SSPoS @ Lorentz Center 2014

Mechanisms for science: Leasons learned from modeling peer review

Francisco Grimaldo, Juan Bautista Cabotà (U.València)
Mario Paolucci (LABSS- ISTC- CNR)
Flaminio Squazzoni (GECS-U. Brescia)

Peer review (PR) process

- PR is a cornerstone of science as it ultimately determines how the resources of the science system are allocated.
- Scrutinizes scientific contributions before they are made available to the community.
- Used in conferences, journals, granting agencies for project evaluations...
- As any social process, it should be evaluated with respect to a series of parameters [LiquidPub project]:
  - Efficiency, effectiveness, fairness, fraud detection, innovation promotion...

State of the art

- Diffuse dissatisfaction of scientists towards the current mechanisms of peer review:
  - Famous papers initially rejected.
  - PR failures due to judgement bias and misconduct.
- Previous studies have found that:
  - PR includes a strong “lottery” component, independent of editor and referee integrity [Neff & Olden, 2006].
  - Numerical evidence on the failures of PR [Casati et al., 2011].
  - “Rational” scientist can corrupt the PR mechanism under certain circumstances [Thurner & Hanel, 2010].
  - Rather low level of agreement [Bornmann, 2014].

Aim of this research

- To create a model (better, a plurality of models) of peer review that takes into account recent theoretical developments in recommender systems and reputation theories and test the proposed innovations.

- Today I will ...
  - ... draw an overview of how we foresee such models.
  - ... present alternative implementations of them.

Two opposing forces

- Simplify to the extreme:
  - Swarm intelligence.
  - Complex systems, economics.
  - Synchronous, interleaved, stepwise.
  - Failures by oversimplification.
- Make it as complex as you can:
  - Cognitive intelligence.
  - Philosophy, sociology, psychology, engineering
  - Asynchronous, concurrent, behevioral-driven.
  - Failures by overfitting and validation.

A pragmatic standpoint

- Focus on mechanisms, thus dealing with:
  - Processes and algorithms.
  - Parameters and distributions.

- GECS - U. Brescia:
  - *A PR model with no name*.
- LABSS-ISTC-CNR:
  - PR-M.

An ABM in NetLogo

- Starting point:
  - *Opening the Black-Box of Peer Review: An Agent-Based Model of Scientist Behaviour*. F. Squazzoni, C. Gandelli. JASSS. 2013.
  - Effect of reciprocity on the quality of PR.
- Extensions:
  - Reviewing behaviors: random, fair, unreliable and strategic (local competition vs. glass ceiling).
  - Author-referee matching policies: random, peer, higher-skilled and lower-skilled.

Overview of the model

- Entities and state variables:
  - Scientists (resources & behavior).
  - Editorial policy (publication rate & matching).
- Process overview and scheduling:
  - Noisy production and evaluation.
  - Resource accumulation and expenses.
  - Evaluation bias, productivity loss & Gini index.
- Submodels:
  - +16 extracted by combining reviewing behaviors with editorial matching policies.

