ONLINE APPENDIX

Analyst coverage, information, and bubbles

Sandro C. Andrade* Jiangze Bian† Timothy R. Burch*

Journal of Financial and Quantitative Analysis, forthcoming

*School of Business Administration, University of Miami.
†School of Banking and Finance, University of International Business and Economics.

1 A simple model of bubbles and analyst coverage: Proofs

Proposition 1. ***There is a bubble in the asset price.***

The asset price at date 0 is

$$P_0 = p\left(1+\frac{(2q-1)\,q\,(1-q)\,(1-p)}{(2q-1)^2\,p\,(1-p)+q\,(1-q)}\right) > p\ .$$

Proof:

The asset pays off a dividend $d=1$ or $d=0$ at date 2. The two groups of risk-neutral investors (A and B) have common priors at date 0: the probability that $d=1$ is equal to p. At date 1, investors receive independent signals a and b about the asset's pay-off:

$$\Pr[a=1|d=1]=\Pr[b=1|d=1]=q>\frac{1}{2};$$
$$\Pr[a=0|d=0]=\Pr[b=0|d=0]=q>\frac{1}{2}.$$

Investor group A only considers signal a, and disregards b, and investor group B only considers b, and disregards a. Investors trade at date 1, after receiving the signals, in a market in which there are short-sale constraints.

There are four scenarios for date 1 signals $(a,b)=(0,0)$, $(a,b)=(1,1)$, $(a,b)=(0,1)$, and $(a,b)=(1,0)$. We first find the updated beliefs given the signals in each of the scenarios, and the resulting equilibrium asset price at date 1. Then we find the probability of each one of the four scenarios from date 0 perspective, and the equilibrium price at date 0.

Using Bayes' Rule, we have the following updated beliefs for investor group A:

$$\Pr[d=1|a=1]=\frac{\Pr[a=1|d=1]\ \Pr[d=1]}{\Pr[a=1|d=1]\ \Pr[d=1]\ +\ \Pr[a=1|d=0]\ \Pr[d=0]}=\frac{qp}{qp+(1-q)\,(1-p)};$$
$$\Pr[d=1|a=0]=\frac{\Pr[a=0|d=1]\ \Pr[d=1]}{\Pr[a=0|d=1]\ \Pr[d=1]\ +\ \Pr[a=0|d=0]\ \Pr[d=0]}=\frac{(1-q)\,p}{(1-q)\,p+q\,(1-p)}.$$

The expressions for the updated beliefs of investor group B are analogous, only substituting signal b for signal a.

When a=b=0 or a=b=1, then, since investors are risk-neutral, the corresponding equilibrium asset prices at date 1, denoted $P_{1,(a,b)}$, are equal to:

$$\begin{aligned} P_{1,(0,0)} &= \Pr[d=1|a=0]\times 1+\Pr[d=0|a=0]\times 0=\frac{(1-q)\,p}{(1-q)\,p+q\,(1-p)}\ ; \\ P_{1,(1,1)} &= \Pr[d=1|a=1]\times 1+\Pr[d=0|a=1]\times 0=\frac{qp}{qp+(1-q)\,(1-p)}\ . \end{aligned}$$

If the signals do *not* coincide, then investors agree to disagree. Because there are short-sale constraints, the equilibrium asset price is determined by the beliefs of the most optimistic investor group, i.e., the one which received the signal $d=1$. Therefore, $P_{1,(0,1)}=P_{1,(1,0)}=P_{1,(1,1)}$.

Next we compute the probability of each one of the four different scenarios for date 1 signals, considering that the signals are independently drawn. For example, for the probability of the a=b=0 scenario is:

$$\begin{aligned} \Pr[a &= 0,b=0] \;=\; \Pr[a=0,b=0|d=0]\;\Pr[d=0]\;+\;\Pr[a=0,b=0|d=1]\;\Pr[d=1] \\ &= \Pr[a=0|d=0]\;\Pr[b=0|d=0]\;\Pr[d=0]\;+\;\Pr[a=0|d=1]\;\Pr[b=0|d=1]\;\Pr[d=1] \\ &= q^2\,(1-p)+\left(1-q^2\right)p\ . \end{aligned}$$

Analogously, we find:

$$\begin{aligned} \Pr[a &= 1,b=1] \;=\; q^2p+\left(1-q^2\right)(1-p)\ ; \\ \Pr[a &= 1,b=0] \;=\; \Pr[a=0,b=1] \;=\; q\,(1-q)\ . \end{aligned}$$

The price at date 0, when both investor groups share the same beliefs, is given by the dot product of date 1 scenario probabilities and corresponding equilibrium prices:

$$P_0=\Pr[a=0,b=0]\,P_{1,(0,0)}+\Pr[a=1,b=1]\,P_{1,(1,1)}+\Pr[a=1,b=0]\,P_{1,(1,0)}+\Pr[a=0,b=1]\,P_{1,(0,1)}$$

Substituting in the expressions, and simplifying, yields the P_0 formula in Proposition 1.■

Proposition 2. ***A stronger public information signal results in a smaller bubble.***

The price at date 0 with analyst coverage is $P_0^{analyst} = p\ (1 + f(r))$ *for*

$$f(r) = (2q-1)\,q\,(1-q)\,(1-p)\,r\,(1-r)\,\frac{pq\,(1-q)\,(1-p)\ +\ r\,(1-r)\ \{p\,(1-p)+q\,(1-q)-8pq\,(1-q)\,(1-p)\}}{\{pq+(1-p-q)\,r\}\ \{qr+(1-q-r)\,p\}\ \{pr+(1-p-r)\,q\}\ \{(1-q)\,(1-r)-(1-q-r)\,p\}}\ .$$

The function f(r) is strictly decreasing in r *for* $\frac{1}{2} \leq r \leq 1$. *When* $r = \frac{1}{2}$ *then*

$$f\left(\frac{1}{2}\right) = \frac{(2q-1)\,q\,(1-q)\,(1-p)}{(2q-1)^2\,p\,(1-p)+q\,(1-q)}\ ,$$

that is, $P_0^{analyst}$ *is maximum and equal to* P_0 *in Proposition 1 when the public information signal is not informative.*

Proof:

Both groups of investors (A and B) observe signal c from a stock analyst, in addition to the a and b signals. Investors believe signal c carries information about the asset's payoff at date 2.

$$\Pr[c = 1|d = 1] = r \geq \frac{1}{2}\ ;$$
$$\Pr[c = 0|d = 0] = r \geq \frac{1}{2}\ .$$

There are 2^3=8 scenarios for date 1, given by the combinations of signals a, b, and c. We calculate the updated beliefs given the signals in each of the scenarios, and the resulting equilibrium asset price at date 1. Then we compute the probability of each one of the eight scenarios as of date 0, and the resulting equilibrium price at date 0.

First, using Bayes' Rule and considering that signals are independent, we compute the updated beliefs at date 1. For example, for investor group A when a=1 and c=1, we have:

$$\begin{aligned}
\Pr[d &= 1|a = 1, c = 1] = \frac{\Pr[a = 1, c = 1|d = 1]\ \Pr[d = 1]}{\Pr[a = 1, c = 1|d = 1]\ \Pr[d = 1]\ \Pr[d = 1]\ +\ \Pr[a = 1, c = 1|d = 0]\ \Pr[d = 0]} \\
&= \frac{\Pr[a = 1|d = 1]\ \Pr[c = 1|d = 1]\ \Pr[d = 1]}{\Pr[a = 1|d = 1]\ \Pr[c = 1|d = 1]\ \Pr[d = 1]\ +\ \Pr[a = 1|d = 0]\ \Pr[c = 1|d = 0]\ \Pr[d = 0]} \\
&= \frac{qrp}{qrp + (1-q)\,(1-r)\,(1-p)}\ .
\end{aligned}$$

Similarly, the other investor A expressions are:

$$\begin{aligned}
\Pr[d &= 1|a=1,c=0] = \frac{q\,(1-r)\,p}{q\,(1-r)\,p+(1-q)\,r\,(1-p)}\ ;\\
\Pr[d &= 1|a=0,c=1] = \frac{(1-q)\,rp}{(1-q)\,rp+q\,(1-r)\,(1-p)}\ ;\\
\Pr[d &= 1|a=0,c=0] = \frac{(1-q)\,(1-r)\,p}{(1-q)\,(1-r)\,p+qr\,(1-p)}\ .
\end{aligned}$$

The expressions for the updated beliefs of investor group B are analogous, only substituting signal b for signal a.

There are eight different scenarios for date 1 signals, depending on combinations of the signals a, b and c. The resulting date 1 equilibrium prices, denoted $P_{1,(a,b,c)}$, in each of these scenarios are:

$$\begin{aligned}
P_{1,(0,0,0)} &= \Pr[d=1|a=0,c=0] = \frac{(1-q)\,(1-r)\,p}{(1-q)\,(1-r)\,p+qr\,(1-p)}\ ;\\
P_{1,(0,0,1)} &= \Pr[d=1|a=0,c=1] = \frac{(1-q)\,rp}{(1-q)\,rp+q\,(1-r)\,(1-p)}\ ;\\
P_{1,(1,1,1)} &= P_{1,(1,0,1)} = P_{1,(0,1,1)} = \Pr[d=1|a=1,c=1] = \frac{qrp}{qrp+(1-q)\,(1-r)\,(1-p)}\ ;\\
P_{1,(1,1,0)} &= P_{1,(1,0,0)} = P_{1,(0,1,0)} = \Pr[d=1|a=1,c=0] = \frac{q\,(1-r)\,p}{q\,(1-r)\,p+(1-q)\,r\,(1-p)}\ .
\end{aligned}$$

Next we compute the probability of each one of the eight different scenarios for date 1 signals. For example, for the probability of the a=b=c=0 scenario is:

$$\begin{aligned}
\Pr[a &= 0,b=0,c=0] = \Pr[a=0,b=0,c=0|d=0]\ \Pr[d=0]\ +\ \Pr[a=0,b=0,c=0|d=1]\ \Pr[d=1]\\
&= \left(\begin{array}{c}\Pr[a=0|d=0]\ \Pr[b=0|d=0]\ \Pr[c=0|d=0]\ \Pr[d=0]\\ +\ \Pr[a=0|d=1]\ \Pr[b=0|d=1]\ \Pr[c=0|d=1]\ \Pr[d=1]\end{array}\right)\\
&= q^2r\,(1-p)+\left(1-q^2\right)(1-r)\,p\ .
\end{aligned}$$

Analogously, we find:

$$\begin{aligned}
\Pr[a &= 1, b = 1, c = 0] = \left(1 - q^2\right) r\,(1 - p) + q^2\,(1 - r)\,p\ ; \\
\Pr[a &= 1, b = 1, c = 1] = \left(1 - q^2\right)(1 - r) + q^2 rp\ ; \\
\Pr[a &= 0, b = 0, c = 1] = q^2\,(1 - r)\,(1 - p) + \left(1 - q^2\right) rp\ ; \\
\Pr[a &= 0, b = 1, c = 0] = \Pr[a = 1, b = 0, c = 0] = q\,(1 - q)\left\{r\,(1 - p) + (1 - r)\,p\right\}\ ; \\
\Pr[a &= 0, b = 1, c = 1] = \Pr[a = 1, b = 0, c = 1] = q\,(1 - q)\left\{(1 - r)\,(1 - p) + rp\right\}\ .
\end{aligned}$$

Finally, the price at date 0, when both investor groups share the same beliefs, is given by the dot product of date 1 scenario s probabilities and corresponding equilibrium prices.

$$P_0 = \sum_s \Pr[s] P_{1,(s)}\ .$$

Substituting in the expression found before, and simplifying, results in the expression in Proposition 2.

