

# The Italian Research Assessment Program

Gianluigi Greco

# Case study: Italian Research Assessment Program

- VQR: ANVUR should evaluate the quality of research of all Italian research structures
- Funds for the structures in the next years depend on the outcome of this evaluation
- Substructures will be also evaluated (departments)

**(1) 2004-2010**

**(2) 2011-2014**

# ANVUR Evaluation



ANVUR Criteria



# ANVUR Evaluation



ANVUR Criteria



Self-evaluations



$score_r(p)$ , for each  $\begin{cases} r \in \mathcal{R} \\ p \in products(r) \end{cases}$

# ANVUR Evaluation



ANVUR Criteria



Self-evaluations



$score_r(p)$



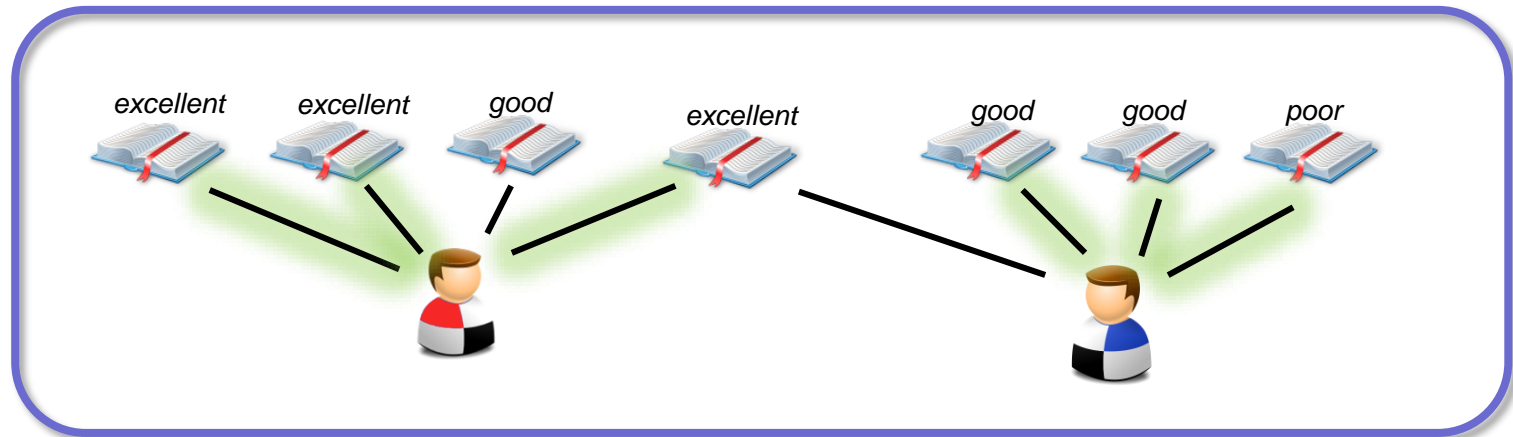
Structures are in charge of selecting the products to submit

# Constraints

- Every researcher has to submit 3 publications
- A publication cannot be allocated to two researchers