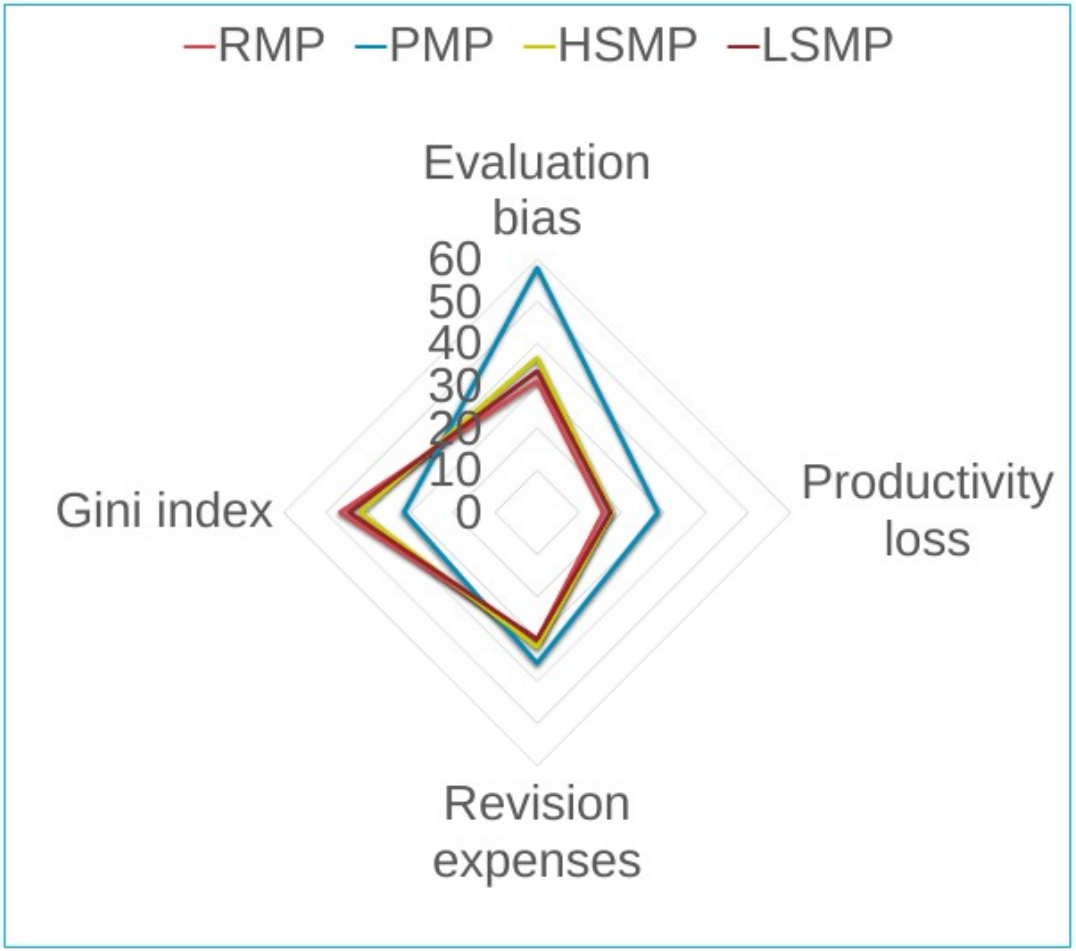
Two research questions

- What is the impact of strategic behaviors by referees on the quality and efficiency of PR?
- Which are the effects of different editorial policies to match referees and authors based on their academic status?

Effects of cheating behaviours

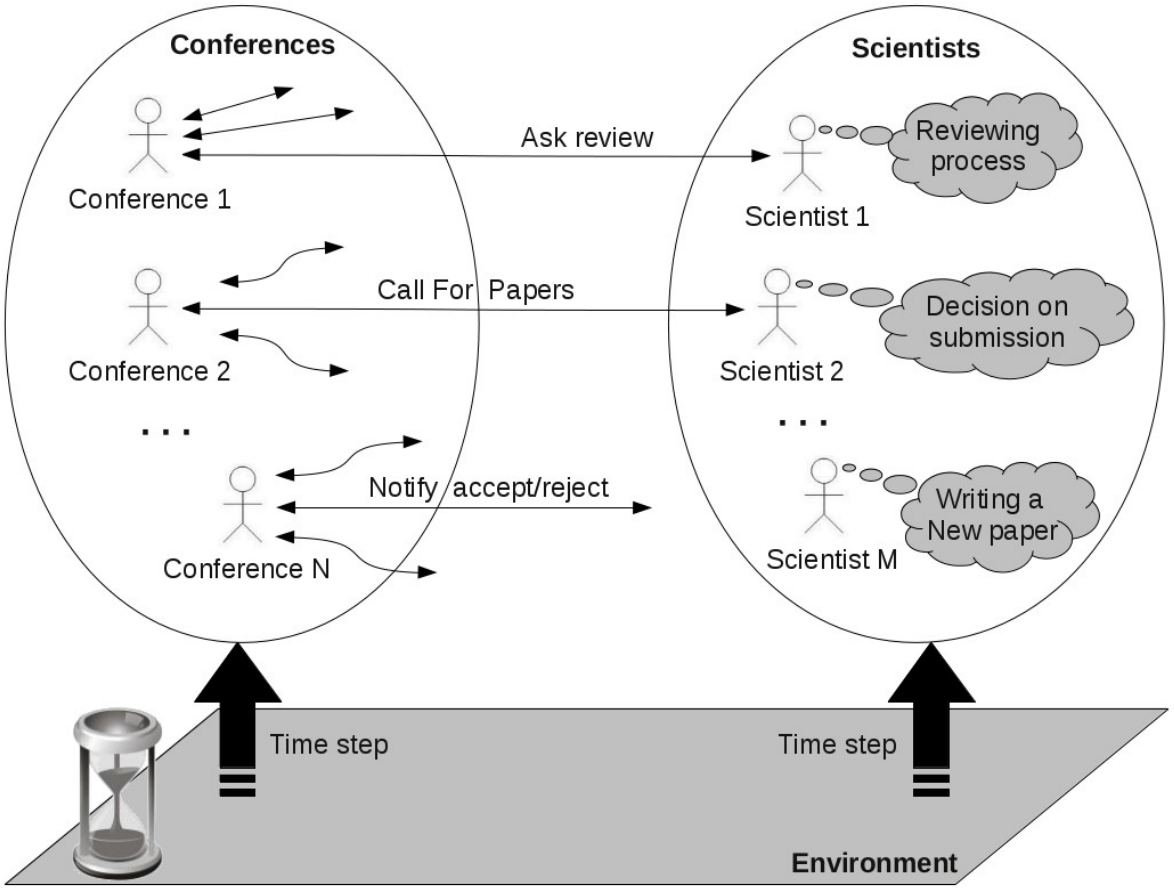
Scenario	Evaluation bias	Productivity loss	Reviewing expenses
Weak selection (75% published submissions)			
Random behaviour	16.51 %	7.68 %	25.98 %
Cheating	20.07 %	4,91 %	21.34 %
Medium-level selection (50% published submissions)			
Random behaviour	25.27 %	14.98 %	30.77 %
Cheating	56.63 %	28.02 %	32.21 %
Strong selection (25% published submissions)			
Random behaviour	29.42 %	15.00 %	29.42 %
Cheating	70.86 %	34.72 %	35.24 %

