
The effects of disinformation in Rating Networks

— by Giulio G. Cantone —

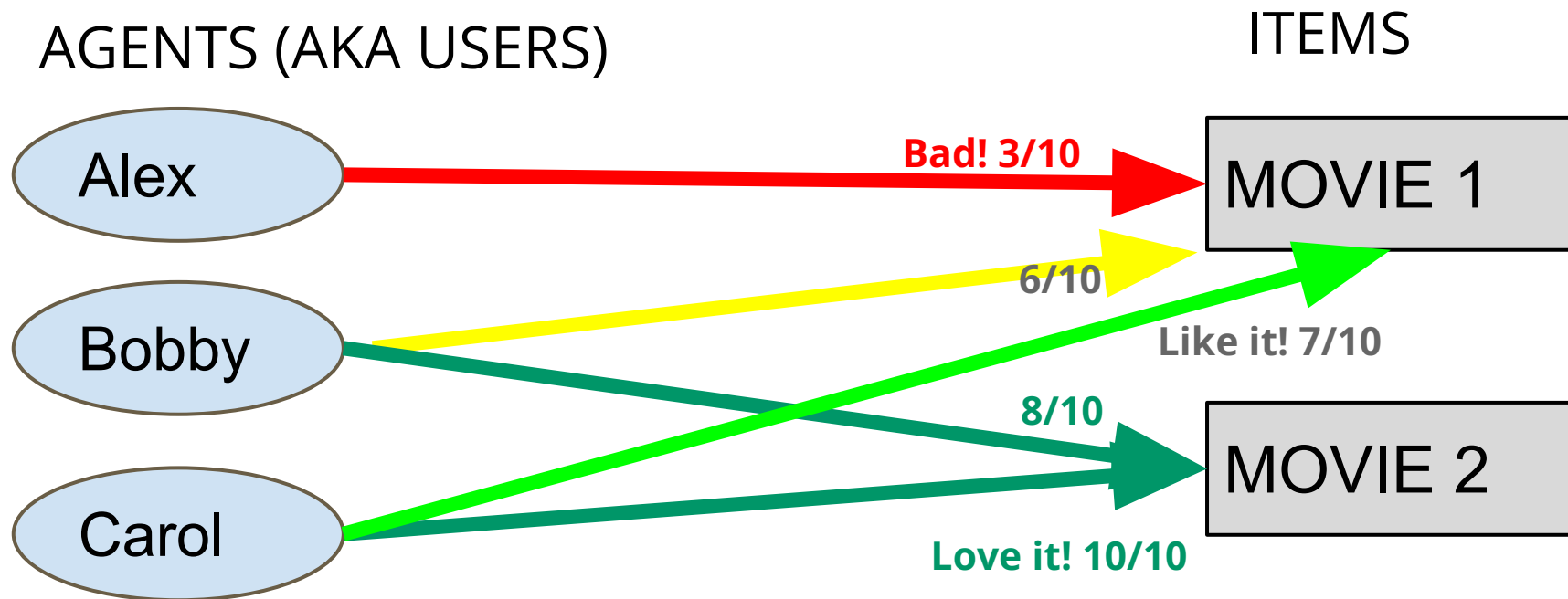
Index

- **Rating Networks**, similarities with Twitter and Youtube
- **Disinformation** in the infamous Metacritic
- **Review Bomb** in The Last of Us Part II
- Reciprocal/simultaneous **causality models**
- De-bias a rating network? **A simulation**

Rating Networks

- Links have a **y**-attribute: **quantitative** opinion/judgement/sentiment, a **rating**

- Bipartite
- Direct



Why causality in Rating Networks is important?

- Assuming that people “observe” summary information on ratings (averages, rankings)...
- ...then we can infer social influence / information cascades.
- “Effects” here is a very broad concept because causality is often cyclical.

Bias #1: Selection and accountability of agents

Entry barriers for agents (open/restricted access).

Business made for collecting information:
most RN are open access, easy to make new account.

