BANCA D'ITALIA
EUROSISTEMA

Temi di Discussione

(Working Papers)

Investment forecasting with business survey data

by Leandro D'Aurizio and Stefano Iezzi

November 2011

Number 832

BANCA D'ITALIA
EUROSISTEMA

Temi di discussione

(Working papers)

Investment forecasting with business survey data

by Leandro D'Aurizio and Stefano Iezzi

Number 832 - November 2011

INVESTMENT FORECASTING WITH BUSINESS SURVEY DATA

by Leandro D'Aurizio* and Stefano Iezzi*

Abstract

Business investment is a very important variable for short- and medium-term economic analysis, but it is volatile and difficult to predict. Qualitative business survey data are widely used to provide indicators of economic activity ahead of the publication of official data. Traditional indicators exploit only aggregate survey information, namely the proportions of respondents who report "up" and "down". As a consequence, neither the heterogeneity of individual responses nor the panel dimension of microdata is used. We illustrate the use of a disaggregate panel-based indicator that exploits all information coming from two yearly industrial surveys carried out on the same sample of Italian manufacturing firms. Using the same sample allows us to match exactly investment plans and investment realisations for each firm and so estimate a panel data model linking individual investment realisations to investment intentions. The model generates a one-year-ahead forecast of investment variation that follows the aggregate dynamics with a limited bias.

JEL Classification: C500, C520, C530
Keywords: investment plans, dynamic panel data model, forecasting.

Contents

* Bank of Italy, Economics, Research and International Relations.

1 Introduction

Any attempt to predict the GDP growth of a country is risky without greater knowledge of its various components. One important component is gross national expenditure: investments (capital expenditure) are one of the major components in terms of both size and variability. Moreover, the economic importance of capital expenditure is greater than that warranted by their simple values, since increases in production capacity produce their effects over many years. Productive over- or under-capacity is among the main determinants of economic cycles. While the size of the aggregate is fairly small in relation to GDP, it overreacts to variations in the level of activity, thus making a significant contribution to variations in GDP (Bernanke, 1983). In addition to exerting a short-term influence on demand, investment makes it possible for firms to increase their physical capital. As a result, current investment efforts have an impact on the future, with consequences in the medium term for corporate supply (Chirinko, 1993).
For all these reasons, surveys collecting firms' investment intentions have been regularly conducted in the major developed countries since the end of World War II and these data are a very important variable in short- and medium-term economic analysis. Business investment, however, is volatile and difficult to predict.

Normally, in business surveys firms are asked whether they plan to increase, maintain or reduce investment spending over a specified period of time using simple categorical questions. Many studies have attempted to understand the capability of these survey responses to anticipate official data on both output and price movements (Nardo, 2003).

These microdata are generally aggregated as frequency distributions and two approaches have been devised in the literature to transform them into quantitative estimates comparable to official data: the probability method of Carlson and Parkin (1975) and the regression method of Pesaran (1984, 1987), used for inflation and output indicators. Other authors have improved these methods over the years, but all the techniques are based on aggregate individual responses: thus, neither the heterogeneity of individual responses nor the panel dimension of microdata is used.

As an alternative, allowing for a degree of heterogeneity among firms might be a more efficient way to draw inferences about the variation in aggregate output. The panel data structure underlying the aggregate responses has so far received little attention, with the exception of Mitchell et al. (2004). They construct a "disaggregate" indicator built around ordered discrete choice models linking individual firms' categorical responses to economy-wide official data. They combine a sample estimate of firms' output growth obtained by a Bayesian quantification of categorical data with past aggregate indicators of output levels. In this way they build parametric and a non-parametric quantitative forecasts of future output growth, which they compare in an out-of-sample exercise with the classical quantifications of categorical data proposed by Carlson-Parkins and Pesaran and with “naïve” autoregressive models of past aggregate data. The non-parametric version turns out to have the best performance. The method is a refinement of an earlier solution proposed by the same authors (Mitchell et al., 2002) consisting of an alternative “semi-disaggregate” indicator based on grouping the firms according to their responses at both time t and time $t-1$.

In this paper we illustrate the use of a disaggregate panel-based indicator that exploits all information coming from two yearly industrial surveys carried out on the same sample of Italian firms. Since 1993 the Bank of Italy has collected data on annual investment plans and investment realisations in the manufacturing sector with two surveys on the same panel sample. Every firm reports investment plans for the following year in qualitative form in a short-term business outlook survey in September. An extended survey, carried out in the first months of the year, collects investment levels for the previous two years, together with a forecast of the current year’s level. Using the same sample allows us to match exactly investment plans and investment realisations for each firm and so estimate a panel data model linking individual investment realisations to investment intentions usable for forecasting. The purpose of the paper is to show the construction of the model and how it performs in predicting one-year-ahead investment variations.

The paper is organized as follows. Section 2 describes the survey data used and provides some descriptive evidence. The following Section 3 sketches the modelling strategy. Some more complex econometric issues are described in Section 4, while in Section 5 results are presented and interpreted. Section 6 concludes.

2 The data

2.1 The Invind survey

The Bank of Italy has conducted an annual business survey (hereafter, *Invind*) since 1972. The interviews are carried out in the first months of the year. They aim to collect quantitative data on the most important variables concerning the firm’s activity: investment, employment, turnover, together with other related indicators such as quota of investment actually realised compared with previous plans, variation in own prices, etc. The questionnaire is accompanied by many categorical variables.

The main characteristic of the survey is that it allows us to compute variations of quantitative variables such as investment within a single survey edition. In fact, quantitative investment data cover the previous two years and include a forecast for the current year: by this means, variations can be computed by using a single cross-section. Estimates of variations obtained from single surveys have proved much more stable than estimates obtained from adjacent surveys, often made unreliable by firms' structural changes, misclassification and measurement problems. Such sources of error are more easily kept under control within the same questionnaire.

The survey design is stratified with a single stage. The design strata are combinations of branches of activity, size classes and geographical areas (referring to the firms' head offices). The sample size is determined by first using the optimum allocation to strata that minimizes the variance of the means and variations of the main variables (employment, turnover and investments) and successively allocating the numbers obtained among regions and branches of activity according to the population size.

The weighting procedure assigns each firm an initial weight, given in each stratum by the ratio of the number of firms in the population to the number of firms in the sample (strata are formed by combinations of branch of activity and size classes). These weights are adjusted by post-stratification in order to align the weights to the geographical distribution of the firm population.

The sample is a panel, continuously updated and revised to take into account the attrition process.[1] Over the years both the sample size and the reference population have been broadened considerably: initially, only manufacturing firms with 50 employees and over were covered. Starting in 1999, the whole industrial sector (i.e. manufacturing plus energy and mining sector) for firms with at least 50 employees was covered. In 2001 industrial firms with 20-49 employees were added and since 2002 firms belonging to the non-financial private sector with at least 20 employees are also included. The sample has grown from the initial size of around 1,000 units to the current 4,000 (3,000 of which belong to industry and 1,000 to the service sector). The following Table 1 provides details of the sample and the reference population in terms of number of firms and number of employees for the period 1994-2007.

1 For further details on the design of the survey, see Bank of Italy, 2008.

Table 1. Bank of Italy's *Invind* survey,
sample and reference population, 1994-2007

(units)

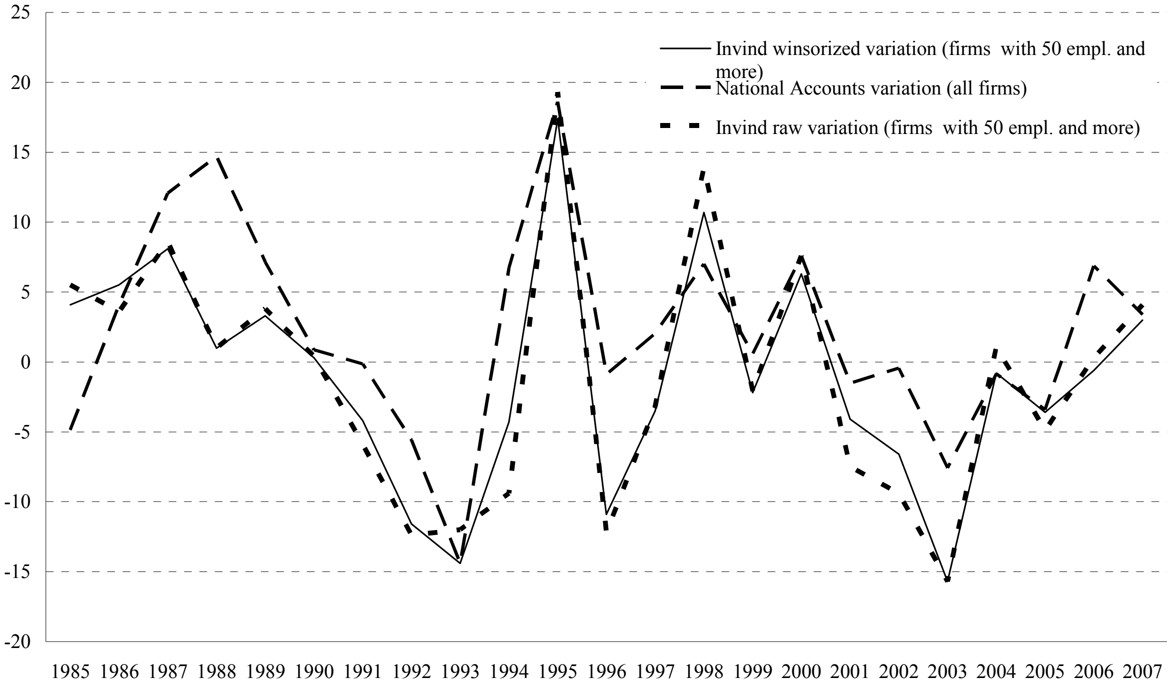
	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Sample size														
Industrial firms with 50 employees and over	---	---	---	---	---	1,135	1,474	1,764	1,852	1,905	1,916	1,950	1,900	1,852
- *Manufacturing*	953	996	1,060	1,002	998	1,107	1,428	1,713	1,797	1,848	1,861	1,890	1,838	1,785
Industrial firms with 20 employees and over (A)	---	---	---	---	---	---	---	2,786	3,046	3,141	3,150	3,227	3,109	2,980
- *Manufacturing*	---	---	---	---	---	---	---	2,713	2,959	3,048	3,065	3,133	3,010	2,878
Service firms with 20 employees and over (B)	---	---	---	---	---	---	---	---	923	994	1,076	1,159	1,143	1,083
Total (A)+(B)	---	---	---	---	---	---	---	---	**3,969**	**4,135**	**4,226**	**4,386**	**4,252**	**4,063**
Reference population size														
Industrial firms with 50 employees and over	---	---	---	---	---	11,708	12,031	12,625	12,025	12,251	1,1953	11,789	11,857	11,778
- *Manufacturing*	11,037	10,880	11,458	11,792	11,609	11,502	11,798	12,383	11,792	11,975	11,676	11,509	11,573	11,483
Industrial firms with 20 employees and over (A)	---	---	---	---	---	---	---	40,131	38,994	38,424	37,638	36,792	36,538	36,923
- *Manufacturing*	---	---	---	---	---	---	---	39,449	38,324	37,688	36,906	36,064	35,782	36,140
Service firms with 20 employees and over (B)	---	---	---	---	---	---	---	---	23,914	26,678	27,566	28,486	29,326	31,154
Total (A)+(B)	---	---	---	---	---	---	---	---	**62,908**	**65,102**	**65,204**	**65,278**	**65,864**	**68,077**

The investment variation for manufacturing firms with 50 employees (only for this type of firm can we produce a long time series) is simply obtained as:

$$\frac{\sum_i I_{it+1} W_i}{\sum_i I_{it} W_i} - 1 \qquad (1)$$

where, for the generic *i-th* firm belonging to the *Invind* survey's sample, W_i is the design weight, I_{it+1} and I_{it} are respectively the investment levels for the years *t+1* and *t*. Investments for the year *t* have been trimmed according to the method known as 'type II Winsorization', used in the official dissemination of the survey results. The method (Kocic and Bell, 1994; Smith et al., 2003) prevents the value of smaller firms that are outliers in terms of per capita investments from influencing the estimates too much. As clearly shown in Figure1, the effect of Winsorization on the estimate of formula (1) is quite limited. The figure also presents the time series of investment variation for all the manufacturing sector, derived from the Italian national accounts: we see that the firms with more than 50 employees belonging to the Bank of Italy sample determine the trend for the whole sector. Finally, we can observe that investment variation is extremely volatile and therefore difficult to forecast.

Figure 1. Italian manufacturing firms. Per cent variation of realised investment from the *Invind* survey (firms with 50 employees or more) and from the Italian national accounts 1985-2007 *(at 2007 constant prices)*



Source: Invind survey, Bank of Italy.