Effect of the matching policy



- Do editors have a silver bullet? An agent-based model of peer review. J.B. Cabotà, F. Grimaldo, F. Squazzoni. ECMS. 2014.

Overview of the MAS in Jason



- Paper intrinsic values are integers in a N-values ordered scale, ranging from strong reject to enthusiastic accept.

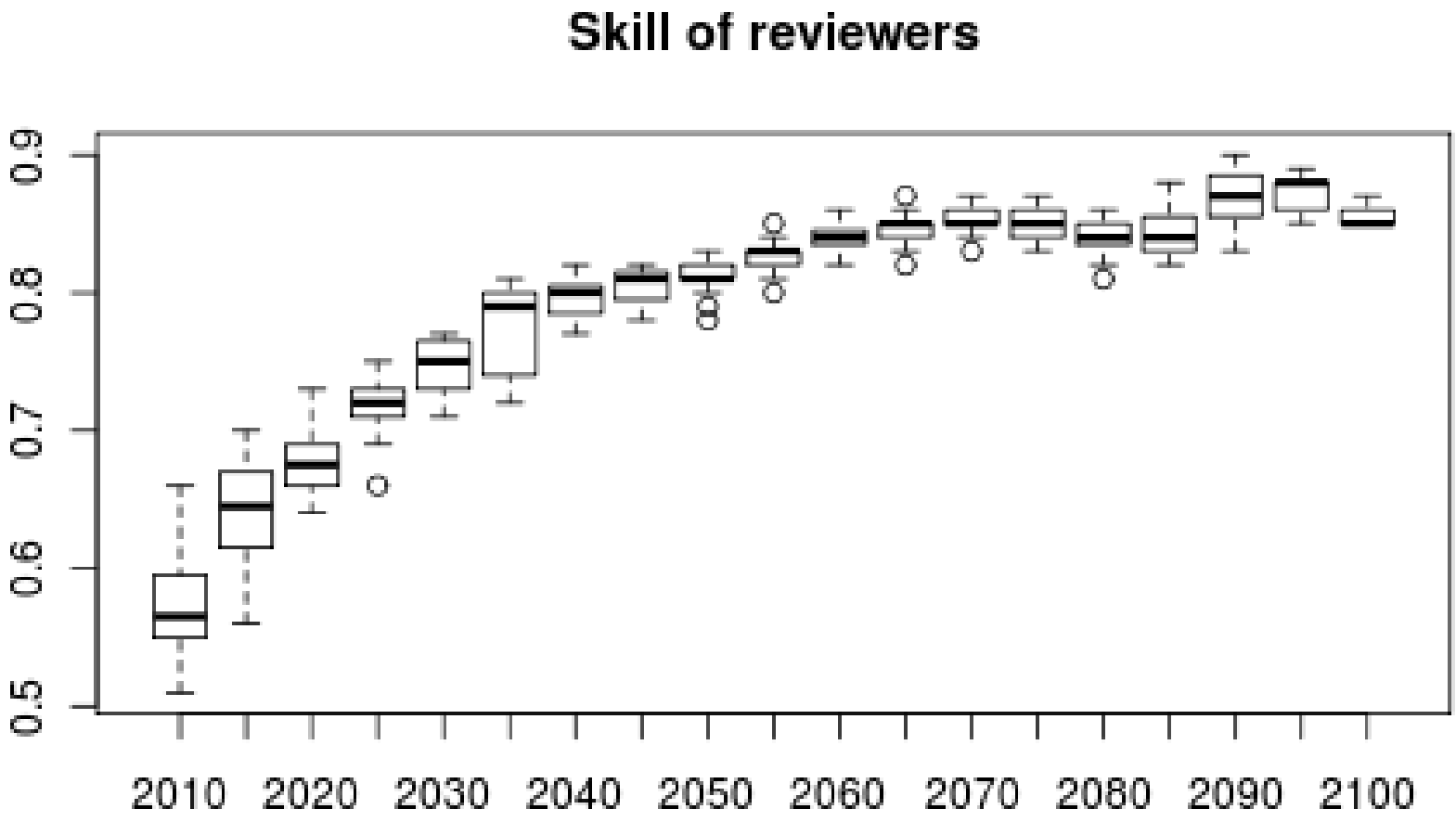
Entities

- Papers:
  - Object level: Any item subject to evaluation (e.g. papers, project proposals...).
  - Object value is noisily perceived.
- Scientists:
  - Writing and reviewing skills and decisions.
  - Strategic behaviours (e.g. rational cheating to elliminate competitors).
- Conferences:
  - Acceptance rate and policy (e.g. unanimity).
  - PC selection based on disagreement.

