A NEW MEASURE FOR DETERMINING THE AMOUNT OF PRIVATE INFORMATION OF COMMON STOCKS

Radu BURLACU
Patrice FONTAINE
Sonia JIMENEZ-GARCES[1]

C.E.R.A.G. – UMR CNRS 5820
Université Pierre Mendès France

Abstract

This paper proposes and tests a new measure of the private information content of common stocks. The measure is designated by "AIM" as "Asymmetric Information Measure". This measure obtains by projecting stock returns on prices and considering the R^2 from the regression. We use the individual stock price, the price of industry portfolios and the price of the market portfolio as information sources allowing uninformed investors to extract information on stock returns. The measure is theoretically founded on a generalisation of the Grossman and Stiglitz (1980) model for multi-securities markets. AIM has a strong positive impact on stock returns and dominates traditional factors of risk like β and the Fama and French factors. Our results are robust to the AIM measure and CAPM model specifications. The results support our approach for determining the amount of private information and the hypothesis that stock returns embed an information-risk premium. The firm-specific return variation, another measure used for private information, seems to measure rather the price informativeness.

JEL classification: G12, G14

Key-words: Information Asymmetry, Price Informativity, Information-risk premium, Specific Risk, Industry, Information extraction

---

[1] Corresponding author: Jimenez-Garcès is PhD student at the CERAG and graduate assistant professor at the High Business School (ESA) of the University UPMF (Université Pierre Mendès France). Professional address: CERAG-UPMF, BP 47, 38040 Grenoble Cedex 9, France; Phone number: (33) 06 70 46 92 86 (personal); E-mail: sonia.jimenez@wanadoo.fr

A NEW MEASURE FOR DETERMINING THE AMOUNT OF PRIVATE INFORMATION OF COMMON STOCKS

Introduction

Information asymmetry (hereafter IA) is a challenging and controversial topic. The hypothesis that investors have heterogeneous beliefs on the returns of the securities traded in the market has been addressed in rational expectations (hereafter RE) models of asset pricing where the information is transmitted or aggregated by prices[2]. These models investigate how investors receive additional information relatively to their own, private information, by observing asset prices. They show that if the market is not noisy, it is overefficient and cannot compensate the information-gathering function (Grossman-Stiglitz paradox). If the market is noisy, the Grossman-Stiglitz paradox is solved and equilibrium asset prices reveal only a limited quantity of information to the other investors. We designate this quantity of information relative to the total quantity of relevant information as "*price informativity*". Another quantity of information is retained by the noise in the market. We designate this quantity of information relative to the total quantity of information as "*information asymmetry*".

Recent empirical studies have tested some fundamental implications of RE models with privately informed investors on the American stock market. Two complementary approaches are relevant here because they are strongly related to the issue developed in our paper. Easley et al. (2002) test the presence of private information on financial markets. They document that stock returns are positively correlated to the extent of private information. An increase of 10% in their proxy for information asymmetry, called "the probability of information-based trading" (PIN) leads to a difference in expected returns of 2.5% per year. The information risk factor is more significant in explaining stock returns than the traditional market factor and the Fama and French factors. The information factor seems to dominate the traditional factors in explaining stock returns. Thus the evidence supports the hypothesis of information asymmetry (IA) on financial markets: investors seem to possess private information on market-traded

[2] Grosmann and Stiglitz (1980), Hellwig (1980), Admati A. (1983, 1985), Wang (1993), Easley and O'Hara, (2002). See also the presidential address of O'Hara (2003), "Liquidity and Price Discovery" for a panorama of the theoretical and empirical literature on this topic.

securities and prices reveal part of it (the market is noisy). Biais et al. (2003) take a different approach. Rather than studying the implications of REM models in terms of asset pricing, they study the implications of these models in terms of portfolio management. The main implication they exploit is that investors possessing private information trade actively, depending on the private information they possess or the public information they observe in prices. These authors propose a price-contingent portfolio strategy based on the extraction of information from prices for an uninformed investor. This strategy strongly outperforms the performance of the market portfolio. Because the signal extraction is *ex-ante*, not *ex-post*, the results of Biais et al. (2003) corroborate the hypothesis that there is significant private information in financial markets and that these markets are noisy.

This paper addresses the issue of measuring the degree of private information embedded in common stocks and the impact of this private information on asset prices. Our main contribution is to propose a new measure of stock private information which obtains directly from RE models with asymmetrically informed investors. As far as we know, the financial literature has not yet put forward an approach based on such models in order to measure the extent of private information. The extent of private information in common stocks is determined by projecting returns on prices. This approach considers the position of an uninformed investor that extract private information from observed prices in order to better anticipate the return distribution. If there is some private information not reflected in prices (information asymmetry), then stock returns and equilibrium prices will be correlated. Indeed, this private information conditions simultaneously the equilibrium price at the beginning of a certain period and also the return on that period. The higher the amount of private information *not* revealed by prices, the higher the degree of correlation between returns and prices. This correlation is captured by the R^2 from the regression of returns on prices.

What prices are relevant for considering in the regression for a relevant extraction of information from prices? Theoretically, the prices of all the stocks on the market reflect some relevant information on the return of any common stock i if stocks are correlated. The correlation may be induced by the stocks' exposure to common factors of risk, by the correlation of the information detained by investors or by the noise terms. Thus, a relevant theoretical approach would be to project returns on the prices of all securities. Practically, this is unachievable. In order to choose the most fruitful sources of information for the return of a certain common stock, we study the price informativeness in the context of a multi-asset generalisation of the Grossman and Stiglitz (1980) model. Moreover, we use a factor approach, allowing to determine the sources of information in a meticulous way.

We propose several specifications for determining our asymmetric information measure (AIM), depending on the number of information sources we use in the regression. Our AIM measure is used for analysing the impact of the information factor on stock returns on the American market over the period 1985 – 2002. The theory teaches us that, under asymmetric information, stock returns embed an information-risk premium. We should thus observe a positive correlation between our AIM measure and stock prices. In addition, we compare our AIM measure with the existing measure based on the firm-specific return variation.

The following section gives a brief panorama of existing approaches for measuring the extent of private information. Then we present our approach based on a generalisation of the Grossman and Stiglitz (1980) model for multi-securities markets and with specific and common factors generating returns. The third section presents our data and descriptive statistics for our private information measures. The fourth section estimates several CAPM specifications in order to determine the impact of private information on stock prices. This section is followed by a conclusion.

Existing measures of stocks' private information

The existing literature proposes several approaches for measuring the degree of stocks' private information. A first category includes somewhat "informal" measures. This category includes measures based on analysts' forecasts about the firm's expected earnings, measures based on the investment opportunity set and the total or the specific return variation. A second category includes more "formal" private information measures, derived from microstructure models.

Informal measures

A large body of the empirical literature has used IA measures based on analysts' forecasts. Some authors assert that firms with high degrees of IA exhibit a high (ex-ante) dispersion of analysts' forecasts or a high (ex-post) forecast error. For example, Blackwell and Dubins (1962) observe that the opinions of the financial analysts tend to converge when the quantity of available information about firms increases. Krishnaswami and Subramaniam (1998), Gilson et al. (1998), D'Mello and Ferris (2000) use the forecast error or the forecast dispersion as a measure of companies' IA. Other authors assert the firms followed by a great number of financial analysts should exhibit a small IA. Brennan, Jegadeesh and Swaminathan (1993) show that companies followed by a great number of financial analysts react more quickly to the disclosure of new information. Elton, Gruber and Gultekin (1984) show that financial analysts are attracted by stocks of firms with high degrees of private information because the value of the information for these stocks is higher.

Some empirical studies show that measures based on analysts' forecasts present significant drawbacks. Chung et al. (1995) observe that the number of financial analysts is proportional to the bid-ask spread. As the spread is positively related to informational asymmetries, the positive correlation between the number of analysts and informational asymmetry is in opposition with the results of the preceding studies. Easley, O' Hara and Paperman (1998) analyze the link between the number of financial analysts and microstructure-based IA variables. They conclude that the number of financial analysts is not relevant for approaching the degree of stocks' private information. The studies of Clarke and Shastri (2001) and Van Ness et al. (2001) achieve similar conclusions. They show that the correlation between measures based on analysts' forecasts and other existing IA measures is not stable and has quite frequently a negative sign.

Other studies use the market-to-book ratio or the PER (price earning ratio) as IA measures (Smith and Watts, 1992, McLaughlin et al., 1998). Managers of firms with high market-to-book or PER ratios have probably more precise information on the cash-flows generated by these investments. Some authors find that the market-to-book ratio seems inappropriate. Clarke and Shastri (2001) and Van Ness et al. (2001) use several alternative IA measures and find non-significant correlations between these measures and the market-to-book ratio. Clarke and Shastri (2001) find even a negative correlation between the PER and microstructure-based IA measures.

