Some General Thoughts on Using and Improving Innovation Surveys for Econometric Studies

Jacques Mairesse

CREST-ENSAE, UNU-MERIT (Maastricht University), Banque de France

Compnet Workshop, Banque de France, 21 Septembre 2012, Paris

Outline

- General overview
  - History of innovation surveys*
  - Content of innovation surveys*
  - Characteristics of innovation survey data
  - Use of innovation surveys
    - Benchmarking and monitoring
    - Understanding innovation
      - e.g. CDM and HJMP models
- How to make progress? Data problems: Suggestions to improve the quality and relevance of Innovation surveys and their specific usefulness for econometric analysis [Mairesse-Mohnen (2010)]

Characteristics of the data

1. Qualitative data
   - Dichotomous, ordered categorical, unordered categorical
   - Latent variables and distributional assumptions
   - Less informative than continuous data
   - Less subject to measurement errors than continuous data

Characteristics of the data (2)

2. Censored variables
   - Potential sample selection
   - Few explanatory variables to discriminate between innovators and non-innovators
3. Subjective data + errors in variables
   - What is new? New to the market?
   - Share of innovative sales
   - Innovation expenditures

Characteristics of the data (3)

4) Endogeneity

- Cross-sectional data
  - Few variables to instrument
- If Panel data,
  - Few years
  - 4 years apart
  - Random sampling
  - Timing

Use of innovation surveys

- Monitoring and benchmarking
- Understanding innovation
- Policy guidance

Monitoring and benchmarking

- Cross-tabulations
  - Innovation-size
  - Innovation-government support
  - Innovation-cooperation
  - Appropriation of innovation
  - ETC…
- Interest:
  - Do firms with government support innovate more or do more R&D?:
  - Are some firms/industries/ conuntries lagging behind or catching p with others ?
  - ETC…
- Caution: Descriptive analysis is extremely instructive and useful? However correlation does not mean causality

Constructing indexes

- Aggregate in one index many indicators

- Examples
  - European innovation scoreboard
  - Global summary innovation index

- Difficulties:
  - Which indicators to include in the index?
  - Weights ?
  - Intertemporal comparison when components change?
  - International comparisons when questions differ?
  - How to aggregate qualitative variables?
  - Importance of complementarities

Understanding innovation

- Determinants
- Effects
  - Productivity
  - Employment
- Persistence
- Complementarities

The « CDM » model

Crépon- Duguet-Mairesse (1998)