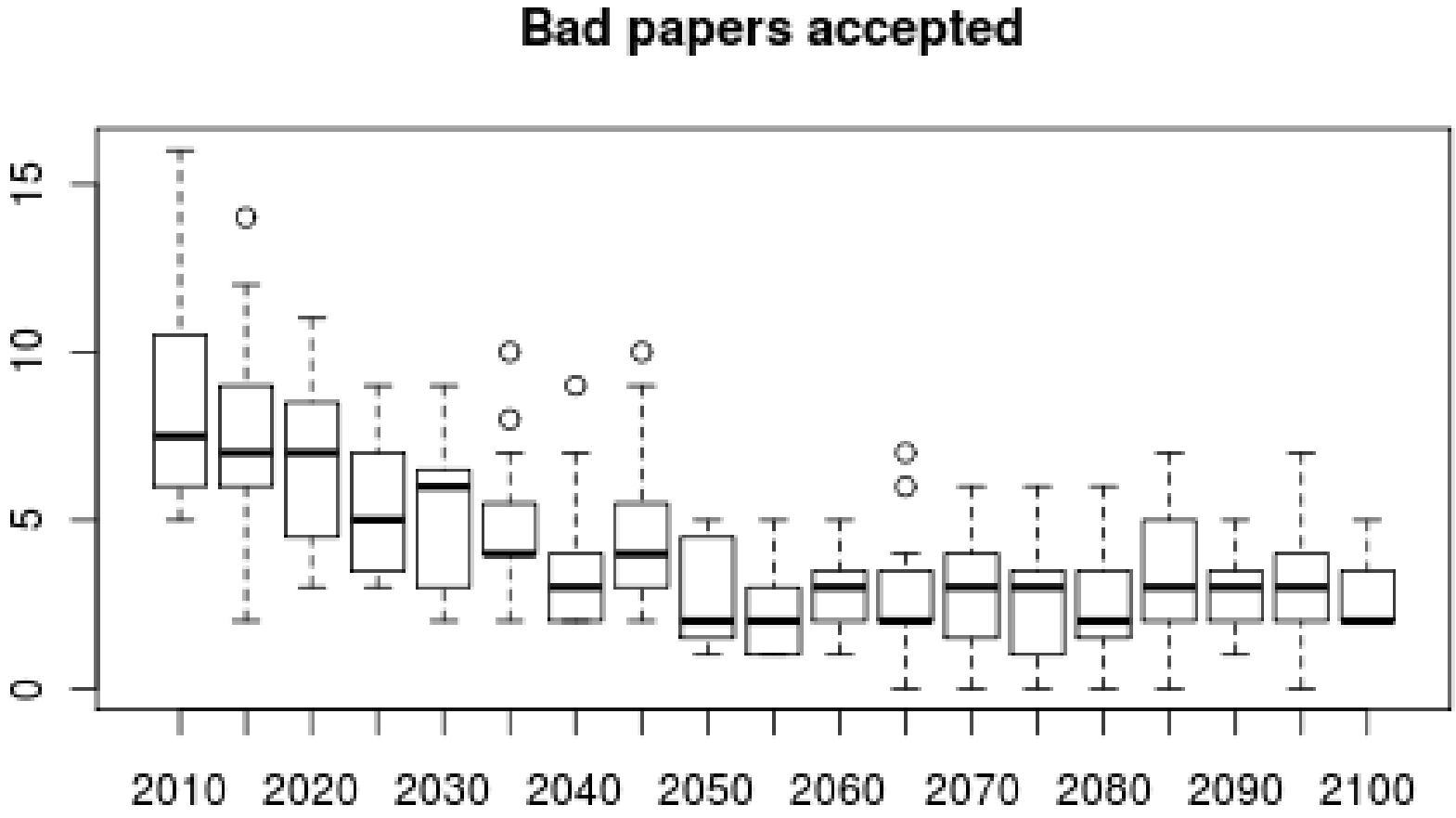
Two research questions

- Can the PR system ensure quality in the face of variable reviewing skills or strategic behaviors, thanks to some selection process of the PC composition that leans on disagreement control?
- Is the rational strategy really detrimental? In which sense and under which circumstances?