Problem seen in Twitter, YouTube, ...

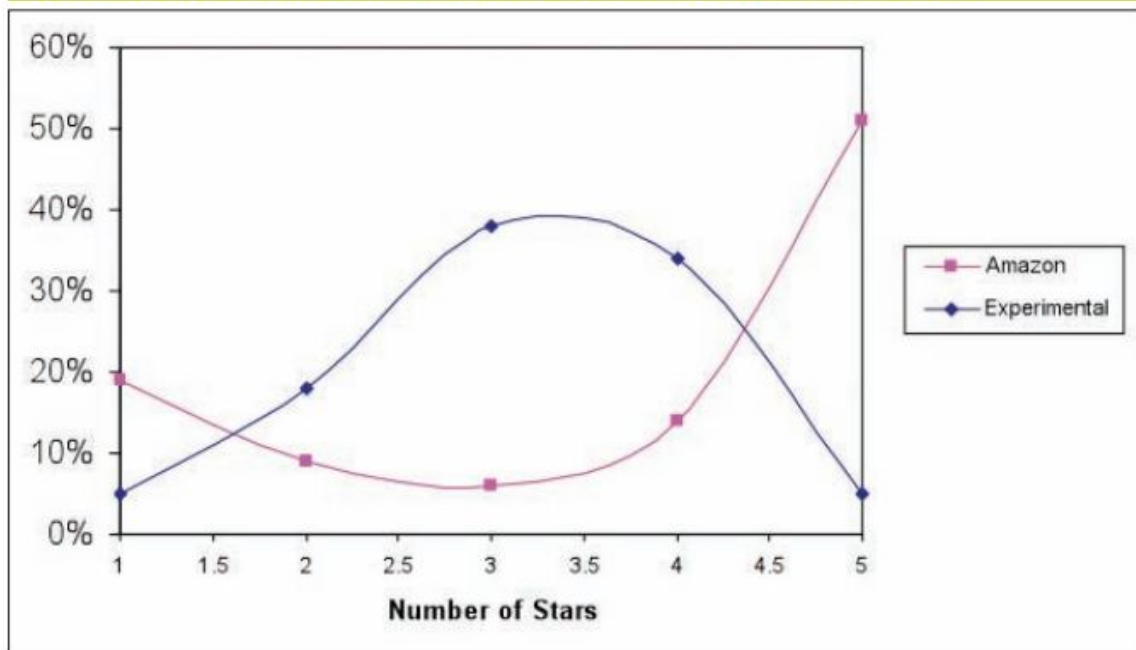
- Same issues for **disinformation**...

Few RN have a small barrier (eg Netflix, Amazon).

Bias #2 and #3: Item availability; unequal link formation

UNIVERSAL J-SHAPE OF RATING NETWORKS

Figure 2. Distribution of Experimental versus Amazon's Ratings for a Music CD



- **Availability bias:** the item must be known
- Unequal likelihood of reviewing after experience.
- **Brag-or-moan dynamics.**
- No direct social influence involved in this model!

Some people are working on correcting these biases

<http://datacolada.org/72>

Wonderful blog on
metascience.

We will discuss this later in the slides.

Let's talk about disinformation

Disinformation: astroturfer bots

10

Arolo

This game is practically awesome.
An easy must have.

Played and adored this masterpiece

32 of 54 users found this helpful

10

lthughy

Probably, it's the recent best platform of the last decade.
The most original, and surprisingly fun to play.
I recommend to everyone who loves the genre.

41 of 73 users found this helpful

10

Ifhyn

One of the best games I've ever played.
Amazing and fun, with awesome gameplay and incredible platform design.

Absolutely gorgeous game

Regular users can spot bots

0

Unholydead

Mar 28, 2021

As a Square Enix fanboy, I cannot in good conscience play more than 5 seconds of this game. If there is a single positive review that WASN'T written by a bot here I'd be surprised. Stay away. This is a scam, dudes.