An alternative IA measure is the firm-specific return variation. Some authors declare that firms with high total return variance (Easley et al., 2003) or high specific return variance (Bhagat et al., 1985, Blackwell et al., 1990, Clark and Shastri, 2001, Van Ness et al., 2001) are more risky, thus less known by the market and more subject to the IA problem. Their results suggest that the firm's specific return variance is a noisy IA measure. Indeed, the specific return variance exhibits positive correlation with alternative microstructure-based IA measures but this correlation is most frequently non-significant. The firm's specific risk is even strongly negatively related with the negative informational component of the bid-ask spread (Clark and Shastri, 2001, Van Ness et al., 2001).

Other authors present an opposite picture of the firm's specific return variance. Morck, Yeung et Yu (2000) and Durnev et al. (2004) assert the firms with high specific return variation are characterised by a more *informative* price, which means less information asymmetry[3]. They show that firms with more specific return variation exhibit more efficient corporate investment. Wurgler (2000) shows that capital resources are more efficiently used in countries where stock returns are less synchronized with market returns. Bushman, Piotroski and Smith (2002) show that stock returns exhibit higher specific risk in countries with more developed financial-information systems and more independent financial press. Collins, Khothari and Rayburn (1987) and Durnev et al. (2001) document that specific return variation is a good predictor of firm's earnings per share in economic sectors where stock returns are less synchronized with market returns.

Microstructure-based IA measures

Some IA measures are derived from microstructure models. These measures basically decompose the bid-ask spread in a transitory component, related to the inventory and ordering costs, and an informational component. Microstructure-based IA measures have a solid theoretical background. They do not need long time-series of data and may thus measure the degree of IA at a given point in time. But these measures have received various critiques from the financial literature on a conceptual and an empirical ground. O'Hara (1995) declare that microstructure measures may concern only a small number of securities and cannot provide a precise IA measure for all of them. She also declares that it is difficult to clearly distinguish between the transitory and informational component of the bid-ask spread.

In line with this, George, Kaul et Nimalendran (1991) use traditional microstructure specifications and find an informational component which is about 10% of the bid-ask spread while Madhavan, Richardson et Roomans (1997) find an informational component of 40%.

[3] Durnev et al. (2004) suggest that *"in a given time interval and all else being equal, higher firm-specific variation stems from more intensive informed trading due to a lower cost of information, and hence indicates a more informative price … In a market with many risky stocks, during any given time interval, information about the fundamental value of some firms might be cheap, while information about the fundamental value of others might be dear. Traders, ceteris paribus, obtain more private information about the former and less about the latter. Consequently, the stock prices of the former, moving in response to informed trading, are both more active and more informative than the stock prices of the latter".*

Van Ness et al (2001) document that the informational component of the spread determined with six alternative microstructure specifications and the spread itself are very strongly correlated. The authors conclude that the models decomposing the spread give only an approximate indication of the private information. They conclude that the informational component of spread is rather a noisy transformation of the total spread. Neal and Wheatley (1998) reach similar conclusions. They declare that microstructure models are poorly specified for measuring the stocks' degree of IA.

An important microstructure-based IA measure is the probability of informed trading (PIN) proposed by Easley et al. (1996). This measure has been shown to explain a certain number of information-based regularities. Easley et al. (2002) show that PIN is significantly and positively correlated with stock returns. This result is consistent with one of the main conclusions of asset pricing models with privately informed investors: investors require an information-risk premium for holding securities on which they are uninformed. The information-risk premium engenders a positive correlation between the degree of private information and stock returns. Another important empirical result is that PIN seems to dominate traditional factors in explaining stock returns. Easley et al. (2002) conclude that informational effects on securities' markets are far more pervasive than previously documented. Chung and Li (2003) document a strong correlation between the informational component of the bid-ask spread and PIN. Chen et al. (2003) show that there is a strong correlation between the optimality of the investment decisions and the degree of private information measured by PIN. This variable is strongly correlated with the specific stock return variation used by Durnev et al. (2004) as an IA measure. Botosan and Plumlee (2003) use PIN as a measurement of information dissemination across investors. They document a strong correlation between information dissemination and the cost of the capital.

In spite of the sophistication inherent to the variable PIN, some analyses seem to disapprove this variable. Easley et al. (2003) declare that there still remain some unanswered questions of whether volume or momentum effects may not be proxying for some underlying components of PIN. Clarke and Shastri (2001) show that the informational component of the spread obtained with several microstructure models is not positively correlated with PIN. Aktas et al. (2003) analyzed the behaviour of PIN before and after acquisition operations that took place in Euronext stock market. Paradoxically, the value of PIN decreases before acquisitions and increases afterwards. This is contradictory with the assumption that this variable measures the degree of information asymmetry. The authors conclude that their results raise doubts about the capacity of PIN to measure the stocks' private information content. Another puzzling result is the one obtained by Chung and Li (2003), previously cited. The positive link between the optimality of investments and the PIN variable, together with the finding of a positive link between the PIN and the specific firm risk suggest that the PIN variable is rather a measure of price informativity and not a measure of information asymmetry. This is problematic because the price informativity is opposed to information asymmetry, which means that existing studies achieve opposing results with the PIN variable.

Our IA measure

While there is still much controversy about the relevance of extant IA measures, we propose a new IA measure derived *directly* from asymmetric-information asset-pricing models. We claim that our measure is less proxying for other unobserved variable because it directly compares homogeneous-beliefs with heterogeneous-beliefs model specifications.

The intuition underlying our measure is the following. In a homogeneous-beliefs economy, informed and uninformed investors possess the same information. At any given moment, the stock price contains no further information allowing anticipating stock returns. Thus, stock returns and prices are not correlated. In a market with asymmetrically informed investors, stock prices contain some private information, which is information not revealed by prices, about stock returns. Because this private information is not revealed by stock prices, it conditions stock returns. Thus, returns and prices are correlated. The degree of correlation between prices and returns may thus be used as a measure of the private information embedded in stocks.

Our approach is fundamentally based on rational expectation models with privately informed investors. One category of such models, which we qualify as "heterogeneous-information" models, consider that all investors possess different and independent pieces of private information about security payoffs (Grossman, 1976, Hellwig, 1980, Diamond and Verrecchia, 1981, Admati, 1985, Xu, 1999). By observing security prices at equilibrium (*public* information), investors receive some additional information on stock returns relatively to their own, private information. At equilibrium, prices *aggregate* investors' pieces of information and reveal all or part of the available information. In another class of models, which we qualify as "asymmetric-information" models, the information is rather *transmitted*, not aggregated (Grossman and Stiglitz, 1980, Wang, 1993, Jones and Slezak, 1999, Easley and O'Hara, 2001, Kodress and Pritsker, 2001). These models consider two categories of investors, informed and uninformed ones. By observing the equilibrium price, uninformed investors deduce all or part of the private information of informed investors.

The existence of private information in the above models is necessarily linked to the existence of noise, materialized by the existence of noise traders or liquidity needs (generating random supply). Grossman (1976) claims that the equilibrium price reveals *all* the relevant information if the market is not noisy[4]. Because there is no more extra-information available to investors relative to that obtained from prices, the stock return will depend on factors that are unrelated to those pertaining to the equilibrium price. Thus, the stock return will be independent of the equilibrium price. In this case, our AI measure takes the value of zero. If the market is noisy, investors will be unable to extract all the relevant information by observing equilibrium prices. Prices will then reveal *partially* the available information and some information will remain private. Because this private information affects concurrently stock prices and returns, these latter are correlated. Thus, our AI measure is not zero. We demonstrate that a higher noise engender a higher R^2 of the regression of stock returns on prices, *ceteris paribus*. Thus, the R^2 of this regression may capture the information asymmetry affecting stocks.

A simplified IA model for multi-securities markets and common factors

In order to illustrate the above discussion, we present a generalisation of the Grossman and Stiglitz (1980) asymmetric information model for a multi-securities markets and multiple factors generating returns. We consider concurrently *specific* and *common* factors for stock returns. This has important implications for determining the information sources to be used in the regressions for determining our AI measure.

