Framing Processes in Social Dilemmas.

Formal Modelling and
Experimental Validation.

Christian Steglich, ICS Groningen

presentation prepared for the Mathematical Sociology meeting, 27 November 2002

Purpose of the presentation

is the validation of (some of) LINDENBERG's ideas about the microfoundations of solidarity

as they are e.g. spelled out in Chapter **3** of Doreian & Fararo (eds.): *The Problem of Solidarity: Theories and Models*, Amsterdam 1998 (Gordon & Breach).

How?

- The theory is an application of framing theory.
- Thus: test it by means of "framing analysis."
- Take ***social dilemmas*** as test domain.

A group faces a ***social dilemma*** when the following two properties hold (DAWES 1980):

- each group member is *worse off when (s)he cooperates* than when (s)he defects, irrespective of what the other group members do:

$$\forall \mathbf{i} : \mathbf{v}_i(\mathbf{c}_i \mid .) < \mathbf{v}_i(\mathbf{d}_i \mid .)$$

- each group members is *better off when everyone cooperates* than when everyone defects:

$$\forall \mathbf{i} : \mathbf{v}_i(\mathbf{c}_i \mid \forall \mathbf{j} : \mathbf{c}_j) > \mathbf{v}_i(\mathbf{d}_i \mid \forall \mathbf{j} : \mathbf{d}_j)$$

A group faces a ***social dilemma*** when the following two properties hold (DAWES 1980):

- each group member is *worse off when (s)he cooperates* than when (s)he defects, irrespective of what the other group members do:

$$\forall \mathbf{i} : \mathbf{v}_i(\mathbf{c}_i \mid .) < \mathbf{v}_i(\mathbf{d}_i \mid .)$$

- each group members is *better off when everyone cooperates* than when everyone defects:

$$\forall \mathbf{i} : \mathbf{v}_i(\mathbf{c}_i \mid \forall \mathbf{j} : \mathbf{c}_j) > \mathbf{v}_i(\mathbf{d}_i \mid \forall \mathbf{j} : \mathbf{d}_j)$$

dilemma

How to *solve* the social dilemma?

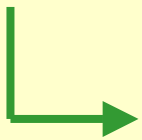
Suggestions in the literature:

- coordinate behaviour by obligatory rules,
- introduce punishments for defection,
- appeal to farsightedness.

How to *solve* the social dilemma?

Suggestions in the literature:

- coordinate behaviour by obligatory rules,
- introduce punishments for defection,
- appeal to farsightedness:



GAME THEORY.

How to *solve* the social dilemma?

Suggestions in the literature:

- coordinate behaviour by obligatory rules,
- introduce punishments for defection,
- appeal to farsightedness:

How to *solve* the social dilemma?

Suggestions in the literature:

- coordinate behaviour by obligatory rules,
- introduce punishments for defection:

→ **SANCTIONING SYSTEMS.**

How to *solve* the social dilemma?

Suggestions in the literature:

- coordinate behaviour by obligatory rules,
- introduce punishments for defection:

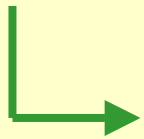
→ **SANCTIONING SYSTEMS.**

→ How are these provided?

How to *solve* the social dilemma?

Suggestions in the literature:

- coordinate behaviour by obligatory rules:

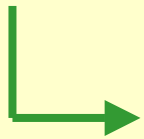


NORMS.

How to *solve* the social dilemma?

Suggestions in the literature:

- coordinate behaviour by *obligatory rules*:



NORMS.

How does obligation work?

Normative behaviour …

… is ***always stabilized*** by sanctions.

Absence of sanctions is a telltale sign that a behavioural rule is ***not*** normative.

… is ***internalized*** .

Actors ***want*** to do what they ***have*** to do.

Sanctions …

… are ***integral part*** of normative behaviour.

Absence of sanctions is a telltale sign that a behavioural rule is ***not*** normative.

… but ***do not*** (directly) ***influence*** behaviour.

Actors ***want*** to do what they ***have*** to do.

(***the "sociologists' dilemma"***)

LINDENBERG's theory of norms:

- when sanctions directly influence behaviour, the actor *"is in a gain frame."*
  (*foreground influence* of sanctions)
- when an actor "*is in a normative frame,*" sanctions only influence the strength of the norm, not its content.
  (*background influence* of sanctions)

LINDENBERG's framing theory

(***Discrimination model*** of framing)**:**

- weakness of a frame leads to random preference, and vice versa.
- weakness of a frame leads to a *frame switch*.

Experimental validation will centre around the dynamic manipulation of frame strength by variation of sanction sizes.

Some hypotheses:

- ***sensitivity to sanction size*** differs between frames:

  normative frame: lower sensitivity,

  gain frame: higher sensitivity.

Some hypotheses:

- ***sensitivity to sanctions*** differs between frames,
- ***attitude towards sanctions*** differs between frames:

  normative frame: positive attitude,

  gain frame: negative attitude.

Some hypotheses:

- ***sensitivity to sanctions*** differs between frames,
- ***attitude towards sanctions*** differs between frames,
- ***behavioural randomness*** depends on frame×sanction interaction:

normative frame: behavioural randomness occurs for low sanctions,

gain frame: behavioural randomness occurs for high sanctions.

Some hypotheses:

- ***sensitivity to sanctions*** differs between frames,
- ***attitude towards sanctions*** differs between frames,
- ***behavioural randomness*** depends on frame×sanction interaction,
- ***stability of frames over time***:

  Actors approach decision situations with the frame they applied in the previous situation.

Some hypotheses:

- ***sensitivity to sanctions*** differs between frames,
- ***attitude towards sanctions*** differs between frames,
- ***behavioural randomness*** depends on frame×sanction interaction,
- ***stability of frames over time***:
  - ***inertia of frames*** and behaviour,
  - ***hysteresis of*** frames and ***behaviour***.

Formal modelling:

- Assume that before making a decision in a social dilemma, actors adopt either a *normative* or a *gain* frame: $\mathbf{F \in \{f_{norm}, f_{gain}\}}$.

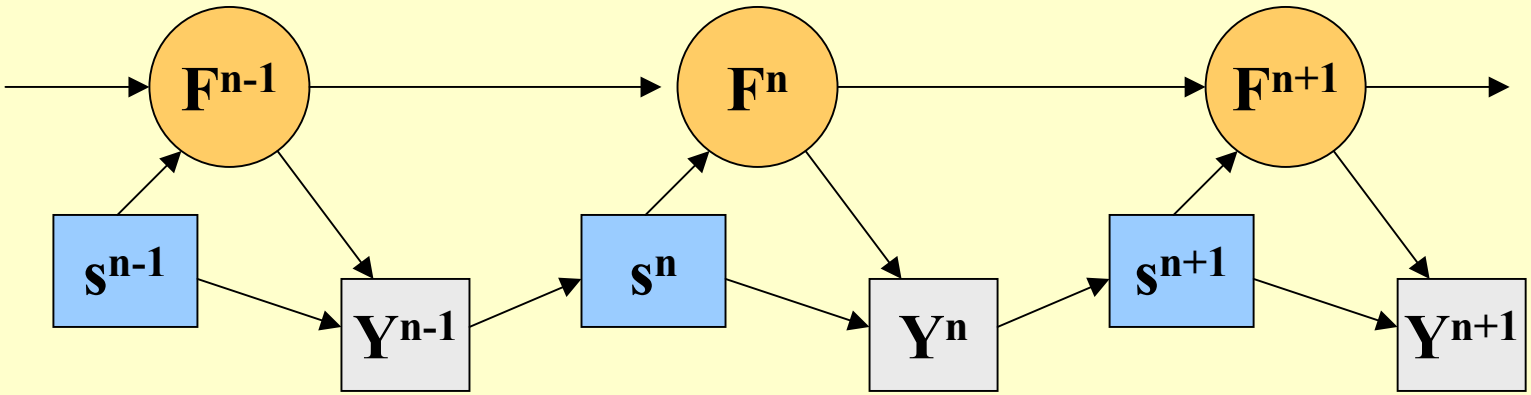
  This framing stage is influenced by situational parameters **s** and the previously used frame.

