A RE-EXAMINATION OF ANALYSTS' DIFFERENTIAL TARGET PRICE FORECASTING ABILITY

Patrice Fontaine et Tristan Roger

Presses universitaires de Grenoble | « Finance »

2020/1 Vol. 41 | pages 53 à 95

ISSN 0752-6180
ISBN 9782706147364

A re-examination of analysts' differential target price forecasting ability

Patrice Fontaine[1], **Tristan Roger**[2]

Abstract

We challenge the view that persistent differences in accuracy across analysts are proof that analysts differ in their ability to forecast stock prices. We show that these persistent differences in target price accuracy are driven instead by stock return volatility. Building upon option pricing theory, we construct a measure of forecast quality that controls for stock return volatility and forecast horizon. Contrary to previous studies, which failed to properly account for differences in stock return volatility, our empirical analysis reveals that analysts do not exhibit differences in their ability to forecast stock prices. We show that the accuracy of a target price strongly depends on the stock return volatility and the forecast horizon.

Keywords: Financial analysts, Target prices, Forecasting abilities, Expected accuracy, Persistence of volatility

Professional investors [...] fail a basic test of skill: persistent achievement. Daniel Kahneman (2011)

Financial analysts' skill in forecasting stock prices has become the focus of significant research interest in recent years. Data on their target prices is newly available, setting target prices is rising in popularity among analysts, and investors seem to care about this type of information. Brav and Lehavy (2003), for example, find significant abnormal returns following revisions to a target price, both unconditionally and conditional on contemporaneously issued recommendations and revisions to earnings forecasts. These findings

1 150 rue de la Chimie, 38040 Grenoble Cedex 9, France.
Email addresses: patrice.fontaine@eurofidai.org (Patrice Fontaine), tristan.roger@dauphine.fr (Tristan Roger)
CNRS (Eurofidai), Léonard de Vinci pôle universitaire, Research Center.

2 Place du Maréchal de Lattre de Tassigny, 75775 Paris Cedex 16, France.
DRM Finance, Université Paris-Dauphine, PSL Research University.
This article was written in part when Tristan Roger was a Visiting Scholar at the Haas School of Business, University of California, Berkeley. We thank Carole Bernard (the editor) and three anonymous referees. The article benefited from comments by Francesco D'Acunto, Brian Ayash, Mark Bradshaw, Otto Clemens, Patricia Dechow, François Degeorge, François Derrien, Michel Dubois, Ivan Guidotti, Isaac Hacamo, Sonia Jimenez-Garces, Marie-Aude Laguna, Alastair Lawrence, Martin Lettau, Roni Michaely, Franck Moraux, Terrance Odean, Richard Sloan, Boris Vallée, and Michael Weber, as well as seminar participants at UC Berkeley, Université Paris-Dauphine, Neuchâtel University, Orléans University, HEC Lausanne, and ESCP Europe Business School.

(confirmed by Asquith et al., 2005) provide evidence that investors believe target prices to be informative.

Another stream of literature examines financial analysts' skill in forecasting stock prices. Bradshaw et al. (2013) and Bilinski et al. (2013) find some analysts to be persistently more accurate than others when issuing target prices, and conclude that analysts exhibit differences in ability. However, Bradshaw et al. (2013) also find that markets do not react more strongly to target price revisions issued by analysts known for their accuracy. The authors put forward two possible explanations for this puzzling result. First, analysts' differences in skill are too weak economically to generate market reactions. Second, these differences are economically meaningful, but markets fail to recognize them.

In this paper, we investigate an alternative explanation. We argue that the persistent differences in analysts' target price accuracy found in previous studies (Bradshaw et al., 2013; Bilinski et al., 2013) result from a failure to properly control for differences in volatility among the stocks covered by the analysts.

Analysts tend to issue target prices for a small pool of stocks. This pool is relatively stable over time, meaning that analysts cover the same stocks for several periods. Some analysts cover stocks with low volatility, while others cover more volatile stocks. Analysts covering low-volatility stocks face an easier task, as these stocks are easier to forecast; such analysts are thus more accurate than their peers who cover high-volatility stocks. These differences in coverage persist over time. Thus, some analysts appear to be persistently more accurate than their peers. However, these persistent differences in accuracy do not necessarily mean that analysts possess differential abilities when it comes to forecasting stock prices. Differences in stock price volatility, instead, most likely drive the differences in accuracy across analysts.

To assess the quality of a forecast, we need to take into account both its accuracy and the general uncertainty surrounding the process. In this paper, we propose a measure of forecast quality that takes into account the difficulty of issuing an accurate forecast, which is a function of the stock return volatility and the forecast horizon. We provide evidence that the persistent differences in accuracy across analysts, as found in previous studies (Bradshaw et al., 2013; Bilinski et al., 2013), result from the stocks they cover having different levels of volatility. We demonstrate that analysts covering low volatility stocks do tend to be more accurate than those covering high volatility stocks. Additionally, we show that these volatility-induced differences in accuracy are persistent, because financial analysts tend to cover the same firms from one period to another. Using our new measure of forecast quality, we show that financial analysts do not exhibit differential abilities to forecast stock prices.

We incorporate the difficulty of issuing an accurate forecast into our measure by estimating, on the target price issue date, the accuracy that is to be expected if the target price does not contain any information (i.e., if it was randomly issued). We define our measure of target price forecast quality (*TPFQ*) as the difference between this expected accuracy and the realized accuracy. The critical issue here is to properly estimate the expected accuracy. We can show, however, that it is similar to calculating the value of a portfolio of options. That is, when target price accuracy is defined as the absolute forecast error, estimating the expected accuracy is the same as calculating the price of a straddle (a portfolio containing a call option and a put option) with a strike price equal to the target price and a maturity corresponding to the target price horizon.

Our empirical analyses show that persistent differences in accuracy cannot be interpreted as differences in analysts' skill in forecasting stock prices. Our approach consists of replacing the actual target prices in our sample with naive forecasts (i.e., target prices that are issued following a mechanical rule). Using naive forecasts implies, by definition, an absence of differential forecasting abilities across analysts. We keep, however, the structure of analyst coverage. Our results using naive forecasts indicate the existence of persistent differences in accuracy, providing direct proof that such differences are due to the structure of analyst coverage. Our analysis further suggests that the results found in previous studies are driven by stock return volatility.

Using our measure of target price forecast quality we show that, although the quality of the information contained in target prices varies across analysts, these differences are not persistent. This result indicates that analysts do not possess differential abilities to forecast stock prices. Because analysts frequently specialize in one or two industries (Boni and Womack, 2006; Kadan et al., 2012), we also test, for each industry, whether some analysts are persistently better than their peers. They are not. Finally, we look at whether persistent differences in forecast quality can be observed at the brokerage house level. The rationale for such a test is that some brokerage houses (i.e., the largest ones) possess superior resources, have better access to information, and can offer better compensation packages to attract the best analysts (Mikhail et al., 1997; Clement, 1999). These characteristics could translate into higher quality forecasts. Our results indicate, however, that persistent differences in forecast quality across brokerage houses do not exist.

The bottom line is that, while target prices are informative, financial analysts do not exhibit differences in their ability to forecast stock prices. Our empirical analyses show that target prices do contain information (i.e., our measure of information quality is positive, on average), which is consistent with earlier findings that market participants react to target price revisions (Brav and Lehavy, 2003; Asquith et al., 2005). However, and contrary to studies of earnings forecasts (Sinha et al., 1997; Park and Stice, 2000), financial analysts do not appear to possess differential abilities to forecast stock prices.

Our paper contributes to the emerging literature on target prices and, more generally, to that on financial analysts. The methodology we have devised to evaluate the difficulty of issuing an accurate forecast is not limited to target prices and may be extended to assess the quality of any kind of forecast (e.g., exchange rates) if estimating the distribution of the underlying stochastic process is possible. Our measure is an improvement over traditional *ex-post* measures, as it can be used in a dynamic setting; that is, we are able to evaluate the quality of a forecast at any moment in time. Finally, we answer the important question of whether analysts' differential abilities to forecast earnings translate into differential abilities to forecast stock prices. Our analysis provides significant evidence that this is not the case.

1. Data and descriptive statistics

We obtain target prices from the I/B/E/S unadjusted detail history target price dataset.[3] Stock prices, returns, and adjustment factors (splits and corporate actions) come from CRSP. Our initial sample consists of 892,922 target prices issued on U.S. companies between 2000 and 2012. For each forecast, we collect the code of the analyst who issues the forecast (and the

3 We use the unadjusted dataset to avoid retroactive stock split rounding effects (Baber and Kang, 2002; Payne and Thomas, 2003).

broker code), the issue date, the horizon (usually 6 or 12 months), and the value of the target price. We remove from our sample forecasts for which the stock price is not available around the dates of issue or of the end of the horizon, and forecasts for which the price history is too short to estimate historical volatility. We then restrict our sample to 12-month-ahead target prices.[4] We also delete forecasts for which the ratio of the target price over the stock price is in the bottom or top 1% of the distribution. Finally, we remove target prices that were likely issued by teams of analysts, rather than by a single individual.[5] Our final sample is composed of 683,995 target prices issued by 9,245 analysts (707 brokers) on 6,955 U.S. stocks.

Table 1 reports for each year the number of forecasts, analysts, and firms; the average, median, and maximum number of analysts per stock; the average, median, and maximum number of stocks covered per analyst; and the average absolute forecast error of the target prices. We observe that the number of forecasts per year triples over our sample period, while the number of analysts remains roughly constant. This trend indicates that including target prices in reports is an increasingly popular practice among financial analysts. The number of different stocks each analyst typically covers increases from 8 at the beginning of our sample period to 13 in the final years. At the same time, the number of analysts covering a given stock increases from 6 to 12. Finally, in the last column of Table 1, we observe that the average absolute forecast error is greatest during the Dotcom crisis and the 2008 financial crisis.

On average, the analysts in our sample revise their forecasts approximately every 6 months (117 trading days). Their target prices are on average 20% higher than the concurrent stock price. This statistic is similar to what can be observed for other periods and/or countries. For instance, Brav and Lehavy (2003) find that target prices on U.S. stocks in 1997-99 are on average 28% higher than the current price, while Kerl (2011) reports an implicit return of 18.07% for German stocks in 2002-2004. Finally, it appears that the analysts in our sample are mostly optimistic about future stock prices, with only 8.7% of the target prices issued below the concurrent price.

4 Previous studies focus on 12-month-ahead target prices. We similarly restrict our sample to facilitate comparisons with their results.

5 We remove analysts whose names contain an ampersand (&), a slash (/), the word DEPARTMENT (or DEPT), the word RESEARCH, the word AND, and the word GROUP.

Table 1. Descriptive statistics

The sample consists of a total of 683,995 target prices issued by 9,245 analysts (707 brokers) on 6,955 U.S. stocks for the 2000-2012 period. The first column indicates the number of target prices issued each year. The second column shows the number of active analysts. The third column provides the number of firms in the sample. The six following columns report the average, median, and maximum number of active analysts per stock and the average, median, and maximum number of stocks covered per analyst. The last column provides the average absolute forecast error (*AFE*) of the target prices issued during the corresponding year.