There are *2* groups of investors, informed and uninformed ones, *n* risky securities and one risk-free security. Agents invest in $t=0$ and consume in $t=1$. The percentage of informed investors is denoted λ. Informed investors have superior information on stock returns which are generated by one specific and *k* common factors. The stock liquidation value writes:

[4] In this case, private information is not useful because asset prices are more informative. This destroys the incentive to buy costly information, an unstable condition (the Grossman and Stiglitz paradox).

$$\tilde{P}^1 = C\tilde{\eta} + \tilde{\varepsilon} \qquad (1)$$

where: $\tilde{P}^1$ is the n-vector of securities' payoffs in $t=1$; $\tilde{\eta} = \left(\tilde{\theta}, \tilde{f}\right)'$ is the $(n+k)$-dimensional column vector obtained by joining the vector of the specific payoff $\tilde{\theta}$ (also designated by specific factor) to the k-vector $\tilde{f}$ of common factors; C is the $(n+k, n)$ block matrix $\left(I_n, B\right)$ obtained by joining the matrix of factor loadings B to the n-dimensional identity matrix I_n; $\tilde{\varepsilon}$ is a random n-vector unknown by all investors, also qualified as an error term. All the variables considered are jointly multivariate normal and $\tilde{\theta}$, $\tilde{f}$ and $\tilde{\varepsilon}$ are independent. Denote Q the variance-covariance matrix of $\tilde{\eta}$, T the variance-covariance matrix of $\tilde{\theta}$, $F=BB'$ the variance-covariance matrix of $\tilde{f}$ (all factors are independent, have zero mean and unit variance) and E the variance-covariance matrix of $\tilde{\varepsilon}$. Because specific factors are independent, the matrix T is diagonal.

As is common in such models, the asset supply per investor is a random vector $\tilde{z}$. This vector is independent and joins the normal distribution of the other variables in the model. The assumption that the supply $\tilde{z}$ is unknown by investors is necessary for making equilibrium prices partially (not perfectly) revealing. The randomness of the supply may be justified by the existence of non-speculative trades or the existence of liquidity-traders. Denote the supply variance-covariance matrix by Z. We assume that E, T and Z are regular, which is necessary for insuring the existence and the uniqueness of the equilibrium price. In addition, we assume that investors have constant absolute risk aversion, implying that their demand is independent of their initial wealth. For simplicity, we also consider that the risk aversion coefficients are identical for all investors[5], say $a>0$.

Informed investors know the realisation of the specific and common factors generating returns. More formally, they know the realisation of the random variable $\tilde{\eta}$. Uninformed investors know only the distribution of $\tilde{\eta}$. Our problem is to find the equilibrium price $\tilde{P}_0$ in $t=0$ for the setting described above. Jimenez (2004) demonstrates that under the above

[5] The benefit of considering different risk aversion coefficients would be too small for understanding the model.

conditions there exists a unique closed-form solution for $\tilde{P}_0$ within the class of linear functions of $\tilde{\eta}$ and $\tilde{z}$ of the form:

$$\tilde{P}_0 = A_0 + A_1\tilde{\eta} - A_2\tilde{z} \tag{2}$$

where A_2 is non-singular and:

$$A_1 = \left(CQC' + E - V_m\right)\left(CQC'\right)^{-1} C \tag{3}$$

$$A_2 = a\left(CQC' + E - V_m\right)\left(\lambda CQC'\right)^{-1} E \tag{4}$$

$$A_0 = \left(I_n - A_1\right)\overline{\eta} + \left(A_2 - aV_m\right)\overline{z} \tag{5}$$

In the above equations, V_m is the market average return variance, $V_m = \left[\lambda V_I^{-1} + \left(1-\lambda\right)V_{NI}^{-1}\right]^{-1}$, where V_I and V_{NI} are the conditional stock returns variance-covariance matrices of the informed and respectively the uninformed investors. Jimenez (2004) shows that:

$$V_m = \left(E + CNC'\right)\left(I_n + \lambda E^{-1}CNC'\right)^{-1}$$

where $N = Var\left(\tilde{\eta} \mid \tilde{P}^0\right) = Q - QK'M^{-1}KQ$ is the conditional variance-covariance matrix of the factor $\tilde{\eta}$ for the uninformed investor, $M = KTK' + Z$ and $K = a^{-1}\lambda E^{-1}C$.

The IA measure: Conceptual ground

Equation (2) shows that the current equilibrium price is a linear function of the informed information $\tilde{\eta}$ and the supply. Because the supply is unknown, the uninformed investor is not able to infer the realisation of $\tilde{\eta}$ and thus some information remains private. We define the degree of information asymmetry for a stock *i* as the amount of private information that is *un-revealed* by prices relative to the total amount of initial private information. The first quantity is the total return variance minus the return variance conditioned by observing the prices of all existing stocks, $Var\left(\tilde{P}_1^i - \tilde{P}_0^i\right) - Var\left(\tilde{P}_1^i - \tilde{P}_0^i \mid \tilde{P}_0\right)$. The return is considered here as the difference between the price in *t=1*, which is the liquidation value, and the equilibrium price in *t=0*. The second quantity is the total, unconditional return variance, $Var\left(\tilde{P}_1^i - \tilde{P}_0^i\right)$. Our measure of private information for the stock *i*, AIM, is thus defined by:

$$\text{AIM} = 1 - \frac{Var\left(\tilde{P}_1^i - \tilde{P}_0^i \mid \tilde{P}_0\right)}{Var\left(\tilde{P}_1^i - \tilde{P}_0\right)} \tag{6}$$

Our measure obtains thus by projection of returns on beginning-of-period prices. AIM is bounded within the interval *[0, 1]*. If AIM equals 0, the conditional and unconditional return variance in (6) are equal. There is no private information on the stock *i* because the price sends no information about stock returns. This may be possible in two cases. The first is when there is simply no private information at all (homogenous information). The second is when all the available information has been revealed by prices and there is no extra-information allowing anticipating stock returns. If AIM is strictly positive, the conditional return variance is lower than the unconditional one in (6). This is because some private information is revealed by prices but some other private information remains in the hands of informed investors. The latter allows revising to a certain extent the return distribution. In other words, returns and equilibrium prices are correlated.

It is important to note that our measure involves *returns* on the period conditioned on prices at the beginning of the period. Our measure *does not consider future prices* $\tilde{P}^1$ conditioned on equilibrium prices $\tilde{P}^0$. The regression of end-of-period prices on beginning-of-period prices is not relevant for measuring the degree of private information. This regression measures rather the price informativity. We have demonstrated that end-of-period prices are always positively correlated with beginning-of-period prices, even in the absence of private information. We have also demonstrated that a higher degree of information asymmetry implies a lower correlation between end-of-period prices and beginning-of-period prices. On the opposite, returns are always negatively related to prices in our model. We have also demonstrated (see Jimenez, 2004) that a higher degree of information asymmetry implies a higher correlation between returns and prices. Thus, the regression of returns on prices is relevant for capturing the degree of information asymmetry.

Empirical specification

One fundamental hypothesis in the model is that prices are normally distributed, which is an unrealistic, yet necessary hypothesis. This hypothesis is fundamental for finding closed-form solutions for asset prices. In order to create a link between our model and the real world, we will rather consider logarithms of prices, which are more close to the hypothesis of normality. With this transformation, the returns in our model equal the logarithm of the ratio between end-of-period and beginning-of-period prices. This definition is close to the real definition of returns. The implementation of our AIM measure is thus obtained by projection of real stock returns on beginning-of-period stock prices[6], as in Biais et al. (2003). More precisely, we propose the following regression model for the projection:

$$R_{i,t} = \alpha_i + \sum_{j=1}^{n} \beta_i^j P_{j,t-1} + \varepsilon_{i,t} \tag{7}$$

where $R_{i,t}$ is the return of security *i* on period *t* and $P_{j,t}$ is the price of security *j* at the beginning of period *t*. The sum of the residual squares from (7) measures the ratio between the conditional and unconditional return variance while the regression R^2 is exactly our measure AIM specified in (6). The higher the regression R^2, the stronger the link between returns and prices, thus the stronger the degree of private information embedded in stocks.

The simplicity of our AIM measure is attributable to the low number of parameters which intervene in its calculation. But this measure involves a practical problem of implementation. Theoretically, the price of any stock *j* correlated with the target stock *i* contains some information on the return of the latter. A judicious extraction of information would thus require the study of the correlations between all stocks. This is impossible because of the large number of regressors and the limited data on prices and returns. For this reason, we choose some portfolios that are strongly correlated with the target security as information sources in the specification (7).