- Assume that then, actors base their behaviour on a frame-dependent decision rule:

  $$\mathbf{Y \sim \varphi (s \mid F)} .$$

Formal modelling:

The model can be summed up visually as follows:



The experimental study:

(January 2001, **124** students, computer experiment.)

Task: Protection of wild animals over **N=21** days,

- cooperation was tied to a reduction in collective housing costs,
- defection meant private gain (and was sanctioned by percentage **s**).

Experimental conditions:

- sanctioning pattern: ∨ versus ∧ ,
- semantic framing **a** (for *accessibility* manipulation): *environmentalist group* versus *leisure time brokers*.

Dependent variables:

- sanctioning attitude **x** (adequate sanctions in %),
- contribution **y** to common task (in hours out of **10h**).

Analytical framework:

- initial frame probabilities: $\boldsymbol{logit}[\Pr(\boldsymbol{F}^0 = \boldsymbol{f}_{norm})] = \alpha^0 + \alpha^1 \boldsymbol{a}$

- rules for frame updating:

$$\boldsymbol{logit}[\Pr(\boldsymbol{F}^n = \boldsymbol{f}_m \mid \boldsymbol{F}^{n-1} = \boldsymbol{f}_m)] = \beta_m^0 + \beta_m^1 \boldsymbol{a} + \beta_m^2 \boldsymbol{s} + \beta_m^3 \boldsymbol{y}^{n-1} + \beta_m^4 \boldsymbol{n}$$

- rules for frame-dependent behaviour:

  Y ~ beta(p,q) with **p** mean contribution:

$$\boldsymbol{logit}(\boldsymbol{p}_f) = \pi_f^o + \pi_f^1 \boldsymbol{a} + \pi_f^2 \boldsymbol{s}$$

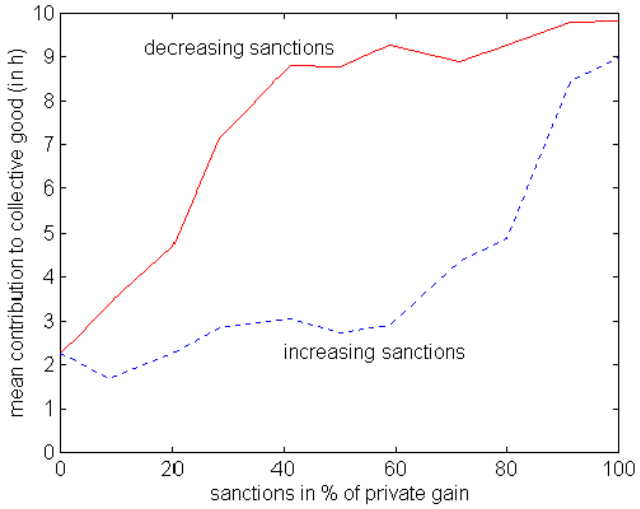
  and **q** corrected variance:

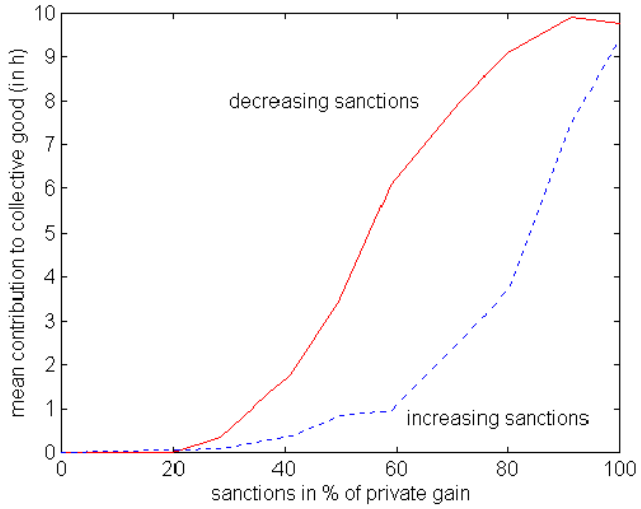
$$\boldsymbol{logit}(\boldsymbol{q}_f) = \kappa_f^o + \kappa_f^1 \boldsymbol{a} + \kappa_f^2 \boldsymbol{s}$$

Descriptive results:

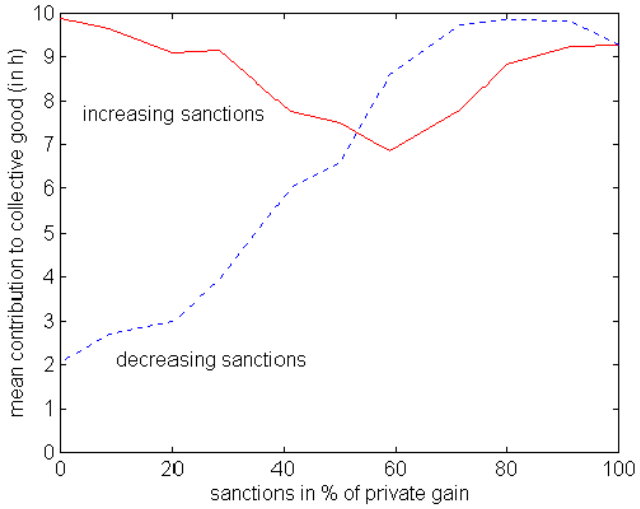
normative semantics | **gain semantics**

