

**RISIS WP20 MEETING**

**January, 28<sup>th</sup> - 2015**

**Roma**

 **POLITECNICO DI MILANO**



**WP20 – Group meeting**

# Agenda

- FGMF database
  - Identification of the population of mid-sized firms
  - Preliminary statistics on FGMFs
- Key issues and next steps of WP20



**RISIS**  
Research infrastructure for research  
and innovation policy studies



# Data Collection

- As a first step, we identified the potential population of **Mid/Intermediary-sized firms**
- Starting point: Turnover / Balance Sheet criteria:
  - turnover lower than €1,5 billion OR
  - balance sheet total lower than €2 billion
- Observation periods start in 2008, 2009 and 2010
- Data source: 2008, 2009, 2010 On-Line versions of ORBIS
- Subsidiaries and parent companies located in Europe 28 + Israel
  
- **Using the above criteria the number of firms identified was 13,034,346**



# Data Collection - cleaning

- All the observations for which the turnover in 2008, 2009 and 2010 is missing have been removed
- All the observations for which the BvD ID code (companies unique identification numbers in ORBIS database, BVD Id nr. = Bureau van Dijk ID number) is missing have been removed
- Final number of firms in the preliminary list was **8,609,209**



# Data Collection

- Main issue for the identification of the population of mid-sized firms:
  - **missing data about employees on about 40% of firms**
- For these firms, we use a predicted value of the number of employees
  - $\text{Log}(\text{Employees}) = f(\text{Log}(\text{Turnover}), \text{age of the company}, \text{industry dummies}, \text{country dummies and year dummies})$
  - Tobit regression



# Data Collection

- Selection of the final population of European Mid/Intermediary sized firms applying the employees threshold:

**50 < Number of Employees < 4999**

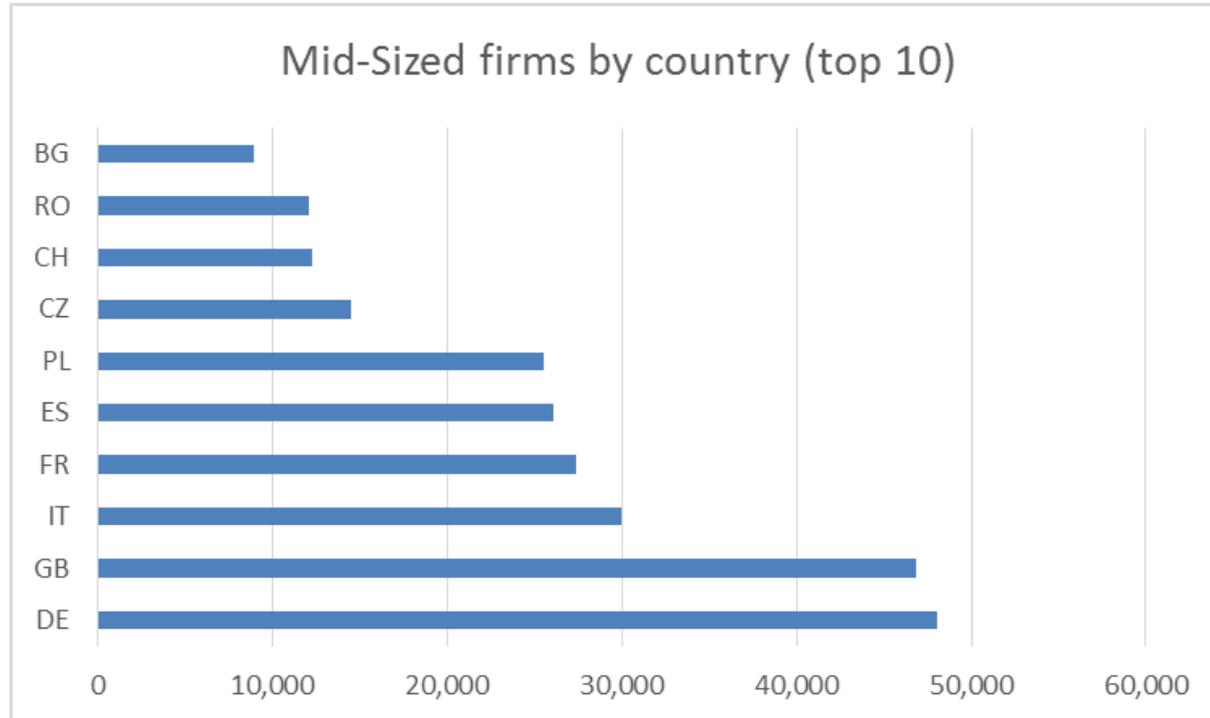
| Year | N. Firms with known employees | N. firms with predicted employees | Total   |
|------|-------------------------------|-----------------------------------|---------|
| 2008 | 210,536                       | 41,044                            | 251,580 |
| 2009 | 227,655                       | 39,177                            | 266,832 |
| 2010 | 230,829                       | 46,089                            | 276,918 |