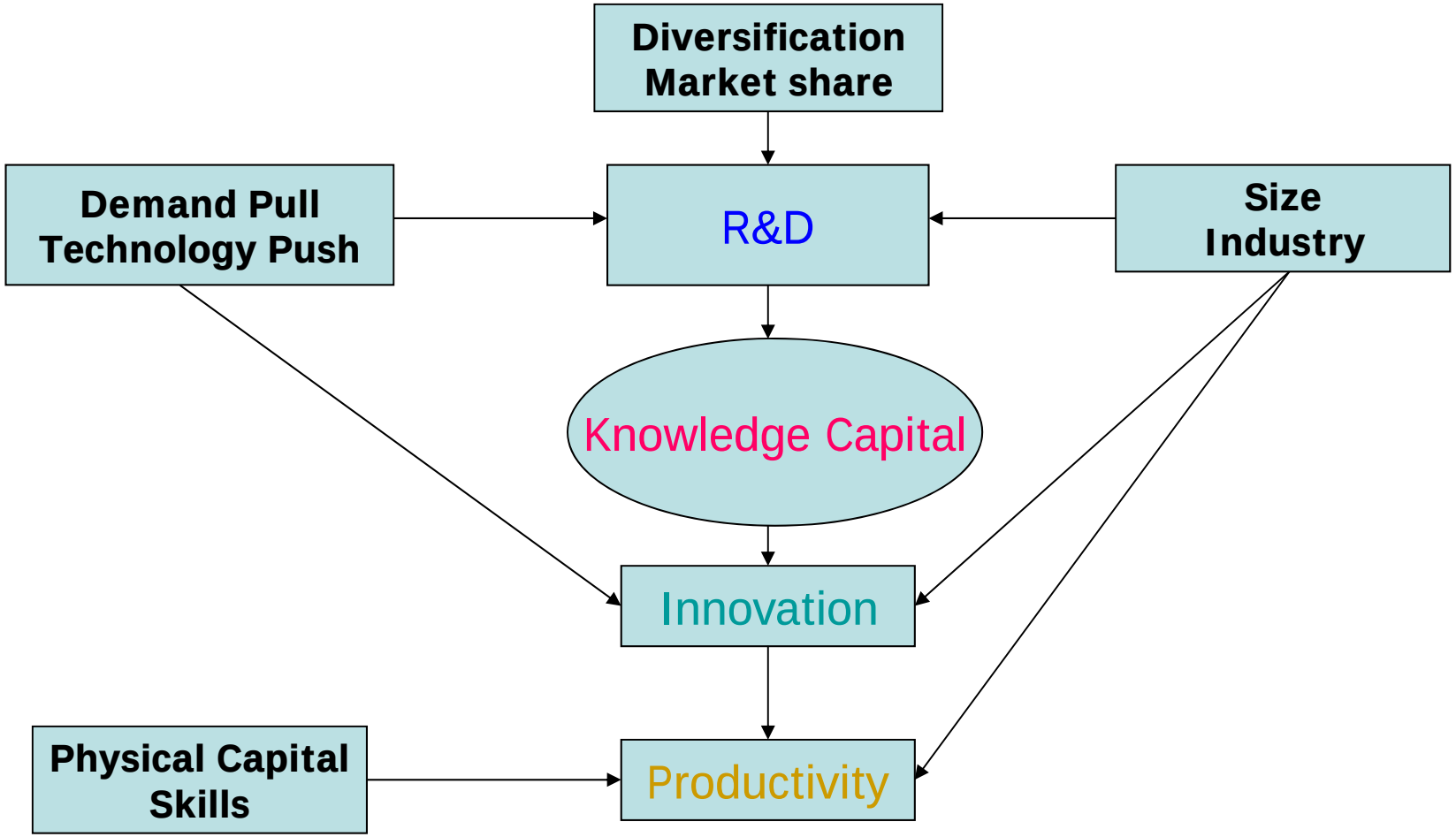
"Research, Innovation and Productivity: An Econometric Analysis at the Firm Level",

Mairesse-Mohnen-Kremp (2005)

"The Importance of R&D and Innovation for Productivity:
A Reexamination in Light of the 2000 French Innovation Survey"

- Brings together the three main fields of investigation in the econometrics of research and innovation
- Proposes a "simple" econometric model articulating innovative and productive activities,
- Uses estimation methods appropriate to the specification of the model and the nature of data
- It is also a framework that can be extended and generalized in various directions
- Takes advantage of the innovation survey information

The « CDM » model



Different estimates of productivity elasticity (in %) with respect to R&D and three innovation intensity variables in High-Tech industries

Estimates of productivity elasticities with respect to:	R&D per employee	Products new to the firm	Products new to the market	Patent holdings
No correction for selectivity and for endogeneity (productivity equation alone - sub-samples)	3.46 (0.51)	0.32 (0.42)	-0.04 (0.51)	0.41 (0.27)
Correcting only for selectivity (partial model - full sample)	3.40 (0.52)	0.34 (0.39)	-0.01 (0.47)	0.52 (0.24)
Correcting only for endogeneity (partial model - sub-samples)	4.50 (1.44)	3.04 (1.82)	1.52 (1.65)	-2.27 (1.00)
Correcting for selectivity and endogeneity (partial model - full sample)	4.28 (0.93)	3.33 (1.26)	2.16 (0.88)	1.57 (0.84)
Correcting only for selectivity (complete model - full sample)	-	0.20 (0.40)	-0.10 (0.49)	0.48 (0.23)
Correcting only for endogeneity (complete model - sub-samples)	-	30.6 (30.1)	11.02 (6.06)	-86.00 (443.50)
Correcting for selectivity and endogeneity (complete model - full sample)	-	**22.54** **(14.50)**	7.00 (2.57)	16.97 (15.66)

