Opaque Financial Reports, R-square, and Crash Risk

Amy P. Hutton*, Alan J. Marcus, and Hassan Tehranian

Boston College

First Draft: September 2007
This Draft: October 2008

Abstract We investigate the relation between the transparency of financial statements and the distribution of stock returns. Using earnings management as a measure of opacity, we find that opacity is associated with higher R-squares, indicating less revelation of firm-specific information. Moreover, opaque firms are more prone to stock price crashes, consistent with the prediction of the Jin-Myers (2006) model. However, these relations seem to have dissipated since the passage of the Sarbanes-Oxley Act, suggesting that earnings management has decreased or that firms can hide less information in the new regulatory environment.

JEL Descriptors G19, D89, M40

* Corresponding Author: ahutton@bc.edu; 617 552-1951 phone; 617 552-6345 fax;
Fulton Hall 520, 140 Commonwealth Ave. Chestnut Hill, MA 02467-3808

We thank workshop participants at Boston College, Harvard Business School, and the University of Wisconsin, Madison for helpful comments. We also thank Ying Duan and Shan He for excellent research assistance. We thank Qiang Cheng and David Farber for their restatement dataset.

1. Introduction

Financial economists and accountants have long viewed stock price changes as tied to new information about firms' prospects. However, Roll (1988) finds that only a relatively small portion of price movements can be explained by contemporaneous public news, and speculates that traders acting on non-public firm-specific information may drive returns. These results have stimulated considerable interest in the relation between information and stock price dynamics, and in particular, between the R-square from a modified index-model regression and the revelation of firm-specific news. For example, Morck, Yeung and Yu (2000) show that R-square is lower (equivalently, firm-specific return variation is higher) in countries with greater property rights accorded to arbitrageurs who trade on firm-specific information. Following up on these insights, several other papers, summarized below, have explored the connection between R-square and various measures of opacity. Their common point of departure is the notion that greater transparency and more complete revelation of firm-specific information should reduce R-square.

Opacity may, however, affect more than the second moment of stock return distributions. For example, Jin and Myers (2006) develop a model with incomplete transparency that predicts occasional (individual firm) stock price crashes as accumulated negative firm-specific information suddenly becomes publicly available. In fact, Jin and Myers find that several cross-country indicators of opacity predict crash risk. But, in comparison to the large R-square literature, the empirical relation between opacity and crashes has been largely ignored.

In this paper, we further consider the empirical link between opacity and the distribution of stock returns. Our approach differs from the existing literature in two respects. First, instead of employing market-wide measures of opacity in cross-country comparisons, we develop a measure of opacity for individual firms based on an indicator of earnings management, specifically the prior three years' moving sum of the absolute value of discretionary accruals. Second, we consider the empirical link between opacity and the crash risk of individual firms, a relation that is predicted by models such as that of Jin-Myers.

We demonstrate that our measure of opacity reliably predicts both R-square and crash risk. While crash risk is associated with earnings management, the incidence of large positive jumps is not; therefore, opacity does not predict fat-tailed distributions per se, but is associated specifically with crashes. Finally, we show that in the post-Sarbanes-Oxley (SOX) years, the relations between discretionary accruals and both R-square and crash risk essentially disappears. This pattern is consistent with the interpretation that opacity associated with earnings management has declined sharply in the post-SOX years.

These findings are significant for several reasons. First, they corroborate earlier research establishing a link between proxies for opacity and R-square, but using more direct measures of opacity of particular firms. Second, they shed light on the process by which information is revealed to the market place. While it is well known that stock prices are more prone to big downward moves than upward ones (French, Schwert, Stambaugh, 1987; Campbell and Hentschel, 1992; Bekaert and Wu, 2000), our findings indicate that this asymmetry is not due entirely to the exogenous stochastic process generating information but likely results as well from the way that firms manage the flow of information to capital markets. In particular, our findings suggest that part of the asymmetry is due to managers who are able to stockpile negative information, hiding it from investors' view until its accumulation reaches a tipping point sufficient to result in a stock price crash (see also Kothari, Shu, and Wysocki, 2007). Third, an understanding of the firm-specific characteristics that can predict extreme outcomes may be useful in portfolio and risk-management applications that focus on tail events, for example, value at risk. Similarly, option pricing depends on skewness and crash risk. So-called smirk curves have characterized the implied volatility of individual stock options as well as index options since the crash of October 1987, and are widely held to reflect risk of future crashes (see Dumas, Fleming and Whaley, 1998 or Bates, 2000). Again, understanding the factors that drive cross-sectional variation in such tail risk would be of obvious importance to market participants and allow for sharper option pricing.

