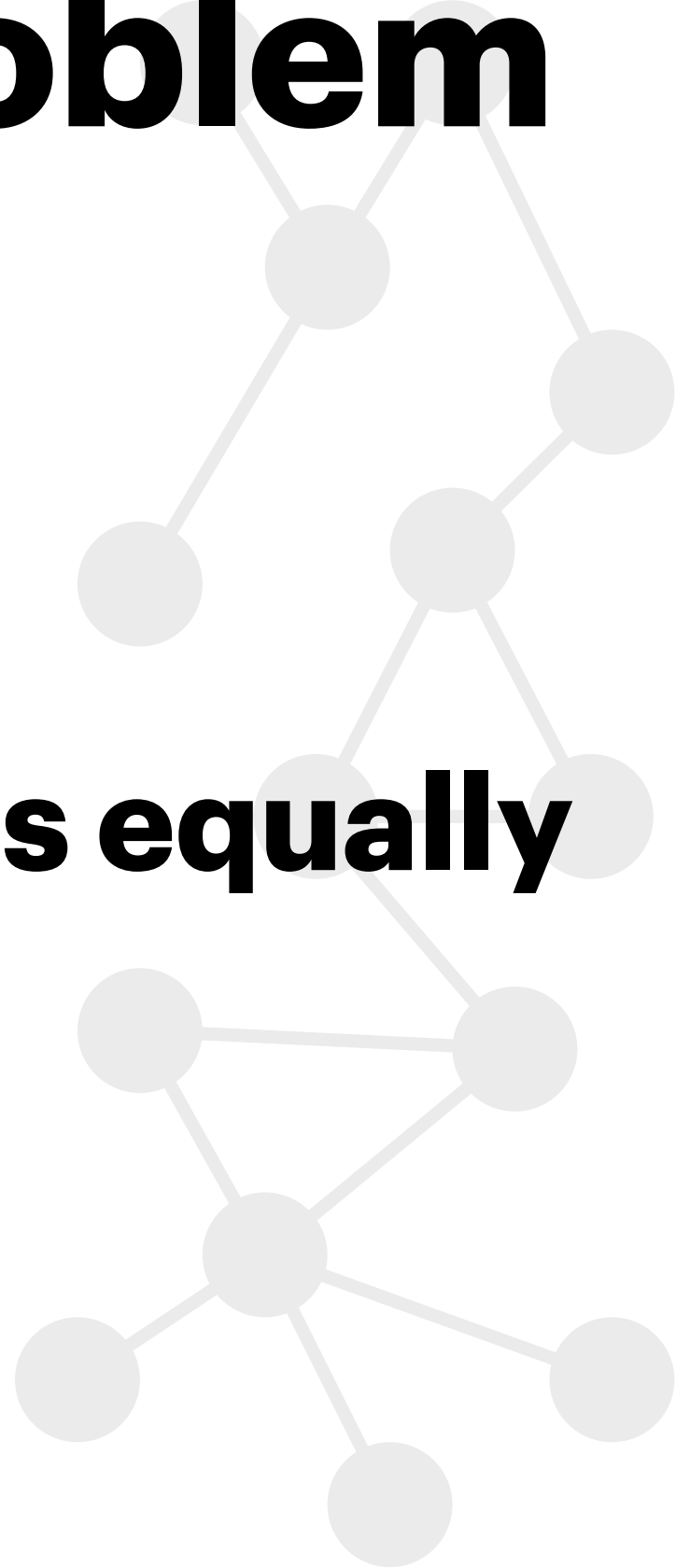
Instrumentalism says theory is an instrument

Are we interested in reproducing the world
or in making predictions?



Identification problem

What if two different models equally reproduce reality?





"To be is to do."

- Socrates



"To do is to be."