Different direct and indirect estimates of productivity elasticities (in %) with respect to R&D in High-Tech industries

Estimates of productivity elasticities with respect to R&D*:	R&D per employee	Through Products new to the firm	Through Products new to the market	Through Patent holdings
No correction (productivity equation alone - sub-samples)	3.46 (0.51)			
Correcting only for selectivity (complete model - full sample)	3.40 (0.52)	0.02 (0.56)	0.00 (0.21)	0.00 (0.36)
Correcting only for endogeneity (complete model – sub-samples)	4.50 (1.44)	3.70 (1.56)	3.01 (1.41)	3.63 (1.60)
Correcting for selectivity and endogeneity (complete model - full sample)	4.28 (0.93)	**4.20** **(1.15)**	2.86 (0.98)	4.39 (1.17)

Corresponding estimates of R&D gross rates of return (in %)
for complete model specifications correcting for selectivity and endogeneity, and using or not the different intensity innovation indicators in High-Tech Industries

Variables	R&D per employee (log)	Products new to the firm (logit)	Products new to the market (logit)	Patent holdings (logit)
Elasticity of output w/t R&D	**4.28**	**4.55**	**4.37**	**4.88**
	(0.93)	**(1.13)**	**(0.99)**	**(1.14)**
Rate of return of R&D				
First quartile	**13.1**	**13.9**	**13.3**	**14.9**
Median	**24.7**	**26.3**	**25.2**	**28.2**
Third quartile	**56.7**	**60.3**	**57.9**	**64.7**
Average	**51.8**	**55.2**	**52.9**	**59.1**

SOME CONCLUSIONS

- Innovation indicators are noisy measures which needs to be instrumented. Taking into account selectivity also helps. Binary indicators seem somewhat less noisy than intensity indicators.
- R&D appears largely "exogenous" to productivity, not to innovation. If one is only interested in R&D productivity (i.e., rate of return), the extended production function does well.
- Innovation indicators are useful to go beyond a strict focus on R&D returns and to investigate the "black box" of innovative activities, for example in a framework à la "CDM".
- …

Employment Growth, Export, Product Innovation and Distance to the Productivity Frontier in China: A Firm Level Comparison across Regions, Industries, Ownership Types and Size Classes

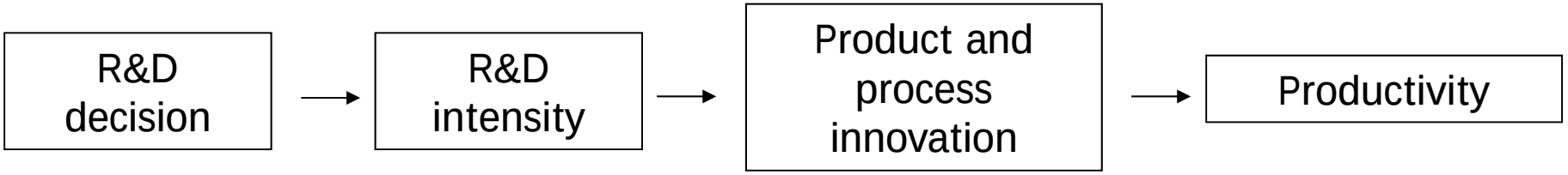
By

Jacques Mairesse, Yilin Wu, Yanyun Zhao and Feng Zhen

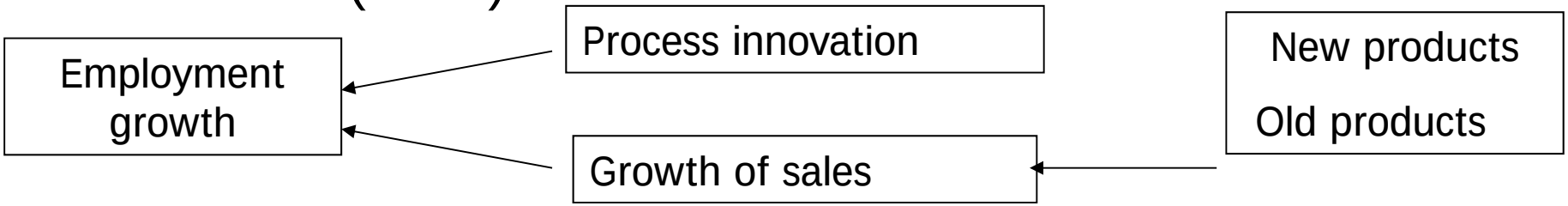
5th Conference on Micro Evidence
on Innovation and Development
27-28 June 2011, San José, Costa Rica