2.2. The Sondel survey

Since 1993 a business outlook survey (*Sondtel*) has also been carried out on the same sample as *Invind* survey. The interviews take place between 20th September and 10th October. Forecasts on the firm's specific activities are collected in qualitative form during a telephone interview lasting 15-20 minutes. The investment plans for the following year *t+1* are collected in terms of investment variation of *t+1* over *t*. Five ordered categories are used: "strong decrease" (less than -10%), "slight decrease" (-10% to -3%), "stable" (-3% to 3%), "slight increase" (+3% to +10%), "strong increase" (more than +10%).

Table 2 shows the information flow across the various survey occasions. For example, the *Sondtel* taking place in 2005 collected categorical data about the planned investment variation between 2005 and 2006. The corresponding realised investment levels for 2005 and 2006 were collected only a year and a half later in the *Invind* between January and April 2007. On the same occasion, planned investment levels for 2007 were also asked.

Table 2. Bank of Italy's *Invind* and *Sondtel* business surveys
Information provided during the years for investments

Reference years of variables	**Surveys**[(1)]									
	……….	*Invind* 2005	*Sondtel* 2005	*Invind* 2006	*Sondtel* 2006	*Invind* 2007	*Sondtel* 2007	*Invind* 2008	*Sondtel* 2008	……….
……….	……….									……….
2003		Realised investment level								
2004		Realised investment level		Realised investment level						
2005		Planned investment level		Realised investment level		Realised investment level				
2006			Categorical planned investment variation	Planned investment level		Realised investment level		Realised investment level		
2007					Categorical planned investment variation	Planned investment level		Realised investment level		
2008							Categorical planned investment variation	Planned investment level		
2009									Categorical planned investment variation	
……….	……….									……….

(1) For the surveys, the years are those of the interviews (first four months of the year for *Invind*, 20 September-10 October for *Sondtel*).

2.3 The relationship between plans and realisations

Sondtel investment plans can be concisely summarized by neglecting the “stable” response category and computing the difference between frequencies of increase and decrease. If we use INC_t to indicate the percentage of answers that report "slight increase" or "strong increase" at time t and DEC_t for the percentage of those reporting "slight decrease" or "strong decrease", the balance statistic is simply: $BAL_t = INC_t - DEC_t$.. It is traditionally used to present business surveys that attempt to forecast the short-term economic outlook (Goldrian, 2007) by simply measuring whether firms planning an increase exceed those planning a decrease.[2]

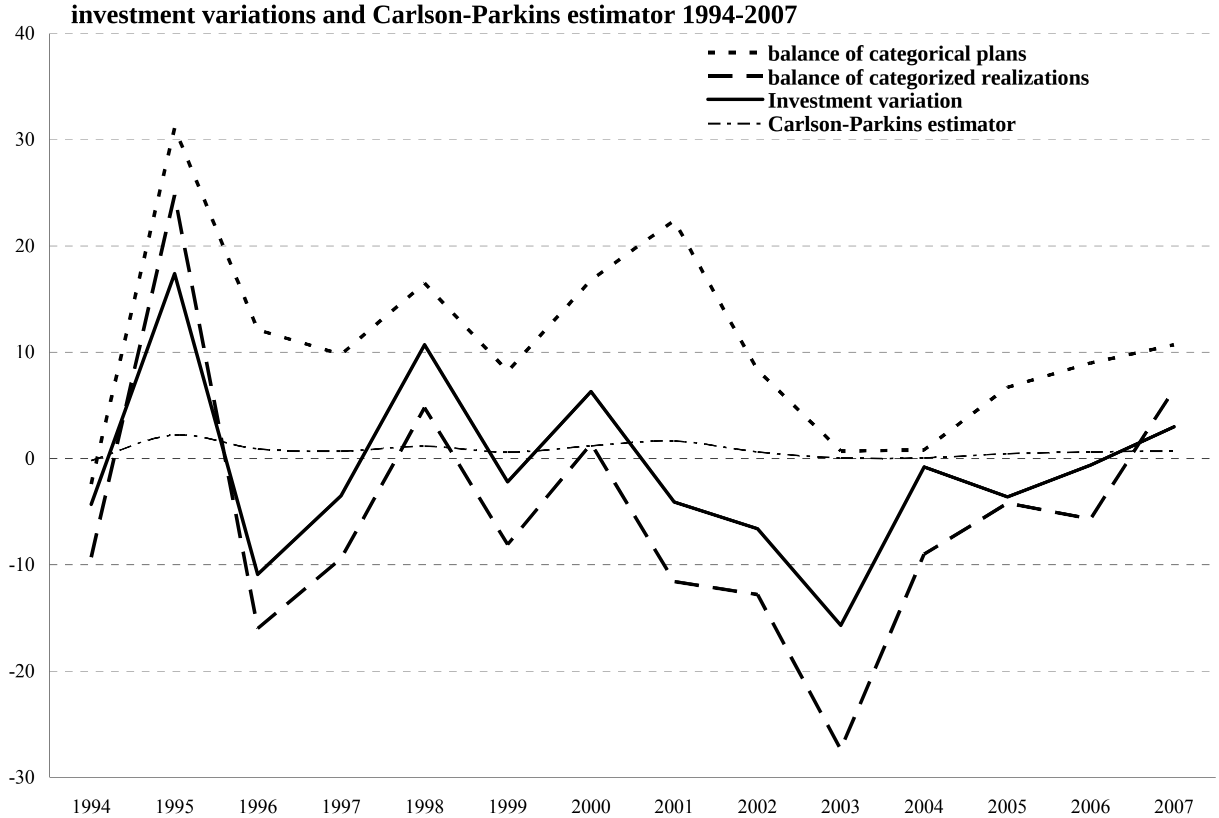
It might seem natural to compare the balance statistics with the corresponding *Invind* realised investment variations over the years. However, the two series are not directly comparable, since they refer to different units of measurement (respectively, difference of frequencies and percentage variations). Nevertheless, balances can provide a rough idea of the direction of investment variation with respect to

2 Going beyond this simple yet useful meaning, a rationale for their use is that the balance statistic is the expected value of a discrete aggregate probability distribution which locates answers in three points: -100, 0, 100 (expressing respectively decrease, stability and increase in percentage variation). The transformation assumes, a priori, the symmetry of answers: the distance between "increase" and "stable" is the same as that between "stable" and "increase".

turning points, accelerations or decelerations. The direct comparison of the two variables requires that quantitative realised investment variations be preliminarily transformed into a categorical variable like the *Sondtel* plans and the corresponding balance statistic be computed. The two balance series obtained can finally be compared directly. An alternative way of comparing categorical and quantitative data is to transform the first into quantities. The literature has proposed many methods (see Pesaran, 1984). We use the classical Carlson-Parkins method (Carlson and Parkins, 1975), suitably modified to take into account the fact that the categorical answers provided in *Sondtel* are associated with numerical intervals (see Appendix for further details).

Figure 2 shows the two balances, together with the quantitative realised investment variation, for the years 1994-2007 and the Carlson-Parkins quantification of *Sondtel* plans. The predictive capability of the balances of the categorical plans can be assessed by looking at the coincidence of its turning points with those of the series of the quantitative realisations: as we can see, discrepancies take place only for years of sharp and unforeseeable recessions, such as 2001. Moreover, if we compare the series of the two balances, plans seem systematically to overestimate realisations. As for the Carlson-Parkins estimator, it seems to lack any informative power because of its structurally limited variability.

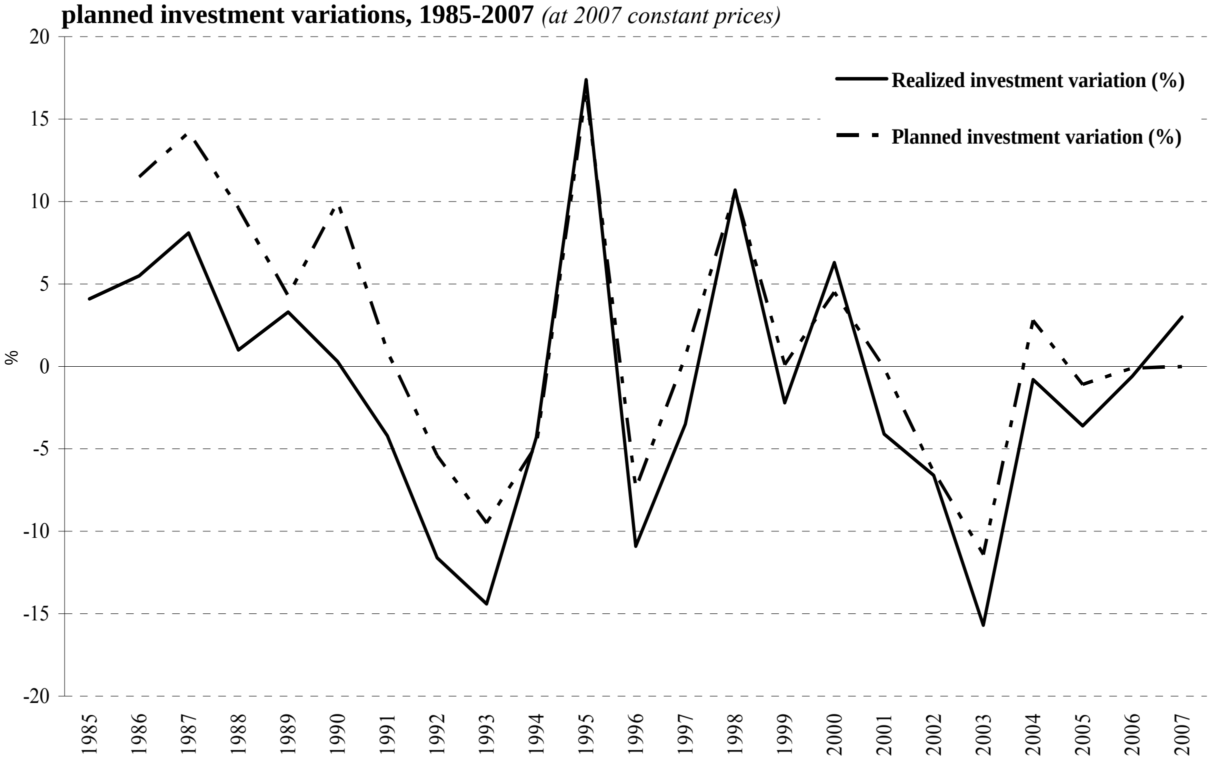
Figure 2. Italian manufacturing firms with 50 employees or more. Quantitative investment variations *(at 2007 constant prices)*, **balances of category investment plans and of categorized investment variations and Carlson-Parkins estimator 1994-2007**



Source: Invind and Sondtel surveys.

If we look at Figure 3 comparing the average realised investment variation with the average planned investment variation, both taken from *Invind* survey, we still detect a positive bias on average, although it seems much smaller, especially in more recent years.

Figure 3. Italian manufacturing firms with 50 employees or more. Quantitative realised and planned investment variations, 1985-2007 *(at 2007 constant prices)*



Source: Invind and Sondtel surveys.

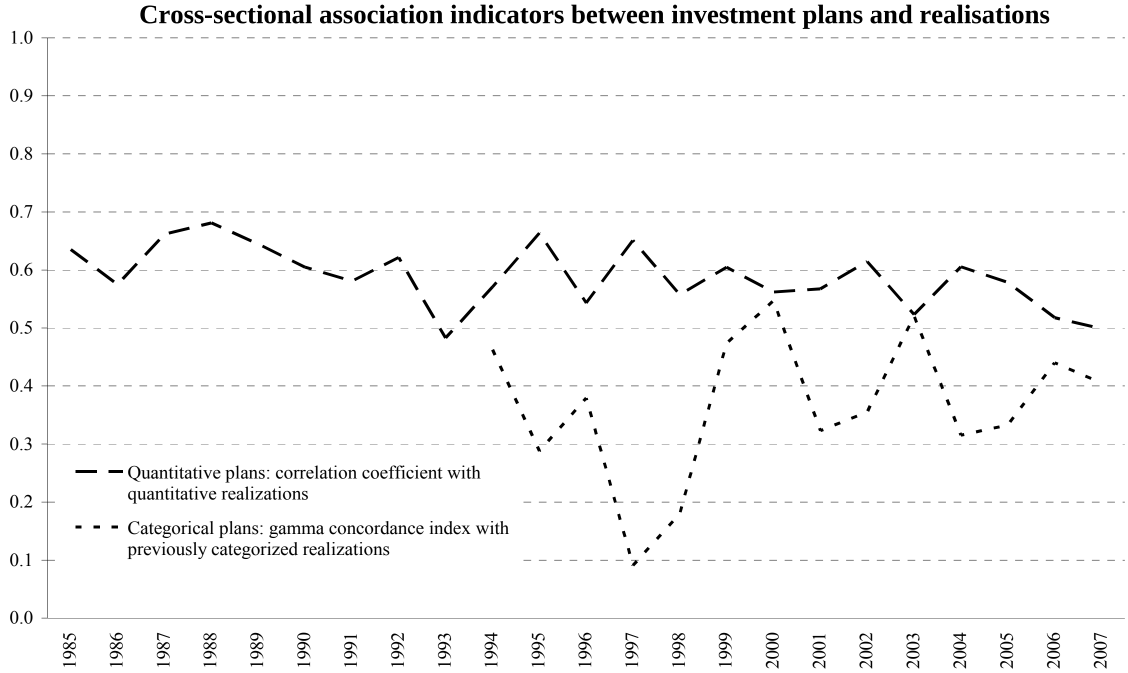
The smaller bias in investment plans in *Invind* survey compared with *Sondtel* is expected, since these plans are collected during the year of interest about six months later than the corresponding *Sondtel* plans.

