Does soft information matter for financial analysts' forecasts? A gravity model approach

Régis Breton
Banque de France, CNRS-Laboratoire d'Economie d'Orléans (LEO)
Banque de France
35-2341 DGO-DSF-SMF
75049 Paris Cedex 01
FRANCE
Regis.Breton@banque-france.fr

Sébastien Galanti, corresponding author
Université d'Orléans, Laboratoire d'Economie d'Orléans (LEO)
Faculté de droit, économie et gestion
Rue de Blois
BP 26739
45067 Orlans cedex 02
FRANCE
Sebastien.Galanti@univ-orleans.fr
Tel +33 2 38 41 73 62, Fax +33 2 38 41 73 80

Christophe Hurlin
Université d'Orléans, Laboratoire d'Economie d'Orléans (LEO)
Faculté de droit, économie et gestion
Rue de Blois
BP 26739
45067 Orl'éans cedex 02
FRANCE
Christophe.Hurlin@univ-orleans.fr

Anne-Gaël Vaubourg
Université Bordeaux IV, LAREFI
Avenue Lon Duguit
33 608 Pessac
FRANCE
Anne-Gael.Vaubourg@u-bordeaux4.fr

Abstract

We study whether financial analysts' concern for maintaining good relationships with firms' managers to preserve their access to 'soft' qualitative information motivates them to issue pessimistic or optimistic forecasts. We apply a gravity model approach to firm-analyst relationships. The decomposition of fixed effects allows us to use the pair-effect as a measure soft information. We find that a low (high) pair-effect is associated with a low (high) forecast error. This observation suggests that pessimism and optimism result from the analysts' concern for preserving access to the soft information released by managers.

Keywords: financial analysts, earnings forecasts, soft information, panel regression, gravity models.

1 Introduction

Financial analysts' role as information producers is crucial for financial markets. By issuing forecasts regarding the value of firms' shares or earnings per share (EPS), they reduce information asymmetries between firms and investors or fund managers. Generally issued on behalf of brokers, forecasts and selling or buying recommendations are widely used by fund managers for taking portfolio allocation decisions. However, many studies have demonstrated that analysts' earnings forecasts can be inaccurate (Brown, 1997), thus increasing corporate agency costs, and reducing the informational efficiency of financial markets.

One strand of literature reveals that analyst forecasts are excessively optimistic[1]. There are at least two reasons for this bias. On the one hand, some firms may refuse to contract with an investment bank if 'sell-side' analysts do not issue favorable forecasts on its EPS (Barber, Lehavy and Trueman 2007; Dugar and Nathan 1995; Hayward and Boeker 1998; Lin, McNichols and O'Brien, 2005). On the other hand, it is in the interest of the analyst to release optimistic recommendations or EPS forecasts to generate buying transactions (Jackson, 2005).

A second set of papers suggests that some analysts' EPS forecasts are pessimistic, i.e., below, instead of above, the actual EPS. To explain such behavior, the literature generally refers to the 'earnings management strategy'. This strategy consists of the discretionary manipulation of earnings by firm managers such that the actual EPS ultimately appears to be higher than the forecast (Payne and Robb, 2000; Matsumoto, 2002; Burgstahler and Eames, 2006). This phenomenon can also involve analysts favoring a company by adjusting their estimates to help managers match or exceed

[1] Throughout this article, we will denote a forecast that exceeds the realized value (forecast-earnings>0) as optimistic and, conversely, denote a forecast that is below the realized value (forecast-earnings<0) as pessimitic.

expectations (Chan, Karceski and Lakonishok 2007). The goal of this strategy is to create a 'positive earnings surprise' on financial markets when the firm's manager reports the actual EPS. If the realized EPS appears higher than its forecast, investors in financial markets may have a favorable reaction, thus increasing the price of firm's shares.

However, analysts do not only produce forecasts. They also provide their brokerage clients, primarily fund managers, with different types of qualitative information about a given firm. To do so, they organize phone calls, one-on-one meetings and they invite fund managers to participate in conference calls with the firm's management (Fogarty and Rogers, 2005; Breton and Taffler, 2001). This access to so-called 'soft information' is only possible if the analyst has a good relationship with the firm's management. In this context, issuing biased forecasts may be a means for analysts to maintain their access to 'soft information' about firms, with the aim of privately transmitting this soft information to their clients (Libby, Hunton, Tan and Seybert, 2008; Pratt, 1993; Gibson, 1995; Womack, 1996; Boni and Womack, 2002). On the one hand, the analyst may be enticed to issue pessimistic forecasts in order to help a firm's manager 'beat the forecast', thus providing a positive earnings surprise on the financial markets when actual EPS is reported. In this case, the need to maintain relationships with firms' top executives results in the analysts committing negative forecast errors (Easterwood and Nutt, 1999; Lim, 2001). On the other hand, issuing optimistic forecasts make it possible to create a favorable reaction on financial markets, thus increasing a firm's share price (Payne and Robb, 2000; Matsumoto, 2002; Burgstahler and Eames, 2006). In this case, the analysts' desire to preserve their access to the soft information released by firm managers leads to positive forecast errors.

This observation suggests that the impact of soft information on forecast errors is specific to the particular relationship that exists between a certain firm and a certain analyst. In other words, high (negative or positive) forecast errors should be associated with a high firm-analyst pair effect.

The goal of this paper is to examine the empirical relevance of this theoretical prediction by investigating whether soft information, or a close relationship between a firm and an analyst, implies optimistic or pessimistic forecasts[2]. Soft information, which is qualitative and is only delivered if an analyst has a friendly relationship with a firm's manager, is based on unobservable characteristics. It is thus a difficult task to provide a direct and explicit measure of the role played by soft information in the forecast process. To the best of our knowledge, this issue has yet to be adressed in the empirical literature on financial analysts' forecasts.

The primary contribution of this paper is to fill this gap in the literature. Using an IBES data set provided by ThomsonReuters, which contains forecasts issued by 4 648 analysts regarding the earnings of 241 French firms between 1997 and 2007, we employ a gravity model approach to investigate financial analysts' forecasts. Our approach is particularly innovative because gravity models are typically used in international economics to analyze relationships between two trading partners. By definition, the effect of soft information on forecast accuracy is not captured by observable determinants. Therefore, when regressing the analysts' forecast accuracy on observable variables, the impact of soft information should be captured by the disturbance term in the regression. More precisely, as soft information reflects the strength of the relationship between *a given analyst* and *a given firm*, the concept of the paper is to compute a measure of soft information based on the firm-analyst specific effect. We then study

[2] In this article, we do not consider insider information (i.e., quantitative, precise, timely information that is likely to affect prices). Firm executives are generally very cautious not to disclose fraudulent insider information, which is prohibited by most coutries' financial regulation.

the relationship between this measure and analysts' forecast to determine whether soft information contributes to biased (pessimistic or optimistic) EPS forecasts.

The paper is organized as follows. Section 2 presents the theoretical and empirical background of our research and our testable assumption. Section 3 presents our empirical investigation while our results are reported in Section 4. Section 5 considers some robustness checks. Section 6 concludes.

2 Literature and Testable Assumption

In this section, we review the literature dedicated to forecast accuracy and its determinants. We first focus on the observable determinants of forecast accuracy (firms', analysts' and pairs' characteristics). Then, we present the role of soft information and state our testable assumption.

2.1 Firms' Characteristics as Observable Determinants of Forecast Accuracy

We first concentrate on firms' observable characteristics as determinants of forecast accuracy.

First, the empirical literature documents that forecast errors are negatively correlated with information availability and earnings predictability. Using a data set of firms from the Value Line Survey between 1989 and 1993, Das, Levine and Sivaramakrishnan (1998) report that the forecast error (the difference between forecasted and realized earnings) increases in firms' profit volatility. This result is confirmed by Lim (2001) using a set of forecasts provided by I/B/E/S (*Institutional Brokers Estimate System*) for the period 1984-1996. Further corroboration is provided by Jackson

(2005) who employs a data set of brokers on the Australian security market over the period 1992-2002.

Another result of Lim (2001)'s study is that the optimism bias is higher for firms with negative past earnings surprise and poor past stock returns. This supports the view that firms with bad past performance should be more reluctant to release public information.

Moreover, as greater public information is available for large firms and those followed by a large number of analysts, optimism is shown to decrease in firm size (Das and Levine, 1998; Lim, 2001; Jackson, 2005) and with analyst coverage (Lim, 2001).