Background

- Crépon, Duguet & Mairesse (1998); Mairesse, Mohnen & Kremp (2005); Griffith, Huergo, Mairesse & Peters (2006); Hall, Lotti & Mairesse (2009)



- Harrison, Jaumandreu, Mairesse & Peters (2005); Hall, Lotti & Mairesse (2007)



Data Source

- Survey conducted by the National Bureau of Statistics
- The annual industry survey, covering (all) 29 manufacturing industries, for **all** State-owned firms, and non-state-owned firms **with sales higher than 5 million RMB Yuan** (i.e., Limited-liability, Share-holding, Private, Hong-Kong, Macao and Taiwan, Foreign)
- Allows the construction of an **unbalanced firm panel over the 8 years 1999-2006**, with usual current account and balance sheet variables, **Total Employment** (EMP), and **New Products Output** (NPV) in all years except 2004, **R&D** expenditures (R&D) in all years 2001 -2006 except 2004, and **Export** (EXP) in all years

Numbers of firms in different periods

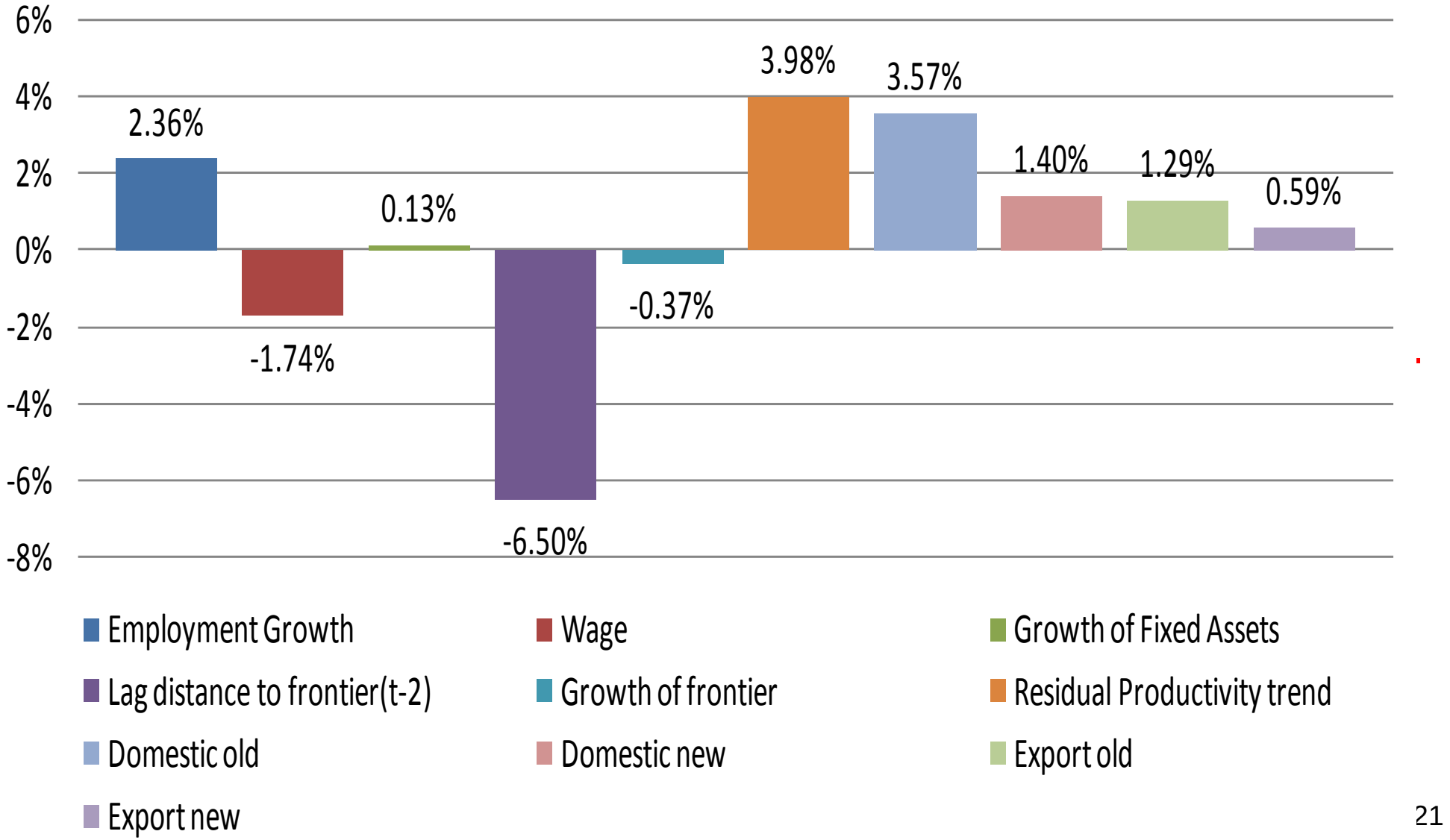
Time Period	Numbers of firms
1999-2001	47531
2000-2002	36831
2001-2003	39628
2002-2004	43083
2003-2005	50537
2004-2006	52193