The sources of positive bias in investment plans are multiple and not easily separable. A source is the tendency for firms to be over-optimistic about the outcome of planned action (Kahneman and Lovallo, 2003). More specifically, the strong tendency to regard every budgeting process as unique prevents planners from considering correctly all the historical data available for risk evaluation (Kahneman and Lovallo, 1993) and, as an aftermath, an optimistic bias of capital investment projects becomes recurrent.

Another source of the bias might stem from the survey timing: the month of September coincides with the start of the budgeting process, when exuberant moods, later revised, could prevail. Another factor could be a mechanism of "social desirability", which pushes the respondent to cast himself/herself in a favourable light by over-reporting a desirable attribute to the interviewer (Cannell et al., 1981). Moreover, a recent strand of industrial economics argues that the formulation of plans to be fulfilled exactly might not be the best entrepreneurial strategy. Misrepresentation could be chosen for strategic reasons (Flyvbjerg, 2003). More specifically, if accurate forecasting comes with heavy costs of information collection, entrepreneurs might deliberately overestimate future realisations so as to be able to diversify over different projects (Rötheli, 1998). Finally, there seem to be idiosyncratic factors in the Italian economy leading to positive forecasting errors in macroeconomic estimates, for example of GDP (see Batchelor, 2007): they could also play a role when dealing with business microdata of firm plans.

Despite the presence of a significant bias in investment plans, especially in qualitative plans formulated in the *Sondtel* survey, individual responses can be used efficiently to provide a one-year-ahead aggregate investment variation forecast. For instance, for every year we have measured the correlation coefficient between the firm-level realised investment variation and the corresponding *Invind* survey planned investment variation. The coefficient holds steady at 60 per cent (Figure 4). The correlation coefficient cannot be computed between *Sondtel* categorical plans and corresponding quantitative realisations. In this case we have calculated the gamma coefficient after categorizing the realised investment variation. This coefficient too is positive and significant (Figure 4), although always smaller than the correlation coefficient.

Figure 4. Italian manufacturing firms with 50 employees or more.
Cross-sectional association indicators between investment plans and realisations



Source: Invind and Sondtel surveys.

3 The forecasting procedure

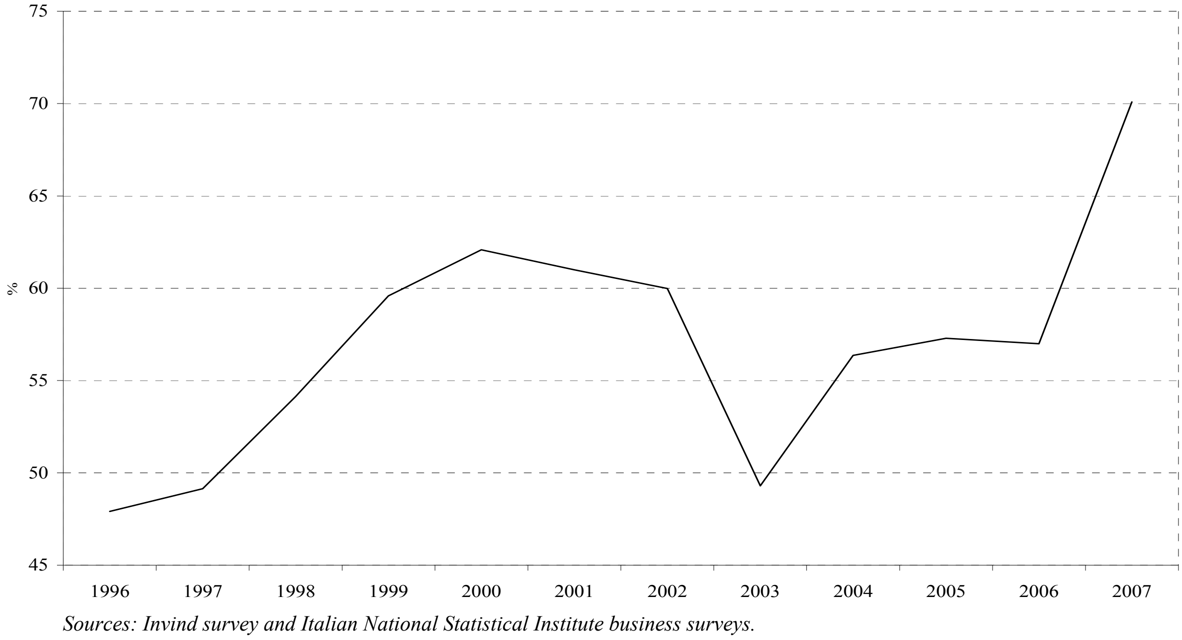
3.1 General features

The above findings show that investment plans can be used to gain insights into the future course of investment activity with respect to turning points, accelerations or decelerations. This section describes the procedure for forecasting investment growth using a panel data model that exploits all the heterogeneity among firms. The procedure should be used at the end of the year, when data from the latest *Sondtel* survey become available, in order to forecast the aggregate investment growth for the following year.

The forecasting procedure is based on three steps: in the first, a dynamic panel data model for the realised investment variation is estimated, with the *Sondtel* qualitative planned investment variation among the covariates; in the second step, the model parameter estimates are employed to produce a one-year-ahead prediction of firm-level investment variations; in the last step an aggregation procedure is used to compute the investment growth forecast for the entire economy.

In order to assure good consistency, the model is estimated on the manufacturing firms with 50 employees and over that have been continuously present in the survey: we therefore neglect the successive extensions of the reference population as they span too few survey editions. Even with this restriction, the estimation maintains an economic significance, as the sub-population of these firms represents on average 56 per cent of the total investment of the Italian industrial sector[3] (see Figure 5).

Figure 5. Share of Italian industrial investment made by manufacturing firms with 50 employees and over



Sources: Invind survey and Italian National Statistical Institute business surveys.

3.2 The model

Let us indicate with y_{it} the yearly investment variation between *t* and *t-1* for firm *i*:

$$y_{it} = \frac{I_{it}}{I_{it-1}}$$

[3] Around 95 per cent of the firms in the industrial sector operate in manufacturing, whereas the rest belong to the energy and extraction sub-sector.

Our starting point is a dynamic model of order *p* for panel data with y_{it} as dependent variable and the planned investment variation among the covariates:[4]

$$y_{it} = \beta_0 + \sum_{j=1}^{p} \beta_j y_{it-j} + \gamma' y_{it-1}^{ed} + \varepsilon_{it} \tag{2}$$

Since the planned investment variations are collected in discrete form in five categories (see Section 2) y_{it-1}^{ed} is a four-dimension vector of binary variables, $y_{it-1}^{ed} = \left(y_{it-1}^{--}, y_{it-1}^{-}, y_{it-1}^{+}, y_{it-1}^{++} \right)$, standing respectively for "strong decrease", "slight decrease", "slight increase" and "strong increase", with "stable" as reference category.

Let us suppose we are at the end of year *t*, when the model is to be employed. As Table 3 has shown, at that time the available data are:

- the planned investment variation between year *t* and *t+1* collected in *Sondtel* in year *t*;
- the planned investment variation between year *t-1* and *t* collected in *Invind* in year *t*;
- the investment variation between year *t-2* and *t-1* collected in *Invind* in year *t*;
- the investment variation between year *t-3* and *t-2* collected in *Invind* in year *t-1*.

At the end of year *t* the investment variation between year *t-1* and *t* is not yet available. The term can therefore be replaced by the corresponding planned investment variation collected in *Invind*. In fact, as we have shown in Section 2.3, the two variables are highly correlated, both across time at the aggregate level and within firms for every cross-section.

This is why we use the following alternative specification, with the one-year lagged investment variation, y_{it-1}, substituted by the corresponding planned investment variation from *Invind*, y_{it-1}^{eq}:

$$y_{it} = \beta_0^* + \beta_1^* y_{it-1}^{eq} + \sum_{j=2}^{p} \beta_j^* y_{it-j} + \gamma^{*\prime} y_{it-1}^{ed} + \varsigma_{it} \tag{3}$$

Equation (3) can also be regarded as the reduced form of a two-equation system, where the first component is equation (2) and the second one models the relationship between realisations and quantitative plans collected in *Invind*:

---

[4] We have chosen to model y_{it} directly instead of its logarithm. This choice was supported by the results of an exploratory analysis that estimated equation (2) with $\log(y_{it})$ instead of y_{it} and then computed predictions expressed as $\exp\left(\hat{y}_{it-1} + \frac{1}{2}\hat{\sigma}^2(1) \right)$, where the second term inside the exp operator is half the error variance of a one-step-ahead forecast. These predictions were considerably less stable than those obtained without the transformation (see also Lutkepohl and Xu, 2009, for evidence supporting these findings in the modelling of monthly inflation data series).

$$\begin{cases} y_{it} = \beta_0 + \sum_{j=1}^{p} \beta_1 y_{it-j} + \gamma' y_{it-1}^{ed} + \varepsilon_{it} \\ y_{it-1} = \alpha_0 + \alpha_1 y_{it-1}^{eq} + \zeta_{it-1} \end{cases} \tag{4}$$

where the error term of equation (3) is $\varsigma_{it} = \varepsilon_{it} + \beta_1 \zeta_{it-1}$.

Equation (3) is our baseline specification (model *M0*) for the forecasting procedure, which we progressively enrich according to the following Table 3. Model *M1* adds to *M0* time-invariant effects referring to economic sector, geographical location of the headquarters and size, so as to control for heterogeneity in the means of the y_{it} series across sectors, geographical areas and to capture investment behaviour and financial constraints differentiating small and large firms. As an alternative to *M1*, *M2* adds to *M0* two binary variables indicating whether in the previous year the investment plans were above or below the realised investment variations: they record the prediction performance of past qualitative plans. *M3* simply combines the regressors of *M1* and *M2* and, finally, *M4* adds the real growth rate of turnover from sales to the set of *M3* regressors. [5]

Table 3. Model specifications

M0	Lagged dependent variable (investment growth rate[a]) Category investment plans: 1. Strong decrease 2. Slight decrease 3. Slight increase 4. Strong increase
*M1**	All the regressors in M0 plus Branch: 1. Textile, clothing, leather, shoes 2. Engineering Area of location[+] 1. Northern Italy 2. Central Italy Number of employees: 1. Between 50 and 199 2. Between 200 and 499
*M2**	All the regressors in M0 plus Prediction error of investment plans in the previous year 1. Over-planning of realised investment 2. Under-planning of realised investment
*M3**	M1+M2
*M4**	All the regressors in M3 plus Growth rate of turnover from sales
MM0	Category investment plans: 1. Strong decrease 2. Slight decrease 3. Slight increase 4. Strong increase

(a): Investment plan for the first lag.
(*): All the variables added in models M1-M4 refer to the year immediately preceding that of the dependent variable.
(+): Geographical area is defined by the location of the firm's headquarters.

[5] We tried many other variables among the regressors, including macroeconomic indicators at national and local level, such as interest rate and growth rate of product (properly lagged), but they all failed to have significant explicative power.

Therefore, five model specifications (*M0-M4*) are compared in terms of forecasting accuracy to determine which produces the best forecast of the one-year ahead aggregate investment variation. We also try a specification *MM0* using only the categorical investment plans from the most recent Sondel in order to gauge their usefulness in the model.

3.3 The weighted estimation and the forecasting procedure

For the estimation of the dynamic panel data model we use a set of weights that control for the survey design:

$$\overline{W}_{it-1} = W_i^* \cdot I_{it-1} \tag{5}$$

where W_i^* is the design weight adjusted for the panel attrition[6] and I_{it-1} is the investment level for the year *t-1* as collected in the cross-section *t*. The investment level at time *t-1* graduates the units' contribution to the estimate according to their investment size.

Moreover, the estimation of standard deviation of the coefficients should take into account the survey design. For this purpose, we apply the DuMouchel and Duncan's procedure (1983). Let us indicate with:

$$y = X\beta^* + \varepsilon \tag{6}$$

and with:

$$\hat{\beta}^* = N_p\left(\beta^*, \Sigma\right) \tag{7}$$

respectively the forecasting model in compact matrix form and the relative estimated coefficients. Moreover, we use the following symbols:

W: diagonal matrix containing the weights on the main diagonal,

I: identity matrix,

k: number of regressors,

n: number of observations used in the model estimation.