Regarding interactions between the determinants of forecast accuracy, Das et al. (1998) establish that analyst coverage mitigates the increasing effect of earnings volatility on forecast optimism.

Another driving factor behind analyst forecast behavior is past optimistic consensus about a firm, which should deter analysts from contradicting the consensus forecast when it is inaccurate and thus increase their forecast error. This is confirmed empirically by Lim (2001).

Finally, Das et al. (1998) find that the need to preserve their relationships with managers entices analysts to issue particularly optimistic forecasts about firms that received unfavorable ratings from the well-known American financial publication Value Line. The same result is obtained by Francis and Philbrick (1993) using a set of Value Line recommendations for 1987, 1988 and 1989.

2.2 Analysts' Characteristics as Observable Determinants of Forecast Accuracy

The literature also investigates how analyst characteristics affect earnings forecasts. While Clement (1999) focuses on forecast accuracy (defined as the *absolute value* of the difference between forecasted and realized earnings), Lim (2001) concentrates on optimism. For Clement (1999) accuracy increases in analyst ability, the resources they devote to analyzing firms and with the incentive to preserve their reputations with their employer. For Lim (2001), analysts who have valuable information about firms are less concerned with maintaining their relationships with top executives. This effect explains why greater analyst ability and resources should reduce optimism.

Again, econometric studies provide some support for these theoretical assertions. Using a set of I/B/E/S forecasts over the period 1983-1994, Clement (1999) finds that, for the full sample period, general experience, measured by the number of years in which an analyst supplied at least one forecast, is associated with accuracy. When general experience increases, two effects are at play. First, the skill of the analyst increases, due to a learning-by-doing process. Second, the analyst is identified as highly capable because low-skilled analysts do not last in the profession. General experience also affects optimistic bias. Lim (2001) provides evidence that highly experienced analysts, who are less concerned with preserving their relationships with firm managers, are generally less optimistic than low-experience ones.

Forecast error also increases in the complexity of the analyst's portfolio (Clement, 1999). When an analyst follows a large number of firms, he devotes less resources to each one. When he follows a large number of industrial sectors, he benefits from sector specialization to a lesser extent.

Another important characteristic is the size of the broker employing the analyst.

This feature affects forecast accuracy and optimism bias. As a large broker can devote more resources to analyzing firms, analysts issue more accurate forecasts (Clement, 1999). Using an I/B/E/S dataset from 1989 to 1998, Clement and Tse (2005) confirm this result. For the same reason, maintaining relationships with firms' top executives becomes less crucial, and hence forecasts are less optimistic (Lim, 2001).

2.3 Pairs' Characteristics as Observable Determinants of Forecast Accuracy

We now turn to firm-analyst characteristics as determinants of forecast accuracy. For the same reasons as above (learning-by-doing and analyst survival effects), specific experience, measured by the number of years in which an analyst supplied at least one forecast on a given firm, should increase forecast accuracy. Clement (1999)'s results are in accordance with this assumption. They reveal a negative relationship between specific experience and the absolute value of the forecast error. Using a set of I/B/E/S forecasts between 1988 and 2000, Clarke and Subramanian (2006) confirm the role of specific experience.

The forecast frequency is a signal of forecast accuracy, in the sense that analysts react more to new information about a given firm when the frequency is high. This result appears in Jacob, Lys and Neale (1999), using a Zacks Investment Research database from 1981 to 1992, and in Clement (1999), Clement and Tse (2005), and Krishnan et al. (2006), who use a dataset from I/B/E/S over the period 1990-2004.

In the same vein, those studies reveal that the time elapsed between two forecasts indicates how outdated the forecast is: the longer a forecast goes unrevised, the less accurate it is. They also provide evidence that the greater the forecast horizon, the less accurate it is.

Finally, Brown (2001), using an I/B/E/S database over 1986-1998 and as Clement and Tse (2005), Krishnan, Lim and Zhou (2006) and Clarke and Subramanian (2006), explore the role played by past accuracy. They find that low past accuracy is associated with a low current absolute value of the difference between forecasted and actual earnings. This result suggests that forecast errors are persistent.

2.4 The Role of Soft Information

In addition to all of the observable variables described above, access to soft information appears to be an important, unobservable, determinant of financial analysts' forecast errors (Libby et al., 2008). Fund managers seem to highly value contacts with the firm's financial directors (Barker, 1998). In exchange for future trading commissions, analysts can provide access to this soft information. Because they forecast firms' earnings, analysts are in frequent contact with firm management. They can organize phone calls, conference calls and one-on-one meetings with corporate managers for their customers (portfolio and fund managers), to inform them about the projects or strategy of a particular firm (Fogarty and Rogers, 2005, Breton and Taffler, 2001). This access to qualitative soft information is only possible if the analyst has a good relationship with the firm's management. The desire for this access explains why analysts are prone to adopt forecasting behavior that satisfies firms' managers.

First, the analyst can issue pessimistic forecasts to allow the firm's manager to 'beat the forecast' and prompt a positive earnings surprise on the market. This scenario implies negative forecasting errors by analysts (Easterwood and Nutt, 1999; Lim, 2001). Second, the analyst can issue optimistic forecasts, with the aim of creating a positive reaction from stock markets (Payne and Robb, 2000; Matsumoto, 2002; Burgstahler and Eames, 2006). This scenario results in positive forecast errors. Pratt (1993), Gibson (1995), Womack (1996) and Boni and Womack (2002) provide several examples of situations in which analysts lost access to a firm's manager due to an unfavorable recommendation or earnings forecast about the firm.

Taken together, these arguments lead us to state the following testable assumption:

H1: High (negative or positive) forecast errors are associated with a close relationship between a firm and an analyst.

The following section presents the econometric methodology that will allow us to verify this testable assumption.

3 Empirical Investigation

Turning to our empirical investigation, we present our data and econometric model.

3.1 Data

We use data provided by ThomsonReuters. These data include I/B/E/S earnings forecasts and additional data from Worldscope. Our sample contains 241 French firms from the largest Paris' stock index *SBF 250*, diversified according to the firms' size and sector. We study one-year ahead EPS forecasts by 4 648 analysts from 1997 to 2007 on a monthly basis. This raw database consists of 265 238 firm-analyst-time observations. Several steps were required to clean the data. First, once issued, a forecast is frequently repeated for several months in the database. We obtained the number of monthly occurrences of each forecast by storing it in a variable called $DURATION$. Then, for each forecast, we dropped repeated occurrences of the same forecast, to avoid artificially counting it several times. Second, the date of realized EPS, i.e., the final day of the fiscal year's, was carefully checked. While some firms do not close their fiscal years prior to the 31st of December, the database systematically reports the realized EPS each month from January to December. Thus, some EPS artificially appear in January in the database although they are issued in march, for example. When a difference was detected, forecast errors were computed using fiscal years and not calendar years. Third, we dropped aberrant observations (for example when there are several different forecasts from the same analyst, on the same day, regarding the same firm, etc). As the reported forecasts are supposed to be one-year ahead earnings forecasts, we created a variable denoted $TERM$, measuring the number of days between the earnings announcement date and the forecast release date. We then dropped forecasts with a negative 'horizon' value, or with a 'horizon' value exceeding 365 days (366 for leap years). Finally we obtain 102 876 firm-analyst-time forecast' observations.

3.2 Econometric Model

3.2.1 General Methodology

The goal of our econometric study is to assess the strength of the relationship between a firm and an analyst and the importance of soft information in analysts' forecasts. The main contribution of our empirical approach is to employ a gravity model approach, that is typically applied in international economics to analyze trade relationships between two countries.

By definition, the effect of soft information on forecast accuracy is not captured by observable determinants. Therefore, when regressing the analysts' forecast accuracy on the variables described in Sections 2.1 to 2.3, the impact of soft information, highlighted in section 2.4, should be captured by the disturbance term of the regression. More precisely, as soft information is produced by *a given analyst* regarding *a given firm*, our empirical strategy is to compute a measure of soft information based on the firm-analyst specific effect. We then study the relationship between this measure and analysts' forecast to determine whether soft information contributes to biased (pessimistic or optimistic) EPS forecasts. This determination requires two steps, that are presented in the two following subsections.

3.2.2 First Step: Estimation Using a Gravity Model Approach

The first step of our investigation consists of estimating the following model:

$$AFE_{i,j,t} = \alpha + \beta X_{i,t} + \gamma Y_{j,t} + \delta Z_{i,j,t} + \lambda_i D_i + \mu_j D_j + \eta_{i,j} + \upsilon_t + \varepsilon_{i,j,t} \tag{1}$$

The dependent variable, denoted $AFE_{i,j,t}$, is the absolute forecast error for the firm i's EPS, forecasted by analyst j at date t.