***After we pooled these six samples together, the numbers of observations in each model is 269803.**

Descriptive statistics of main variables

	Employment Growth	Old and domestic product growth	Old and export product growth	New and domestic product growth	New and export product growth
mean	2.4%	9.9%	2.2%	4.1%	1.1%
sd	0.22	0.42	0.25	0.17	0.08
p5	-29.1%	-42.9%	-24.9%	0.0%	0.0%
p25	-6.9%	-8.1%	0.0%	0.0%	0.0%
p50	0.0%	2.9%	0.0%	0.0%	0.0%
p75	7.9%	23.1%	0.0%	0.0%	0.0%
p95	42.5%	78.3%	37.7%	28.2%	1.2%

Employment Growth Decomposition



Employment Growth Decomposition of Different Ownerships

%	Employment Growth	Wage	Growth of Fixed Assets	Lag distance to frontier (t-2)	Growth of frontier	Domestic old	Domestic new	Export old	Export new	Residual Productivity trend	Numbers of Obs
Descriptive Statistics											
State Own	0.13	15.13	6.25	162.96	11.74	8.65	0.50	5.31	0.83	-	46.21
Private	3.87	18.18	13.16	148.08	12.81	15.82	2.34	3.14	1.04	-	26.14
Foreign	4.65	15.12	4.75	131.67	7.99	6.25	4.91	2.95	1.74	-	27.66
Decomposition base on whole sample model											
State Own	0.13	-1.65	0.11	-7.04	-0.39	3.13	0.32	1.67	0.43	3.56	46.21
Private	3.87	-1.98	0.22	-6.40	-0.43	5.73	1.49	0.99	0.53	3.71	26.14
Foreign	4.65	-1.65	0.08	-5.69	-0.27	2.27	3.13	0.93	0.90	4.96	27.66
Decomposition base on separate sub ownership samples models											
State Own	**0.13**	-1.56	0.16	-6.56	-0.30	2.33	0.37	1.28	0.45	3.97	46.21
Private	**3.87**	-1.72	0.16	-7.64	-0.61	5.28	1.56	1.00	0.53	5.31	26.14
Foreign	**4.65**	-1.83	0.10	-5.11	-0.24	2.98	2.57	1.12	0.81	4.26	27.66
Total	**2.36**	**-1.74**	**0.13**	**-6.50**	**-0.37**	**3.57**	**1.40**	**1.29**	**0.59**	**3.98**	100.00

Conclusions

- The growth of output for the domestic market plays a major role for employment growth (5%), about more than 2 times larger than for exports (2%)
- Innovation has a positive effect on employment growth, but a modest one: 2% (1.4% for the domestic market , 0.6% for export)
- The growth of wage has a significant negative effect (-1.7%) while growth of fixed assets has negligible effect
- Catching up to the productivity frontier corresponding to process and organizational innovation has the largest effect (-6.5%)

How to make progress?

