Indicators for Universities and Higher Education

Andrea Bonaccorsi, Cinzia Daraio***

*University of Pisa - Italy

** IIT-CNR and Scuola Superiore Sant'Anna, Pisa - Italy

Second Workshop on Indicators in the Knowledge Economy

Maastricht, 6-7 October 2005

Outline

- Methodological setting
- The AQUAMETH project
- The comparability issue: problems and possible solutions
- Exploiting the database
- Bridging the gap: towards “meso” data on the European Higher Education System

Methodological setting

To address the data constraint in University and Higher Education some methodological discussions and choices have been done:

- University as the appropriate unit of analysis (*microfoundation*, universities as *strategic units*, Bonaccorsi and Daraio, 2006).
- Profiling of universities
- Using techniques which allow the estimation of *complementarities* (non-linearities, external influence), e.g.,
  - trade-offs between research and teaching,
  - between undergraduate and postgraduate teaching,
  - between publication and patenting,
  - between research and third mission activities:

  here we need to estimate substitution vs. complementarity effects that may not be stable across all the relevant distribution.

- Traditional econometric techniques (production function as the method used to parametrically estimate an input-output relation, connected to a linear model of innovation) are unsatisfactory.

Methodological setting (2)

- Universities are multi-input multi-output production units in a non-market context. We do not have a theory which tells us how to weight the several inputs and outputs

- We adopt techniques which are:
  - Nonparametric (we do not want to specify a functional relation between inputs and outputs)
  - Robust with respect to the curse of dimensionality and outliers problem
  - Allow multiple-inputs multiple-outputs comparison
  - Allow the measurement of conditional effects, and the impact of external factors (which may affect the performance- but we do not know in advance if they are input or output) in the assessment of the performance
  - We choose a robust nonparametric approach in frontier analysis (Cazals, Florens and Simar, 2002, JE; Daraio and Simar, 2005 JPA)

The AQUAMETH project

1. A simple framework for the production of HEI

2. Data collection strategy (a compromise between methodological choices and the data availability)

3. Classification of the variables at microlevel

4. Coverage
   - *Countries* : Italy, Spain, Norway, Portugal, Switzerland and UK (270 universities)
   - *Data availability*: 1995-2003 (with some breaks)
   - *Planned extension* to France, Germany, Netherlands, Hungary (2006) under Aquameth 2

5. Potential for addressing policy relevant issues

Ongoing research projects

CHINC

Institute for Prospective Technology Studies (IPTS), funding agency.

A project aimed at collecting quantitative data and qualitative observation on **university funding** across Europe.

Sample of universities in each of the 10 European countries covered. Direct survey with University top officials.

Classification of the variables at micro level

AREA	CATEGORIES
General information	• year of foundation • city, province, region (NUTS) • number and type of faculties/schools/disciplines covered • governance (public, private) • type (university, technical college) • other relevant historical information.
Revenues	• Total revenues of the university. • General budget of the university (in federal countries divided between national and regional appropriations). • Tuition and Fees. • Grants and contracts, if possible divided between government, international, private and private non-profit. • Other revenues.
Expenditures	• Total expenditures (excluding investments and capital costs). • Personnel expenditures, if possible divided between personnel categories. • Other expenditures.
Personnel	• Total staff (FTE or headcount). • Professors. • Other academic staff. • Technical and administrative staff.
Education production	• Number of undergraduate students. • Number of undergraduate degrees. • Number of PhD students. • Number of PhD degrees.
Research and technology production	• ISI publications. • Technological production indicators.

The comparability issue

problems (and possible solutions)

Three main categories of comparability problems:

1. Differences in the organization and governance structure of national HE systems (**institutional context**)
    1. Dual systems (some European countries have an HES which includes universities and a range of non-phd awarded institutions- separate analysis)
    2. Private vs. public universities (dummy variable)
    3. Public research organizations (blurred boundaries with universities, e.g. CNRS)
        1. Possible solution: analysis of co-publications, estimation of coefficients, sensitivity analysis

The comparability issue (2)

problems (and possible solutions)

4. Age and structure of universities

Possible solutions:

- Add qualitative information in the database and consider in interpretation
- sensitivity analysis
- external variables

5. Funding pattern

possible solutions:

- composition ratios: percentages of each category of total funds

The comparability issue (4)
problems (and possible solutions)

Possible solutions for subject mix:

1. Dummy variables

2. Categorization (a concentration index by computing the distribution of students in four broad disciplinary areas (Human and Social Sciences, Technical Sciences, Natural Sciences, Medicine) C1 is the concentration index for the largest discipline, C2 for the first two, etc. We define Specialist a university with C1>= 0.50 or C2>= 0.75 and Generalist otherwise. Other specifications can be explored)

3. External variables

4. Test of hypothesis

5. Multi-layer models (Sarrico and Dyson, 2004)

The comparability issue (3)

problems (and possible solutions)

2. **Heterogeneity of the individual HEI**

subject mix

- cost structure
- research intensity
- bibliometric indicators

are very different according to the domain

In AQUAMETH we focus on four broad disciplinary areas:

- Human and Social Sciences,
- Technical Sciences,
- Natural Sciences,
- Medicine

The comparability issue (5)
problems and possible solutions

3. **Data problems**

1. Expenditures
    1. Inclusion of annex services, social security payments, separation of healthcare expenditures in universities with hospital
    2. Capital expenditures (in some countries included in the state accounts)
2. Physical capital stock
    1. limited data on seats in classroom, linguistic and computer laboratories
    2. no data on laboratories equipment
3. Funding
    1. contract funding (e.g. UK)

The comparability issue (6)
problems and possible solutions

4. Staff (classification: broad definition)
   - in some countries only headcount,
   - not all staff is full time,
   - PhD students' role

5. Research and technology production
   - PhD,
   - limits of ISI data on publications

6. Education production
   - quality of education

Possible solutions

- Expert judgment
- Multiple variables
- Estimation of proportions and sensitivity analysis

