

AN OVERVIEW OF THE BANCA D'ITALIA BUSINESS SURVEYS

LAURA BARTILORO SAMPLE SURVEYS DIVISION – FIRMS SECTOR

CENTRAL BANK BUSINESS SURVEYS AND LIAISON PROGRAMS

MIAMI, NOVEMBER 28TH 2017



A SHORT DESCRIPTION OF BANCA D'ITALIA BUSINESS SURVEYS

- ➢ 4 SURVEYS
- **THE PLAYERS**
- DATA RELEASE (BIRD)

PROBLEMS?

- **ORGANIZATIONAL ISSUES**
- **FIRMS STATISTICAL BURDEN**

THE WAY FORWARD

- **Reshape**, ONCE IN A WHILE!
- **Use of Administrative Data**
- **Research Data Center**



STRUCTURAL SURVEYS (QUANTITATIVE DATA)



Not all the Bol's surveys are conducted by the Sample Surveys division (i.e. **Direct Reporting**)



SAMPLE DESIGN

The units of analysis are Italian firms of the following sectors: industry (incl. construction) and services. For the survey on the housing market the sample refers to real estate agents. The list of contacted firms is extracted from Cerved (Central Balance Sheet Database); that of RE agents from FIAIP (an association of RE agents). The sample is a panel sample, refreshed once in a while.

Spring Survey	Business Outlook	Expectations	Housing
3,000 industry	3,000 industry	400 industry	
1,000 service	1,000 service	400 service	
500 construction	500 construction	200 construction	
4,500 total	4,500 total	1,000 total	Aprox 1,400 RE agents

Samples are stratified, with strata consisting of combinations of economic sector (industry, service, construction), number of employees (10/20/50-199, 200-999, over 999) and geographical area (North-West, North-East, Center, South and Islands).

BANCA D'ITALIA EUROSISTEMA

BUSINESS SURVEYS

	Survey of Industrial and Service Firms	Business Outlook of Industrial and Service Firms	Survey on inflation and growth expectations	Italian Housing Market Survey
Complexity	Very high	Medium	Low	Low
Partners	No	No	Sole-24-ore	Agenzia del territorio & TecnoBorsa
Collection of data	Regional branches	Regional branches	External agency	External agency
Mode of interview	Mixed (Email /Web /Tel /Pers)	Mixed (Email/Web)	Web	Mixed (Web/Tel.)

... but participation is not compulsory for any of these surveys!



Annual Business Surveys

(percentages by participation modes)





INTERVIEW MODES

Email message to the firm asking for its participation in the survey (with a link to Banca d'Italia survey managing application GestInvind).

1. The firm may log in the dedicated web site and **directly load** the data required (preferred option).

2. Firms may either **fill in the dynamic-Pdf file** downloadable in *GestInvind* and send it back by email to the BoI local branch.

3. In some other cases firms may **answer on the phone** to the questions posed by Bol's local branch's employees.

4. Around 30% of the firms participating in the sample, are **directly visited** by the Bol's local staff in case of the Spring Survey.



THE INTERVIEWERS

Highly specialized bank personnel (often economists) with:

- deep knowledge of the economic condition of their local area
- often good and friendly relationship with interviewees
- mutual trust climate fostered over the years

As a consequence:

- ✓ questions can involve sensitive topics (long-term strategies, trade credit policies, sentiment towards the banking sector, etc.)
- ✓ participation rate quite high (70-80%)
- ✓ partial non response rate kept within reasonable limits



THE QUESTIONNAIRE

The questionnaire varies according to firm size, economic activity and rotation.

Around 50% of the questionnaire content is renewed at every wave.

A selection process of the new questions is carried out:



All economists of the Bol Research Department and of the Local Branches are asked to propose new questions for the survey



The Sample Surveys Division

- ✓ examines feasibility of all proposed questions
- ✓ suggests rewording (if needed)
- ✓ estimates total firms' statistical burden

In order to satisfy as many economists' requests as possible, some variables are collected only for randomized half-samples representing 50% of the total sample. <u>Spring survey questionnaire</u> <u>Business outlook questionnaire</u> <u>Inflation and growth exp. questionnaire</u>



→ the press and the general public

The reports usually are an important reference for the domestic political debate on the economic conditions of firms.



Main results

The survey conducted in September 2017 on a sample of Italian firms with 50 or more employees suggests a continuing and widespread improvement in the assessments of the current general economic statusion and in expectations for the next three months, confirming the signals of firms' optimism about the outlook for demand.