	Number of forecasts	Number of analysts	Number of firms	Number of analysts covering a stock			Number of stocks covered per analyst			Average AFE
				Mean	Median	Max	Mean	Median	Max	
2000	24,811	2,594	3,096	7.35	6	25	9.48	8	59	0.5553
2001	31,880	3,067	3,125	8.76	8	33	9.62	9	59	0.5314
2002	38,212	2,919	3,203	9.14	8	32	10.76	10	54	0.4546
2003	42,595	2,604	3,318	8.81	8	32	11.43	10	58	0.3600
2004	46,732	2,771	3,424	9.22	8	31	11.47	10	58	0.3115
2005	47,589	2,824	3,615	8.95	8	32	11.98	11	68	0.2884
2006	50,365	2,830	3,696	9.11	8	32	12.27	11	66	0.2909
2007	53,434	2,819	3,608	9.36	8	28	12.72	11	63	0.4677
2008	65,658	2,828	3,607	9.60	9	32	12.88	12	74	0.5642
2009	65,843	2,772	3,420	10.65	10	32	13.31	12	75	0.4410
2010	68,192	3,152	3,358	12.46	12	45	13.16	12	63	0.3139
2011	75,632	3,250	3,348	13.11	12	48	13.43	13	66	0.3305
2012	73,052	3,191	3,288	13.11	12	45	13.51	13	65	0.3177

2. Target price accuracy and information content

In this section, we show that the absolute forecast error and other such measures of target price accuracy are not good proxies for the quality of the information contained in target prices or for measuring analysts' performance. We provide both empirical and theoretical evidence that target price accuracy is impacted by two factors: (1) stock return volatility, and (2) forecast horizon.[6] We show that these two factors influence attempts to measure the ability of financial analysts to forecast stock prices and can bias economic findings. Our analysis focuses on the absolute forecast error, as it is the most popular measure of forecast accuracy. We note, however, that these two factors also impact other measures of accuracy.[7] We define the absolute forecast error (*AFE*) as:

$$AFE_t = \frac{\left|S_T - TP_{t,T}\right|}{S_t}, \tag{1}$$

where $TP_{t,T}$ is the value of a target price issued at time t with horizon T, S_T is the stock price at the end of the forecast horizon, and S_t is the stock price at the time the target price was issued.

2.1. Target price accuracy and stock return volatility

To examine whether stock return volatility impacts target price accuracy, for each year we assign target prices to five quintiles with respect to the volatility of the underlying stock. Table 2 provides the average AFE of the target prices, per quintile, for 2000-2012. Panel A reports the results using actual target prices. Panels B and C provide the average AFE using two types of naive forecasts. We use naive forecasts to eliminate the possibility that the relationship between AFE and stock return volatility ensues entirely from financial analysts being particularly good at forecasting stock prices for low volatility firms. In Panel B, we use "naive price forecasts". Naive price forecasts are built so that the implied stock return $\left(TP_{t,T} - S_t\right) / S_t$ of a 12-month horizon target price is equal to the 12-month risk free rate.

6 A negative relationship exists between target price accuracy and forecast horizon. We do not report our results on this relationship, as most studies consider only target prices with a 12-month horizon.

7 Although the absolute forecast error is the most popular measure of accuracy, previous studies have also considered: (1) whether the actual closing price, at the end of the 12-month forecast horizon, is at or above the target price (TPMETEND); (2) whether the target price is met at any time during the 12-month forecast horizon (TPMETANY). Our conclusions remain unchanged if we consider these alternative measures of accuracy. We provide empirical results in Section 5.

Table 2. Relationship between volatility and absolute forecast error (*AFE*)

Each year, the target prices in the sample are assigned to five quintiles with respect to the six-month historical volatility of the underlying stock estimated at the issue date. The table reports, for each year and each quintile, the average target price accuracy (average absolute forecast errors). Panel A reports the average absolute forecast errors using the actual target prices from our sample. Panel B reports the average absolute forecast errors using naive price forecasts (each target price is replaced by a forecast that is equal to the price at the issue date, capitalized by the risk-free rate). Panel C reports the average absolute forecast errors using naive return forecasts (each target price is replaced by a forecast that is equal to the concurrent stock price, times one plus the previous 12-month stock return). The statistical significance of the difference across top and bottom quintiles is computed using a t-test. ***/**/* represent significance at the 0.01, 0.05, and 0.1 level.

Panel A: Absolute forecast errors computed using actual target prices													
Volatility quintiles	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
1 (Low volatility)	0.2520	0.2285	0.2113	0.2093	0.1891	0.1595	0.1651	0.3223	0.4183	0.2635	0.1875	0.1776	0.1837
2	0.3691	0.3250	0.3095	0.2554	0.2404	0.2345	0.2348	0.4110	0.5096	0.3798	0.2802	0.2531	0.2393
3	0.5264	0.4897	0.4193	0.3261	0.3175	0.2894	0.2966	0.4784	0.5697	0.4302	0.3497	0.3458	0.3041
4	0.7387	0.7036	0.5995	0.4714	0.4040	0.3742	0.3847	0.5667	0.6413	0.5028	0.4131	0.4640	0.4031
5 (High volatility)	1.0585	0.9883	0.8521	0.6724	0.5656	0.5454	0.5062	0.7141	0.7633	0.7648	0.4626	0.6176	0.6172
Diff (5-1)	0.8064	0.7598	0.6407	0.4631	0.3765	0.3859	0.3411	0.3918	0.3450	0.5013	0.2751	0.4400	0.4335
Mean t-test	90.9300***	110.4600***	57.0200***	41.2200***	82.1600***	73.1800***	53.7300***	77.3400***	56.7700***	48.0400***	75.2500***	121.7600***	90.3300***

Panel B: Absolute forecast errors computed using naive price forecasts													
Volatility quintiles	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
1 (Low volatility)	0.2333	0.1815	0.1790	0.2421	0.1846	0.1472	0.1617	0.2807	0.3412	0.2968	0.2036	0.1774	0.2189
2	0.3158	0.2367	0.2622	0.3043	0.2526	0.2241	0.2258	0.3493	0.4204	0.4192	0.2770	0.2024	0.2754
3	0.4076	0.3338	0.3451	0.3644	0.3054	0.2723	0.2643	0.4015	0.4742	0.4662	0.3325	0.2611	0.3189
4	0.5046	0.4546	0.4880	0.4986	0.3362	0.3537	0.3368	0.4688	0.5326	0.5350	0.3777	0.3390	0.3851
5 (High volatility)	0.6720	0.5952	0.7190	0.6659	0.3901	0.4643	0.4104	0.5265	0.6191	0.8158	0.3820	0.4038	0.5240
Diff (5-1)	0.4387	0.4137	0.5401	0.4238	0.2054	0.3172	0.2487	0.2458	0.2779	0.5190	0.1784	0.2263	0.3051
Mean t-test	70.2000***	92.7700***	45.8200***	35.4300***	49.3600***	58.5500***	40.2800***	62.200***	46.7400***	46.9800***	50.7500***	77.5200***	61.3700***

Panel C: Absolute forecast errors computed using naive return forecasts													
Volatility quintiles	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
1 (Low volatility)	0.3766	0.2909	0.2724	0.2631	0.2166	0.1946	0.2143	0.3499	0.4432	0.4487	0.2623	0.2334	0.2287
2	0.5228	0.4015	0.3978	0.3937	0.3530	0.3002	0.3243	0.4307	0.5447	0.6466	0.4317	0.3318	0.3095
3	0.8622	0.5622	0.5323	0.5667	0.5227	0.4501	0.4077	0.5063	0.6362	0.7627	0.5978	0.4594	0.4016
4	1.6078	0.6414	0.7977	0.9231	0.7776	0.5501	0.5734	0.6359	0.7243	0.8963	0.8268	0.5753	0.5090
5 (High volatility)	2.8843	0.4781	1.0175	1.6514	1.3814	0.8115	0.8437	0.8223	0.8786	1.3646	1.2567	0.7816	0.7452
Diff (5-1)	2.5078	0.1872	0.7450	1.3883	1.1648	0.6169	0.6295	0.4725	0.4354	0.9159	0.9944	0.5482	0.5165
Mean t-test	49.1900***	21.4700***	48.4000***	59.4300***	66.6100***	61.1200***	65.4600***	57.3300***	49.4600***	64.2300***	73.2900***	87.7500***	79.7200***

In Panel C, we use "naive return forecasts", defined by Bilinski et al. (2013) as the concurrent stock price times one plus the previous 12-month stock return. Our findings indicate a strong monotonic relationship between *AFE* and stock return volatility, both for actual target prices and for naive forecasts.

In multivariate analyses, the usual way to account for a dependence between two variables is to incorporate a control variable. However, we show that stock return volatility cannot be used as a control variable in our context, as the relationship between it and AFE is highly nonlinear. To demonstrate this nonlinearity, we perform the following regression:

$$AFE_{j,t} = \alpha + \sum_{k=1}^{10} \beta_k 1_{j,t}^k \sigma_{j,t} + \varepsilon_{j,t}, \tag{2}$$

where $AFE_{j,t}$ is the absolute forecast error of a target price on firm j issued at time t, $\sigma_{j,t}$ is the six-month historical stock return volatility of stock j measured at time t, and $1_{j,t}^k$ is an indicator variable that takes a value of 1 if the stock return volatility $\sigma_{j,t}$ belongs to the k-th volatility decile, and 0 otherwise.

In the case of linearity, all the coefficients β_k take the same value. If the relationship between AFE and volatility is nonlinear, however, the coefficients β_k will take different values. If the coefficient β_k decreases (increases) with k, the relationship between volatility and AFE is concave (convex). Table 3 reports the results of this regression. The coefficient β_k is shown to increase with k, indicating that the relationship between AFE and volatility is nonlinear and convex. In Appendix A, we provide theoretical evidence of the nonlinearity of the relationship between AFE and volatility.

This nonlinearity rules out the use of simple controls, such as including volatility in multivariate regressions,[8] and is one reason why relative measures of accuracy are unsuitable when evaluating analysts' performance. (See Appendix B for a more detailed discussion on the problems in using relative measures of accuracy). As a result of this nonlinearity, we introduce a new target price quality measure.

8 An alternative would be to use nonlinear controls, such as volatility decile dummies. However, additional analyses show that standard nonlinear controls do not fully eliminate the link between volatility and accuracy. Our results are available upon demand.

Table 3. Nonlinear relationship between volatility and absolute forecast errors (*AFE*)

This table shows the coefficient estimates from the following OLS regression:

$$AFE_{j,t} = \alpha + \sum_{k=1}^{10} \beta_k 1_{j,t}^k \sigma_{j,t} + \varepsilon_{j,t},$$

where $AFE_{j,t}$ is the absolute forecast error of a target price on firm *j* issued by any analyst at time *t*, $\sigma_{j,t}$ is the stock return volatility of stock *j* measured at time *t*, and $1_{j,t}^k$ is an indicator that takes the value 1 if the stock return volatility $\sigma_{j,t}$ belongs to the *k*-th volatility decile, and 0 otherwise. We denote the variable $1_{j,t}^k \sigma_{j,t}$ as the *k*-th decile. ***/**/* represent significance at the 0.01, 0.05, and 0.1 level. P-values are computed using robust standard errors.

	Absolute forecast errors			
k-th decile = $1_{j,t}^k \sigma_{j,t}$	Coefficient (β_k)	Standard error	t-statistic	p-value
1st decile (Low volatility)	0.2413***	0.0112	21.62	0.00
2nd decile	0.3717***	0.0083	44.60	0.00
3rd decile	0.4056***	0.0071	57.43	0.00
4th decile	0.4266***	0.0062	68.92	0.00
5th decile	0.4553***	0.0055	83.02	0.00
6th decile	0.4842***	0.0049	99.36	0.00
7th decile	0.5122***	0.0043	118.98	0.00
8th decile	0.5426***	0.0038	144.79	0.00
9th decile	0.5669***	0.0032	179.48	0.00
10th decile (High volatility)	0.5926***	0.0022	269.07	0.00
Number of observations	683,995			
R-squared	0.1525			

2.2. Persistent differences in accuracy

In this subsection, we show that the persistent differences in absolute forecast error found in previous studies are driven by factors other than information and cannot be interpreted as evidence of analysts exhibiting differential forecasting abilities. To show that the differences seen are instead driven by persistent differences in the firms covered by the analysts, we again make use of naive forecasts. We expect the differences in absolute forecast error to be mainly driven by persistent differences in volatility.