We define:

$$D = \left(X'WX\right)^{-1} X'W - \left(X'X\right)^{-1} X' \tag{8}$$

and

$$K = I - X\left(X'X\right)^{-1} X' - D'\left(D'D\right)^{-1} D \tag{9}$$

By using (8) and (9) an estimator for the variance of the residuals is:

[6] Subsection 4.1 is dedicated to an exhaustive discussion of the panel attrition problem in our dataset.

$$\hat{s}_E^2 = \frac{y'Ky}{(n-2k)} \tag{10}$$

and the variance/covariance matrix of the model coefficients *Σ* is accordingly estimated as:

$$\hat{\Sigma} = (X'WX)^{-1}(X'W^2X)(X'WX)^{-1}\hat{s}_E^2 \tag{11}$$

Once the distribution of $\hat{\beta}^* = N_p(\beta^*, \Sigma)$ is estimated, Choleski's decomposition $\Sigma = TT'$ generates 5,000 drawings from the distribution of *β*. The individual prediction for unit *i* of the investment variation between times *t* and *t+1* can then be expressed as:

$$\hat{y}_{it+1} = \frac{1}{5000}\sum_{j=1}^{5000} Z_{it}\hat{\beta}^{(j)} \tag{12}$$

where Z_{it} indicates the model regressors.

As a consequence, a consistent predictor for the aggregate investment variation is obtained as the weighted average of equation (12) over all the units:

$$\frac{\sum_i \hat{y}_{it+1}\overline{W}_{it}}{\sum_i \overline{W}_{it}} \tag{13}$$

The estimator (13) can now be compared with the realised investment variation to be forecast:[7]

$$\frac{\sum_i I_{it+1}W_i^*}{\sum_i I_{it}W_i^*} \tag{14}$$

which can also be written as:

$$\frac{\sum_i \frac{I_{it+1}}{I_{it}}\overline{W}_{it}}{\sum_i \overline{W}_{it}} \tag{15}$$

The similarity between (15) and (13) is now unambiguous: the expression (13) clearly estimates (14) by replacing individual planned investment variations $\frac{I_{it+1}}{I_{it}}$ with corresponding individual estimates, as expressed in (12).

[7] The differences between the design weights corrected for the sample attrition for the years *t* and *t+1* are negligible, since the population is stable in the two years.

4 From data to model

4.1 Panel attrition

Panel attrition can derive either from a decision of the firm, which is no longer willing to participate, or from the firm leaving the reference population. This latter can happen for a variety of reasons (mergers or acquisitions, number of employees dropping below the threshold level, economic activity no longer within those envisaged in the target population, bankruptcy, etc.).

Since the attrition process affecting the two surveys determines incomplete information for every sample unit over the years, two strategies can be followed for model estimation.

According to the first, the model could be estimated on the pooled sample of firms. However, we cannot make a simple data pooling. In fact, in case of a dynamic panel model of order 1, every investment plan from *Sondtel* needs to be exactly matched with the corresponding realisation and the one-year lagged quantitative plan, coming from two consecutive *Invind* editions: therefore every unit in the pooled dataset would need complete data from three surveys. If the order of the autoregressive model is higher (for example three), for every unit we would have to recover data from five surveys (four consecutive *Invind* editions and one *Sondtel* edition). Therefore, even these short panels entail a non-negligible loss of information deriving from the inevitable attrition.

The alternative strategy uses a sample of firms for which plans and realisations can be found without gaps over a long time span. The advantage of this approach is that we would have standard panel datasets at the cost of an additional loss of information caused by the smaller number of units considered.

In terms of estimation procedures, the econometric literature suggests numerous techniques for estimating, testing and validating models for panel data. We can find various estimation methods for the balanced one-way or two-way random effects model (Baltagi and Song, 2006), but in the case of unbalanced panels, the available literature is more parsimonious. Moreover, few estimation procedures for dynamic panel data models are feasible for unbalanced panels: such methods are quite uncommon (Bruno, 2005) as well as quite complex and based on strong assumptions (Moffitt, 1993; Collado, 1997; Lokshin, 2008). Verbeek and Vella (2005) have also shown that these assumptions are not trivially satisfied in applied works.

For these reasons, and also because our model has to be used for generating one-year-ahead forecasts of aggregate investment variation, we prefer to employ robust estimation methods and therefore opt for balanced panels.

The following Table 4 shows the shrinkage of the sample size in terms of number of units and investment value if we use the balanced panels needed for our model.

Table 4. Attrition process for the *Invind* survey
Firms and total investment in cross-sections and corresponding share for balanced panels (1)
(Manufacturing firms with 50 employees and over)

Cross-sections		Panel length				
		4	5	6	7	8
	Number of firms	Sample size of the balanced panels (% of the cross-sectional value)				
1994	953	26.3	22.1	17.7	14.7	13.0
1995	996	26.7	21.7	17.9	15.5	13.2
1996	1,060	24.5	20.4	17.9	14.9	12.9
1997	1,002	25.2	22.0	18.4	15.7	13.8
1998	998	24.9	20.6	17.6	15.3	12.7
1999	1,107	22.7	19.0	16.0	13.6	11.4
2000	1,428	17.6	14.7	12.5	10.4	9.2
	Investment value (million of euro)	Investment value of the balanced panels (% of the cross-sectional value)				
1994	5,467	46.7	44.1	38.9	33.6	32.5
1995	6,491	55.6	48.7	37.3	35.3	33.1
1996	6,432	51.7	40.9	39.3	35.9	33.0
1997	6,096	42.4	40.1	36.0	33.5	32.3
1998	6,871	44.0	36.8	34.1	32.8	27.3
1999	7,779	41.0	38.4	36.4	31.1	28.5
2000	10,477	38.5	36.9	33.8	31.7	31.0

(1) The balanced panels are composed of units found in the cross-section of the year in every row and beyond, for a number of years equal to the panel length

The balanced panel is composed of the units found both in the cross-section of the year in every row and beyond, for a number of years equal to the panel length. The attrition has a heavy impact on the number of firms, but is less damaging in terms of share of total investment: this occurs because big firms, which are the most important for the kind of estimate we perform, tend to remain in the panel.

As for panel length, there is a clear trade-off between short panels with a large cross-sectional dimension and long ones, with a relatively small number of units. Short panels could not feature the regularities needed for reliable forecasts and represent behaviour too idiosyncratic of single years. Longer panels, however, could be distorted by sample selection mechanisms. A balance must therefore be struck between these two extremes.

In this regard, we carry out an analysis of panel attrition caused by the utilization of balanced panels of different lengths. An easy way to do this is to run a simple dummy regression with realised investment variation as dependent and a dummy indicating whether the unit belongs to a panel, together with the complete sample design variables acting as control covariates.[8] More formally, for each separate cross-sectional survey, we estimate the following equation:

$$y_{it} = \pi_0 + \pi_1 d_{it} + \pi_2 Z_{it}$$

The sub-script i indicates the generic unit belonging to the cross-section for year t and d_{it} is simply:

[8] They are: geographical area of the firm's administrative headquarters (North-West, North-East, Centre, South and Islands), class size (50-99, 100-199, 200-499, 500-999, 1000 and more, in number of employees), sector of economic activity (food products, beverages and tobacco; textiles, clothing, hides and leather; chemicals, rubber and plastic; non-metal minerals; engineering; other manufacturing). In accordance with the survey design, the class sizes and the sectors of activity are interacted.

$$d_{it} = \begin{cases} 1 & \text{if the unit } i \text{ belongs to the panel} \\ 0 & \text{otherwise} \end{cases}$$

Z_{it} is the vector of dummies representing the survey-design variables for unit i.[9] This procedure is simpler and more intuitive than the classical Heckman two-step procedure, which produces similar results however, not shown here for brevity.[10] The balanced panel would therefore be a source of bias if the coefficient π_1 were significant: in such a case the selection mechanism would not be controlled by the survey design variables. The number of cases where such a coefficient is significant for every sample cross-section is quite limited (Table 5).

A significant panel attrition is present in 1996, 1999 and 2001: in the latter two years, considerable increases in the sample size took place (see Table 1), which certainly contributed to produce this effect.[11] However, using longer panels does not increase the risk of panel attrition.

We therefore opt for eight-year panels. This length allows the implementation of complex statistical tests and estimators for dynamic panel data models. The first seven years of each panel are required for model estimation, whereas the last year is set aside for the evaluation of out-of-sample forecasting.

Table 5. *Invind* and *Sondtel* surveys: number of panels affected by sample selection bias and numbers of panels considered for every combination of panel length and survey year

Survey year	Panel length					Total
	4	5	6	7	8	
	Number of panels with sample selection and number of panels considered (1)					
1994	1 , 0	0 , 1	0 , 1	0 , 1	0 , 1	**0 , 5**
1995	2 , 0	0 , 2	0 , 2	0 , 2	0 , 2	**0 , 10**
1996	3 , 1	1 , 3	1 , 3	0 , 3	0 , 3	**3 , 15**
1997	4 , 0	0 , 4	0 , 4	0 , 4	0 , 4	**0 , 20**
1998	4 , 0	0 , 5	0 , 5	0 , 5	0 , 5	**0 , 24**
1999	4 , 3	4 , 5	4 , 6	1 , 6	0 , 6	**12 , 27**
2000	4 , 0	0 , 5	0 , 6	1 , 7	1 , 7	**2 , 29**
2001	4 , 0	1 , 5	1 , 6	3 , 7	4 , 7	**9 , 29**
2002	4 , 0	0 , 5	1 , 6	0 , 6	0 , 6	**1 , 27**
2003	4 , 0	1 , 5	0 , 5	0 , 5	1 , 5	**2 , 24**
2004	4 , 1	0 , 4	0 , 4	0 , 4	0 , 4	**1 , 20**
2005	3 , 0	0 , 3	0 , 3	0 , 3	1 , 3	**1 , 15**
2006	2 , 0	0 , 2	0 , 2	0 , 2	0 , 2	**0 , 10**
2007	1 , 0	0 , 1	1 , 1	0 , 1	0 , 1	**1 , 5**

(1) Significance at the 5 per cent level.

[9] Every unit is weighted by the product of the design weight and the investment level at time $t-1$: in this way firms are scaled according to the size of their contribution to the time $t-1$ estimated total investment level.

[10] With the Heckman approach we modelled a first equation for the inclusion in one of the panel samples and a second one having the realised investment variation as dependent.

[11] Results for the years 1996, 1999 and 2007 are computed after an outlier detection: units lying to the left of the 2nd percentile or to the right of the 98th one of the distribution of the dfbeta indicator for the coefficient π_1 are excluded. The dfbeta indicator is a commonly used regression diagnostic indicator that ranks the observations according to their contribution to the coefficient size. It is obtained as the difference between the regression coefficient calculated for all of the data and the regression coefficient calculated with the observation deleted, scaled by the standard error calculated with the observation.

4.2 Econometric issues

Given the high degree of similarity between y_{it} and y_{it-1}^{eq}, discussed at length in Section 3, we can treat equation (3) as a classical dynamic panel data model, where the first lag of the dependent variable is replaced by its quantitative forecast.

We explicitly write the disturbance term of equation (3) as the sum of an individual-specific time-invariant effect μ_i and a pure disturbance term θ_{it} : $\varsigma_{it} = \mu_i + \theta_{it}$.

If individual effects exist, the use of *GMM* would be required, after first-differencing the equation to solve. On the contrary, if individual effects follow a degenerate distribution, *OLS* estimators on the original equation are consistent and more efficient than *GMM* ones, on the assumption that the error term is serially uncorrelated. Testing for the presence of individual effects is therefore a necessary step. The fact that the dependent variable is a variation, instead of a level, is already a good clue to the fact that the individual effects might be absent.

Holtz-Eakin (1988) proposed a very simple Sargan-difference test for the presence of individual effects for the purely first-order autoregressive model, which can be generalized to account for the presence of additional lagged values of the dependent variable and both endogenous, predetermined and time-invariant regressors. Through Montecarlo simulations Jimenez-Martin (1998) showed that the test lacks power when the coefficient of the lagged dependent variable tends to unity, whereas additional regressors sharply increase the power of the test (Jimenez, 1998).

Since the Holtz-Eakin's test is based on the assumption that the error term θ_{it} is serially uncorrelated, this assumption must first be tested. The complex structure of the error term (see equation 7), obtained by the linear combination of two disturbance terms separated by a time lag, justifies this caution. For this purpose Arellano and Bond (1991) propose a simple direct test, based on the error term of the model expressed in first-differences: the consistency of the *GMM* estimators relies upon the assumption that $E\left(\Delta\theta_{it} \cdot \Delta\theta_{it-2}\right) = 0$. A test for lack of second-order serial correlation in the first-difference residuals can be done in two ways: 1) by using residuals of the model on differences, and 2) by exploiting residuals of the equations in differences of the system model. We prefer the first solution, since the second is more efficient but is conditional on the assumption of absence of individual effects.

The Holtz-Eakin's test is carried out for all the combinations of lags for the dependent variable and model specification, in order to implement it on an appropriate number of lags and avoid committing a type II error. Given the limited panel length, the number of lags *p* for the dependent variable can be *1*,*2*, or *3*.