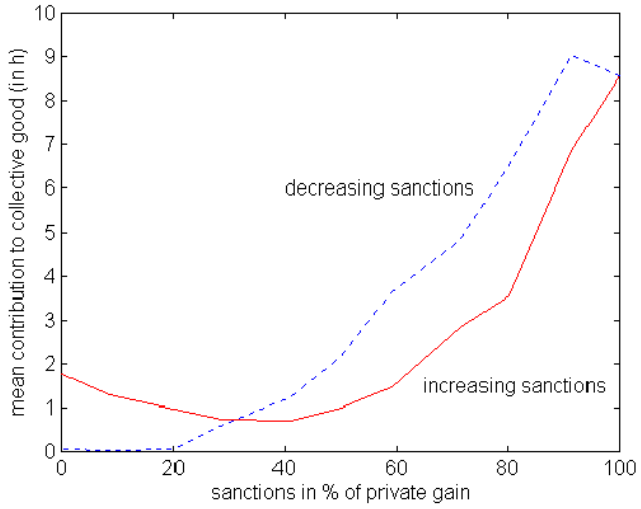
pattern \/





pattern /

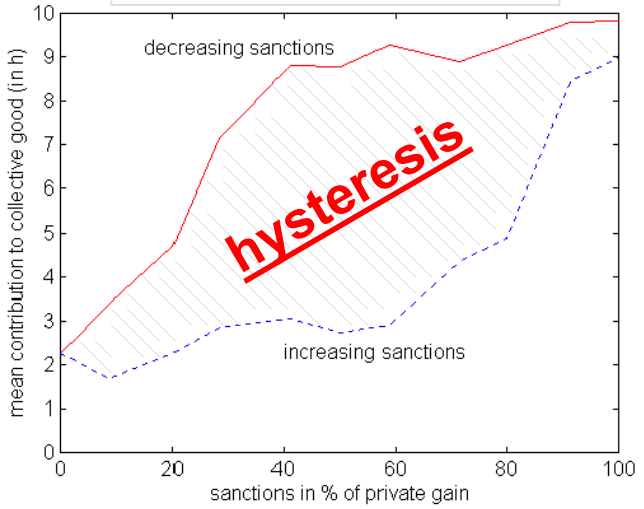


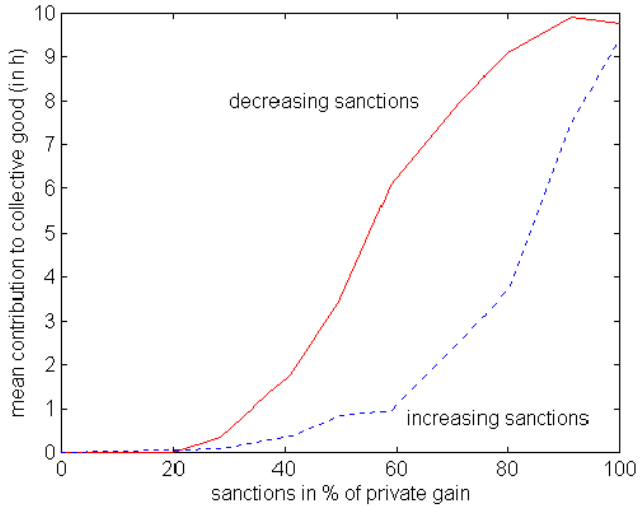


Descriptive results: *hysteresis hypothesis* confirmed.

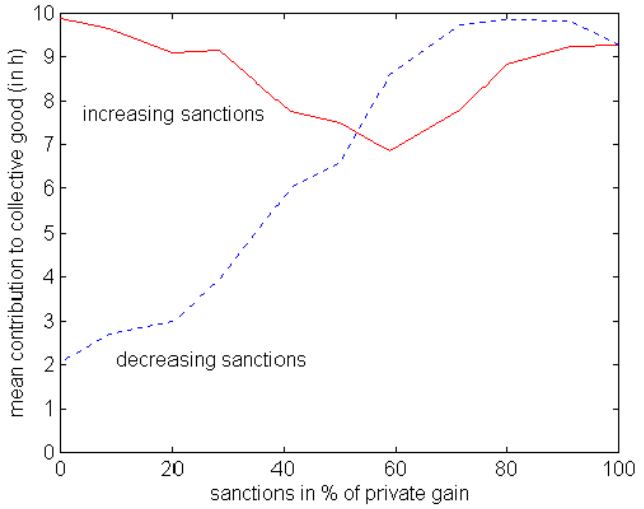
normative semantics | **gain semantics**

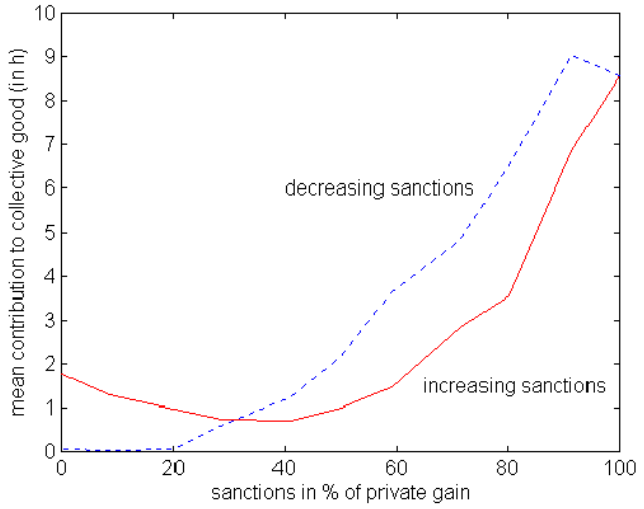
pattern \/





pattern /

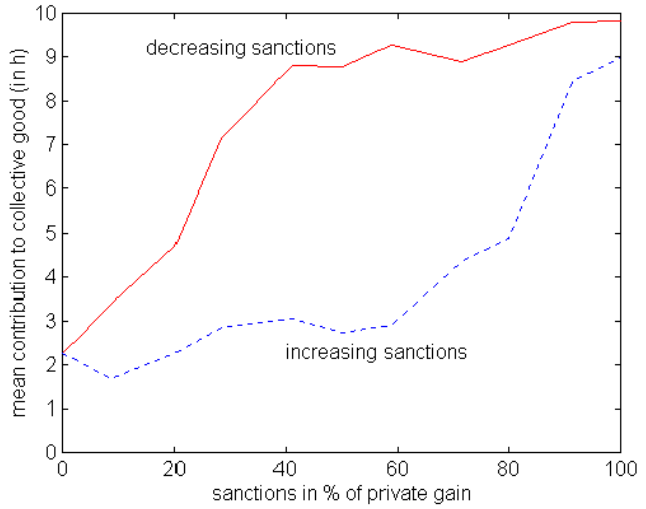


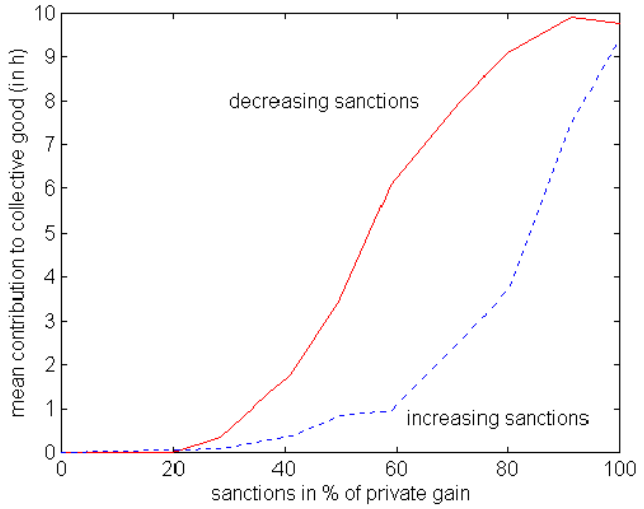


Descriptive results: *semantic framing* successful.