We define an analyst's performance as in Bradshaw et al. (2013) and Bilinski et al. (2013). For a given period $]t-1;t]$, we evaluate the analyst's performance as the average of the AFE of the target prices she issued during that period. As we are also interested in the volatility of the stocks she covered, for each target price she issued during the $]t-1;t]$ period we calculate the historical stock return volatility for the six months preceding the target price issue date. We then take the average of the volatilities of all stocks for which she issued target prices during that period.

Following Bradshaw *et al.* (2013), we assign analysts to five quintiles with respect to their performance over the measurement period, $]t-1;t]$. We next measure persistence by estimating their performance over a test period, conditional on their ranking over the measurement period. The test period is defined as $]t+\theta;t+\theta+1]$, where $\theta = 12$ months. A lag $\theta = 12$ months is added between the measurement period and the test period to ensure that the two periods do not overlap.[9] We therefore avoid mechanically inducing a positive relation between current and subsequent analysts' AFE. We observe persistent differences in accuracy if the most (least) accurate analysts over the measurement period are ranked in the best (worst) performance quintile for the test period and if the difference between the AFE of the first quintile (best) and that of the fifth quintile (worst) is statistically different from zero.

Our results are described in Table 4. Panel A uses the actual target prices in our sample. As our database is similar to that of Bradshaw et al. (2013), we obtain similar results. Analysts with a better average AFE in the measurement period also have a better average AFE in the test period. Analysts in the best (worst) quintile exhibit an average AFE of 0.1418 (0.8318) in the measurement period and an average AFE of 0.3087 (0.5131) in the test period. However, the stock return volatility is lowest for target prices issued by the analysts in the first quintile, and this volatility increases from performance quintile 1 to quintile 5 , both in the measurement period and in the test period. Thus, when we observe persistent differences in *AFE*, we also observe persistent differences in volatility.

9 If the test period is $]t;t+1]$, then the accuracy of the target prices issued at the end of the measurement period and the accuracy of those issued at the beginning of the test period would be artificially correlated. Indeed, if we consider two target prices issued at time $t-\varepsilon$ (in the measurement period) and time $t+\varepsilon$ (in the test period), there would be an overlap in the period $]\,t+\varepsilon;t-\varepsilon+0]$. This autocorrelation would cause an artificial persistence.

Table 4. Persistent differences in absolute forecast errors (*AFE*)

This table reports the analysts' performance in the test period $]t+\theta;\, t+\theta+1]$, conditional on their performance in the measurement period $]t-1;\, t]$. θ is a 12-month lag which ensures that the measurement period and the test period do not overlap. An analyst's performance is measured, for a given period, as the average of the absolute forecast errors (*AFE*) on the target prices she issued during that period. The measurement period is quarterly. Panel A corresponds to the actual target prices in the sample. Panel B (Panel C) corresponds to naive price forecasts (naive return forecasts). We rank analysts in quintiles based on their performance in the measurement period, and we compute the corresponding performance in the test period. We also report the volatility of the stocks covered, computed as the average of the six-month historical stock return volatilities associated with the target prices issued by the analyst during the period under consideration. Conditional on the ranking made during the measurement period $]t-1;\, t]$, we report, for the test period $]t+\theta;\, t+\theta+1]$, both the analysts' performance and the volatility of the stocks covered. We test the difference of means across the top and bottom quintiles using a t-test. ***/**/* represent significance at the 0.01, 0.05, and 0.1 level.

	Panel A: Actual target prices				
		Measurement period $]t-1;t]$		Test period $]t+\theta;t+\theta+1]$	
Performance quintile (measurement period)	Number of observations	Analysts' performance (average *AFE*)	Volatility of stocks covered	Analysts' performance (average *AFE*)	Volatility of stocks covered
1 (Best)	12,768	0.1418	0.3673	0.3087	0.3538
2	12,798	0.2570	0.4094	0.3379	0.3910
3	12,802	0.3580	0.4551	0.3760	0.4341
4	12,798	0.4914	0.5139	0.4311	0.4811
5 (Worst)	12,779	0.8318	0.6125	0.5131	0.5480
Diff (5-1)		0.6901	0.2452	0.2044	0.1942
Mean t-test		18.57***	6.87***	7.90***	6.55***

		Panel B: Naive price forecasts			
		Measurement period $]t-1;t]$		Test period $]t+\theta;t+\theta+1]$	
Performance quintile (measurement period)	Number of observations	Analysts' performance (average *AFE*)	Volatility of stocks covered	Analysts' performance (average *AFE*)	Volatility of stocks covered
1 (Best)	12,768	0.1231	0.3832	0.2935	0.3678
2	12,798	0.2252	0.4182	0.3108	0.3954
3	12,802	0.3097	0.4618	0.3348	0.4343
4	12,798	0.4178	0.5075	0.3671	0.4742
5 (Worst)	12,779	0.7113	0.5875	0.4145	0.5364
Diff (5-1)		0.5883	0.2044	0.1210	0.1686
Mean t-test		15.28 ***	5.70 ***	4.42 ***	5.45 ***

		Panel C: Naive return forecasts			
		Measurement period $]t-1;t]$		Test period $]t+\theta;t+\theta+1]$	
Performance quintile (measurement period)	Number of observations	Analysts' performance (average *AFE*)	Volatility of stocks covered	Analysts' performance (average *AFE*)	Volatility of stocks covered
1 (Best)	12,759	0.1808	0.3851	0.4265	0.3800
2	12,790	0.3389	0.4171	0.4707	0.4034
3	12,788	0.4790	0.4548	0.5161	0.4317
4	12,790	0.6806	0.5018	0.5662	0.4637
5 (Worst)	12,769	1.3648	0.5982	0.6706	0.5281
Diff (5-1)		1.1840	0.2131	0.2441	0.1480
Mean t-test		11.25 ***	5.88 ***	5.44 ***	4.56***

Panels B and C of Table 4 provide the results of our analyses using naive forecasts. Even though these target prices are issued following a mechanical rule, we continue to find evidence of persistent differences in accuracy. For naive price forecasts, the average AFE ranges from 0.2935 to 0.4145 in the test period; this difference between the first and fifth quintiles is highly statistically significant. The difference in AFE between quintiles one and five for naive return forecasts in the test period is equal to 0.2441; it is highly statistically significant as well. These results indicate that accuracy is unsuited to evaluating the ability of financial analysts to forecast stock prices; the persistent differences in accuracy are most likely driven by persistent differences in volatility.

3. Target price forecast quality

3.1. Ex-post measure of target price forecast quality

As shown in the previous section, the absolute forecast error strongly depends on stock return volatility. It follows that, when assessing an analyst's ability to forecast stock prices, we must take the volatility into account. The impact of stock return volatility (and of the forecast horizon) can be captured by estimating the expected absolute forecast error as of the issue date. If we state AFE simply as $\left|S_T - TP_{t,T}\right|$ (that is, we set $S_t = 1$ and adjust S_T and $TP_{t,T}$ accordingly), then the expectation of AFE, stated as $E_t\left[\left|S_T - TP_{t,T}\right|\right]$, corresponds to the forecast difficulty associated with the issued target price. This expected value of AFE can be seen as the accuracy that is likely to be achieved if the target price does not contain any information. We therefore define target price forecast quality as the difference between the forecast difficulty, $E_t\left[\left|S_T - TP_{t,T}\right|\right]$, and the forecast accuracy, $\left|S_T - TP_{t,T}\right|$. If the target price accuracy is higher than the expected accuracy, then we can conclude that the analyst provided useful additional information to the market participants.

The important issue here is how to estimate the expected value, $E_t\left[\left|S_T - TP_{t,T}\right|\right]$. We can assume that stock prices follow a log-normal distribution, but the distribution of $\left|S_T - TP_{t,T}\right|$ is unknown. However, the problem of estimating such an expected value has already been studied (and solved) in the literature on option pricing . Indeed, we note that AFE

corresponds to the final payoff of a straddle with a strike price equal to $TP_{t,T}$, that is, a portfolio containing a call option and a put option on the same underlying stock; the two options are characterized by the same strike price and the same maturity. The price of the straddle at time t is equal to $e^{-r(T-t)}E_t\left[\left|S_T - TP_{t,T}\right|\right]$. It follows that we can estimate the expected value, $E_t\left[\left|S_T - TP_{t,T}\right|\right]$, the same way we would compute the capitalized straddle price. We define our *ex-post* measure of target price forecast quality ($TPFQ$) as follows.

Definition 1 *The ex-post forecast quality $TPFQ_{t,T}$ of a target price $TP_{t,T}$ issued at time t on a stock S, with an horizon equal to $T - t$, is defined as:*

$$\begin{aligned} TPFQ_{t,T} &= E_t\left[\left|S_T - TP_{t,T}\right|\right] - \left|S_T - TP_{t,T}\right| \\ &= (C_t + P_t)e^{r(T-t)} - (C_T + P_T), \end{aligned} \tag{3}$$

where C_t (P_t) is the price at time t of a call (put) option on the stock S with maturity date T and strike price $TP_{t,T}$.

As we need to be able to compare the forecast quality of target prices issued on stocks that have different price levels, we require our measure of forecast quality to be homogeneous of degree 1 (i.e., we do not want a stock's price level to influence our measure of forecast quality). We thus set the stock price equal to 1 at the time the target price is issued. We write $S_t = 1$, and adjust the target price, $TP_{t,T}$, and the stock price at time T, S_T , accordingly.

Assuming that the stock price follows a geometric Brownian motion, the *ex-post* forecast quality, $TPFQ_{t,T}$, of a target price issued at time t can be calculated according to the Black and Scholes (1973) model (see Appendix C).[10]

3.2. Properties

Our measure has two components: (1) the expected value of the AFE (which estimates the difficulty of issuing an accurate target price), and (2) the *ex-post* AFE (which is the traditional estimation of the accuracy of a target price). The difficulty of issuing an accurate target price is positively

10 We compute the value of the straddle using the Black and Scholes (1973) model for the sake of simplicity. However, our measure of target price forecast quality could be extended to more complex models of option pricing.

related to the stock return volatility, σ_t (e.g., it is more difficult to forecast the future price of a stock with high volatility than one with low volatility) and to the length of the forecast horizon, $T - t$ (e.g., issuing an accurate target price with a 24-month horizon is more difficult than with a 12-month horizon). Therefore, our measure of forecast difficulty, $E_t\left[\left|S_T - TP_{t,T}\right|\right]$, must satisfy the following two requirements: (1) it must increase with the stock return volatility σ_t, and (2) it must increase with the length of the forecast horizon $T - t$.

Proposition 1 *For a given final stock price S_T and a given target price $TP_{t,T}$, the forecast quality $TPFQ_{t,T}$ is an increasing function of the stock return volatility σ_t and of the length of the horizon $T - t$.*

We provide a proof for this proposition in Appendix E.

3.3. Target price forecast quality in a dynamic setting

We propose a dynamic setting in which the target price forecast quality can be estimated at any point in time. The $TPFQ$ at time $t + \tau$, $\tau \in 0;T - t]$ is simply equal to $E_t\left[\left|S_T - \Phi_{t,T}\right|\right] - E_{t+\tau}\left[\left|S_T - \Phi_{t,T}\right|\right]$. The estimation, at time $t + \tau$, of the expected value of AFE is equivalent to computing the price, at time $t + \tau$, of a straddle with strike price equal to $\Phi_{t,T}$ and maturity equal to $T - (t + \tau)$.