Substituting r=$\frac{1}{2}$ and r=1 we find:

$$\begin{aligned}
f\left(r = \frac{1}{2}\right) &= \frac{(2q - 1)\,q\,(1 - q)\,(1 - p)}{(2q - 1)^2\,p\,(1 - p) + q\,(1 - q)\,r}\ ; \\
f\,(r = 1) &= f\,(r = 0) = 0\ .
\end{aligned}$$

The first derivative of f(r) is equal to

$$\frac{df}{dr}(r) = K \begin{pmatrix} \frac{1}{(p-q)(pr-q(p+r-1))^2} + \frac{1}{(p+q-1)(p(q+r-1)+(q-1)(r-1))^2} \\ -\frac{1}{(p-q)(qr-p(q+r-1))^2} - \frac{1}{(p+q-1)(-r(p+q)+pq+r)^2} \end{pmatrix},$$

where K is a positive constant. Substituting r=$\frac{1}{2}$ and r=1 we find

$$\begin{aligned}
\frac{df}{dr}\left(r = \frac{1}{2}\right) &= 0\ ; \\
\frac{df}{dr}(r = 1) &= -\frac{1}{q\,(1 - q)\,p\,(1 - p)} < 0\ .
\end{aligned}$$

By inspection, we observe that $\frac{df}{dr}(r)<0$ for $\frac{1}{2}< r \leq 1$.

2 Further robustness for Table 3 regressions

2.1 Different controls for firm size

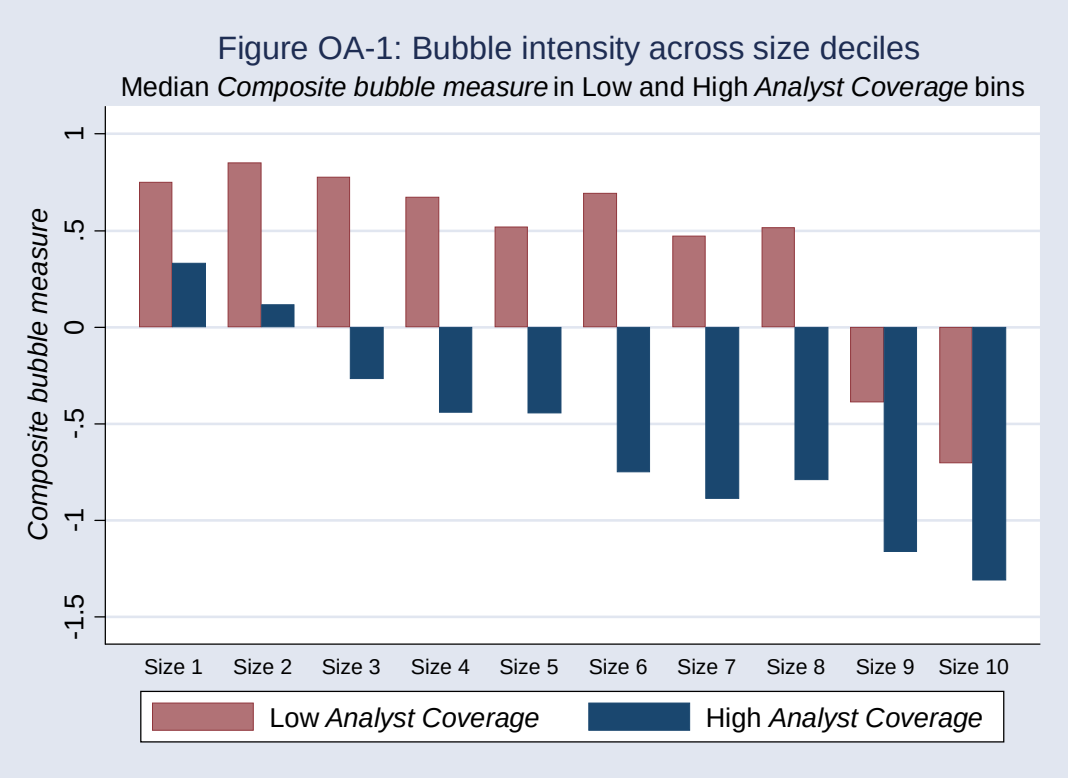
Figure OA-1 provides non-parametric evidence that our key finding is not driven by a positive correlation between analyst coverage and firm size. We first sort stocks into 10 deciles based on market capitalization. Within each size decile, we further sort stocks into two groups based on whether their *Analyst coverage* is above or below the median *Analyst coverage* in each size decile. The figure shows that, within each size decile, the median *Composite bubble measure* is much higher in stocks with low analyst coverage than in stocks with high analyst coverage.

In Panel 1 of Table OA-1 we revisit the regressions in Table 3 while changing the way we control for size. For ease of comparison, in the first row we repeat the baseline results in Table 3. In the second row, we use *Analyst coverage* orthogonalized with respect to *Log of market capitalization* instead of raw *Analyst coverage*. In the third row, we include the square and the cube of *Log of market capitalization* as regressors, in addition to *Log of market capitalization* itself. We also include interactions between *Log of market capitalization* and all the other control variables in Table 3. In the fourth and fifth rows we use market capitalization at the beginning or at the end of the six-month reference period from November 29, 2006 to May 29, 2007, instead of using the average market capitalization within that period. In the next to last row we use only tradable shares to compute market capitalization.[1] Finally, in the last row we control for firm size using total assets rather than market capitalization. Panel 1 shows that the coefficient on *Analyst coverage* is statistically significant at the 1% level in all 28 regressions. All 28 regressions show smaller bubbles in stocks with more analyst coverage, after controlling for firm size.

TABLE OA – 1

In Panel 2 of Table OA-1 we report the linear and Spearman rank correlations between *Analyst coverage* and *Composite bubble measure* within each of 10 size deciles (measuring size as *Log of market capitalization* as in the baseline results) We find an economically large

[1] See Li et al. (2011) for discussion of tradable versus non-tradable shares, as well as the split-share structure reform in Chinese stocks. Of our 623 sample firms, 598 underwent the reform before the beginning of our reference period in November 29, 2006 and 5 were not eligible because all shares were always tradable.



and statistically significant correlation in each size decile. Within each size decile, greater analyst coverage is associated with smaller bubbles.

2.2 I/B/E/S data

In Panel 1 of Table OA-2 we repeat the regressions in Columns (2), (4), (6), and (8) of Table 3 while replacing the Resset-derived *Analyst coverage* with *I/B/E/S analyst coverage*. We define *I/B/E/S analyst coverage* as the number of analysts issuing earnings-per-share forecasts during the reference period of November 29, 2006 to May 29, 2007 according to the I/B/E/S Chinese dataset. This data source is used by Chan and Hameed (2006), among others. We find that the analyst coverage data on I/B/E/S is much less comprehensive than on Resset. Specifically, 250 of our sample stocks are reported with at least one analyst in the *I/B/E/S* data, whereas 453 stocks have at least one analyst covering them according to the Resset data. The correlation between *I/B/E/S analyst coverage* and *Analyst coverage* is 0.78, however, and Panel 1 in Table OA-2 shows that all of our conclusions are robust to using *I/B/E/S analyst coverage* rather than the more comprehensive *Analyst coverage*

variable.

TABLE OA – 2

2.3 Outliers

Panel 2 of Table OA-2 summarizes regressions addressing the concern that our results are driven by outliers in *Analyst coverage*. We repeat the regressions in Columns (2), (4), (6), and (8) of Table 3 while replacing the *Analyst coverage* variable with two dummy variables based on *Analyst coverage*. We define *Any coverage dummy* as an indicator variable set to one when *Analyst coverage* is greater than zero and set to zero otherwise, and similarly define *Many analysts dummy* based on whether the stock is followed by more than six brokerage firms (which is the median *Analyst coverage* for stocks with non-zero coverage). These two dummies partition firms in three groups: 170 stocks with *Analyst coverage* equal to zero, 227 stocks with *Analyst coverage* between 1 and 6, and 226 stocks with *Analyst coverage* greater than 6. Panel 2 shows that the two dummies are positive and statistically significant in all specifications but that explaining *Cumulative return*, in which only the *Many analysts dummy* is statistically significant.

Panel 3 of Table OA-2 summarizes regressions addressing the concern that our results are driven by outliers in the dependent variables. We repeat the specifications in Columns (2), (4), (6), and (8) of Table 3 while using median regressions rather than ordinary least squares. We find that the coefficient on *Analyst coverage* remains statistically significant at the 1% level in all four regressions.

2.4 Additional control variables

In Panel 4 of Table OA-2 we summarize the results of adding a number of explanatory variables to our baseline specification explaining *Composite bubble measure*. We find that *Analyst coverage* remains highly statistically and economically significant in all of the seven specifications.

In the first specification we add *Ratio non-tradable/tradable*, the average ratio of non-tradable to tradable shares in each stock in the reference period. This variable accounts for the fact

that a considerable number of outstanding shares are not tradable in the secondary market in China, and addresses the concern that it is not clear on which basis (all shares or tradable shares) one should define *Market capitalization* as a control variable for *Analyst coverage*. In the second column we add *Share float*, the average number of tradable shares in the reference period of November 29, 2006 to May 29, 2007, in billions (results are also robust to using the log of *Share float*). This is motivated by Hong, Scheinkman, and Xiong (2006), who propose a theory in which bubble magnitudes are negatively related to a stock's float.

In the third specification we add *Contemporaneous return volatility*, the average daily return squared during the reference period of November 29, 2006 to May 29, 2006. Scheinkman and Xiong's (2003) theory predicts larger bubbles in stocks with more volatile fundamentals, so this variable controls for the possibility that for some reason analysts are less likely to cover more volatile stocks. In the next specification we add *Turnover trend*. This variable is calculated as the slope coefficient on a regression of daily turnover in the reference period on a time trend variable. This addresses the concern that turnover is non-stationary during the six-month reference period (see Figure 2 in the main paper), and hence that its population average is not well defined and may be misrepresented by the sample average. In the fifth specification we add *Number of trades per day*, an alternative measure of trading activity. In the sixth column we add the loadings of three empirical factors constructed from daily returns.[2] In the last specification we include all of the additional explanatory variables.[3]

2.5 Placebo periods

To investigate whether our results obtain in all periods rather than only in the bubble period we study, we repeat our main specifications in placebo, non-overlapping six-month periods far away from May 30, 2007. To make sure these placebo periods are "normal" and thus not part of the bubble inflating-deflating phenomenon, we discard the six-month periods immediately before and immediately after our reference period of November 29, 2006 to May 29, 2007. We examine four placebo periods, two earlier ones and two later ones. Both

[2] To construct the factor loadings, we perform a factor analysis of the daily returns of the sample stocks in the pre-tax-increase period, and retain loadings on the first three factors (Roll and Ross, 1980). The first factor is overwhelmingly dominant, accounting for 39.4% of the covariation in the data. The second and third factors account for 3.1% and 2.1%, respectively, with additional factors individually accounting for less than 2.1%.

[3] The correlation between *Share float* and *Analyst coverage* is 0.42, while that between *Share float* and *Log of market capitalization* is 0.64. The correlation between *Daily turnover* and *Turnover trend* is 0.68. The correlation between *Loading on empirical factor 1* and *Market beta* is 0.72, and is the highest correlation between the two betas and the empirical factor loadings.

dependent and independent variables are redefined with data during the time period being studied.

We focus on specifications (2) and (4) of Table 3 that explain *Cumulative return* and *P/E ratio*, respectively. We do not repeat the specification in Column (6) because it concerns announcement returns following the May 30, 2007 transaction tax tripling. Also, we do not estimate the specification in Column (8) because it is not appropriate to define the first principal component of *Cumulative return* and *P/E ratio* during the placebo periods we use. In three of the four placebo periods the correlation coefficients between *Cumulative return* and *P/E ratio* are small and negative (ranging from -0.067 to -0.096), which leads to problems in how to interpret the correlation between the first principal component and *Analyst coverage*.[4,5] In contrast, during the reference period of November 29, 2006 to May 29, 2007, *Cumulative return* and *P/E ratio* are strongly positively correlated ($\rho = 0.316$), and hence interpreting the correlation between their first principal component and *Analyst coverage* is straightforward.

Panel 5 of Table OA-2 shows that the reference period is the only period in which regressions explaining both *Cumulative return* and *P/E ratio* have negative and statistically significant coefficient estimates for *Analyst coverage*. In contrast to the reference period, the coefficient on *Analyst coverage* in regressions explaining *P/E ratio* is statistically insignificant in all four placebo periods. For the regressions explaining *Cumulative return,* we find that the coefficient on *Analyst coverage* is actually *positive* in three of the four placebo periods (significantly so in the first), and negative and statistically significant in only the placebo periods that begins 12 months after the reference period. However, as we explain below, this result is not particularly robust.