In the next section, we review the literature on opacity and stock returns and develop our hypotheses. Section 3 provides an overview of our data, and Section 4 presents tests of our hypotheses. Section 5 summarizes and concludes.

2. Opacity and Stock Return Distributions

The link between opacity and R-square has been extensively discussed in the literature. When less firm-specific information is publicly available, there are fewer observable reasons for individual stock returns to depart from broad market indexes, and stock-market synchronicity increases. The link between opacity and crash risk has received comparatively much less attention.

Several mechanisms may engender crash risk or more generally, negative skewness in returns. For example, it is well known that trading among investors who have different opinions may reveal the private signals of others and move prices even in the absence of new fundamental information (e.g., Romer, 1993). In Hong and Stein (2003) this process, combined with short sale constraints, imparts an asymmetry in which market declines differentially reveal the private signals of relatively pessimistic investors. Such revelation may lead other investors to downgrade their assessments of a firm's prospects, thereby reinforcing the decline.

Other sources of negative skewness focus on volatility feedback effects (e.g., French, Schwert, and Stambaugh, 1987 or Campbell and Hentschel, 1992). For example, big price movements may cause investors to reassess market volatility and increase required risk premia. An increased risk premium will reduce equilibrium prices, which reinforces the impact of bad news but offsets the impact of good news, thus generating negative skewness.

Our empirical tests, however, are motivated by far simpler models of the information structure. They most directly follow the model of Jin and Myers (2006) and require neither disagreement among investors nor time variation in risk premia. Instead, the model envisions firm managers controlling at least a portion of the public access to fundamental information about the firm. Managers have incentives to "stockpile" bad news, but in some circumstances those incentives collapse, leading to a sudden release of accumulated negative information and a stock-price crash. The management of firm-specific information simultaneously increases both R-square and crash risk.

In Jin and Myers (2006), lack of full transparency concerning firm performance enables managers to capture a portion of cash flow, in the process absorbing (and making non-visible) part of the variation in firm-specific performance. This increases R-square.

Managers are willing to personally absorb losses due to temporary bad performance in order to protect their jobs. However, following a run of sufficiently bad news, they will be unwilling or unable to absorb any more losses—in other words, they have an abandonment option. If they abandon their positions, all of the hitherto-unobserved negative firm-specific shocks become public at once, resulting in a crash. This process can give rise to long left tails in the return distribution.

Whereas the precise nature of opacity in Jin and Myers is largely irrelevant, Kirschenheiter and Melumad (2002) focus on a particular source of imperfect information, earnings smoothing. In their model, higher reported earnings increase the inferred level of permanent earnings, and therefore, the value of the firm. This effect is greater when reported earnings are perceived as more precise, and therefore managers generally have an incentive to smooth earnings, under-reporting earnings in response to positive surprises, and over-reporting following negative surprises. But when news is particularly bad, the manager will under-report earnings to the greatest extent possible, partially to reduce the inferred precision of the bad news, and partially to enable shifting of discretionary income to future periods. This gives rise to occasional "big baths," or stock-price crashes. However, while the Kirschenheiter-Melumad model is consistent with stock price crashes, it does not predict a reduction in R-square. Jin and Myers note that for R-square to decrease, managers (or interpreting their model more generally, employees) must absorb some risk from shareholders. In the Kirschenheiter-Melumad framework, the timing of information releases is affected, but not the amount of risk borne by shareholders over extended periods.

Jin and Myers' cross-country empirical results indicate that measures of opacity in a given national market are associated with both higher average R-square and higher crash risk of the firms in that country. But virtually all other empirical work on opacity has focused entirely on R-square. Li et al. (2004) also focus on country-level data, finding that capital market openness and better legal systems are associated with lower R-square. Piotroski and Roulstone (2005) find that the presence of informed market participants such as analysts and institutional investors is associated with lower R-square. Durnev et al. (2003) show that stock returns better predict future earnings changes when R-square is lower (i.e., when there is greater implied transparency). Wei and Zhang

(2004) relate the secular decline in average R-square between 1976 to 2000 to changes in the level and volatility of firm-specific information, particularly return on equity. Ferreira and Laux (2007) show that firms with fewer anti-takeover provisions display higher trading activity, better information about future earnings in stock prices, and higher levels of idiosyncratic returns.