As shown in Table 6, the presence of individual effects can be ruled out for all the specifications and all the lags of the autoregressive component. The hypothesis of a lack of second-order serial correlation in the first-difference residuals is not always supported by data, however: it is strongly rejected for most of the specifications, especially for the middle panels (1997-2003 and 1998-2004) . The result is explained, however, by the large increase in the sample size in the years 1999-2001 which brought some instability. the strongest element supporting the validity of the Holtz-Eakin test is that no model specification rejects, for all the possible panels, the hypothesis of a lack of second-order serial correlation in the first-difference residuals.

5 The forecasting performance

We have previously identified six model specifications (*M0-M4, MM0*) that can be employed for the one-year-ahead forecasting. All but *MM0* use a maximum of three lags for the dependent variable, with a total of 16 different forecasting models (15 specifications for *M0-M4* and one for *MM0*). The objective of this section is to select the best one to forecast the one-year-ahead investment variation. In order to do so, we compare the out-of-sample forecasting performance of the 16 models in terms of bias.

Since our main concern is to remove the bias of investment plans, unbiasedness matters more than efficiency in our appraisal of model forecasting power: we therefore select the models with the smallest bias, provided they are also satisfactory, albeit not optimal, in terms of the forecast standard error.

We use the first seven years of each panel for OLS estimation and the last year only for the out-of-sample forecasting performance analysis. The forecasting period refers to the years 2001-2007.

Results are reported in Table 7. The last row at the bottom of each of the two sections of the table shows the overall forecasting performance of the 16 model specifications across all the panels.

In terms of bias, if we consider only models with $p=1$, the best specification turns out to be *M1*, which reduces the squared bias by 9 to 22 per cent compared with the other specifications. Adding more lags never produces any improvement in terms of bias, and model specification *M1* remains the best one. We therefore manage to contain the model bias simply by using the proxy of the first lag of the dependent variable, together with the survey design variables.

We therefore choose this specification because it minimizes the bias with a manageable increase of the standard error in comparison with the best specification from this point of view (*MM0*).

Table 8 presents, for brevity, the parameter estimates for model (*M1*, $p=1$) for all the rolling panels. The most important variables are those relating to the investment plans, with the most important role played by the categories “strong decrease” and “strong increase”.

There is a significant heterogeneity in the estimated parameter values across the panels. The first order autoregressive coefficient varies between -0.19 and -0.01 and its trend is decreasing over time (i.e. across the panels). At the same time, the magnitude of the coefficients relating to categorical investment plans becomes slightly higher over time. This result might be explained by the growing uncertainty making investment dynamics more erratic (also confirmed by the decreasing values of the adjusted R^2), with categorical investment plans possessing a greater predictive power than lagged investment variations.

Table 6. Results from Holtz-Eakin test for the presence of unobserved individual effects

	M0			*M1*			*M2*			*M3*			*M4*		
Panel	*p=1*	*p=2*	*p=3*	*p=1*	*p=2*	*p=3*	*p=1*	*p=2*	*p=3*	*p=1*	*p=2*	*p=3*	*p=1*	*p=2*	*p=3*
1994-2000	0.115 (-1.473)	0.022 (0.974)	-0.007 (-1.815)	0.006 (-0.394)	-0.188 -1.379	0.021 (-3.006)	0.100 (-1.210)	0.069 (0.550)	0.044 (-0.125)	0.006 (-0.394)	-0.188 -1.379	0.021 (-3.006)	0.101 (-1.210)	0.070 (0.551)	0.045 (-0.124)
1995-2001	0.012 (0.249)	0.147 (2.280)	0.072 (1.672)	0.095 (2.834)	0.010 (2.450)	0.061 (1.881)	0.206 (0.328)	0.052 (2.920)	-0.042 (1.614)	0.096 (2.833)	0.010 (2.450	0.061 (1.880)	0.206 (0.328)	0.053 (2.092)	-0.042 (1.614)
1996-2002	0.000 (-1.362)	0.000 (-1.013)	0.146 (-0.250)	0.152 (-2.573)	0.002 (-0.119)	0.000 (-1.642)	0.010 (-2.162)	0.203 (-1.395)	0.000 (-0.693)	0.153 (-2.572)	0.003 (-0.118)	0.001 -1.643	0.011 (-2.162)	0.204 (-1.394)	0.001 (-0.692)
1997-2003	0.051 (-2.466)	0.097 (-2.106)	0.092 (-2.453)	0.004 (-4.517)	-0.088 (-2.525)	-0.030 (-1.842)	0.052 (-8.112)	0.113 (-2.804)	0.117 (-3.723)	0.004 (-4.517)	-0.088 (-2.525)	-0.029 (-1.842)	0.053 (-8.112)	0.113 (-2.803)	0.118 (-3.723)
1998-2004	0.135 (-2.155)	0.034 (-2.336)	0.002 (-2.140)	0.334 (-2.505)	0.011 (-3.212)	-0.016 (-2.792)	0.011 (-3.285)	0.379 (-2.694)	0.243 (-4.402)	0.335 (-2.504)	0.012 (-3.211)	-0.015 (-2.792)	0.011 (-3.284)	0.380 (-2.693)	0.244 (-4.401)
1999-2005	0.061 (-0.379)	0.940 (0.030)	0.071 (-0.787)	0.016 (-1.473)	0.007 (-1.406)	0.002 (-1.341)	0.090 (-2.707)	0.036 (-1.838)	0.000 (-2.161)	0.017 (-1.473)	0.007 (-1.405)	0.003 (-1.340)	0.090 (-2.707)	0.037 (-1.837)	0.000 (-2.161)
2000-2006	0.140 (0.594)	0.045 (1.311)	0.004 (1.845)	0.178 (-0.898)	0.077 (-0.257)	0.108 (1.050)	0.041 (1.742)	0.077 (0.811)	0.098 (2.753)	0.179 (-0.898)	0.078 (-0.256)	0.108 (1.050)	0.041 (-1.742)	0.077 (0.811)	0.098 (2.753)
d.f.	7	6	5	7	6	5	7	6	5	7	6	5	7	6	5
χ^2_{df} at 5 %	14.067	12.592	11.070	14.067	12.592	11.070	14.067	12.592	11.070	14.067	12.592	11.070	14.067	12.592	11.070

Sample size: 124 (1994-2000), 131 (1995-2001), 137 (1996-2002), 138 (1997-2003), 127 (1999-2005), 131 (2000-2006).

Table 7. Squared bias, and Standard Error of one-year-ahead forecasts

		Specifications															
		M0			*M1*			*M2*			*M3*			*M4*			*MM0*
		number of lags			number of lags			number of lags			number of lags			number of lags			
		1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	
Years	Sample size	Squared Bias															
2001	124	0.0002	0.0016	0.0000	0.0004	0.0026	0.0003	0.0031	0.0043	0.0010	0.0019	0.0059	0.0022	0.0018	0.0077	0.0035	0.0018
2002	131	0.0176	0.0177	0.0299	0.0159	0.0189	0.0288	0.0270	0.0434	0.0394	0.0297	0.0403	0.0414	0.0347	0.0396	0.0444	0.0152
2003	137	0.0061	0.0077	0.0118	0.0028	0.0037	0.0093	0.0063	0.0036	0.0108	0.0036	0.0057	0.0138	0.0022	0.0039	0.0110	0.0058
2004	138	0.0128	0.0092	0.0093	0.0135	0.0048	0.0041	0.0135	0.0060	0.0099	0.0117	0.0030	0.0030	0.0131	0.0036	0.0132	0.0136
2005	127	0.0013	0.0001	0.0001	0.0000	0.0025	0.0016	0.0022	0.0001	0.0015	0.0002	0.0010	0.0005	0.0000	0.0017	0.0001	0.0005
2006	126	0.0027	0.0017	0.0076	0.0012	0.0016	0.0016	0.0030	0.0023	0.0044	0.0020	0.0008	0.0017	0.0020	0.0001	0.0031	0.0040
2007	131	0.0312	0.0399	0.0422	0.0233	0.0272	0.0306	0.0185	0.0247	0.0322	0.0137	0.0191	0.0206	0.0116	0.0144	0.0196	0.0344
Average		**0.0103**	**0.0111**	**0.0144**	**0.0082**	**0.0088**	**0.0109**	**0.0105**	**0.0121**	**0.0142**	**0.0090**	**0.0108**	**0.0119**	**0.0093**	**0.0101**	**0.0136**	**0.0108**
Years	Sample size	Standard Error															
2001	124	0.3632	0.3393	0.3195	0.4754	0.5094	0.4875	0.3119	0.4036	0.4918	0.4416	0.5272	0.5892	0.4503	0.5088	0.8664	0.2500
2002	131	0.3370	0.3138	0.3130	0.3795	0.4102	0.4816	0.2702	0.3596	0.4834	0.3873	0.4374	0.5246	0.3833	0.4545	0.5387	0.2417
2003	137	0.3915	0.3641	0.3214	0.4213	0.4523	0.4770	0.3521	0.3829	0.4837	0.4257	0.4658	0.5333	0.4273	0.4672	0.6463	0.3069
2004	138	0.3633	0.3132	0.2563	0.3419	0.3724	0.3996	0.3455	0.3685	0.4042	0.4182	0.4481	0.4989	0.4084	0.4329	0.4641	0.2506
2005	127	0.3962	0.3666	0.3184	0.3791	0.4085	0.4614	0.3713	0.4062	0.4366	0.4493	0.4885	0.5441	0.4450	0.4969	0.5364	0.2536
2006	126	0.3971	0.3699	0.3292	0.3503	0.3754	0.4486	0.3426	0.3826	0.4472	0.4259	0.4609	0.5118	0.4165	0.4471	0.5225	0.2775
2007	131	0.3808	0.3514	0.3137	0.3333	0.3538	0.3975	0.3339	0.3719	0.4523	0.3867	0.4098	0.4822	0.6545	0.7192	0.5765	0.2629
Average		**0.3756**	**0.3455**	**0.3102**	**0.3830**	**0.4117**	**0.4505**	**0.3325**	**0.3822**	**0.4570**	**0.4192**	**0.4625**	**0.5263**	**0.4551**	**0.5038**	**0.5930**	**0.2633**

Table 8. Parameter estimates of model *M1*, *p=1*

Panel	***1994 2000***	***1995 2001***	***1996 2002***	***1997 2003***	***1998 2004***	***1999 2005***	***2000 2006***
Intercept	1.2279***	1.1245***	0.9309***	0.9119***	0.7504***	0.9563***	0.9358***
	(0.1972)	(0.1723)	(0.2243)	(0.1852)	(0.1983)	(0.1882)	(0.181)
Quantitative plan for y_{t-1}	-0.1951**	-0.0997	-0.0717	-0.0118	-0.0197	-0.0236	-0.0173
	(0.0812)	(0.0689)	(0.0930)	(0.0803)	(0.0882)	(0.0933)	(0.0965)
Sondtel plans:							
Strong decrease	-0.2586	-0.1709	-0.3928**	-0.3404**	-0.3761**	-0.3942**	-0.4063**
	(0.2104)	(0.1418)	(0.1614)	(0.1575)	(0.1606)	(0.1618)	(0.1608)
Slight decrease	-0.0761	-0.1080	-0.1131	-0.1325	-0.0767	0.0430	0.0148
	(0.3817)	(0.2125)	(0.2430)	(0.1797)	(0.1685)	(0.1639)	(0.1527)
Slight increase	0.2222	0.1399	0.1051	0.1533	0.2089	0.1796	0.1367
	(0.2648)	(0.2388)	(0.2612)	(0.2412)	(0.2297)	(0.2300)	(0.2236)
Strong increase	0.5063**	0.4960***	0.5666***	0.5621***	0.7357***	0.6039***	0.5912***
	(0.2024)	(0.1722)	(0.1902)	(0.1717)	(0.1428)	(0.1473)	(0.1358)
Branch:							
Textile, clothing, etc.	-0.1026	0.0417	0.0825	0.1382	-0.0088	0.0170	0.0576
	(0.1367)	(0.1123)	(0.1223)	(0.1150)	(0.1098)	(0.1017)	(0.1033)
Engineering	-0.1928	-0.0896	-0.0592	-0.0503	-0.0475	-0.1033	-0.0861
	(0.1291)	(0.1104)	(0.1129)	(0.1061)	(0.1024)	(0.1011)	(0.0995)
Area of location:							
North	0.1150	0.0076	0.0938	0.0218	0.1844*	0.0011	-0.0099
	(0.1293)	(0.1168)	(0.1090)	(0.1028)	(0.1025)	(0.1039)	(0.0917)
Centre	0.0474	0.0393	0.1182	0.1343	0.2172*	0.1036	0.0721
	(0.1536)	(0.1369)	(0.1488)	(0.1356)	(0.1307)	(0.1226)	(0.1119)
Firm size:							
Between 50 and 199	0.0203	0.0905	0.3395***	0.2373**	0.1904*	0.1594	0.1438
	(0.1443)	(0.1123)	(0.1188)	(0.1012)	(0.1138)	(0.1063)	(0.1034)
Between 200 and 499	0.2834	0.1597	0.2084	0.0966	0.2033	0.0790	0.1418
	(0.1490)	(0.1197)	(0.1361)	(0.1314)	(0.1292)	(0.1128)	(0.1107)
Adj. R^2	0.3825	0.4721	0.3583	0.3105	0.3994	0.1657	0.0451
Durbin-Watson Test	1.4718	1.9849**	1.7321	1.8245	1.7098	1.9219**	1.8176