Our empirical model contains three sets of explanatory variables. The variable $X_{i,t}$ denotes firm characteristics which are invariant across analysts in t. Symmetrically, $Y_{j,t}$ denotes analyst characteristics which are invariant across firms in t. Finally, $Z_{i,j,t}$ contains a set of variables which are specific both to firm i and analyst j in t. The dummy variable D's indicating a specific analyst j or a firm i[3].

As in all gravity models[4], the disturbance effect is decomposed into three effects:

[3]It is true that our model does not control for analyst-industry effects: how well an analyst understands the industry to which the firm belongs might impact the quality of forecasts. Also, information at the industry level may serve as an 'anchor' to the analyst, and as such, can have an impact on the forecasts errors (Cen, Hilary and Wei, 2012). However, as indicated below in the summary statistics, each financial analyst in our data set covers on average two sectors such that analyst-industry effects are primarily captured either by the analyst dummy or the variable recording the number of sectors covered; furthermore, we ran industry-level regressions that did not qualitatively change the results .

[4]Our problem differs from standard gravity models in the sense that firms and analysts are not as

the firm-specific effect λ_i, the analyst-specific effect μ_j, and the pair-specific effect $\eta_{i,j}$. Finally, υ_t denotes the time-specific effect. We assume that these effects are fixed (non stochastic). However, it is well known that the use of a within approach (OLS on demeaned variables) in the presence of invariant and/or rarely changing variables may lead to inefficiency and incorrect inferences. In model (1), this issue could arise not only in the firm's dimension i, but also in the analyst's dimension j. Indeed, many of our explanatory variables $Y_{j,t}$, $X_{i,t}$ and/or $Z_{i,j,t}$ exhibit very small variation in one of these dimensions. Therefore we propose the use the Fixed Effect Vector Decomposition (FEVD) methodology proposed by Plümper and Troeger (2007). This approach consists of a three-stage procedure (similar to that proposed by Mundlack (1978), for the random effects model): the first stage of the estimator runs a fixed-effects model to obtain the unit effects, the second stage decomposes the unit effects into a component explained by the time-invariant and/or rarely changing variables and an error term, and the third stage reestimates the first stage by pooled OLS including the time-invariant variables and the error term from stage 2, which then accounts for the unexplained component of the unit effects.

When estimating model (1), we consider three categories of explanatory variables. Variables denoted $X_{i,t}$ refer to firms' characteristics. In line with Jackson (2005), we first consider three determinants of the analysts' forecasting error. The absolute forecast error is expected to increase in earnings predictability (denoted $EPSPREV_{i,t}$), decrease in firm size (denoted $SIZE_{i,t}$) and decrease in analysts coverage (denoted $COVER_{i,t}$). Following Das et al. (1998), we also consider two interactive terms. The

symmetric as two countries are. Therefore there are no dyadic variables in our database. However, this does not hamper our results, because we focus on the three-way decomposition of the disturbance term. Moreover, the concept of the 'distance' between an analyst and a firm is what we want to capture, although the distance in question is social rather than geographical. Similar to geography in standard gravity models, the forecasting technology of the analysts (which depends on their training and reflects their global financial analysis 'philosophy'), is assumed to be time-invariant.

term $EPSPREV_{i,t}.SIZE_{i,t}$ accounts for interactions between EPS predictability and firm size. As firm size should mitigate impact of EPS variability, its coefficient should be negative. The term $EPSPREV_{i,t}.COVER_{i,t}$ stands for interactions between EPS predictability and coverage. For the same reason as above, its coefficient should be negative. Finally, as in Lim (2001), we include past optimistic consensus about a firm, denoted $PASTMDFE_{i,t}$. As past optimistic consensus encourages analysts to follow the crowd by also being optimistic, the expected sign of this variable is positive.

The second category of explanatory variables refer to analysts' characteristics, denoted $Y_{j,t}$. As underlined by Clement (1999), absolute forecast errors are expected to: decrease in ($GENEXP_{j,t}$), which stands for the general experience of analyst j, increase with $NBFIRM_{j,t}$ and $NBSEC_{j,t}$, the number of firms and the number of sectors followed by the analyst respectively, and decrease in $BROKER_{j,t}$, the size of the brokerage house.

The third category of variables relates to determinants specific both to firm i and analyst j in t, $Z_{i,j,t}$. First, as in Clarke and Subramanian (2006) and Clement (1999), the error should decrease with $SPECEXP_{i,j,t}$, the specific experience of analyst i with firm j. Following Clement and Tse (2005) and Krishnan, Lim and Zhou (2006), absolute forecast errors should: be greater if the forecast is far from the end of the fiscal year (denoted $TERM_{i,j,t}$), decrease if the analyst frequently revises his forecasts (denoted $FREQ_{i,j,t}$), and hence increase in the forecast's lifetime in the database (denoted $DURATION_{i,j,t}$) and increase in past errors (denoted $PASTAFE_{i,j,t}$).

Table 1, in the Appendix, reports the list of regression variables mentioned above and how they are computed while Tables 2 and 3, also in the Appendix, provide summary statistics and correlation coefficients, respectively. The coefficients reported in Table 3 are generally consistent with the expected correlations. The forecast error of analyst j for firm i is positively correlated with the degree of firm i EPS predictability, the past median forecast error regarding firm i, the number of firms and sectors covered by analyst j, the number of days to the fiscal year end and the latest forecast for the firm-analyst pair. It is negatively correlated with the size of firm i, the general and the specific experience of analyst j, the size of the broker and the forecast's lifetime in the database. Some other correlations are also noteworthy. First, specific and general experience are positively correlated. Moreover, both variables are positively linked to the size of the broker and the frequency of forecasts. Finally, there is also a negative correlation between the size of the firm and its degree of EPS variability, which lends some support to the notion that $SIZE$ may mitigate the positive impact of $EPSPREV$ on the forecast error.

3.2.3 Second Step: Analyzing the Pair-Specific Effect

In a second step, we focus on the pair-specific effect of our estimate to produce a measure of soft information, i.e., the information obtained within the framework of a privileged relationship between the firm and the analyst that is not captured in the observable determinants defined in the previous subsection.

Our methodology decomposes the fixed effect of the panel regression into three components: the firm effect (λ_i), the analyst effect (μ_j), and the firm-analyst' pair effect ($\eta_{i,j}$). Each accounts for the unobservable determinants influencing the absolute forecast error.

Having estimated our model, we are able to compute $C_{i,j}$, the 'contribution of the pair-specific effect', defined as follows:

$$C_{i,j} = \frac{\mid \eta_{i,j} + \alpha \mid}{\mid \lambda_i \mid + \mid \mu_j \mid + \mid \eta_{i,j} + \alpha \mid}$$

, where α is the unconditional constant of the regression. The pair-specific effect, computed by the STATA procedure of Plümper and Troeger (2007), is centered on a mean calculated over all observations. Adding the unconditional constant α in both the denominator and the numerator of $C_{i,j}$ makes it possible to correct for this bias and construct a consistent indicator.

$C_{i,j}$, defined as the pair-specific effect in absolute value relative to the absolute value of the total fixed effect, measures the importance of the pair-specific effect as a determinant of $AFE_{i,j,t}$, relative to firm-specific and analyst-specific effects. For a given forecast, the greater the contribution of the pair-specific effect, the greater the effect of soft information and the closer the relationship between a particular firm and a particular analyst. When $C_{i,j}$ is low, the unexplained component of the forecast error is primarily due to either the firm alone or to the analyst alone. When $C_{i,j}$ is high, the unexplained component of the forecast error is primarily due to the specific firm i-analyst j pair.

Finally, we rank observations by 'twentiles' of the relative median forecast error $MedFE_{i,j}$, defined as the median, over the full sample period, of the difference between the EPS forecasts and the EPS realization of each firm i by each analyst j for a forecast issued in t. We obtain 20 groups from the 5% most pessimistic AFE to the 5% most optimistic AFE. For each observation of $MedFE_{i,j}$, we are able to match the $C_{i,j}$ value that concerns the analyst and the firm involved in this observation. We then compute the mean value of $C_{i,j}$ for each 'twentile' (from the 5% most pessimistic to the 5% most optimistic). We plot the median forecast error 'twentiles' against the pair-effect contribution. Such a graph should allow us to verify our testable assumption H1.