Admittedly, not all of the literature is consistent. Ashbaugh, Gassen and LaFond (2006) find that the Durnev et al. (2003) results do not generalize in an international setting. Some papers suggest that lower R-square may reflect greater non-information related noise in returns rather than more firm-specific information (e.g., West, 1988). Barberis, Shleifer and Wurgler (2005) find that the simple addition or deletion of a firm to the S&P 500 index can significantly change its R-square.

In contrast to most prior work, we employ a direct measure of opacity based on measures of earnings (specifically accruals) management. Considerable evidence indicates that accruals management obscures at least some information about firm fundamentals (see e.g., Sloan, 1996), and is thus a direct, firm-specific measure of opacity. In addition, aggressive earnings management is likely to proxy for management's general proclivity to hide information from the capital market, and thus captures less easily quantifiable or observable aspects of opacity. This is an advantage of our empirical framework compared to that of Jin and Myers. In their cross-country setting, opacity is measured using characteristics of the broad capital market in which the firm happens to be situated. In contrast, our observations are at the firm level, and allow us to construct proxies for the level of opacity *chosen* by each particular firm.

We take this choice as given and do not model cross-sectional differences in the underlying motivation of firms to manage their earnings, but note that this is a topic that has received considerable attention in the corporate governance literature. Dechow, Sloan and Sweeney (1996) examine firms subject to enforcement actions by the SEC and show that compared to a matched control sample these firms have weaker corporate governance. For example, they are less likely to have audit committees, less likely to have outside blockholders, more likely to have insiders holding a majority of the board seats, and more likely to have a CEO who is also the Chairman of the Board. These authors also attempt to understand *why* firms manipulate earnings and conclude that "important motivations for earnings manipulation are the desire to raise external financing at low cost and to avoid debt covenant restrictions."

More recent research confirms many of these findings. For instance, Klein (2002) documents a negative relation between audit committee / board independence and earnings management. Cornett, Marcus, and Tehranian (2008) find that large stock ownership by institutional investors also constrains earnings management. Conversely, they show that economic rewards to a high stock price (in the form of incentive-based compensation such as option grants) encourage earnings management. Other studies also find that earnings management responds to payoffs associated with a temporarily high stock price (Bergstresser, Desai, and Rauh, 2006; Bergstresser and Philippon, 2006; Teoh, Welch, and Wong, 1998a, 1998b). More generally, firms seem to respond to incentives to manipulate reported earnings to retain healthy market valuations. Specifically, Skinner and Sloan (2002) and Kinney, Burgstahler, and Martin (2002) document dramatic stock price declines associated with even small negative earnings surprises; Matsumoto (2002) and Burgstahler and Eames (2006) provide evidence of earnings management to meet analysts' forecasts.[1]

We hypothesize that when a firm's financial reports are more opaque, there is less firm-specific information available to affect its stock returns. Thus, we expect such firms to have stock returns with less idiosyncratic volatility relative to market-wide variation. In other words:

> *H*1*: Firms with opaque financial reports have stock returns that are more synchronous with the market.*

[1] While corporate governance and incentive-based compensation may be related to the propensity to manage earnings, we do not see an obvious concern with endogeneity. There is no obvious reason for firms with naturally higher or lower R-square due to their lines of business to be more or less inclined to manage earnings. Therefore, we are not overly concerned about endogeneity or "reverse feedback" effects. In fact, while it is a stretch, one might argue that firms with high R-square might be, if anything, *less* prone to manage earnings, which would cut against our hypothesis. This argument might follow from Collin and Kothari (1987), who find that high-beta firms (which tend to be high R-square firms) exhibit lower price responses to earnings announcements, possibly because those earnings are capitalized at higher discount rates. The lower price response would reduce the incentive to manage earnings. In any event, as we discuss below, our results stand up even when we measure earnings management relative to other firms in the same industry, for which all firms ought to have similar stock market synchronicity. This robustness also argues against a problem with endogeneity.

However, these firms also face a higher likelihood of extreme outcomes when the firm-specific information is finally revealed to the market. Specifically, we hypothesize that firms managing reported earnings shelter *bad* information up to a point, but that once a threshold is crossed, the information comes out in one fell swoop, at which point we observe a price crash.

H2: Firms with opaque financial reports face greater crash risk.

In contrast, managers do not face incentives to shelter *good* information. So, we do not expect opaque financial reporting to be related to the likelihood of positive jumps.