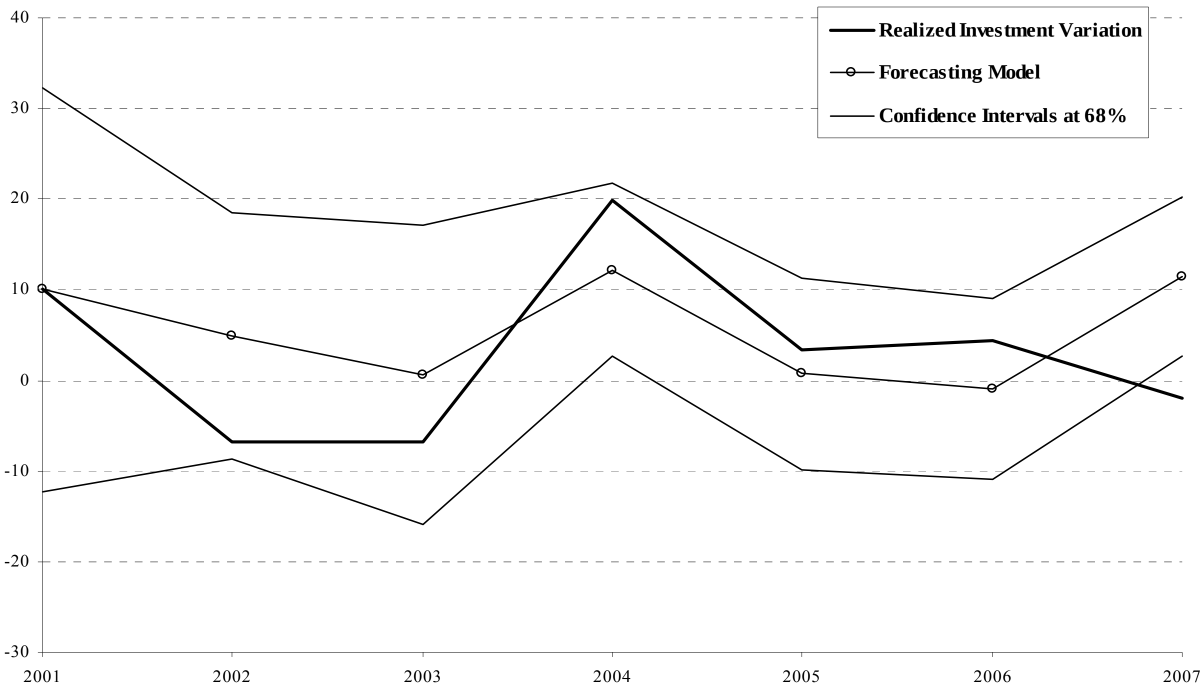
***: Significant at 1 per cent; **: Significant at 5 per cent; *: Significant at 10 per cent.
Standard errors in parentheses.

Among the time-invariant fixed effects, the number of employees dominates the sector specific and geographical effects, even though its parameters are rarely statistically significant. However, the utilization of time-invariant fixed effects, albeit non-significant, is helpful in providing a model (*M1*) with superior forecasting capabilities, as shown in Table 7: this is consistent with the fact that models with more variables than those strictly needed for simple data explanation are often preferable for prediction (Burnham and Anderson, 2002).

Finally, the hypothesis of no serial correlation of residuals tends to be confirmed by the Durbin-Watson statistic adapted for panel data (see last row of Table 8). The hypothesis is rejected only for two panels (1995-2001 and 1999-2005), for which the Durbin-Watson statistic is well below the threshold level of 1.86 (Bhargava et al., 1982).

Finally, we use the model parameter estimates to predict the one-year-ahead investment variations at firm level. A consistent predictor for the aggregate investment variation is obtained as the weighted average of individual predictions over all units (equation 13) and can be compared with the aggregate realised investment variation (equation 14) computed on the same panel sets. Figure 6 plots the two series, together with the forecast confidence interval at 68 per cent. The model generates a one-year-ahead forecast that follows the aggregate dynamics without bias, even if for small-size variations the sign of the prediction is not always the right one.

Figure 6. Forecast and realised investment variation (variation index on 8-year panel data)



Source: Invind and Sondtel surveys.

6 Conclusions

We enrich the available instruments of short-term economic analysis by examining a sample of Italian manufacturing firms for which qualitative investment plans and investment levels are collected in two separate yearly surveys. The peculiar feature of this panel sample is that categorical and quantitative investment plans and quantitative realised investments are collected for the same firms on different occasions. By relying on exactly matched data on plans and realisations, our model significantly enriches the information set obtained from simple categorical variables on investment plans.

We provide a tool that makes full use of the heterogeneity of disaggregated individual responses, together with the microdata panel dimension. These characteristics have only recently begun to be explored in the econometric literature.

We plan an empirical extension of the model to all the target population covered by the survey, including the industrial sectors outside manufacturing, as well as private non-financial services. This generalization will become feasible in a couple of years, once an adequate number of repeated cross-section surveys are available.

Appendix - The Carlson-Parkins method

The method postulates that whenever respondents answer a simple categorical question about a forecast with three response items (1= goes down, 2= stays stable, 3= goes up), all have the same indifference interval (a,b), with a < 0, and accordingly answer 1 if their quantitative forecast y_e is below a, 3 if it is above b and 2 otherwise. y_e is assumed to be distributed according to a cumulative standardized distribution G^*.

If we indicate with μ_e and σ^2_e the mean and variance of y_e and with D and U the fractions of respondents respectively declaring a negative and a positive variation, we can write:

$$d = G^{*-1}(D) = \frac{a - \mu_e}{\sigma_\varepsilon} \tag{a1}$$

$$u = G^{*-1}(1 - D) = \frac{b - \mu_e}{\sigma_\varepsilon} \tag{a2}$$

For a given form of G^*, the system formed by (a1) and (a2) is solvable only if the indifference interval is symmetric around zero, with $-a=c=b$

Elementary algebra produces the following quantitative estimation of the aggregate forecast y_e expressed in terms of percentage variation:

$$y_e^{cp} = 100c\frac{d+u}{d-u} \tag{a3}$$

The original proposal by Carlson and Parkin does not make use of a specific value for c and relies on the additional hypothesis that expectations and realisations are the same over T past periods (for which all data are available) in order to get an estimate for c. We do not need this limitation, since for the *Sondtel* survey we have $-a=c=b=3$, after collapsing the categories ("strong decrease" [less than -10%], "slight decrease"[(-10% to -3%]) into “decrease” (less than -3%) and ("slight increase" [+3% to +10%], "strong increase" [more than +10%]) into “increase” (more than 3%).

References

Arellano, M., Bond, S., (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment. *The Review of Economic Studies*, 58, 277- 297.

Baltagi B. H., Song S. H. (2006). Unbalanced panel data: a survey, *Statistical Papers*, 47, 493-523.

Bank of Italy (2008). Survey of Industrial and Service Firms (Year 2007). *Supplements to the Statistical Bulletin, Sample Surveys*, XVIII, 42.

Batchelor R (2007). Bias in macroeconomic forecasts. *International Journal of Forecasting,* 23, 2, 189-203.

Bernanke B. (1983). Irreversibility, Uncertainty and Cyclical Investments. *The Quarterly Journal of Economics*, 98, 85-106.

Bhargava A., Franzini L., Narendranathan W. (1982). Serial Correlation and the Fixed Effects Model. *Review of Economic Studies*, XLIX, 533-549.

Bruno G. S. F. (2005). Approximating the bias of the LSDV estimator for dynamic unbalanced panel data models. *Economics Letters*, 87, 361-366.

Burnham K.P., Anderson D.R. (2002), *Model Selection and Multimodel Inference (2nd Edition),* Springer-Verlag, New York.

Cannell C., Miller, P., Oksenberg L. (1981). Research on Interviewing Techniques. In *Sociological Methodology* (Leinhardt, S. ed.), 389-437, San Francisco: Jossey-Bass.

Carlson J. A., Parkins M. (1975), Inflation expectations. *Economica*, 42, 123-138.

Chirinko R. (1993). Business Fixed Investment Spending: Modeling Strategies, Empirical Results, and Policy Implications. *Journal of Economic Literature*, XXXI, 1875-1911.

Collado M.D. (1997). Estimation dynamic models from time series of independent cross-sections*, Journal of Econometrics* 82, 37-62.

DuMouchel W. H., Duncan, G. (1983). Using Sample Survey Weights in Multiple Regression Analyses of Stratified Samples. *Journal of the American Statistical Association*, 78, 535-543.

Flyvbjerg B. (2003). Delusions of Success: Comment on Dan Lovallo and Daniel Kahneman, *Harvard Business Review*, December Issue, 121-122.

Holtz-Eakin D. (1988). Testing for individual effects in autoregressive models. *Journal of Econometrics*, 39, 297-307.

Jimenez-Martin S. (1998). On the testing of heterogeneity effects in dynamic unbalanced panel data models. *Economics Letters*, 58, 157-163.

Goldrian G. (ed.) (2007). Handbook of survey-based business cycle analysis, Edward Elgar Publishing Limited.

Kahneman D., Lovallo, D. (1993) Timid Choices and Bald Forecasts: A Cognitive Perspective on Risk Taking. *Management Science*, 39, n. 1, 17-31.

Kahneman D., Lovallo D. (2003). Delusions of Success: How Optimism Undermines Executives Decisions, *Harvard Business Review,* July Issue, 56-63.

Kocic P. N., Bell P. A. (1994). Optimal Winsorized Cutoffs for a Stratified Finite Population. *Journal of Official Statistics*, 10, 419-435.

Lokshin B. (2008). Further results on bias in dynamic unbalanced panel data models with an application to firm R&D investment. *UNU-MERIT Working Papers*, ISSN 1871-9872.

Mitchell J., Smith R. J., Weale M. R (2002), Quantification of qualitative firm-level survey data, *Economic Journal*, 112, C117-C135.

Mitchell J., Smith R. J., Weale M. R. (2004), Aggregate versus disaggregate survey-based indicators of economic activity. *Revision of National Institute of Economic and Social Research Discussion Paper* No. 194.

Moffitt R. (1993). Identification and estimation of dynamic models with a time series of repeated cross-sections, *Journal of Econometrics* 59, 99-123.

Nardo M, (2003). The Quantification of Qualitative Survey Data: A Critical Assessment. *Journal of Economic Surveys*, 17, 645-668.

Pesaran M. H. (1984), Expectations formation and macroeconomic modelling. In *Contemporary Macroeconomic Modelling* (eds. Malgrange, P. and P. Muet), Blackwell.

Pesaran M. H. (1987), The limits to rational expectations, Basil Blackwell.

Rötheli T. F. (1998). Forecasting among alternative strategies in the management of uncertainty, *Managerial and Decision Economics*, 19, n. 3, 179-187.

Smith P., Pont, M., Jones, T. (2003). Developments in business survey methodology in the Office for National Statistics, 1994-2000. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 52, 257-286.

Verbeek M., Vella F. (2005). Estimating dynamic models from repeated cross-sections, *Journal of Econometrics*, 127, 1, 83-102.

N. 811 – *Schooling and youth mortality: learning from a mass military exemption*, by Piero Cipollone and Alfonso Rosolia (June 2011).

N. 812 – *Welfare costs of inflation and the circulation of US currency abroad*, by Alessandro Calza and Andrea Zaghini (June 2011).

N. 813 – *Legal status of immigrants and criminal behavior: evidence from a natural experiment*, by Giovanni Mastrobuoni and Paolo Pinotti (June 2011).

N. 814 – *An unexpected crisis? Looking at pricing effectiveness of different banks*, by Valerio Vacca (July 2011).

N. 815 – *Skills or culture? An analysis of the decision to work by immigrant women in Italy*, by Antonio Accetturo and Luigi Infante (July 2011).

N. 816 – *Home bias in interbank lending and banks' resolution regimes*, by Michele Manna (July 2011).

N. 817 – *Macroeconomic determinants of carry trade activity*, by Alessio Anzuini and Fabio Fornari (September 2011).

N. 818 – *Leaving home and housing prices. The experience of Italian youth emancipation*, by Francesca Modena and Concetta Rondinelli (September 2011).

N. 819 – *The interbank market after the financial turmoil: squeezing liquidity in a "lemons market" or asking liquidity "on tap"*, by Antonio De Socio (September 2011).

N. 820 – *The relationship between the PMI and the Italian index of industrial production and the impact of the latest economic crisis*, by Valentina Aprigliano (September 2011).