- N. of firms = **332,553 unique BvD IDs**



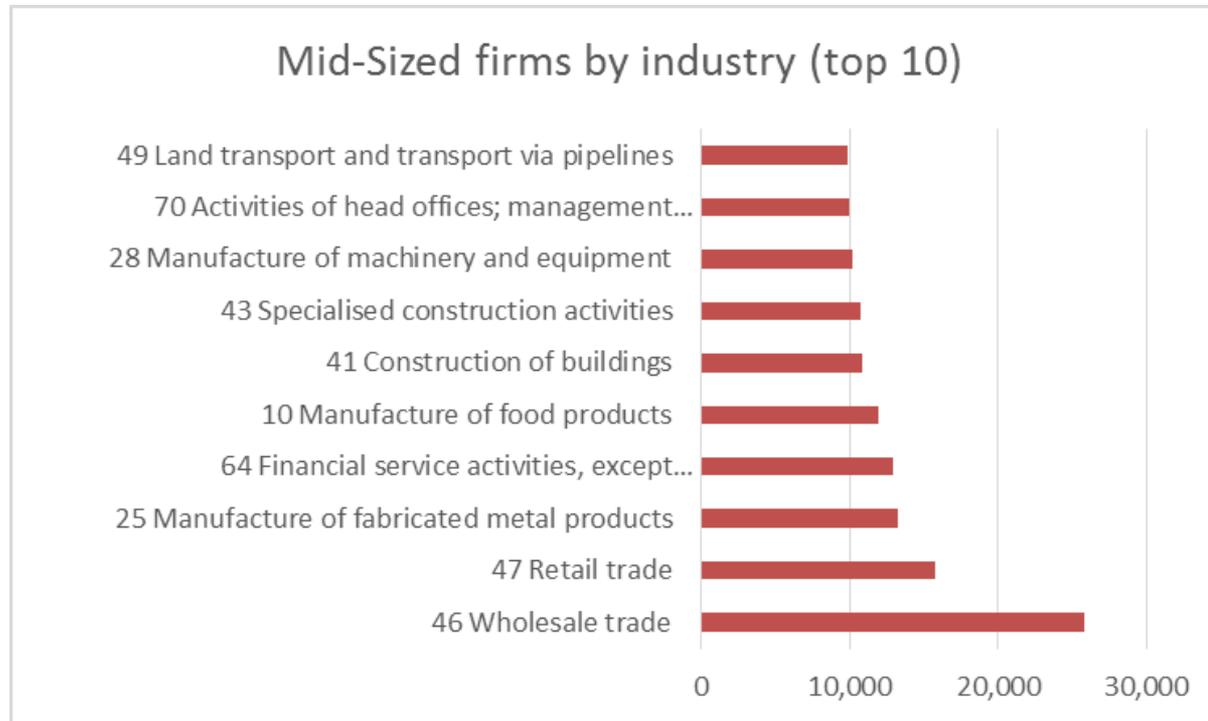
**RISIS**  
Research infrastructure for research  
and innovation policy studies