Consumer price inflation expectations remain moderate across all time horizons, as do the past and expected trands in firms' own prices. The expected growth of firms' list prices is likely to be boosted by the increase in raw material prices, and to a lesser extent by labour costs and intermediate inputs, demand-side pressure is expected to be negligible, despite a consolidated improvement in firms' assessments.

Firms' investment plans for the current year continue to grow in all the main sectors, reflecting the overall favourable opinions of firms on investment conditions. Expectations of an increase in employment are confirmed in industry excluding construction and to a lesser extent in the services sector; in construction the negative situation is mitigated.



(1) The shaded areas represent the 95% confidence intervals. Construction firms are included from 2013 Q3 See tabb. s1 and s4 of the statistical appendix. Each report contains:

- A description of the main results (and a summary for the press)
- A set of statistical tables
- A methodological appendix
- The questionnaire





→ BI researchers, regional branches, decisionmakers and academics

EUROSISTEMA

- The Annual Report
- Regional notes
- Policy-relevant research projects
- Speeches by the Governor
- Articles on scientific journals
- Books





An example of our in-depth analysis

(L.Bartiloro, M.Bottone, A.Rosolia, 2017, *What does the heterogeneity of inflation expectations of Italian firms tell us?*, mimeo)



Figure 1: Businesses' inflation expectations and disagreement.



→ BIRD (Banca d'Italia Remote access to micro Data)

Firms microdata have to be anonimysed \rightarrow remote processing solution, using internet, without direct access to individual data.

In March 2008 the Banca d'Italia started a system for remote processing of data collected for her Survey of Industrial and Service Firms, known as BIRD (Banca d'Italia Remote access to micro Data).

http://www.bancaditalia.it/statistiche/bird

BIRD allows researchers to process data on industrial firms and nonfinancial private service firms collected since 1984, ensuring that individual information cannot be visualized.

The sistem also covers data from the Business Outlook Survey and was extended in 2013 also to microdata from the Survey on Inflation and Growth Expectations and on the Italian Housing Market Survey. The researcher performs statistical and econometric analyses by submitting source programs by email and receiving output by email. Input and output alike are subjected to automatic and manual checks of their content and logic (for example, a command attempting to visualize individual data is automatically rejected)

Statistical packages supported by the system: Stata and R

How BIRD works:

- The user sends a message containing authentication info and the statements to be submitted
- 2. Input message is parsed and checks are performed
- 3. If no error/security violation @ submit statements
- 4. Output is parsed (automatically / manually)
- 5. If no security violation @ forward to the user via e-mail



THE DIMENSIONS OF THE PROBLEM:

- STRESSED" FIRMS
- QUESTIONNAIRE
 - ✓ COMPLEXITY
 - ✓ LENGTH
- FIRMS' INCENTIVE





Number of waves to which firms have participated in 3 years (2 surveys per year)

1	2	3	4	5	6
526	475	703	686	1121	2778

Firms estimate of the statistical burden

(Average of assessments expressed on a range from 1=low to 10=high; Spring Survey, 2016)

burden		# quest.	# people	difficulty	mis.opt.	dif. opt.
Low-med	56%	4.6	3.8	2.8	2.8	3.0
High-exc	36%	7.4	6.5	4.3	3.6	4.2

BANCA D'ITALIA EUROSISTEMA

Spring Survey

Business Outlook







The number of variables increased at the end of the Nineties, and it is stable since then ...



... but the number of pages increased in the last 8 years!

MISSING VALUES OVER TIME



In the last 20 years the share of firms that answer to (almost) the whole questionnaire is pretty stable, while that of those that answer to fewer questions has increased.





Firm distribution for % of non missing variables

The amount of missing values and questionnaires length seem to be positively correlated ...



BANCA D'ITALIA

EUROSISTEMA

... and the correlation resists even after controlling for the economic cycle.

Logistic regression (pooled cross-section)

dep. var. = non-response to at least one of exp. turnover, investments or employment Time dummies among the expl. vars.