N. 821 – *Inside the sovereign credit default swap market: price discovery, announcements, market behaviour and corporate sector*, by Alessandro Carboni (September 2011).

N. 822 – *The demand for energy of Italian households*, by Ivan Faiella (September 2011).

N. 823 – *Sull'ampiezza ottimale delle giurisdizioni locali: il caso delle province italiane*, by Guglielmo Barone (September 2011).

N. 824 – *The public-private pay gap: a robust quantile approach*, by Domenico Depalo and Raffaela Giordano (September 2011).

N. 825 – *Evaluating students' evaluations of professors*, by Michele Braga, Marco Paccagnella and Michele Pellizzari (October 2011).

N. 826 – *Do interbank customer relationships exist? And how did they function in the crisis? Learning from Italy*, by Massimiliano Affinito (October 2011).

N. 827 – *Foreign trade, home linkages and the spatial transmission of economic fluctuations in Italy*, by Valter Di Giacinto (October 2011).

N. 828 – *Healthcare in Italy: expenditure determinants and regional differentials*, by Maura Francese and Marzia Romanelli (October 2011).

N. 829 – *Bank heterogeneity and interest rate setting: what lessons have we learned since Lehman Brothers?*, by Leonardo Gambacorta and Paolo Emilio Mistrulli (October 2011).

N. 830 – *Structural reforms and macroeconomic performance in the euro area countries: a model-based assessment*, by Sandra Gomes, Pascal Jacquinot, Matthias Mohr and Massimiliano Pisani (October 2011).

N. 831 – *Risk measures for autocorrelated hedge fund returns*, by Antonio Di Cesare, Philip A. Stork and Casper G. de Vries (October 2011).

(*) Requests for copies should be sent to:
Banca d'Italia – Servizio Studi di struttura economica e finanziaria – Divisione Biblioteca e Archivio storico – Via Nazionale, 91 – 00184 Rome – (fax 0039 06 47922059). They are available on the Internet www.bancaditalia.it.

2008

P. ANGELINI, *Liquidity and announcement effects in the euro area*, Giornale degli Economisti e Annali di Economia, v. 67, 1, pp. 1-20, **TD No. 451 (October 2002).**

P. ANGELINI, P. DEL GIOVANE, S. SIVIERO and D. TERLIZZESE, *Monetary policy in a monetary union: What role for regional information?*, International Journal of Central Banking, v. 4, 3, pp. 1-28, **TD No. 457 (December 2002).**

F. SCHIVARDI and R. TORRINI, *Identifying the effects of firing restrictions through size-contingent Differences in regulation*, Labour Economics, v. 15, 3, pp. 482-511, **TD No. 504 (June 2004).**

L. GUISO and M. PAIELLA,, *Risk aversion, wealth and background risk*, Journal of the European Economic Association, v. 6, 6, pp. 1109-1150, **TD No. 483 (September 2003).**

C. BIANCOTTI, G. D'ALESSIO and A. NERI, *Measurement errors in the Bank of Italy's survey of household income and wealth*, Review of Income and Wealth, v. 54, 3, pp. 466-493, **TD No. 520 (October 2004).**

S. MOMIGLIANO, J. HENRY and P. HERNÁNDEZ DE COS, *The impact of government budget on prices: Evidence from macroeconometric models*, Journal of Policy Modelling, v. 30, 1, pp. 123-143 **TD No. 523 (October 2004).**

L. GAMBACORTA, *How do banks set interest rates?,* European Economic Review, v. 52, 5, pp. 792-819, **TD No. 542 (February 2005).**

P. ANGELINI and A. GENERALE, *On the evolution of firm size distributions*, American Economic Review, v. 98, 1, pp. 426-438, **TD No. 549 (June 2005).**

R. FELICI and M. PAGNINI, *Distance, bank heterogeneity and entry in local banking markets,* The Journal of Industrial Economics, v. 56, 3, pp. 500-534, **No. 557 (June 2005).**

S. DI ADDARIO and E. PATACCHINI, *Wages and the city. Evidence from Italy*, Labour Economics, v.15, 5, pp. 1040-1061, **TD No. 570 (January 2006).**

S. SCALIA, *Is foreign exchange intervention effective?,* Journal of International Money and Finance, v. 27, 4, pp. 529-546, **TD No. 579 (February 2006).**

M. PERICOLI and M. TABOGA, *Canonical term-structure models with observable factors and the dynamics of bond risk premia,* Journal of Money, Credit and Banking, v. 40, 7, pp. 1471-88, **TD No. 580 (February 2006).**

E. VIVIANO, *Entry regulations and labour market outcomes. Evidence from the Italian retail trade sector*, Labour Economics, v. 15, 6, pp. 1200-1222, **TD No. 594 (May 2006).**

S. FEDERICO and G. A. MINERVA, *Outward FDI and local employment growth in Italy,* Review of World Economics, Journal of Money, Credit and Banking, v. 144, 2, pp. 295-324, **TD No. 613 (February 2007).**

F. BUSETTI and A. HARVEY, *Testing for trend*, Econometric Theory, v. 24, 1, pp. 72-87, **TD No. 614 (February 2007).**

V. CESTARI, P. DEL GIOVANE and C. ROSSI-ARNAUD, *Memory for prices and the Euro cash changeover: an analysis for cinema prices in Italy*, In P. Del Giovane e R. Sabbatini (eds.), The Euro Inflation and Consumers' Perceptions. Lessons from Italy, Berlin-Heidelberg, Springer, **TD No. 619 (February 2007).**

B. H. HALL, F. LOTTI and J. MAIRESSE, *Employment, innovation and productivity: evidence from Italian manufacturing microdata*, Industrial and Corporate Change, v. 17, 4, pp. 813-839, **TD No. 622 (April 2007).**

J. SOUSA and A. ZAGHINI, *Monetary policy shocks in the Euro Area and global liquidity spillovers,* International Journal of Finance and Economics, v.13, 3, pp. 205-218, **TD No. 629 (June 2007).**

M. DEL GATTO, GIANMARCO I. P. OTTAVIANO and M. PAGNINI, *Openness to trade and industry cost dispersion: Evidence from a panel of Italian firms,* Journal of Regional Science, v. 48, 1, pp. 97-129, **TD No. 635 (June 2007).**

P. DEL GIOVANE, S. FABIANI and R. SABBATINI, *What's behind "inflation perceptions"? A survey-based analysis of Italian consumers,* in P. Del Giovane e R. Sabbatini (eds.), *The Euro Inflation and Consumers' Perceptions. Lessons from Italy,* Berlin-Heidelberg, Springer, **TD No. 655 (January 2008).**

R. BRONZINI, G. DE BLASIO, G. PELLEGRINI and A. SCOGNAMIGLIO, *La valutazione del credito d'imposta per gli investimenti*, Rivista di politica economica, v. 98, 4, pp. 79-112, **TD No. 661 (April 2008).**

B. BORTOLOTTI, and P. PINOTTI, *Delayed privatization*, Public Choice, v. 136, 3-4, pp. 331-351, **TD No. 663 (April 2008).**

R. BONCI and F. COLUMBA, *Monetary policy effects: New evidence from the Italian flow of funds*, Applied Economics , v. 40, 21, pp. 2803-2818, **TD No. 678 (June 2008).**

M. CUCCULELLI, and G. MICUCCI, *Family Succession and firm performance: evidence from Italian family firms*, Journal of Corporate Finance, v. 14, 1, pp. 17-31, **TD No. 680 (June 2008).**

A. SILVESTRINI and D. VEREDAS, *Temporal aggregation of univariate and multivariate time series models: a survey*, Journal of Economic Surveys, v. 22, 3, pp. 458-497, **TD No. 685 (August 2008).**

2009

F. PANETTA, F. SCHIVARDI and M. SHUM, *Do mergers improve information? Evidence from the loan market*, Journal of Money, Credit, and Banking, v. 41, 4, pp. 673-709, **TD No. 521 (October 2004).**

M. BUGAMELLI and F. PATERNÒ, *Do workers' remittances reduce the probability of current account reversals?*, World Development, v. 37, 12, pp. 1821-1838, **TD No. 573 (January 2006).**

P. PAGANO and M. PISANI, *Risk-adjusted forecasts of oil prices*, The B.E. Journal of Macroeconomics, v. 9, 1, Article 24, **TD No. 585 (March 2006).**

M. PERICOLI and M. SBRACIA, *The CAPM and the risk appetite index: theoretical differences, empirical similarities, and implementation problems*, International Finance, v. 12, 2, pp. 123-150, **TD No. 586 (March 2006).**

U. ALBERTAZZI and L. GAMBACORTA, *Bank profitability and the business cycle*, Journal of Financial Stability, v. 5, 4, pp. 393-409, **TD No. 601 (September 2006).**

S. MAGRI, *The financing of small innovative firms: the Italian case*, Economics of Innovation and New Technology, v. 18, 2, pp. 181-204, **TD No. 640 (September 2007).**

V. DI GIACINTO and G. MICUCCI, *The producer service sector in Italy: long-term growth and its local determinants*, Spatial Economic Analysis, Vol. 4, No. 4, pp. 391-425, **TD No. 643 (September 2007).**

F. LORENZO, L. MONTEFORTE and L. SESSA, *The general equilibrium effects of fiscal policy: estimates for the euro area*, Journal of Public Economics, v. 93, 3-4, pp. 559-585, **TD No. 652 (November 2007).**

R. GOLINELLI and S. MOMIGLIANO, *The Cyclical Reaction of Fiscal Policies in the Euro Area. A Critical Survey of Empirical Research*, Fiscal Studies, v. 30, 1, pp. 39-72, **TD No. 654 (January 2008).**

P. DEL GIOVANE, S. FABIANI and R. SABBATINI, *What's behind "Inflation Perceptions"? A survey-based analysis of Italian consumers*, Giornale degli Economisti e Annali di Economia, v. 68, 1, pp. 25-52, **TD No. 655 (January 2008).**

F. MACCHERONI, M. MARINACCI, A. RUSTICHINI and M. TABOGA, *Portfolio selection with monotone mean-variance preferences*, Mathematical Finance, v. 19, 3, pp. 487-521, **TD No. 664 (April 2008).**

M. AFFINITO and M. PIAZZA, *What are borders made of? An analysis of barriers to European banking integration*, in P. Alessandrini, M. Fratianni and A. Zazzaro (eds.): The Changing Geography of Banking and Finance, Dordrecht Heidelberg London New York, Springer, **TD No. 666 (April 2008).**

A. BRANDOLINI, *On applying synthetic indices of multidimensional well-being: health and income inequalities in France, Germany, Italy, and the United Kingdom*, in R. Gotoh and P. Dumouchel (eds.), Against Injustice. The New Economics of Amartya Sen, Cambridge, Cambridge University Press, **TD No. 668 (April 2008).**

G. FERRERO and A. NOBILI, *Futures contract rates as monetary policy forecasts*, International Journal of Central Banking, v. 5, 2, pp. 109-145, **TD No. 681 (June 2008).**

P. CASADIO, M. LO CONTE and A. NERI, *Balancing work and family in Italy: the new mothers' employment decisions around childbearing*, in T. Addabbo and G. Solinas (eds.), Non-Standard Employment and Qualità of Work, Physica-Verlag. A Sprinter Company, **TD No. 684 (August 2008).**

L. ARCIERO, C. BIANCOTTI, L. D'AURIZIO and C. IMPENNA, *Exploring agent-based methods for the analysis of payment systems: A crisis model for StarLogo TNG*, Journal of Artificial Societies and Social Simulation, v. 12, 1, **TD No. 686 (August 2008).**

A. CALZA and A. ZAGHINI, *Nonlinearities in the dynamics of the euro area demand for M1*, Macroeconomic Dynamics, v. 13, 1, pp. 1-19, **TD No. 690 (September 2008).**

L. FRANCESCO and A. SECCHI, *Technological change and the households' demand for currency*, Journal of Monetary Economics, v. 56, 2, pp. 222-230, **TD No. 697 (December 2008).**

G. ASCARI and T. ROPELE, *Trend inflation, taylor principle, and indeterminacy*, Journal of Money, Credit and Banking, v. 41, 8, pp. 1557-1584, **TD No. 708 (May 2007).**

S. COLAROSSI and A. ZAGHINI, *Gradualism, transparency and the improved operational framework: a look at overnight volatility transmission*, International Finance, v. 12, 2, pp. 151-170, **TD No. 710 (May 2009).**

M. BUGAMELLI, F. SCHIVARDI and R. ZIZZA, *The euro and firm restructuring*, in A. Alesina e F. Giavazzi (eds): Europe and the Euro, Chicago, University of Chicago Press, **TD No. 716 (June 2009).**

B. HALL, F. LOTTI and J. MAIRESSE, *Innovation and productivity in SMEs: empirical evidence for Italy*, Small Business Economics, v. 33, 1, pp. 13-33, **TD No. 718 (June 2009).**

2010

A. PRATI and M. SBRACIA, *Uncertainty and currency crises: evidence from survey data*, Journal of Monetary Economics, v, 57, 6, pp. 668-681, **TD No. 446 (July 2002).**