Exploiting the database

1. Comparative analysis of funding sources
    1. Total expenditures per student in real terms
    2. Composition of revenues
2. Learning from a national case (Filippini and Lepori, 2006)

3. Some empirical evidence on the Italian case using the robust nonparametric approach in efficiency analysis (for the comparative analysis read the AQUAMETH book!):

    1. Size effects (Bonaccorsi, Daraio and Simar, 2005a)
    2. Trade-off between research output and third mission output (Bonaccorsi, Daraio and Simar, 2005a)
    3. PhD students: inputs or outputs? (Bonaccorsi, Daraio and Simar, 2005b)

Data description

INPUTS	***Description***
Human capital	
TOTDOC	Sum of full professors, associate professors assistant professors and researchers
	(average - academic years 1995/96-97/98)
TECHADM	Number of Technical and Administrative staff
	(average - academic years 1995/96-97/98)
Financial capital	
CUMEXP	Total cumulated expenses years 1995-1999
	(in million of Italian lire)
CUMEXp100iscr	Cumulated expenses per 100 enrolled students
CUMEXp1doc	Cumulated expenses per 1 scholar
Physical capital	
SPACE	Number of places in the lecture-halls
	(average - academic years 1995/96-97/98)
OUTPUTS	
Research	
PUB	Cumulated sum of publications years 1995-1999
CIT	Cumulated sum of citations years 1995-1999
PUBp100doc	No. of publications per 100 scholars (average values)
Teaching	
LAUCUM	Cumulated sum of graduated (and with *diploma*) academic years 1995/96-97/98
LAUCUMp100iscr	No. of graduated per 100 enrolled students (average values)
EXTERNAL FACTORS	
ISCR	No. of enrolled students - average academic years 1995/96-97/98
FACULT	No. of schools within the university
LOAD	No. of curricula (or *courses of specialisation*) activated per 100 scholars
CITPUB	Ratio of CIT over PUB
TRASFPRIV	Percentage of private contracts over total university budget

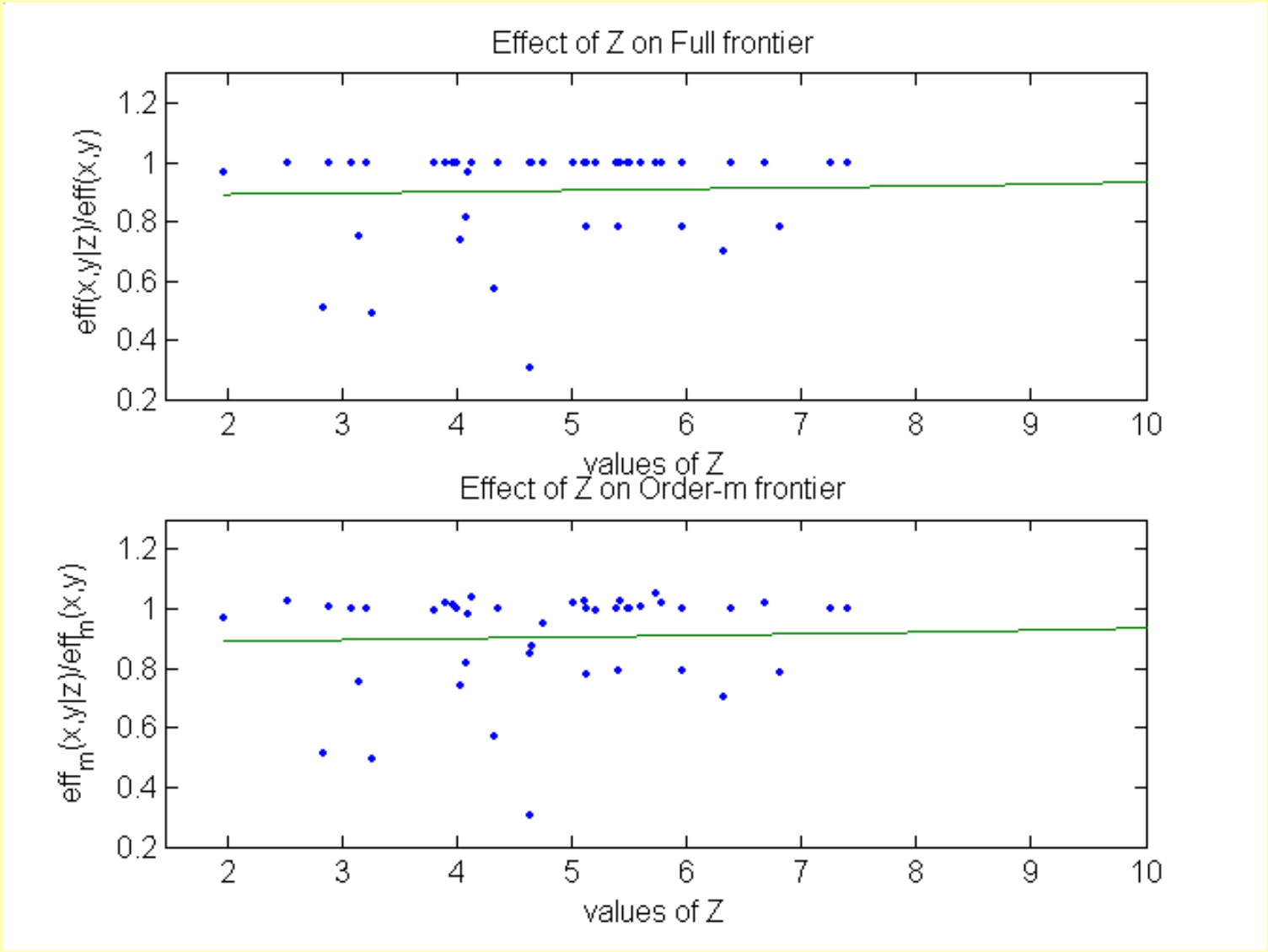
Advantages of Robust nonparametric techniques