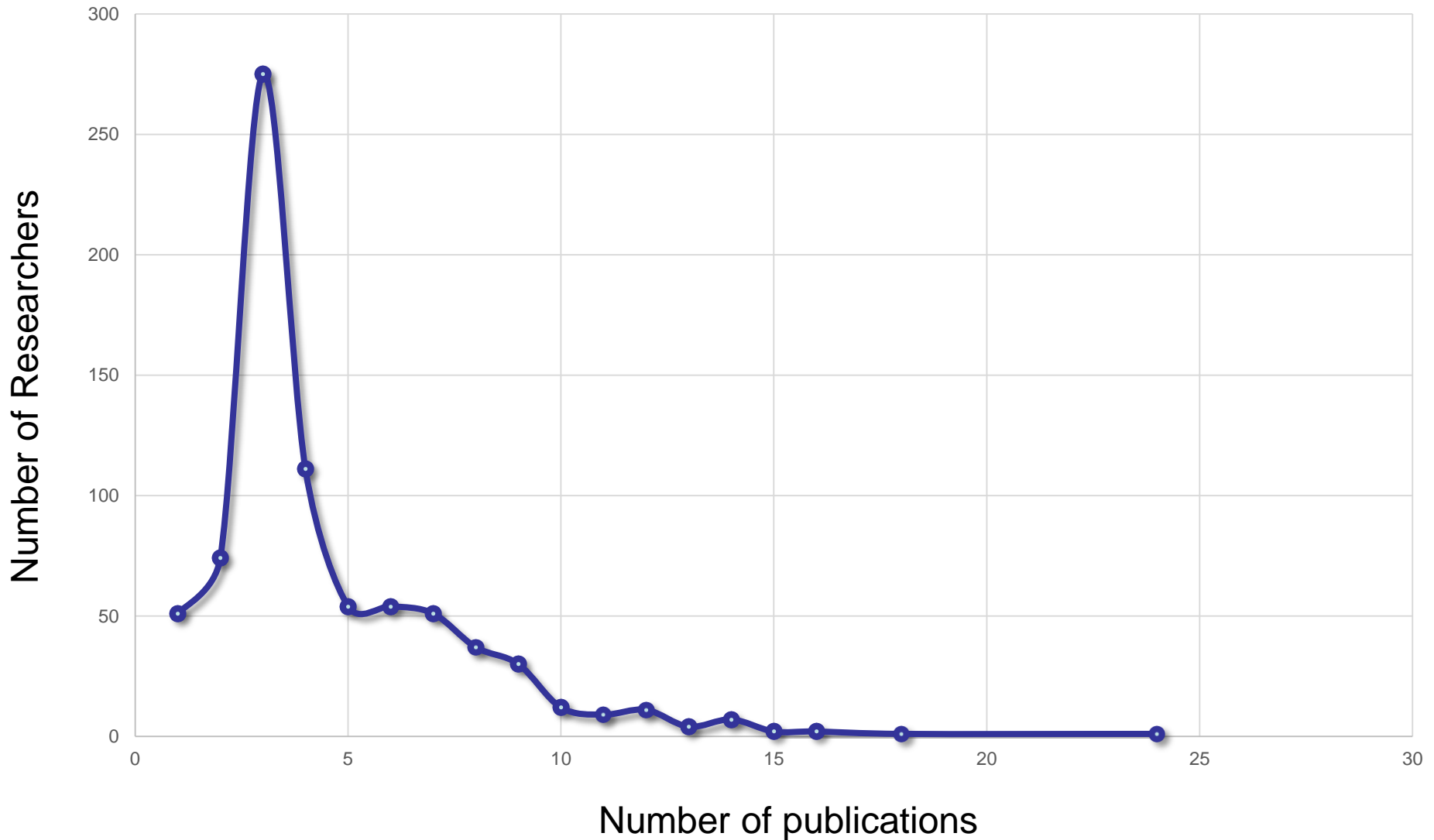


## Allocation Problem



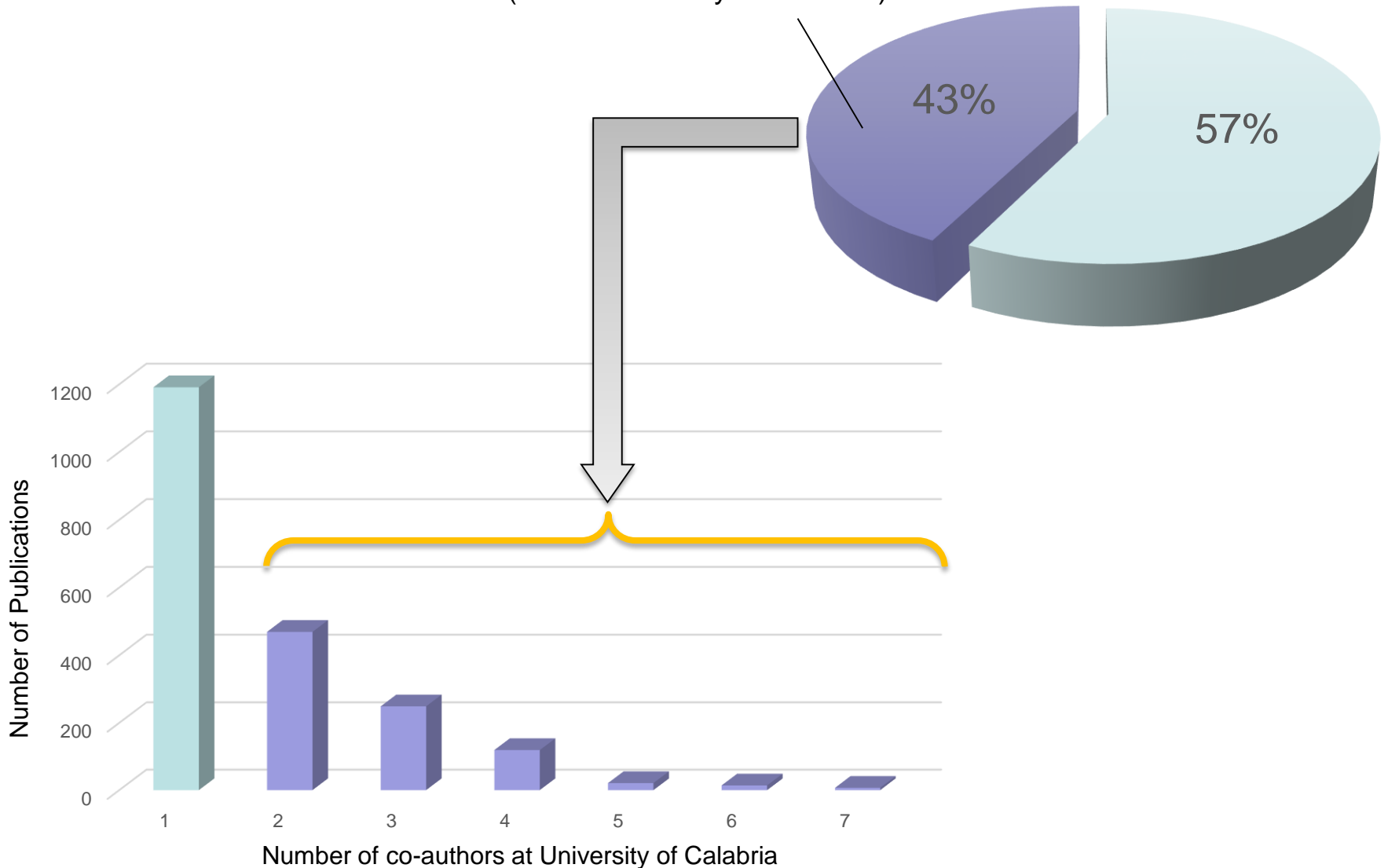
(based on **declared** values, i.e., not necessarily true!)

# Co-Autorships at University of Calabria



# Co-Autorships at University of Calabria

Co-authored (within University of Calabria)





# ANVUR Evaluation



ANVUR Criteria

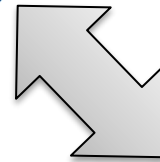


Self-evaluations



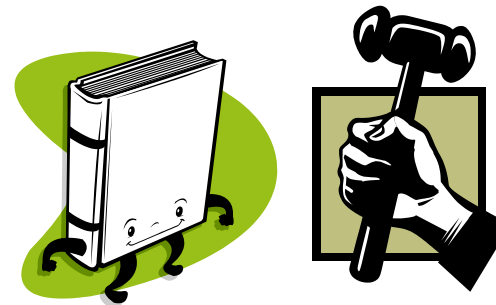
$score_r(p)$

ANVUR Evaluation



$score_{VQR}(p)$

Selected publications



# ANVUR Evaluation



ANVUR Criteria



Self-evaluations



$score_r(p)$



Selected publications



$score_{VQR}(p)$



Evaluation



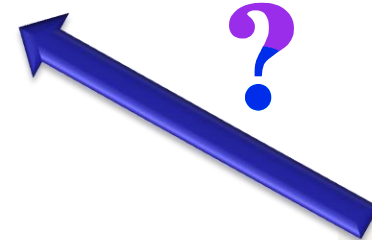
# ANVUR Evaluation



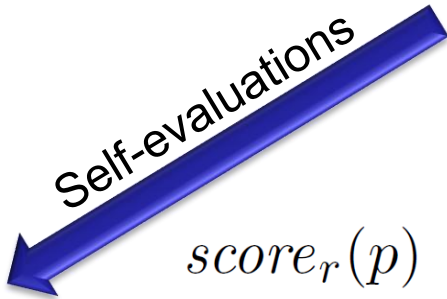
ANVUR Criteria



Division Rules



Self-evaluations



$score_r(p)$



Selected publications



$score_{VQR}(p)$



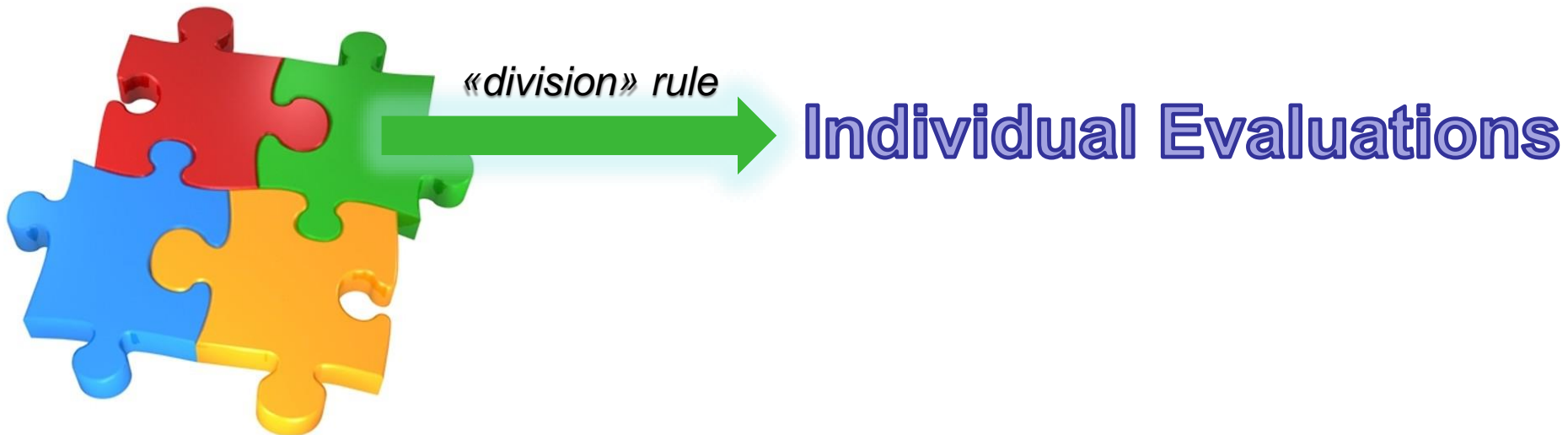
Evaluation



# Issues

- Allocation Problem
- Valuations are declared
- The program is meant to evaluate the structures...
  - ...but outcomes are used to evaluate researchers, too

## Global Evaluation



# Desiderata for Division Rules

- (P1) **“budget-balance”**: A division rule  $\gamma$  must completely distribute the VQR score of  $R$  over all its members, i.e.,  $\sum_{r \in \mathcal{R}} \gamma_r(\psi^*) = \text{score}_{VQR}(R)$ .
- (P2) **“fairness”**: A division rule  $\gamma$  must be indifferent w.r.t. the optimal allocation being selected, i.e., for each  $r \in \mathcal{R}$ , and for each pair of optimal allocations  $\psi^*$  and  $\hat{\psi}^*$ ,  $\gamma_r(\psi^*) = \gamma_r(\hat{\psi}^*)$ .
- (P3) **“implementability”**: A division rule  $\gamma$  must be indifferent w.r.t. the scores (possibly cheats) declared for unverified products, that is, for products not occurring in the selected allocation.
- (P4) **“truthfulness”**: A division rule  $\gamma$  must provide no incentive in misreporting the score of the research products.
- (P5) **“no punishment”**: A division rule  $\gamma$  must be such that, for each  $r \in \mathcal{R}$  and each allocation  $\psi^*$ , the value  $\gamma_r(\psi^*)$  is indifferent w.r.t. self-assessed scores, in particular, w.r.t. discrepancies possibly emerging between such scores and VQR ones.