Definition 2 *The forecast quality $TPFQ_{t,t+\tau}$ at time $t + \tau$, $\tau \in [0;T - t]$ of a target price $TP_{t,T}$ issued at time t on a stock S, with a horizon equal to $T - t$, is defined as*

$$\begin{aligned} TPFQ_{t,t+\tau} &= E_t\left[\left|S_T - TP_{t,T}\right|\right] - E_{t+\tau}\left[\left|S_T - TP_{t,T}\right|\right] \\ &= (C_t + P_t)e^{r\tau} - (C_{t+\tau} + P_{t+\tau}), \end{aligned} \tag{4}$$

where C_t (P_t) is the price at time t of a call (put) option on the stock S with maturity date T and strike price $TP_{t,T}$.

When the target price is issued ($\tau = 0$), $TPFQ_{t,t}$ is equal to 0. This property expresses the idea that, at the time the target price is issued, one does not know yet whether it is a good or a bad forecast. Note that when $\tau = T - t$, we retrieve the *ex-post* measure of forecast quality defined in Equation 3.

Remark 1 *Consistent with the assumptions of the Black and Scholes (1973) model, the stock return volatility remains constant for a given forecast. That is, once the target price is set, we use the stock return volatility at the time the forecast was made to estimate* $E_{t+\tau}\left[\left\|S_T - TP_{t,T}\right\|\right]$. *We use this same volatility until a new target price (a revision) is issued. With this method, we can distinguish between target prices issued on stocks with different volatilities (the cross-section), while preventing the variations of volatility over time from influencing the target price forecast quality (the time-series).*

3.4. Target price revisions

In practice, financial analysts often revise their target prices before the end of the horizon. We consider the initial forecast and the revision to be two separate forecasts. Once a revision occurs at time $t+\tau$, the first forecast is no longer active. However, we need to evaluate the forecast quality of the initial target price over the period $]t;t+\tau]$. It follows from Definition 2 that the forecast quality of the initial target price at time $t+\tau$ is simply equal to $E_t\left[\left\|S_T - TP_{t,T}\right\|\right] - E_{t+\tau}\left[\left\|S_T - TP_{t,T}\right\|\right]$.

Let us consider the following example. An analyst issues a target price $TP^1_{t,T}$ at time t. She then revises her forecast, at time $t+\tau$, and issues a target price $TP^2_{t+\tau,T+\tau}$. The forecast quality over the period $]t;T+\tau]$ is then equal to

$$TPFQ_{t,T+\tau} = TPFQ_{t,t+\tau}\left(TP^1_{t,T},\sigma_t\right) + TPFQ_{t+\tau,T+\tau}(TP^2_{t+\tau,T+\tau},\sigma_{t+\tau}), \quad (5)$$

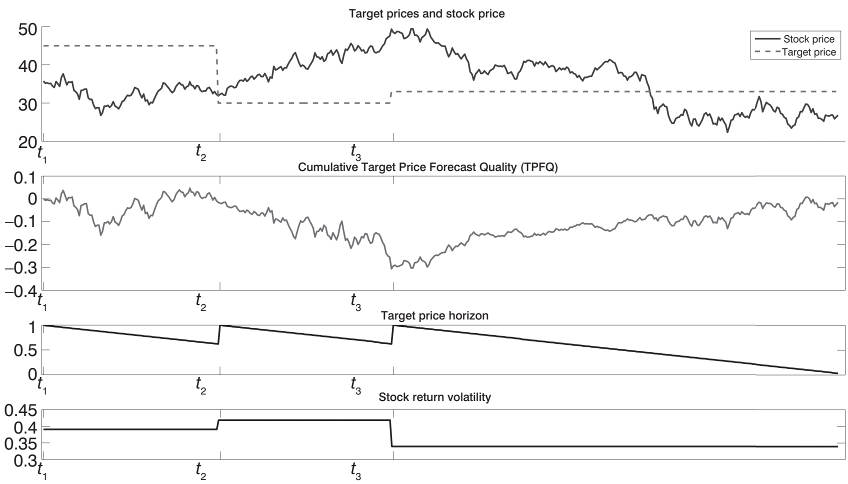
where $TPFQ_{t,t+\tau}$ is the forecast quality, estimated at time $t+\tau$ of the initial target price $TP_{t,T}$, issued at time t with a horizon equal to $T-t$, and $TPFQ_{t+\tau,T+\tau}$ is the forecast quality estimated at time $T+\tau$ of the revised target price $TP_{t+\tau,T+\tau}$, issued at time $t+\tau$ with a horizon equal to $(T+\tau)-(t+\tau)$.

3.5. TPFQ : An illustration

To provide a clearer understanding of how we compute forecast quality, we present in Figure 1 an example of three 12-month-ahead target prices made by a single analyst. In this illustration, the risk-free rate is equal to 0. The first target price TP_{t_1,t_1+1}, equal to \$45, is issued at time t_1 (when the actual price of the stock is equal to \$35.76). The second target price TP_{t_2,t_2+1} (first revision), equal to \$30, is issued at time t_2 (when the stock's actual price is \$31.88).

Finally, at time t_3, the analyst revises her forecast again and announces a target price TP_{t_3,t_3+1} of \$33 (when the stock price is equal to \$49.34).

Figure 1 An illustration of the TPFQ measure



At time t_1, we estimate the forecast difficulty for the first target price of \$45. This forecast difficulty, $E_{t_1}[|S_{t_1+1} - TP_{t_1,t_1+1}|]$, is equal to 0.4091 (with a six-month historical volatility of $\sigma_{t_1} = 0.3905$). Because t_2 comes before the end of the horizon of the first target price, we do not know the realized accuracy of the initial forecast. Therefore, we estimate the accuracy at time t_2 as $E_{t_2}[|S_{t_1+1} - TP_{t_1,t_1+1}|]$. This forecast accuracy is equal to 0.4249. The target price forecast quality is then obtained by taking the difference between the forecast difficulty and the forecast accuracy. For the period $]t_1;t_2]$, we have:

$$TPFQ_{t_1,t_2} = E_{t_1}[|S_{t_1+1} - TP_{t_1,t_1+1}|] - E_{t_2}[|S_{t_1+1} - TP_{t_1,t_1+1}|] = -0.0158. \quad (6)$$

Similarly, for the period $]t_2;t_3]$, we have:

$$TPFQ_{t_2,t_3} = E_{t_2}[|S_{t_2+1} - TP_{t_2,t_2+1}|] - E_{t_3}[|S_{t_2+1} - TP_{t_2,t_2+1}|] = -0.2901. \quad (7)$$

Note that these two expected values, $E_{t_2}[|S_{t_2+1} - TP_{t_2,t_2+1}|]$ and $E_{t_3}[|S_{t_2+1} - TP_{t_2,t_2+1}|]$, are computed with a volatility σ_{t_2} equal to 0.4181.

For the period $]t_3;t_3+1]$, no revision is made before the end of the target price horizon. Therefore, the target price forecast quality $TPFQ_{t_3,t_3+1}$ can be written as:

$$\begin{aligned} TPFQ_{t_3,t_3+1} &= E_{t_3}[|S_{t_3+1} - TP_{t_3,t_3+1}|] - \left|S_{t_3+1} - TP_{t_3,t_3+1}\right|. \\ &= 0.2905. \end{aligned} \tag{8}$$

Finally, the target price forecast quality $TPFQ_{t_1,t_3+1}$ over the whole period is equal to:

$$\begin{aligned} TPFQ_{t_1,t_3+1} &= TPFQ_{t_1,t_2} + TPFQ_{t_2,t_3} + TPFQ_{t_3,t_3+1}. \\ &= -0.0154. \end{aligned} \tag{9}$$

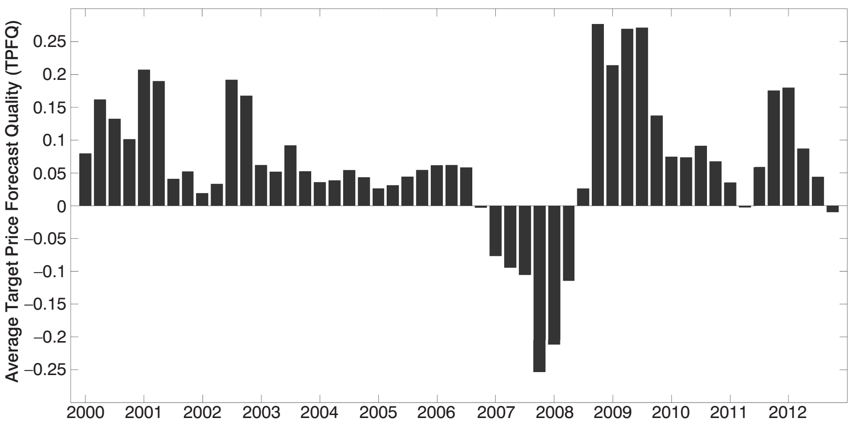
4. Analysts' ability to forecast stock prices

4.1. Information content of target prices

Our measure of target price forecast quality captures the information content of target prices. As shown in Figure 2, the forecast quality of target prices from 2000 to 2012 is, on average, positive. This result is consistent with the idea that target prices do contain information and thus supports earlier findings that market participants react to target price revisions (Brav and Lehavy, 2003; Asquith et al., 2005). That Figure 2 indicates a negative target price forecast quality for prices issued during the global financial crisis (in 2007 and during the first two quarters of 2008) suggests that financial analysts: (1) failed to anticipate the financial crisis, and (2) failed to adjust their target prices accordingly.

Our main concern in this article is to assess whether financial analysts exhibit genuine skill in forecasting stock prices. A positive value for target price forecast quality is not sufficient to prove the existence of such a skill. The positive value can be due to skill, to luck, or simply to the fact that, when analysts are overly optimistic the forecast quality increases when the market rises. One way to determine if forecasting skill exists is to check not only whether analysts differ in the quality of their forecasts, but if that difference persists. As stated by Kahneman (2011), "the diagnostic for the existence of any skill is the consistency of individual differences in achievement". We will thus consider an analyst to be skilled if she manages to consistently beat the other analysts.

Figure 2 The information content of target prices



4.2. Persistent differences in analysts' forecast performance

To determine whether financial analysts persistently differ in their ability to forecast prices, we conduct the same analysis as before, but use our measure of target price forecast quality (*TPFQ*), instead of target price accuracy (*AFE*), to evaluate the analysts' performance. Since *TPFQ* takes into account how difficult it is to make a forecast, a persistent difference in *TPFQ* implies that different analysts possess different levels of skill. To allow for a direct comparison with our previous results, we use the *ex-post* version of *TPFQ* (defined in Equation 3). For a given period, we define an analyst's *ex-post* performance (*exAFP*) as the average of the *ex-post TPFQ* of the target prices she issued during that period.

As Table 5 shows, when we account for differences in volatility in this way, the persistent difference among analysts' performance vanishes. In the measurement period, the analysts in the first quintile (i.e., best performance) exhibit an average forecast quality of 0.3255 , while the analysts in the fifth quintile (i.e., worst performance) exhibit an average forecast quality of

-0.2366 . In the test period, the difference in forecast quality between the first and fifth quintiles is equal to 0.0088 and is not significant. We obtain this result both for quarterly and semiannual periods.

Table 5. Test of forecasting abilities using the *ex-post* measure of target price forecast quality (*TPFQ*)

This table presents the analysts' *ex-post* forecast performance (*exAFP*) in the test period $]t+\theta;\, t+\theta+1]$, conditional on their forecast performance in the measurement period $]t-1;\, t]$. θ is a 12-month lag which ensures that the measurement period and the test period do not overlap. An analyst's *ex-post* forecast performance (*exAFP*) is measured, for a given period, as the average of the *ex-post* target price forecast quality on the target prices she issued during that period. The *ex-post* target price forecast quality (*TPFQ*) is measured as the expected value of the absolute forecast error, estimated at the time the target price is issued, minus the realized absolute forecast error measured at the end of the 12-month horizon. The measurement periods are quarterly (Panel A) and semiannual (Panel B). We rank analysts in quintiles based on their *ex-post* forecast performance (*exAFP*) in the measurement period, and we obtain the corresponding forecast performance in the test period. Conditional on the ranking made during the measurement period $]t-1;\, t]$, we report, for the test period $]t+\theta;\, t+\theta+1]$, the analysts' *ex-post* forecast performance (*exAFP*). We test the difference of means across the top and bottom quintiles using a t-test. ***/**/* represent significance at the 0.01, 0.05, and 0.1 level.