More precisely, our information sources are the price of the target stock, the price of the portfolio of all stocks belonging to the same two-digit SIC industrial classification, the price of the portfolio of all stocks belonging to the same three-digit SIC industrial classification and the price of the market portfolio. These two industrial portfolios are among the most used in the financial literature (see Durnev et al., 2001, 2004, Chen et al., 2003). The price of the target stock *i* is the most informative on the return generated by the stock's specific factor. The prices of industry portfolios are the most informative on industrial factors affecting the returns of the target stock. Jimenez (2004) has shown that portfolios of stocks that share the

---

[6] Considering logarithms of stock prices instead of stock prices as explanatory variables does not change fundamentally our conclusions.

same common factor are more informative than individual stocks from these portfolios[7]. Finally, the price of the market portfolio is the most informative on macro-economic factors affecting all stocks. Jimenez(2004) has demonstrated in a rigorous manner the relevance of industry portfolios as fruitful information sources allowing uninformed investors to anticipate stock returns.

To check the robustness of our results we use several specifications and test the sensitivity of the results obtained. The simplest specification is based on the projection of returns on the prices of the target security and the closest, three-digit SIC portfolio:

$$R_{i,t} = \beta_{i,0} + \beta_{i,1} P_{i,t-1} + \beta_i^{SIC3} P_{i,t-1}^{SIC3} + \varepsilon_{i,t} \qquad (8)$$

where $R_{i,t}$ is the return of security *i* on the period *t*, $P_{i,t-1}$ is the security *I* price at the beginning of the period *t* and $P_{i,t-1}^{SIC3}$ is the price of the industrial portfolio containing all stocks belonging to the same three-digit SIC classification. In order to eliminate spurious correlations between regressors, we eliminate the stock *i* from the corresponding industrial portfolio. The consideration of additional sources of information should increase the precision of the calculation of the degree of private information. In order to enrich the extraction of information, we also considered the two-digit industrial portfolio. The specification becomes:

$$R_{i,t} = \beta_{i,0} + \beta_{i,1} P_{i,t-1} + \beta_i^{SIC2} P_{i,t-1}^{SIC2} + \beta_i^{SIC3} P_{i,t-1}^{SIC3} + \varepsilon_{i,t} \qquad (9)$$

where $P_{i,t-1}^{SIC2}$ is the price of the industrial portfolio containing all stocks belonging to the same two-digit SIC classification. In order to eliminate spurious correlations between regressors, we eliminate the stock *i* from the corresponding industrial portfolios and the three-digit SIC portfolio from the two-digit SIC portfolio. Alternatively, we consider the price of the market portfolio instead of the two-digit industrial portfolio. This allows extracting the information on economic factors that are common to all securities. The two-digit industrial portfolio has been excluded in order to diminish the number of regressors for accurate regression estimation. The third specification becomes:

[7] Of course, the projection of returns on all portfolio stocks would be more informative but, as we have mentioned, it is impossible to realize practically.

$$r_{i,t} = \beta_{i,0} + \beta_{i,1} P_{i,t-1} + \beta_i^{SIC3} P_{i,t-1}^{SIC3} + \beta_i^m P_{i,t-1}^m + \varepsilon_{i,t} \tag{10}$$

where $P_{i,t-1}^m$ the price of the market portfolio. As above, the market portfolio excludes the stocks of the two-digit SIC portfolio. In order to ameliorate the econometric properties of our AIM measures which are bounded in the *[0, 1]* interval, we apply a logistic transformation to the R^2 from the above regressions. Our AIM measures for a stock *i* will thus be:

$$AIM_i = \ln\left(\frac{R_i^2}{1-R_i^2}\right) \tag{11}$$

In order to compare our AIM measures derived from the specifications above with traditional ones, we use in addition the firm-specific return variation as in Durnev et al. (2001), Durnev et al. (2004) and Chen et al. (2003). This return variation is obtained by controlling the total return variation for market and industrial movements. Durnev et al. (2004) propose the following specification:

$$R_{i,t} = \beta_{i,0} + \beta_i^{SIC3} R_{i,t}^{SIC3} + \beta_i^m R_t^m + \varepsilon_{i,t} \tag{12}$$

where $R_{i,t}^{SIC3}$ is the return of the three-digit industrial portfolio excluding the target security *i* and R_t^m is the return of the market portfolio. The regression (12) breaks up the total return variation into a systematic part $\beta_i^{SIC3} r_{i,t}^{SIC3} + \beta_i^m r_t^m$ and a specific part $\varepsilon_{i,t}$. If the specific variation is important, it means that there is significant firm-specific information and a significant part of this information is transmitted by prices. The degree of private information is denoted by *RS* and is measured by the residual variance relative to the total variance, which is precisely $1-R^2$. It is important to note that our approach is totally different because we use prices at the beginning of the period instead of returns on the period and we consider the R^2 not the $1-R^2$ from the corresponding regressions. In order to ameliorate the econometric properties of the RS measures, we apply a logistic transformation to $1-R^2$ from the regression (12).

$$RS_i = \ln\left(\frac{1-R_i^2}{R_i^2}\right) \tag{13}$$

It follows that $RS_i = \ln\left(\sigma^2_{\text{résid},i}\right) - \ln\left(\sigma^2_{\text{expl},i}\right)$ where $\sigma^2_{\text{résid},i}$ is the firm-specific return variation and $\sigma^2_{\text{expl},i}$ is the explained, firm-systematic return variation. The reader can easily notice that our AIM measures reverse the position of the explained and residual variance.

Data and sample description

Data

The data were extracted from CRSP US Daily Database Stock. We calculate our AIM measures with weekly returns for the analysis period which is from January 1, 1985 and this until December 31, 2002 (17 years). The regressions (8) - (10) are estimated with weekly returns on a period of two years, which represents 104 weekly returns. The resulting AIM measure represents the degree of private information of the firm in the middle of this period. The two-year periods are moved month by month, allowing us to determine AIM for each month on the corresponding analysis period. AIM data is thus available for 6 988 common stocks representing are 584 347 "stocks/month" observations on the period between January 1, 1986 and this until December 31, 2001 (15 years). These stocks cover 79 sectors (two-digit SIC) and 408 industries (three-digit SIC). We retain only class A ordinary stocks by eliminating all stocks which CUSIP code does not terminate with figure 10.

The use of weekly returns for estimating AIM is a compromise. The use of daily returns allows more accurate estimation of AIM since the information is captured more rapidly and the available data allows using smaller estimation periods. The problem with using such data is “thin-trading”, engendering identical prices in CRSP (which frequently reports the last price if it is unavailable for a given day) and zero returns. This may result in a strong over-estimation of AIM for small securities, more subject to the problem of “thin-trading”. At the other extreme, the use of monthly data would strongly reduce the number of observations and would enlarge the estimation period for AIM, thus making the calculation of the latter less precise.

Stocks’ and portfolios’ prices have been determined by compounding adjusted returns on the estimation period. For each stock or portfolio, their price represent the value of a buy-and-hold strategy consisting in investing one dollar at the beginning of the estimation period, the dividends being reinvested each week. The sector/industry portfolios include all NYSE, AMEX and NASDAQ stocks pertaining to the same sector/industry. In order to choose the type of stocks in these portfolios, we have followed the same criteria as the one used by CRSP in constructing the CRSP Value-Weighted Index. In addition, larger portfolios exclude stocks from smaller portfolios in order to diminish the spurious correlations between regressors in the specifications (8) - (10) for our AIM measure and in (12) for the RS measure. This correlation may be important if the target security has a high equity market-value or if there is a small number of firms in the corresponding industry.

Sample description

The following table presents descriptive statistics for our three AIM measures based on the specifications (8) - (10) and for the RS measure based on (12). We have provided the results for the raw measures and for their logistic transformations.

Table 1
Descriptive statistics for private information measures

The table presents the number of observations, the value-weighted average, the median, the minimum, the maximum, the first quartile, the third quartile as well as the standard deviation of private information measures. RS indicates the firm-specific return variation. AIM measures are obtained by projecting returns on prices. AIM1, AIM2 and AIM3 correspond respectively to the specifications (8) - (10). L(RS), L(AIM1), L(AIM2), L(AIM3) are obtained by logistic transformation of initial measures. In Part A of the table, the observations represent couples of stocks/months. In Part B, the observations are stocks. The private information for each stock is determined as the mean of private information during the entire stock analysis period.