# Mid-sized firms by country



**RISIS**  
Research infrastructure for research  
and innovation policy studies

# Mid-sized firms by industry



**RISIS**  
Research infrastructure for research  
and innovation policy studies

# Selection errors

- We report observations when both actual and predicted values of employees are known in order to estimate the proportion of observations that have been predicted correctly (or not)

| <b>A</b> | N. observations <b>in the threshold</b> according to both actual and predicted employees                      | 326,367   |
|----------|---|-----------|
| <b>B</b> | N. observations <b>outside the threshold</b> according to both actual and predicted employees                 | 9,588,338 |
| <b>C</b> | N. observations <b>in the threshold</b> according to predicted employees (but <b>NOT</b> to actual employees) | 66,457    |
| <b>D</b> | N. observations <b>in the threshold</b> according to actual employees (but <b>NOT</b> to predicted employees) | 300,754   |

- $A + B =$  good prediction
- $C + D =$  bad prediction



**RISIS**  
 Research infrastructure for research  
 and innovation policy studies



# Identifying FGMFs

- OECD/EUROSTAT definition:

*All enterprises with **average annualised growth greater than 20% per annum**, over a **three year period** should be considered as high-growth enterprises. Growth can be measured by the number of **employees** or by **turnover**.*



**RISIS**  
Research infrastructure for research  
and innovation policy studies

# Identifying FGMFs

- **Calculation:** Size threshold of 10 or more employees at the beginning of the period (t-3)
  - Turnover growth:  $\left(\frac{\text{turnover}_t}{\text{turnover}_{t-3}}\right)^{\frac{1}{3}} - 1 > 0.2$
  - Employees growth:  $\left(\frac{\text{employees}_t}{\text{employees}_{t-3}}\right)^{\frac{1}{3}} - 1 > 0.2$
- **Exclusions:**
  - Firms born in t-3 (the foundation date can be at the end of the year)
  - If growth in the number of employees or turnover was due to M&A, the firm should not be considered a HGF



| TURNOVER  | Number of Fast Growing Firms |                          |
|-----------|------------------------------|--------------------------|
|           | Total                        | Cases without imputation |
| 2008-2011 | 25,512                       | 19,535                   |
| 2009-2012 | 31,306                       | 24,311                   |
| 2010-2013 | 17,182                       | 13,504                   |

| EMPLOYEES | Number of Fast Growing Firms |        |
|-----------|------------------------------|--------|
|           | 2008-2011                    | 17,383 |
| 2009-2012 | 16,252                       |        |
| 2010-2013 | 10,719                       |        |