- Kant

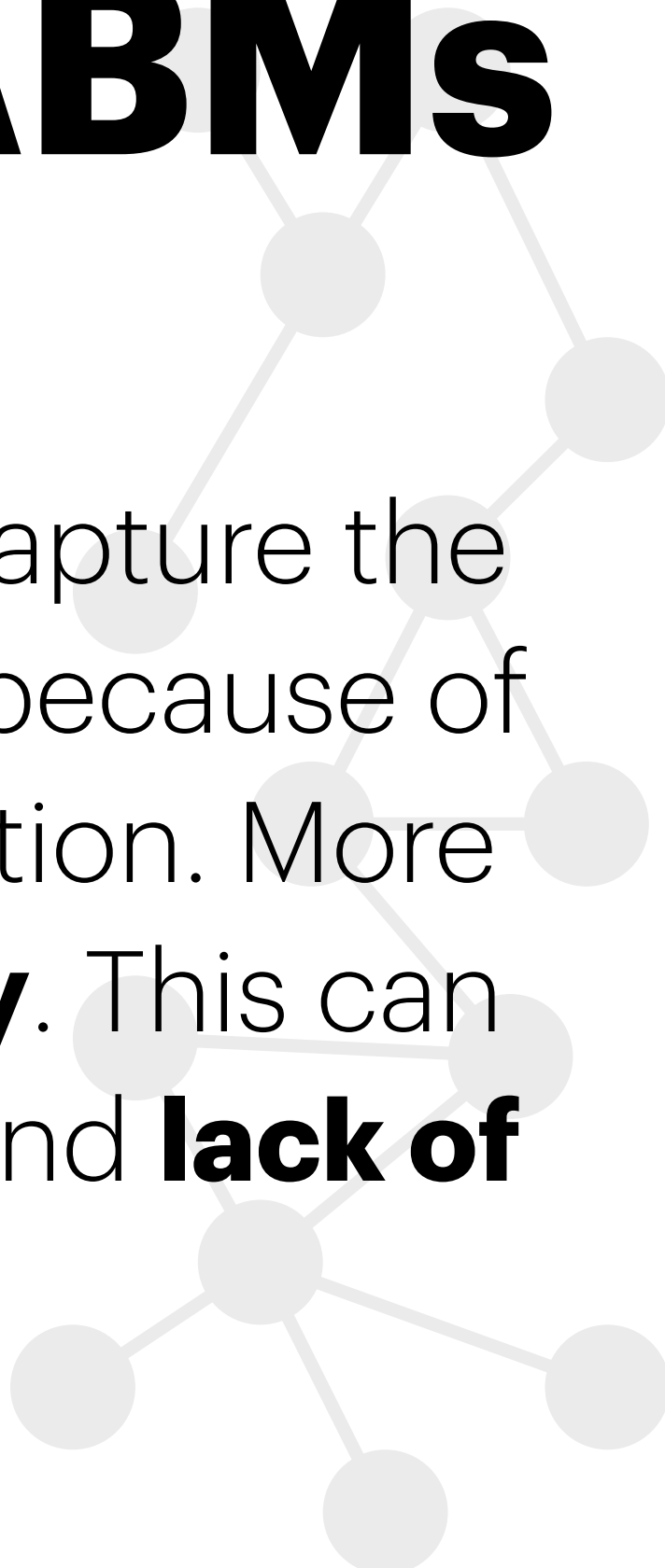


"Do be do be do."

- Scooby-Doo

Problems of ABMs

ABMs are flexible and allow us to capture the complexity of a system. However, because of this, we need **more data** for validation. More data also means **more complexity**. This can cause issues such as **overfitting** and **lack of interpretability**.



