



# The classification of Research Organisations: Taxonomical explorations with a set of “dedicated” Research organisations

**Cristina Peñasco, Luis Sanz-Menéndez, Laura Cruz-Castro, Catalina Martínez**

**Institute of Public Policies and Goods (IPP).  
Spanish National Research Council (CSIC)  
Madrid, Spain.**

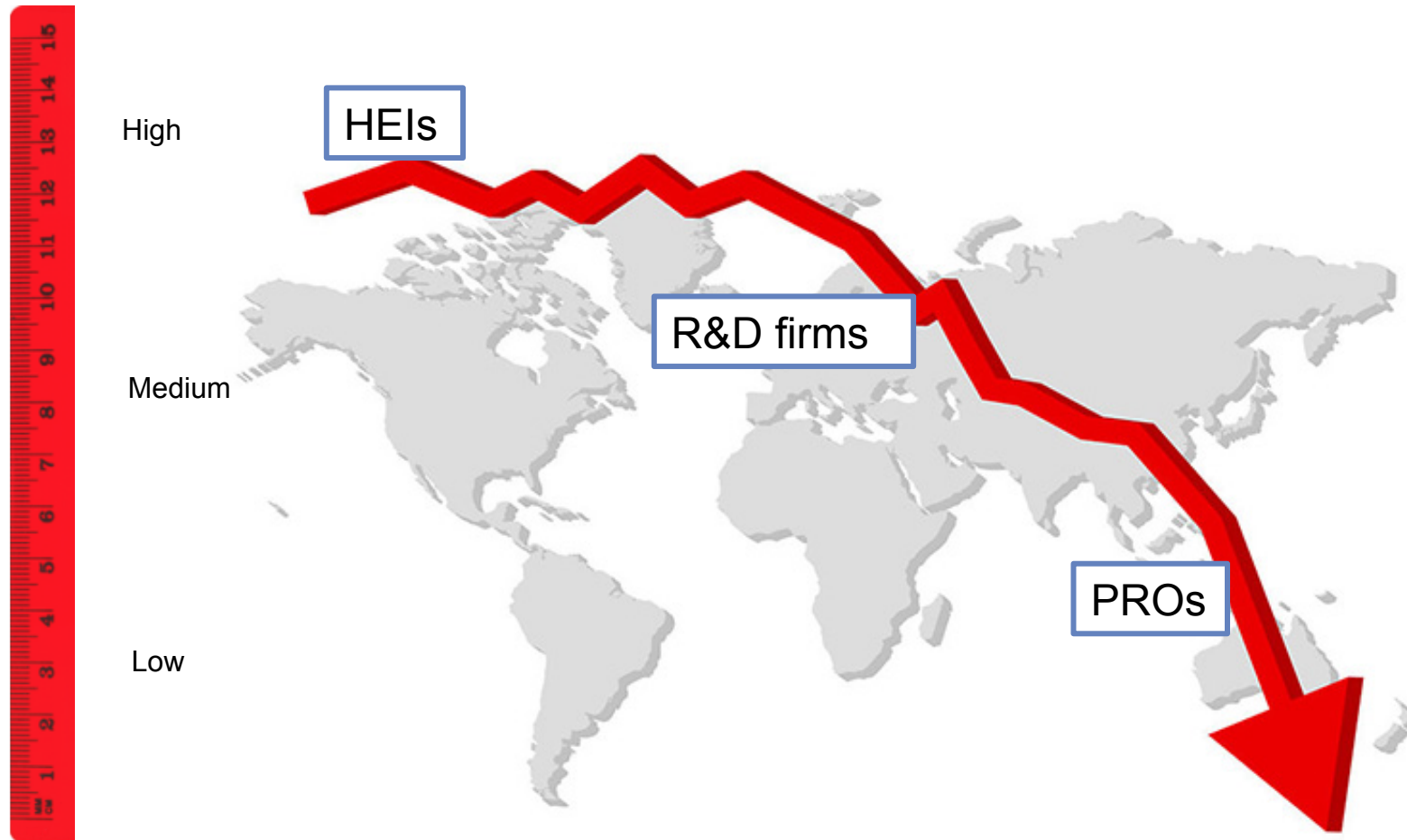
*Annual Conference of EU-SPRI Forum.  
6-8 June 2017- Paris*

# OUTLINE

- 1. Introduction**
- 2. Analytical background**
- 3. Data and methodology**
- 4. Results**
- 5. Conclusions**