If we observe that $C_{i,j}$ is greater for the most extreme (negative or positive) forecast error 'twentiles', this means that the pair-effect accounts for most of the residuals *when the forecast error is high*. In other words, controlling for all observed variables (including analyst and firm dummies), the unobservable factors *specifically related to the firm i-analyst j pair* play a greater role than analyst- or firm-specific unobservable factors, when the forecast error is high. We interpret this result as the demonstration of the soft information channel between the analyst and the firm. It captures the fact that the need to maintain access to soft information provide analysts with an incentive to issue pessimistic or optimistic forecasts. Conversely, we should observe that $C_{i,j}$ is weaker for central 'twentiles' (i.e. forecast error around zero). If this prediction is true, H1 is validated.

4 Results

We now present the results of panel regressions. We first present our results concerning the estimation of model [1]. Second, we comment on the test of our assumption concerning the relationship between the pair-effect contribution and the forecast error.

4.1 Results of Panel Regressions with Vector Decomposition

The results of panel regressions with vector decomposition for model [1] are reported in Table 4, in the Appendix.

First, we focus on variables $X_{i,t}$, that capture firm-specific characteristics. We observe that $EPSPREV$ has a positive impact on the dependent variable. In line with the results obtained by Das et al. (1998), Lim (2001) and Jackson (2005), this suggests that the more difficult it is to predict the firm's EPS, the larger the analyst's forecast error. $COVER$ has a negative impact on the dependent variable. This observation is consistent with Lim (2001): as public information availability is enhanced for firms followed by a large number of analysts, a higher analyst coverage decreases forecast error. The coefficient for the interaction term $EPSPREV.COVER$ also has the expected (negative) sign. This finding indicates that analyst coverage mitigates the impact of EPS variability (Das et al., 1998). Finally, following Lim (2001), the coefficient for $PASTMDFE$ has the expected positive sign. The greater the past optimistic consensus, the larger the analysts' forecast error. Table 4 also indicates that, in contrast to the prior empirical literature, a firm's size has no effect on the dependent variable. Finally, the coefficient for the interaction term $EPSPREV.SIZE$ is positive, suggesting that firm size amplifies the impact of EPS variability.

We now turn to variables $Y_{j,t}$, that represent analysts' characteristics. General experience ($GENEXP$) reduces forecast error. In line with Clement (1999) and Lim (2001), the greater the general experience of the analyst, the lower the analysts' forecast error. $NBFIRM$ has a significant and positive sign. This supports theoretical predictions and Clement (1999)'s empirical finding: when an analysts follows a large number of firms, he dedicates fewer resources to each of them such that the forecast error is higher. Table 4 also reveals that the number of sectors followed by the analyst has no impact on the dependent variable because the coefficient for $NBSECT$ is insignificant. Although this observation contradicts the theoretical intuition, it is consistent with many of the annual regression results obtained by Clement (1999). Interestingly, as expected, the coefficient for $BROKER$ has a significant and negative sign. This result, which is in line with Lim (2001) and Clement (1999), suggests that being employed by a large broker allows an analyst to dedicate more resources to prediction and produce more accurate forecasts.

Finally, we comment on our results concerning variables $Z_{i,j,t}$, that are specific both to firms i and analysts j. First, the coefficient for $SPECEXP$ exhibits a negative sign:

as expected (and in line with the literature), the absolute forecast error decreases in the specific experience of analysts. Table 4 also indicates that in line with theory, the coefficient for $TERM$ is significant and positive: the further from the end of the fiscal year, the less accurate the analyst's forecast. While the coefficient for $DURATION$ is not significant, the coefficient for $FREQ$ is positive, which is not the expected impact on the dependent variable. Finally, $PASTAFE$ has the expected positive sign, which indicates inertia in forecast dynamics.

4.2 Graphs of pair-effect contribution $C_{i,j}$ by 'twentiles' of median forecast error $MedFE_{i,j}$

Graph 1 provides interesting representations of $C_{i,j}$ by 'twentiles' of $MedFE_{i,j}$. The first 'twentile' represents the most pessimistic forecasts (FE approximately -2), the 10th 'twentile' represents the most accurate forecasts (approximately 0) and the 20th 'twentile' represents the most optimistic forecast (FE approximately +9). We observe that there is a non-linear relationship between $C_{i,j}$ and $MedFE_{i,j}$. This is represented by a convex curve. $C_{i,j}$ is at its lowest level when $MedFE_{i,j}$ lies in the tenth 'twentile'. When forecast error is weak, the pair-effect only accounts for approximately 30% of the total fixed effect. When the forecast error is high, the pair effect account for over 40% of the fixed effect. This result is interesting for at least two reasons. First, it suggests that the relationship between $C_{i,j}$ and $MedFE_{i,j}$ can be represented by a non-linear curve. Second, this finding indicates that the contribution of the pair-specific effect reaches its minimum for intermediate values of $MedFE_{i,j}$ while is at its maximum for both the most optimistic and for the most pessimistic forecast 'twentile'. This observation means that the contribution of the pair-specific effect is more important when the forecast error is high. Our results thus validate our testable assumption H1 and provide evidence that some analysts attempt to maintain friendly relationships with some firms' managers by intentionally biasing their EPS forecasts.

5 Robustness Checks

In this section, we propose three robustness checks for our findings. We first estimate several variants of our model, inspired by the literature. We then discuss the FEVD estimator. We finally test for the relationship between the pair-effect contribution and forecast errors.

5.1 Variants of the Model

First, we estimate several variants of model [1]. The goal of this subsection is to 'reproduce' the estimates conducted in some papers in the empirical literature and to examine whether they result in the same representations of $C_{i,j}$ by 'twentiles' of $MedFE_{i,j}$ as model [1].

In variant [2], we refer to the estimate by Jackson (2005), which only considers earnings predictability ($EPSPREV_{i,t}$), firm size ($SIZE_{i,t}$) and analysts coverage ($COVER_{i,t}$).

Specification [3] relates to the study by Das et al. (1998), which considers the three variables mentioned above and the two interactive terms, $EPSPREV_{i,t}.SIZE_{i,t}$ and $EPSPREV_{i,t}.COVER_{i,t}$.

In variant [4], we follow Lim (2001) by including all variables contained in specification [3] and past optimistic consensus ($PASTMDFE_{i,t}$).

Variant [5] refers to the study by Clement (1999), which considers the general experience of analyst j ($GENEXP_{j,t}$), the number of firms ($NBFIRM_{j,t}$) and sectors

($NBSEC_{j,t}$), brokerage house size ($BROKER_{j,t}$) as well as the specific experience of analyst i with firm j ($SPECEXP_{i,j,t}$).

Finally, specification [6] is inspired by Clement and Tse (2005), who add four additional variables to those included in variant [5]: how far the forecast is from the end of the fiscal year ($TERM_{i,j,t}$), analyst revision frequency ($FREQ_{i,j,t}$), the forecast's lifetime in the database ($DURATION_{i,j,t}$) and past errors ($PASTAFE_{i,j,t}$).

The results concerning these variants are reported in Table 5, in the Appendix. Concentrating on the variables $X_{i,t}$, we observe that in all specifications in which it is included ([2], [3] and [4]), $EPSPREV$ has a positive impact on the dependent variable. These findings, which reinforce the results obtained in Section 4, are in line with the empirical literature (Lim, 2001, Das et al., 1998, Jackson, 2005). While it was not significant in model [1], the coefficient for $SIZE$ is significant with the expected sign (i.e. negative) in specifications [3] and [4]. This result is consistent with the empirical literature. Reflecting the diversity of findings obtained in the literature, results concerning $COVER$ are less robust. In line with Lim (2001), its coefficient has the expected (i.e. negative) sign in specification [4] but has a positive sign in specification [2]. Moreover, following Das et al. (1998) and Jackson (2005), it is insignificant in specification [3]. Turning to the interactions terms $EPSPREV.SIZE$ and $EPSPREV.COVER$, one observes that their respective coefficients vary across specifications [3] and [4]. In the former, $EPSPREV.COVER$ has a positive impact on AFE (this was also the case in model [1]) but, in line with theory and the empirical results of Das et al. (1998), its coefficient is negative in the latter. Moreover, while its coefficient is never significant in Das et al. (1998), $EPSPREV.COVER$ has a negative impact on FE in variant [3] but a positive impact in variant [4]. Finally, following Lim (2001) and our previous estimates, the coefficient for $PASTMDFE_{i,t}$ has the expected positive sign (variant

[4]).