21 of 35 users found this helpful

All this user's reviews

0

B-Bog

Mar 27, 2021

DO NOT TRUST THESE POSITIVE USER REVIEWS. THIS GAME IS ABSOLUTE GARBAGE.

Seems like Squirt-Enix are resorting to the cheapest tactics possible and are trying to buy their way out of this mess by flooding this site with fake user reviews (look at the profiles of these people, none of them have reviewed anything else besides this POS). For which, shame on you, you f**ers! Don't you have ANY dignity left?

There was a reason as to why they didn't hand out any review copies: This game could not even be considered decent as a free-to-play mobile game. To release something this bad as a fully-priced console title borders on fraud and is absolutely shameless.

If my review gets marked as "not helpful", you know who's to blame (hint: it's not the actual players). **Collapse** ▲

25 of 49 users found this helpful

All this user's reviews

Disinformation: insincerity

- **Giving 0 as a reaction to bots is still insincerity!**
- Sincerity: objectively evaluating the product “as it is”.
Tactical voting to counter-manipulate: still “rational insincerity”.
Huge literature on this on Theory of rational voting.
- (Political) Disinformation could be driven by perception that “society is unfair” -> extremists spread fake news.
- Insincerity coordinates (by chance/by design) -> **escalation**. (I see you, **4chan**)

Submitted 2 papers on a huge case of Review Bomb

Review Bombs are:

- Escalations of highly polarizing votes.
 - Class actions in the information Era.
- Distort summary info in order to get media attention on your point of view.

Often are only protests against commercial practices, unrelated to cultural wars.

The Last of Us Part II was the biggest Review Bomb in history. It was rooted in identity politics + perception that “journalists are paid to give positive reviews”.

The Last of Us Part II (PS4)

Release Date: June 19, 2020



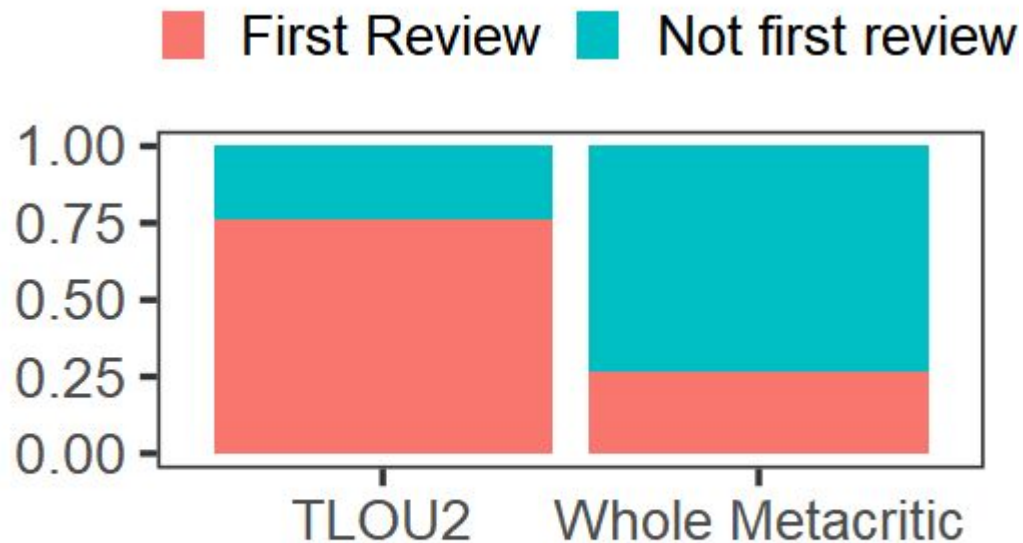
Metascore
94 reviews



User Score
5510 ratings

**12 hours after
release**

Descriptive paper was accepted



n=78,219

n = 1,469,459

SOLUTION:

if people can pre-register their legitimate biases ("I saw the trailer, I didn't like it at all") we could trust them much more.