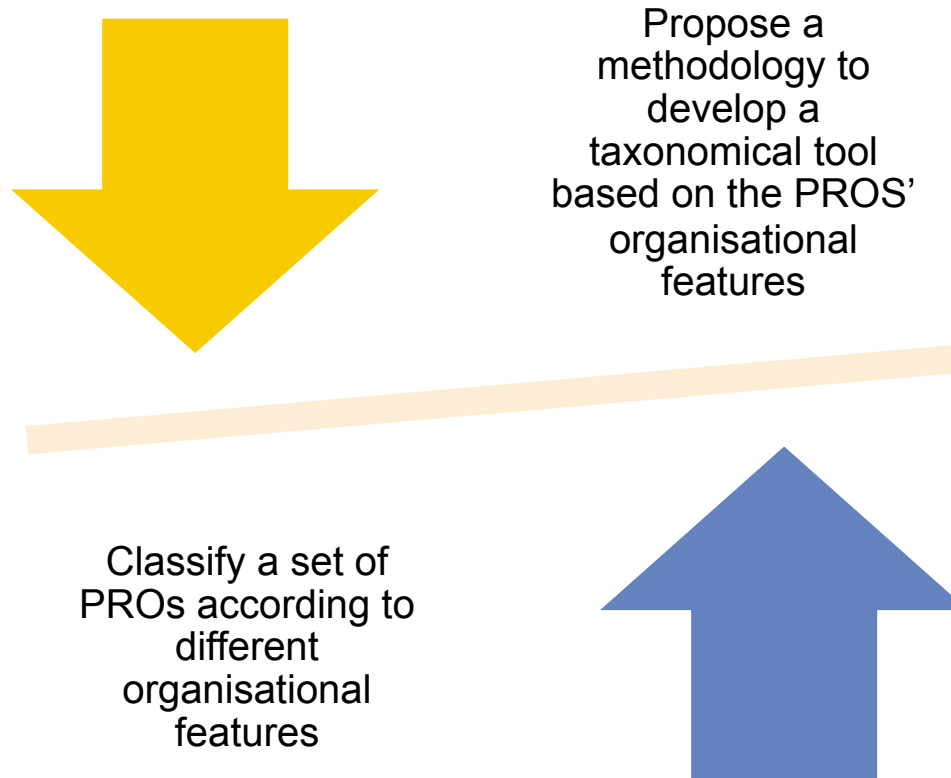
# 1. INTRODUCTION

## Empirical and analytical work of Science and Innovation Studies



# 1. INTRODUCTION

## AIM



# 1. INTRODUCTION

## CONTRIBUTION

- A systematic study about the role played by PROs and their evolution is needed.
- The changing context forces a further analysis and exploration.
- Increasing heterogeneity in the population of research centres in Europe
- New organisational forms

**This work tries to close this gap by analysing:**

How taxonomies developed in the past could be improved if informed by organisational theory.

Set up the foundations for monitoring, analysing and studying the dynamics of PROs from a comparative perspective.

# THE VALUE OF TYPOLOGIES (IDEAL TYPES) CLASSIFICATION OF (IDEAL TYPES) PRO ACCORDING TO ATTRIBUTES A (EXTERNAL AUTONOMY) AND B (INTERNAL AUTHORITY)

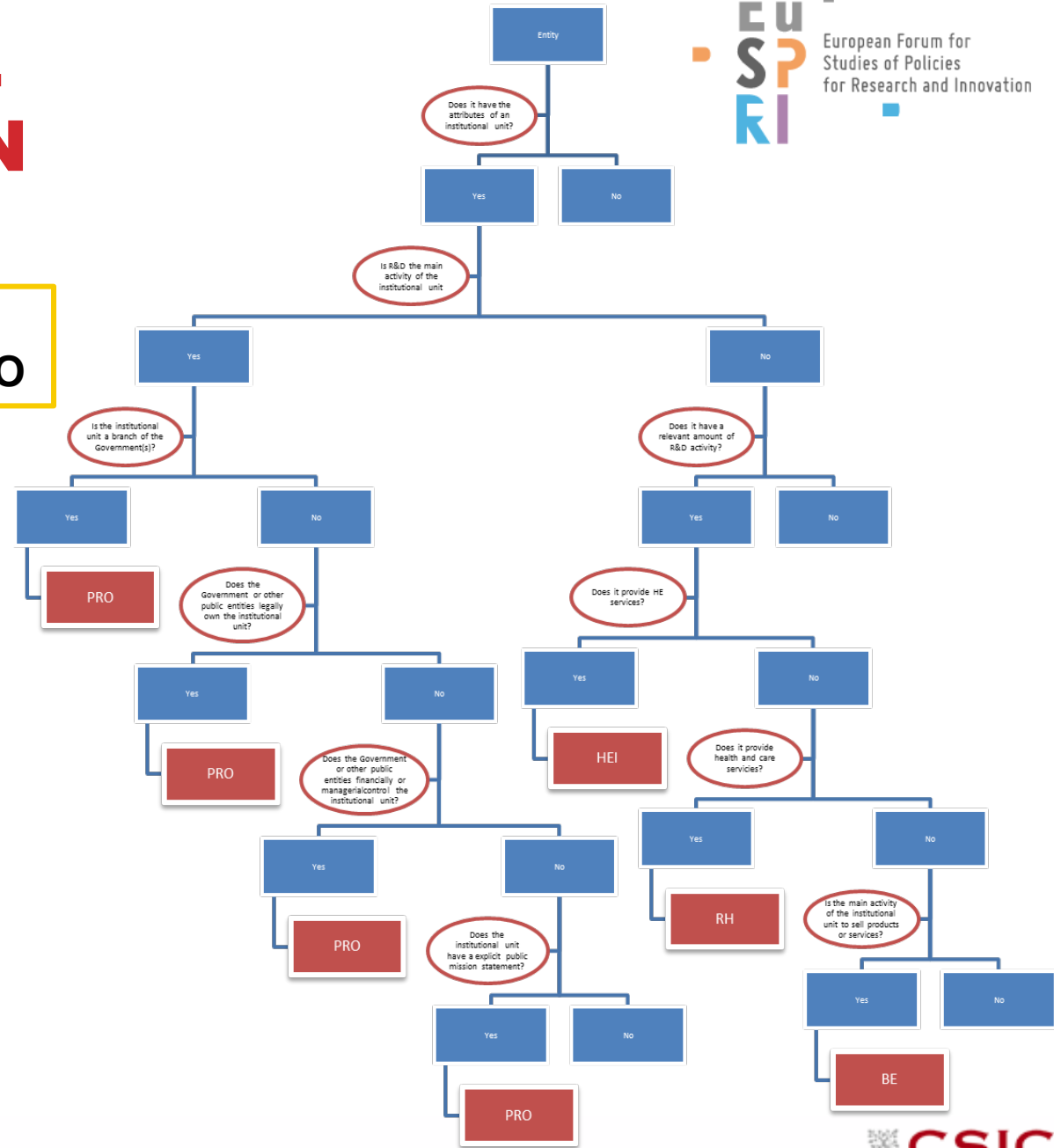
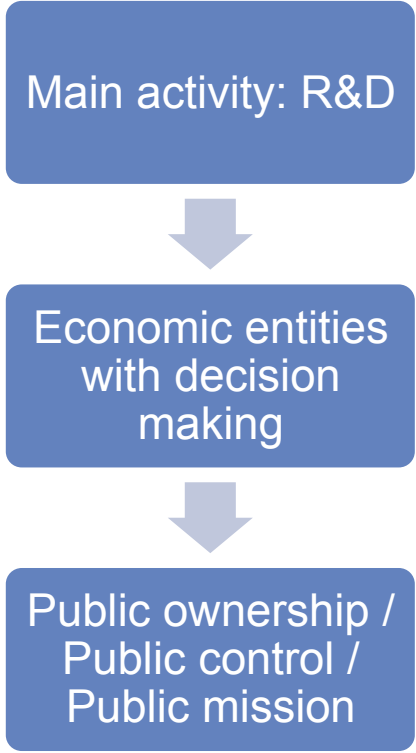
| PRO Management    |   | Internal authority |     |
|-------------------|---|--------------------|-----|
|                   |   | +                  | -   |
| External autonomy | + | RTO                | IRI |
|                   | - | MOC                | PRC |