Mairesse-Mohnen (2010)

"Using Innovation Surveys for Econometric Analysis", in *Handbook of the Economics of Innovation*
B. H. Hall and N. Rosenberg eds, Elzevier, 1129-1156

- Enforce a set of recommendations concerning the design and implementation of the innovation surveys that could be useful for better information both in the form of descriptive statistics, indicators and scoreboards and for better econometric analyses

(1) Harmonize strictly across countries and across waves a core of basic questions in the innovation surveys

- The survey questionnaire should be split into three parts:
  - the core permanent part, which might be relatively short but should be as strictly stable over time and identical across countries as possible;
  - a part carefully harmonized across countries but possibly varying from one survey to another to analyze specific or new aspects;
  - an optional part in response to country special interest.
- If the sampling procedure cannot be identical across countries, information should be provided allowing to compare performances across countries and industries.
- Experiments should be conducted on the sensitivity of the survey responses to the wording and the order of questions, and the functional role of the respondents in the enterprise.

(2) Ease access to innovation data

- Methods to give access to the firm level information to researchers, such as granting secure remote access to the raw data or providing micro-aggregated or otherwise noise-contaminated data that hide the firms' identity, should be generalized.
- They should also be extended to allow researchers to access the data from various countries and do international comparisons.
- The French system of "Comité du secret", which allows selected researchers in academic institutions to have more flexible access to firm level information in the surveys for a specific research project, a limited time and with strict confidentiality obligations, should also be promoted.

(3) Merge innovation survey data with other data

- In order to explain the choice of innovating or not, or to correct for potential selectivity in explaining the intensity of innovation, little analysis can be done with the innovation survey data alone, because few variables are usually collected for all firms (including non-innovators) in the innovation surveys.
- One solution is to collect more data about non-innovators in the innovation surveys themselves. Another one is to be able to merge the innovation survey data with data from other surveys.
- This will also offer a much larger choice of instruments to correct for endogeneity and measurement errors, and more generally provide more explanatory variables to consider in the models and increase their relevance and explanatory power.
- Innovation and R&D surveys should be merged systematically.

(4) Create longitudinal datasets

- If a panel of firms could be constructed that was followed over at least a few years, it would be possible to study the dynamics of innovation (i.e., the time lags in the determinants and the effects of innovation), and to correct for firm-specific effects (i.e., individual unobserved heterogeneity)
- Having a panel should also help in addressing the difficulties arising from the fact that firms over time enter and exit, and can radically change shape over time by mergers and acquisitions
- It is hard to infer strong conclusions regarding causality using only cross-sectional data
- Innovation surveys (jointly with R&D surveys) should be collected and treated as panel data all along by official statisticians

(5)-A: Organize a close collaboration between statisticians and economists

- The elaboration and appropriate implementation of all the above recommendations will greatly benefit from an active and organized collaboration between official between economists working on research and innovation issues and statisticians responsible of the innovation surveys.
- Based on the experience of different organizations in different countries to execute the innovation surveys and exploit their first statistical results, one could even think of implementing a framework of shared responsibilities, following the distinction of three different groups of questions in innovation surveys as suggested in our first recommendation above

(5)-B: Organize a close collaboration between statisticians and economists

- Statistical offices and professional statisticians would be in charge of the core component of the innovation surveys, which should be mandatory in all countries; they could also be responsible for the country harmonized component, or could entrust research institutes with this task on a long term basis.
- The third optional (and usually changing) component that responds to a country's specific interest might be delegated to a research group, or to the professional or private organization best capable of realizing it well.
- Such an organization might be more efficient overall, alleviating some of the various costs involved in doing and making use of the innovation surveys, while contributing to increase their overall usefulness

(6) 0ther

- Collect data on groups and especially on multinationals
- Adapt surveys for providing information at the regional level developing countries
- Adapt surveys for developing countries
- …