BANCA D'ITALIA

EUROSISTEMA

	Odds-ratio	Std. Error	T value	Pr(> t)
# empl.	1.000	0.000	14.319	0.000
Energy	1.032	0.008	3.801	0.000
Services	1.021	0.003	6.175	0.000
North-East	0.966	0.004	-8.077	0.000
Center	0.947	0.004	-13.093	0.000
South & Is.	0.927	0.004	-20.015	0.000
Loss (from profit)	1.009	0.001	7.148	0.000



Corr = 0.67





- **1.** DOES THE ACTUAL SETUP STILL FIT?
 - LOOK AT THE NOWADAYS NEEDS
 - DELETE POSSIBLE OVERLAPPING
- 2. REDUCE FIRMS' STATISTICAL BURDEN
 - ✓ The ESCB has endorsed the principles set up by the *European Statistics Code* of *Practice* of the European Commission for the statistical production.
 - limit respondents' statistical burden to the extent necessary for the users' needs
 - o cost-benefit analysis of the collection, processing and release of data
 - ✓ Quality (and validity) of the survey
- **3. INCREASE EFFICIENCY** (i.e. reduce internal costs)

e.g. Business Outlook vs Growth exp.: a matter of sinergy or doubling?



✓ THE SAMPLE

- THERE'S A RELATIONSHIP BETWEEN SAMPLE DESIGN AND SCOPE OF THE SURVEY
- REGIONAL GOODNESS OF FIT VS STATISTICAL BURDEN (FOR FIRMS AND INTERVIEWERS) TRADE
 OFF
- ✓ THE QUESTIONNAIRE
- ✓ THE INTERVIEWERS

.... AND COMING SOON

Use of Administrative Data

Research Data Center





Laura Bartiloro Banca d'Italia Statistical Analysis Department Sample Surveys Division <u>laura.bartiloro@bancaditalia.it</u> +39-06-47924165



Survey of Industrial and Service Firms (since 1984) Frequency: annual (spring) Business Outlook Survey of Industrial and Service Firms (since 1993) Frequency: annual (autumn)

- Sample: 4,500 firms
 - o 3,000 industrial firms (with 20 employees and more)
 - o 1,000 service firms (with 20 employees and more)
 - o 500 construction firms (with 10 employees and more)
- > Topics:
 - o Stable: Investments, Turnover, Employment
 - **Special**: i.e. Uncertainty (proposed by the economists of the Research Department and of regional branches)



THE PANEL STRUCTURE

Local branches are given a detailed quantitative target. In order to reach their target, they try to contact the firms that participated in the previous edition.

Some of these firms are not available for temporary or permanent reasons:





Firms are routinely replaced with other firms (replacements may even have participated in past editions)

- replacements are chosen from exhaustive lists (yearly updated) and are as similar as possible to the firms they replace
- a set of replacement rules is provided to the local branches



WEB – BASED DATA COLLECTION

• Data are loaded on a dedicated web architecture. <u>https://gestinvind.bancaditalia.it</u>

- The software development was conducted entirely within the Bank's information systems departments (programming tasks were outsourced within broad contractual arrangements with software firms).
- The platform is continuously updated to take into account of new needs and always evolving requirements. (*i.e. multiple emails*)



THE TIMELINE

- ✓ November-December: the new questionnaires are discussed and prepared (there is also a pilot study for the new questions proposed).
- ✓ **February-May:** data are loaded on the web platform.
- ✓ From April: microdata are available for Bol's research personnel and continuously updated.
- ✓ **Mid-May:** field closes.
- ✓ End of May: the main estimates are computed for the Bank of Italy's annual report, released at the end of the month, and for the report on the Economies Developments of the Italian Regions (released at the end of July).
- ✓ End of June/beginning of July: a detailed report on the survey result is published on Bol website and microdata are released for external users.



USE AND USERS OF FIRMS SURVEY DATA: BIRD

Utenti per anno		
Anno di registrazione	frequenza	frequenza cumulata
2008	12	12
2009	4	16
2010	4	20
2011	8	28
2012	6	34
2013	12	46
2014	26	72
2015	10	82
2016	20	102
2017	14	116
Lavori sottomessi (job)		
semestre	numero job	
l 2008	235	
II 2008	15	
l 2009	280	
ll 2009	7	
l 2010	163	
II 2010	95	
2011	155	
ll 2011	12	
l 2012	243	
II 2012	88	
l 2013	665	
ll 2013	196	
l 2014	373	
ll 2014	291	
l 2015	281	
ll 2015	363	
l 2016	325	
ll 2016	244	
l 2017	94	
Totale	4125	



MULTI-LEVEL DATA CHECKS

- 1. First level
 - **Controls on the questionnaire** (online or dynamic pdf file).
 - **Controls by Bol's colleagues** in charge of the interviews.
- 2. Second level
 - Data controls conditional to the other firms in the same strata (**Score**)
 - Special treatment reserved for data on firms affected by extraordinary events, such as mergers or splits.
- 3. Third level
 - **Outlier detection** carried out at the central level.
 - **Macro editing** carried out at the central level before the production of the final estimates.