## PREVIOUS APPROACHES AND EMPIRICAL EFFORTS TO CHARACTERISE PROs

| Research   | Dimensions / criteria   | Typologies   |
|--|---|--|
| Crow and Bozeman (1987)                                  | Publicness and R&D market influence   | <ol style="list-style-type: none"> <li>1. Public generic</li> <li>2. Public multi-market</li> <li>3. Public market</li> <li>4. Quasi-public generic</li> <li>5. Quasi-public multi-market</li> <li>6. Quasi-public market</li> <li>7. Independent generic</li> <li>8. Independent multimarket</li> <li>9. Independent market</li> </ol>                                      |
| Cruz-Castro et al (2012)                                 | Publicness and R&D market influence   | <ol style="list-style-type: none"> <li>1. Public science</li> <li>2. Public science and technology</li> <li>3. Public technology</li> <li>4. Mixed-source science</li> <li>5. Mixed-source science and technology</li> <li>6. Mixed-source technology</li> <li>7. Private niche science</li> <li>8. Private science and technology</li> <li>9. Private technology</li> </ol> |
| Arnold et al., (2010)                                    | History, evolution and current attributes                                     | <ol style="list-style-type: none"> <li>1. Government laboratories</li> <li>2. Academic</li> <li>3. Research and technology organisations</li> </ol>  |
| Sanz-Menéndez et al., (2011) / Cruz-Castro et al. (2015) | Publicness and main research purpose (authority and dependence relationships) | <ol style="list-style-type: none"> <li>1. Public research centers and councils (PRC)</li> <li>2. Mission oriented centers (MOC)</li> <li>3. Research technology organisations (RTO)</li> <li>4. Independent research institutes (IRI)</li> </ol>   |

# CONCEPTUAL DELIMITATION

**Multidimensional and inclusive definition of a PRO**





# 3. . THEORETICAL APPROACHES TO DEFINING ORGANIZATIONAL FORMS

| Focus with respect to organizational boundaries |  |  |
|---|--|--|
| Focus with respect to role of perception        | Internal   | External                                 |
|   | <i>'Blueprints'</i>  | <i>'Resource Niches'</i>                 |
| <b>Objective</b>                                | (Hannan and Freeman 1977; McKelvey 1982; Pentland and Rueter 1994) | (Hannan and Carroll 1995; DiMaggio 1986) |
|   | <i>'Organizational Identities'</i>                                 | <i>'Cultural Codes'</i>                  |
| <b>Subjective</b>                               | (Albert and Whetten 1985; Martin 2002)                             | (Pólos et al. 2002; Zuckerman 1999)      |

# WHY THOSE COUNTRIES?

|   | Size of the country | Location (North, central, South) | Level of R&D Expenditure | Share of the Government sector in total | Degree of political centralization Central versus Federal | Policy Making Traditions (Castles 1998)<br>Scandinavian, Continental, South | Public Administration Traditions (Napoleonic versus Rechtsstaat) |
|---|---------------------|----------------------------------|--------------------------|---|---|---|--|
|   | DE, FR, IT          | NO                               | DE, AT                   | PO, ES                                  | FR, PO, NL, NO  | NO  | FR, ES, IT   |
| n | ES, PO              | DE, FR, AT, NL, PO               | FR, NL, NO               | NO, DE, IT, FR, NL                      | IT  | DE, FR, AT, NL, PO  | PO, NO   |
|   | NL, AT, NO          | ES, IT                           | IT, ES, PO               | AT                                      | DE, ES, AT  | ES, IT  | DE, AT, NL   |

# 3. DATA AND METHODOLOGY

## DATASET AND DESCRIPTIVE STATISTICS

Secondary  
institutional  
sources of  
information /  
Validation

PROs in AT,  
DE, ES, FR, IT,  
NL, NO, PL.