H3: Firms with opaque financial reports face no greater likelihood of positive discrete jumps in stock prices than more transparent firms.

Finally, if Sarbanes-Oxley has been effective in lowering the level of earnings management via accrual manipulation, then our measure of opacity ought to be less highly related to R-square and crash risk in the post-SOX era.

3. Sample Development, Variable Measurement, and Research Design

3.1 The sample

We combine firms' weekly stock return data from CRSP with annual financial data from Compustat. Weekly stock returns are assigned to each firm's fiscal year so as to match the time period of its reported financial data. Our sample period begins with fiscal year 1991, the first fiscal year for which we are able to estimate three annual lags of discretionary accruals using the Statement of Cash Flow method; our sample period ends with the last year for which we have complete CRSP and Compustat data, fiscal year 2005. We begin with all firm-years on Compustat between 1991 and 2005 (109,026 firm-years). We exclude: low-priced stocks (average price for the year less than $2.50); firm fiscal years with less than 26 weeks of stock-return data (22,777 firm-years are excluded); financial services firms and utilities (28,051 firm-years); and firm-years with insufficient financial data to calculate both three lags of discretionary accruals (9,903 firm-years) as well as control variables (7,413 firm-years). We are left with a final sample of 40,882 firm-years. The sample includes 43 of the 49 Fama/French industry definitions; the sample firm-years are approximately evenly distributed across our sample period. See Table 1 for details.

3.2 Accrual manipulation and opacity of financial reports

Under an accrual accounting system the primary measure of firm performance is Earnings or Net Income, where Earnings is an estimate of all the current and expected future net cash flows associated with economic transactions transpiring during a given reporting period. Underlying the Earnings estimate are accruals, such as accounts receivable and depreciation, which are used to allocate revenues to the period in which they are earned and to match expenses to the revenues they generate. In contrast to Earnings, cash flows do not represent the creation of value but rather the distribution of value and thus they are not the primary measure of firm performance. Accruals must be added to actual cash flows to obtain the primary estimate of firm performance, Earnings.

This implies that reported earnings are necessarily *estimates* of firm performance, the accuracy of which depends on the quality of the accruals used to estimate the expected future net cash flows associated with past economic transactions. Earnings and underlying accruals are management's estimates of expected future net cash flows, as management has primary responsibility for developing the estimate of earnings reported to shareholders. Auditors attest that management's estimate have been developed within the confines of generally accepted accounting principles (GAAP). However, there is latitude within GAAP, and in some cases (e.g., WorldCom, Enron, etc.), management teams have ignored even those boundaries.

Thus, while *over the life* of a firm, accruals must sum to zero and earnings must equal net cash flow, *during particular periods* earnings may deviate substantially from net cash flow and total accruals may reach significant magnitudes. In some instances, disparities between earnings and cash flows are reasonable, for example, resulting from growth in credit sales. These unbiased accruals are generally followed by cash flow realizations that remove or reverse the initial accrual (e.g., account receivables are decreased when cash is received from customers). In other instances, disparities between earnings and net cash flow result from earnings manipulation, and the inaccurate accruals are eventually reversed out by oppositely-signed accruals rather than by cash flow realizations.

To distinguish between normal and discretionary accruals, we employ the modified Jones model (Dechow, et al., 1995). Specifically, we estimate the following cross-sectional regression equation using the firms in each Fama/French industry for each fiscal year between 1988 and 2005:

$$\frac{TA_{jt}}{Assets_{jt-1}} = \alpha_0 \frac{1}{Assets_{jt-1}} + \beta_1 \frac{\Delta Sales_{jt}}{Assets_{jt-1}} + \beta_2 \frac{PPE_{jt}}{Assets_{jt-1}} + \varepsilon_{jt} \tag{1}$$

where TA_{jt} denotes total accruals for firm j during year t, $Assets_{jt}$ denotes total assets for firm j at the end of year t, $\Delta Sales_{jt}$ denotes change in sales for firm j in year t, and PPE_{jt} denotes property, plant, equipment for firm j at the end of year t.[2]

Discretionary annual accruals as a fraction of lagged assets for firm j during year t ($DiscAcc_{jt}$) are then calculated using the parameter estimates from Eq. (1):

$$DiscAcc_{jt} = \frac{TA_{jt}}{Assets_{jt-1}} - \left(\hat{\alpha}_0 \frac{1}{Assets_{jt-1}} + \hat{\beta}_1 \frac{\Delta Sales_{jt} - \Delta Receivables_{jt}}{Assets_{jt-1}} + \hat{\beta}_2 \frac{PPE_{jt}}{Assets_{jt-1}}\right) \tag{2}$$

where hats denote estimated values from regression Eq. (1). The inclusion of $\Delta Receivables_{jt}$ in Eq. (2) is the standard "modification" of the Jones model. This variable attempts to capture the extent to which a change in sales is due to aggressive recognition of questionable sales.