We now comment on the coefficients for variables $Y_{j,t}$. As expected, variants [5] and [6] in Table 5 indicate that the forecast error decreases in general experience ($GENEXP$). This observation is in line with the results obtained in Table 4 and the empirical literature (Lim, 2001; Clement, 1999). As in our previous estimates, $NBFIRM$ has a significant and positive sign (specification [5]). This finding supports the theoretical and the empirical literature (Clement, 1999). Specification [5] also indicates that the number of sectors followed by the analyst ($NBSECTOR$) has no impact on the dependent variable. This observation is consistent with our previous findings and Clement (1999). Interestingly, as expected, the coefficient for $BROKER$ has a significant and negative sign in specifications [5] and [6]. This result, which confirms the findings obtained in Table 4, is in line with Lim (2001) and Clement (1999).

Finally, we turn to the results concerning variables $Z_{i,j,t}$. In line with Clement (1999)'s findings, the specific experience of analysts ($SPECEXP$) has a significant and positive impact on the dependent variable (column [5]). However, its coefficient is insignificant in column [6]. Variant [6] in Table 5 also indicates that in line with our previous results, the coefficient for $TERM$ is significant and positive. Consistently with the findings presented in Table 4, the coefficient for $PASTAFE$ has the expected positive sign, aindicating inertia in the forecast dynamics. Finally, as in model [1], $DURATION$ and $FREQ$ do not have the expected impact on the dependent variable.

Graphs 2 to 6 provide interesting representations of $C_{i,j}$ by 'twentiles' of $MedFE_{i,j}$ for specifications [2] to [6] respectively. All graphs exhibit a non-linear relationship between $C_{i,j}$ and $MedFE_{i,j}$. In all specifications except variant [5], it is represented by a convex curve. These findings suggest that the results obtained in Section 4.2

are quite robust. Except in specification [5], one observes the same phenomenon as in Section 4.2: whatever the variant and the number of characteristics included in the estimation, the contribution of the pair-specific effect is at its minimum for intermediate values of $MedFE_{i,j}$ and at its maximum for both the most optimistic and for the most pessimistic forecast 'twentile'[5]. Overall, these findings reinforce the notion that maintaining good relationships with firm managers to preserve their access to 'soft' information incentivizes some analysts to issue pessimistic or optimistic forecasts.

5.2 Discussion on the FEVD Estimator

In this section, we discuss the FEVD estimator used to estimate model [1] and variants [2]-[6]. It is worth noting that the FEVD estimator is equivalent to a standard instrumental variables approach, for a specific set of instruments as recently shown by Breusch et al. (2011). Greene (2011) argues that the FEVD approach does not provide an estimator for the coefficients for time invariant variables in a fixed effects model: that component of the parameter vector remains unidentified. However in the presence of slowly changing variables as in our context, this estimator remains consistent, even if the efficiency gains (compared to Hausman and Taylor's, 1981, approach) are controversial. One advantage of the FEVD is that, as the Fixed Effects (Within) estimator, and unlike that of Hausman and Taylor, it does not require specifying the exogeneity status of the explanatory variables. For Greene (2011), the FEVD estimator simply reproduces (identically) the linear fixed effects (dummy variable) estimator, and then

[5] Concerning specification [5], remember that it represents the regression only including analyst-specific variables and one pair-specific observable determinant. A concave function may then illustrate the fact that, because firm-specific and other pair-specific determinants are excluded, firm- and pair-effects capture all of the effect of omitted variables. This is why $C_{i,j}$ represents approximately 50% of the total fixed effect for median values of $MedFE_{i,j}$. Morevoer, this is not the case for specification [2] when the regression contains only firm-specific variables, because the firm-effect seems to be more important than the analyst effect.

substitutes an inappropriate covariance matrix for the correct one. This is why we propose an additional estimate of the covariance matrix here.

The covariance matrix discussed in Subsection 4.1 is defined in the context of the three stages estimation procedure by Plümper and Troeger (2007). In Table 6 (in Appendix), we compute the value of the standard errors obtained using the covariance matrix proposed by Greene (2011), which corresponds to the matrix estimated in the first stage of the Plümper and Troeger (2007) procedure, for model [1] and for variants [2] to [6]. Our goal is to compare both standard errors to show that Greene's (2011) argument has limited relevance for our purposes, which is to investigate the relationship between pair effects and forecast errors. Our results indicate that in most cases, standard errors using the covariance matrix from the first stage of the estimation procedure of Plümper and Troeger (2007), as is proposed by Greene (2011), are higher than those obtained through the third stage. Although this observation modifies the significance of the coefficients reported in Section 4, the coefficients and residuals remain unchanged. Therefore our use of the decomposition of the fixed effect is appropriate for our data and purposes.

5.3 Examining the Relationship Between the Pair-Effect and Forecast Error

In this subsection, we test for the relationship between $C_{i,j}$ and $MedFE_{i,j}$ to determine whether the contribution of the pair effect differs according to the size of the forecast error. We first conduct a median-test that allows us to investigate whether the median of $C_{i,j}$ in a given quantile (here, 'twentile') equals the median of the full sample. We then conduct the Bartlett-test that tests for the equality of variance of $C_{i,j}$ across quantiles. Finally, we use the Krusal-Wallis equality-of-population rank test, which determines whether the rank sum of each observation ranked by $C_{i,j}$ differs across quantiles. The results of these three robustness checks are reported in Table 7, in the Appendix. They indicate that for each test, testable assumption H1 is still validated: the contribution of the pair-effect differs according to the size of the forecast error. This result seems particularly robust since its holds for each of our six specifications.

6 Conclusion

The goal of this paper was to determine whether the analysts' concern for preserving their access to soft information provides them with incentives to issue optimistic or pessimistic forecasts. We employed a Thomson Reuters data set that contains the forecasts issued by 4 648 analysts concerning the earnings of 243 French firms over the period 1997-2007.

One important innovation of our approach is to employ the gravity model methodology to examine relationships between firms and analysts. The second interest of our paper is to propose a measure of soft information. Having regressed analysts forecast error on observable firm-specific, analyst-specific and pair-specific characteristics, we decompose the disturbance effect to extract a pair-specific effect. This effect provides a measure of soft information, allowing us to determine whether soft information contributes to analysts' pessimism or optimism. Finally, we provide interesting evidence that the need to preserve their relationships with firm managers and their access to soft information prompts some analysts to issue pessimistic forecasts while prompting some others to issue optimistic forecasts about firms' EPS.

Our results undoubtedly call for further research. Of course, our work could be extended to the case of other countries, to determine whether the effect of soft information has a national dimension. More ambitiously, it would be interesting to examine the consequences of analysts' forecast on portfolio investment strategies. For example, this investigation could be performed by studying what type of forecasting profile (accurate, pessimistic or optimistic) leads to the most profitable investment recommendations for asset managers.

References

Barber, B., Lehavy, R., Trueman, B., 2007. Comparing the stock recommendation performance of investment banks and independent research firms. Journal of Financial Economics. 85, 490-517.

Barker, R.G., 1998. The market for information : evidence from finance directors, analysts and fund managers. Accounting and Business Research. 29, 3-20.

Boni, L., Womack, K.L., 2002. Wall Street credibility problem: misaligned incentives and dubious fixes. Brookings Wharton Papers on Financial Services, 93-130.

Breton, G., Taffler, R., 2001. Accounting information and analyst stock recommendation decisions: a content analysis approach. Accounting and Business Research. 31, 91-101.

Breusch, T., Ward, M., Nguyen, H., Kompas T., 2011. On the fixed-effects vector decomposition. Political Analysis. 19, 129-134

Brown, L., 1997. Analyst forecasting error: additional evidence. Financial Analysts Journal. 53, 81-88.

Brown, L., 2001. How important is past analyst forecast accuracy. Financial Analysts Journal. 57, 46-49.

Burgstahler, D., and Eames, M., 2006. Management of earnings and analysts' forecasts to achieve zero and small positive earnings surprises. Journal of Business Finance and Accounting. 33, 633-652.

Cen, L., Hilary, G., Wei, K.C.J., 2012. The role of anchoring bias in the equity market: evidence from analysts' earnings forecasts and stock returns. Journal of Financial and Quantitative Analysis, in press, CJO2012. doi:10.1017/S0022109012000609.

Chan, L.K.C., Karceski, J., Lakonishok, J., 2007. Analysts' conflicts of interest and biases in earnings forecasts. Journal of Financial and Quantitative Analysis. 42, 893-914.