Computer-intensive techniques used to account for:

- 1. Design weights and post-stratification adjustments;
- 2. Imputation;
- 3. Deflation;
- 4. Winsorization;
- 5. Statistics.



- **Design weights** are computed as the inverse of the probability of inclusion in each stratum. ($w_i = N_i/n_i$, but remember the autorepresentative)
- Poststratification procedure (**raking**) is performed using the known information on the population distribution.
- Several **weight available**:
 - «peso»: based on the distribution of firms;
 - «pesocc»: based on the distribution of employees (ad-hoc analyses requiring a scale factor);
 - «pesoa» / «pesob»: from 2012 for the two randomized half samples.
- Weights are updated as soon as updates of the distribution of the reference population become available (official sources on the target population, ASIA, have a two-year lag compared with the survey).



Only for investment survey

- The imputation process is carried out either with **strata averages** (in terms of per capita, since the number of employees is never missing) or with simple linear models (also including a randomized component) producing a predicted per-capita value.
- The imputations techniques also take the **panel structure** into account. This method gives an estimate of a level per employee that is obtained by combining a general cross-section mean and an average based on the firm's time series.

Yes, ok, but... how to recognize an imputed variable???

Look at the flag! (fxxx = 1)



- Price change are explicitly asked in the questionnaire (industry and services).
- Inflation rates are calculated by sector of activities.
- Variables at constant price are always referred to the last year price level.
- For construction firms, total production at constant prices is calculated on the basis of Istat's latest deflator for construction.

Yes, ok, but... how to recognize variables at constant prices???



DATA TREATMENTS: WINSORIZATION

Preliminary data treatment for micro-variables

$$X_{i}^{t^{wins}} = \begin{cases} f^{*}X_{i}^{t} + (1-f)^{*}J & if \quad X_{i}^{t} < J \\ f^{*}X_{i}^{t} + (1-f)^{*}K & if \quad X_{i}^{t} > K \\ & X_{i}^{t} & otherwise \end{cases}$$

 $f = \frac{1}{w_i}$ sampling fraction

- J and K are cut-off values generally chosen at the 5th and 95th percentiles.
- The method is used **for investment and turnover** and prevents erratic per-capita values from excessively influencing the ratio estimates.

Yes, ok, but... how to recognize a winsorized variable???

Look at the "w" extension! (xxxw)



The Investment survey collects both qualitative and quantitative questions

TYPICAL ESTIMATES FROM CATEGORICAL VARIABLES

Main indicators: weighted frequencies and balances B.

B=P-N, where P is the frequency of positive answers and N the frequency of negative answers

TYPICAL ESTIMATES FROM QUANTITATIVE VARIABLES

Main indicators: Weighted means, totals, and ratios.

- Yearly variations for investment, turnover, employment, etc, computed with a single survey, by asking values for three years.
- Variables at the numerator (t, t+1) are winsorized, variables at the denominator are not.



TYPICAL ESTIMATES FROM CATEGORICAL VARIABLES

$$se(x) = \sqrt{p * (1-p)} * fpc$$
 con $fpc = \sqrt{(N-n)/N}$

TYPICAL ESTIMATES FROM QUANTITATIVE VARIABLES

- Compulsory, non imputed, variables (investment, turnover and employment at t)
 Jacknife (leave-one-out) method.
- Non compulsory, imputed, variables (investment, turnover and employment at t+1)
 - Multiple imputation: 25 bootstrap samples of complete observations extracted from the original one, preserving the original sampling design. Imputation process repeated after reproducing the weights including post-stratification



KEY ISSUES

- ✓ documentation of the archive structure
- ✓ documentation of the available variables
- ✓ some examples of data utilization (mainly useful for external users)
- ✓ need to update and translate in English this documentation



Surveys are however similar in terms of...

- Sample selection (random sample, panel)
- Data treatments
- Production
 - Questionnaire design (pre-test)
 - □ Sample design and weights
 - Missing data and imputation
 - Selective editing and outliers
 - Standard errors and robust estimators

... and participation is not compulsory for any of these surveys!



There is a strong **cooperation** between our division and the other units of the Research Department as well as the Bol regional branches.

- Preparation of the questionnaires
- Realization of interviews
- Analysis of data (data quality, economic analisys)