It turns out that the sign and significance of *Analyst coverage* in this placebo period varies, depending on which control variables are used. For example, as reported in Table OA3, in this placebo period the coefficient on *Analyst coverage* is positive and statistically insignificant when the log of *Market capitalization* is the unique control variable, and positive and statistically significant at the 5% level (t-statistic=2.41) when both the log of *Market cap-*

[4] If *P/E ratio* and *Cumulative return* are negatively related, then we can represent the first principal component between them as FPC = α*P/E ratio* - β*Cumulative return,* where α and β are positive. Therefore, if we find a negative correlation between FPC and *Analyst coverage*, it could actually be the result of a positive correlation between *Cumulative return* and *Analyst coverage,* which in any event would make it necessary to examine the correlations between *Cumulative return, P/E ratio* and *Analyst coverage* separately for interpretation guidance.

[5] The correlation between Cumulative return and P/E ratio in the fourth placebo period is small and positive (ρ = 0.068).

italization and *Turnover* are included as control variables. This stands in contrast to the reference period, in which *Analyst coverage* is consistently negative and statistically significant for all specifications that we tried. Overall, we conclude that the results in the reference period are not reproduced in the placebo periods.

TABLE OA – 3

2.6 Different combinations of control variables

In Table OA-4 we report some of the intermediate specifications that include less than the full set of control variables. Regressions in which we sequentially add control variables in their listed order are reported in Panels 1, 3, 5 and 7. In the remaining panels (2, 4, 6 and 8) we begin with the full list of control variables and then sequentially remove control variables in their listed order. *Analyst Coverage* remains strongly statistically and economically significant across all regressions.

TABLE OA – 4

3 Further robustness for Table 4 regressions

3.1 First stage regressions

Table OA-5 shows results of the first stage regressions associated with the 2SLS estimation of Table 4. The first stage results indicate that Table 4 regressions do not suffer from a weak instrument problem.

TABLE OA – 5

3.2 Other bubble intensity measures

Table 4 of the paper only reports 2SLS regressions explaining *Composite bubble measure*. In Table OA-6 we repeat our analyses using the other bubble measures (*Cumulative return, P/E ratio*, and *Announcement return).* Columns (1) through (3) show that *Analyst coverage in 2005* remains statistically significant at the 1% level in the regressions. Columns (4) through (6) show that *Analyst coverage* remains statistically significant in the 2SLS regressions. Therefore, based on the results of instrumental variable estimations, we conclude that it is unlikely that our results are driven by an omitted, slow-moving bubble-proneness variable with which *Analyst coverage* is endogenously correlated.

TABLE OA – 6

3.3 One instrument at a time

In Table OA-7 we present two-stage least squares regressions of *Cumulative return*, *P/E ratio*, *Announcement return*, and *Composite bubble measure* in which *Analyst coverage* is instrumented by one instrumental variable at a time (either *Trading volume in 2005* (Panel 1) or *Mutual fund ownership in June 2005* (Panel 2)). *Analyst coverage* remains statistically significant in all cases, except for the *P/E ratio* regression in which *Trading volume in 2005* is the sole instrumental variable.

TABLE OA – 7

3.4 Adding instruments to the RHS of Table III regressions

In Table OA-8 we show that *Analyst coverage* remains strongly statistically significant in OLS regressions in which both *Trading volume in 2005* and *Mutual fund ownership in June 2005* are added as regressors. These variables are statistically insignificant in those regressions, except for *Mutual fund ownership in June 2005*, which is borderline statistically significant (t-value=-1.67) in the regressions explaining *Cumulative return*.

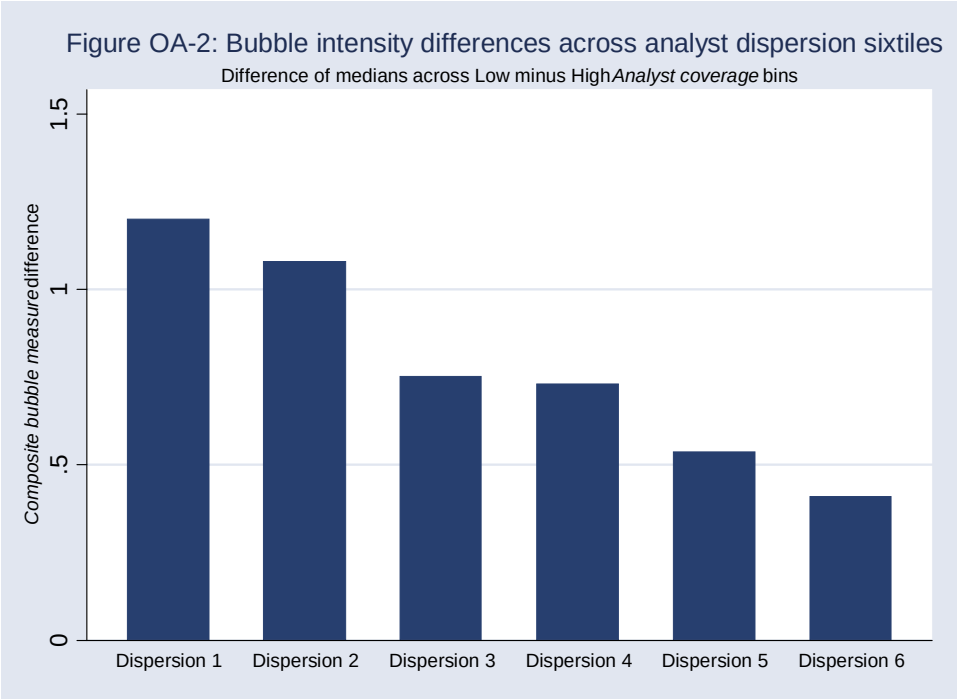
TABLE OA – 8

4 Further robustness for Table 6 regressions

4.1 Figure illustrating main regression result

Figure OA-2 illustrates that analyst coverage is indeed less effective in reducing bubble intensity when there is greater disagreement among analysts. We first sort stocks into sixtiles based *Dispersion among analysts*. Within each sixtile we further categorize stocks into high and low analyst coverage groups, based on whether the stock's analyst coverage is above or below the overall sample median. We then compute the median *Composite bubble measure* for each analyst coverage group within the sixtile and plot the difference between the medians. For example, the bar for dispersion group 1 (the smallest analyst dispersion group) is the median *Composite bubble measure* for its high analyst coverage subgroup minus the median *Composite bubble measure* for its low analyst coverage subgroup.

Figure OA-2 shows that the difference of bubble intensity across low and high analyst coverage bins is positive in all analyst dispersion sixtiles, which confirms our key finding that stocks with high analyst coverage develop smaller bubbles. The additional finding the figure illustrates is that the difference in bubble intensity among Low and High *Analyst coverage* bins decreases mononotically as the level of disagreement among analysts increases from sixtile 1 to sixtile 6. That is, analyst coverage is less effective in reducing bubble intensity when there is greater disagreement among analysts.



4.2 Other bubble intensity and dispersion measures

In Table OA-9 we show that the interaction term results using *Dispersion among analysts* obtains for two of the other three full sample bubble intensity measures (*Cumulative return* and *Announcement return*, but not *P/E ratio*).

TABLE OA – 9

In Table OA-10 we report results of *Composite bubble measure* regressions in which we interact *Analyst coverage* either with *Dispersion of analysts' earnings forecasts* (Panel 1) or with *Dispersion of analysts' recommendations* (Panel 2), rather than with *Dispersion among analysts*. In both cases we observe that the interaction term is positive and statistically and economically significant. This shows that our conclusion that *Analyst coverage* is less effective in mitigating bubbles when there is high disagreement among analysts is robust to measuring disagreement among analysts by using only their earnings forecasts or their buy/sell recommendations.

TABLE OA – 10

In columns (4) through (6) of Table OA-10 we report results of regressions that explain *Turnover* in which we interact *Analyst coverage* either with *Dispersion of analysts' earnings forecasts* (Panel 1) or with *Dispersion of analysts' recommendations* (Panel 2), rather than with *Dispersion among analysts*. In both cases we observe that the interaction term is positive and economically significant. The interaction coefficients are statistically significant both for *Dispersion of analysts' earnings forecasts* and for *Dispersion of analysts' recommendations* when the control variables are omitted (model (5) in Panels 1 and 2). As shown in model (6), the interaction term remains statistically significant when all control variables are included in a regression using *Dispersion of analysts' earnings forecasts* (the interaction term's t-statistic is 1.85), but not when we use *Dispersion of analysts' recommendations* (the interaction term's t-statistic is 1.32).

4.3 Full sample of stocks

Because of our use of the *Dispersion among analysts* variable, our Table 6 *Turnover* regressions are limited to a subsample of 364 stocks with *Analyst coverage* of 2 or more. In Table OA-11 we present *Turnover* regressions for the full sample of 623 stocks. These regressions show that the greater *Analyst coverage* is associated with lower *Turnover*, and that the effect is statistically and economically significant.

TABLE OA – 11

5 Additional figures

Figure 2 of the paper suggests a reference period ending May 29, 2007, based on P/E ratios, turnover, cumulative returns, and two measures of retail investor enthusiasm (Google searches and account openings). For completeness, here we plot two additional figures. We show price indices for our sample of 623 A-shares in the Shanghai Stock Exchange. To calculate the indices, for each stock we first accumulate the gross return since January 2005, normalizing to 1 on November 28 2006, right before our reference period begins. We then calculate both the median and value-weighted average across all 623 stocks.

As Figure 2, Figure OA-3 suggests a regime change on May 30, 2007. Though the peak for median price levels is on January 2008, it is clear that not only did the average rate of price appreciation slow substantially after May 30, 2007, but in addition prices did not display a clear upward trend as they did beforehand.

Figure OA-4 plots the value-weighted average P/E ratio of Shanghai stocks, in addition to the median P/E ratio previously plotted in Figure 2 of the paper. Figure OA-4 also shows a regime change after May 30, 2007, with both median and value-weighted P/E ratios declining thereafter.[6]

[6] One argument for placing greater emphasis on plots of median (as opposed to value-weighted) prices and valuation ratios is that, as we show later, bubble magnitudes are negatively correlated with firm size. Hence, value-weighted plots present a somewhat skewed picture in the sense of not being representative of a randomly picked firm.

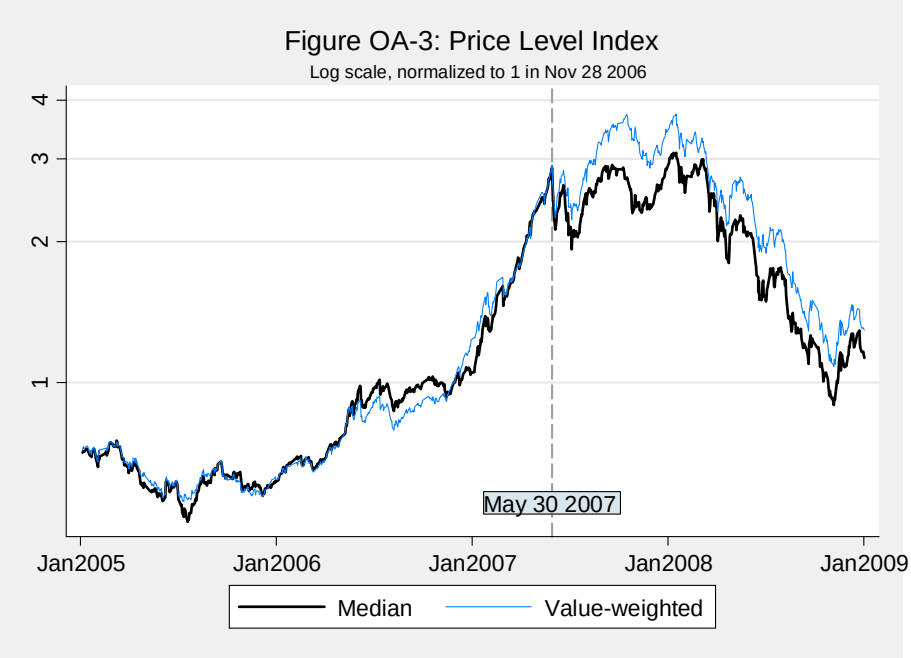


Figure OA-3: Price Level Index
Log scale, normalized to 1 in Nov 28 2006
4
3
2
1
May 30 2007
Jan2005
Jan2006
Jan2007
Jan2008
Jan2009
Median
Value-weighted