normative semantics | **gain semantics**

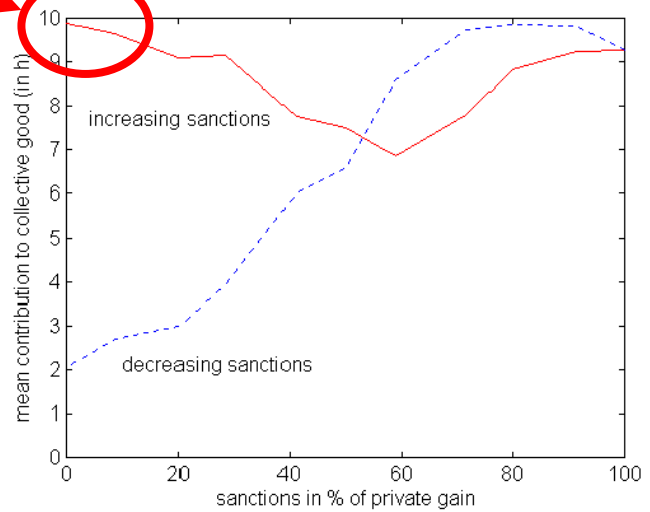
pattern \/

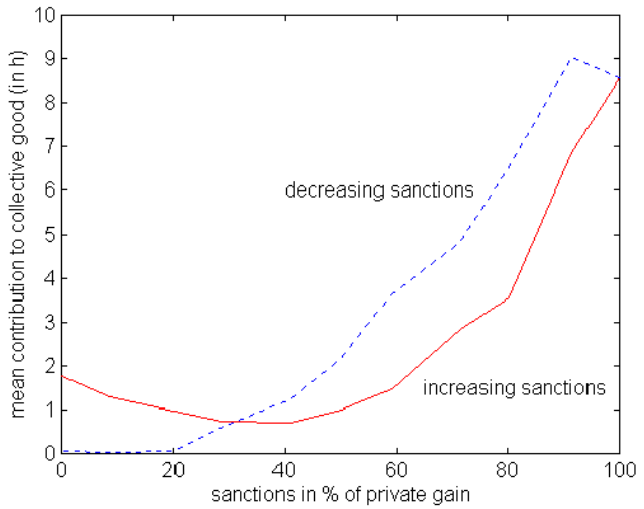




semantics

pattern /





Model estimates:

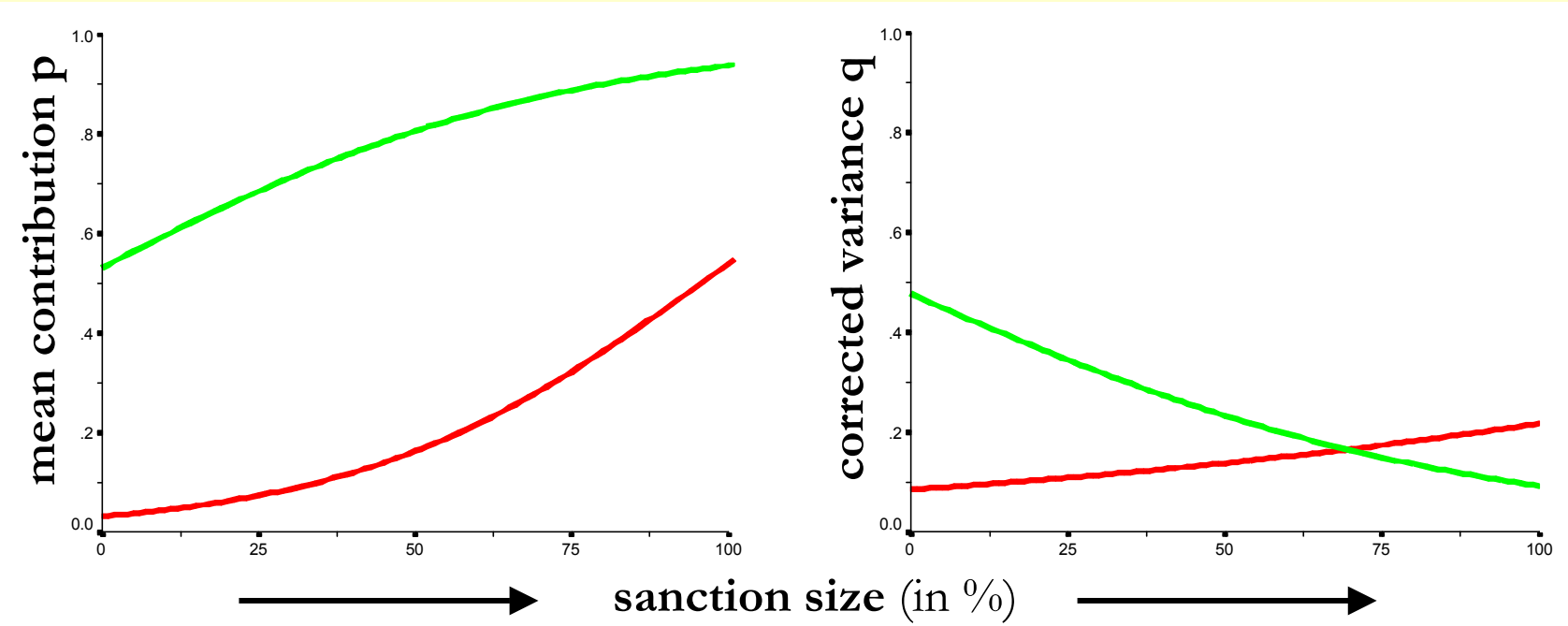
gain frame

normative frame

Behavioural rules per frame

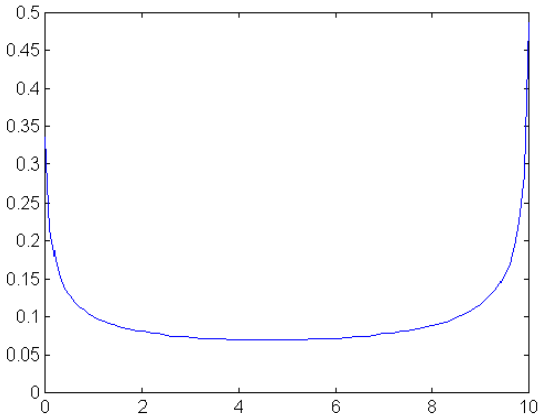
sensitivity hypothesis confirmed

behavioural variation hypothesis confirmed

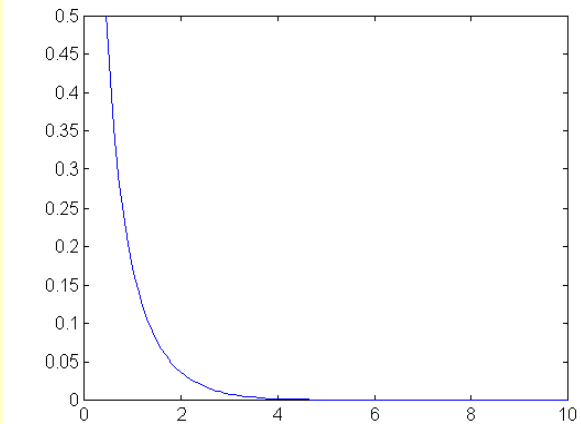


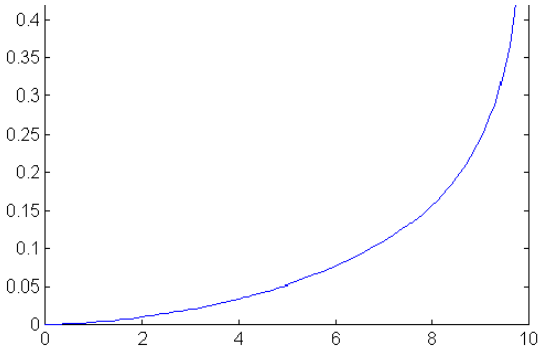
Model estimates: Distribution of contributions per frame