Clarke, J., Subramanian, A., 2006. Dynamic forecasting behavior by analysts: theory and evidence. Journal of Financial Economics. 80, 81-113.

Clement, M., 1999. Analyst forecast accuracy: do ability, resources, and portfolio complexity matter? Journal of Accounting and Economics. 27, 285-303.

Clement, M., Tse S., 2005. Financial analyst characteristics and herding behavior in forecasting Journal of Finance. 60, 307-341.

Das, S., Levine, C., Sivaramakrishnan C., 1998. Earnings predictibility and bias in analysts' earnings forecasts. Accounting Review. 73, 277-294.

Dugar, A., Nathan, S., 1995. The effect of investment banking relationships on financial analyst's earnings forecasts and investment recommendations. Contemporary Accounting Research. 12, 131-160.

Easterwood, J., Nutt, S., 1999. Inefficiency in analysts' earnings forecasts: systematic misreaction or systematic optimism. Journal of Finance. 54, 1777-1797.

Fogarty, T., Rogers, K., 2005. Financial analysts' reports: an extended institutional theory evaluation. Accounting, Organization and Society. 30, 331-356.

Francis, J., Philbrick, D., 1993. Analysts' decisions as product of multi-task environment. Journal of Accounting Research. 31, 216-230.

Gibson, R., 1995. Analysts take heat for 'sell' rating on Archer-Daniels. Wall Street Journal. august 31.

Greene, W., 2011. Fixed effects vector decomposition: A magical solution to the problem of time invariant variables in fixed effects models? Political Analysis. 19, 135-146.

Hausman, J., Taylor, W., 1981. Panal data and unobservable individual effects. Econometrica. 49, 1377-1398.

Hayward, M., Boecker, W., 1998. Power and conflicts of interest in professional firms: evidence from investment banking. Administrative Science Quarterly. 43, 1-22.

Jackson, A., 2005. Trade generation, reputation and sell-side analysis. Journal of Finance. 60, 673-717.

Jacob, J., Lys, T.Z., Neale, M.A., 1999. Expertise in forecasting performance of security analysts. Journal of Accountings and Economics. 28, 51-82.

Krishnan, M., Lim, S., Zhou, P., 2006. Analysts' herding propensity: theory and evidence from earnings forecasts. Working paper, http://ssrn.com/abstract=929467.

Libby, R., Hunton, J., Seybert, N., 2008. Relationship incentives and the optimistic/pessimistic pattern in analysts' forecasts. Journal of Accountings Research. 46, 173-198.

Lim, T., 2001. Rationality and analysts' forecast bias. Journal of Finance. 56, 369-385.

Lin, H.W., McNichols, M., O'Brien, P., 2005. Analyst impartiality and investment banking relationships. Journal of Accounting Research. 43, 623-650.

Matsumoto, D., 2002. Management's incentives to avoid negative earnings surprises. Accounting Review. 77, 483-514.

Mundlak, Y., 1978. On the pooling of times series and cross section data. Econometrica, 46, 69-85.

Payne, J., Robb, S., 2000. Earnings management: The effect of ex-ante earnings expectations. Journal of Accounting, Auditing and Finance. 15, 371-92.

Plümper, T., Troeger, V., 2007. Efficient estimation of time-invariant and rarely changing variables in finite sample panel analyses with unit fixed effects. Political Analysis, 15, 124-139.

Pratt, T., 1993. Wall Street four letter word. Investment Dealers' Digest, march 29.

Womack, K.L., 1996. Do brokerage analysts' recommendations have investment value? Journal of Finance. 51, 137-167.

Appendix

Table 1: List of regression variables

Variables described in Table 1 are defined on a set of one-year-ahead EPS forecasts issued by 4 648 analysts concerning 241 French firms (1997-2007).

	DEPENDENT VARIABLE
AFE	Absolute forecast error (absolute difference between the EPS forecast and the EPS realization of each firm i by each analyst j at date t of forecast issue)
	INDEPENDENT VARIABLES AND EXPECTED SIGNS
	Firm's characteristics $\mathbf{X}_{i,t}$
$EPSPREV$ (+)	Predictibility of EPS (volatility of firm i's EPS over the last 3 years)
$SIZE$ (-)	Size of the firm (log of the market capitalization of firm i in t)
$COVER$ (-)	Coverage of the firm in t (number of analysts who follow firm i in t)
$PASTMDFE$ (+)	Consensus surprise (for each firm i, the median of the difference between the consensus and the realized EPS in the previous year
	Analysts' characteristics $\mathbf{Y}_{j,t}$
$GENEXP$ (-)	General experience of analyst j (in t, number of days since the analyst's first forecast)
$NBFIRM$ (+)	Number of firms followed by analyst j in t
$NBSECT$ (+)	Number of sectors followed by analyst j in t
$BROKER$ (-)	Size of the broker (number of analysts working for analyst j's broker in t)
	Firm-analysts' characteristics $\mathbf{Z}_{i,j,t}$
$SPECEXP$ (-)	Specific experience of the analyst (in t, number of days since the first forecast by analyst j about firm i)
$TERM$ (+)	Number of days from t to fiscal year end for a forecast issued by analyst j on firm i
$FREQ$ (-)	Frequency of forecasts (number of forecasts per year by analyst j on a firm i in t)
$PASTAFE$ (+)	forecast error of analyst j on firm i in the previous year
$DURATION$ (+)	forecast lifetime in the database in months (by analyst j on a firm j in t)

Table 2: Statistical summary for regression variables (1997-2007)

Table 2 presents descriptive statistics for variables. The sample period is from 1997 to 2007. The statistics are both cross-sectional and cross-period. AFE is the absolute difference between the EPS forecast and the EPS realization of each firm i by each analyst j at date t of forecast issue. $EPSPREV$ is the volatility of firm i's EPS over the last 3 years. $SIZE$ is the log of the market capitalization of firm i in t. $COVER$ is the number of analysts who follow firm i in t. $PASTMDFE$ is, for each firm i, the median of the difference between the consensus and the realized EPS in the previous year. $GENEXP$ is, in t, the number of days since the analyst's first forecast. $NBFIRM$ is the number of firms followed by analyst j in t. $NBSECT$ is the number of sectors followed by analyst j in t. $BROKER$ is the number of analysts working for analyst j's broker in t. $SPECEXP$ is, in t, number of days since the first forecast by analyst j about firm i. $TERM$ is tumber of days from t to fiscal year end for a forecast issued by analyst j on firm i. $FREQ$ is the number of forecasts per year by analyst j on a firm i in t. $PASTAFE$ is the forecast error of analyst j on firm i in the previous year. $DURATION$ is the forecast lifetime in the database in months by analyst j on a firm j in t.

Variables	Mean	Standard deviation	Max	Min	Nonmissing observations
AFE	2.46	6.12	162.57	0	102 876
$EPSPREV$	1.79	4.72	85.17	1.15	94 231
$SIZE$	21.93	1.77	25.95	15.30	102 627
$COVER$	19.49	9.60	47	1	102 876
$PASTMDFE$	2	6.54	75.97	-39.90	94 595
$GENEXP$	1 344.29	1 141.75	6 434	0	102 875
$NBFIRM$	4.61	3.49	24	1	102 876
$NBSECT$	2.22	1.39	9	1	102 876
$BROKER$	17.93	10.15	62	1.	102 876
$SPECEXP$	849.65	912.73	6 253	0	102 875
$TERM$	187.88	102.26	365	0	102 875
$FREQ$	4.01	1.98	13	1	102 876
$PASTAFE$	2.48	6.16	127.27	0	69336
$DURATION$	2.82	2.14	18	1	102 876

Table 3: Correlation coefficients of regression variables

Table 3 presents correlation coefficients of variables. The sample period is from 1997 to 2007. Coefficients are cross-sectional and cross-period. For each pair of variables, the correlation coefficient is calculated as the ratio between the covariance of both variables and the product of each variable's standard errors.