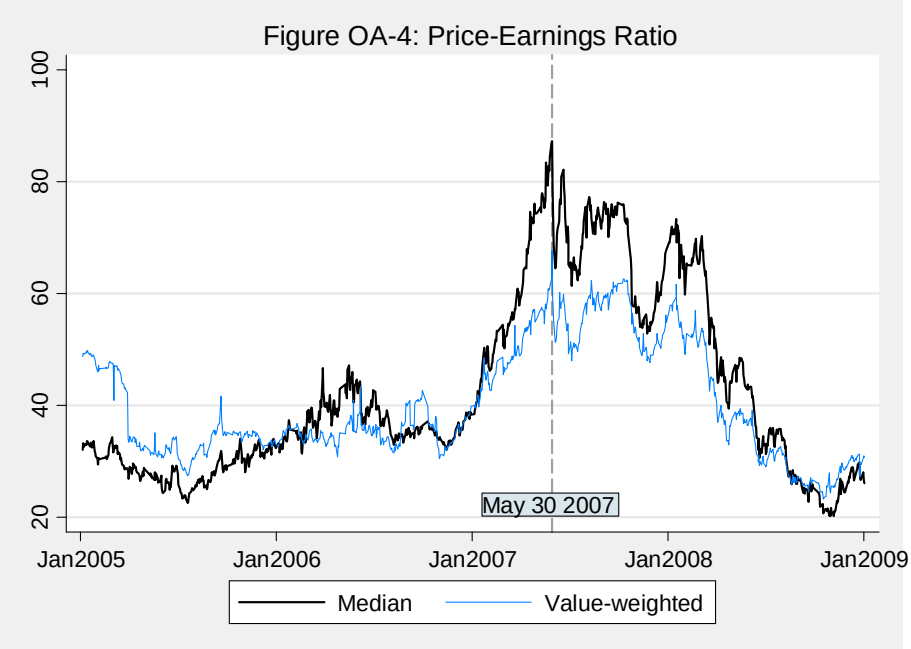


Figure OA-4: Price-Earnings Ratio
100
80
60
40
20
May 30 2007
Jan2005
Jan2006
Jan2007
Jan2008
Jan2009
Median
Value-weighted

6 Brokerages in RESSET data

Table OA-12 lists the Chinese brokerage firms providing earnings-per-share forecasts for the sample stocks during the six-month reference period.

TABLE OA – 12

Table OA-1. Size related robustness checks

Panel 1 summarizes key results from robustness regressions for Table 3specifications that explain bubble intensity measures in a sample of 623 Shanghai A-shares. We report the coefficient on *Analyst coverage* across 28 regressions. Each regression has a different type of control for size. In the first row we repeat our baseline results in Columns (1), (3), (5), and (7) of Table 3, which use size as the log of the average market capitalization (using total number of shares) in the six-month reference period of November 29, 2006 to May 29, 2007. In the second row we first orthogonalize Analyst coverage with respect to log of market capitalization before including it as an explanatory variable. In the third row we include the square and the cube of log of market capitalization, as well as its interactions with all the other control variables in Table 3. In the fourth and fifth rows we measure market capitalization at the beginning or at the end of the reference period, rather than the average across the period. In the penultimate row we only use tradable shares when computing market capitalization. In the last row we use total assets rather than market capitalization. The coefficients on control variables (and the constant term) are not reported for brevity. We report heteroskedasticity-robust t-statistics in parentheses beneath variable coefficients. Panel 2 reports the linear and the Spearman rank correlations between Analyst coverage and Composite bubble measure within each log of market capitalization decile. ***, **, * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel 1: Coefficient on *Analyst Coverage* in different regressions related to size

	Dependent Variable			
Test	Cumul. return	P/E ratio	Ann. Return	Comp. bubble meas.
Baseline in Table 3 - Columns (1), (3), (5), and (7)	-5.392 *** (-13.37)	-5.086 *** (-15.91)	0.997 *** (21.46)	-0.097 *** (-26.84)
Analyst coverage orthogonalized with respect to Log of market Capitalization	-8.124 *** (-10.03)	-5.021 *** (-9.46)	0.908 *** (10.71)	-0.105 *** (-14.39)
Including Log of market capitalization & and its powers & interactions with all other control vars. in Table 3	-5.407 *** (-8.68)	-4.243 ** (-7.94)	0.779 *** (10.27)	-0.083 *** (-13.99)
Log of market capitalization measured at beginning of ref. period rather than average	-5.581 *** (-7.98)	-4.721 *** (-8.95)	0.888 *** (12.44)	-0.091 *** (-15.57)
Log of market capitalization measured at end of ref. period rather than average	-9.657 *** (-12.98)	-5.251 *** (-10.99)	0.949 *** (15.41)	-0.115 *** (-20.50)
Log of market capitalization measured using tradable rather than total shares	-9.009 *** (-11.21)	-4.396 *** (-8.28)	0.921 *** (14.75)	-0.106 *** (-17.52)
Log of total assets as the size variable (instead of Log of market capitalization)	-7.311 *** (-13.04)	-4.306 *** (-10.53)	1.030 *** (18.58)	-0.103 *** (-22.27)

Panel 2: Correlations between *Composite bubble measure* and *Analyst Coverage*

Log of market capitalization deciles	Linear correlation	Rank correlation
Decile 1	-0.33 ***	-0.25 ***
Decile 2	-0.30 ***	-0.35 ***
Decile 3	-0.46 ***	-0.51 ***
Decile 4	-0.63 ***	-0.69 ***
Decile 5	-0.58 ***	-0.61 ***
Decile6	-0.68 ***	-0.80 ***
Decile 7	-0.61 ***	-0.61 ***
Decile 8	-0.67 ***	-0.72 ***
Decile 9	-0.57 ***	-0.59 ***
Decile 10	-0.63 ***	-0.56 ***

Table OA-2. Robustness checks for regressions explaining bubble intensity measures

This table summarizes key results from robustness regressions for Table 3 specifications that explain bubble intensity measures in a sample of 623 Shanghai A-shares. In Panels 1 through 6, unless otherwise noted, all variables are averages across the six-month reference period of November 29, 2006 to May 29, 2007, calculated from daily data. In Panel 5 the variables are averages in other non-overlapping six month periods. The regressions include, but we do not report below, all of the other variables included in the Table 3 model to which each panel. *I/B/E/S Analyst coverage* is the number of brokerage firms issuing earnings-per-share forecasts during the reference period according to I/B/E/S. *Any coverage dummy* equals 1 if the Resset-derived *Analyst coverage* exceeds 0 (and equals 0 otherwise), while *Many analysts dummy* equals 1 if *Analyst coverage* exceeds 6 (and equals 0 otherwise). *Log of market cap.: Higher order and interactions* denotes the inclusion of nine additional control variables: the square and the cube of *Log of market capitalization*, and interactions between *Log of market capitalization* and the other seven control variables in Table 3. *Share float* is the number of tradable shares (in billions). *Return volatility* is the (annualized) standard deviation of daily stock returns in the reference period. *Turnover trend* is the slope coefficient of a regression of daily turnover on a time trend and a constant, during the reference period. *Number of traders per day* is the number of recorded trades per day. *Loading on empirical factors 1* (or *2* or *3*) are coefficients on regressions of daily returns in the pre-tax-increase reference period onto the first three factors obtained from a factor analysis of returns during the reference period. We report heteroskedasticity-robust t-statistics in parentheses beneath variable coefficients, and ***, **, * denote statistical significance at the 1%, 5%, and 10% levels, respectively. All regressions have 623 observations.

Panel 1: Define analyst coverage according to the the number of analysts issuing EPS forecasts as reported in the I/B/E/S dataset for China.

	Dependent Variable			
	Cumul. return	P/E ratio	Ann. Return	Comp. bubble meas.
I/B/E/S Analyst coverage	-6.811 ***	-1.212 **	0.897 ***	-0.077 ***
	(-3.82)	(-2.04)	(3.33)	(-4.63)
Other expl. variables in Table III	yes	yes	yes	yes
Adjusted-R^2	0.43	0.86	0.45	0.71

Panel 2: Use analyst coverage indicator variables for the Reset-derived analyst coverage to address outlier concerns. *Any coverage dummy* = 1 when at least one analyst issues coverage (and = 0 otherwise), and *Many analysts dummy* = 1 when more then six analysts isssue coverage (and = 0 otherwise).

	Dependent Variable			
	Cumul. return	P/E ratio	Ann. Return	Comp. bubble meas.
Any coverage dummy	-9.819	-6.997 *	2.743 ***	-0.204 ***
	(-1.16)	(-1.92)	(3.50)	(-3.40)
Many analysts dummy	-41.012 ***	-10.218 ***	6.488 ***	-0.527 ***
	(-4.84)	(-3.24)	(7.13)	(-8.36)
Other expl. variables in Table III	yes	yes	yes	yes
Adjusted-R^2	0.45	0.87	0.49	0.73

Panel 3: Use median regressions to address outlier concerns.

	Dependent Variable			
	Cumul. return	P/E ratio	Ann. Return	Comp. bubble meas.
Analyst coverage	-3.341 ***	-0.328 ***	0.811 ***	-0.056 ***
	(-5.03)	(-2.72)	(9.79)	(-11.27)
Other expl. variables in Table III	yes	yes	yes	yes
Pseudo-R^2	0.31	0.71	0.36	0.54

Table OA-2
(continued)

Panel 4: Include additional control variables (below, coefficients and t-statistics for *Analyst coverage* as well as for the additional control variables are shown).

Dep. var: Comp. bubble measure	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Analyst coverage	-0.054 ***	-0.059 ***	-0.053 ***	-0.059 ***	-0.054 ***	-0.049 ***	-0.035 ***
	(-10.85)	(-12.07)	(-10.99)	(-11.75)	(-10.06)	(-9.71)	(-6.87)
Ratio non-tradable/tradable	0.033 **						0.041 ***
	(2.09)						(2.90)
Share float		0.086 *					0.128 *
		(1.74)					(1.84)
Contemporaneous return volatility			0.037 ***				0.046 ***
			(8.03)				(10.47)
Turnover trend				-0.351			3.881 *
				(-0.16)			(1.95)
Number of trades per day					0.489 ***		0.775 ***
					(2.85)		(4.34)
Loading on Empirical Factor 1						0.168	-0.929 ***
						(0.71)	(-3.83)
Loading on Empirical Factor 2						-0.628 ***	-0.564 ***
						(-3.30)	(-2.91)
Loading on Empirical Factor 3						-0.754 ***	-0.218
						(-3.49)	(-1.21)
Other expl. variables in Table III	yes	yes	yes	yes	yes	yes	yes
Adjusted-R^2	0.75	0.75	0.78	0.64	0.75	0.76	0.81

Panel 5: Placebo test - Use alternative non-overlapping 6-month periods

	6-month period starting 18 months after reference period begins		6-month period starting 12 months after reference period begins		**Reference period (Nov 29, 2006 to May 29, 2007)**		6-month period starting 12 months after reference period ends		6-month period starting 18 months after reference period ends	
	Dep. Var.		Dep. Var.		**Dep. Var.**		Dep. Var.		Dep. Var.	
	Cum. return	P/E ratio	Cum. return	P/E ratio	**Cum. return**	**P/E ratio**	Cum. return	P/E ratio	Cum. return	P/E ratio
Analyst coverage	0.489 **	-0.702	0.393	-0.115	**-4.628 *** **	**-0.843** ***	-0.694 ***	0.023	0.049	0.020
	(1.97)	(-1.42)	(0.82)	(-0.44)	**(-7.56)**	**(-3.59)**	(-3.09)	(0.09)	(0.35)	(0.06)
Other expl. variables in Table III	yes	yes	yes	yes	**yes**	**yes**	yes	yes	yes	yes
Observations	612	612	612	612	**623**	**623**	620	620	619	619
Adjusted-R^2	0.24	0.62	0.23	0.81	**0.47**	**0.87**	0.43	0.84	0.35	0.69