normative frame

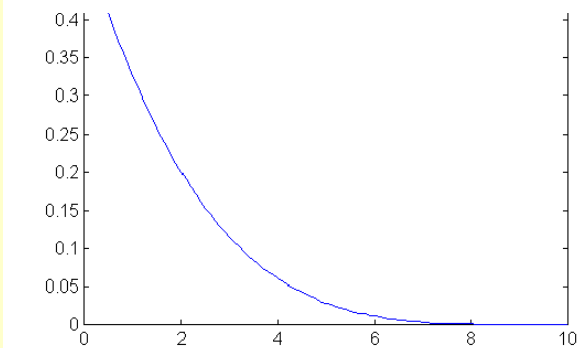


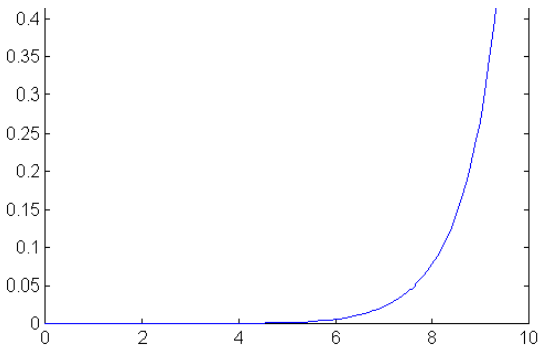
no sanctions



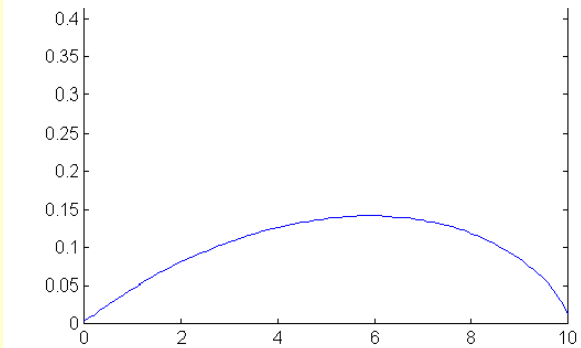


50 % sanctions





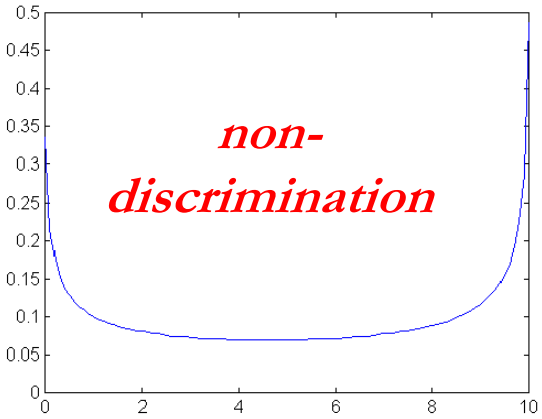
100 % sanctions



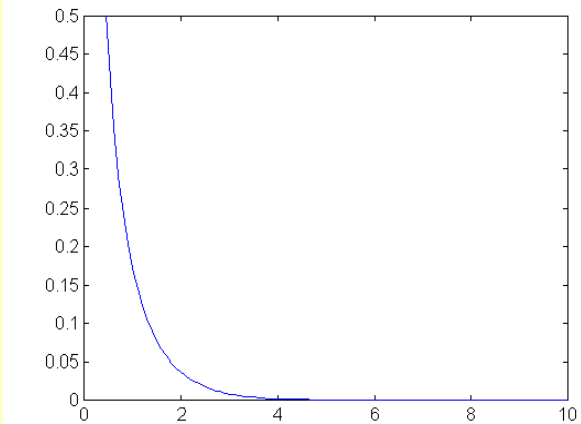
gain frame

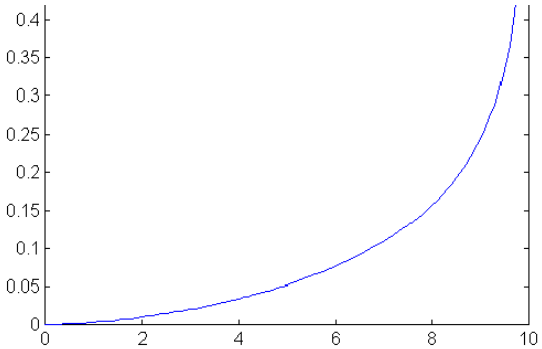
Model estimates: once more *behavioural variation*