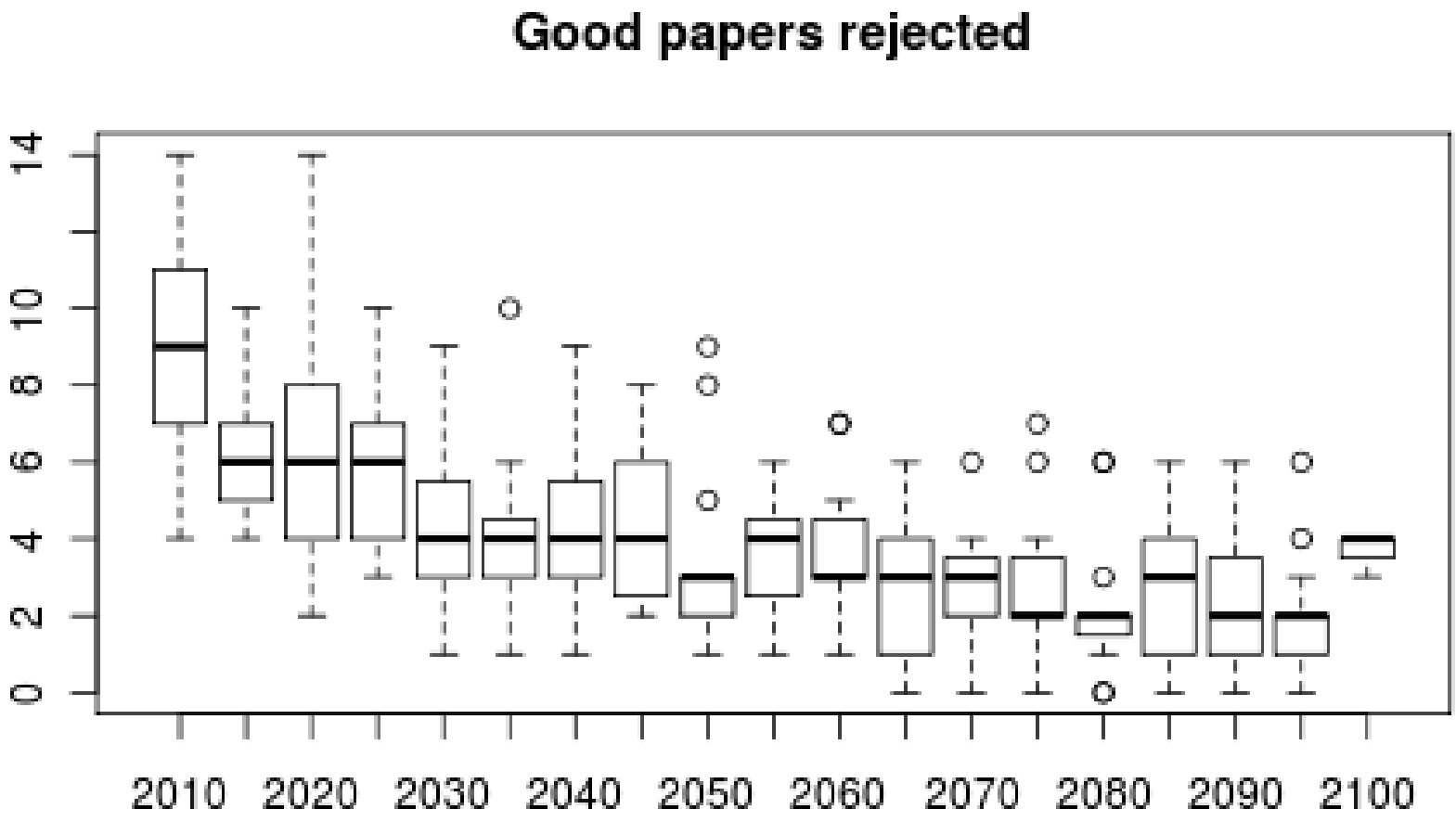
Efficiency


Skill of reviewers
0.5
0.6
0.7
0.8
0.9
2010
2020
2030
2040
2050
2060
2070
2080
2090
2100

Fairness (Type I errors)



Fairness (Type II errors)