# Three «Natural» Division Rules

- $\text{proj}_r(\psi^*) = \sum_{p \in \psi^*(r)} \text{score}_{\text{VQR}}(p)$

# Three «Natural» Division Rules

- $\text{proj}_r(\psi^*) = \sum_{p \in \psi^*(r)} \text{score}_{\text{VQR}}(p)$
- $\text{owner}_r(\psi^*) =$  assign to each author the sum of the “normalized” scores of the submitted products (s)he has co-authored, where by normalization we just mean here dividing the score of any product by the number of its authors

# Three «Natural» Division Rules

- $\text{proj}_r(\psi^*) = \sum_{p \in \psi^*(r)} \text{score}_{\text{VQR}}(p)$
- $\text{owner}_r(\psi^*) =$  assign to each author the sum of the “normalized” scores of the submitted products (s)he has co-authored, where by normalization we just mean here dividing the score of any product by the number of its authors
- $\text{all}_r(\psi^*) = \frac{\sum_{p \in \text{products}(r)} \text{score}_r(p)}{\sum_{r \in \mathcal{R}} \sum_{p \in \text{products}(r)} \text{score}_r(p)} \times \sum_{p \in \psi^*(r)} \text{score}_{\text{VQR}}(p)$



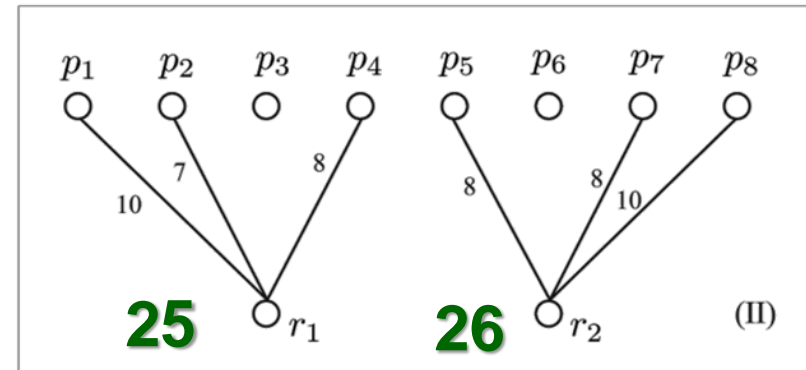
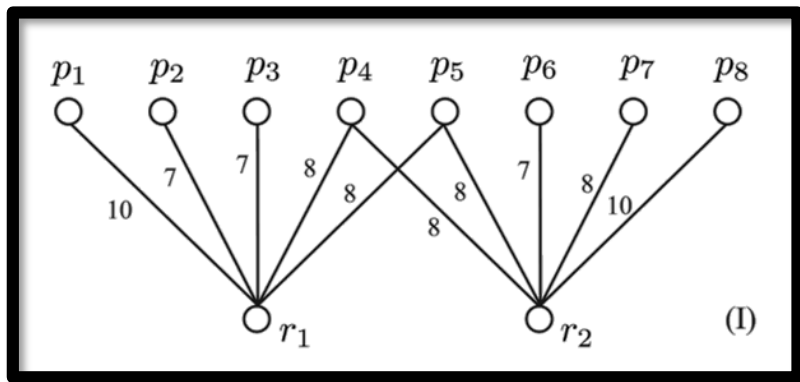
# Three «Natural» Division Rules

- $\text{proj}_r(\psi^*) = \sum_{p \in \psi^*(r)} \text{score}_{\text{VQR}}(p)$
- $\text{owner}_r(\psi^*) =$  assign to each author the sum of the “normalized” scores of the submitted products (s)he has co-authored, where by normalization we just mean here dividing the score of any product by the number of its authors
- $\text{all}_r(\psi^*) = \frac{\sum_{p \in \text{products}(r)} \text{score}_r(p)}{\sum_{r \in \mathcal{R}} \sum_{p \in \text{products}(r)} \text{score}_r(p)} \times \sum_{p \in \psi^*(r)} \text{score}_{\text{VQR}}(p)$

## Are they good division rules?

The last one is clearly not implementable, because it depends on publications without any evaluation by ANVUR. What about the others?

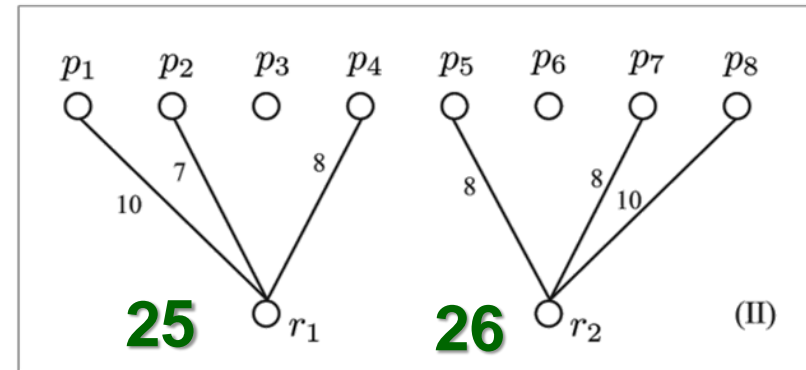
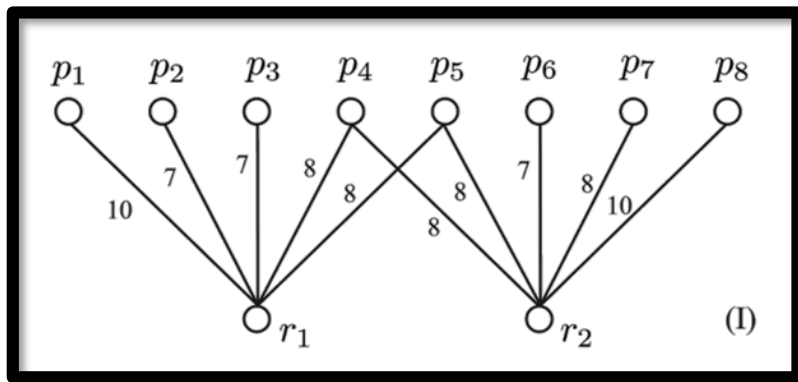
# Division Rule: proj



---

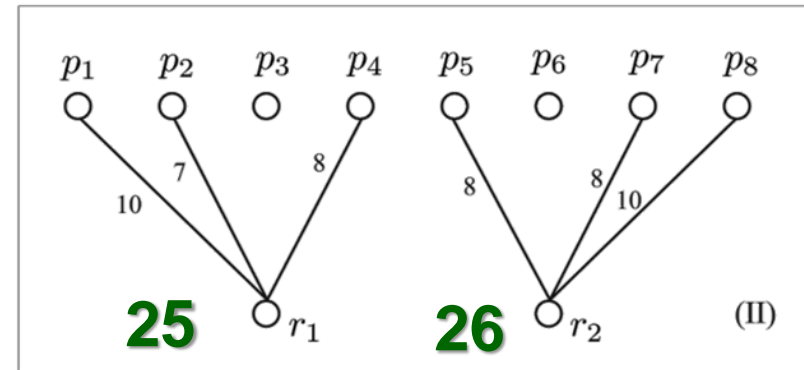
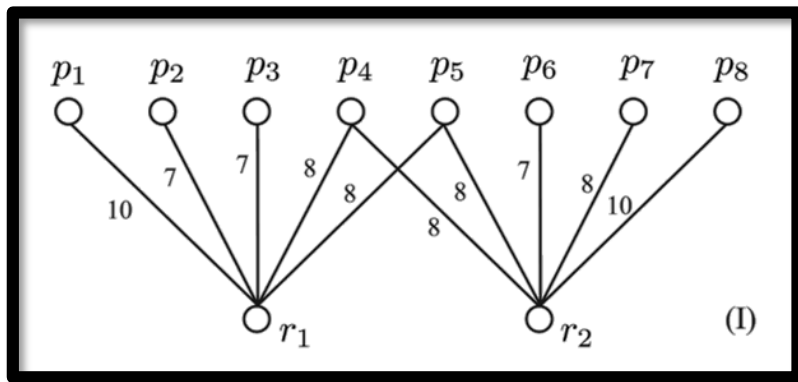
(P1) “budget-balance”: A division rule  $\gamma$  must completely distribute the VQR score of  $R$  over all its members, i.e.,  $\sum_{r \in R} \gamma_r(\psi^*) = \text{score}_{VQR}(R)$ .