- Nonparametric: no need of a functional specification of the input-output relation
- Robust to extremes/noise in the data (we can set the level of robustness, m dual meaning)
- Good statistical properties: rate of convergence as parametric estimators ($n^{1/2}$), no curse of dimensionality (Cazals, Florens and Simar, 2002JE)
- Investigation on the global influence of external-environmental variables on overall efficiency (Daraio and Simar, 2005aJPA)
- Allow the decomposition of conditional efficiency at individual level (Daraio, 2003;Daraio and Simar, 2005aJPA, 2005b)
- Permit parametric approximations that lead robust estimates of elasticity of substitutions among inputs and other economic parameters (Florens, Simar 2005JE)

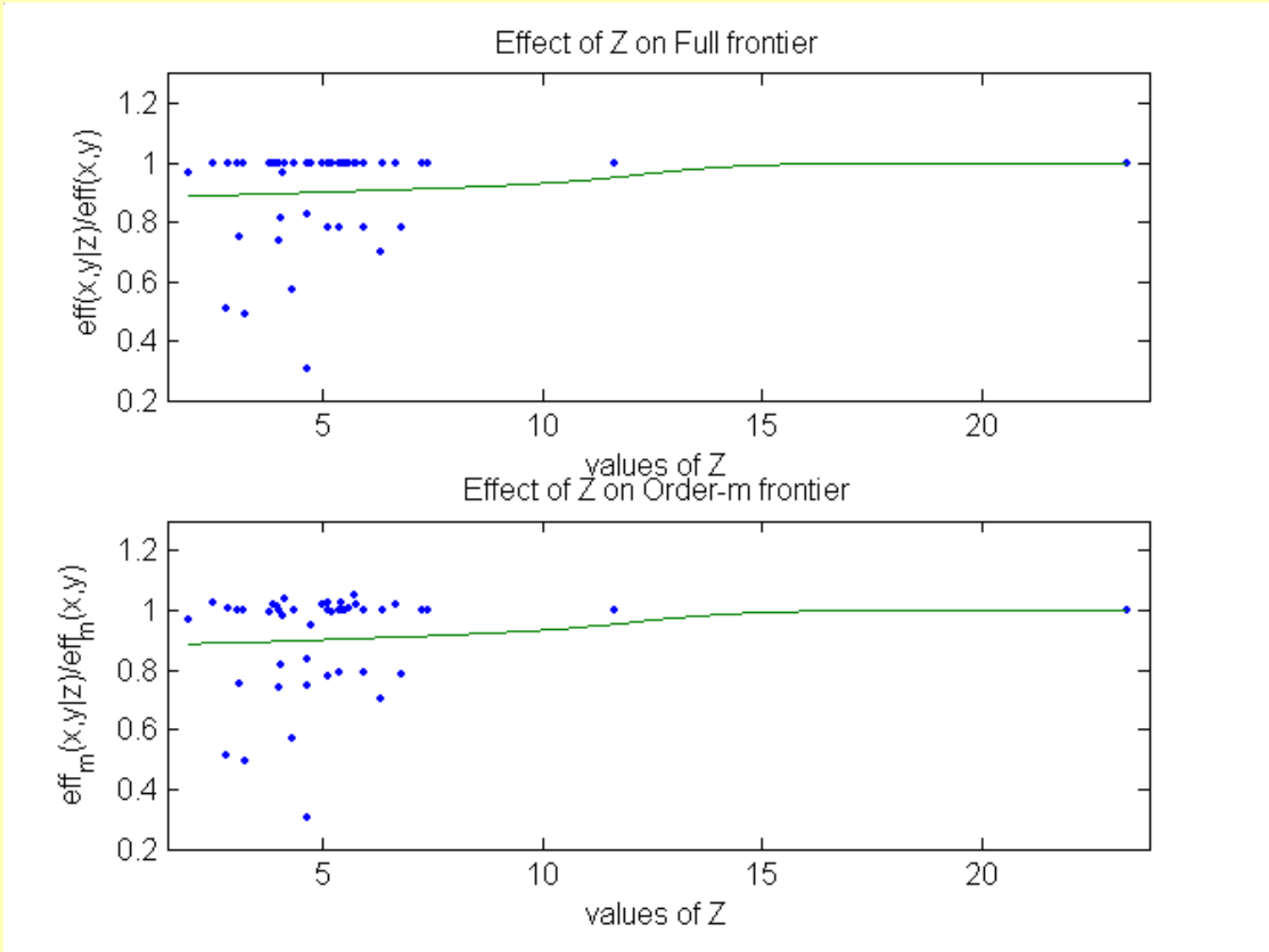
Main results

- ***Empirical evidence on Economics of scale and scope:***
  - They are not significant factors in explaining research and education productivity
- ***Trade-off research vs. teaching:***
  - We find that increasing scientific quality (cit/pub) improves educational efficiency (laucum)
  - On the other hand, a good educational efficiency (high laucum) does not deteriorate research efficiency (pub)
- ***Trade-off publication vs. industry oriented research:***
  - We find that complementarity/rivalry relation has *local effects* (inverted U-shaped pattern), i.e., initially, collaboration with industry may improve productivity (pub), but beyond a certain level the compliance with industry expectations may be too demanding and deteriorate the publication profile