Table OA-3. Intermediate specifications for regressions explaining *Cumulative return* in a placebo period starting 12 months after the end of the reference period

This table is related to Panel 5 in Table A.3 in Appendix A. We report ordinary least squares regressions that explain *Cumulative return* in the placebo 6-month period that starts 12 months after the end of the paper's reference period (November 29, 2006 to May 29, 2007). All variables are defined in the placebo period. We report heteroskedasticity-robust t-statistics in parentheses beneath variable coefficients, and ***, **, * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable: Cumulative return	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Analyst coverage	-0.303 **	0.127	0.509 **	-0.053	-0.184	-0.353	-0.393 *	-0.695 ***
	(-2.17)	(0.57)	(2.41)	(-0.23)	(-0.79)	(-1.54)	(-1.74)	(-3.09)
Log of market capitalization		-3.598 ***	-0.492	0.986 ***	0.716	8.301 ***	7.942 ***	4.438 **
		(-2.94)	(-0.41)	(0.75)	(0.54)	(4.08)	(4.07)	(2.36)
Turnover			12.921 ***	13.219 ***	13.355 ***	17.286 ***	17.413 ***	13.787 ***
			(8.829)	(9.47)	(9.35)	(10.67)	(11.28)	(8.05)
Lagged return volatility				-0.466 ***	-0.424 ***	-0.312 ***	-0.076	-0.058
				(-5.04)	(-4.47)	(-3.47)	(-0.73)	(-0.56)
Lagged P/E ratio					-0.040 ***	-0.067 ***	-0.076	-0.054
					(-2.84)	(-4.78)	(-5.25)	(-3.88)
Effective Spread						1.462 ***	1.202 ***	0.907 ***
						(5.91)	(4.93)	(3.39)
Depth						-15.407 *	-20.039 ***	-8.680
						(-1.94)	(-5.32)	(-1.17)
Market beta							-30.686 ***	-39.301 ***
							(-3.95)	(-4.85)
Liquidity beta							42.341 ***	50.130 ***
							(4.99)	(5.64)
ΔTurnover								-2.365
								(-1.32)
ΔEffective spread								-1.281 ***
								(-7.95)
Industry effects	no	no	no	yes	yes	yes	yes	yes
Constant	-5.428 ***	-1.676 ***	-28.4 ***					
	(-3.94)	(-0.98)	(-9.12)					
Observations	623	623	623	623	623	623	623	623
Adjusted-R^2	0.01	0.02	0.20	0.27	0.28	0.33	0.36	0.43

Table OA-4 Explaining bubble intensity measures with additional specifications from Table 3

This table is related to Table 3 in the main paper, and reports ordinary least squares regressions that explain four different measures of bubble intensity for a sample of 623 Shanghai A-shares. The variable definitions are the same as in Table 3. We report heteroskedasticity-robust t-statistics in parentheses beneath variable coefficients, and ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel 1: Dependent variable is *Cumulative return* (adding control variables in their listed order)

Dependent Variable: Cumulative return	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Analyst coverage	-5.392 ***	-8.124 ***	-7.068 ***	-7.325 ***	-6.395 ***	-6.014 ***	-5.788 ***	-4.628 ***
	(-13.37)	(-10.38)	(-9.78)	(-10.29)	(-8.75)	(-8.35)	(-8.46)	(-7.56)
Log of market capitalization		23.788 ***	29.877 ***	35.165 ***	34.706 ***	59.391 ***	53.252 ***	28.228 ***
		(4.26)	(5.10)	(5.82)	(5.77)	(8.69)	(8.50)	(4.01)
Turnover			19.509 ***	20.063 ***	19.732 ***	32.296 ***	28.048 ***	21.586 ***
			(5.45)	(5.62)	(5.65)	(9.08)	(8.12)	(4.48)
Lagged return volatility				-0.413 *	-0.742 **	-0.624 **	-0.601 **	-0.296
				(-1.71)	(-2.49)	(-2.28)	(-2.25)	(-1.44)
Lagged P/E ratio					0.214 ***	0.039	0.058	0.070
					(4.05)	(0.74)	(1.11)	(1.50)
Effective Spread						7.404 ***	7.463 ***	3.933 ***
						(8.10)	(8.32)	(4.31)
Depth						-40.779 ***	-40.779 ***	-29.630 ***
						(-5.34)	(-5.40)	(-4.65)
Market beta							40.104 **	21.878
							(1.99)	(1.12)
Liquidity beta							39.396 ***	41.564 ***
							(2.70)	(3.04)
ΔTurnover								4.397
								(0.67)
ΔEffective spread								-6.733 ***
								(-6.17)
Industry effects	no	no	no	yes	yes	yes	yes	yes
Constant	237.1 ***	223.4 ***	156.6 ***					
	(47.42)	(44.13)	(11.66)					
Observations	623	623	623	623	623	623	623	623
Adjusted-R^2	0.16	0.19	0.24	0.26	0.29	0.39	0.41	0.47

Table OA-4
(continued)

Panel 2: Dependent variable is *Cumulative return* (removing control variables in their listed order)

Dependent Variable: Cumulative return	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Analyst coverage	-4.628 ***	-3.232 ***	-4.166 ***	-4.215 ***	-4.461 ***	-5.244 ***	-5.061 ***	-5.392 ***
	(-7.56)	(-6.51)	(-9.28)	(-9.44)	(-10.10)	(-14.50)	(-14.30)	(-13.37)
Log of market capitalization	28.228 ***							
	(4.01)							
Turnover	21.586 ***	10.982 ***						
	(4.48)	(2.94)						
Lagged return volatility	-0.296	-0.203	-0.270					
	(-1.44)	(-0.99)	(-1.30)					
Lagged P/E ratio	0.070	0.098 **	0.112 **	0.080 *				
	(1.50)	(2.13)	(2.42)	(1.84)				
Effective Spread	3.933 ***	1.850 **	1.386 *	1.379 *	1.780 **			
	(4.31)	(2.32)	(1.70)	(1.79)	(2.35)			
Depth	-29.630 ***	-8.435 **	-7.762 *	-12.661 ***	-11.936 ***			
	(-4.65)	(-2.18)	(-1.95)	(-3.66)	(-3.23)			
Market beta	21.878	22.929	30.054	35.493 *	31.514	14.620		
	(1.12)	(1.15)	(1.51)	(1.79)	(1.58)	(0.71)		
Liquidity beta	41.564 ***	46.404 ***	45.563 ***	43.538 ***	44.854 ***	46.809 ***		
	(3.04)	(3.28)	(3.19)	(3.14)	(3.25)	(3.53)		
ΔTurnover	4.397	-3.462	-15.238 ***	-14.909 ***	-15.681 ***	-14.889 ***	-16.167 ***	
	(0.67)	(-0.56)	(-3.20)	(-3.10)	(-3.22)	(-3.04)	(-3.41)	
ΔEffective spread	-6.733 ***	-8.976 ***	-9.558 ***	-9.778 ***	-9.782 ***	-10.296 ***	-10.998 ***	
	(-6.17)	(-9.49)	(-10.18)	(-10.78)	(-10.69)	(-12.20)	(-12.27)	
Industry effects	yes	yes	yes	no	no	no	no	no
Constant				144.5 ***	147.0 ***	201.8 ***	207.6 ***	237.1 ***
				(5.22)	(5.26)	(10.08)	(38.12)	(47.42)
Observations	623	623	623	623	623	623	623	623
Adjusted-R^2	0.47	0.46	0.45	0.43	0.42	0.41	0.39	0.16

Table OA-4
(continued)

Panel 3: Dependent variable is *P/E ratio* (adding control variables in their listed order)

Dep. Variable: P/E ratio	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Analyst coverage	-5.086 ***	-5.021 ***	-4.874 ***	-4.615 ***	-0.908 ***	-0.820 ***	-0.835 **	-0.843 ***
	(-15.91)	(-9.84)	(-9.14)	(-8.94)	(-4.00)	(-3.65)	(-3.71)	(-3.59)
Log of market capitalization		-0.562	0.283	2.138	0.309	3.962 ***	4.544 **	9.455 ***
		(-0.16)	(0.08)	(0.52)	(0.20)	(2.01)	(2.27)	(3.66)
Turnover			2.709	2.794	1.472	3.528 ***	4.157 ***	8.954 ***
			(0.88)	(0.94)	(1.24)	(2.74)	(3.00)	(4.03)
Lagged return volatility				1.391 ***	0.080	0.110	0.116	0.098
				(3.97)	(0.99)	(1.33)	(1.33)	(1.23)
Lagged P/E ratio					0.853 ***	0.820 ***	0.814 ***	0.811 ***
					(52.37)	(44.97)	(43.51)	(43.12)
Effective spread						1.341 ***	1.264 ***	1.911 ***
						(4.93)	(4.50)	(4.78)
Depth						-4.468 **	-4.655 **	-7.023 ***
						(-2.24)	(-2.27)	(-3.02)
Market beta							-10.171 ***	-7.100
							(-1.01)	(-0.68)
Liquidity beta							1.466	-0.227
							(0.26)	(-0.04)
ΔTurnover								8.127 ***
								(2.80)
ΔEffective spread								1.151 **
								(2.37)
Industry effects	no	no	no	yes	yes	yes	yes	yes
Constant	125.1 ***	125.4 ***	116.1 ***					
	(29.16)	(26.39)	(9.96)					
Observations	623	623	623	623	623	623	623	623
Adjusted-R^2	0.20	0.20	0.20	0.30	0.86	0.87	0.87	0.87

Table OA-4
(continued)

Panel 4: Dependent variable is *P/E ratio* (removing control variables in their listed order)

Dependent Variable: P/E ratio	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Analyst coverage	-0.843 ***	-0.376 *	-0.836 ***	-0.875 ***	-3.436 ***	-4.743 ***	-5.028 ***	-5.086 ***
	(-3.59)	(-1.80)	(-5.61)	(-5.96)	(-8.26)	(-15.19)	(-15.45)	(-15.91)
Log of market capitalization	9.455 ***							
	(3.66)							
Turnover	8.954 ***	5.403 ***						
	(4.03)	(3.00)						
Lagged return volatility	0.098	0.129	0.096					
	(1.23)	(1.44)	(1.06)					
Lagged P/E ratio	0.811 ***	0.821 **	0.828 ***	0.837 ***				
	(43.12)	(44.41)	(45.34)	(49.42)				
Effective Spread	1.911 ***	1.213 ***	0.985 ***	1.037 ***	5.218 ***			
	(4.78)	(4.05)	(3.40)	(3.81)	(6.99)			
Depth	-7.023 ***	0.076	0.478	1.237	8.805			
	(-3.02)	(0.05)	(0.33)	(1.01)	(1.64)			
Market beta	-7.100	-6.748	-3.243	-3.123	-44.642 ***	-102.598 ***		
	(-0.68)	(-0.64)	(-0.31)	(-0.30)	(-2.70)	(-7.23)		
Liquidity beta	-0.227	1.393 ***	0.980	0.187	13.930	33.751 ***		
	(-0.04)	(0.24)	(0.17)	(0.03)	(1.29)	(3.10)		
ΔTurnover	8.127 ***	5.495 **	-0.299	-0.721	-8.771 **	-9.774 **	-2.924	
	(2.80)	(2.07)	(-0.17)	(-0.41)	(-2.09)	(-2.09)	(-0.61)	
ΔEffective spread	1.151 **	0.400 ***	0.113 ***	0.193	0.112	-2.254 ***	-1.844 **	
	(2.37)	(1.05)	(0.31)	(0.54)	(0.14)	(-3.06)	(-2.43)	
Industry effects	yes	yes	yes	no	no	no	no	no
Constant				17.493	43.887	215.7 ***	119.9 ***	125.1 ***
				(1.26)	(1.55)	(14.44)	(20.15)	(29.16)
Observations	623	623	623	623	623	623	623	623
Adjusted-R^2	0.87	0.87	0.87	0.87	0.34	0.27	0.21	0.20

Table OA-4
(continued)

Panel 5: Dependent variable is *Announcement return* (adding control variables in their listed order)