# Division Rule: proj



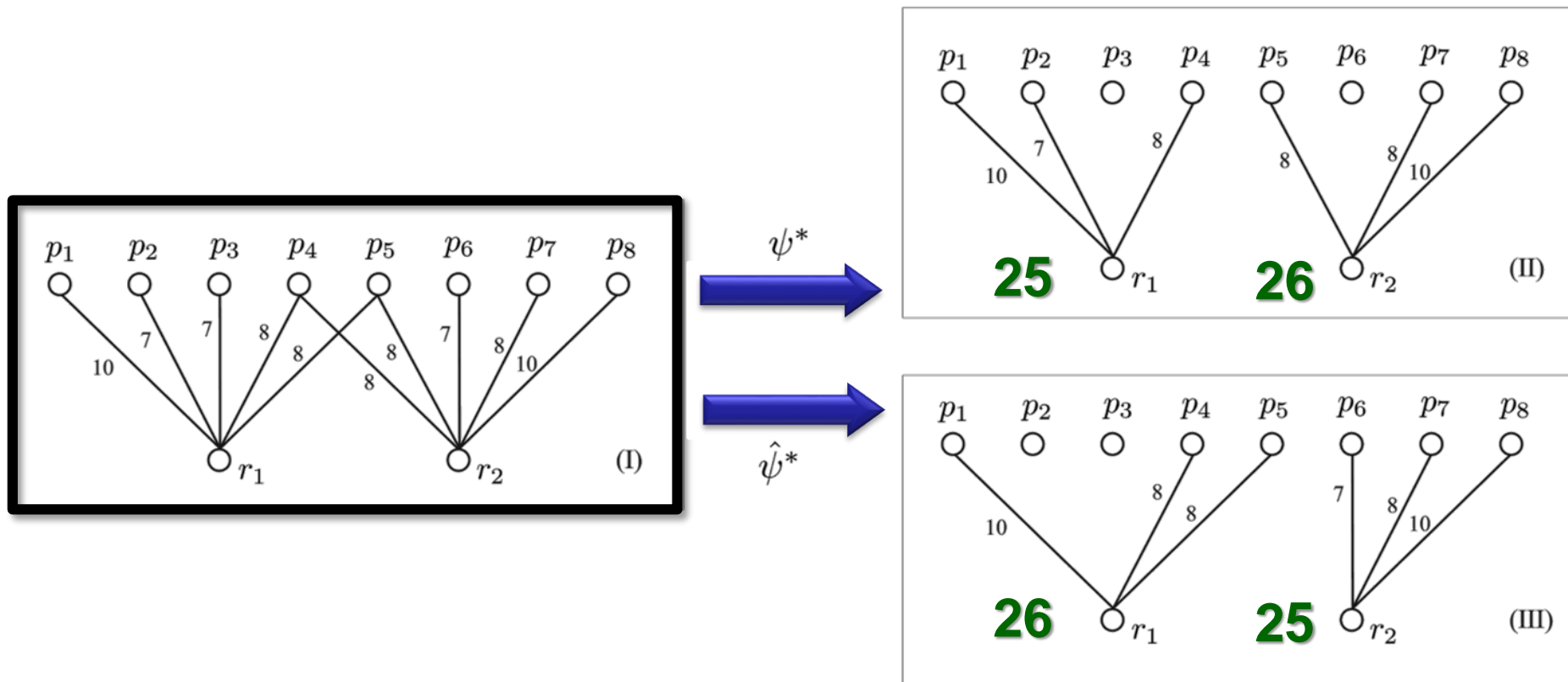
(P1) “budget-balance”: A division rule  $\gamma$  must completely distribute the VQR score of  $R$  over all its members, i.e.,  $\sum_{r \in R} \gamma_r(\psi^*) = \text{score}_{VQR}(R)$ .

# Division Rule: proj



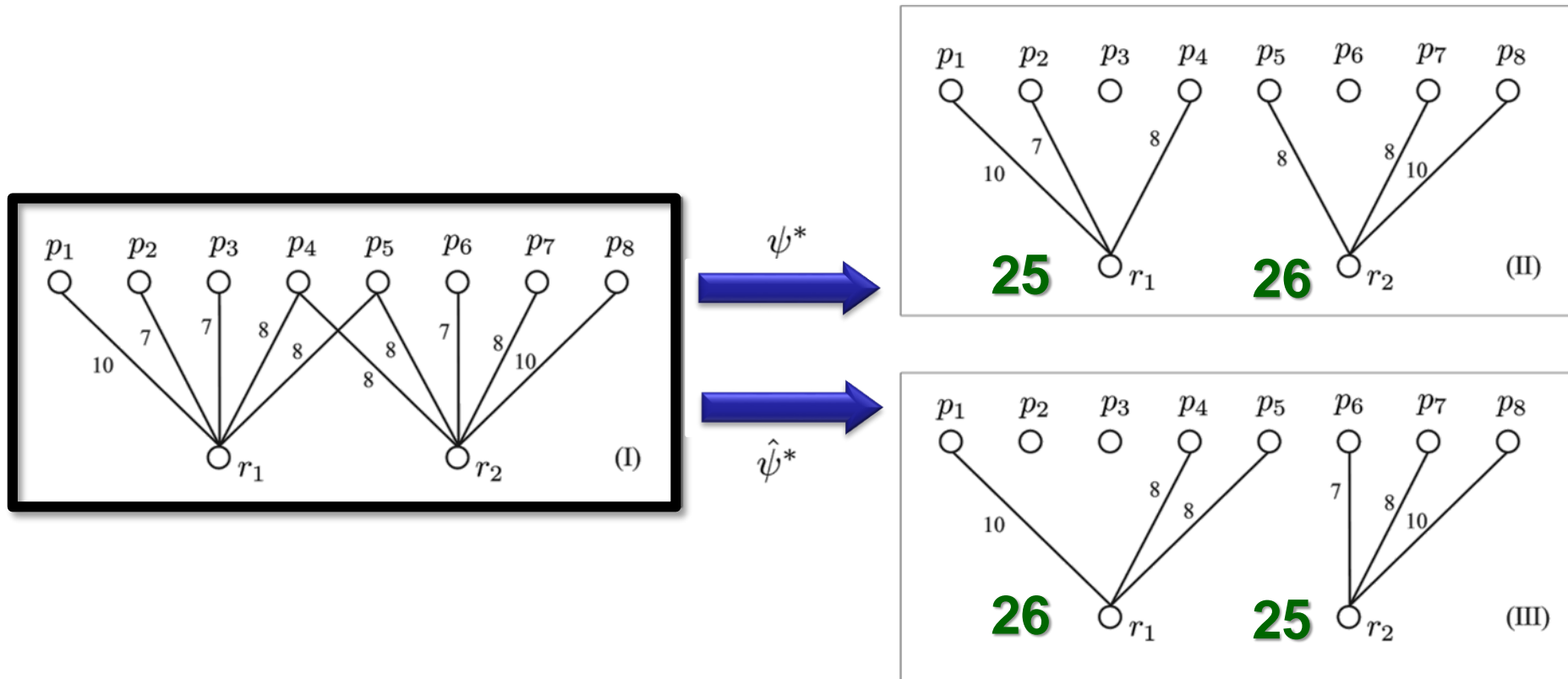
(P2) “fairness”: A division rule  $\gamma$  must be indifferent w.r.t. the optimal allocation being selected, i.e., for each  $r \in \mathcal{R}$ , and for each pair of optimal allocations  $\psi^*$  and  $\hat{\psi}^*$ ,  $\gamma_r(\psi^*) = \gamma_r(\hat{\psi}^*)$ .

# Division Rule: proj



(P2) “fairness”: A division rule  $\gamma$  must be indifferent w.r.t. the optimal allocation being selected, i.e., for each  $r \in \mathcal{R}$ , and for each pair of optimal allocations  $\psi^*$  and  $\hat{\psi}^*$ ,  $\gamma_r(\psi^*) = \gamma_r(\hat{\psi}^*)$ .

# Division Rule: proj



~~(P2) “fairness”: A division rule  $\gamma$  must be indifferent w.r.t. the optimal allocation being selected, i.e., for each  $r \in \mathcal{R}$ , and for each pair of optimal allocations  $\psi^*$  and  $\hat{\psi}^*$ ,  $\gamma_r(\psi^*) = \gamma_r(\hat{\psi}^*)$ .~~