Issue with 2nd paper, on inference:

Deep research question:

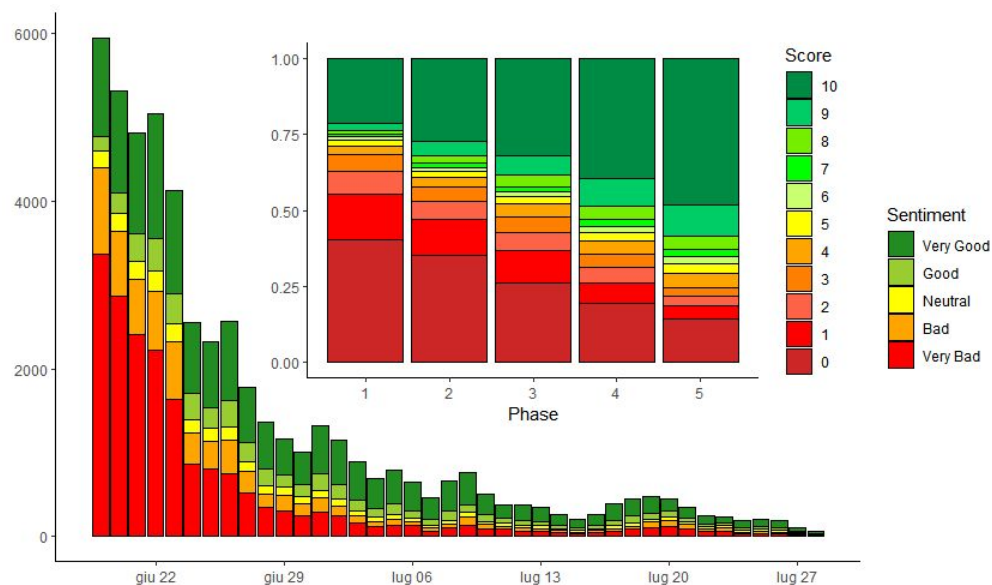
One sees

journalists praising an item,

but **users' review super negative,**

opinion of **A** would be **influenced?** How?

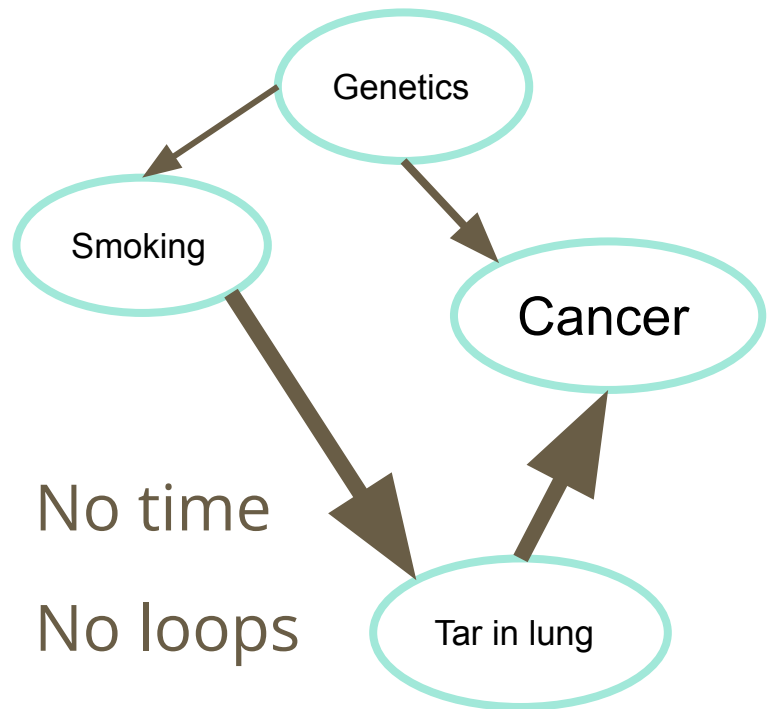
For many reason, I think that this question has no real answer in the observational data.