Part A

	Number of obs.	Average	Median	Min	Max	First quartile	Third quartile	Variation type
RS	584 347	0.866	0.922	0.000	1.000	0.811	0.972	0.147
AIM1	584 347	0.052	0.045	0.000	0.573	0.026	0.070	0.036
AIM2	584 347	0.071	0.064	0.000	0.574	0.042	0.093	0.041
AIM3	584 347	0.073	0.066	0.000	0.581	0.044	0.094	0.041
L(RS)	584 339	2.577	2.464	-8.740	11.513	1.454	3.542	1.605
L(AIM1)	584 315	-3.186	-3.057	-11.513	0.295	-3.609	-2.592	0.912
L(AIM2)	584 341	-2.752	-2.681	-11.513	0.297	-3.134	-2.280	0.708
L(AIM3)	584 341	-2.716	-2.655	-10.414	0.328	-3.088	-2.268	0.673

Part B

	Number of obs.	Average	Median	Min	Max	First quartile	Third quartile	Variation type
RS	6989	0.882	0.924	0.082	1.000	0.841	0.964	0.117
AIM1	6989	0.051	0.048	0.001	0.338	0.038	0.060	0.021
AIM2	6989	0.069	0.067	0.003	0.347	0.055	0.081	0.024
AIM3	6989	0.071	0.069	0.004	0.352	0.057	0.082	0.024
L(RS)	6989	2.737	2.755	-3.980	8.290	1.834	3.661	1.309
L(AIM1)	6989	-3.230	-3.189	-7.930	-0.891	-3.475	-2.920	0.536
L(AIM2)	6989	-2.795	-2.766	-6.257	-0.805	-2.996	-2.547	0.435
L(AIM3)	6989	-2.754	-2.731	-5.683	-0.794	-2.949	-2.523	0.410

The specific return variation for the companies considered in our sample accounts for on average 87% of the total risk. Roll (1988) found a value of 76% for its sample over the period 1982–1986. Over this period, the number of small companies was relatively weaker relative to our recent analysis period. These companies have generally a higher specific risk, which may explain the difference between the results. The studies of Durnev et al. (2001) or Durnev et al. (2004) find an average firm-specific return variation of about 80% but these studies consider also companies of bigger size. Indeed, their samples situates at the intersection between the CRSP and Compustat databases, the latter being less exhaustive on small companies.

The results for our measures AIM show that the information extracted from prices makes it possible to reduce the uncertainty about future returns of about 7% with the specification including the sector and industry portfolio. The addition of the market portfolio does not allow a significant increase in the information extracted from prices. We also observe that the traditional RS measure is more dispersed than our AIM measures. From this point of view, AIM is closely related to the PIN measure of Easley et al. (2002) who observes that this one has a weak dispersion across companies and a low variability in time.

Tableau 2 presents the correlations between the private information measures, the firm size and the volume of transactions. The firm size is approached by the logarithm of the market value of equity. The volume of transactions is approached by the logarithm of number of security traded each month (VOL1) and turnover, which is the monthly volume divided by shares outstanding at the end of month.

Tableau 2
Correlation between private information measures

Correlation between the logarithm of market value of firm equity (SIZE), the logarithm of the volume of transactions (VOL1), turnover (VOL2), the firm-specific return variation (RS) and our measures of private information AIM1, AIM2 and AIM3. RS indicates the firm-specific return variation. AIM measures are obtained by projecting returns on prices. AIM1, AIM2 and AIM3 correspond respectively to the specifications (8) - (10). Turnover is the monthly volume divided by shares outstanding at the end of month. The observations represent 584 347 stocks/months in Panel A and 6 988 stocks in Panel B. Non-significant correlations are presented in italic.

Pannel A

	SIZE	VOL1	VOL2	RS	AIM1	AIM2	AIM3
SIZE	1.00	0.78	0.11	-0.64	-0.06	-0.05	-0.06
VOL1	0.78	1.00	0.34	-0.56	-0.08	-0.09	-0.10
VOL2	0.11	0.34	1.00	-0.11	-0.04	-0.05	-0.06
RS	-0.64	-0.56	-0.11	1.00	0.07	0.06	0.08
AIM1	-0.06	-0.08	-0.04	0.07	1.00	0.75	0.70
AIM2	-0.05	-0.09	-0.05	0.06	0.75	1.00	0.72
AIM3	-0.06	-0.10	-0.06	0.08	0.70	0.72	1.00

Pannel B

	SIZE	VOL1	VOL2	RS	AIM1	AIM2	AIM3
TAILLE	1.00	0.81	0.21	-0.73	-0.05	-0.03	-0.05
VOL1	0.81	1.00	0.45	-0.64	-0.12	-0.12	-0.15
VOL2	0.21	0.45	1.00	-0.24	-0.12	-0.12	-0.15
RS	-0.73	-0.64	-0.24	1.00	0.07	0.06	0.08
AIM1	-0.05	-0.12	-0.12	0.07	1.00	0.81	0.77
AIM2	-0.03	-0.12	-0.12	0.06	0.81	1.00	0.77
AIM3	-0.05	-0.15	-0.15	0.08	0.77	0.77	1.00

Tableau 2 presents encouraging results. All the correlations are highly significant and present the anticipated signs. The majority of the correlations become more significant by considering observations aggregated by stocks (Pannel B). The correlations between private information measures and size are negative. This result is traditional: large companies, for their majority, are better known by the market. As expected, the correlation between private information measures and the volume of transactions is negative, whichever the proxy used for volume. Easley et al. (1996), Datar, Naik and Radcliffe (1998), Chordia, Subrahmanyam and Anshuman (2001) and Easley et al. (2202) obtain similar results. All these authors assert that a high volume of transactions is indicative of more transparent information.

The correlations between the alternative private information measures are all positive. We note also a high correlation of the traditional RS measure with size and volume, especially with size. This result may also be explained by large firms having lower total risk and higher systematic risk. We have controlled the size effect by calculating partial correlations between private information and volume. Unexposed results show that private information is significantly negatively correlated with volume, as expected, and that all private information measures are positively correlated. After controlling for size effects, the correlation between private information and volume becomes quite constant for the alternative measures used.

Asset pricing tests

This section considers the impact of private information on asset prices. We want to assess the significance of our AIM measure and its importance relative to traditional factors affecting stock returns. To allow the comparison with former studies and across different specifications, we will consider several CAPM specifications. The most traditional considers only one factor, the systematic risk factor β. Other more sophisticated versions include the size (SMB) and book-to-market (HML) factors as suggested by Fama and French (1992). The loadings for these factors will be denoted β^{SMB} and β^{HML} . The theory teaches us that we should observe a positive correlation between stock returns and these factor loadings.

Our tests are based on the well-known methodology of Fama and MacBeth (1973). We will estimate traditional CAPM versions and compare them with our informational version including the private information factor. According to the theory, we should also observe a positive correlation between the degree of private information and expected returns because the latter contain an information-risk component. Each month, the cross-section of stock returns is regressed on factor loadings and stocks' extent of private information. The time-series mean of the estimated coefficients from these regressions allow determining if investors price the corresponding factors.

Informal tests

We first check the reasonableness of our AIM measures by computing excess returns and other parameters for 12 portfolios of stocks sorted independently according to their extent of private information (four portfolios) and size (three portfolios). The private information is approached by the traditional firm-specific return variation (RS) and our three AIM measures. The results are presented in the following table. The size is approached by the logarithm of firm equity capitalization at the end of the month. We rebalance our portfolios for each month on the analysis period between January, 1986 and December, 2001 then we compute equally-weighted averages of parameters on this period.

A first group of returns are computed in excess of the risk-free rate and the other group are computed in excess of the risk-free rate and the systematic risk premium (alpha). According to the theory, we should observe a positive correlation between excess returns and the extent of private information. The classification by size allows analyzing the link between excess returns and private information distinctly for small and large firms[8]. This link should be stronger for small, unknown firms. The other parameters are the volume of transactions and the traditional factor loadings β, β^{SMB} and β^{HML}. Because these factor loadings are estimated variables, we have computed them by using the Fama French (1992) methodology in order to diminish the error-in-variable (EIV) bias[9]. The results are presented in the following table.

Tableau 3

The table presents arithmetic mean of several parameters for 6 988 stocks between 01/01/1986 and 31/12/2001 on the American markets NYSE, AMEX and NASDAQ. The parameters are: returns in excess to the risk-free rate approached by the rate of return on 30 days T-Bills (Panel A); the stocks alpha (Panel B); the volume of transactions approached by the stock turnover ratio, i.e. the ratio of volume on shares outstanding (Panel C); the loadings of the systematic, size and market-to-book factor β, β^{SMB} and β^{HML} computed with the methodology proposed by Fama and French (1992) (Panel D, E and F). The observations represent 584347 stocks/months. The size is approached by the logarithm of the equity market value. The category 1 represents the smallest size and respectively the smallest degree of private information. RS indicates the firm-specific return variation. AIM measures are obtained by projecting returns on prices. AIM1, AIM2 and AIM3 correspond respectively to the specifications (8) - (10).