normative frame

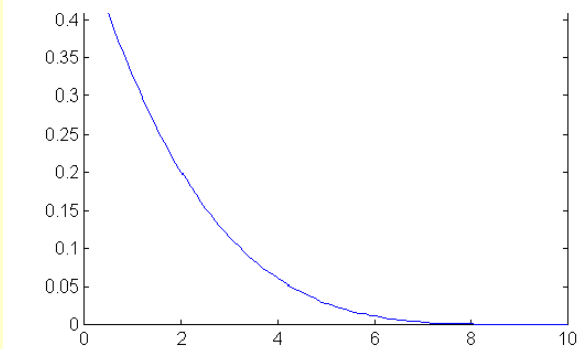


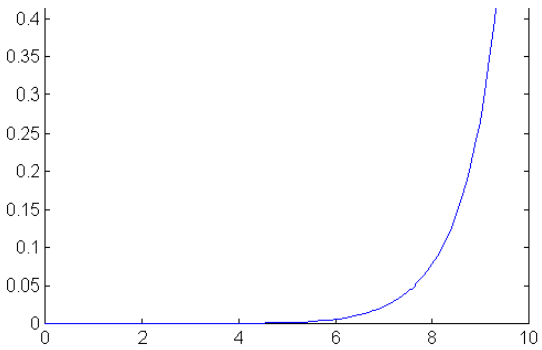
no sanctions



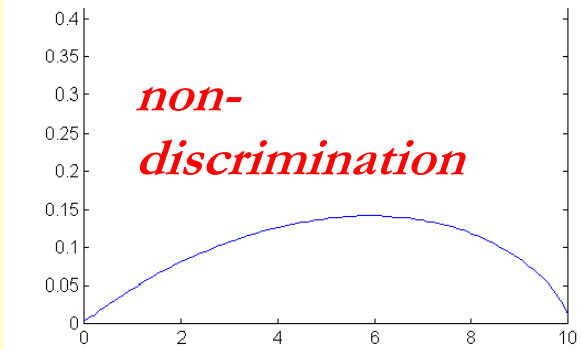


50 % sanctions





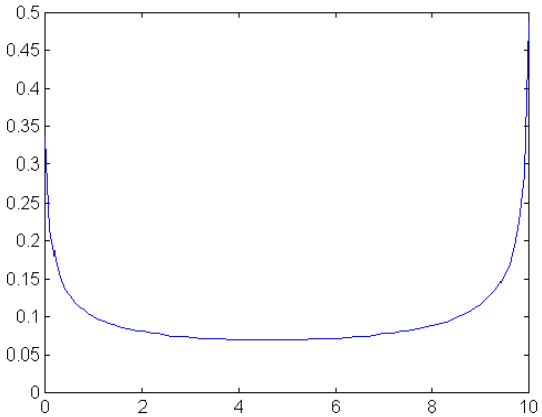
100 % sanctions



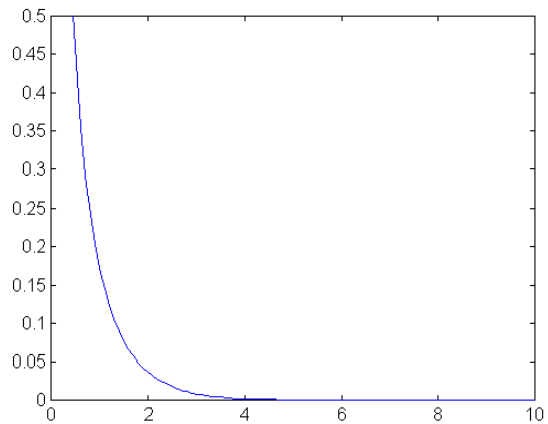
gain frame

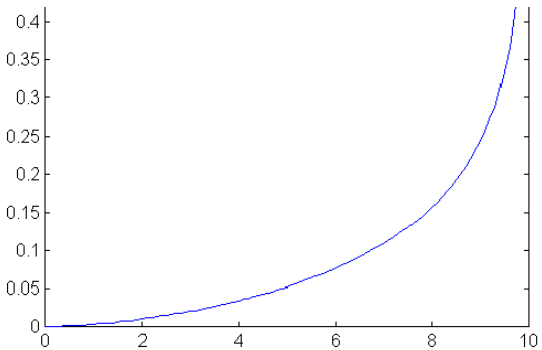
Model estimates: "rationality" of gain frame's rule ?

normative frame

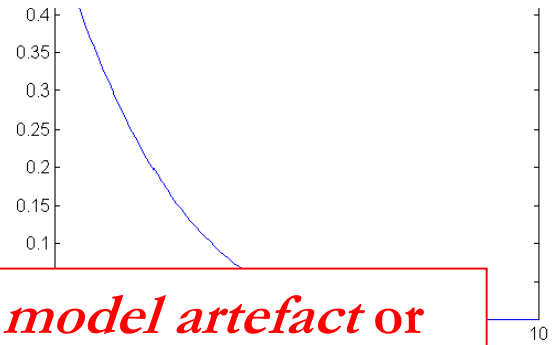


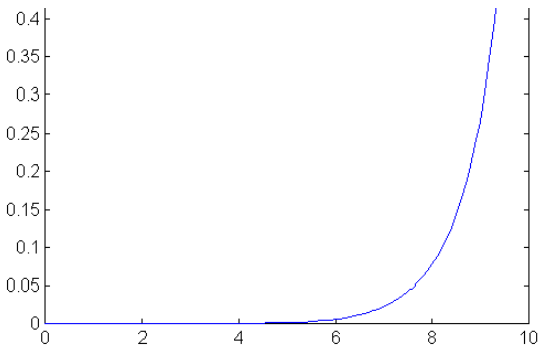
no sanctions



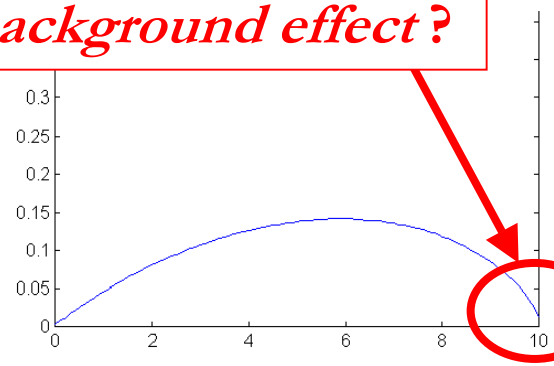


50 % sanctions





100 % sanctions



gain frame

¿ *model artefact* or *background effect* ?

Model estimates:

Frame updating:

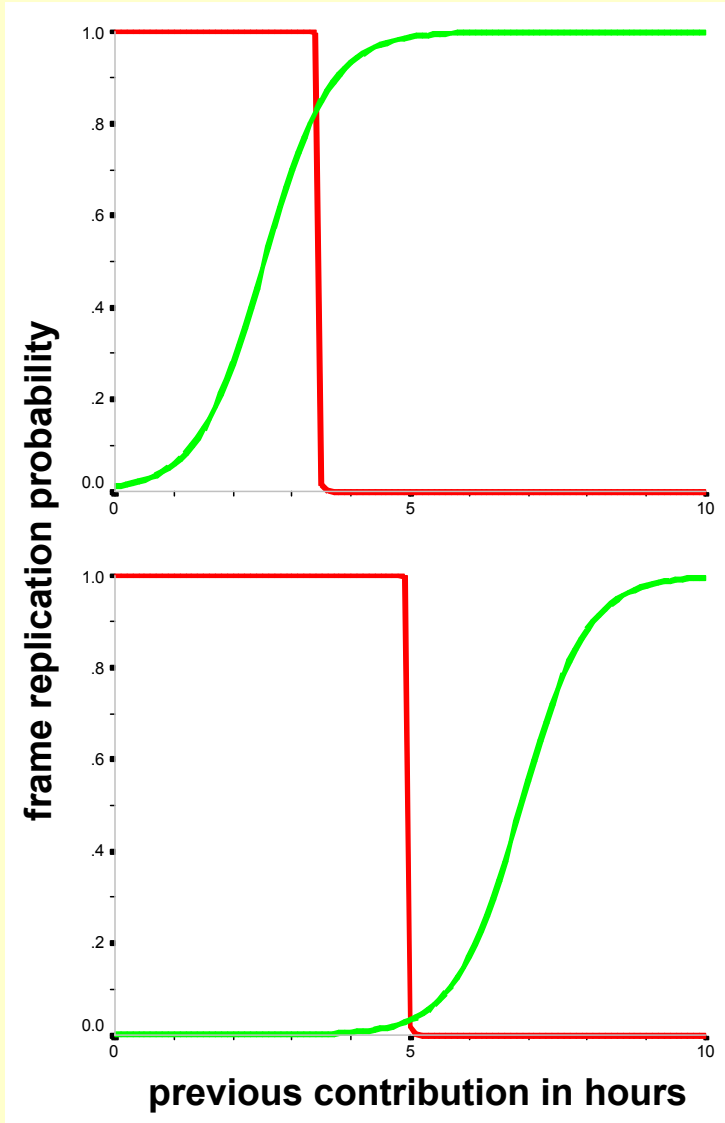
regions of frame stability

gain frame

normative frame

The ***inertia hypothesis*** is partly confirmed by *threshold shape* :