	Panel A: Quarterly period		
		Measurement period $]t-1;t]$	Test period $]t+\theta;t+\theta+1]$
Performance quintile (measurement period)	Number of observations	Analysts' *ex-post* forecast performance (*exAFP*)	Analysts' *ex-post* forecast performance (*exAFP*)
1 (Best)	12,779	0.3255	0.0630
2	12,798	0.1638	0.0624
3	12,802	0.0828	0.0591
4	12,798	–0.0049	0.0558
5 (Worst)	12,768	–0.2366	0.0542
Diff (1-5)		0.5621	0.0088
Mean t-test		18.97 ***	0.3600

	Panel B: Semiannual period		
		Measurement period $]t-1;t]$	Test period $]t+\theta;t+\theta+1]$
Performance quintile (measurement period)	Number of observations	Analysts' *ex-post* forecast performance (*exAFP*)	Analysts' *ex-post* forecast performance (*exAFP*)
1 (Best)	7,254	0.3082	0.0584
2	7,265	0.1572	0.0607
3	7,267	0.0824	0.0558
4	7,265	0.0015	0.0555
5 (Worst)	7,252	–0.21136	0.0528
Diff (1-5)		0.5195	0.0056
Mean t-test		13.33 ***	0.1800

4.3. Persistent differences in a dynamic setting

The main limitation of the *ex-post* TPFQ measure is the need to introduce a 12-month lag time between the measurement period and the test period. A second limitation is that the target price forecast quality is evaluated using the stock price at the end of the 12-month horizon. In practice, when a revision occurs the first forecast becomes inactive and only the revision is taken into account.

When used in a dynamic setting, however, our TPFQ measure allows us to consider revisions and to estimate the variations in analysts' performance on a daily basis. In this case, an analyst's forecast performance $AFP_{t,t+\tau}$ over the period $]t;t+\tau]$ is defined as the sum of her daily forecast performance over this period.

Contrary to *ex-post* measures, the analyst's forecast performance can now be measured using only information from this same period.[11] To evaluate the persistence of difference over the short run, we set the measurement period to $]t-1;t]$ and the test period to $]t;t+1]$ (we no longer need to add a lag between the measurement period and the test period) and restrict our sample period to 2001-2012.[12]

The results in Table 6 show that, even over the short run, analysts do not exhibit persistent differences in forecast quality. Using both quarterly and semiannual frequencies, we observe no significant differences, within the test period, between the forecast performance (*AFP*) of analysts in the first quintile and those in the fifth quintile.

4.4. Persistent differences by industry

Since financial analysts frequently specialize by industry (Boni and Womack, 2006; Kadan et al., 2012), another way to uncover differences in forecasting ability is to test for persistent differences in forecast quality within each industry. Taking this approach, we ensure that our test of persistence is not affected by unpredictable industry-wide shocks and unobservable industry-specific factors.

11 When using ex-post measures, one needs to have information up to $T+12$ in order to assess the analyst's forecast performance.

12 For 2000, we do not observe the target prices issued in 1999, which could still be outstanding.

Table 6. Test of forecasting abilities in a dynamic setting

This table reports the analysts' forecast performance (*AFP*) in the test period $]t; t+1]$, conditional on their forecast performance in the measurement period $]t-1; t]$. An analyst's forecast performance (*AFP*), for a given period, is defined as the average of the target price forecast quality (*TPFQ*) on her outstanding target prices during that period. The measurement periods are quarterly (Panel A) and semiannual (Panel B). We rank the analysts in quintiles based on their forecast performance in the measurement period, and we obtain the corresponding forecast performance in the test period. Conditional on the ranking made during the measurement period $]t-1; t]$, we report the analyst's forecast performance (*AFP*) for the test period $]t; t+1]$. The statistical significance of the difference across top and bottom quintiles is computed using a t-test. ***/**/* represent significance at the 0.01, 0.05, and 0.1 level.

	Panel A: Quarterly period		
		Measurement period $]t-1;t]$	Test period $]t;t+1]$
Performance quintile (measurement period)	Number of observations	Analysts' forecast performance (*AFP*)	Analysts' forecast performance (*AFP*)
1 (Best)	27,154	0.1237	0.0150
2	27,174	0.0519	0.0158
3	27,174	0.0204	0.0166
4	27,174	–0.0134	0.0148
5 (Worst)	27,145	–0.1094	0.0108
Diff (5-1)		0.2330	0.0042
Mean t-test		18.16***	0.37

	Panel B: Semiannual period		
		Measurement period $]t-1;t]$	Test period $]t;t+1]$
Performance quintile (measurement period)	Number of observations	Analysts' forecast performance (*AFP*)	Analysts' forecast performance (*AFP*)
1 (Best)	13,399	0.1611	0.0265
2	13,408	0.0727	0.0291
3	13,409	0.0333	0.0257
4	13,408	–0.0087	0.0215
5 (Worst)	13,397	–0.1324	0.0129
Diff (5-1)		0.2934	0.0135
Mean t-test		13.49***	0.75

As in Fama and French (1997), we use the four-digit SIC codes to define 48 industries.[13] Table 7 provides analysts' *ex-post* forecast performance in the test period, conditional on their performance in the measurement period. Column 7 gives the difference in *exAFP* between analysts who rank in the best performance quintile in the measurement period and those who rank in the worst. Column 8 (column 9) reports the t-statistics (p-values). Our results indicate that differences in forecast quality are not persistent within industries. The reported t-statistics indicate that none of the differences are statistically significant.[14] This additional test by industry confirms our previous findings that analysts do not differ in their ability to forecast stock prices.

4.5. Persistent differences across brokerage houses

While we do not find differences in forecast ability across analysts, such differences may still exist across brokerage houses. Large houses possess superior resources and have better access to information, which could translate into better target price forecasts. Additionally, large brokers may offer better compensation packages to their employees and thus attract the best analysts. Stickel (1995) finds that stock recommendations issued by analysts employed by large brokerage houses generate a stronger market reaction. Similarly, Clement (1999) and Mikhail et al. (1997) show that analysts employed by large brokerage houses are more accurate in their earnings forecasts.

The results of our analysis appear in Table 8. We conduct the same test as in Table 5, but consider brokerage houses instead of individual analysts. We find that the best brokerage houses in the measurement period do not perform better than the other houses in the test period.

13 We point out, however, that our results are robust to different industry specifications.

14 With the exception of the Shipbuilding and Railroad Equipment industry.

Table 7. Test of persistent differences in analysts' forecast performance by industry

This table presents, for each industry, the analysts' *ex-post* forecast performance (*exAFP*) in the test period $]t+\theta;\, t+\theta+1]$, conditional on their forecast performance in the measurement period $]t-1;\, t]$. θ is a 12-month lag which ensures that the measurement period and the test period do not overlap. An analyst's *ex-post* forecast performance (*exAFP*) is measured, for a given period and a given industry, as the average of the *ex-post* target price forecast quality on the target prices she issued during that period for the given industry. We define 48 industries from four-digit SIC codes, as in Fama and French (1997). The *ex-post* target price forecast quality (*TPFQ*) is measured as the expected value of the absolute forecast error, estimated at the time the target price is issued, minus the realized absolute forecast error measured at the end of the 12-month horizon. The measurement periods are semiannual. We rank analysts in quintiles based on their *ex-post* forecast performance (*exAFP*) in the measurement period, and we obtain the corresponding forecast performance in the test period. Conditional on the ranking made during the measurement period $]t-1;\, t]$, we report, for the test period $]t+\theta;\, t+\theta+1]$, the analysts' *ex-post* forecast performance (*exAFP*). We test the difference of means across the top and bottom quintiles using a t-test.

	Average *exAFP* in the test period $]t+\theta;\, t+\theta+1]$							
	Best analysts (highest quintile of past *exAFP*)	(2)	(3)	(4)	Worst analysts (lowest quintile of past *exAFP*)	Diff (5-1)	Mean t-test	P-value
Agricultural production	0.0775	0.1018	0.0757	0.0271	0.1323	–0.0548	–0.0300	0.9772
Food Products	0.0854	0.0686	0.0660	0.0831	0.0611	0.0244	0.5200	0.6078
Candy and Soda	0.0505	0.0528	0.0463	0.0205	0.0511	–0.0005	0.1100	0.9158
Beer and Liquor	0.0538	0.0194	0.0513	0.0094	0.1261	–0.0722	–0.3500	0.7260
Tobacco Products	0.0205	0.0289	0.0049	0.0330	0.0628	–0.0423	–1.2000	0.2379
Recreation	0.0374	0.0429	0.0796	–0.0008	0.0446	–0.0072	0.5600	0.5778
Entertainment	–0.0847	–0.0228	–0.0117	0.0518	0.0210	–0.1058	–0.7700	0.4462
Printing and Publishing	–0.0157	0.0626	0.0462	0.0565	0.0963	–0.1120	–1.4300	0.1594
Consumer Goods	0.0101	0.0319	0.0573	0.0590	0.0312	–0.0211	–0.5500	0.5857
Apparel	0.0312	0.0461	0.0458	0.0772	0.0734	–0.0423	–0.8500	0.3999
Healthcare	0.0735	0.0816	0.0843	0.0690	0.0936	–0.0201	–0.5600	0.5762
Medical Equipment	0.0856	0.0655	0.0274	0.0687	0.0952	–0.0096	–0.2100	0.8326
Pharmaceutical Products	0.0581	0.0331	0.0431	0.0245	0.0463	0.0118	0.2800	0.7788
Chemicals	0.0729	0.0574	0.0379	0.0763	0.0595	0.0134	0.5300	0.5963
Rubber and Plastic Products	0.1116	0.0999	0.0886	0.1077	0.1833	–0.0717	–1.1500	0.2584
Textiles	0.1173	0.0195	0.0892	0.0337	–0.0156	0.1329	1.2000	0.2373
Construction Materials	0.0872	0.0408	0.0481	0.0345	0.0501	0.0371	0.6200	0.5358
Construction	0.1382	0.1080	0.0882	0.0941	0.1069	0.0313	0.2900	0.7732