Sample size:  
201 PROs

Reference year  
2013

| Country         | Freque<br>ncy | Valid<br>Percent |
|-----------------|---------------|------------------|
| AUSTRIA         | 15            | 7,6              |
| GERMANY         | 28            | 14,2             |
| SPAIN           | 51            | 25,9             |
| FRANCE          | 29            | 14,7             |
| ITALY           | 29            | 14,7             |
| NETHERLA<br>NDS | 19            | 9,6              |
| NORWAY          | 16            | 8,1              |
| POLAND          | 10            | 5,1              |
| <b>Total</b>    | <b>197</b>    | <b>100,0</b>     |

# 3. DATA AND METHODOLOGY

## DATASET AND DESCRIPTIVE STATISTICS

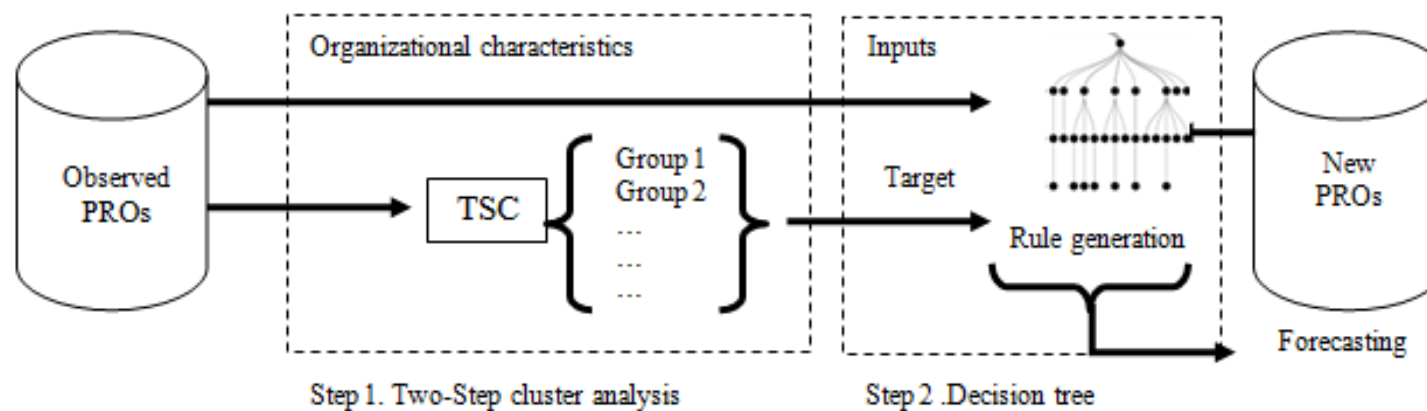
| VARIABLES  | DEFINITION/EXPLANATION  | Mean (Std.) | Frequencies   |
|--|---|-------------|---------------|
| <b><i>G1. Internal institutional characteristics</i></b> |   |             |               |
| Legal status   | 1= Public   | 2.01 (1.02) | 1=93 (47.2%)  |
|  | 2= Private  |             | 2=18 (9.1%)   |
|  | 3= Not for profit (NPI)   |             | 3=78 (39.6%)  |
|  | 4= Other  |             | 4=8 (4.1%)    |
| Main Owner   | 1=Central government  | 1.61 (0.92) | 1=114 (57.9%) |
|  | 2=Regional or local government  |             | 2=49 (24.9%)  |
|  | 3=Private Company   |             | 3=19 (9.6%)   |
|  | 4=Higher Education  |             | 4=3 (1.5%)    |
|  | 5=Other   |             | 5=5 (2.5%)    |
|  |   |             | m=7 (3.6%)    |
| <b><i>G2. Resource niche of each RO</i></b>              |   |             |               |
| Number FOS   | Number of sectors of research activity (FOS-2007) classification in which the PRO is involved as a proxy of multidisciplinary | 2.22 (1.29) | 1=68 (34.5%)  |
|  |   |             | 2=67 (34.0%)  |
|  |   |             | 3=34 (17.3%)  |
|  |   |             | 4=14 (7.1%)   |
|  |   |             | 5=7 (3.6%)    |
|  |   |             | 6=7 (3.6%)    |
| Main Orientation   | Main orientation of the R&D activity  | 1.71 (0.69) | 1=84 (42.6)   |
|  | 1=Basic Research  |             | 2=87 (44.2)   |
|  | 2=Applied Research  |             | 3=26 (13.2)   |
|  | 3=Experimental Development  |             |               |
| <b><i>G3. Subjective claims of identity</i></b>          |   |             |               |
| Main mission   | 1=Develop knowledge   | 1.82 (0.78) | 1=80 (40.6)   |
|  | 2=Contribute to solve public policy issues  |             | 2=72 (36.5)   |
|  | 3=Generate economic value   |             | 3=44 (22.3%)  |
|  |   |             | m= 1 (0.5%)   |
| Association  | 0=PRO does not belong to EARTO  | 0.12 (0.33) | 0=173 (87.8%) |
|  | 1=PRO belongs to EARTO  |             | 1=24 (12.2%)  |