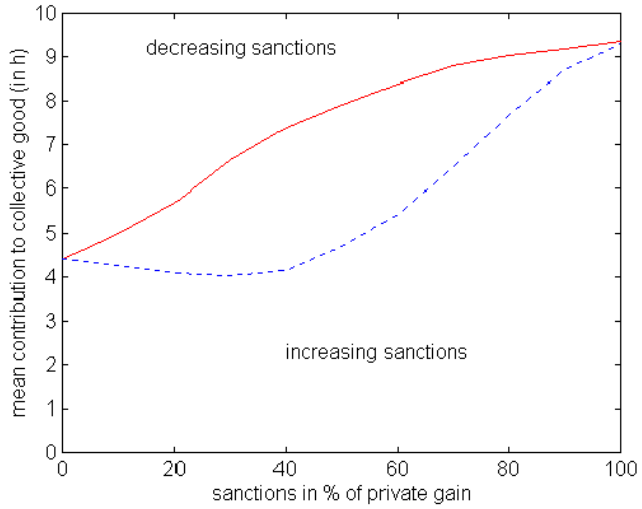
Frames are stable in the region of compatible behaviour.

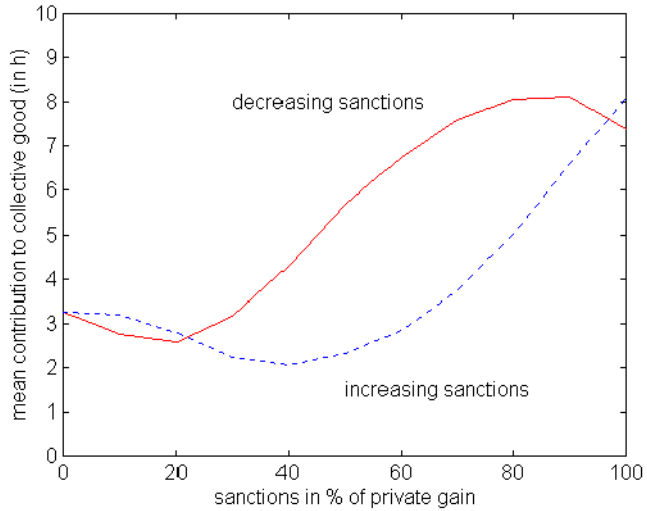


Model-derived simulations: *goodness of fit* visualised.

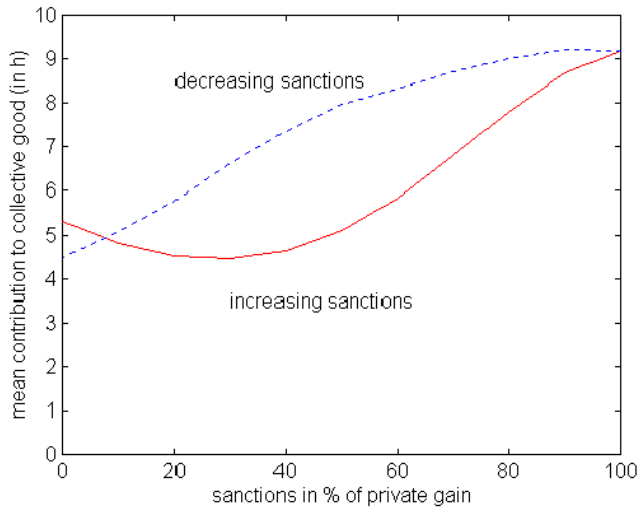
normative semantics | **gain semantics**