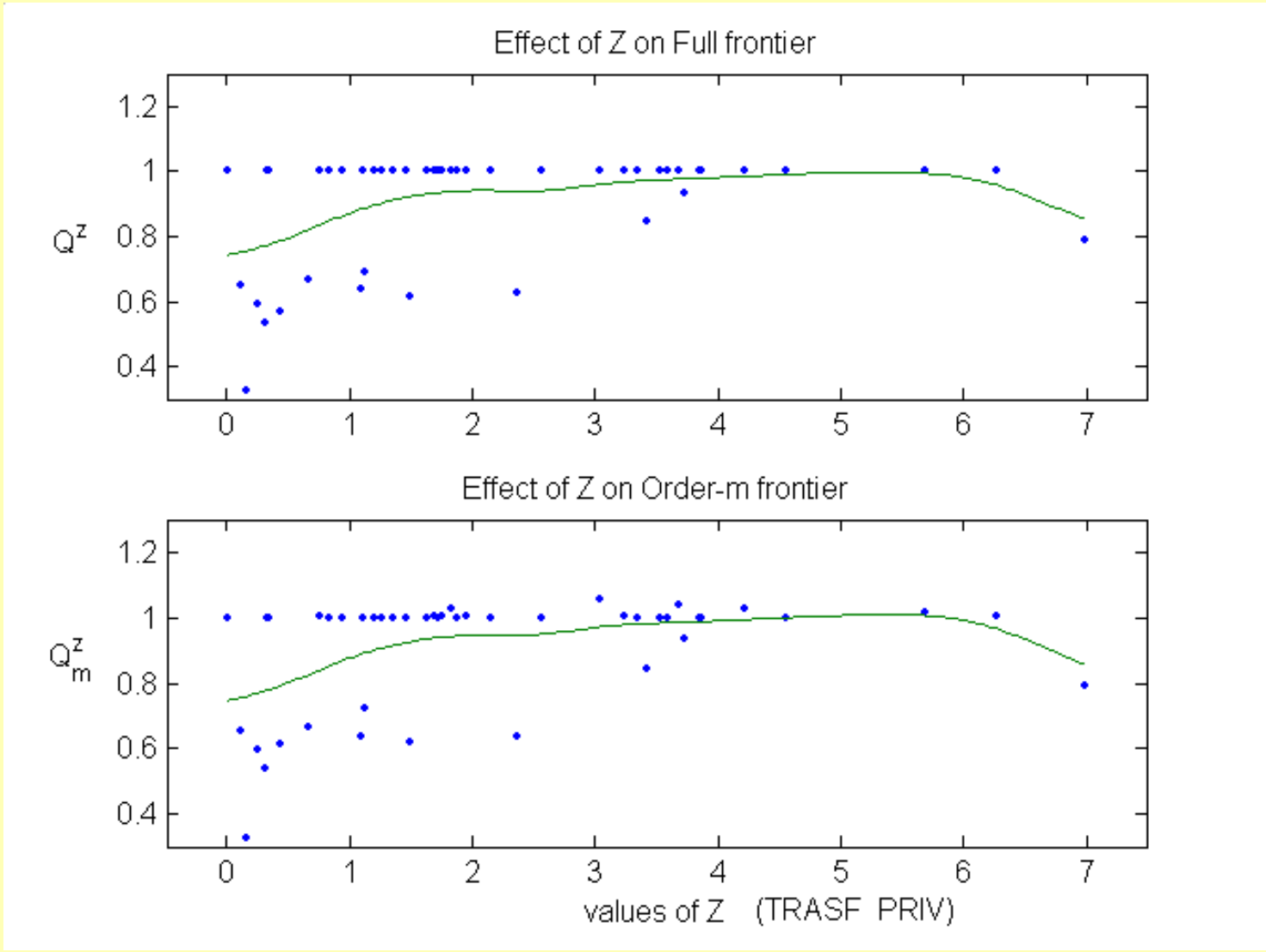
Mod teach Z facult



Mod teach Z cit/pub



Trade-off publication vs. industry oriented research



Trade-off publication vs. contract research

Empirical evidence by discipline

Year 2000-2001 (average)