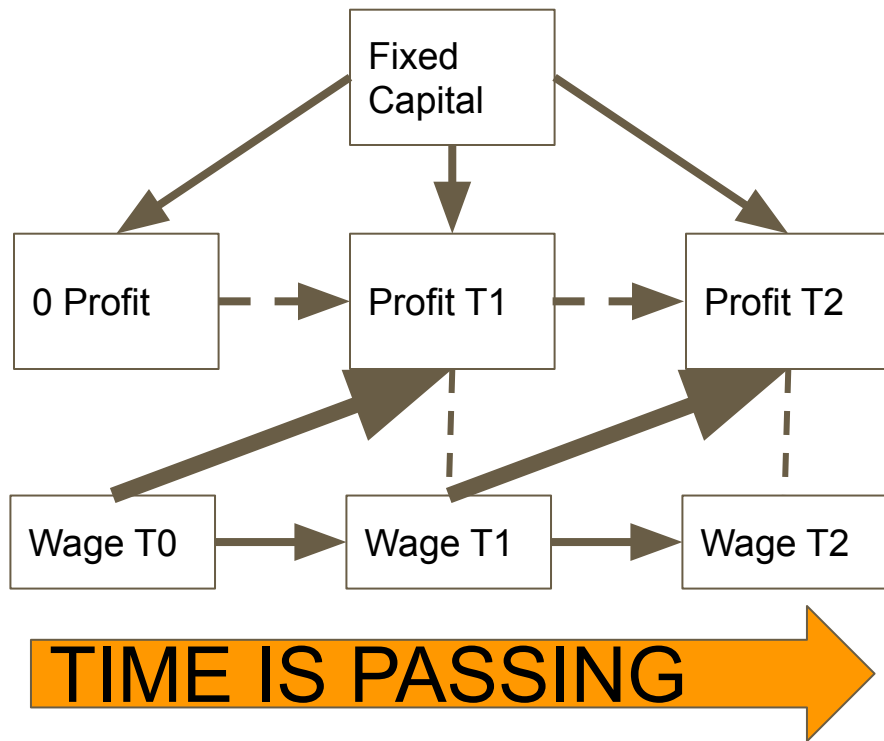
Small digression on causal methods, to illustrate something that could be useful

There are 2 macro-kinds of observational causal models

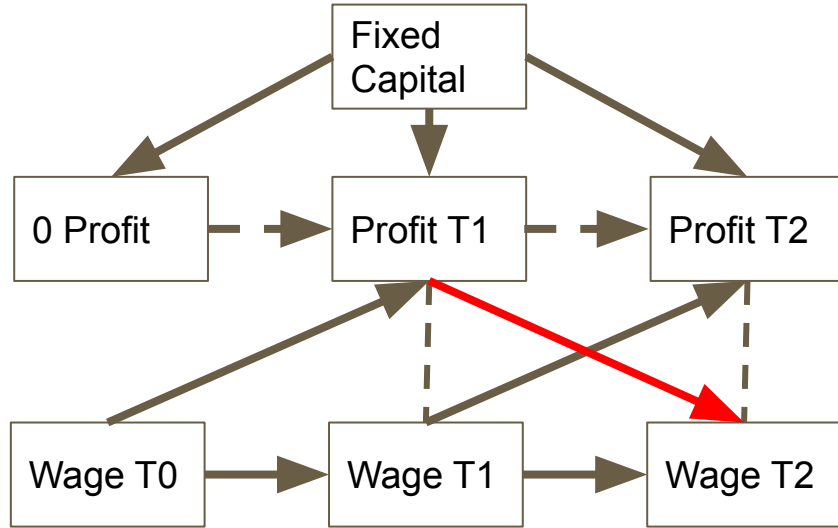
DAGS, the gold standard



Wiener-Granger (people hate this)