# Back on the Desiderata...

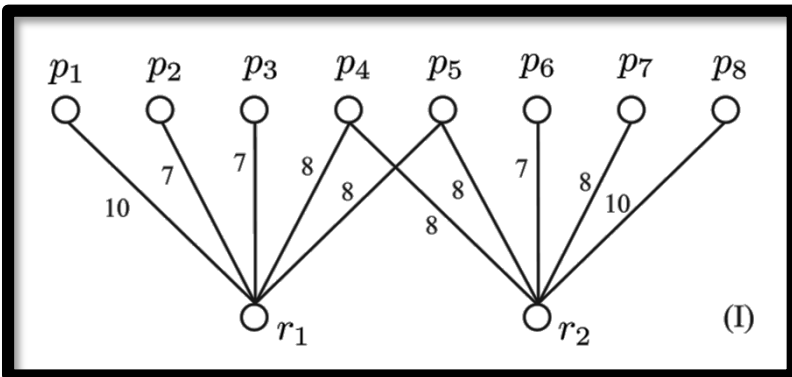
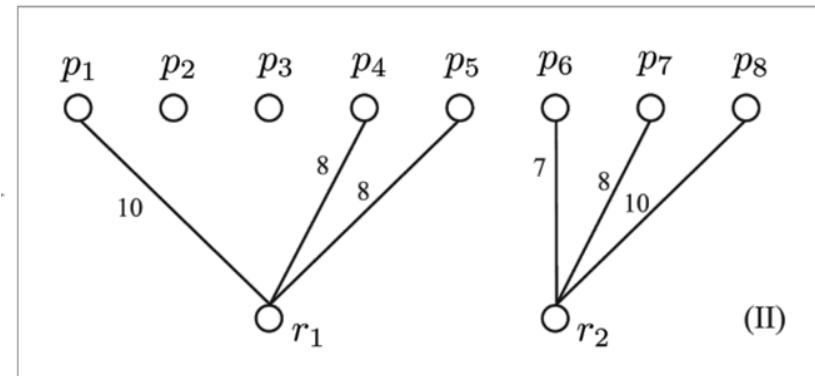
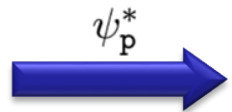
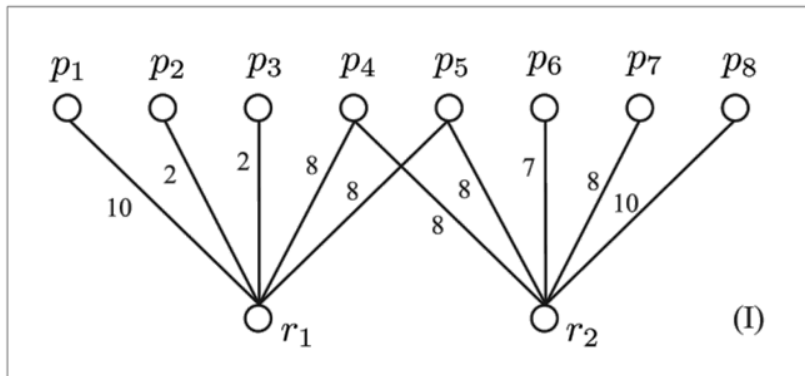
- (P1) **“budget-balance”**: A division rule  $\gamma$  must completely distribute the VQR score of  $R$  over all its members, i.e.,  $\sum_{r \in \mathcal{R}} \gamma_r(\psi^*) = \text{score}_{VQR}(R)$ .
- (P2) **“fairness”**: A division rule  $\gamma$  must be indifferent w.r.t. the optimal allocation being selected, i.e., for each  $r \in \mathcal{R}$ , and for each pair of optimal allocations  $\psi^*$  and  $\hat{\psi}^*$ ,  $\gamma_r(\psi^*) = \gamma_r(\hat{\psi}^*)$ .
- (P3) **“implementability”**: A division rule  $\gamma$  must be indifferent w.r.t. the scores (possibly cheats) declared for unverified products, that is, for products not occurring in the selected allocation.
- (P4) **“truthfulness”**: A division rule  $\gamma$  must provide no incentive in misreporting the score of the research products.
- (P5) **“no punishment”**: A division rule  $\gamma$  must be such that, for each  $r \in \mathcal{R}$  and each allocation  $\psi^*$ , the value  $\gamma_r(\psi^*)$  is indifferent w.r.t. self-assessed scores, in particular, w.r.t. discrepancies possibly emerging between such scores and VQR ones.

# Back on the Desiderata...

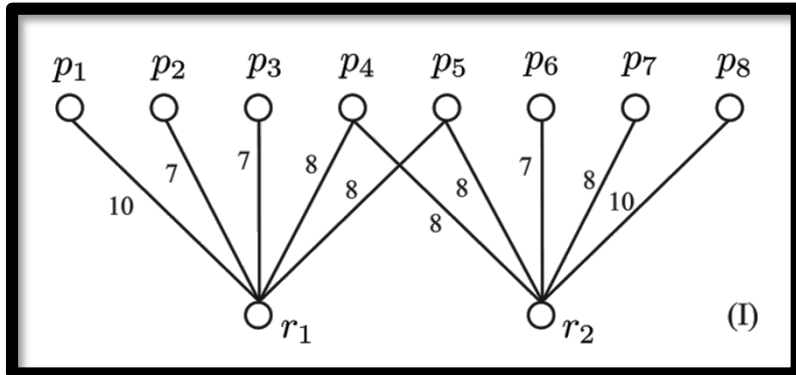
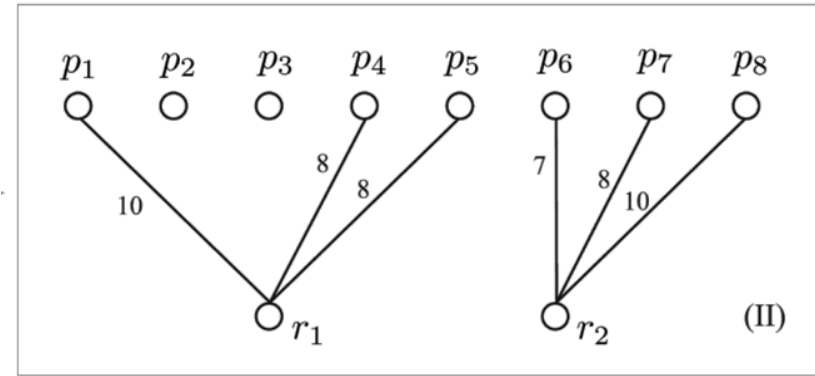
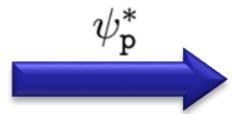
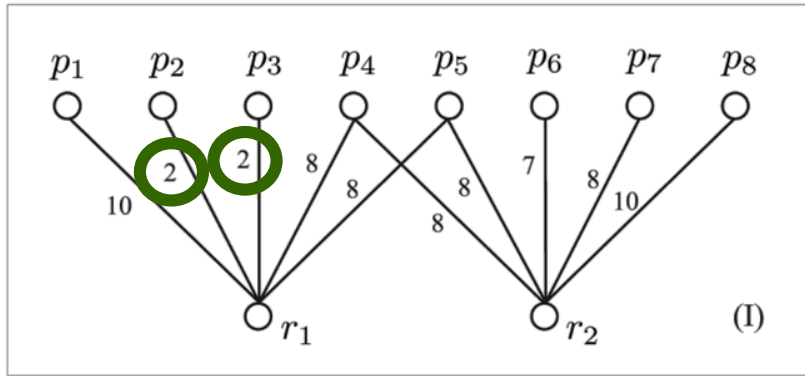
- (P1) “budget-balance”: A division rule  $\gamma$  must completely distribute the VQR score of  $R$  over all its members, i.e.,  $\sum_{r \in \mathcal{R}} \gamma_r(\psi^*) = \text{score}_{VQR}(R)$ .
- (P2) “fairness”: A division rule  $\gamma$  must be indifferent w.r.t. the optimal allocation being selected, i.e., for each  $r \in \mathcal{R}$ , and for each pair of optimal allocations  $\psi^*$  and  $\hat{\psi}^*$ ,  $\gamma_r(\psi^*) = \gamma_r(\hat{\psi}^*)$ .
- (P3) “implementability”: A division rule  $\gamma$  must be indifferent w.r.t. the scores (possibly cheats) declared for unverified products, that is, for products not occurring in the selected allocation.
- (P4) “truthfulness”: A division rule  $\gamma$  must provide no incentive in misreporting the score of the research products.
- (P5) “no punishment”: A division rule  $\gamma$  must be such that, for each  $r \in \mathcal{R}$  and each allocation  $\psi^*$ , the value  $\gamma_r(\psi^*)$  is indifferent w.r.t. self-assessed scores, in particular, w.r.t. discrepancies possibly emerging between such scores and VQR ones.