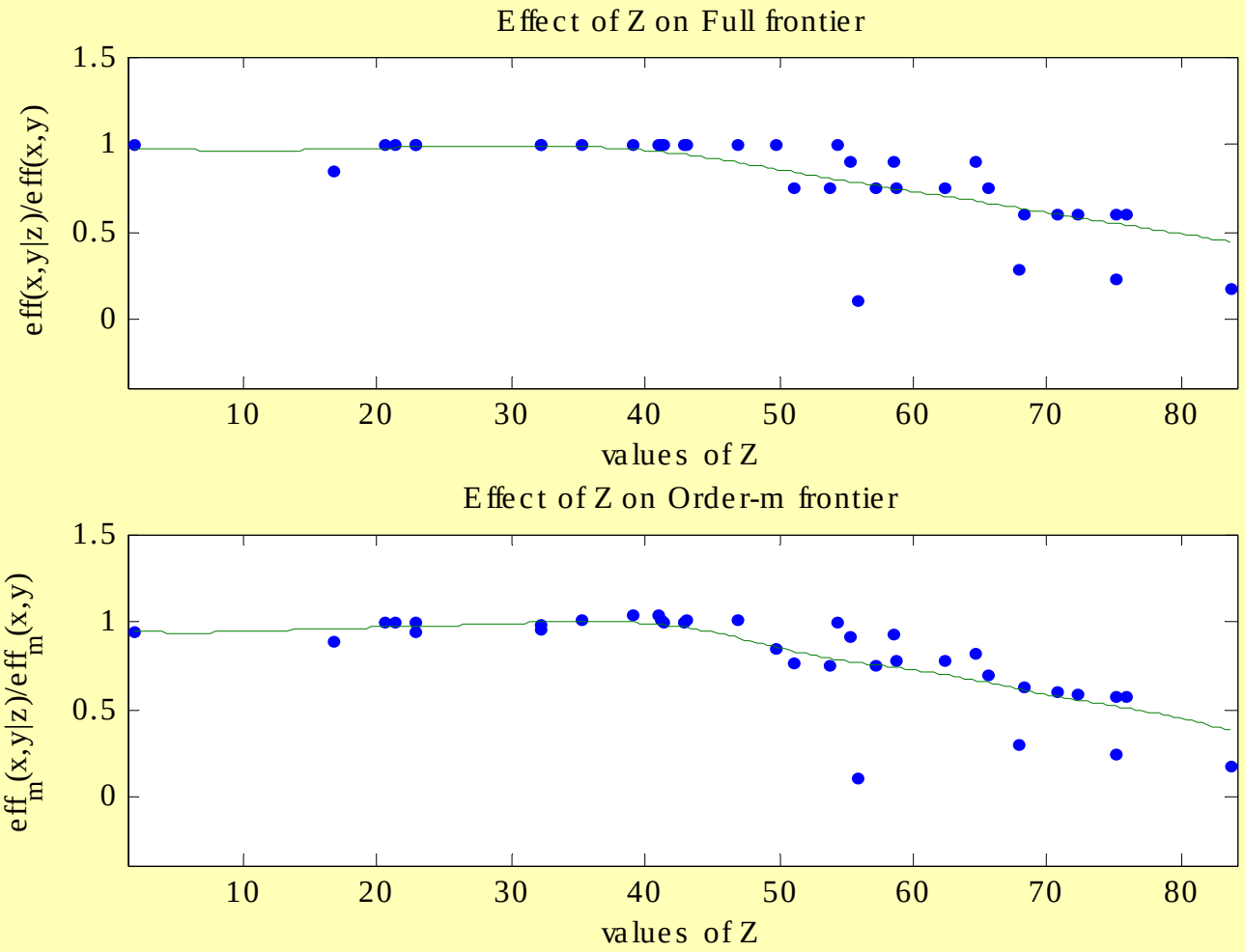
Source: CNVSU

Transfpriv includes all research funding different from government funding (e.g. government call, EU research, private contracts, grants)

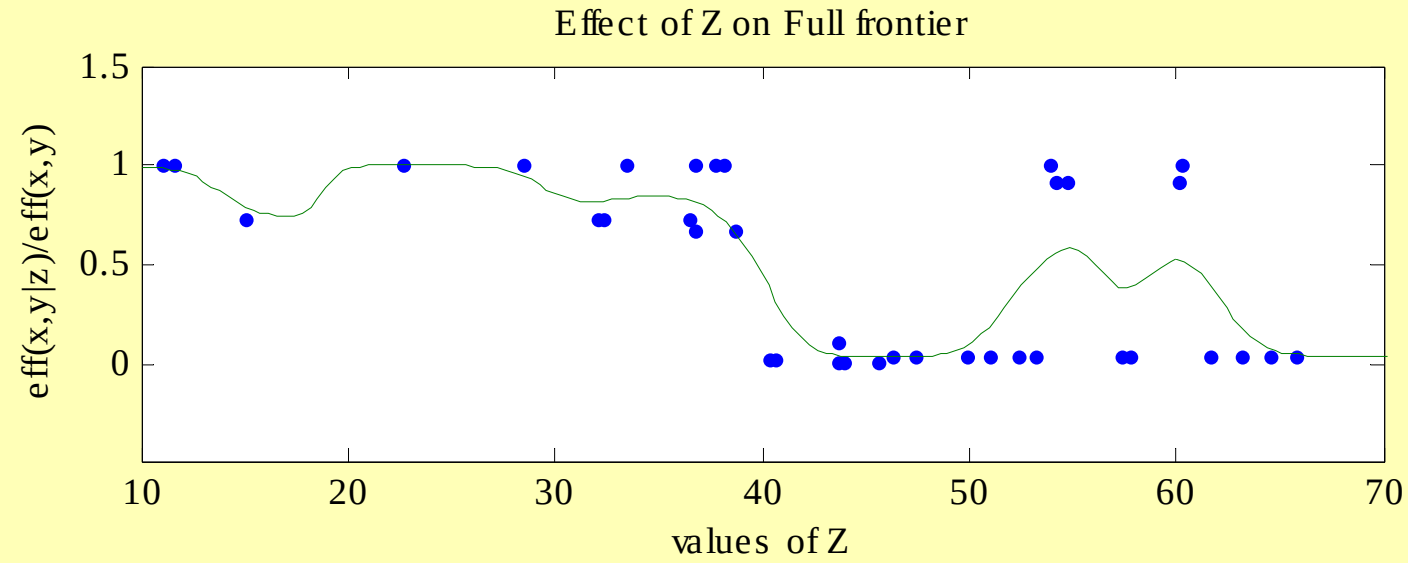
- We find that complementarity/rivalry relation has *local effects*
- Maybe, initially, collaboration with industry may improve productivity (pub), but beyond a certain level the compliance with industry expectations may be too demanding and deteriorate the publication profile
- This inverted U-shaped pattern has to be confirmed adding more evidence
- New data from CRUI (survey Network)

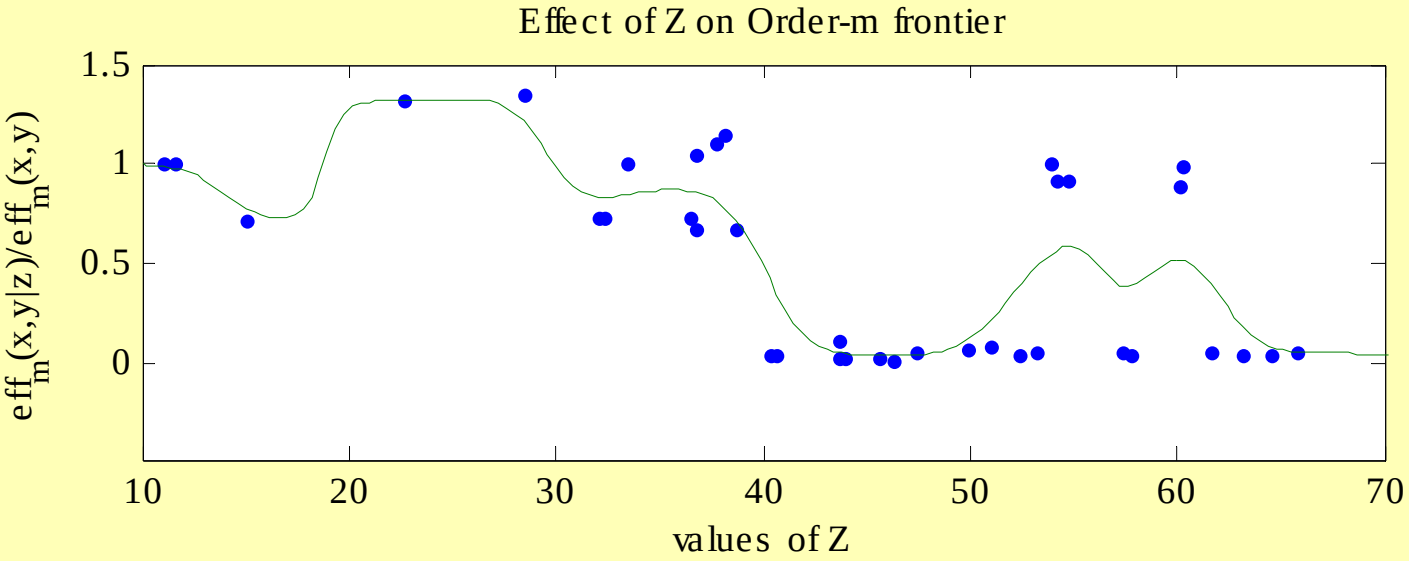
Source: Bonaccorsi e Daraio (2006)