	AFE	*EPSPREV*	*SIZE*	*COVER*	*PASTMDFE*	*GENEXP*	*NBFIRM*	*NBSECT*	*BROKER*
AFE	1								
EPSPREV	0.187	1							
SIZE	-0.024	-0.040	1						
COVER	0.063	0.016	0.741	1					
PASTMDFE	0.391	0.007	-0.046	0.127	1				
GENEXP	-0.054	0.004	-0.028	-0.073	-0.078	1			
NBFIRM	-0.004	-0.015	-0.379	-0.272	0.017	0.249	1		
NBSECT	-0.03	-0.010	-0.360	-0.305	0.013	0.190	0.7068	1	
BROKER	-0.065	0.008	-0.141	-0.217	-0.088	0.124	0.021	0.026	1
SPECEXP	-0.041	0.005	0.126	0.094	-0.061	0.728	0.093	0.025	0.078
TERM	0.012	-0.005	0.016	-0.257	0.001	-0.016	-0.078	0.002	0.0399
FREQ	-0.002	-0.014	-0.113	0.140	0.018	-0.024	0.023	0.031	-0.052
PASTAFE	0.462	0.039	-0.037	0.096	-0.772	0-.068	0.019	0.093	-0.065
DURATION	-0.008	-0.016	-0.119	-0.138	0.019	-0.035	0.025	0.029	-0.049
	SPECEXP	*TERM*	*FREQ*	*PASTAFE*	*DURATION*				
SPECEXP	1								
TERM	-0.007	1							
FREQ	-0.037	0.055	1						
PASTAFE	-0.054	-0.005	-0.002	1					
DURATION	-0.044	0.129	-0.451	0.010	1				

Table 4: Results of panel regression with vector decomposition for model [1]

Table 4 presents the results of panel regression for the following model: $AFE_{i,j,t} = \alpha + \beta X_{i,t} + \gamma Y_{j,t} + \delta Z_{i,j,t} + \lambda_i D_i + \mu_j D_j + \eta_{i,j} + \upsilon_t + \varepsilon_{i,j,t}$. The variable $X_{i,t}$ denotes firm characteristics which are invariant across analysts in t. The variable $Y_{j,t}$ denotes analyst characteristics which are invariant across firms in t. The variable $Z_{i,j,t}$ contains a set of variables which are specific both to firm i and analyst j in t. The dummy variable D's indicating a specific analyst j or a firm i. We use the Fixed Effect Vector Decomposition (FEVD) methodology proposed by Plümper and Troeger (2007). * and ** denote significance at the 5% and 1% levels, respectively.

Variables (expected sign)	
Firms' characteristics $\mathbf{X}_{i,t}$	
EPSPREV (+)	2.359**
	(0.081)
SIZE (-)	0.016
	(0.041)
COVER (-)	-0.021**
	(0.004)
EPSPREV.COVER (-)	-0.112**
	(0.004)
EPSPREV.SIZE (-)	0.012**
	(0.000)
PASTMDFE (+)	0.088**
	(0.004)
Analysts' characteristics $\mathbf{Y}_{j,t}$	
GENEXP (-)	-0.000**
	(0.000)
NBFIRM (+)	0.077**
	(0.009)
NBSECT (+)	0.000
	(0.336)
BROKER (-)	-0.009**
	(0.002)
Firm-analysts' characteristics $\mathbf{Z}_{i,j,t}$	
SPECEXP (-)	-0.000**
	(0.000)
TERM (+)	0.001**
	(0.000)
DURATION (+)	-0.003
	(0.008)
FREQ (-)	0.145**
	(0.011)
PASTAFE (+)	0.091**
	(0.004)
Sector dummies	yes
Analyst dummies	yes
Firm dummies	yes
Nb. obs.	65 586

Graph 1: $C_{i,j}$ by 'twentiles' of $MedFE_{i,j}$

$C_{i,j}$, the 'contribution of the pair-specific effect', is defined as follows: $C_{i,j} = \frac{|\eta_{i,j}+\alpha|}{|\lambda_i|+|\mu_j|+|\eta_{i,j}+\alpha|}$, where α is the unconditional constant of the regression reported in Table 4, λ_i the firm effect, μ_j the analyst effect, and $\eta_{i,j}$ the firm-analyst' pair effect. $MedFE_{i,j}$ is the median, over the full sample period, of the difference between the EPS forecasts and the EPS realization of each firm i by each analyst j for a forecast issued in t. Ranking observations by 'twentiles' of the (relative) median forecast error, we obtain 20 groups from the 5% most pessimistic AFE to the 5% most optimistic AFE. For each observation of $MedFE_{i,j}$, we match the $C_{i,j}$ value that concerns the analyst and the firm involved in this observation. The mean value of $C_{i,j}$ is then computed for each 'twentile' (from the 5% most pessimistic to the 5% most optimistic).

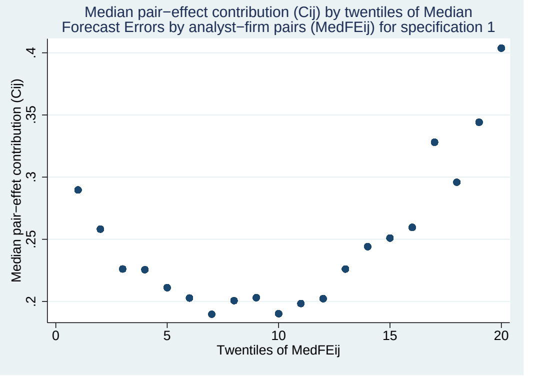


Table 5: Results of panel regression with vector decomposition for variants [2] to [6]

Table 5 presents the results obtained when we 'reproduce', using our own sample, the estimates conducted in some papers in the empirical literature. Variants [2], [3], [4], [5] and [6] refer to the estimates by Jackson (2005), Das et al. (1998), Lim (2001), Clement (1999) and Clement and Tse (2005), respectively. * and ** denote significance at the 5% and 1% levels, respectively.

Variables (expected sign)	Specifications				
	(2)	(3)	(4)	(5)	(6)
Firms' characteristics $\mathbf{X}_{i,t}$					
EPSPREV (+)	0.208**	1.677**	3.521**		
	(0.003)	(0.070)	(0.088)		
SIZE (-)	-1.066	-1.084**	-0.224**		
	(0.033)	(0.032)	(0.037)		
COVER (-)	0.023**	-0.002	-0.016**		
	(0.002)	(0.003)	(0.003)		
EPSPREV.COVER (-)		0.012**	-0.163**		
		(0.006)	(0.004)		
EPSPREV.SIZE (-)		-0.081**	0.008**		
		(0.003)	(0.000)		
PASTMDFE (+)			0.130**		
			(0.003)		
Analysts' characteristics $\mathbf{Y}_{j,t}$					
GENEXP (-)				-0.003**	-0.000**
				(0.000)	(0.000)
NBFIRM (+)				0.070**	0.053**
				(0.008)	(0.010)
NBSECT (+)				0.162	-0.010**
				(0.002)	(0.002)
BROKER (-)				-0.016**	-0.010**
				(0.002)	(0.002)
Firm-analysts' characteristics $\mathbf{Z}_{i,j,t}$					
SPECEXP (-)				0.002**	0.000
				(0.000)	(0.000)
TERM (+)					0.001**
					(0.000)
DURATION (+)					-0.005
					(0.008)
FREQ (-)					0.165**
					(0.003)
PASTAFE (+)					0.134**
					(0.003)
Sector dummies	yes	yes	yes	yes	yes
Analyst dummies	yes	yes	yes	yes	yes
Firm dummies	yes	yes	yes	yes	yes
Nb. obs.	85 398	94 026	79 778	102 875	69 336

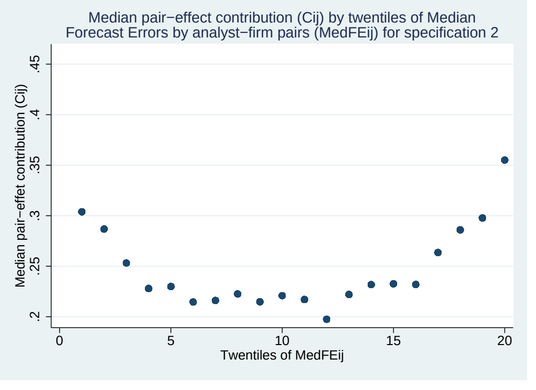
Table 6: Standard deviations using the FEVD procedure and using the covariance matrix proposed by Greene (2011) for model [1] and variants [2] to [6]

Table 6 reports the value of the standard errors obtained using the covariance matrix proposed by Greene (2011), which corresponds to the matrix estimated in the first stage of the Plümper and Troeger (2007) procedure. In parentheses: standard error using the FEVD procedure. In italics: the standard errors using the covariance matrix proposed by Greene (2011).