# 3. EXAMPLES OF RESEARCH ORGANISATIONS IN SOME OF THE ATTRIBUTES

|    |                         | CSIC (ES)          | INRA (FR)          | AIT (AT)           |
|----|-------------------------|--------------------|--------------------|--------------------|
|    |                         |                    |                    |                    |
| G1 | Legal status            | Public             | Public             | Private            |
|    | Main owner              | Central government | Central government | Central government |
|    |                         |                    |                    |                    |
| G2 | Number of FOS Filed     | Six                | Four               | Two                |
|    | Main orientation in R&D | Basic research     | Applied research   | Applied research   |
|    |                         |                    |                    |                    |
| G3 | Main Mission            | Produce Knowledge  | Solving Problems   | Creating value     |
|    | EARTO membership        | NO                 | NO                 | YES                |

# 3. DATA AND METHODOLOGY

## MODEL AND ESTIMATION STRATEGY



# 3. DATA AND METHODOLOGY

## MODEL AND ESTIMATION STRATEGY

Two-Step cluster analysis to identify PRO's profiles

(Zhang et al., 1996; Chiu et al., 2001)

### Why...

1. Clustering is one of the most relevant unsupervised learning techniques
2. To determine the optimal number of clusters
3. Use of continuous and categorical variables

# 3. DATA AND METHODOLOGY

## MODEL AND ESTIMATION STRATEGY

Decision tree technique (Breiman et al. 1983)

Exhaustive CHAID proposed by Biggs et al. (1991). A CHAID model will build non-binary trees

## Why...

1. Non-parametric data mining tool that are not constrained to the use of specific types of variables (Attewell and Monaghan, 2015)
2. DT can handle small dataset (Markham et al., 2000)
3. More accurate predictions than regression. No problems with degrees-of-freedom in the dataset (Attewell and Monaghan, 2015)
4. DT aims at making a predictive model for the target variable according to some predictors or input variables