Steel Works	0.0167	0.0459	0.0799	0.0624	0.0294	–0.0128	–0.4900	0.6233
Fabricated Products	–0.0253	0.0069	0.0883	0.1263	0.1259	–0.1512	–1.7800	0.0903
Machinery	0.1025	0.0712	0.0769	0.0876	0.0943	0.0082	0.2800	0.7834
Electrical Equipment	0.0846	0.0818	0.0968	0.0932	0.1102	–0.0256	–0.6700	0.5049
Automobiles and Trucks	0.0723	0.0250	0.0294	0.0319	0.0684	0.0038	0.1600	0.8713
Aircraft	0.1173	0.0583	0.0581	0.0676	0.0955	0.0218	0.4100	0.6843
Shipbuilding, Railroad Equipment	0.0861	0.1303	0.1044	0.2093	0.3196	–0.2334	–2.9400	0.0083
Defense	–0.0094	0.0359	0.0119	0.0346	0.0087	–0.0181	0.0200	0.9835
Precious Metals	0.0712	0.0708	0.0447	0.0361	0.0624	0.0087	0.0200	0.9841
Non Metallic and Industrial Metal Mining	–0.0167	0.0278	–0.0016	–0.0075	–0.0312	0.0145	–0.7300	0.4699
Coal	–0.0570	0.0205	0.0445	0.0211	0.0129	–0.0700	–0.9500	0.3467
Petroleum and Natural Gas	0.0572	0.0635	0.0711	0.0784	0.0651	–0.0079	–0.1500	0.8797
Utilities	0.0647	0.0452	0.0414	0.0501	0.0595	0.0052	0.1400	0.8928
Communication	0.0574	0.0220	0.0168	0.0281	0.0471	0.0103	–0.0600	0.9509
Personal Services	0.0503	0.0930	0.0962	0.0649	0.0833	–0.0329	–0.6000	0.5533
Business Services	0.0903	0.0737	0.0799	0.0650	0.0653	0.0250	0.7700	0.4455
Computers	0.0665	0.0616	0.0384	0.0471	0.0539	0.0126	0.6500	0.5218
Electronic Equipment	0.0477	0.0427	0.0550	0.0540	0.0537	–0.0060	–0.1200	0.9081
Measuring and Control Equipment	0.0925	0.0781	0.0729	0.0654	0.0851	0.0074	0.0200	0.9814
Business Supplies	0.0511	0.0192	0.0645	0.0719	0.0679	–0.0167	–0.4200	0.6752
Shipping Containers	0.0999	0.1244	0.1099	0.0524	0.0393	0.0606	1.3100	0.2036
Transportation	0.0723	0.0726	0.0788	0.0991	0.0677	0.0046	–0.1700	0.8639
Wholesale	0.0781	0.0661	0.0465	0.0734	0.0865	–0.0083	–0.0500	0.9595
Retail	0.0556	0.0445	0.0509	0.0620	0.0641	–0.0085	–0.2300	0.8159
Restaurants, Hotels, Motels	0.1248	0.0645	0.0600	0.0452	0.0312	0.0936	1.5900	0.1197
Banking	0.0727	0.0659	0.0637	0.0711	0.0659	0.0068	0.2300	0.8223
Insurance	0.0627	0.0677	0.0547	0.0608	0.0707	–0.0080	–0.2000	0.8410
Real Estate	–0.0196	0.0900	0.0074	0.0744	0.0433	–0.0629	–0.6300	0.5312
Trading	0.0713	0.0521	0.0478	0.0586	0.0723	–0.0010	–0.1500	0.8789
Other	0.0890	0.0678	0.0719	0.0815	0.1027	–0.0137	–0.2200	0.8246

Table 8. Test of persistent differences in forecast quality across brokerage houses

This table presents the brokerage houses' *ex-post* forecast performance (*exAFP*) in the test period $]t+\theta;\ t+\theta+1]$, conditional on their forecast performance in the measurement period $]t-1;\ t]$. θ is a 12-month lag which ensures that the measurement period and the test period do not overlap. A brokerage house's *ex-post* forecast performance (*exAFP*) is measured, for a given period, as the average of the *ex-post* target price forecast quality on the target prices issued by all its analysts during that period. The *ex-post* target price forecast quality (*TPFQ*) is measured as the expected value of the absolute forecast error, estimated at the time the target price is issued, minus the realized absolute forecast error measured at the end of the 12-month horizon. The measurement periods are quarterly (Panel A) and semiannual (Panel B). We rank brokerage houses in quintiles based on their *ex-post* forecast performance (*exAFP*) in the measurement period, and we obtain the corresponding forecast performance in the test period. Conditional on the ranking made during the measurement period $]t-1;\ t]$, we report, for the test period $]t+\theta;\ t+\theta+1]$, the brokers' *ex-post* forecast performance (*exAFP*). We test the difference of means across the top and bottom quintiles using a t-test. ***/**/* represent significance at the 0.01, 0.05, and 0.1 level.

		Panel A: Quarterly period	
		Measurement period $]t-1;t]$	Test period $]t+\theta;t+\theta+1]$
Performance quintile (measurement period)	Number of observations	Brokers' *ex-post* forecast performance (*exAFP*)	Brokers' *ex-post* forecast performance (*exAFP*)
1 (Best)	1,445	0.2552	0.0516
2	1,464	0.1195	0.0562
3	1,467	0.0690	0.0594
4	1,464	0.0151	0.0580
5 (Worst)	1,436	–0.1786	0.0492
Diff (1-5)		0.4338	0.0024
Mean t-test		14.72 ***	0.17

		Panel B: Semiannual period	
		Measurement period $]t-1;t]$	Test period $]t+\theta;t+\theta+1]$
Performance quintile (measurement period)	Number of observations	Brokers' *ex-post* forecast performance (*exAFP*)	Brokers' *ex-post* forecast performance (*exAFP*)
1 (Best)	767,000	0.2465	0.0422
2	775,000	0.1153	0.0544
3	778,000	0.0692	0.0630
4	775,000	0.0189	0.0550
5 (Worst)	764,000	–0.1673	0.0461
Diff (1-5)		0.4139	–0.0040
Mean t-test		10.44 ***	–0.01

5. Robustness checks

5.1. Impact of learning

One reason why we might not observe any persistent differences in analysts' forecast performance is that financial analysts learn over time and subsequently improve their forecast quality. If experience influences target price forecast quality, then the younger and more inexperienced analysts will be ranked in the poorer-performing quintiles when they enter the sample period. They will then gradually move toward the best quintile as they acquire experience. These individuals could therefore add noise to our analysis of persistent differences in analysts' forecast performance.

In order to control for the effects of learning, in each period we rank analysts into three terciles by level of experience. We then conduct our analysis on the three subsamples. Following Clement (1999), we measure an analyst's experience in year t by counting the number of previous years for which the analyst supplied at least one EPS forecast.[15]

The results, presented in Table 9, show that the learning process cannot explain the absence of persistent differences in forecast performance. The differences observed in the test period are still not statistically significant when we identify subsamples by level of experience.

5.2. Teams of analysts

When creating our initial sample, as noted in Section 1, we removed target prices issued under names that seemed to indicate a team of analysts rather than a single individual. However, any name reported in I/B/E/S may correspond to the lead analyst of a team, rather than to a single individual. Another way to detect forecasts issued by teams is to check the number of firms covered per year. For this robustness check, we assume that analysts covering more than 15 firms per year are teams, not individuals, and we remove these observations from our sample.

15 We use EPS forecasts, rather than target prices, to calculate analysts' experience because no data for target prices is available prior to 2000. Also, some analysts might have only started issuing target prices after 2000, even though they were active prior to that year.

Table 9. Impact of learning on forecasting abilities

This table presents the analysts' *ex-post* forecast performance (*exAFP*) in the test period $]t+\theta;\, t+\theta+1]$, conditional on their forecast performance in the measurement period $]t-1;\, t]$. θ is a 12-month lag which ensures that the measurement period and the test period do not overlap. An analyst's *ex-post* forecast performance (*exAFP*) is measured, for a given period, as the average of the *ex-post* target price forecast quality on the target prices she issued during that period. In panel A, the analysis is done on a sub-sample of inexperienced analysts (first tercile). Panel B corresponds to analysts who belong to the second tercile of experience, and Panel C corresponds to experienced analysts. The *ex-post* target price forecast quality is measured as the expected value of the absolute forecast error, estimated at the time the target price is issued, minus the realized absolute forecast error measured at the end of the 12-month horizon. We rank analysts in quintiles based on their *ex-post* forecast performance (*exAFP*) in the measurement period, and we obtain the corresponding forecast performance in the test period. Conditional on the ranking made during the measurement period $]t-1;\, t]$, we report, for the test period $]t+\theta;\, t+\theta+1]$, the analysts' *ex-post* forecast performance (*exAFP*). We test the difference of means across the top and bottom quintiles using a t-test. ***/**/* represent significance at the 0.01, 0.05, and 0.1 level.

	Panel A: Analysts in the first tercile of experience (low experience)		
		Measurement period $]t-1;t]$	Test period $]t+\theta;t+\theta+1]$
Performance quintile (measurement period)	Number of observations	Analysts' *ex-post* forecast performance (*exAFP*)	Analysts' *ex-post* forecast performance (*exAFP*)
1 (Best)	4.955	0.3478	0.0542
2	4.973	0.1719	0.0611
3	4.971	0.0835	0.0560
4	4.973	–0.0163	0.0527
5 (Worst)	4.944	–0.2830	0.0467
Diff (1-5)		0.6308	0.0075
Mean t-test		19.61 ***	0.22

	Panel B: Analysts in the second tercile of experience		
		Measurement period $]t-1;t]$	Test period $]t+\theta;t+\theta+1]$
Performance quintile (measurement period)	Number of observations	Analysts' *ex-post* forecast performance (*exAFP*)	Analysts' *ex-post* forecast performance (*exAFP*)
1 (Best)	3.768	0.3122	0.0682
2	3.791	0.1598	0.0666
3	3.791	0.0813	0.0642
4	3.791	–0.0014	0.0545
5 (Worst)	3.762	–0.2205	0.0646
Diff (1-5)		0.5327	0.0036
Mean t-test		18.6 ***	0.24

Performance quintile (measurement period)	Panel C: Analysts in the third tercile of experience (high experience)		
		Measurement period $]t-1;t]$	Test period $]t+\theta;t+\theta+1]$
	Number of observations	Analysts' *ex-post* forecast performance (*exAFP*)	Analysts' *ex-post* forecast performance (*exAFP*)
1 (Best)	3.524	0.2921	0.0694
2	3.544	0.1532	0.0596
3	3.550	0.0807	0.0595
4	3.544	0.0022	0.0602
5 (Worst)	3.518	–0.1793	0.0552
Diff (1-5)		0.4714	0.0142
Mean t-test		16.29 ***	0.59

Table 10 provides the results of our analysis for this restricted sample. We still do not observe persistent differences in forecast performance. This additional test indicates that the potential presence of teams of analysts in our sample does not impact our results.

5.3. Slow adjustment of target prices

Target prices are usually embedded in analysts' reports and, because writing a report is a long and arduous task, the prices may not be adjusted as often as they should be. For example, an analyst may change her mind about the future price of a stock a month after her initial forecast, but she might wait for publication of the next report to officially revise her target price. This fact might cause analysts to appear less skilled than they actually are. To test this hypothesis, we restrict the validity of the target prices to a shorter period of time (e.g., one month). That is, for a given stock and a given target price, we compute TPFQ only for the first month following the issue date. In other words, we consider forecasts to become inactive after one month. We then compute the analysts' forecast performance (*AFP*) using these short-validity target prices. We conduct the same analysis as before to test for the existence of differential abilities. Our (unreported) results – using one month, three months, and six months for the validity of the target prices – confirm our findings that analysts do not differ in their ability to forecast stock prices.

Table 10. Impact of the existence of teams of analysts on tests of differential abilities

In this table, the sample is restricted to analysts who cover no more than 15 firms per year. Analysts who cover more than 15 firms per year are assumed to be teams of analysts and are removed from the sample. This table presents the analysts' *ex-post* forecast performance (*exAFP*) in the test period $]t + \theta;\ t + \theta + 1]$, conditional on their forecast performance in the measurement period $]t - 1;\ t]$. θ is a 12-month lag which ensures that the measurement period and the test period do not overlap. An analyst's *ex-post* forecast performance (*exAFP*) is measured, for a given period, as the average of the *ex-post* target price forecast quality on the target prices she issued during that period. The *ex-post* target price forecast quality (*TPFQ*) is measured as the expected value of the absolute forecast error, estimated at the time the target price is issued, minus the realized absolute forecast error measured at the end of the 12-month horizon. The measurement periods are quarterly (Panel A) and semiannual (Panel B). We rank analysts in quintiles based on their *ex-post* forecast performance (*exAFP*) in the measurement period, and we obtain the corresponding forecast performance in the test period. Conditional on the ranking made during the measurement period $]t - 1;\ t]$, we report, for the test period $]t + \theta;\ t + \theta + 1]$, the analysts' *ex-post* forecast performance *exAFP*. We test the difference of means across the top and bottom quintiles using a t-test. ***/**/* represent significance at the 0.01, 0.05, and 0.1 level.