L. MONTEFORTE and S. SIVIERO, *The Economic Consequences of Euro Area Modelling Shortcuts*, Applied Economics, v. 42, 19-21, pp. 2399-2415, **TD No. 458 (December 2002).**

S. MAGRI, *Debt maturity choice of nonpublic Italian firms* , Journal of Money, Credit, and Banking, v.42, 2-3, pp. 443-463, **TD No. 574 (January 2006).**

R. BRONZINI and P. PISELLI, *Determinants of long-run regional productivity with geographical spillovers: the role of R&D, human capital and public infrastructure*, Regional Science and Urban Economics, v. 39, 2, pp.187-199, **TD No. 597 (September 2006).**

E. IOSSA and G. PALUMBO, *Over-optimism and lender liability in the consumer credit market*, Oxford Economic Papers, v. 62, 2, pp. 374-394, **TD No. 598 (September 2006).**

S. NERI and A. NOBILI, *The transmission of US monetary policy to the euro area*, International Finance, v. 13, 1, pp. 55-78, **TD No. 606 (December 2006).**

F. ALTISSIMO, R. CRISTADORO, M. FORNI, M. LIPPI and G. VERONESE, *New Eurocoin: Tracking Economic Growth in Real Time*, Review of Economics and Statistics, v. 92, 4, pp. 1024-1034, **TD No. 631 (June 2007).**

A. CIARLONE, P. PISELLI and G. TREBESCHI, *Emerging Markets' Spreads and Global Financial Conditions*, Journal of International Financial Markets, Institutions & Money, v. 19, 2, pp. 222-239, **TD No. 637 (June 2007).**

U. ALBERTAZZI and L. GAMBACORTA, *Bank profitability and taxation*, Journal of Banking and Finance, v. 34, 11, pp. 2801-2810, **TD No. 649 (November 2007).**

M. IACOVIELLO and S. NERI, *Housing market spillovers: evidence from an estimated DSGE model*, American Economic Journal: Macroeconomics, v. 2, 2, pp. 125-164, **TD No. 659 (January 2008).**

F. BALASSONE, F. MAURA and S. ZOTTERI, *Cyclical asymmetry in fiscal variables in the EU*, Empirica, **TD No. 671**, v. 37, 4, pp. 381-402 **(June 2008).**

F. D'AMURI, O. GIANMARCO I.P. and P. GIOVANNI, *The labor market impact of immigration on the western german labor market in the 1990s*, European Economic Review, v. 54, 4, pp. 550-570, **TD No. 687 (August 2008).**

A. ACCETTURO, *Agglomeration and growth: the effects of commuting costs*, Papers in Regional Science, v. 89, 1, pp. 173-190, **TD No. 688 (September 2008).**

S. NOBILI and G. PALAZZO, *Explaining and forecasting bond risk premiums*, Financial Analysts Journal, v. 66, 4, pp. 67-82, **TD No. 689 (September 2008).**

A. B. ATKINSON and A. BRANDOLINI, *On analysing the world distribution of income*, World Bank Economic Review , v. 24, 1 , pp. 1-37, **TD No. 701 (January 2009).**

R. CAPPARIELLO and R. ZIZZA, *Dropping the Books and Working Off the Books*, Labour, v. 24, 2, pp. 139-162 ,**TD No. 702 (January 2009).**

C. NICOLETTI and C. RONDINELLI, *The (mis)specification of discrete duration models with unobserved heterogeneity: a Monte Carlo study*, Journal of Econometrics, v. 159, 1, pp. 1-13, **TD No. 705 (March 2009).**

L. FORNI, A. GERALI and M. PISANI, *Macroeconomic effects of greater competition in the service sector: the case of Italy*, Macroeconomic Dynamics, v. 14, 5, pp. 677-708, **TD No. 706 (March 2009).**

V. DI GIACINTO, G. MICUCCI and P. MONTANARO, *Dynamic macroeconomic effects of public capital: evidence from regional Italian data*, Giornale degli economisti e annali di economia, v. 69, 1, pp. 29-66, **TD No. 733 (November 2009).**

F. COLUMBA, L. GAMBACORTA and P. E. MISTRULLI, *Mutual Guarantee institutions and small business finance*, Journal of Financial Stability, v. 6, 1, pp. 45-54, **TD No. 735 (November 2009).**

A. GERALI, S. NERI, L. SESSA and F. M. SIGNORETTI, *Credit and banking in a DSGE model of the Euro Area,* Journal of Money, Credit and Banking, v. 42, 6, pp. 107-141, **TD No. 740 (January 2010).**

M. AFFINITO and E. TAGLIAFERRI, *Why do (or did?) banks securitize their loans? Evidence from Italy*, Journal of Financial Stability, v. 6, 4, pp. 189-202, **TD No. 741 (January 2010).**

S. FEDERICO, *Outsourcing versus integration at home or abroad and firm heterogeneity,* Empirica, v. 37, 1, pp. 47-63, **TD No. 742 (February 2010).**

V. DI GIACINTO, *On vector autoregressive modeling in space and time,* Journal of Geographical Systems, v. 12, 2, pp. 125-154, **TD No. 746 (February 2010).**

S. MOCETTI and C. PORELLO, *How does immigration affect native internal mobility? new evidence from Italy,* Regional Science and Urban Economics, v. 40, 6, pp. 427-439, **TD No. 748 (March 2010).**

A. DI CESARE and G. GUAZZAROTTI, *An analysis of the determinants of credit default swap spread changes before and during the subprime financial turmoil,* Journal of Current Issues in Finance, Business and Economics, v. 3, 4, pp., **TD No. 749 (March 2010).**

P. CIPOLLONE, P. MONTANARO and P. SESTITO, *Value-added measures in Italian high schools: problems and findings,* Giornale degli economisti e annali di economia, v. 69, 2, pp. 81-114, **TD No. 754 (March 2010).**

A. BRANDOLINI, S. MAGRI and T. M SMEEDING, *Asset-based measurement of poverty,* Journal of Policy Analysis and Management, v. 29, 2 , pp. 267-284, **TD No. 755 (March 2010).**

G. CAPPELLETTI, *A Note on rationalizability and restrictions on beliefs,* The B.E. Journal of Theoretical Economics, v. 10, 1, pp. 1-11,**TD No. 757 (April 2010).**

S. DI ADDARIO and D. VURI, *Entrepreneurship and market size. the case of young college graduates in Italy,* Labour Economics, v. 17, 5, pp. 848-858, **TD No. 775 (September 2010).**

A. CALZA and A. ZAGHINI, *Sectoral money demand and the great disinflation in the US,* Journal of Money, Credit, and Banking, v. 42, 8, pp. 1663-1678, **TD No. 785 (January 2011).**

2011

S. DI ADDARIO, *Job search in thick markets*, Journal of Urban Economics, v. 69, 3, pp. 303-318, **TD No. 605 (December 2006).**

E. CIAPANNA, *Directed matching with endogenous markov probability: clients or competitors?,* The RAND Journal of Economics, v. 42, 1, pp. 92-120, **TD No. 665 (April 2008).**

L. FORNI, A. GERALI and M. PISANI, *The Macroeconomics of Fiscal Consolidation in a Monetary Union: the Case of Italy,* in Luigi Paganetto (ed.), Recovery after the crisis. Perspectives and policies, VDM Verlag Dr. Muller, **TD No. 747 (March 2010).**

A. DI CESARE and G. GUAZZAROTTI, *An analysis of the determinants of credit default swap changes before and during the subprime financial turmoil*, in Barbara L. Campos and Janet P. Wilkins (eds.), The Financial Crisis: Issues in Business, Finance and Global Economics, New York, Nova Science Publishers, Inc., **TD No. 749 (March 2010).**

G. GRANDE and I. VISCO, *A public guarantee of a minimum return to defined contribution pension scheme members,* The Journal of Risk, v. 13, 3, pp. 3-43, **TD No. 762 (June 2010).**

P. DEL GIOVANE, G. ERAMO and A. NOBILI, *Disentangling demand and supply in credit developments: a survey-based analysis for Italy,* Journal of Banking and Finance, v. 35, 10, pp. 2719-2732, **TD No. 764 (June 2010).**

M. TABOGA, *Under/over-valuation of the stock market and cyclically adjusted earnings*, International Finance, v. 14, 1, pp. 135-164, **TD No. 780 (December 2010).**

S. NERI, *Housing, consumption and monetary policy: how different are the U.S. and the Euro area?*, Journal of Banking and Finance, v.35, 11, pp. 3019-3041, **TD No. 807 (April 2011).**

FORTHCOMING

M. BUGAMELLI and A. ROSOLIA, *Produttività e concorrenza estera*, Rivista di politica economica, **TD No. 578 (February 2006).**

G. DE BLASIO and G. NUZZO, *Historical traditions of civicness and local economic development,* Journal of Regional Science, **TD No. 591 (May 2006).**

F. CINGANO and A. ROSOLIA, *People I know: job search and social networks*, Journal of Labor Economics, **TD No. 600 (September 2006).**

F. SCHIVARDI and E. VIVIANO, *Entry barriers in retail trade*, Economic Journal, **TD No. 616 (February 2007).**

G. FERRERO, A. NOBILI and P. PASSIGLIA, *Assessing excess liquidity in the Euro Area: the role of sectoral distribution of money,* Applied Economics, **TD No. 627 (April 2007).**

P. E. MISTRULLI, *Assessing financial contagion in the interbank market: maximun entropy versus observed interbank lending patterns*, Journal of Banking & Finance, **TD No. 641 (September 2007).**

Y. ALTUNBAS, L. GAMBACORTA and D. MARQUÉS, *Securitisation and the bank lending channel*, European Economic Review, **TD No. 653 (November 2007).**

M. BUGAMELLI and F. PATERNÒ, *Output growth volatility and remittances*, Economica, **TD No. 673 (June 2008).**

V. DI GIACINTO e M. PAGNINI, *Local and global agglomeration patterns: two econometrics-based indicators*, Regional Science and Urban Economics, **TD No. 674 (June 2008).**

G. BARONE and F. CINGANO, *Service regulation and growth: evidence from OECD countries*, Economic Journal, **TD No. 675 (June 2008).**

S. MOCETTI, *Educational choices and the selection process before and after compulsory school*, Education Economics, **TD No. 691 (September 2008).**

P. SESTITO and E. VIVIANO, *Reservation wages: explaining some puzzling regional patterns*, Labour, **TD No. 696 (December 2008).**

P. PINOTTI, M. BIANCHI and P. BUONANNO, *Do immigrants cause crime?*, Journal of the European Economic Association, **TD No. 698 (December 2008).**

R. GIORDANO and P. TOMMASINO, *What determines debt intolerance? The role of political and monetary institutions*, European Journal of Political Economy, **TD No. 700 (January 2009).**

F. LIPPI and A. NOBILI, *Oil and the macroeconomy: a quantitative structural analysis,* Journal of European Economic Association, **TD No. 704 (March 2009).**

F. CINGANO and P. PINOTTI, *Politicians at work. The private returns and social costs of political connections*, Journal of the European Economic Association, **TD No. 709 (May 2009).**

Y. ALTUNBAS, L. GAMBACORTA, and D. MARQUÉS-IBÁÑEZ, *Bank risk and monetary policy*, Journal of Financial Stability, **TD No. 712 (May 2009).**

P. ANGELINI, A. NOBILI e C. PICILLO, *The interbank market after August 2007: What has changed, and why?*, Journal of Money, Credit and Banking, **TD No. 731 (October 2009).**

G. BARONE and S. MOCETTI, *Tax morale and public spending inefficiency*, International Tax and Public Finance, **TD No. 732 (November 2009).**

L. FORNI, A. GERALI and M. PISANI, *The macroeconomics of fiscal consolidations in euro area countries*, Journal of Economic Dynamics and Control, **TD No. 747 (March 2010).**

G. BARONE, R. FELICI and M. PAGNINI, *Switching costs in local credit markets*, International Journal of Industrial Organization, **TD No. 760 (June 2010).**

G. BARONE and S. MOCETTI, *With a little help from abroad: the effect of low-skilled immigration on the female labour supply*, Labour Economics, **TD No. 766 (July 2010).**

S. MAGRI and R. PICO, *The rise of risk-based pricing of mortgage interest rates in Italy*, Journal of Banking and Finance, **TD No. 778 (October 2010).**

A. ACCETTURO and G. DE BLASIO, *Policies for local development: an evaluation of Italy's "Patti Territoriali"*, Regional Science and Urban Economics, **TD No. 789 (January 2006).**

E. COCOZZA and P. PISELLI, *Testing for east-west contagion in the European banking sector during the financial crisis*, in R. Matoušek; D. Stavárek (eds.), Financial Integration in the European Union, Taylor & Francis, **TD No. 790 (February 2011).**

S. NERI and T. ROPELE, *Imperfect information, real-time data and monetary policy in the Euro area*, The Economic Journal, **TD No. 802 (March 2011).**