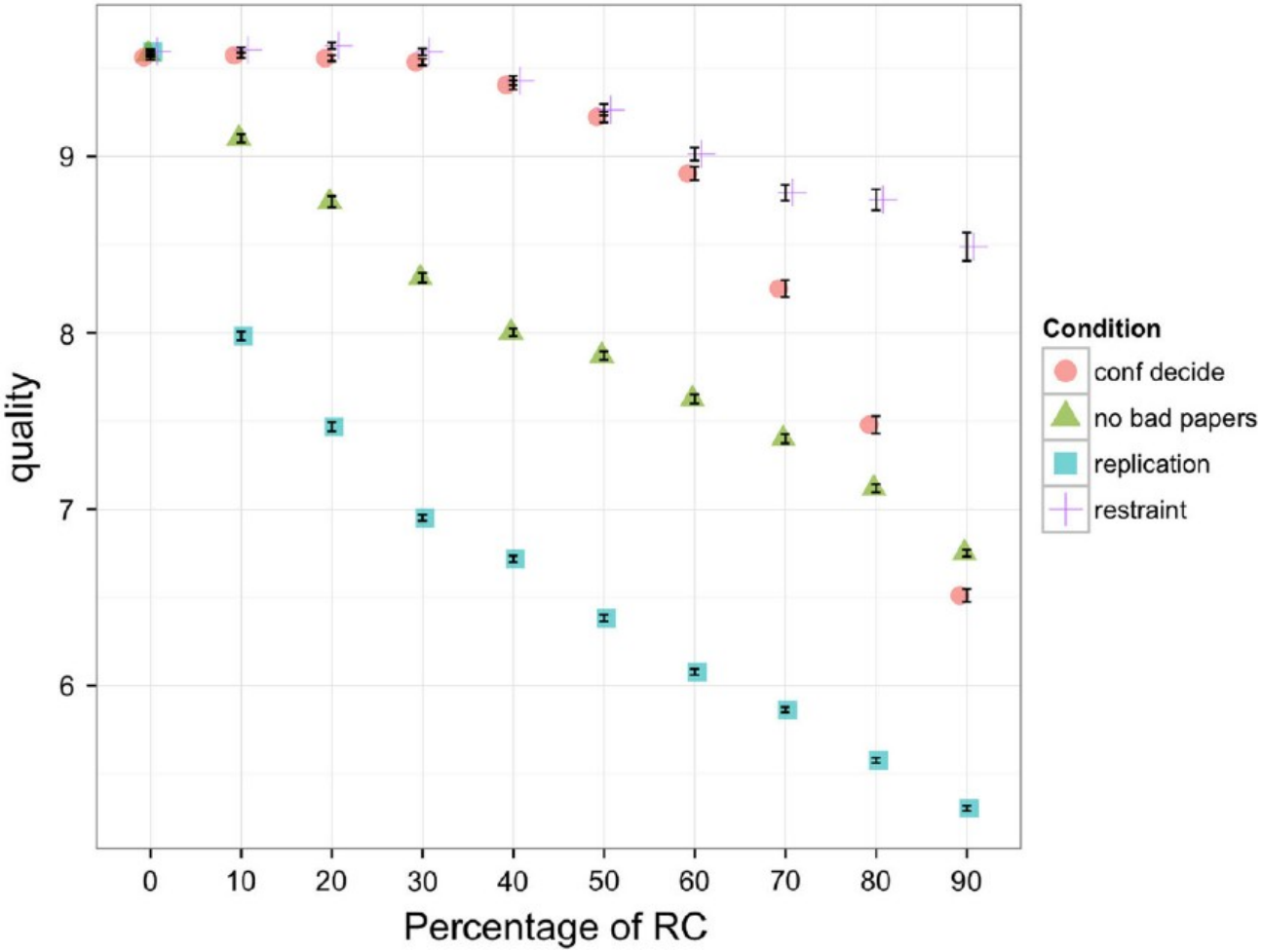


Effectiveness

Conference	% Initial disagreements	% Final disagreements	% Disagreement reduction
National	18.25	-	-
Summer School	10.71	-	-
International	5.41	-	-
Intl. Core C	5.0	-	-
Intl. Core B	0.0	-	-
Hom-0%RC	4.3	2.9	32.6
Hom-10%RC	6.1	4.5	26.2
Hom-30%RC	11.9	5.6	52.9
Het-0%RC-LQ	4.7	3.6	23.4
Het-0%RC-MQ	3.4	1.7	50.0
Het-0%RC-HQ	4.2	3.8	9.5
Het-10%RC-LQ	9.4	4.2	55.3
Het-10%RC-MQ	8.6	5.5	36.1
Het-10%RC-HQ	5.2	2.4	53.9
Het-30%RC-LQ	46.0	11.8	74.4
Het-30%RC-MQ	16.0	6.1	61.9
Het-30%RC-HQ	3.9	2.8	28.2

- A simulation of disagreement for control of rational cheating in peer review. F. Grimaldo, M. Paolucci. Advances in Complex Systems. 2013.

Effect of rational cheaters



- Mechanism change in a simulation of peer review: from junk support to elitism. M. Paolucci, F. Grimaldo. Scientometrics. 2014.

Some conclusions

- PR outcomes are sensitive to how scientists identify their competitors (e.g. local competition reduces negative effects)
- Editorial counteractions to reduce the impact of referee misbehavior
  - Avoid peer matchings under local competition.
  - Select referees considering disagreements.
- PR and strategic behavior show a complex interaction:
  - It can cause a quality collapse or even a slight quality increase depending on the mechanisms.

Ongoing and future work

- Ground model assumptions:
  - Game theoretical description and analysis.
  - Calibration from experiments.
- Adding networks:
  - Co-author, citation and behavioral networks.
  - Network dynamics.
- Data analysis and validation...

PEERE: New Frontiers of Peer Review

European Cooperation in Science and Technology

- Improve efficiency, transparency and accountability of PR
- Kick-off meeting: May 12th, 2014.
- Working groups:
  - Theory, analysis and models of PR.
  - Data sharing and testing:
    - Elsevier & Springer on board.
  - Research and implementation agenda.
- http://www.cost.eu/domains_actions/TDP/Actions/TD1306

Thank you!

Mechanisms for science: Leasons learned from modeling peer review

Francisco Grimaldo, Juan Bautista Cabotà (U.València)
Mario Paolucci (LABSS- ISTC- CNR)
Flaminio Squazzoni (GECS-U. Brescia)

Some more details…

JaCaMo system overview