Panel A: Excess return (%)

Measure	Category	Small	Medium	Large
RS	1	3.978	2.322	0.962
	2	3.316	1.093	0.343
	3	2.016	0.297	-0.022
	4	1.542	-0.384	-0.449
AIM 1	1	0.761	0.266	1.091
	2	1.903	0.792	0.680
	3	2.392	0.720	0.412
	4	2.438	0.829	0.514

[8] Existing studies show that private information effects are less obvious for large firms.

[9] We first determine *pre-ranking* stock betas for each month *t* by regressing excess returns on the return of the CRSP Value-Weighted Index on a minimum 5 years period before month *t*. Our analysis period starts in January 1980 in order to have sufficient data for subsequent analyses. Then, for each month *t* we sort 10 portfolios based on size and then, inside each size deciles we sort 10 portfolios based on the *pre-ranking* stock betas. For each month we thus form 100 portfolios based on size and *pre-ranking* stock betas. Then we compute the return of each portfolio *i* at month *t* as the equally-weighted average of stock returns that enters portfolio *i* at month *t-1*. The second step is to regress portfolio excess returns on the excess returns of the CRSP Value-Weighted Index on the entire analysis period. This procedure allows us determining *post-ranking* portfolio betas. Thus, we have 100 betas estimated for each one of the 100 portfolios on the entire analysis period. The beta of each stock *i* and for each month *t* is the beta of the portfolio that contained the stock *i* during month *t*. Because the portfolio compositions change each month, individual stocks betas vary over time. We have used the same procedure to compute the factor loadings for the size and book-to-market factor.

AIM 2	1	0.821	0.180	0.975
	2	2.015	0.808	0.607
	3	2.498	0.926	0.530
	4	2.245	0.685	0.596
AIM 3	1	0.744	0.181	1.031
	2	2.174	0.645	0.625
	3	2.506	0.987	0.522
	4	2.145	0.780	0.505

Panel B: Stocks' Alpha (%)

Measure	Category	Small	Medium	Large
RS	1	3.102	1.496	0.247
	2	2.503	0.301	-0.340
	3	1.374	-0.387	-0.635
	4	0.998	-1.028	-1.026
AIM 1	1	0.227	-0.405	0.287
	2	1.265	0.034	-0.063
	3	1.724	-0.078	-0.199
	4	1.841	0.142	-0.093
AIM 2	1	0.202	-0.503	0.191
	2	1.411	0.034	-0.112
	3	1.856	0.131	-0.124
	4	1.666	0.023	-0.008
AIM 3	1	0.170	-0.536	0.272
	2	1.542	-0.092	-0.101
	3	1.858	0.218	-0.133
	4	1.558	0.088	-0.116

Panel C: Volume of transactions (stock turnover)

Measure	Category	Small	Medium	Large
RS	1	12.049	13.690	10.640
	2	9.169	10.022	10.044
	3	6.266	7.961	9.021
	4	4.948	6.614	8.032
AIM 1	1	6.866	10.136	10.675
	2	6.437	9.708	10.943
	3	5.754	8.843	10.276
	4	4.928	7.469	9.033
AIM 2	1	7.019	10.229	10.873
	2	6.216	9.597	10.820
	3	5.886	8.927	10.202
	4	4.846	7.456	9.019
AIM 3	1	7.180	10.632	11.227
	2	6.079	9.287	10.854
	3	5.868	8.727	10.021
	4	4.885	7.572	8.727

Panel D: Beta

Measure	Category	Small	Medium	Large
RS	1	0.951	1.233	1.091
	2	1.006	1.157	1.049
	3	0.895	1.015	0.925
	4	0.770	0.857	0.791
AIM 1	1	0.882	1.078	1.074
	2	0.876	1.079	1.085
	3	0.841	1.066	1.061
	4	0.774	0.990	1.003

AIM 2	1	0.894	1.088	1.083
	2	0.864	1.082	1.081
	3	0.839	1.068	1.062
	4	0.775	0.977	0.997
AIM 3	1	0.900	1.102	1.080
	2	0.872	1.074	1.076
	3	0.836	1.052	1.060
	4	0.769	0.987	1.007

Panel E: Beta SMB

Measure	Category	Small	Medium	Large
RS	1	0.964	1.048	0.259
	2	1.080	1.013	0.522
	3	0.950	0.942	0.510
	4	0.852	0.837	0.470
AIM 1	1	0.907	0.982	0.341
	2	0.928	0.974	0.383
	3	0.916	0.953	0.354
	4	0.896	0.912	0.336
AIM 2	1	0.905	0.990	0.349
	2	0.907	0.967	0.371
	3	0.908	0.949	0.357
	4	0.922	0.917	0.338
AIM 3	1	0.952	0.992	0.319
	2	0.907	0.959	0.372
	3	0.896	0.945	0.370
	4	0.894	0.926	0.358

Panel F : Beta HML

Measure	Category	Small	Medium	Large
RS	1	0.230	0.092	0.092
	2	0.260	0.169	0.223
	3	0.264	0.255	0.312
	4	0.301	0.272	0.327
AIM 1	1	0.241	0.202	0.113
	2	0.253	0.198	0.125
	3	0.304	0.209	0.156
	4	0.324	0.233	0.222
AIM 2	1	0.245	0.192	0.105
	2	0.266	0.203	0.137
	3	0.303	0.200	0.166
	4	0.313	0.246	0.205
AIM 3	1	0.244	0.180	0.066
	2	0.270	0.211	0.151
	3	0.299	0.216	0.181
	4	0.313	0.234	0.220

The data reveal a significant link between size, private information and excess stock returns or alpha. With some rare exceptions, stock excess returns decrease with the size, which confirms the results obtained by preceding studies. This reduction seems more marked for the companies with strong degrees of private information. An important result is that the link between excess returns and the degree of private information is different according to whether this last is measured by the firm-specific return variation (RS) or our AIM measure. The correlation between RS and excess returns is negative, opposite to that expected. On the other hand, the correlation between our AIM measures and excess returns is positive, as expected, for small and medium firms. In addition, the rise in excess returns is more important for small firms as we move from low to high degrees of private information, which is consistent with the private information hypothesis. For large firms, the correlation is however ambiguous. Our results resemble those obtained by Easley et al. (2002) with the PIN variable. These authors affirm that for large firms the private information may have a lower impact on stock prices.

The relation between private information and volume is negative whichever the measure used and whichever the firm size. This relation seems stronger for the RS measure. In addition, volume is positively correlated with size whichever the degree of private information. All these results support the private information measures. The factor loadings are also, to a certain extent, related to private information. The systematic risk factor loading β is higher when there is more private information. For the RS measure, this relationship may be independent of the private information content because this measure is by definition negatively correlated to systematic risk. For the AIM measures, our results confirm the intuition of Durnev et al. (2004). These authors assert the firms affected by large information asymmetries are less sensitive to specific information and more sensitive to systematic return variations. It may explain why the systematic risk is higher in emerging countries. Finally, the size and market-to-book factor loadings support also the private information hypothesis. Firms having higher degrees of private information are smaller and exhibit higher market-to-book ratios.

Traditional CAPM tests

This section presents the parameter estimates for several traditional CAPM specifications. This analysis is intended to check the robustness of the results and compare these results with those obtained in other similar studies. The second objective is to compare the corresponding results with those obtained with the specification considering the private information factor. The first model we estimate is the traditional CAPM of Sharpe and Lintner. For each month over the period of analysis between 01/01/1986 and 31/12/2001, we estimate the following regression model:

$$R_{it} = \gamma_{ot} + \gamma_{1t}\beta_{i,t-1} + \varepsilon_{it} \quad (14)$$

where R_{it} is the excess return of security *i* at month *t* and $\beta_{i,t-1}$ is the stock's beta determined with the methodology of Fama and French (1992). The regression (14) is estimated for each cross-section of monthly returns and then the coefficient estimates are averaged through time with the Fama-MacBeth (1973) procedure. As in Easley et al. (2002), we apply the correction of Litzenberger and Ramaswamy (1979) who suggest weighting the coefficient estimates in time-series averages by their precision in the cross-sectional regressions. This correction improves the Fama-MacBeth (1973) procedure which is inefficient under time-varying volatility. The coefficient of interest in this specification is γ_{1t}. According to the theory, this coefficient should be positive and significant.