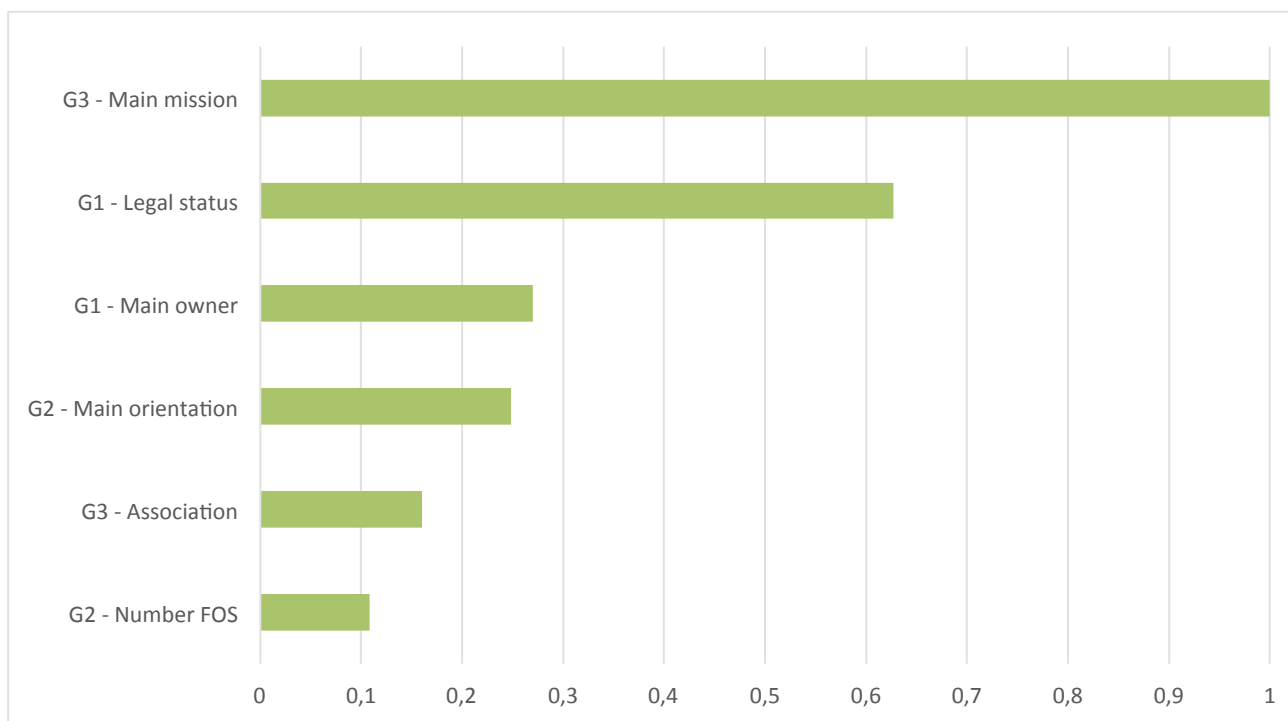
# 4. RESULTS

## TWO-STEP CLUSTER. Cluster distribution

|   |                |          | N   | % of<br>Combined | % of Total |
|---|----------------|----------|-----|------------------|------------|
| Internal institutional<br>characteristics (G1)  | Clusters       | 1        | 78  | 41.1%            | 39.6%      |
|   |                | 2        | 26  | 13.7%            | 13.2%      |
|   |                | 3        | 41  | 21.6%            | 20.8%      |
|   |                | 4        | 45  | 23.7%            | 22.8%      |
|   |                | Combined | 190 | 100.0%           | 96.4%      |
|   | Excluded Cases |          | 7   |                  | 3.6%       |
|   | Total          |          | 197 |                  | 100.0%     |
| (G1)+Resource niche of<br>each RO (G2)          | Clusters       | 1        | 87  | 45.8%            | 44.2%      |
|   |                | 2        | 103 | 54.2%            | 52.3%      |
|   |                | Combined | 190 | 100.0%           | 96.4%      |
|   | Excluded cases |          | 7   |                  | 3.6%       |
|   | Total          |          | 197 |                  | 100%       |
| (G1)+(G2)+Subjective<br>claims of identity (G3) | Clusters       | 1        | 46  | 24.3%            | 23.4%      |
|   |                | 2        | 47  | 24.9%            | 23.9%      |
|   |                | 3        | 41  | 21.7%            | 20.8%      |
|   |                | 4        | 55  | 29.1%            | 27.9%      |
|   |                | Combined | 189 | 100.0%           | 95.9%      |

# 4. RESULTS

## TWO-STEP CLUSTER Predictor variables importance



Least Important

Most important

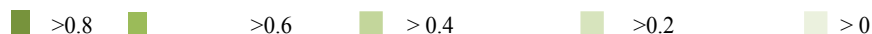
# 4. RESULTS

## TWO-STEP CLUSTER

### Cluster differences

| Cluster |                       | 1                                   | 2                            | 3                                | 4  |
|---------|-----------------------|-------------------------------------|------------------------------|----------------------------------|--|
| Size    | Variables             | 24.3% (46)                          | 24.9% (47)                   | 21.7% (41)                       | 29.1 (55)  |
|         | Main mission (G3)     | Develop knowledge<br>67.4%          | Develop knowledge<br>91.5%   | Generate economic value<br>95.1% | Contribute to solve public policy issues<br>100% |
|         | Legal Status (G1)     | NPI<br>100%                         | Public<br>72.3%              | NPI<br>61%                       | Public<br>98.2%                                  |
|         | Main Owner (G1)       | Regional-local government<br>(58.7) | Central government<br>(78.7) | Private Company<br>(36.6)        | Central Government<br>(92.7)                     |
|         | Main orientation (G2) | BR 60.9%                            | BR 59.6%                     | ED 51.2%                         | AR 56.4%   |
|         | Association (G3)      | 0<br>97.8%                          | 0<br>93.6%                   | 0<br>56.1%                       | 0<br>96.4%                                       |
|         | Number of FOS (G2)    | 2<br>50%                            | 1<br>31.9%                   | 1<br>43.9%                       | 3<br>29.1%                                       |

Predictor importance



# 4. RESULTS

## TWO-STEP CLUSTER Taxonomy of PROs

|                             | <b>Cluster 1 – Hybrids<br/>(HYB)</b> | <b>Cluster 2 - Research<br/>councils<br/>(REC)</b> | <b>Cluster 3 – Technology<br/>oriented centers<br/>(TOC)</b> | <b>Cluster 4 – Government<br/>laboratories<br/>(GOL)</b> |
|-----------------------------|--------------------------------------|--|--|--|
| Main mission                | Develop knowledge                    | Develop knowledge                                  | Generate Economic value                                      | Contribute to solve public policy issues                 |
| Legal status                | Non-profit institution               | Public   | Non-profit institution                                       | Public   |
| Main owner                  | Regional Governments                 | Central government                                 | Private companies  | Central Government                                       |