Dep. Variable: Announcement return	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Analyst coverage	0.997 ***	0.908 ***	0.805 ***	0.816 ***	0.775 ***	0.761 ***	0.750 ***	0.748 ***
	(21.46)	(13.27)	(11.66)	(11.62)	(10.52)	(10.26)	(10.13)	(9.88)
Log of market capitalization		0.779 *	0.187	0.104	0.124	0.663	0.937 *	0.592
		(1.73)	(0.43)	(0.22)	(0.27)	(1.27)	(1.74)	(0.86)
Turnover			-1.897 ***	-1.891 ***	-1.876 ***	-1.746 ***	-1.566 ***	-1.951 ***
			(-7.30)	(-7.10)	(-7.01)	(-6.06)	(-5.00)	(-4.29)
Lagged return volatility				0.054 **	0.068 **	0.061 **	0.060 **	0.060 **
				(2.22)	(2.46)	(2.40)	(2.42)	(2.36)
Lagged P/E ratio					-0.009 **	-0.008 *	-0.009 *	-0.008 *
					(-2.15)	(-1.71)	(-1.86)	(-1.75)
Effective spread						-0.018	-0.018	-0.063
						(-0.24)	(-0.24)	(-0.64)
Depth						-2.043 ***	-2.035 ***	-1.867 ***
						(-3.43)	(-3.30)	(-2.94)
Market beta							-1.537	-1.745
							(-0.84)	(-0.95)
Liquidity beta							-1.960	-1.823
							(-1.43)	(-1.33)
ΔTurnover								-0.694
								(-0.97)
ΔEffective spread								-0.079
								(-0.79)
Industry effects	no	no	no	yes	yes	yes	yes	yes
Constant	-30.0 ***	-30.4 ***	-23.9 ***					
	(-77.13)	(-64.77)	(-23.60)					
Observations	623	623	623	623	623	623	623	623
Adjusted-R^2	0.46	0.46	0.50	0.51	0.52	0.52	0.52	0.53

Table OA-4
(continued)

Panel 6: Dependent variable is *Announcement return* (removing control variables in their listed order)

Dependent Var.: Announcement return	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Analyst coverage	0.748 ***	0.777 ***	0.962 ***	0.976 ***	0.996 ***	0.990 ***	0.977 ***	0.997 ***
	(9.88)	(12.51)	(17.89)	(17.80)	(18.53)	(21.09)	(20.81)	(21.46)
Log of market capitalization	0.592							
	(0.86)							
Turnover	-1.951 ***	-2.173 ***						
	(-4.29)	(-5.35)						
Lagged return volatility	0.060 **	0.062 **	0.075 ***					
	(2.36)	(2.36)	(2.93)					
Lagged P/E ratio	-0.008 *	-0.008 *	-0.011 **	-0.006				
	(-1.75)	(-1.66)	(-2.27)	(-1.49)				
Effective Spread	-0.063	-0.107	-0.014 *	-0.041	-0.073			
	(-0.64)	(-1.31)	(-0.19)	(-0.54)	(-1.02)			
Depth	-1.867 ***	-1.422 **	-1.584 ***	-1.150 **	-1.208 **			
	(-2.94)	(-2.56)	(-2.91)	(-2.35)	(-2.43)			
Market beta	-1.745	-1.723	-3.133 *	-2.991	-2.672	-1.653		
	(-0.95)	(-0.94)	(-1.65)	(-1.63)	(-1.46)	(-0.94)		
Liquidity beta	-1.823	-1.721 ***	-1.556 ***	-1.499 ***	-1.605	-2.231 *		
	(-1.33)	(-1.26)	(-1.10)	(-1.11)	(-1.18)	(-1.69)		
ΔTurnover	-0.694	-0.858	1.472 ***	1.521 ***	1.583 ***	1.679 ***	1.806 ***	
	(-0.97)	(-1.24)	(2.81)	(2.97)	(3.10)	(3.33)	(3.74)	
ΔEffective spread	-0.079	-0.126 ***	-0.010 ***	0.045	0.046 ***	0.100	0.141 ***	
	(-0.79)	(-1.59)	(-0.14)	(0.59)	(0.60)	(1.49)	(2.14)	
Industry effects	yes	yes	yes	no	no	no	no	no
Constant				-24.4 ***	-24.6 ***	-27.3 ***	-28.4 ***	-30.0 ***
				(-9.26)	(-9.27)	(-16.09)	(-47.83)	(-77.13)
Observations	623	623	623	623	623	623	623	623
Adjusted-R^2	0.53	0.53	0.50	0.48	0.48	0.41	0.39	0.46

Table OA-4
(continued)

Panel 7: Dependent variable is *Composite bubble measure* (adding control variables in their listed order)

Dep. Variable: Composite bubble measure	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Analyst coverage	-0.097 ***	-0.105 ***	-0.095 ***	-0.095 ***	-0.068 ***	-0.065 ***	-0.064 ***	-0.058 ***
	(-26.86)	(-17.43)	(-16.39)	(-15.87)	(-13.44)	(-12.49)	(-12.55)	(-12.00)
Log of market capitalization		0.069 *	0.127 ***	0.165 ***	0.152 ***	0.259 ***	0.223 ***	0.154 ***
		(1.73)	(3.18)	(3.64)	(4.13)	(5.95)	(5.41)	(3.05)
Turmover			0.186 ***	0.189 ***	0.179 ***	0.241 ***	0.218 ***	0.233 ***
			(6.34)	(6.39)	(7.90)	(9.92)	(8.84)	(6.64)
Lagged return volatility				0.004 **	-0.006 ***	-0.005 ***	-0.005 ***	-0.003 **
				(2.18)	(-2.80)	(-2.57)	(-2.59)	(-2.20)
Lagged P/E ratio					0.006 ***	0.005 ***	0.005 ***	0.005 ***
					(19.84)	(15.92)	(16.19)	(16.62)
Effective spread						0.042 ***	0.041 ***	0.031 ***
						(6.41)	(6.45)	(4.22)
Depth						-0.118 ***	-0.119 ***	-0.090 **
						(-2.73)	(-2.71)	(-2.36)
Market beta							0.189	0.135
							(1.38)	(0.99)
Liquidity beta							0.271 ***	0.265 ***
							(2.67)	(2.68)
ΔTurnover								0.096 **
								(1.97)
ΔEffective spread								-0.020 ***
								(-2.65)
Industry effects	no	no	no	yes	yes	yes	yes	yes
Constant	0.588 ***	0.548 ***	-0.089 ***					
	(14.24)	(12.64)	(-0.80)					
Observations	623	623	623	623	623	623	623	623
Adjusted-R^2	0.46	0.46	0.51	0.52	0.71	0.73	0.74	0.75

Table OA-4
(continued)

Panel 8: Dependent variable is *Composite bubble measure* (removing control variables in their listed order)

Dependent Variable: Composite bubble measure	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Analyst coverage	-0.058 ***	-0.051 ***	-0.066 ***	-0.067 ***	-0.083 ***	-0.094 ***	-0.094 ***	-0.097 ***
	(-12.00)	(-12.60)	(-18.80)	(-19.14)	(-18.37)	(-27.06)	(-27.31)	(-26.86)
Log of market capitalization	0.154 ***							
	(3.05)							
Turnover	0.233 ***	0.175 ***						
	(6.64)	(5.96)						
Lagged return volatility	-0.003 **	-0.003 **	-0.004 ***					
	(-2.20)	(-2.03)	(-2.77)					
Lagged P/E ratio	0.005 ***	0.005 ***	0.006 ***	0.005 ***				
	(16.62)	(17.50)	(18.01)	(18.21)				
Effective Spread	0.031 ***	0.020 ***	0.012 **	0.014 **	0.041 ***			
	(4.22)	(3.35)	(2.12)	(2.47)	(5.32)			
Depth	-0.090 **	0.025	0.038	0.001	0.050			
	(-2.36)	(0.70)	(1.06)	(0.03)	(0.95)			
Market beta	0.135	0.140	0.254 *	0.273 **	0.005	-0.444 ***		
	(0.99)	(1.01)	(1.79)	(1.97)	(0.03)	(-2.96)		
Liquidity beta	0.265 ***	0.291 ***	0.278 ***	0.262 **	0.350 ***	0.499 ***		
	(2.68)	(2.91)	(2.64)	(2.56)	(2.92)	(4.27)		
ΔTurnover	0.096 **	0.053	-0.135 ***	-0.138 ***	-0.190 ***	-0.196 ***	-0.168 ***	
	(1.97)	(1.14)	(-3.78)	(-3.95)	(-4.44)	(-4.43)	(-3.81)	
ΔEffective spread	-0.020 ***	-0.032 ***	-0.042 ***	-0.045 ***	-0.045 ***	-0.063 ***	-0.066 ***	
	(-2.65)	(-5.34)	(-7.04)	(-7.71)	(-5.81)	(-8.88)	(-9.09)	
Industry effects	yes	yes	yes	no	no	no	no	no
Constant				-0.682 ***	-0.512 *	0.826 ***	0.358 ***	0.588 ***
				(-3.36)	(-1.93)	(5.50)	(6.45)	(14.24)
Observations	623	623	623	623	623	623	623	623
Adjusted-R^2	0.75	0.74	0.73	0.73	0.59	0.56	0.55	0.46

Table OA-5. First stage partial regression for *Analyst coverage* and its instruments

This table is related to Table 4 in the main paper, and reports the first stage ordinary least squares regressions of the 2SLS regressions reported in Table 4. The first stage regressions explain *Analyst coverage* for a sample of 623 Shanghai A-shares. *Analyst coverage* is the number of brokerage firms issuing earnings-per-share forecasts during the six-month reference period of November 29, 2006 to May 29, 2007. *Trading volume in 2005* is the average daily trading volume during 2005, in millions of yuans. *Mutual fund ownership in June/2005* is the fraction of tradable shares owned by Chinese mutual funds on June 30, 2005. See Table 3 for other variables. We report heteroskedasticity-robust t-statistics in parentheses beneath variable coefficients, and ***, **, * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Dep. Variable: Analyst coverage	(1)	(2)
Trading volume in 2005	0.303 ***	0.093 ***
	(7.96)	(3.34)
Mutual fund ownership in June 2005	0.282 ***	0.155 ***
	(16.17)	(10.02)
Log of market capitalization		3.233 ***
		(9.09)
Turnover		-1.037 ***
		(-4.44)
Lagged return volatility		-0.010
		(-0.93)
Lagged P/E ratio		-0.010 ***
		(-4.86)
Effective spread		0.011
		(0.18)
Depth		-0.119
		(-0.36)
Market beta		1.303
		(1.54)
Liquidity beta		-0.838
		(-1.32)
ΔTurnover		-0.974 ***
		(-3.11)
ΔEffective spread		0.146 **
		(2.57)
Industry effects	no	yes
Constant	2.967 ***	
	(12.69)	
Observations	623	623
Adjusted-R^2	0.48	0.73