	Panel A: Quarterly period		
		Measurement period $]t-1;t]$	Test period $]t+\theta;t+\theta+1]$
Performance quintile (measurement period)	Number of observations	Analysts' *ex-post* forecast performance (*exAFP*)	Analysts' *ex-post* forecast performance (*exAFP*)
1 (Best)	10,416	0.3386	0.0625
2	10,436	0.1700	0.0611
3	10,432	0.0848	0.0587
4	10,436	–0.0073	0.0560
5 (Worst)	10,408	–0.2530	0.0537
Diff (1-5)		0.5917	0.0088
Mean t-test		19.48 ***	0.37

	Panel B: Semiannual period		
		Measurement period $]t-1;t]$	Test period $]t+\theta;t+\theta+1]$
Performance quintile (measurement period)	Number of observations	Analysts' *ex-post* forecast performance (*exAFP*)	Analysts' *ex-post* forecast performance (*exAFP*)
1 (Best)	6,026	0.3196	0.0568
2	6,037	0.1625	0.0580
3	6,033	0.0838	0.0551
4	6,037	–0.0013	0.0557
5 (Worst)	6,018	–0.2267	0.0519
Diff (1-5)		0.5463	0.0049
Mean t-test		13.84 ***	0.15

5.4. Alternative measures of accuracy

In this paper, we use the absolute forecast error (*AFE*) as our main measure of target price accuracy since it is the most popular such measure. Our measure of forecast difficulty is therefore equal to $E[AFE]$ and our measure of target price forecast quality can be written as $E[AFE]-AFE$.

However, alternative ways to measure accuracy exist in the literature. For instance, Bradshaw et al. (2013) measure whether the actual closing price, as of the end of the 12-month forecast horizon, is at or above the target price ($TPMETEND$). The forecast difficulty, in this case, equals $E[TPMETEND]$ and the measure of target price forecast quality can be written as $E[TPMETEND]-TPMETEND$. Estimating the expected value of the binary variable $TPMETEND$ is thus equivalent to computing the price of a digital cash-or-nothing option.

In addition, Bradshaw et al. (2013) define target price accuracy as whether the target price is met before the end of the horizon ($TPMETANY$). The probability of meeting the target price depends on the volatility of the underlying stock, the forecast horizon, and the magnitude of the change predicted by the analyst. The forecast difficulty equals $E[TPMETANY]$ and the measure of target price forecast quality can be written as $E[TPMETANY]-TPMETANY$. In this case, estimating the expected value of the binary variable $TPMETANY$ is equivalent to computing the price of a cash-or-nothing up-and-in (or down-and-in) digital barrier option.

To investigate the impact of using different measures of accuracy, we repeated our tests of persistence in analysts' performance using these two measures. Our results, presented in Table 11, once again show that analysts do not exhibit persistent differences in forecasting ability. We conclude that our results do not depend on how target price accuracy is defined.

Table 11. Test of forecasting abilities using alternative measures of target price accuracy

This table presents the analysts' *ex-post* forecast performance (*exAFP*) in the test period $]t + \theta; t + \theta + 1]$, conditional on their forecast performance in the measurement period $]t - 1; t]$. θ is a 12-month lag which ensures that the measurement period and the test period do not overlap. An analyst's *ex-post* forecast performance (*exAFP*) is measured, for a given period, as the average of the *ex-post* target price forecast quality on the target prices she issued during that period. In panel A, the *ex-post* target price forecast quality (*TPFQ*) is measured as *TPMETEND – E[TPMETEND]*, where *TPMETEND* is a binary variable that takes the value 1 if the actual closing price as of the end of the 12-month forecast horizon is at or above the target price, and 0 otherwise. In panel B, the *ex-post* target price forecast quality (*TPFQ*) is measured as *TPMETANY – E[TPMETANY]*, where *TPMETANY* is a binary variable that takes the value 1 if the target price is met before the end of the horizon, and 0 otherwise. The measurement periods are quarterly. We rank analysts in quintiles based on their *ex-post* forecast performance (*exAFP*) in the measurement period, and we obtain the corresponding forecast performance in the test period. Conditional on the ranking made during the measurement period $]t - 1; t]$, we report the analyst's *ex-post* forecast performance (*exAFP*) for the test period $]t + \theta; t + \theta + 1]$. We test the difference of means across the top and bottom quintiles using a t-test. ***/**/* represent significance at the 0.01, 0.05, and 0.1 level.

	Panel A: Forecast quality measured as *TPMETEND – E[TPMETEND]*		
		Measurement period $]t - 1; t]$	Test period $]t + \theta; t + \theta + 1]$
Performance quintile (measurement period)	Number of observations	Analysts' forecast performance	Analysts' forecast performance
1 (Best)	12,779	0.4960	0.0860
2	12,798	0.1761	0.0722
3	12,802	0.0166	0.0702
4	12,798	–0.1156	0.0649
5 (Worst)	12,768	–0.2767	0.0594
Diff (5-1)		0.7726	0.0266
Mean t-test		28.07***	0.86

	Panel B: Forecast quality measured as *TPMETANY – E[TPMETANY]*		
		Measurement period $]t - 1; t]$	Test period $]t + \theta; t + \theta + 1]$
Performance quintile (measurement period)	Number of observations	Analysts' forecast performance	Analysts' forecast performance
1 (Best)	12,779	0.3835	0.0158
2	12,798	0.1577	0.0158
3	12,802	0.0176	0.0085
4	12,798	–0.1299	0.0081
5 (Worst)	12,768	–0.3931	0.0004
Diff (5-1)		0.7766	0.0153
Mean t-test		36.80***	0.69

6. Conclusion

This paper provides a new framework for evaluating the performance of financial analysts. We show that absolute forecast error cannot be used as a proxy for the quality of information contained in target prices, because the accuracy of a stock price forecast is affected by both stock return volatility and the forecast horizon. We find evidence of a strong nonlinear relationship between stock return volatility and absolute forecast error, which implies that analysts who cover low volatility firms are more accurate.

We introduce a new measure of target price forecast quality, which takes into consideration the difficulty of making an accurate forecast, by incorporating differences in stock return volatility and forecast horizon. Building on option-pricing theory, we capture forecast difficulty by estimating the expected value of the absolute forecast error, i.e., the accuracy that is to be expected if the target price was mechanically issued. We define our measure as the difference between the forecast difficulty and the realized accuracy.

We conclude that target prices do contain information (i.e., our measure of information quality is positive, on average), which is consistent with the evidence that market participants react to target price revisions (Brav and Lehavy, 2003; Asquith et al., 2005). However, financial analysts do not exhibit differences in their ability to forecast stock prices.

References

Asquith, P., Mikhail, M.B., Au, A.S., 2005. Information content of equity analyst reports. *Journal of Financial Economics* 75, 245–282.

Baber, W.R., Kang, S.H., 2002. The impact of split adjusting and rounding on analysts' forecast error calculations. *Accounting Horizons* 16, 277–289.

Bilinski, P., Lyssimachou, D., Walker, M., 2013. Target price accuracy: International evidence. *The Accounting Review* 88, 825–851.

Black, F., Scholes, M.S., 1973. The pricing of options and corporate liabilities. *Journal of Political Economy* 81, 637–654.

Boni, L., Womack, K.L., 2006. Analysts, industries, and price momentum. *Journal of Financial and Quantitative Analysis* 41, 85–109.

Bradshaw, M.T., Brown, L.D., Huang, K., 2013. Do sell-side analysts exhibit differential target price forecasting ability? *Review of Accounting Studies* 18, 930–955.

Brav, A., Lehavy, R., 2003. An empirical analysis of analysts' target prices: Short-term informativeness and long-term dynamics. *Journal of Finance* 58, 1933–1967.

Clement, M.B., 1999. Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter? *Journal of Accounting and Economics* 27, 285–303.

Fama, E.F., French, K.R., 1997. Industry costs of equity. *Journal of Financial Economics* 43, 153–93.

Hong, H., Kubik, J.D., Solomon, A., 2000. Security analysts' career concerns and herding of earnings forecasts. RAND *Journal of Economics* 31, 121–144.

Kadan, O., Madureira, L., Wang, R., Zach, T., 2012. Analysts' industry expertise. *Journal of Accounting and Economics* 54, 95–120.

Kahneman, D., 2011. Thinking, fast and slow. New York: Farrar, Straus, and Giroux. Kerl, A.G., 2011. Target price accuracy. *BuR - Business Research* 4, 74–96.

Mikhail, M.B., Walther, B.R., Willis, R.H., 1997. Do security analysts improve their performance with experience? *Journal of Accounting Research* 35, 131–157.

Park, C.W., Stice, E.K., 2000. Analyst forecasting ability and the stock price reaction to forecast revisions. *Review of Accounting Studies* 5, 259–272.

Payne, J.L., Thomas, W.B., 2003. The implications of using stock-split adjusted I/B/E/S data in empirical research. *The Accounting Review* 78, 1049–1067.

Sinha, P., Brown, L.D., Das, S., 1997. A re-examination of financial analysts' differential earnings forecast accuracy. *Contemporary Accounting Research* 14, 1–42.

Stickel, S.E., 1995. The anatomy of the performance of buy and sell recommendations. *Financial Analysts Journal* 51, 25–39.

Appendix A

Assuming, as is common in the financial literature, that stock prices follow a Geometric Brownian Motion, we have:

$$\log(S_T) \sim \mathcal{N}\left(\log(S_t) + \left(\mu - \frac{1}{2}\sigma^2\right)(T-t), \sigma^2(T-t)\right), \tag{10}$$

where $\mathcal{N}()$ is the normal distribution, $\propto$ is the drift, and σ is the volatility.

The probability that the stock price ends up inside an interval $[b_l;b_u]$ at the end of a given horizon is equal to:

$$\begin{aligned}
&\Pr[\log(b_l) < \log(S_T) < \log(b_u)] \\
&=\Pr\left[\frac{\log(b_l / S_t) - (\mu - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}} < z < \frac{\log(b_u / S_t) - (\mu - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}\right] \\
&=\Pr[b_l^* < z < b_u^*] \\
&= \Phi(b_u^*) - \Phi(b_l^*),
\end{aligned} \tag{11}$$

where b_l^* and b_u^* are defined by:

$$b_l^* = \frac{\log(b_l / S_t) - (\mu - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}} \text{ and } b_u^* = \frac{\log(b_u / S_t) - (\mu - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}, \tag{12}$$

where $\Phi()$ is the cumulative distribution function of a standard Gaussian random variable and z is a standard Gaussian variable.

The probability of the stock price ending up inside a given interval at the end of a determined horizon is a nonlinear function of both the volatility and the horizon. It follows, by extension, that the expected value of the absolute forecast error is a nonlinear function of both the stock return volatility and the target price horizon.

Appendix B

The literature on earnings forecasts uses relative measures of accuracy to account for differences in predictability (i.e., differences in earnings volatility). That is, the accuracy of a forecast is determined with respect to the accuracy of other forecasts issued under similar conditions (i.e., forecasts issued on the same firm and during the same period of time). Clement (1999) proposes to measure an analyst's performance by comparing the analyst's absolute forecast error to the average absolute forecast error of other analysts following the same stock during the same time period. Hong et al. (2000) propose an alternative way to control for differences in earnings predictability: For a given firm and a given year, they rank analysts with respect to the absolute forecast error of their most recent forecast. These rankings are then transformed into scores.