# Strategic Manipulations: proj



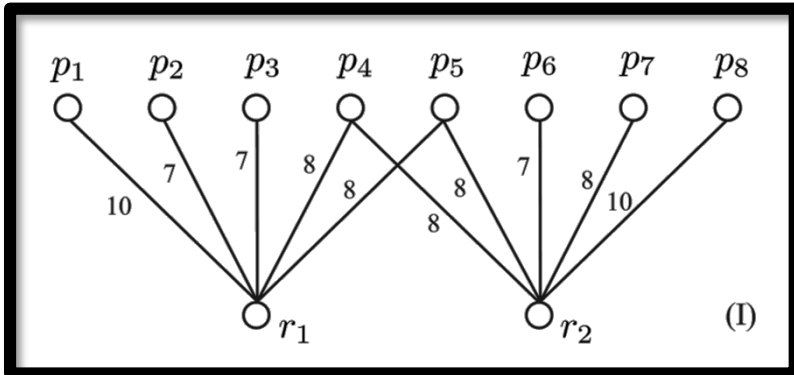
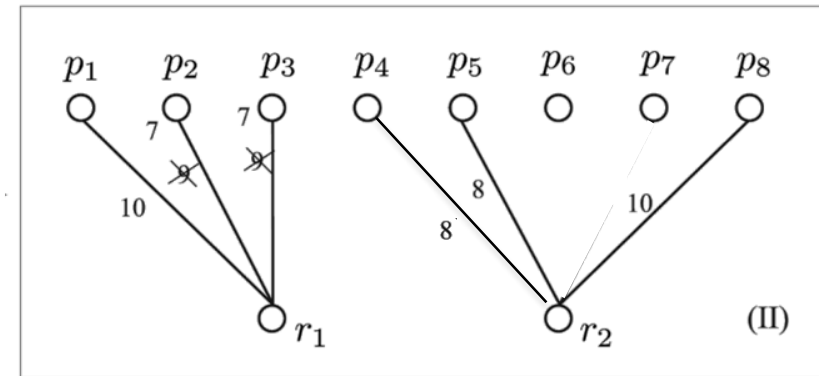
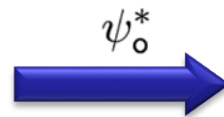
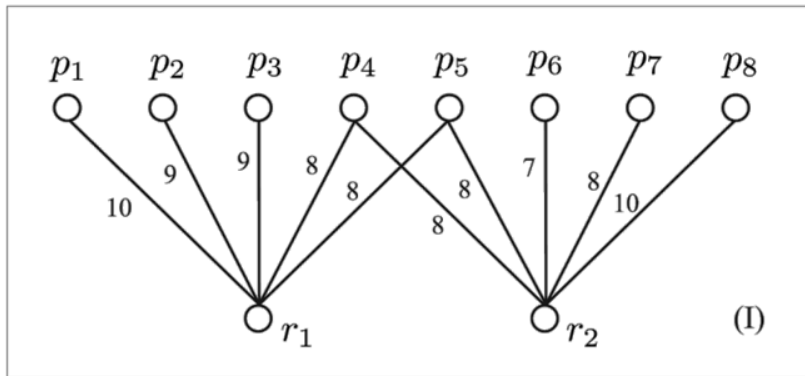
# Strategic Manipulations: proj



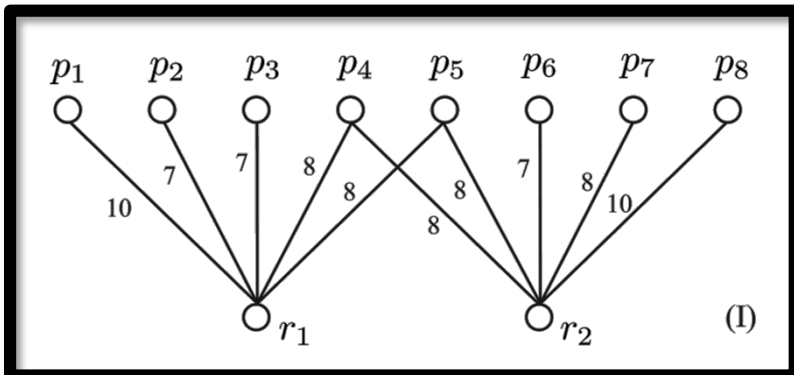
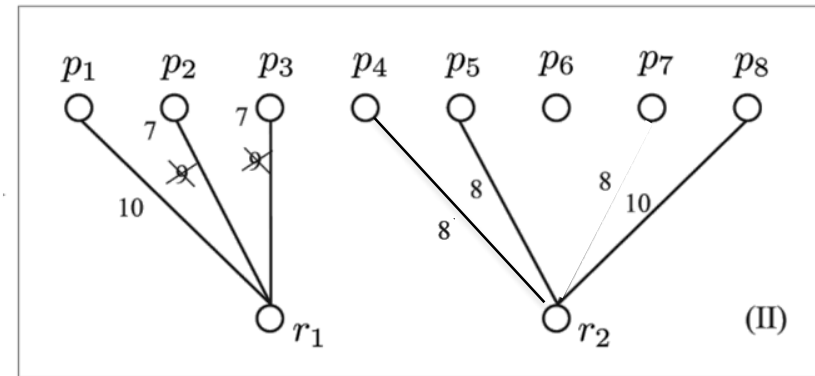
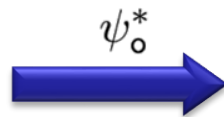
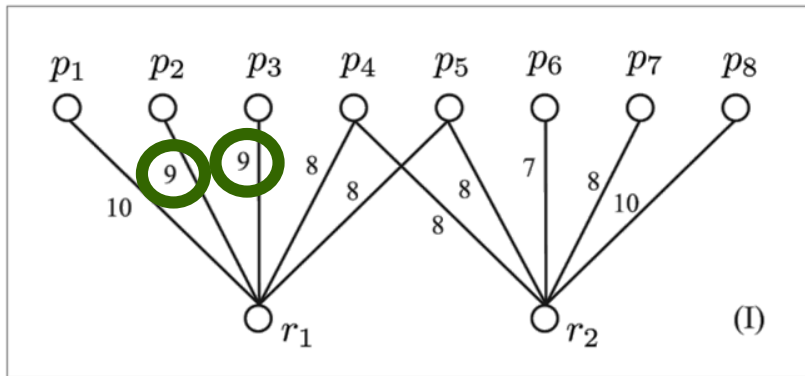
## Under-estimation



# Strategic Manipulations: owner



# Strategic Manipulations: owner



## Over-estimation

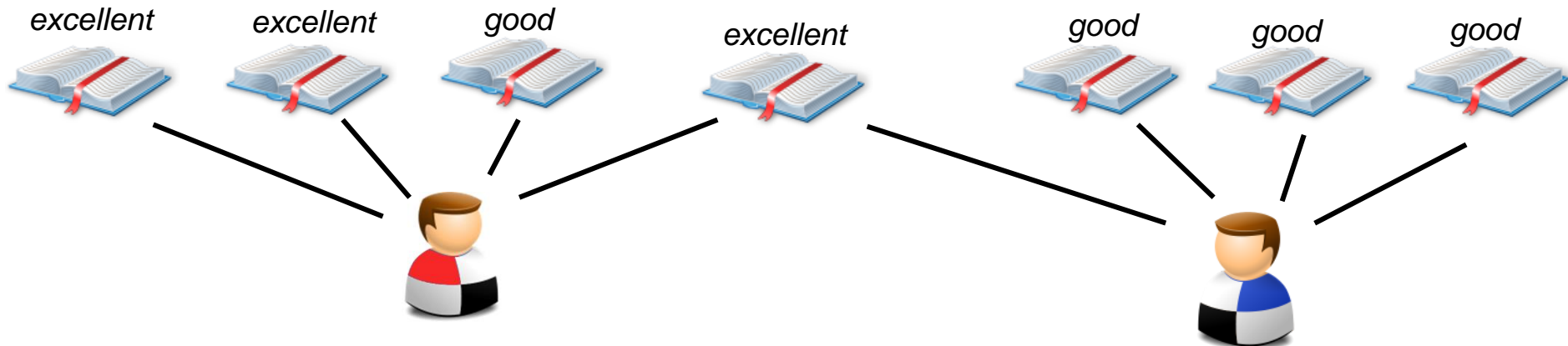


The optimal solution is missed!!!