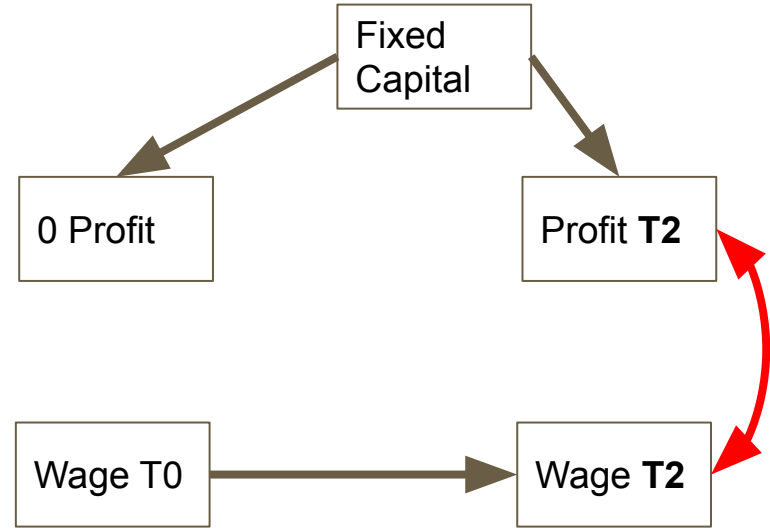


Reciprocal Causality



Why reciprocal causality happens
anticipation of causality in agents.

Simultaneous Causality



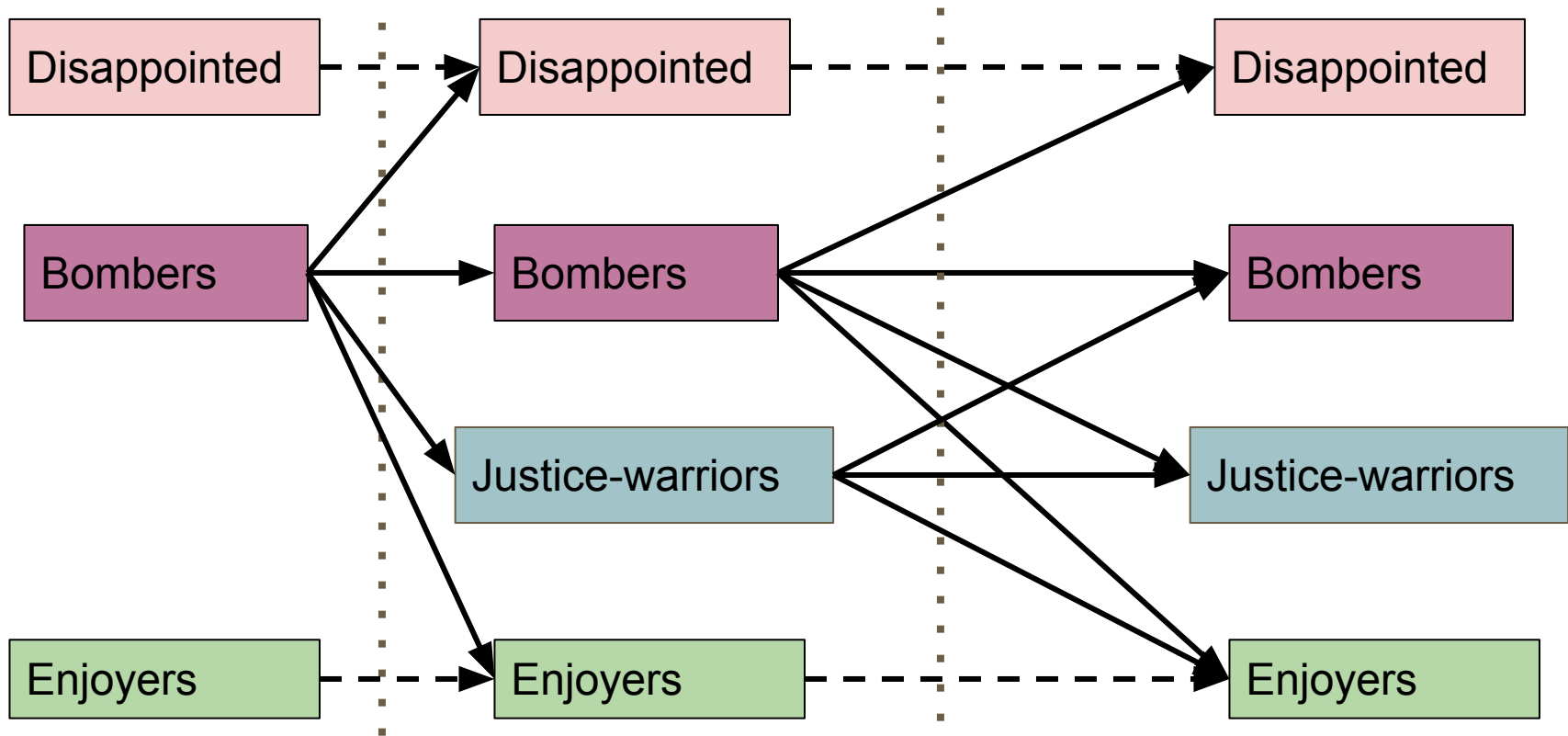
Simultaneous models are less valid, but
perform better with granular data