Dechow, Sloan, and Sweeney (1996) document the pattern of discretionary accruals for known manipulators, specifically firms subject to enforcement actions by the SEC (see their Figure 1 on page 18). Discretionary accruals gradually increase as the alleged year of earning manipulation approaches, and then exhibit a sharp decline. The increase in discretionary accruals is consistent with earnings manipulation; the decline with the reversal of prior accrual overstatements. Interestingly, the large, positive discretionary accruals recorded by these firms are followed by large, negative

[2] Total annual accruals equal income before extraordinary items minus cash flow from operating activities adjusted for extraordinary items and discontinued operations, specifically TA_{jt} = Compustat annual data items #123 – [Compustat annual data item #308 – Compustat annual data item #124].

discretionary accruals rather than by the positive cash flow realizations that would have tended to follow unbiased accrual practice. Dechow, Sloan and Sweeney also document that these firms generally manipulate reported earnings from one to three years before being detected (see their Table 3), and that the overstated accruals of these firms typically reverse fairly quickly, with negative discretionary accruals following the prior positive ones in the years immediately following the periods of earnings manipulation.

Taken together, these facts motivate a simple, but intuitively appealing, measure of opacity in financial reports, the three-year moving sum of the absolute value of annual discretionary accruals:

$$OPAQUE = \text{AbsV}(DiscAcc_{t-1}) + \text{AbsV}(DiscAcc_{t-2}) + \text{AbsV}(DiscAcc_{t-3}) \tag{3}$$

The straightforward idea behind this measure is that firms with consistently large absolute values of discretionary accruals are more likely to be managing reported earnings, thus revealing less firm-specific information to investors. For example, the characteristic pattern of manipulation documented in Dechow, Sloan and Sweeney—large positive followed by large negative abnormal accruals—would result in a high measure of earnings management, as both the positive and negative abnormal values would contribute to the moving sum of *absolute* discretionary accruals. We use a three-year moving sum (rather than a one-year value) to capture the multi-year effects of earnings management and because the moving sum is more likely to reflect an underlying policy of the firm to manage earnings.[3]

Figures 1a and 1b plot the cross-sectional means of AbsV(*DiscAcc*) and *OPAQUE* over the sample period. We also include the mean weekly return on the CRSP value-weighted market index in each year to examine the relation between levels of earnings management and market returns.

Figure 1a shows that the mean of AbsV(*DiscAcc*) is usually between six and eight percent of lagged total assets. However, in four years: 1996, 2000, 2001 and 2002, the

[3] We repeated our analysis substituting latest-year absolute discretionary accruals for the moving sum (see Table 6, Panel B below). While results were qualitatively similar, the moving sum is in fact a more powerful predictor of both R-square as well as crash propensity.

mean level of AbsV(*DisAcc*) exceeds this typical range, and in three out of four of these high-manipulation years the mean weekly return on the market index is negative (aggregating to annual returns of −3.6%, −10.5% and −19.3% in 2000, 2001 and 2002, respectively). Thus, Figure 1a suggests that the level of earnings management may be related to overall market performance. Figure 1b demonstrates that our primary measure of earnings management, *OPAQUE*, is less highly correlated with overall market performance. Nevertheless, we control for the level of the overall market performance in our analyses below.[4]

3.3 Measuring idiosyncratic risk

R-squares and residual returns are calculated from an expanded index model regression:

$$r_{j,t} = \alpha_j + \beta_{1,j}r_{m,t\text{-}1} + \beta_{2,j}r_{i,t\text{-}1} + \beta_{3,j}r_{m,t} + \beta_{4,j}r_{i,t} + \beta_{5,j}r_{m,t+1} + \beta_{6,j}r_{i,t+1} + \varepsilon_{j,t} \qquad (4)$$

where $r_{j,t}$ is the return on stock j in week t, $r_{m,t}$ is the CRSP value-weighted market index, and $r_{i,t}$ is the Fama/French value-weighted industry index. We allow for non-synchronous trading by including lead and lag terms for the market and industry indexes (Dimson, 1979).[5]