Table OA-6. Robustness regressions addressing endogeneity: other bubble measures

This table is related to Table 4 in the main paper, and reports ordinary least squares and two-stage least square (2SLS) regressions that explain *Cumulative return*, *P/E ratio*, and *Announcement return* for a sample of 623 Shanghai A-shares. The 2SLS regressions use *Trading volume in 2005* (average daily trading volume in 2005) and *Mutual fund ownership in June/2005* (the percent of tradable shares owned by mutual funds at the end of June/2005) as instruments for *Analyst coverage*. Unless otherwise noted, all variables are averages across the six-month reference period of November 29, 2006 to May 29, 2007, calculated from daily data. *Analyst coverage* is the number of brokerage firms issuing earnings-per-share forecasts during the reference period. *Market capitalization* is the stock price times the number of tradable shares. *Turnover* is the number of shares traded divided by the total number of tradable shares. *Effective spread* is the average of twice the difference between the transaction price and mid-point, divided by the midpoint. *Depth* is the average of one half times the sum of the monetary quantities associated with the best bid and best ask offers. *ΔTurnover* and *ΔEffective spread* are defined as the averages in the six-month post-tax-increase period of May 31, 2007 to November 30, 2007 minus the average in the six-month pre-tax-increase reference period. *Market beta* and *Liquidity beta*, respectively, are the coefficients on the value-weighted market return and on an aggregate liquidity factor in a regression of daily stock returns, where the aggregate liquidity factor is defined as the daily innovation on the average effective spread across all stocks. *Industry effects* are dummy variables based on thirteen industries defined by the China Securities Regulatory Commission (CSRC). We report heteroskedasticity-robust t-statistics in parentheses beneath variable coefficients, and ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Ordinary least squares			Two-stage least squares		
	Dependent Variable			Dependent Variable		
	Cum. ret. (1)	P/E ratio (2)	Ann. ret. (3)	Cum. ret. (4)	P/E ratio (5)	Ann. ret. (6)
Analyst coverage in 2005	-3.182 ***	-0.698 ***	0.301 ***			
	(-5.90)	(-4.10)	(4.49)			
Analyst coverage				-6.514 ***	-0.717 *	0.715 ***
				(-4.97)	(-1.93)	(5.22)
Log of market capitalization	20.370 ***	8.480 ***	2.788 ***	37.162 ***	8.854 ***	0.749
	(3.01)	(3.34)	(4.35)	(4.17)	(3.04)	(0.87)
Turnover	25.148 ***	9.420 ***	-2.617 ***	19.716 ***	9.080 ***	-1.984 ***
	(5.05)	(4.32)	(-5.52)	(4.11)	(4.08)	(-4.30)
Lagged return volatility	-0.470	0.089	0.075 **	-0.362 *	0.103	0.059 **
	(-1.48)	(0.85)	(2.44)	(-1.71)	(1.31)	(2.35)
Lagged P/E ratio	0.093 **	0.816 ***	-0.014 ***	0.046	0.813 ***	-0.009 *
	(2.02)	(44.42)	(-2.88)	(0.97)	(42.78)	(-1.78)
Effective spread	3.759 ***	1.888 ***	-0.036	3.992 ***	1.906 ***	-0.062
	(3.94)	(4.68)	(-0.34)	(4.52)	(4.86)	(-0.64)
Depth	-36.444 ***	-8.180 ***	-1.443 **	-31.564 ***	-6.893 ***	-1.900 ***
	(-4.61)	(-2.96)	(-2.05)	(-4.75)	(-3.05)	(-3.00)
Market beta	20.223	-7.825	-1.620	22.197	-7.120	-1.740
	(1.03)	(-0.74)	(-0.81)	(1.16)	(-0.69)	(-0.96)
Liquidity beta	44.418 ***	0.391	-2.249	40.449 ***	-0.152	-1.842
	(3.24)	(0.07)	(-1.53)	(2.99)	(-0.03)	(-1.36)
ΔTurnover	7.435	8.431 ***	-1.355 *	2.266	8.270 ***	-0.731
	(1.12)	(2.93)	(-1.83)	(0.35)	(2.84)	(-1.03
ΔEffective spread	-7.359	1.055 **	0.039	-6.374 ***	1.127 **	-0.072
	(-6.65)	(2.16)	(0.39)	(-5.81)	(2.37)	(-0.72)
Industry effects	yes	yes	yes	yes	yes	yes
Observations	623	623	623	623	623	623
Adjusted-R^2	0.46	0.87	0.45	0.47	0.87	0.52
Sargan chi-square				0.196	2.154	0.236
(p-value)				(0.66)	(0.14)	(0.63)

Table OA-7. Addressing endogeneity concerns using either *Trading volume* or *Mutual fund ownership June/2005* as the unique instruments

This table is related to Table 4 in the main paper, and reports two-stage least square (2SLS) regressions that explain bubble intensity measures for a sample of 623 Shanghai A-shares. In Panel 1 we use *Trading volume in 2005* (average daily trading volume in 2005) as instrument for *Analyst coverage*. In Panel 2 we use *Mutual fund ownership June/2005* (percent of tradable shares owned by mutual funds at the end of June/2005) as instrument for *Analyst coverag.* Unless otherwise noted, all variables are averages across the six-month reference period of November 29, 2006 to May 29, 2007, calculated from daily data. *Composite bubble measure* is the first principal component of *Cumulative return*, *P/E ratio*, and *Announcement return,* normalized to have mean zero and variance equal to 1. *Analyst coverage* is the number of brokerage firms issuing earnings-per-share forecasts during the reference period. *Market capitalization* is the stock price times the number of tradable shares. *Turnover* is the number of shares traded divided by the total number of tradable shares. *Effective spread* is the average of twice the difference between the transaction price and mid-point, divided by the midpoint. *Depth* is the average of one half times the sum of the monetary quantities associated with the best bid and best ask offers. *ΔTurnover* and *ΔEffective spread* are defined as the averages in the six-month post-tax-increase period of May 31, 2007 to November 30, 2007 minus the average in the six-month pre-tax-increase reference period. *Market beta* and *Liquidity beta*, respectively, are the coefficients on the value-weighted market return and on an aggregate liquidity factor in a regression of daily stock returns, where the aggregate liquidity factor is defined as the daily innovation on the average effective spread across all stocks. *Industry effects* are dummy variables based on thirteen industries defined by the China Securities Regulatory Commission (CSRC). We report heteroskedasticity-robust t-statistics in parentheses beneath variable coefficients, and ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel 1: Using *Trading volume in 2005* as unique instrument for *Analyst coverage*

	Dependent Variable							
	Cumulative return		P/E ratio		Announcement return		Composite bubble measure	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Analyst coverage	-3.172 ***	-5.338 **	-3.681 ***	0.490	0.829 ***	0.842 ***	-0.071 ***	-0.059 ***
	(-3.55)	(-2.31)	(-6.93)	(0.56)	(8.10)	(2.67)	(-10.56)	(-3.31)
Log of market capitalization		32.538 **		3.144		0.148		0.157
		(2.44)		(0.66)		(0.09)		(1.58)
Turnover		20.683 ***		10.274 ***		-1.858 ***		0.233 ***
		(4.05)		(4.33)		(-3.48)		(6.33)
Lagged return volatility		-0.328		0.145 *		0.064 **		-0.003 **
		(-1.50)		(1.71)		(2.33)		(-2.13)
Lagged P/E ratio		0.058		0.829 ***		-0.007		0.005 ***
		(1.06)		(37.94)		(-1.18)		(13.70)
Effective spread		3.396 ***		1.869 ***		-0.066		0.031 ***
		(4.46)		(4.53)		(-0.68)		(4.30)
Depth		-30.563 ***		-5.656 **		-1.770 ***		-0.091 **
		(-4.41)		(-2.58)		(-2.65)		(-2.00)
Market beta		22.032		-7.324		-1.761		0.135
		(1.15)		(-0.69)		(-0.97)		(1.01)
Liquidity beta		41.028 ***		0.560		-1.767		0.265 ***
		(3.05)		(0.10)		(-1.29)		(2.71)
ΔTurnover		3.369		9.632 ***		-0.589		0.096 *
		(0.49)		(3.10)		(-0.77)		(1.89)
ΔEffective spread		-6.560 ***		0.897 *		-0.096		-0.020 **
		(-5.75)		(1.77)		(-0.85)		(-2.44)
Industry effects	no	yes	no	yes	no	yes	no	yes
Constant	223.6 ***		116.4 ***		-29.0 ***		0.433 ***	
	(33.53)		(24.06)		(-42.12)		(7.92)	
Observations	623	623	623	623	623	623	623	623
Adjusted-R^2	0.13	0.47	0.19	0.87	0.45	0.53	0.43	0.75

Table OA-7
(continued)

Panel 2: Using *Mutual fund ownership June/2005* as unique instrument for *Analyst coverage*

	Dependent Variable							
	Cumulative return		P/E ratio		Announcement return		Composite bubble measure	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Analyst coverage	-5.572 ***	-6.622 ***	-5.313 ***	-0.850 **	0.988 ***	0.701 ***	-0.099 ***	-0.065 ***
	(-10.35)	(-4.87)	(-13.23)	(-2.18)	(14.38)	(5.15)	(-18.90)	(-6.50)
Log of market capitalization		37.673 ***		9.485 ***		0.815 ***		0.187 ***
		(4.18)		(3.21)		(0.94)		(2.84)
Turnover		19.608 ***		8.948 ***		-1.998 ***		0.226 ***
		(4.07)		(4.01)		(-4.32)		(6.41)
Lagged return volatility		-0.366 *		0.098		0.059 **		-0.004 **
		(-1.72)		(1.25)		(2.34)		(-2.30)
Lagged P/E ratio		0.044		0.811 ***		-0.009 *		0.005
		(0.93)		(42.61)		(-1.81)		(15.89)
Effective spread		3.996 ***		1.911 ***		-0.061 ***		0.031 ***
		(4.52)		(4.89)		(-0.64)		(4.36)
Depth		-31.675 ***		-7.030 ***		-1.914 ***		-0.097 **
		(-4.77)		(-3.07)		(-3.01)		(-2.28)
Market beta		22.216		-7.098		-1.737 **		0.136
		(1.16)		(-0.69)		(-0.96)		(1.02)
Liquidity beta		40.386 ***		-0.231		-1.851		0.261 ***
		(2.98)		(-0.04)		(-1.37)		(2.67)
ΔDaily turnover		2.144		8.120 ***		-0.747		0.089 *
		(0.33)		(2.79)		(-1.05)		(1.82)
ΔEffective spread		-6.354 ***		1.152 **		-0.070		-0.019 **
		(-5.87)		(2.43)		(-0.69)		(-2.43)
Industry effects	no	yes	no	yes	no	yes	no	yes
Constant	238.4 ***		126.4 ***		-29.92 ***		0.598 ***	
	(44.45)		(28.13)		(-59.51)		(12.96)	
Observations	623	623	623	623	623	623	623	623
Adjusted-R^2	0.16	0.46	0.21	0.87	0.46	0.53	0.46	0.75

Table OA-8. Including *Analyst coverage* instruments in OLS regression

This table is related to Table 4 in the main paper, and reports ordinary least squares regressions explaining four different measures of bubble intensity for a sample of 623 Shanghai A-shares. The variables used as instruments for *Analyst coverage* in Table 4 are added as explanatory variables in the regressions explaining bubble intensity measures. These variables are *Trading volume in 2005* (average daily trading volume in 2005) and *Mutual Fund Ownership in June/2005* (percent of tradable shares owned by mutual funds at the end of June/2005). *Analyst coverage* is the number of brokerage firms issuing earnings-per-share forecasts during the six-month reference period of November 29, 2006 to May 29, 2007. Unless otherwise noted, all variables are averages across the six-month reference period of November 29, 2006 to May 29, 2007, calculated from daily data. Definitions of other explanatory variables can be found in Table 3. We report heteroskedasticity-robust t-statistics in parentheses beneath variable coefficients, and ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Dependent Variable			
	Cumulat. return	P/E ratio	Announc. return	Composite bubble meas.
Analyst coverage	-4.076 ***	-0.880 ***	0.758 ***	-0.057 ***
	(-5.81)	(-3.16)	(8.92)	(-10.21)
Mutual fund ownership in June 2005	-0.419 *	-0.025	-0.020	-0.149
	(-1.67)	(-0.31)	(-0.46)	(-0.79)
Trading Volume in 2005	-0.004	0.290	0.025	0.001
	(-0.01)	(1.46)	(0.38)	(0.15)
Log of market capitalization	28.839 ***	8.841 ***	0.553	0.155 ***
	(3.99)	(3.41)	(0.79)	(2.97)
Turnover	21.897 ***	8.480 ***	-1.984 ***	0.233 ***
	(4.42)	(3.76)	(-4.22)	(6.41)
Lagged return volatility	-0.337	0.103	0.060 **	-0.004 **
	(-1.58)	(1.28)	(2.34)	(-2.22)
Lagged P/E ratio	0.070	0.811 ***	-0.008 *	0.005 ***
	(1.51)	(43.04)	(-1.76)	(16.60)
Effective spread	3.978 ***	1.924 ***	-0.060	0.031 ***
	(4.32)	(4.82)	(-0.62)	(4.22)
Depth	-31.465 ***	-7.147 ***	-1.919 ***	-0.097 **
	(-4.77)	(-3.19)	(-3.04)	(-2.23)
Market beta	19.053	-6.864	-1.791	0.125
	(0.97)	(-0.65)	(-0.97)	(0.91)
Liquidity beta	42.325 ***	-0.499	-1.828	0.267 ***
	(3.10)	(-0.09)	(-1.34)	(2.70)
ΔTurnover	4.487	7.922 ***	-0.709	0.096 *
	(0.68)	(2.73)	(-0.99)	(1.95)
ΔEffective spread	-6.740 ***	1.139 **	-0.080	-0.020 ***
	(-6.17)	(2.34)	(-0.80)	(-2.65)
Industry effects	yes	yes	yes	yes
Observations	623	623	623	623
Adjusted-R^2	0.47	0.87	0.52	0.75