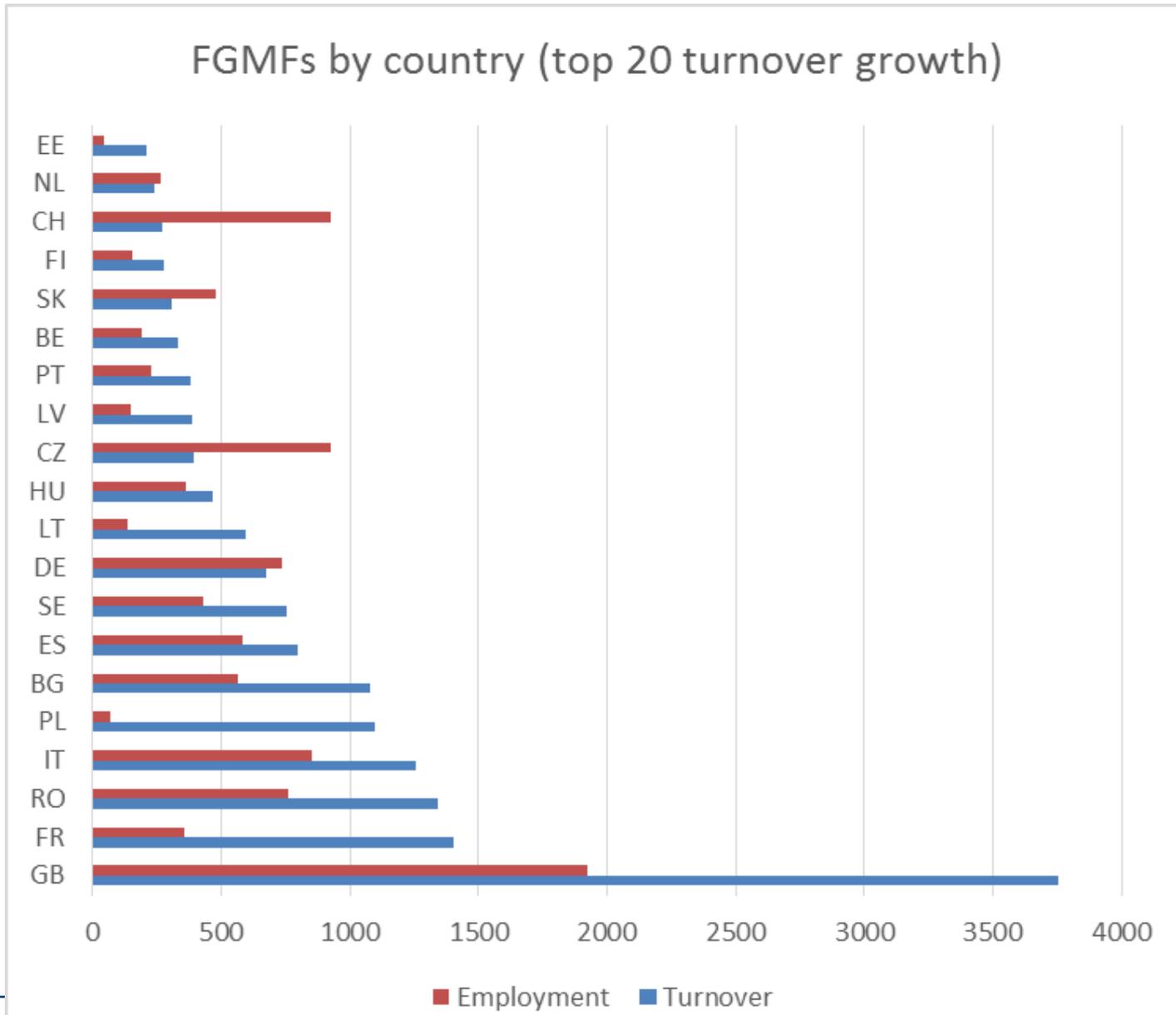


# FGMFs - Correlations

|   |                       | 1             | 2             | 3             | 4      | 5      | 6      |
|---|-----------------------|---------------|---------------|---------------|--------|--------|--------|
| 1 | Turnover Growth 08-11 | 1.0000        |               |               |        |        |        |
| 2 | Turnover Growth 09-12 | 0.4829        | 1.0000        |               |        |        |        |
| 3 | Turnover Growth 10-13 | 0.1211        | 0.4215        | 1.0000        |        |        |        |
| 4 | Empl. Growth 08-11    | <b>0.3101</b> | 0.225         | 0.1207        | 1.0000 |        |        |
| 5 | Empl. Growth 09-12    | 0.2232        | <b>0.3547</b> | 0.2923        | 0.4643 | 1.0000 |        |
| 6 | Empl.Growth 10-13     | 0.1061        | 0.2444        | <b>0.4039</b> | 0.1803 | 0.4567 | 1.0000 |



**RISIS**  
Research infrastructure for research  
and innovation policy studies



## Key issues and next steps of WP20



**RISIS**  
Research infrastructure for research  
and innovation policy studies

# Key issues

- Previous statistics report:
  - Consolidated data for groups
  - Data on subsidiaries
  - Possible duplications
- Unit of analysis of the FGMF database is the **business group** (i.e. subsidiaries need to be consolidated)
- In the October meeting in Milan, we agreed to keep in the FGMF database:
  - Consolidated data for groups with parent company's headquarter in an European country (+ Israel)
  - Data at the subsidiary-level only for firms whose parent company is headquartered outside Europe
- **Need to define a methodology for the identification of business groups**



# Key issues

- Different definitions of high-growth (by industry?)
- Data on scientific publications (UoS)
- Next steps according to 2015 activities:
  - Data collection on legal and financial information for FGMFs (POLIMI)
  - Data collection on VC-backed firms in Israel (SNI)
  - Data collection on legal and financial information for VICO firms (POLIMI)
  - Data collection on legal and financial information for CIB firms (POLIMI)
  - Data collection on patents for VICO, CIB and FGMFs (UPEMLV)
  - Data collection on scientific publications for VICO, CIB and FGMFs (UoS)
  - Harmonization of geographical information and names for VICO, CIB and FGMFs according to WP8 and WP9 (UoS)