To improve the performance of the model and to allow the comparison with former studies, we add the size of the companies as an explanatory variable as suggested by Fama and French. This two-factor specification is as follows:

$$R_{it} = \gamma_{ot} + \gamma_{1t}\beta_{i,t-1} + \gamma_{3y}Size_{i,t-1} + \varepsilon_{it} \quad (15)$$

where $Size_{i,t-1}$ represents the logarithm of the firm's equity market value at the end of the month *t*. Next, we explore the three factors Fama-French model including the market-to-book factor. Because we do not possess data on the market-to-book ratio, we use the specification considering factor loadings. Fama and French (1996) and Jagannathan and Wang (1996) declare that this approach is more relevant for capturing the impact of factor risk-premiums on stock returns. We follow the literature and estimate the following three-factor model:

$$R_{it} = \gamma_{ot} + \gamma_{1t}\beta_{it} + \gamma_{3y}\beta_{it}^{SIZE} + \gamma_{4t}\beta_{it}^{BM} + \varepsilon_{it} \quad (16)$$

where $\beta_{i,t-1}^{SIZE}$ et $\beta_{i,t-1}^{BM}$ are the factor loadings for the size and market-to-book factors. These loadings have been determined for each stock with the Fama-French procedure (see the previous section). The results are exposed in the following table.

Tableau 4

The table presents the average coefficients from the regression of stock excess returns on traditional factors of risk. The methodology is that proposed by Fama and MacBeth (1973). The regressions cover 538884 stocks/months over the period between 01/01/1986 and 31/12/2001 on the NYSE, AMEX and NASDAQ markets. The cross-sectional regressions are corrected for the heteroscedasticity. The results correspond to the specifications (14), (15) and (16).

Coefficients	γ_o	γ_1	γ_2	γ_3	γ_4	R^2 (%)
Estimation	1.38	0.04				
t-value	5.49	0.19				1.27
p-value	0.00	0.84				
Estimation	2.56	0.12		-0.25		
t-value	5.38	0.50		-3.55		2.39
p-value	0.00	0.61		0.00		
Estimation	1.31	-0.13		0.45	-0.15	
t-value	6.14	-0.90		1.45	-0.68	5.6
p-value	0.00	0.36		0.14	0.49	

The traditional specification (14) has a low explanatory power, with an R^2 of only 1,27%. The mean coefficient of β is only 0,04 with a probability of 0,84. As in the former studies, the traditional MEDAF is seriously called into question. As an example, Jagannathan and Wang (1996) find a mean R^2 of only 1,35%. The Fama-French two factor specification has a higher explanatory power, with an R^2 of 2,39% but the β coefficient is still not significant. On the contrary, the effect of the size factor on excess returns is negative and highly significant, with a Student *t* of –3,55. We thus find the traditional result that expected returns are higher for small companies. Finally, the three-factor Fama-French model increases quite substantially the model R^2 but the coefficients on the three factor loadings are all non-significant. These results are different from those obtained by previous studies but are close to those obtained in more recent studies. For example, Easley et al. (2002) obtain a positive and significant value for the coefficient of the SMB factor loading and a non-significant value for that of the HMB factor loading. These results seem to raise questions about the validity of the traditional CAPM and its factor Fama-French versions.

CAPM tests considering private information

This section analyzes the impact of private information on stock returns. The simplest specification is directly deduced from our model. It breaks up the stock returns into a systematic and an informational component:

$$R_{it} = \gamma_{ot} + \gamma_{1t}\beta_{i,t-1} + \gamma_{2t}INFO_{i,t-1} + \varepsilon_{it} \qquad (17)$$

where $INFO_{i,t-1}$ represents the degree of private information for the security *i* during the month *t-1*. The second specification we propose controls for the size factor:

$$R_{it} = \gamma_{ot} + \gamma_{1t}\beta_{i,t-1} + \gamma_{2t}INFO_{i,t-1} + \gamma_{3y}Size_{i,t-1} + \varepsilon_{it} \qquad (18)$$

The parameter estimates are presented in the following table.

Tableau 5

The table presents average coefficients from the regression of stock excess returns on factors of risk. The methodology is that proposed by Fama and MacBeth (1973). The regressions cover 538 884 stocks/months for which data are available over the period between 01/01/1986 and 31/12/2001 on the NYSE, AMEX and NASDAQ American markets. The estimates are, in order, the averages of the coefficients estimated for the constant term, β, the informational factor, the SMB and HML factor loadings. The factor loadings are determined following the methodology suggested by Fama and French. The cross-sectional regressions are corrected for heteroscedasticity.

Panel A : information approached by the specific risk (RS)

Coefficients	γ_o	γ_1	γ_2	γ_3	γ_4	R^2 (%)
Estimation	1.60	0.01	-0.08			
t-value	4.75	0.07	-0.97			2.13
p-value	0.00	0.94	0.33			
Estimation	5.65	-0.03	-0.61	-0.54		
t-value	8.19	-0.15	-7.64	-6.71		2.90
p-value	0.00	0.88	0.00	0.00		

Panel B : information approached by AIM1

Coefficients	γ_o	γ_1	γ_2	γ_3	γ_4	R^2 (%)
Estimation	2.3	0.07	0.30			
t-value	8.21	0.32	5.31			1.47
p-value	0.00	0.74	0.00			
Estimation	3.40	0.15	0.26	-0.26		
t-value	7.89	0.63	4.80	-3.62		2.58
p-value	0.00	0.52	0.00	0.00		

Panel C : information approached by AIM2

Coefficients	γ_o	γ_1	γ_2	γ_3	γ_4	R^2 (%)
Estimation	2.43	0.07	0.38			
t-value	8.55	0.31	6.07			1.44
p-value	0.00	0.75	0.00			
Estimation	3.48	0.15	0.33	-0.25		
t-value	7.91	0.61	5.34	-3.53		2.56
p-value	0.00	0.54	0.00	0.00		

Panel D : information approached by AIM3

Coefficients	γ_o	γ_1	γ_2	γ_3	γ_4	R^2 (%)
Estimation	2.43	0.07	0.38			
t-value	7.76	0.30	4.82			1.47
p-value	0.00	0.76	0.00			
Estimation	3.43	0.14	0.32	-0.25		
t-value	7.79	0.59	4.11	-3.51		2.58
p-value	0.00	0.54	0.00	0.00		

The coefficient of β is non-significant in all regressions. As expected, the size has a negative and highly significant impact on excess returns. These results join those obtained previously for the CAPM specifications not considering the private information. The private information approached by RS has no significant impact on returns when considered alone (17) and exhibit a highly significant, negative coefficient, when considered with the size factor (18). The private information approached by AIM has a negative impact on returns whichever the model and private information specification. Our AIM measures are all highly significant and present the expected signs. In addition, AIM measures have a strong stability, contrary to the RS measure. AIM2 is the most significant variable, with a Student *t* of 6,07 when the factor size is not considered and of 5,34 when the size effect is controlled. AIM1 and AIM3 present equivalent explanatory significance. Their Student *t* exceeds *4*, whether the size effect is controlled or not. The superiority of AIM2 relative to AIM1 means that adding additional sources of information raises the accuracy for determining the amount of private information. The superiority of AIM2 relative to AIM1 means that industrial effects are more pervasive for extracting information than market effects.

To complete our analyses, we consider the SMB and HML factors loadings. The objective is to determine if our informational factor dominates or not these traditional factors. The regression suggested is as follows:

$$R_{it} = \gamma_{ot} + \gamma_{1t}\beta_{i,t-1} + \gamma_{2t}INFO_{i,t-1} + \gamma_{3t}\beta_{i,t-1}^{Size} + \gamma_{4t}\beta_{i,t-1}^{BM} + \varepsilon_{it} \tag{19}$$

where $\beta_{i,t-1}^{Size}$ and $\beta_{i,t-1}^{BM}$ are the SMB and HML factor loadings coefficients for the stock *i* during month *t-1*. The results are presented in the following table.

Tableau 6

The table presents average coefficients from the regression of stock excess returns on factors of risk. The methodology is that proposed by Fama and MacBeth (1973). The regressions cover 538 884 stocks/months for which data are available over the period between 01/01/1986 and 31/12/2001 on the NYSE, AMEX and NASDAQ American markets. The estimates are, in order, the averages of the coefficients estimated for the constant term, β, the informational factor, the SMB and HML factor loadings. The factor loadings are determined following the methodology suggested by Fama and French. The cross-sectional regressions are corrected for heteroscedasticity.