Trade-off publications vs. contract research
res ENG &TECH (Z Trasfpriv)

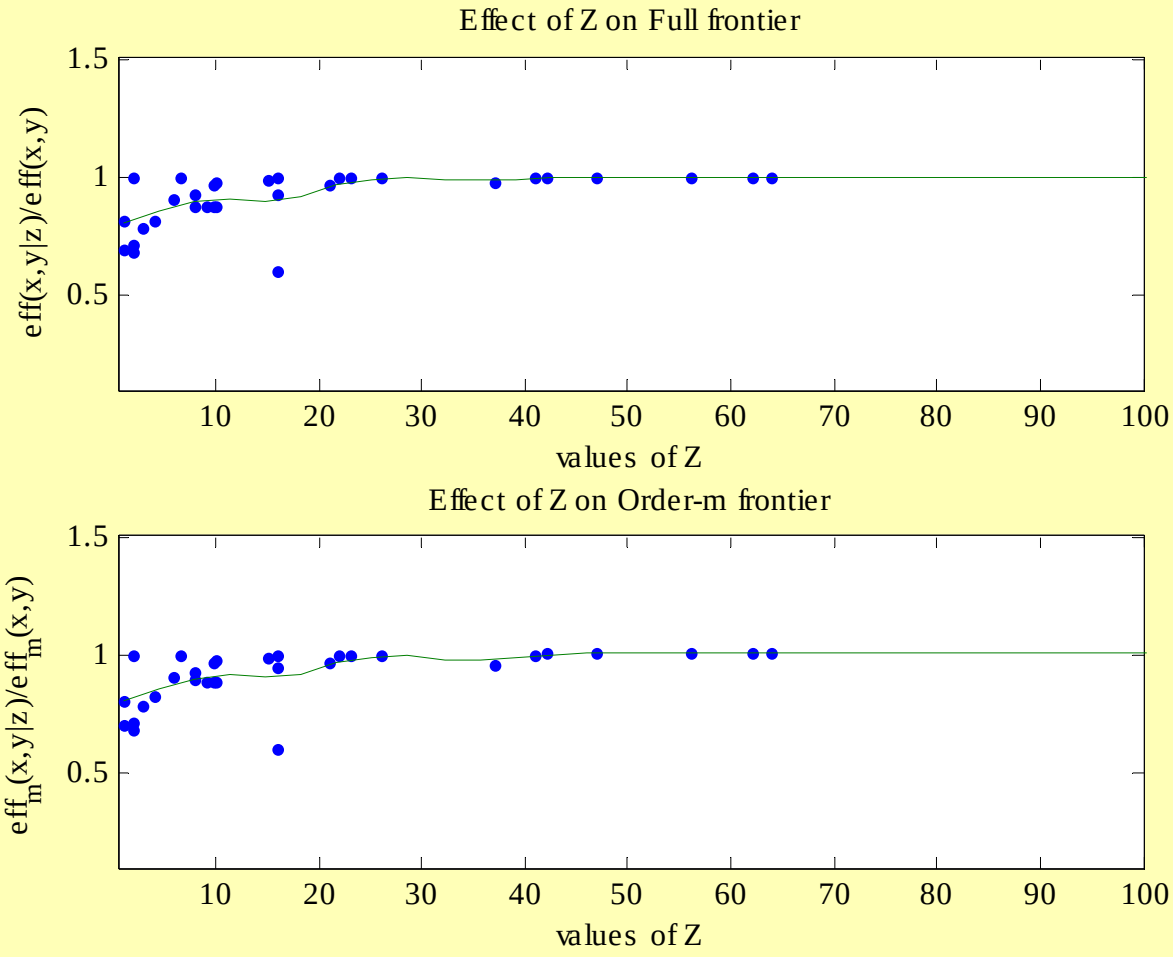


Trade-off publications
vs. industry oriented research
MEDICAL SCIENCES (Z Trasfpriv)