Output validation

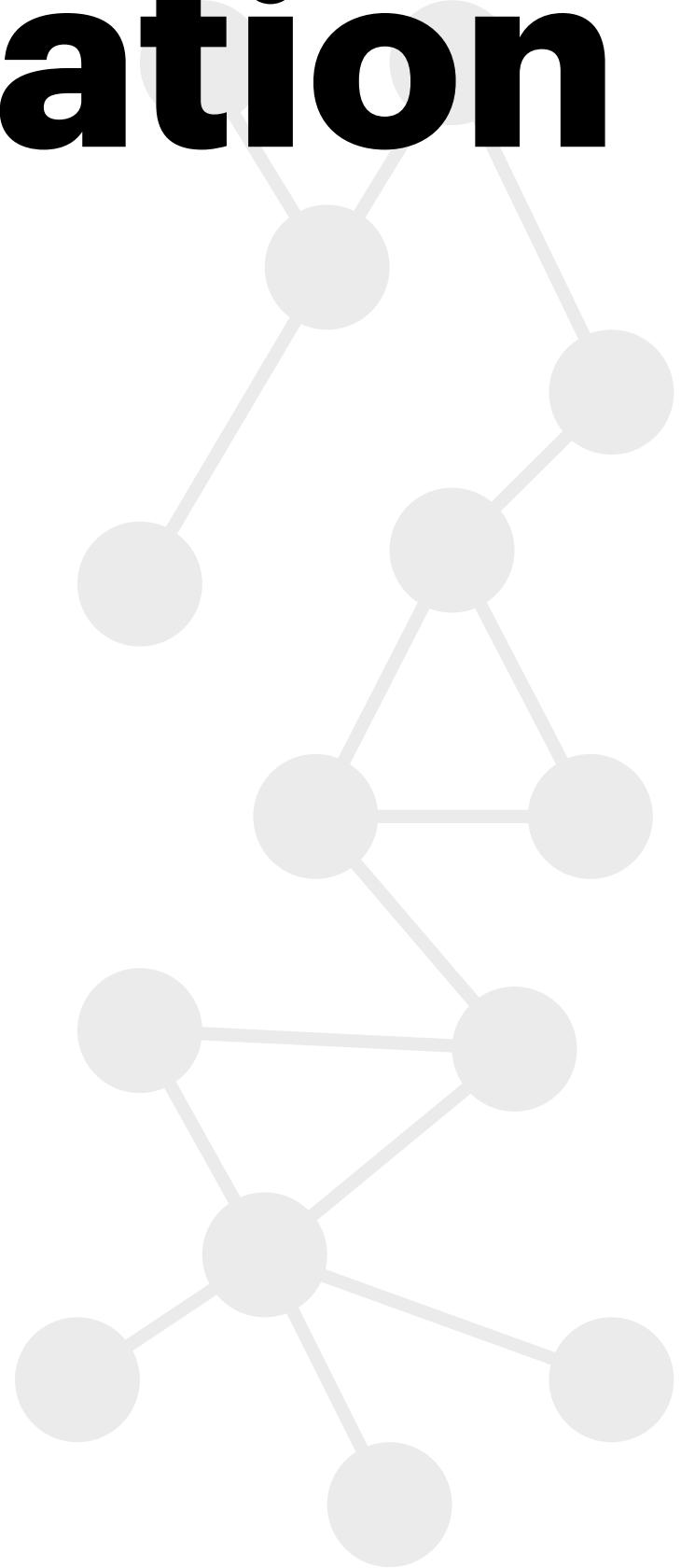
My suggestion:

**Use as much data as
you can!!!**

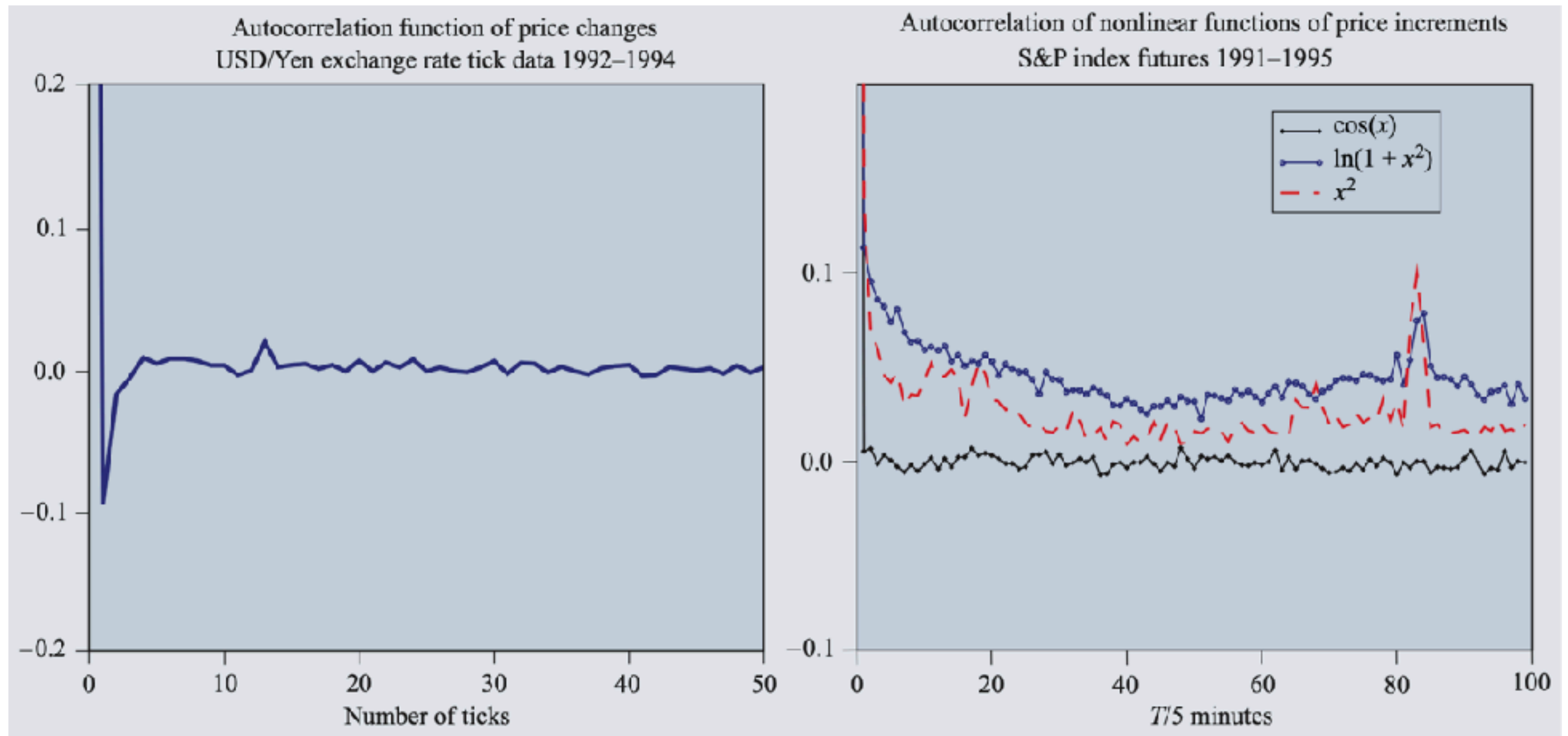


Indirect calibration

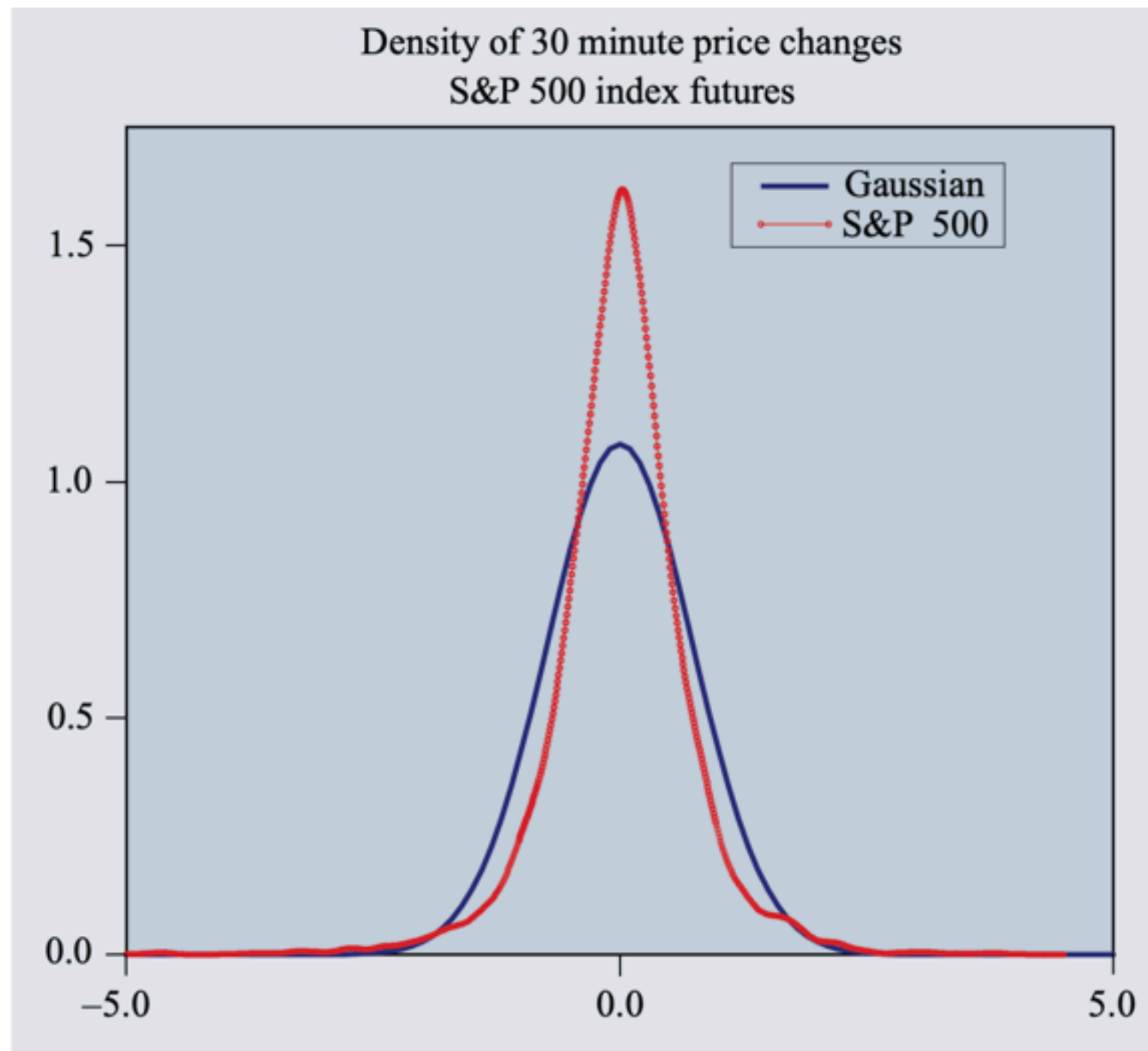
Stylised facts



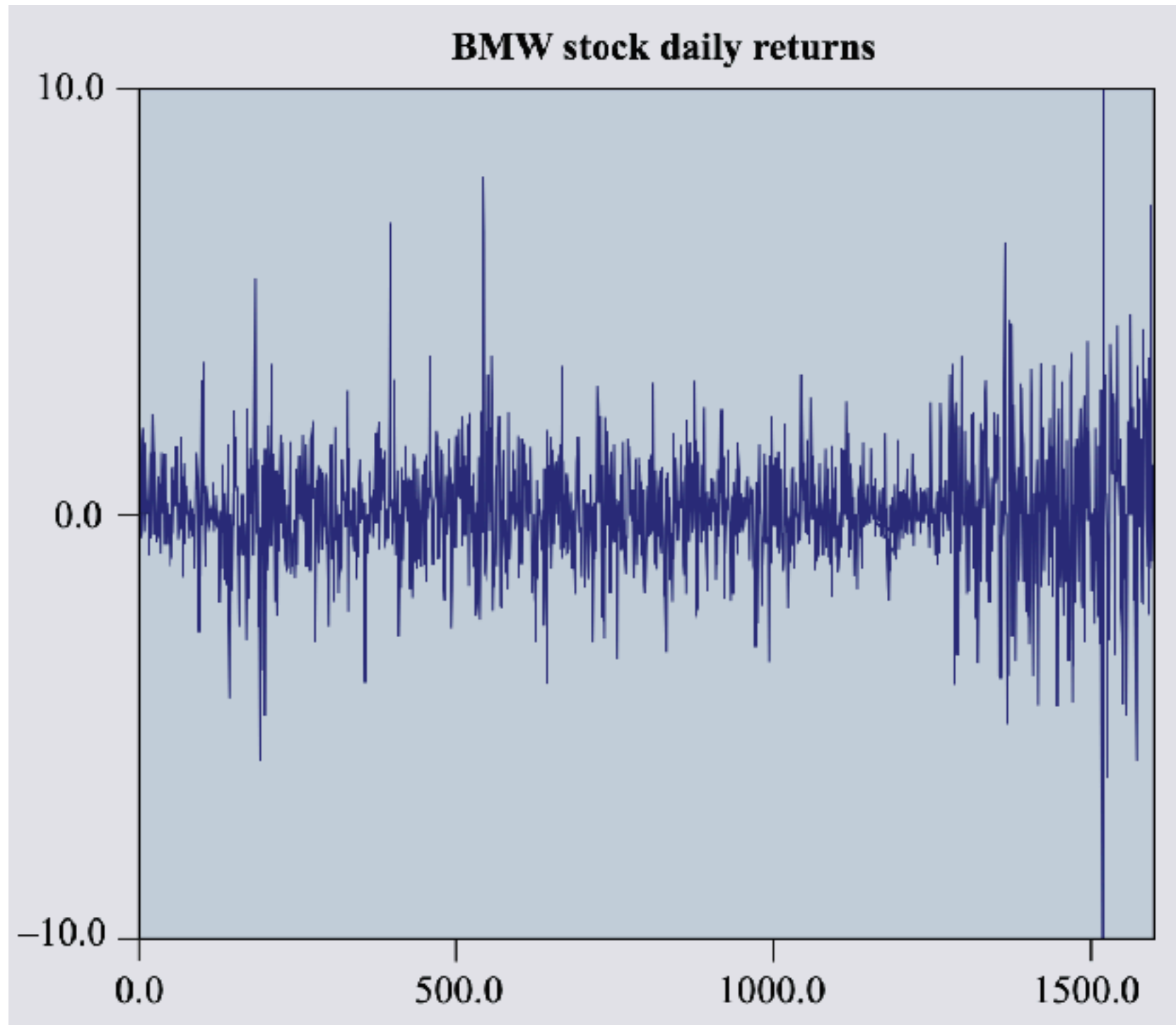
Indirect calibration



Indirect calibration



Indirect calibration



Tesco example - Recap

Two types of customer agents with heterogeneous variables*

One type of worker agent

Variables: time spent in store, expenditure, bias on deals, etc.

Actions: buy, recommend (last year people also suggested “steal”)

Interactions: recommend products to others, interact with store



Tesco example - Recap

Two types of customer agents with heterogeneous variables*

One type of worker agent

Variables: time spent in store, expenditure, bias on deals, etc.

Actions: buy, recommend (last year people also suggested “steal”)

Interactions: recommend products to others, interact with store



What data would you use for validation?

What could be used as stylised facts?

Tesco example - Recap

Two types of customer agents with heterogeneous variables*

One type of worker agent

Variables: time spent in store, expenditure, bias on deals, etc.

Actions: buy, recommend (last year people also suggested “steal”)

Interactions: recommend products to others, interact with store



What kind of validation would you perform?
With what data?