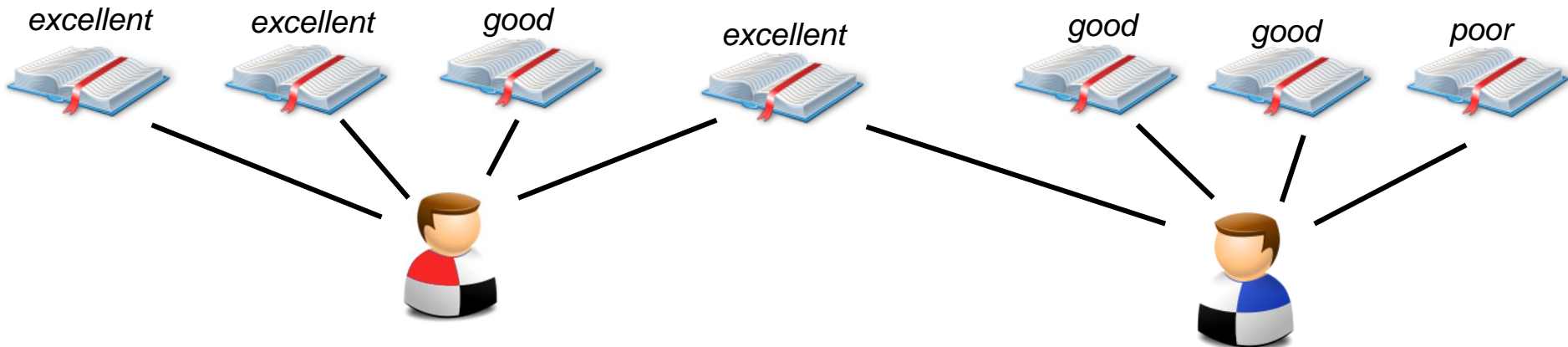
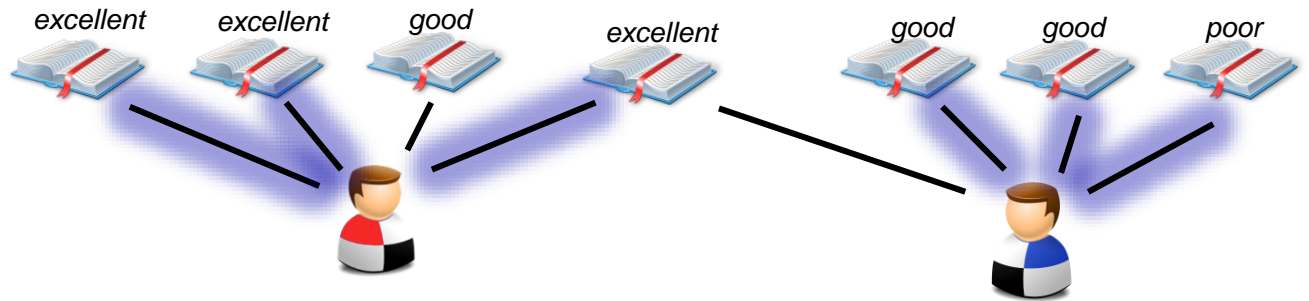
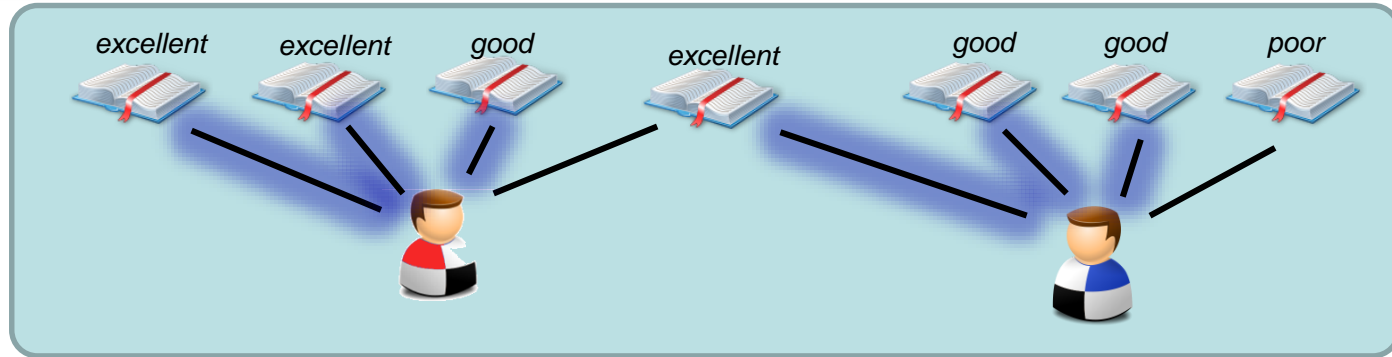
## A Closer Look

# Efficiency VS Fairness



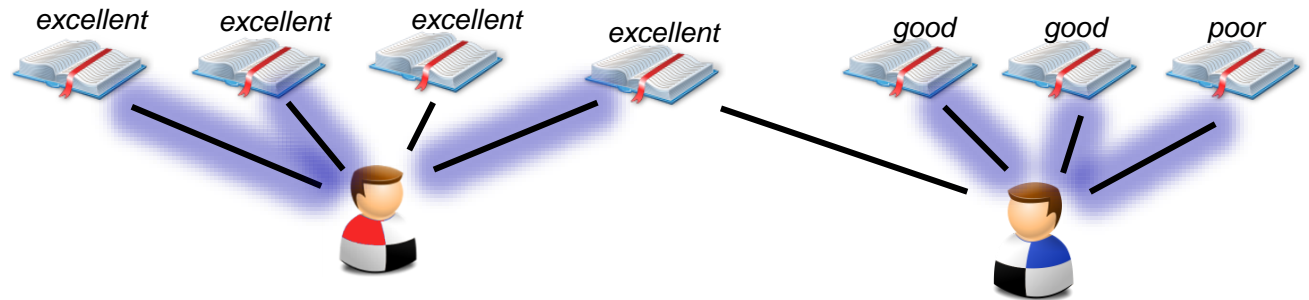
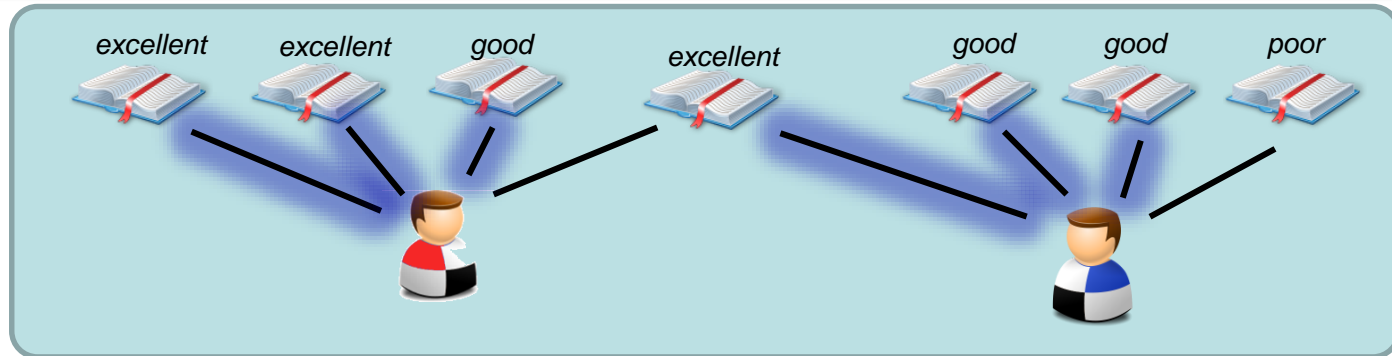
# A Closer Look

*Optimal Allocation*



# A Closer Look

*Optimal Allocation*

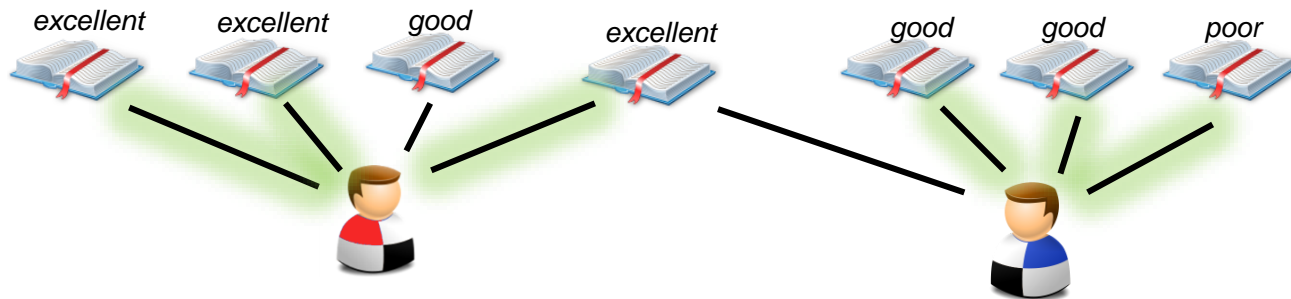


❖ «Penalizing»  is not fair!

❖ Unless it is clear that no penalization will occur,  will act «strategically»

# The Story....

- ANVUR did not specify a division rule
- Reserchers considered *proj* as «the rule»
- Researchers submitted (rated) only the minimum number of publications required (by default 3), thus implicitly under-estimating all their other products
- To avoid overlapping submissions, «agreements» have been made



- Conflicts resolved «strategically», «hierarchically», ...