Despite their popularity, relative measures of accuracy like these present a number of issues which prevent them from being of use in the context of target prices. First, the end of the forecast horizon for a target price depends on the issue date. Contrary to earnings forecasts, for which analysts forecast end-of-year (or end-of-quarter) earnings, target prices will have different horizons if the issue dates are separated in time. The economic meaning of comparing the accuracy of target prices issued at different points in time is not clear.

Second, when the number of analysts covering a stock is low, or when the number of firms covered by an analyst is too small (as in Hong et al., 2000), relative measures of accuracy may be fairly noisy or even impossible to compute.

Third, relative measures of accuracy control for firm effects (by considering the average absolute forecast error across analysts for a given stock), without distinguishing between forecast difficulty and the analysts' shared biases. Yet these firm effects, as captured by the differences in average *AFE* across firms, are a function of both the difficulty of making a forecast and the biases shared by analysts covering the same firm. In terms of forecast difficulty, the higher the stock return volatility, the greater the average *AFE* will be. In the case of shared biases we can distinguish two types that affect *AFE*, as the following examples illustrate.

The first type of shared bias results from analysts' failure to take into account the impact of firm-specific factors on future returns. According to the literature, high accrual firms tend to earn low future returns. Because accruals have an impact on the expected return, analysts can take this information into account, when forecasting future stock prices, by issuing higher target prices for low accrual firms and lower target prices for high accrual firms. Assume, however, that analysts who cover high accrual firms fail to recognize the impact of accruals on future returns. These analysts will issue upwardly biased target prices. We will then observe that the average AFE on high accrual firms is higher than the average AFE on low accrual firms. In relative measures of accuracy, this difference in average AFE will be captured as a firm effect, and analysts covering high accrual firms will therefore not be penalized for their failure to account for the accruals factor. Yet it is the analysts' task to assess the potential impact of firm-specific factors (such as accruals, firm size, book-to-market, momentum, liquidity, etc.) on future returns. Failure to do so weakens the quality of the information they provide to market participants. A measure of analysts' performance should take into account this inability to correctly assess the impact of firm-specific factors, even if it is shared by the majority of analysts.

The second type of shared bias arises when the analysts covering a given stock are too optimistic (or too pessimistic). Let us assume that there are two distinct groups of analysts G_A and G_B, and two identical firms, A and B. Analysts in the G_A group cover firm A, while analysts in the G_B group cover firm B. We assume that analysts covering firm A issue unbiased target prices (their target prices are equal to $S_T + \varepsilon$ where ε is a white noise process), while analysts covering firm B exhibit an optimistic bias equal to Δ (their target prices are equal to $S_T + \Delta + \varepsilon$). The difference in average AFE between firm A and firm B is therefore equal to Δ. When using relative measures of accuracy, the optimism bias exhibited by analysts covering firm B is thus captured by firm effects; these measures would not penalize the analysts covering firm B for their unjustified optimism. Yet again, the role of analysts is to provide the best possible estimation of future returns. They must take into account all the factors (such as market capitalization, book-to-market, momentum, liquidity, etc.) that may impact future returns and adjust their target price accordingly. Their estimation of future returns should not be biased by a shared optimism (or pessimism). Forecast difficulty relates to the second order moment (i.e., the volatility) of the stock price process, but not to the first order moment (i.e., the expected return). Therefore, a measure of analysts' performance should penalize financial analysts for their shared biases.

Appendix C

Assuming that the stock price follows a geometric Brownian motion, the *ex-post* forecast quality $TPFQ_{t,T}$ of a target price issued at time t can be calculated according to the Black and Scholes (1973) model as:

$$\begin{aligned} TPFQ_{t,T} &= e^{r(T-t)}\left[\Phi(d_{1,t}) - \Phi(-d_{1,t}) - TP_{t,T}e^{-r(T-t)}\left(\Phi(d_{2,t}) - \Phi(-d_{2,t})\right)\right] - \left|S_T - TP_{t,T}\right| \\ &= e^{r(T-t)}\left[2\Phi(d_{1,t}) - 1\right] - TP_{t,T}\left[2\Phi(d_{2,t}) - 1\right] - \left|S_T - TP_{t,T}\right|, \end{aligned} \tag{13}$$

with

$$d_{1,t} = \frac{\ln\left(\frac{1}{TP_{t,T}}\right) + \left(r + \frac{1}{2}\sigma_t^2\right)(T-t)}{\sigma_t\sqrt{T-t}} \tag{14}$$

$$d_{2,t} = d_{1,t} - \sigma_t\sqrt{T-t}, \tag{15}$$

where $\Phi()$ is the cumulative distribution function of a standard Gaussian random variable, t is the time at which the forecast is issued, and σ_t is the stock return volatility estimated at time t. The assumption $S_t = 1$ explains the way $d_{1,t}$ is written.

Our approach implies that we do not distinguish between under- and over-achievement. If we consider two forecasts, $TP^1_{t,T} = S_t - \Delta$ and $TP^2_{t,T} = S_t + \Delta$, we should obtain the same forecast quality if, at the end of the horizon, we have $\left|S_T - TP^1_{t,T}\right| = \left|S_T - TP^2_{t,T}\right|$. However, because $\ln(\frac{S_t}{S_t + \Delta}) \neq -\ln(\frac{S_t}{S_t - \Delta})$, this is not the case. In order to solve this issue, we apply a simple transformation (see Appendix D).

Appendix D

Consider two forecasts, $TP^1_{t,T} = S_t - \Delta$ and $TP^2_{t,T} = S_t + \Delta$. As we do not distinguish between under- and over-achievement, we should have $\left|S_T - TP^1_{t,T}\right| = \left|S_T - TP^2_{t,T}\right| \Rightarrow TPFQ^1_{t,T} = TPFQ^2_{t,T}$. However, because $\ln(\frac{S_t}{S_t + \Delta}) \neq -\ln(\frac{S_t}{S_t - \Delta})$, we have

$$\left[C_t\left(TP^1_{t,T}\right) + P_t\left(TP^1_{t,T}\right)\right]e^{r(T-t)} < \left[C_t\left(TP^2_{t,T}\right) + P_t\left(TP^2_{t,T}\right)\right]e^{r(T-t)}. \tag{16}$$

It follows that $TPFQ^1_{t,T} < TPFQ^2_{t,T}$. Even though the absolute deviation of the target price from the stock price S_t is the same for both target prices, $TP^1_{t,T}$ and $TP^2_{t,T}$, and the absolute forecast errors $\left|S_T - TP^1_{t,T}\right|$ and $\left|S_T - TP^2_{t,T}\right|$ are the same at the end of the horizon, we do not obtain the same quality for the two forecasts. We apply a simple transformation to correct this.

When a target price is below the concurrent stock price, we consider the symmetric of the stock price with respect to the target price. That is, we set the target price equal to 1 and consider the concurrent stock price to be equal to $1 + \left|S_t - TP_{t,T}\right|$. However, when there is a positive drift $\propto = r > 0$, the probability of reaching a target price of $TP_{t,T} = S_t - \Delta'$ is lower than that of reaching a target price of $TP_{t,T} = S_t + \Delta'$. Therefore, we need to consider the symmetric of the price with respect to the discounted target price. The consequence of defining the stock price as a function of the discounted target price is that the risk-free rate in the Black and Scholes (1973) model is equal to 0.

Definition 2 *We consider the function f, which measures the discounted deviation of the stock price from the target price. We write:*

$$f(S_{t+\tau}, TP_{t,T}, r) = 1 + \left|S_{t+\tau} - TP_{t,T} e^{-r(T-t-\tau)}\right|. \tag{17}$$

The forecast quality of a target price issued at time t with horizon $T - t$ becomes:

$$\begin{aligned} TPFQ_{t,T} &= (C_t + P_t) e^{r(T-t)} - \left|f(S_T, TP_{t,T}, r) - 1\right| \\ &= e^{r(T-t)} \big(f(S_t, TP_{t,T}, r)[\Phi(d_{1,t}) - \Phi(-d_{1,t})] - [\Phi(d_{2,t}) - \Phi(-d_{2,t})] \big) \\ &\quad - \left|f(S_T, TP_{t,T}, r) - 1\right|, \end{aligned} \tag{18}$$

with

$$d_{1,t} = \frac{\ln\big(f(S_t, TP_{t,T}, r)\big) + \left(\frac{1}{2}\sigma_t^2\right)(T-t)}{\sigma_t \sqrt{T-t}} \tag{19}$$

$$d_{2,t} = d_{1,t} - \sigma_t \sqrt{T-t}, \tag{20}$$

where $\Phi()$ is the Gaussian cumulative distribution function, t is the time at which the forecast was issued, C_t is the value of the call option at time t, P_t is the value of the put option at time t, σ_t is the stock return volatility estimated at time t, r is the risk-free rate, and $T-t$ is the horizon of the target price.

Appendix E

Proposition 1 *For a given final stock price S_T and a given target price $TP_{t,T}$, the forecast quality $TPFQ_{t,T}$ is an increasing function of the stock return volatility σ_t and of the length of the horizon $T-t$.*

Proof. For a given final stock price S_T and a given target price $TP_{t,T}$, the sensitivity of the forecast quality $TPFQ_{t,T}$ to the volatility σ_t is written as:

$$\begin{aligned}\frac{\partial TPFQ_{t,T}}{\partial \sigma_t} &= \left[\frac{\partial C_t}{\partial \sigma_t} + \frac{\partial P_t}{\partial \sigma_t}\right] e^{r(T-t)} \\ &= 2e^{r(T-t)} S_t \sqrt{T-t}\,\Phi'(d_{1,t}) > 0, \end{aligned} \tag{21}$$

with $\Phi'(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$. For a given final stock price S_T and a given target price $TP_{t,T}$, the sensitivity of the forecast quality $TPFQ_{t,T}$ to the horizon $T-t$ becomes:

$$\frac{\partial TPFQ_{t,T}}{\partial (T-t)} = \left[\frac{\partial C_t}{\partial (T-t)} + \frac{\partial P_t}{\partial (T-t)}\right] e^{r(T-t)} + re^{r(T-t)}(C_t + P_t). \tag{22}$$

The sensitivity of a straddle to the maturity $T-t$ is written as:

$$\begin{aligned}\frac{\partial C_t}{\partial (T-t)} + \frac{\partial P_t}{\partial (T-t)} &= S_t\Phi'(d_{1,t})\frac{\partial d_{1,t}}{\partial (T-t)} - e^{-r(T-t)}TP_{t,T}\Phi(d_{2,t})\frac{\partial d_{2,t}}{\partial (T-t)} \\ &\quad + S_t\Phi'(-d_{1,t})\frac{\partial d_{1,t}}{\partial (T-t)} + e^{-r(T-t)}TP_{t,T}\Phi'(d_{2,t})\frac{\partial d_{2,t}}{\partial (T-t)} \\ &= \frac{\sigma}{\sqrt{T-t}} S_t\Phi'(d_{1,t}) + r\Phi_{t,T}e^{-r(T-t)}\left[\Phi(d_{2,t}) - \Phi(-d_{2,t})\right], \end{aligned} \tag{23}$$

with $\Phi'(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$.

The sensitivity of a call option to the maturity $T - t$ is always positive. The sensitivity of a put option to the maturity $T - t$ is also positive, except when the option is deep in the money. The transformation we apply (see Appendix D) implies that the put option is never in the money. Thus, the sensitivity of the straddle to the horizon $T - t$ is always positive. We then have:

$$\frac{\partial TPFQ_{t,T}}{\partial(T - t)} = \left[\frac{\partial C_t}{\partial(T - t)} + \frac{\partial P_t}{\partial(T - t)}\right] e^{r(T-t)} + re^{r(T-t)}(C_t + P_t) > 0. \quad (24)$$