| Orientation of R&D activity | Basic/applied research               | Basic/applied<br>Research                          | Experimental development                                     | Applied research   |
| Association RTOs            | No                                   | No   | Yes  | No   |
| Multidisciplinarity         | Yes                                  | Yes  | No   | Yes  |

# 4. RESULTS

## TWO-STEP CLUSTER Taxonomy of PROs



European Forum for  
Studies of Policies  
for Research and Innovation

|                       |                                       | Group 1-<br>HYB     | Group 2-<br>REC   | Group 3-<br>TOC     | Group 4-<br>GOL   | Total          |             |
|-----------------------|---------------------------------------|---------------------|-------------------|---------------------|-------------------|----------------|-------------|
|                       |                                       | Count/<br>mean      | Count/<br>mean    | Count/<br>mean      | Count/<br>mean    | Count<br>/mean | P-<br>value |
| G1 - Legal status     | Public                                | 0 <sup>3</sup>      | 34 <sub>d</sub>   | 3                   | 54 <sub>c,d</sub> | 91             | <0.001      |
|                       | Private                               | 0 <sup>3</sup>      | 5                 | 12 <sub>c</sub>     | 0 <sup>3</sup>    | 17             |             |
|                       | NPI                                   | 46 <sup>3</sup>     | 2                 | 25 <sub>c,e</sub>   | 1                 | 74             |             |
|                       | Other                                 | 0 <sup>3</sup>      | 6                 | 1                   | 0 <sup>3</sup>    | 7              |             |
| G1 - Main owner       | Central government                    | 12                  | 37 <sub>b,d</sub> | 13                  | 51 <sub>b,d</sub> | 113            | <0.001      |
|                       | Regional-local government             | 27 <sub>c,d,e</sub> | 8                 | 10                  | 4                 | 49             |             |
|                       | Private company                       | 4                   | 0 <sup>3</sup>    | 15 <sub>b</sub>     | 0 <sup>3</sup>    | 19             |             |
|                       | HEI                                   | 0 <sup>3</sup>      | 1                 | 2                   | 0 <sup>3</sup>    | 3              |             |
|                       | Other                                 | 3                   | 1                 | 1                   | 0 <sup>3</sup>    | 5              |             |
| G2 - Main orientation | Basic Research                        | 28 <sub>d</sub>     | 28 <sub>d</sub>   | 1                   | 22 <sub>d</sub>   | 89             | <0.001      |
|                       | Applied Research                      | 18                  | 16                | 19                  | 31                | 84             |             |
|                       | Experimental Development              | 0 <sup>3</sup>      | 3                 | 21 <sub>c,e</sub>   | 2                 | 26             |             |
| G2 - Number FOS       |                                       | 2                   | 3 <sub>b</sub>    | 2                   | 3 <sub>b</sub>    | 2.5            | 0.004       |
| G3 - Main mission     | Develop knowledge                     | 31                  | 43 <sub>b</sub>   | 0 <sup>3</sup>      | 0 <sup>3</sup>    | 74             | <0.001      |
|                       | Contribute solve public policy issues | 15 <sub>d</sub>     | 0 <sup>3</sup>    | 2                   | 55 <sup>3</sup>   | 72             |             |
|                       | Generate economic value               | 0 <sup>3</sup>      | 4                 | 39 <sub>c</sub>     | 0 <sup>3</sup>    | 43             |             |
| G3 - Association      | 0                                     | 45 <sub>d</sub>     | 44 <sub>d</sub>   | 23                  | 53 <sub>d</sub>   | 165            | <0.001      |
|                       | 1                                     | 1                   | 3                 | 18 <sub>b,c,e</sub> | 2                 | 24             |             |