**The optimum has been missed!**  
**No fairness at all!**



## ...and the reaction

- ANVUR declared that VQR has not to be used to evaluate researchers, but only structures

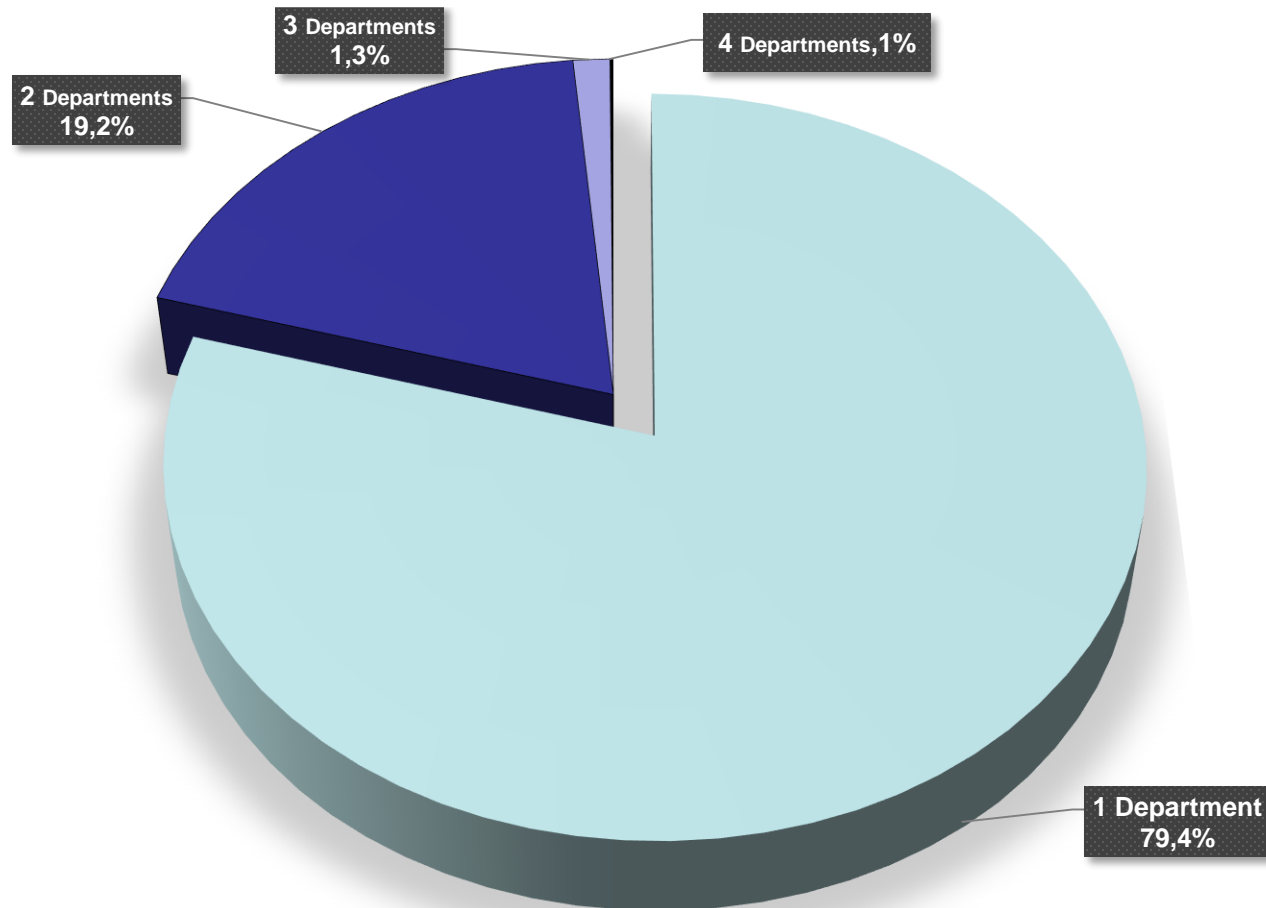


*Waste of money... and even just false!*

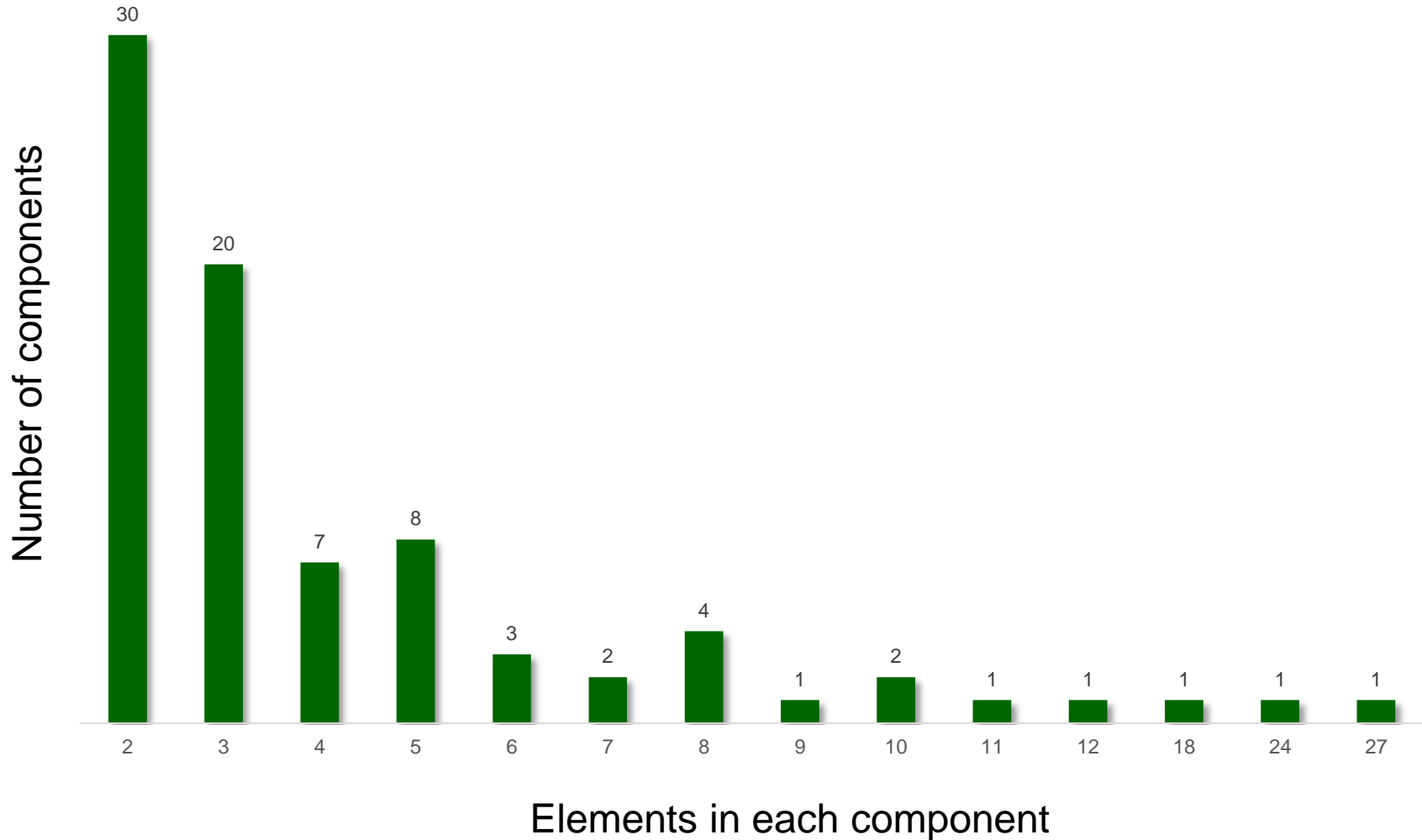


Moreover,  
what about Departments?

# Distribution at University of Calabria



# Components at University of Calabria



# Implementation Strategies

- Sampling
- Structural properties



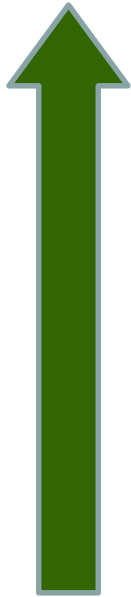
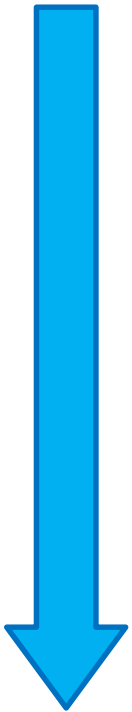
**Very particular interaction graph**

---

**Treewidth is bounded!**

# Lessons?

GAME THEORY, AI, ...



SOCIETY



Thank you!