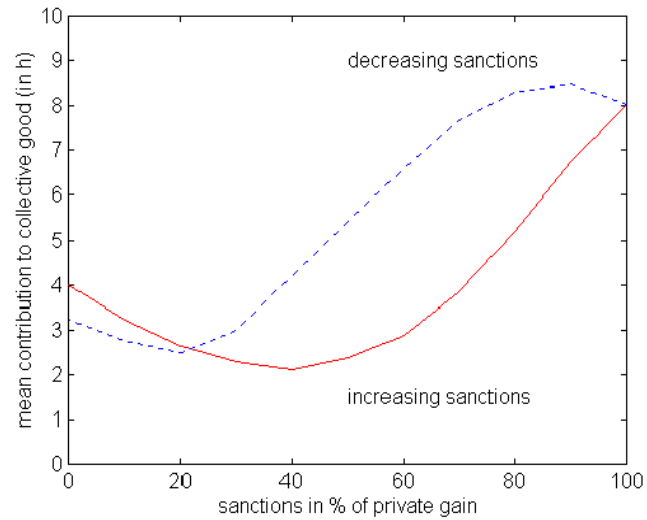
pattern \/





pattern /

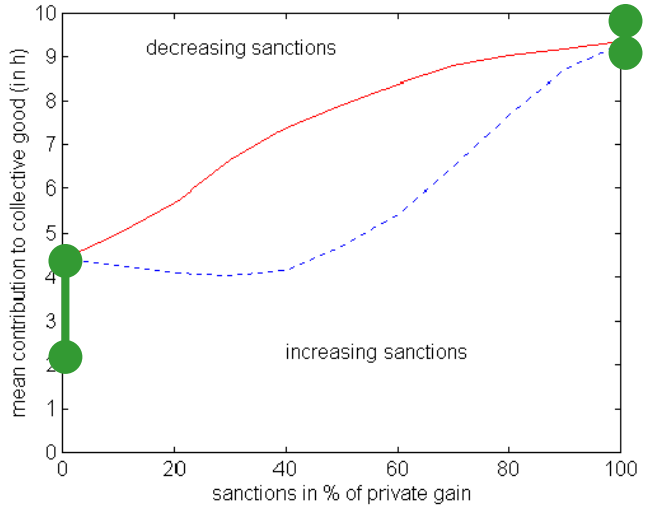


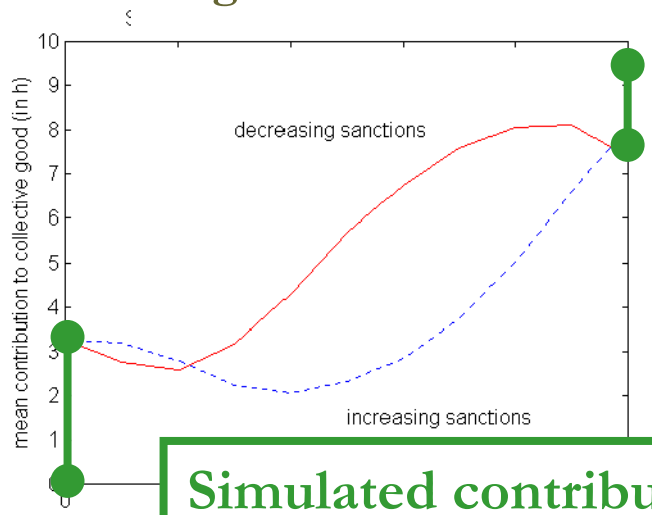


Model-derived simulations: fit problems

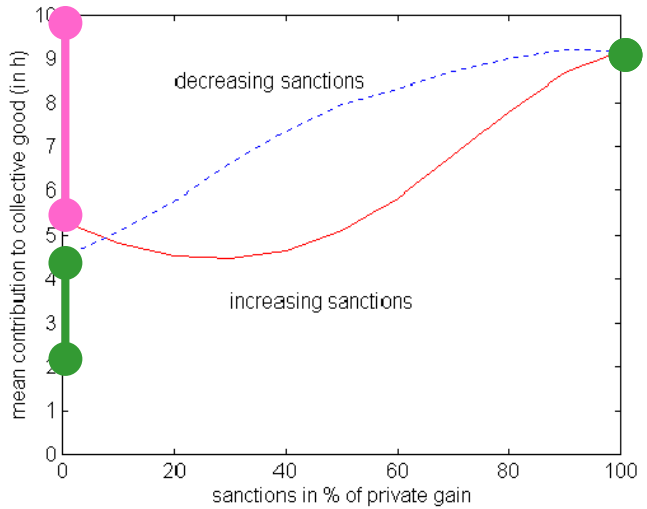
normative semantics | **gain semantics**

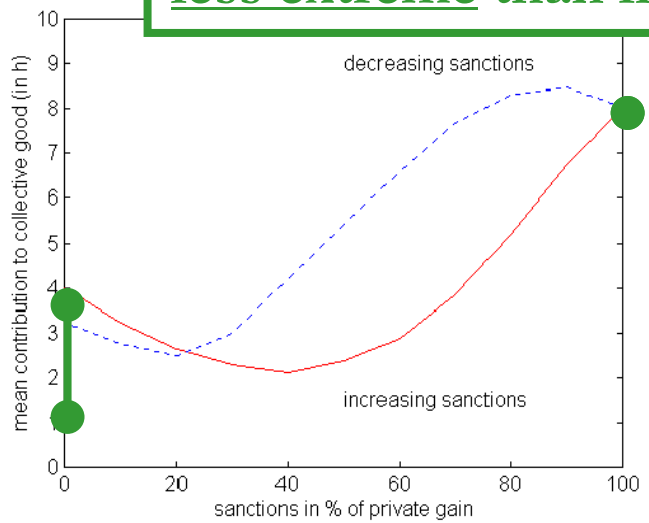
pattern \/





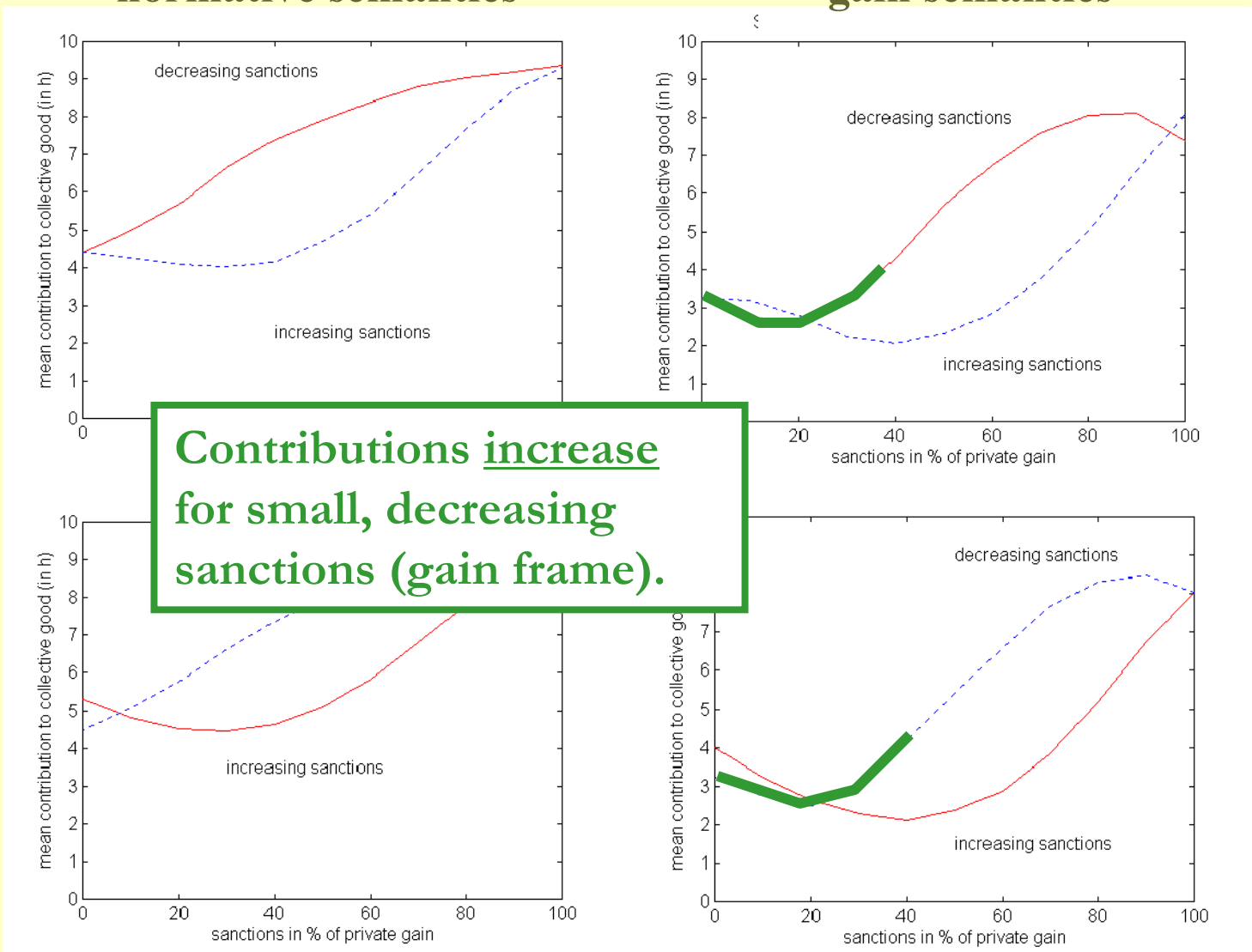
pattern /





Simulated contributions less extreme than in data.

Model-derived simulations: fit problems



External validity of the model:

The estimates are solely based on actors' *behaviour* **Y**.

Model-derived frames can now be compared to the other dependent variable *sanction attitude* **X**:

estimated frame	sanction attitude: positive	sanction attitude: negative
normative	1129	183
gain	198	1094

The ***sanction attitude*** hypothesis is confirmed.

Conclusions:

- ***Framing theory*** gives a valid account of behaviour.
- The ***theory of normative behaviour*** is confirmed.
- The ***model-fitting procedure*** does a good job

(*but suffers from rigid specifications*).