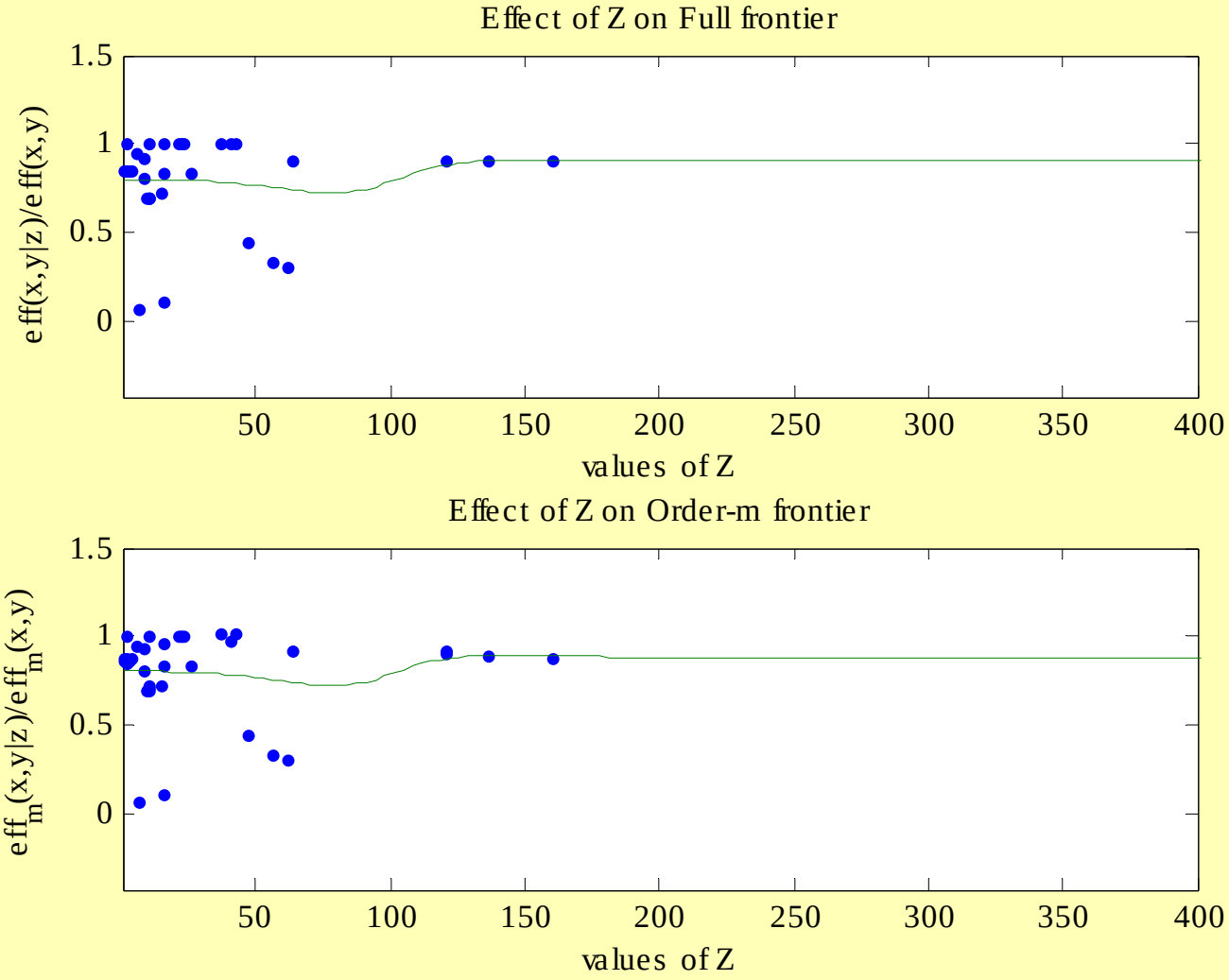




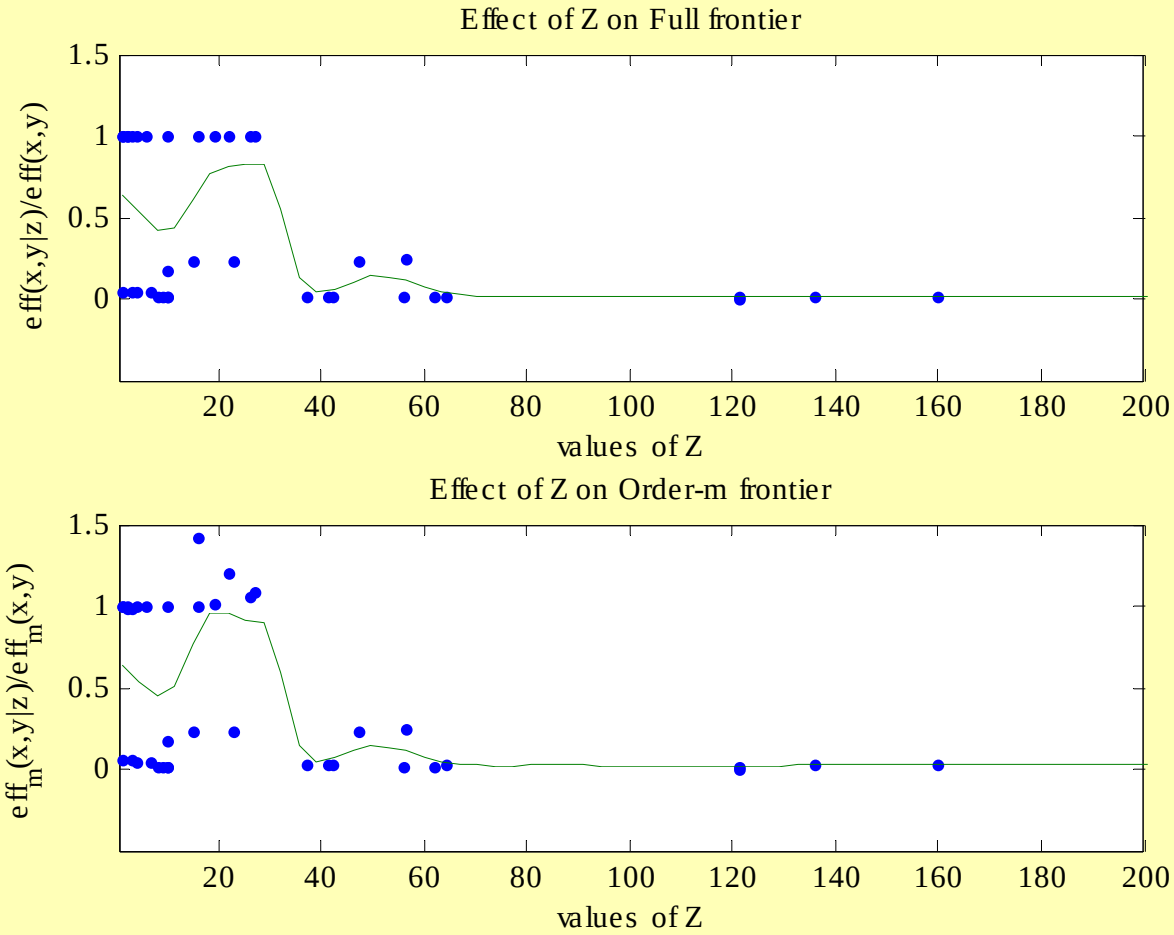
Trade-off teaching vs. applied development
ENG &TECH (Z patents -Balconi)