START

$T=1, T=2 \dots$

END

Hypothesis



What the results are supposed to tell us at the end?

Quantify the strength of the arrows, as regression coefficients.



The coefficient of the structural equation model tells if:

- negative review bombs are effective in manipulation of the general opinion, or...
- if they backfire because justice-warriors will fight back.

Same could happen for contrasting astroturfers.

Lot of unseen confounders...

Reciprocal causality connects Granger to DAGs

- To solve a reciprocal causality model:
multi-stage regression.

Eg, **2-stage Linear Regression (2SLR)**,
that is used in DAGs, too.

<https://mda.gesis.org/index.php/mda/article/view/2017.07>

Agent-Based Models and Estimation Techniques

Simmons' Theory

- **Professional critics have no time to rate everything**, they maximise the utility of their review (**Principle of Least Effort**), so...
- ...they avoid super-niche products, **unless** they really like them and want to promote these (Collider bias of selection).

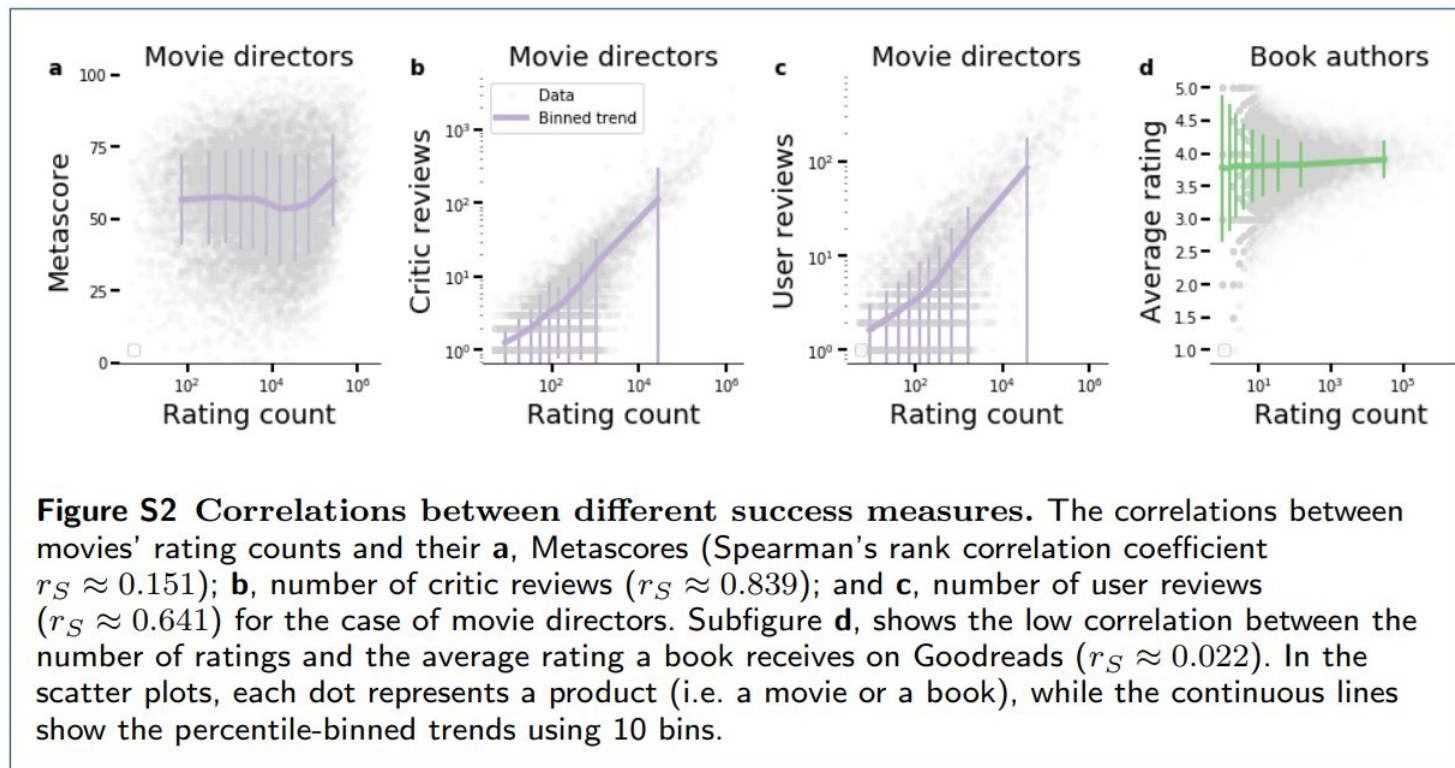
Effect:

- There is no correlation between n of reviews and average score

But: if we **random assign** critics to products, **fixed the observed n** ,

- Then correlation happens, because on average non-niche products are a bit better.

Data support this Theory



Simmons' Proposal

$$n_{max} - n_i = \nu_i$$

$$\theta(i) = \sum_1^{n_i} \frac{x_i}{n_i + \nu_i} + \sum_1^{\nu_i} \frac{\zeta}{n_i + \nu_i}$$

In Simmons' Proposal, ζ is a constant.

Elements of agent-based network

- Set **I** Items, each Item has a latent attribute of quality
- Set **U** of Agents. Different classes: honest agents, dishonest agents...
- A mechanic of connection $U \rightsquigarrow I$
- A mechanic of scoring: once $U \rightsquigarrow I$, then how score Y is determined?
This depends by the class of the agent.
- Estimator: function that converts vectors of scores into a single value.
- Diagnostic tools:
 - errors between estimator and latent qualities
 - error only in the ranking of estimators vs latent qualities

Scenarios: error attribution

Disinformation
Agents

Preferential attachment

	NO	YES
NO	Random Bipartite Rating Network	PA Bipartite Rating Network
YES	Manipulated Network (linear error growth)	Manipulated Network (complex error growth)

A lot of details can be modeled...

Preferential attachment comes in few varieties:

- latent quality attribute
- observable k
- both?

A lot of way to model agents:

- Honest
- Honest but very influenced by previous votes (conformist, contrarian)
- Tactically dishonest: hidden agenda, reactive on previous votes
- Astroturfers: they will target specific items and always give max score
- others

Scoring mechanics are a mystery (that cannot be unfolded in a laboratory)

The general idea is that there are:

- Latent quality α in Items
- Personality β of agents (that depends on class, too)
- Scoring function $f(\alpha, \beta) = y_{u \rightarrow i}$

Why no results?

- I coded the simulation without disinformation, and I measured a relatively small effect of Preferential Attachment.
- But I wanted to change some parts in the code, possibly simplify...
- ...and I needed more focus

Only when I am 100% sure of this code I will move into coding disinformation agents.

Possible final outcome: a Shiny App that would work like a NetLogo interface, with a back-end coded in R (or Python)

The End...