# 4. RESULTS

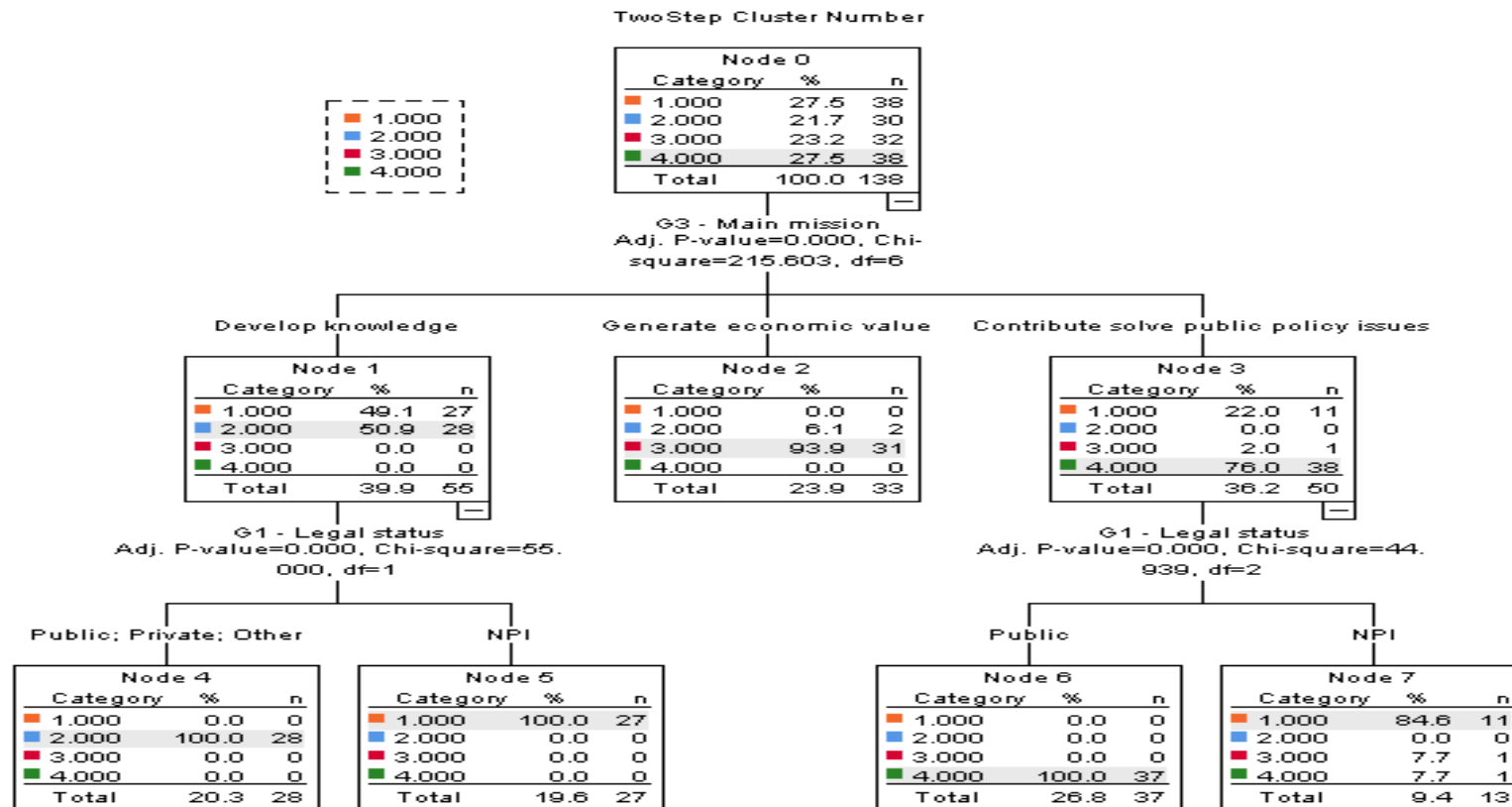
## EXHAUSTED CHAID DECISION TREE

### Classification results

| Sample | Type of PRO Observed | Predicted |         |         |         |          | Total | Risk       |  |
|--------|----------------------|-----------|---------|---------|---------|----------|-------|------------|--|
|        |                      | 1 - HYB   | 2 - REC | 3 - TOC | 4 - GOL | Estimate |       | Std. Error |  |
| ing    | 1 - HYB              | 38        | 0       | 0       | 0       | 100.0%   | 0.029 | 0.014      |  |
|        | 2 - REC              | 0         | 28      | 2       | 0       | 93.3%    |       |            |  |
|        | 3 - TOC              | 1         | 0       | 31      | 0       | 96.9%    |       |            |  |
|        | 4 - GOL              | 1         | 0       | 0       | 37      | 97.4%    |       |            |  |
|        | Overall %            | 29.0%     | 20.3%   | 23.9%   | 26.8%   | 97.1%    |       |            |  |
|        | 1 - HYB              | 8         | 0       | 0       | 0       | 100.0%   | 0.098 | 0.042      |  |
|        | 2 - REC              | 2         | 13      | 2       | 0       | 76.5%    |       |            |  |
|        | 3 - TOC              | 1         | 0       | 8       | 0       | 88.9%    |       |            |  |
|        | 3 - GOL              | 0         | 0       | 0       | 17      | 100.0%   |       |            |  |
|        | Overall%             | 21.6%     | 25.5%   | 19.6%   | 33.3%   | 90.2%    |       |            |  |

There are significant differences between means (metric variables) and categories (categorical variables) in all the variables in the analysis at 99% regarding the taxonomy of centres.

# EXHAUSTED CHAID DECISION TREE (BASED ON THE TRAINING DATASET)



Note: The categories of the dependent variables have been identified in the clustering step as 1: HYB, 2: RECs, 3: TOCs, 4: GOLs

# 5. CONCLUSIONS

Growing heterogeneity among public organisations which develop R&D

Traditional categories of research organizations are not useful to address and represent the dynamism of this population.

Establishing new categories allows to capture a detail representation of the PROs and to open a window for partial membership.

Policy-makers can not propose tailored policies for homogenous groups when traditional typologies are not able to identify the current panorama of public research organizations.

The design of an updated taxonomy opens a venue for improving the adjustment between policy-making and the actors subjected to the policy.



# 5. CONCLUSIONS

Design a classification tool for PROs to derive homogeneous categories;



Describe the attributes (organisational features) that characterised the different types

- Taxonomy according to organisational features to make comparisons with theoretical categorisations: structure, role played by the institution, R&D main orientation, legal status, ownership, FOS... **Two-step Cluster analysis:**

Group 1 – GOL | Group 2 – NAH | Group 3 – REC | Group 4 - TOC | Group 5 –REH

- **Exhaustive CHAID classification tree** to check the power of the classification.

# 5. CONCLUSIONS

| REC   | TOC   | REH   | GOL   | NAH   |
|---|---|---|---|---|
| <input type="checkbox"/>                              | <input type="checkbox"/>  | <input type="checkbox"/>                            | <input type="checkbox"/>  | <input type="checkbox"/>                            |
| <input type="checkbox"/> Largest organisations        | <input type="checkbox"/> Small organisation                                       | <input type="checkbox"/> Medium organisations       | <input type="checkbox"/> Medium organisation                      | <input type="checkbox"/> Small organisations        |
| <input type="checkbox"/> Direct power of central Gov. | <input type="checkbox"/> Important role of other actors: Reg. Gov /HEIs / Private | <input type="checkbox"/> Reg. Gov                   | <input type="checkbox"/> Central Gov.                             | <input type="checkbox"/> Direct power central Gov.  |
| <input type="checkbox"/> Basic and applied research   | <input type="checkbox"/> ED   | <input type="checkbox"/> Basic and applied research | <input type="checkbox"/> Basic and applied research               | <input type="checkbox"/> Basic and applied research |
| <input type="checkbox"/> Public legal status          | <input type="checkbox"/> NPI  | <input type="checkbox"/> NPI                        | <input type="checkbox"/> Public                                   | <input type="checkbox"/> Public / NPI               |
| <input type="checkbox"/> Develop knowledge            | <input type="checkbox"/> Generate economic value                                  | <input type="checkbox"/> Develop knowledge          | <input type="checkbox"/> Contribute to solve public policy issues | <input type="checkbox"/> Develop knowledge          |
| <input type="checkbox"/> R&D performer & funder       | <input type="checkbox"/> Performer  | <input type="checkbox"/> Performer                  | <input type="checkbox"/> Performer                                | <input type="checkbox"/> Performer                  |
| <input type="checkbox"/> Multidisciplinary            | <input type="checkbox"/> Non-multidisciplinar                                     | <input type="checkbox"/> Non-multidisciplinar       | <input type="checkbox"/> Non-multidisciplinar                     | <input type="checkbox"/> Non-multidisciplinar       |
| <input type="checkbox"/> Old organisations            | <input type="checkbox"/> Young organisations                                      | <input type="checkbox"/> Young organisations        | <input type="checkbox"/> Old organisations                        | <input type="checkbox"/> Old organisations          |

**Good tool for classification based on decision trees with a rate of recognition about 87% of the cases in the sample**