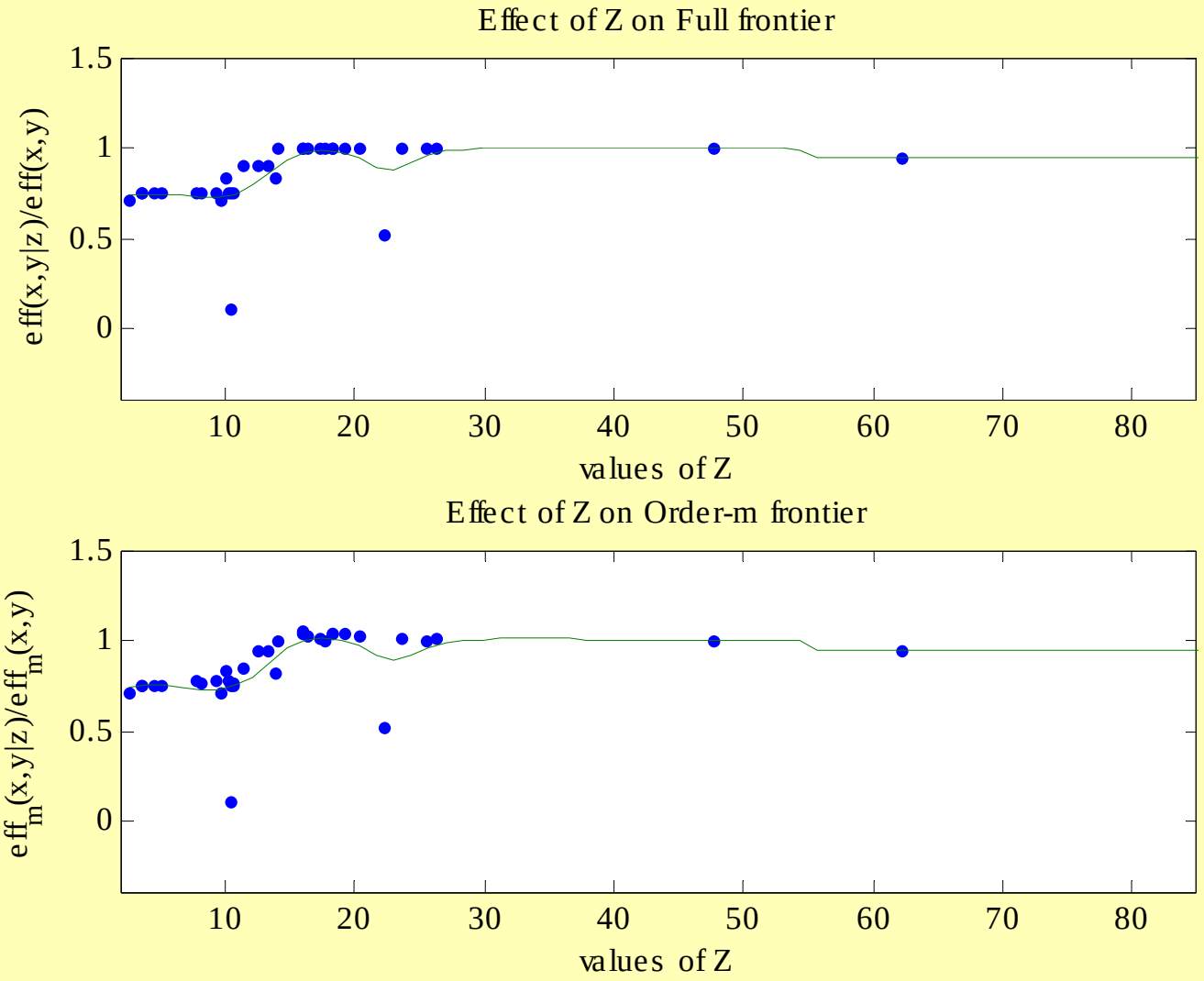
Trade-off publications vs. applied development
ENG &TECH (Z patents -Balconi)



Trade-off publications vs. applied development
MEDICAL SCIENCES (Z patents -Balconi)



Research activity:
ENG&TECH (Z PhDp100T) Source: Bonaccorsi, Daraio, Simar (2005b)



Exploiting the database (2)

Final remarks

In AQUAMETH we demonstrated with very limited resources the feasibility of building a database at the level of individual HEI containing at least some indicators for each main category considered.

This is already an important progress and leads to some new and interesting results.

However, cautious and conscious use of the database: some data are more reliable than others and the data allow answering to some questions, but at the moment not to others. Expert knowledge and the comparison with qualitative information are essential in interpreting the results.

Towards a “meso” data platform on the European Higher Education System

Interesting development in the study of higher education institutions for two aspects:

- The attempt to integrate *quantitative information* on higher education units, allowing systematic comparisons between countries and institutions, as well as the use of robust *econometric methods*, with a *qualitative approach* to characterize both national systems and individual institutions. Within this framework, expert knowledge is a fundamental requisite for any meaningful quantitative assessment.

- The analysis of the *population of individual higher education institution*s going beyond the dichotomy between individual case studies and the study of national higher education systems as whole. This could lead, with further development, to the construction of a typology of the European higher education institutions.

Towards a “meso”data platform on the European Higher Education System (2)

There are several arguments which support the effort done by this project:

1. We address issues not covered by aggregated statistics, at OECD, Eurostat or DG Research level.
2. The project did its job with limited (even too limited, actually) financial resources. This shows the potential for rigorous data gathering and data analysis exercises at European level.
3. The integration of micro-data at European level, taking into account all possible comparability issues, is a **fundamental requisite** for the European Research Area.

Extensions of the database (to other countries):

- Coverage (new countries: France, Hungary, Netherlands, -Germany)
- Regional indicators
- Disciplinary composition
- Quality
- Third mission
- Interactive database for dynamic variables