Table OA-9. Interaction with *Dispersion among analysts* – other bubble measures

This table is related to Table 6 in the main paper, and reports ordinary least squares regressions that *Cumulative return*, *P/E ratio*, and *Announcement return* for a subsample of 367 Shanghai A-shares that are followed by at least two analysts. Unless otherwise noted, all variables are averages across the six-month reference period of November 29, 2006 to May 29, 2007, calculated from daily data. *Analyst coverage* is the number of brokerage firms issuing earnings-per-share forecasts during the reference period. *Market capitalization* is the stock price times the number of tradable shares. *Turnover* is the number of shares traded divided by the total number of tradable shares. *Effective spread* is the average of twice the difference between the transaction price and mid-point, divided by the midpoint. *Depth* is the average of one half times the sum of the monetary quantities associated with the best bid and best ask offers. *ΔTurnover* and *ΔEffective spread* are defined as the averages in the six-month post-tax-increase period of May 31, 2007 to November 30, 2007 minus the average in the six-month pre-tax-increase reference period. *Market beta* and *Liquidity beta*, respectively, are the coefficients on the value-weighted market return and on an aggregate liquidity factor in a regression of daily stock returns, where the aggregate liquidity factor is defined as the daily innovation on the average effective spread across all stocks. *Industry effects* are dummy variables based on thirteen industries defined by the China Securities Regulatory Commission (CSRC). We report heteroskedasticity-robust t-statistics in parentheses beneath variable coefficients, and ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Dependent Variable								
	Cumulative return			P/E ratio			Announcement return		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Analyst coverage	-4.960 ***	-4.754 ***	-4.794 ***	-2.437 ***	-2.396 ***	-0.457 *	0.852 ***	0.837 ***	0.585 ***
	(-9.05)	(-8.81)	(-6.21)	(-7.17)	(-7.06)	(-1.86)	(13.26)	(13.41)	(6.44)
Dispersion among analysts		-12.230 *	-11.831 **		-6.090	-0.399		0.607	0.710
		(-1.89)	(-2.22)		(-1.34)	(-0.20)		(1.13)	(1.42)
Analyst cov. * Disp. among analysts		3.241 ***	1.724 ***		0.171	0.079		-0.265 ***	-0.187 ***
		(5.10)	(2.99)		(0.36)	(0.39)		(-4.31)	(-2.93)
Log of market capitalization			23.892 ***			5.436 **			0.664
			(3.05)			(2.12)			(0.80)
Turnover			13.531 *			6.893 **			-2.289 ***
			(1.90)			(2.18)			(-3.28)
Lagged return volatility			-0.003			0.277 *			0.150 ***
			(-0.01)			(1.65)			(3.42)
Lagged P/E ratio			0.039			0.822 ***			-0.020 ***
			(0.44)			(19.90)			(-2.75)
Effective spread			1.802 *			0.788 **			-0.055
			(1.66)			(2.13)			(-0.48)
Depth			-22.333 ***			-2.732			-1.536 **
			(-3.77)			(-1.25)			(-2.22)
Market beta			26.787			-23.472 **			0.542
			(1.04)			(-2.17)			(0.20)
Liquidity beta			37.683 **			-2.464			-1.281
			(2.20)			(-0.39)			(-0.68)
ΔTurnover			-16.792			6.971			-0.245
			(-1.62)			(1.63)			(-0.21)
ΔEffective spread			-6.445 ***			0.478			0.029
			(-4.86)			(0.81)			(0.22)
Industry effects	no	no	yes	no	no	yes	no	no	yes
Constant	231.1 ***	228.2 ***		86.7 ***	85.6 ***		-27.9 ***	-27.65 ***	
	(28.06)	(27.80)		(15.72)	(15.87)		(-37.72)	(-37.77)	
Observations	364	364	364	364	364	364	364	364	364
Adjusted-R^2	0.15	0.20	0.45	0.10	0.11	0.81	0.35	0.38	0.50

Table OA-10. Regressions including interactions with the dispersion of analysts' earnings forecasts

This table is related to Table 6 in the paper, and reports ordinary least squares regressions explaining *Composite bubble measure* and *Turnover* for a subsample of 364 Shanghai A-shares that are followed by at least two analysts. Here we use either *Dispersion of analysts' earnings forecasts* (Panel 1) or *Dispersion of analysts' recommendations* instead of *Dispersion among analysts.* All other variables are identical to those in Table 6. We report heteroskedasticity-robust t-statistics in parentheses beneath variable coefficients, and ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel 1: Using *Dispersion of analysts' earnings forecasts*

	Dependent Variable					
	Composite bubble measure			Turnover		
	(1)	(2)	(3)	(4)	(5)	(6)
Analyst coverage	-0.074 ***	-0.071 ***	-0.050 ***	-0.084 ***	-0.081 ***	-0.024 ***
	(-15.79)	(-15.45)	(-8.38)	(-11.83)	(-11.34)	(-2.82)
Dispersion of analysts' earnings forecasts		-0.061	-0.021		0.101	0.093
		(-1.55)	(-0.60)		(1.59)	(1.56)
Analyst cov. * Dispersion of anal. earnings forec.		0.018 ***	0.009 **		0.019 ***	0.011 *
		(4.04)	(2.40)		(2.48)	(1.85)
Log of market capitalization			0.102 *			-0.760 ***
			(1.82)			(-10.77)
Turnover			0.198 ***			
			(3.54)			
Lagged return volatility			-0.005			0.001
			(-1.46)			(0.19)
Lagged P/E ratio			0.006 ***			0.004 ***
			(10.58)			(3.57)
Effective spread			0.014 *			-0.092 ***
			(1.74)			(-10.02)
Depth			-0.041			0.721 ***
			(-1.06)			(4.01)
Market beta			-0.046			0.736 ***
			(-0.26)			(2.60)
Liquidity beta			0.221 *			0.200
			(1.81)			(1.14)
ΔTurnover			-0.023			
			(-0.29)			
ΔEffective spread			-0.028 ***			
			(-2.76)			
Industry effects	no	no	yes	no	no	yes
Constant	0.247 ***	0.229 ***		3.164 ***	3.136 ***	
	(3.72)	(3.45)		(31.88)	(31.86)	
Observations	364	364	364	364	364	364
Adjusted-R^2	0.37	0.40	0.68	0.23	0.28	0.58

Table OA-10
(continued)

Panel 2: Using *Dispersion of analysts' recommendations*

	Dependent Variable					
	Composite bubble measure			Turnover		
	(1)	(2)	(3)	(4)	(5)	(6)
Analyst coverage	-0.074 ***	-0.074 ***	-0.051 ***	-0.084 ***	-0.086 ***	-0.027 ***
	(-15.89)	(-15.97)	(-8.99)	(-11.83)	(-12.13)	(-3.03)
Dispersion of analysts' recommendations		-0.098 **	-0.092 ***		-0.035	-0.037
		(-2.07)	(-2.60)		(-0.45)	(-0.55)
Analyst cov. * Dispersion of analysts recommendations		0.023 ***	0.014 ***		0.019 *	0.011
		(3.34)	(2.64)		(1.66)	(1.32)
Log of market capitalization			0.118 **			-0.757 ***
			(2.09)			(-10.46)
Turnover			0.199 ***			
			(3.54)			
Lagged return volatility			-0.005			-0.001
			(-1.63)			(-0.26)
Lagged P/E ratio			0.006 ***			0.004 ***
			(10.78)			(3.51)
Effective spread			0.014 *			-0.094 ***
			(1.71)			(-10.24)
Depth			-0.053			0.736 ***
			(-1.34)			(3.76)
Market beta			0.011			0.947 ***
			(0.06)			(3.32)
Liquidity beta			0.231 *			0.197
			(1.92)			(1.07)
ΔTurnover			-0.048			
			(-0.58)			
ΔEffective spread			-0.028 ***			
			(-2.71)			
Industry effects	no	no	yes	no	no	yes
Constant	0.247 ***	0.240 ***		3.164 ***	3.164 ***	
	(3.72)	(3.64)		(31.88)	(31.82)	
Observations	364	364	364	364	364	364
Adjusted-R^2	0.37	0.39	0.67	0.23	0.24	0.56

Table OA-11. Regressions explaining *Turnover* (full sample of stocks)

This table is related to Table 6 in the main paper, and reports ordinary least squares regressions that explain *Turnover* for a sample of 623 Shanghai A-shares. *Turnover* is the average number of shares traded divided by total number of tradable shares during the reference period (November 29, 2006 to May 29, 2007). *Analyst coverage* is the number of brokerage firms issuing earnings-per-share forecasts during the reference period. *Market capitalization* is the average of stock price times the number of tradable shares during the reference period. *Effective spread* is the average of twice the difference between the transaction price and mid-point, divided by the midpoint. *Depth* is the average of one half times the sum of the monetary quantities associated with the best bid and best ask offers. *Market beta* and *Liquidity beta* are, respectively, the coefficients on the value-weighted market return and on an aggregate liquidity factor in a regression of daily stock returns; where the aggregate liquidity factor is defined as the daily innovation on the average effective spread across all stocks. *Industry effects* are dummy variables based on thirteen industries defined by the China Securities Regulatory Commission (CSRC). We report heteroskedasticity-robust t-statistics in parentheses beneath variable coefficients, and ***, **, * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Dep. Variable: Turnover	(1)	(2)	(3)
Analyst coverage	-0.090 ***	-0.054 ***	-0.035 ***
	(-17.05)	(-5.87)	(-4.33)
Log of market capitalization		-0.311 ***	-0.714 ***
		(-5.22)	(-10.07)
Lagged return volatility		0.001	-0.005
		(0.21)	(-1.55)
Lagged P/E ratio			0.002 ***
			(4.20)
Effective spread			-0.064 ***
			(-7.15)
Depth			0.890 ***
			(3.30)
Market beta			1.592 ***
			(7.11)
Liquidity beta			-0.227
			(-1.45)
Industry effects	no	no	yes
Constant	3.25 ***	3.43 ***	
	(56.84)	(57.54)	
Observations	623	623	623
Adjusted-R^2	0.26	0.29	0.49

Table OA-12. Brokerage firms providing EPS forecasts in the RESSET database

The table lists the 48 Chinese brokerage firms providing earnings-per-share forecasts for the sample stocks during the six-month reference period (November 29, 2006 to May 29, 2007), according to the RESSET database. We report the total number of sample firms covered by each brokerage firm. The total number of covered firms across all brokerage firms is 3,782, and hence the sample average of the variable *Analyst coverage* used in the study is 3,782/623=6.071

Brokerage firm	Number of covered firms
Anxin Securities	24
BOC International	90
Bohai Securities	16
Capitalcare Securities	1
Central China Securities Holdings	54
Century Securities	14
Changjiang Securities	140
China Galaxy Securities	78
China International Capital Corporation	144
China Merchants Securities	220
China Minzu Securities	8
China National Investment	193
China Southwest Securities	42
Citic Construction Securities	181
Citic Securities Company	193
Dongguan Securities	14
Everbright Securities	160
Firstcapital Securities	9
Fortune Securities	16
Gaohua Securities	3
Gf Securities	101
Golden Sun Securities	3
Goldstate Securities	26
Gongdong Securities	1
Great Wall Securities	87
Guodu Securities	106
Guojin Securities	95
Guolian Securities	9
Guotai Junan Securities	204
Guoxin Securities	136
Guoyuan Securities	10
Haitong Securities Company	164
Huatai Securities	131
Huaxi Securities	48
Industrial Securities	129
Jiangnan Securities	3
Kaiji Securities	38
Masterlink Securities	16
Nanjing Securities	1
Northeastern Securities	8
Orient Securities Company	134
Ping An Securities Company	140
Sealand Securities	33
Shandong Qilu Securities Broking	14
Shanghai Securities	169
Shenyin & Wanguo Securities	236
United Securities	122
Zhejiang Securities	18