	Specifications					
	[1]	[2]	[3]	[4]	[5]	[6]
Firms' characteristics $\mathbf{X}_{i,t}$						
EPSPREV	(0.081)	(0.003)	(0.070)	(0.088)		
	0.093	*0.003*	*0.079*	*0.078*		
SIZE	(0.041)	(0.033)	(0.032)	(0.037)		
	0.052	*0.040*	*0.041*	*0.044*		
COVER	(0.004)	(0.002)	(0.003)	(0.003)		
	0.004	*0.003*	*0.003*	*0.003*		
EPSPREV.COVER	(0.004)		(0.006)	(0.004)		
	0.004		*0.004*	*0.004*		
EPSPREV.SIZE	(0.000)		(0.003)	(0.000)		
	0.000		*0.000*	*0.000*		
PASTMDFE	(0.004)			(0.003)		
	0.005			*0.003*		
Analysts' characteristics $\mathbf{Y}_{j,t}$						
GENEXP	(0.000)				(0.000)	(0.000)
	n.a.				*0.002*	*0.000*
NBFIRM	(0.009)				(0.008)	(0.010)
	0.012				*0.010*	*0.012*
NBSECT	(0.336)				(0.002)	(0.002)
	0.039				*0.031*	*0.038*
BROKER	(0.002)				(0.002)	(0.002)
	0.002				*0.002*	*0.002*
Firm-analysts' characteristics $\mathbf{Z}_{i,j,t}$						
SPECEXP	(0.000)				(0.000)	(0.000)
	0.000				*0.000*	*n.a.*
TERM	(0.000)					(0.000)
	0.000					*0.000*
DURATION	(0.008)					(0.008)
	0.009					*0.009*
FREQ	(0.011)					(0.003)
	0.013					*0.004*
PASTAFE	(0.004)					(0.003)
	0.006					*0.004*
Sector dummies	yes	yes	yes	yes	yes	yes
Analyst dummies	yes	yes	yes	yes	yes	yes
Firm dummies	yes	yes	yes	yes	yes	yes
Nb. obs.	65 586	85 398	94 026	79 778	102 875	69 336

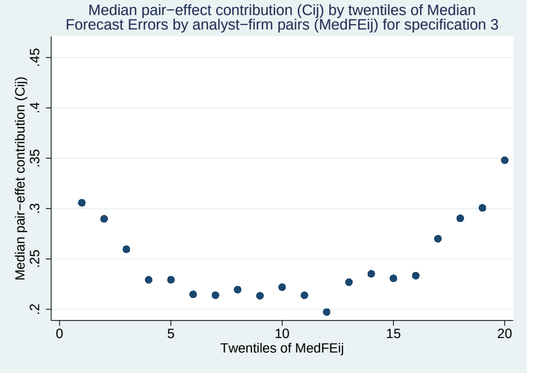
Graph 2: $C_{i,j}$ by 'twentiles' of $MedFE_{i,j}$ for specification [2]

$C_{i,j}$, the 'contribution of the pair-specific effect', is calculated as: $C_{i,j} = \frac{|\eta_{i,j}+\alpha|}{|\lambda_i|+|\mu_j|+|\eta_{i,j}+\alpha|}$, where α is the unconditional constant of specification [2], λ_i the firm effect, μ_j the analyst effect, and $\eta_{i,j}$ the firm-analyst' pair effect.



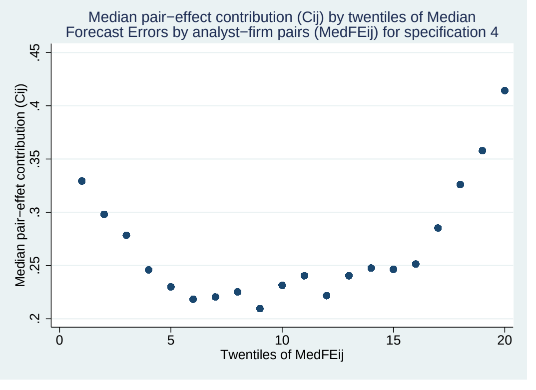
Graph 3: $C_{i,j}$ by 'twentiles' of $MedFE_{i,j}$ for specification [3]

$C_{i,j}$, the 'contribution of the pair-specific effect', is calculated as: $C_{i,j} = \frac{|\eta_{i,j}+\alpha|}{|\lambda_i|+|\mu_j|+|\eta_{i,j}+\alpha|}$, where α is the unconditional constant of specification [3] reported in Table 5, λ_i the firm effect, μ_j the analyst effect, and $\eta_{i,j}$ the firm-analyst' pair effect.



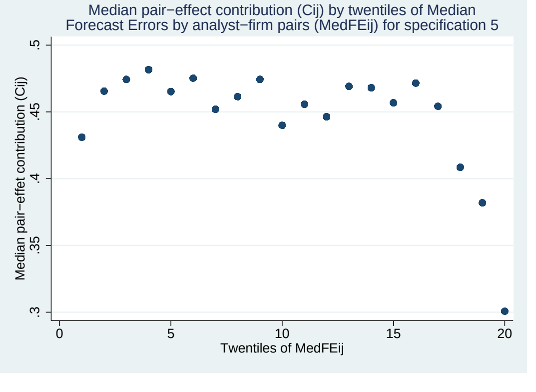
Graph 4: $C_{i,j}$ by 'twentiles' of $MedFE_{i,j}$ for specification [4]

$C_{i,j}$, the 'contribution of the pair-specific effect', is calculated as: $C_{i,j} = \frac{|\eta_{i,j}+\alpha|}{|\lambda_i|+|\mu_j|+|\eta_{i,j}+\alpha|}$, where α is the unconditional constant of specification [4] reported in Table 5, λ_i the firm effect, μ_j the analyst effect, and $\eta_{i,j}$ the firm-analyst' pair effect.



Graph 5: $C_{i,j}$ by 'twentiles' of $MedFE_{i,j}$ for specification [5]

$C_{i,j}$, the 'contribution of the pair-specific effect', is calculated as: $C_{i,j} = \frac{|\eta_{i,j}+\alpha|}{|\lambda_i|+|\mu_j|+|\eta_{i,j}+\alpha|}$, where α is the unconditional constant of specification [5] reported in Table 5, λ_i the firm effect, μ_j the analyst effect, and $\eta_{i,j}$ the firm-analyst' pair effect.



Graph 6: $C_{i,j}$ by 'twentiles' of $MedFE_{i,j}$ for specification [6]

$C_{i,j}$, the 'contribution of the pair-specific effect', is calculated as: $C_{i,j} = \frac{|\eta_{i,j}+\alpha|}{|\lambda_i|+|\mu_j|+|\eta_{i,j}+\alpha|}$, where α is the unconditional constant of specification [6] reported in Table 5, λ_i the firm effect, μ_j the analyst effect, and $\eta_{i,j}$ the firm-analyst' pair effect.

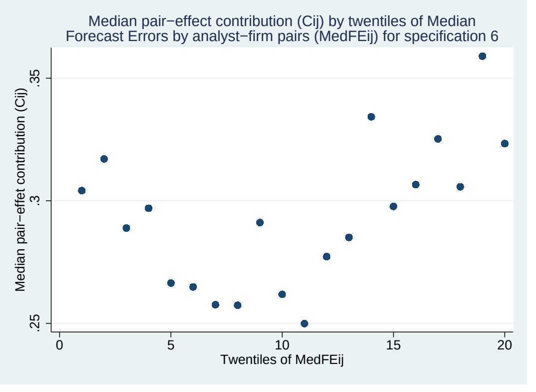


Table 7: Robustness tests for model [1] and variants [2] to [6]

The median-test investigates whether the median of $C_{i,j}$ in a given 'twentile' equals the median of the full sample. The Bartlett-test tests for the equality of variance of $C_{i,j}$ across quantiles. The Krusal-Wallis is an equality-of-population rank test, which determines whether the rank sum of each observation ranked by $C_{i,j}$ differs across quantiles. * and ** denote significance at the 5% and 1% levels, respectively.

Median test (Chi^2 stat)					
(1)	(2)	(3)	(4)	(5)	(6)
233**	1 100**	244**	369**	495**	81**

The null is the equality of the median of $C_{i,j}$ in each 'twentile' to the median of the whole sample.

Bartlett test (Chi^2 stat)					
(1)	(2)	(3)	(4)	(5)	(6)
256**	1 600**	334**	377**	114**	82**

The null is the equality of variances of $C_{i,j}$ across 'twentiles'.

Krusal-Wallis test (Chi^2 stat)					
(1)	(2)	(3)	(4)	(5)	(6)
341**	1693**	314**	524**	524**	99**

The null is the equality of the rank-sum of each observation ranked by $C_{i,j}$ across 'twentiles'.