Mesure	Coefficients	γ_o	γ_1	γ_2	γ_3	γ_4	R^2 (%)
	Estimation	1.74	-0.20	-0.14	0.50	-0.16	
RS	*t*-value	5.43	-1.62	-1.90	1.60	-0.74	6.35
	p-value	0.00	0.10	0.05	0.11	0.45	
	Estimation	2.23	-0.11	0.29	0.47	-0.15	
AIM 1	*t*-value	9.58	-0.74	5.67	1.51	-0.70	5.77
	p-value	0.00	0.45	0.00	0.13	0.48	
	Estimation	2.34	-0.10	0.37	0.46	-0.15	
AIM 2	*t*-value	9.79	-0.73	6.44	1.47	-0.70	5.76
	p-value	0.00	0.46	0.00	0.14	0.48	
	Estimation	3.42	-0.10	0.39	0.45	-0.15	
AIM 3	*t*-value	9.68	-0.73	5.78	1.46	-0.69	5.77
	p-value	0.00	0.46	0.00	0.14	0.48	

The explanatory power of the information factor remains weak *(p=0.05)* if this one is approached by the specific risk, and presents the contrary sign relative to the theory. Our AIM measures remain highly significant and present the expected signs. As for the preceding regressions, variable AIM 2 remains most significant with a Student *t* equal to 6,44. The market, SMB and HML factors present the same signs and the same explanatory capacities as the preceding regressions. These factors are very largely dominated by our informational factor.

Our results are not different from those obtained by Easley et al. (2002). These authors propose a regression of stock excess returns on β, PIN, size measured by the logarithm of equity market value and the logarithm of the market-to-book ratio. The coefficient of the market factor loading β paradoxically presents a negative sign which becomes highly significant with the correction suggested by Litzenberger-Ramaswamy (1979). PIN presents a positive correlation with returns, with a Student *t* equal to 2,49 without the correction of Litzenberger-Ramaswamy (1979) and equal to 4,36 with this correction. The size effect is paradoxically positive whereas the effect of growth approached by the market-to-book ratio is non-significant. The authors affirm that their results support the hypothesis of private information.

Conclusion

Our paper proposes a new measure for calculating the degree of private information for stocks traded on the market, designated by AIM. Our measure derives directly from asset-pricing models with asymmetrically informed investors. Recent studies have demonstrated that common stocks contain private information and this information is not perfectly revealed by prices because the market is noisy. We continue this line of reasoning and affirm that, if a stock has private information, its price at the beginning of the target period should be correlated with the return on that period. The higher the correlation, the higher the information asymmetry affecting the stock. Thus, by projecting returns on prices and considering the explanatory power of the regression (the R^2) we can obtain a proxy for that information asymmetry. Because our measure is directly related to the existence of private information on the market, it circumvents the problems associated with existing information-asymmetry measures, mainly that they may proxy for other unobserved variables.

We have proposed several specifications for determining the degree of private information in common stocks. These specifications considered the private information on specific and common factors generating returns. The consideration of common factors as private information sources is new in the literature. Existing rational expectation models cannot solve the equilibrium prices when securities are affected by such factors (Admati, 1985). Jimenez (2004) has proposed a generalisation of the Grossman and Stiglitz (1980) model for multi-securities markets and with multiple factors generating returns and has found a closed-form solution for asset prices. The results show that individual prices are the most informative on specific factors. On the other side, portfolios that are sensitive to a certain factor are more informative on such factors than individual securities. These results justify our approach that determines the private information amount by projecting returns on individual prices and sector/industrial portfolios. In our view, these portfolios are sensitive to economic factors driving returns.

Our results show that AIM private information measures based on projection of returns on prices are relevant. These measures are correlated as expected with the size, volume of transactions and traditional factors. The impact of AIM on stock returns is positive and highly significant. This impact is robust across several specifications for measuring AIM and several CAPM specifications, as the traditional Sharpe and Lintner one or those including Fama-French factors. Traditional CAPM specifications seem not relevant for explaining stock returns, a result similar to those obtained by other recent studies. The informational CAPM specifications have a much stronger explanatory power. Including additional private information sources in the AIM specifications adds more explanatory power to the link between stock returns and private information. These results support our approach.

We have also compared our measure with a traditional one, based on the firm-specific return variation. The latter exhibits expected signs when correlated with size, volume of transactions our AIM measures. This seems to support the hypothesis that the specific risk measure is a proxy for private information. But this measure exhibits a high instability across different CAPM specifications and has a negative impact on stock returns in most cases. These results support the hypothesis, already put forward by Durnev et al. (2004), that the specific risk is rather a measure of price informativeness and not a measure of information asymmetry. We conclude that, as other existing private information measures in the literature, the specific-risk measures other effects, adding noise to the private information effect.

References

Admati, A., 1985, A Noisy Rational Expectations Equilibrium for Multi-Asset Securities Markets, *Econometrica* 53, 629 – 657

Biais, B., Bossaerts, P., and C. Spatt, 2002, Equilibrium Asset Pricing under Heterogeneous Information, *Working paper Toulouse University and CEPR*, 31p.

Blackwell, D. and L. Dubins, 1962, “Merging of opinions with increasing information,” Annals of Mathematical Statistics 38, 882-886.

Botosan, C., and M. Plumlee. “A re-examination of disclosure level and the expected cost of equity capital.” Journal of Accounting Research 40 (2002): 21-40.

Clarke, J., and K. Shastri, 2001, On Information Asymmetry Metrics, *Working Paper, Dupree College of Management, Georgia Institute of Technology* 01-14, 49 p.

D’Mello, R. and S. Ferris, 2000, “The Information Effects of Analyst Activity as the Announcement of New Equity Issues,” Financial Management 29, 78-95.

Durnev, Artyom, Randall Morck, and Bernard Yeung 2000, “Does Firm-specific Information in Stock Prices Guide Capital Allocation? ”, NBER WP. 8993.

Easley, D., and M. O’Hara, 2001, Information and The Cost of Capital, *Working Paper, Johnson Graduate School of Management, Cornell University*

Easley, D., Hvidkjaer, S., and M. O’Hara, 2002, Is Information Risk a Determinant of Asset Returns?, *The Journal of Finance* 57, 2185-2222

Easley, D., N. Kiefer, M. O’Hara, and J. Paperman, 1996, “Liquidity, information, and infrequently traded stocks,” Journal of Finance 51, 1405-1436.

Elton, E., M. J. Gruber, and M. N. Gultekin, 1984, “Professional expectations: Accuracy and diagnosis of errors,” Journal of Financial and Quantitative Analysis 19, 351- 363.

Gilson, S., P. Healy, C.F. Noe, and K.G. Palepu, 1997 “Information effects of spin-offs, equity carve-outs, and targeted stock offerings,” Working paper, Harvard Business School.

Grossman, S. J., and J. E. Stiglitz, 1980, On The Impossibility of Informationally Efficient Markets, *The American Economic Review* 70, 393 – 408

Hellwig, M. F., 1980, On The Aggregation of Information in Competitive Markets, *Journal of Economic Theory* 22, 477 – 498

Jimenez-Garcès, S., 2004, Information Transmission in Multi-Asset Securities markets, *Working Paper CERAG,* 15 p.

Krishnaswami, S. and V. Subramaniam, 1998, "Information asymmetry, valuation, and the corporate spin-off decision", *Journal of Financial Economics* 53, 73-112.

Madhavan, A., M. Richardson, and M. Roomans, 1997, "Why do security prices change? A transaction-level analysis of NYSE stocks," The Review of Financial Studies 10, 1035-1064.

McLaughlin, R., A. Safieddine, and G. Vasudevan, 1998, "The information content of corporate offerings of seasoned securities: An empirical analysis," Financial Management 27, 31-45.

Morck, Randall, Bernard Yeung, and Wayne Yu, 2000, "The Information Content of Stock Markets: Why Do Emerging Markets Have Synchronous Stock Price Movements?" Journal of Financial Economics 59, 215-260.

O'Hara, M., 2003, Presidential Address: Liquidity and Price Discovery", *Journal of Finance* 53, 1335-1354

Roll, R., 1977, A Critique of the Asset Pricing Theory's Tests; Part I: On Past and Potential Testability of Theory, *Journal of Financial Economics* 4, 129-76

Smith, C. and R. Watts, 1992, "The investment opportunity set and corporate financing, dividend, and compensation policies," Journal of Financial Economics 1992, 263- 292.

Van Ness, Bonnie F. Van Ness, Robert A. Warr, Richard S., How well do adverse selection components measure adverse selection?, *Financial Management*, September 22, 2001;

Wang, J., 1993, A Model of Intertemporal Asset Prices under Asymmetric Information, *Review of Economic Studies* 60, 249-282