The residuals from Eq. (4) are highly skewed. We transform them to a roughly symmetric distribution by defining the *Firm-Specific Weekly Return* as the log of one plus the residual return from Eq. (4). This transformation allows us to define crash and positive jumps symmetrically, as residual returns corresponding to a threshold number of standard deviations either above or below the mean.[6]

Notice that we define crash and jump events using firm-specific returns, i.e., using residuals from Eq. (4). Using actual returns would result in an abundance of "crashes"

---

[4] Additionally, to ensure that our findings are not sensitive to prevailing market conditions, we re-run all regression analyses using a Fama-MacBeth approach with Pontiff (1996) adjusted standard errors. The findings are qualitatively unchanged.

[5] Our findings below are qualitatively unchanged if we include either zero leads and lags or two leads and two lags in the expanded index model. Finally, using bi-weekly returns instead of weekly returns also leaves our findings qualitatively unchanged.

[6] This transformation does in fact seem to make the distribution essentially symmetric, even in the tails, which is our greatest concern. For example, the 99th percentile of the distribution is .186 while the 1st percentile is –.184.

during broad market declines and "jumps" during advances. The firm-specific crashes and jumps of interest to us require that we define extreme events using residual returns.

$1 - R^2$ is a natural measure of firm-specific volatility or (lack of) market synchronicity, but it is bounded between 0 and 1, which creates complications for empirical estimation. We follow common practice (e.g., Morck et al., 2000) in defining idiosyncratic risk using a logistic transformation of R^2, which can range from negative to positive infinity.

$$IDIOSYN = \ln\left(\frac{1 - R^2}{R^2}\right) \tag{5}$$

Thus, a high value for *IDIOSYN* indicates a high level of idiosyncratic risk.

As discussed in Section 2, we expect firms with more opaque financial reports to have stock returns that are more synchronous with the market. Opacity implies that less firm-specific information is available to affect the firm's stock returns, and thus the firm will exhibit lower values of *IDIOSYN*. However, while opaque firms may generally tend to follow the market more closely, they also present a higher likelihood of an extreme outcome or crash when accumulated negative firm-specific information is finally revealed to investors.

3.4 Measuring the frequency of crashes and jumps

We construct measures of the likelihood of crashes or positive jumps based on the number of the *Firm-Specific Weekly Returns* exceeding 3.09 standard deviations below or above its mean value, respectively, with 3.09 chosen to generate a frequency of 0.1% in the normal distribution.[7] An indicator variable, *CRASH*, is set equal to one for a firm-year if the firm experiences one or more *Firm-Specific Weekly Returns* falling 3.09 standard deviations below the mean weekly firm-specific return for that fiscal year; otherwise *CRASH* is set equal to zero. Similarly, *JUMP* is an indicator variable defined as one if a firm experiences one or more *Firm-Specific Weekly Return* 3.09 standard deviations above the mean value for that fiscal year, and zero otherwise.

[7] While the mean residual from the index model regression, Eq. (4), is zero by construction, the mean value of ln(1 + residual return) is not. Therefore, we calculate deviations from the firm-year mean of ln(1 + residual return) to define crashes and positive jumps.

This definition of a crash implies a substantial weekly return. For example, the average standard deviation of ln(1 + residual return) across our sample is 5.8 percent. Using this value, a negative 3.09-sigma event translates into an abnormal weekly return of −18 percent.

Given our definition of a crash, if firm-specific returns were actually normally distributed, one would expect to observe 0.1% of the sample firms crashing in any week. A similar number of jumps would be expected. The probability of a crash or a jump over the course of a year would then be $1 - (1 - .001)^{52} = .0507$. In fact, we observe considerably greater frequency of both crashes and jumps than this benchmark.[8] Table 2, Panel A, indicates that 17.1% of the firm-years in our sample experience at least one *CRASH* (6,950 firm-years); 22.0% experience at least one *JUMP* (9,029 firm-years).

In Table 2, Panel B, we highlight the mean, median, and variance of raw weekly returns for three sub-samples of firm-weeks: *CRASH* weeks, *JUMP* weeks, and *ALL OTHER* weeks. The first set of columns presents statistics for individual firms. The mean weekly return for *CRASH* weeks is −22.74 percent, for *JUMP* weeks 33.27 percent, and for *ALL OTHER* weeks 0.26 percent. Median returns demonstrate a similar pattern. As expected, the variance of weekly returns is dramatically higher for *CRASH* and *JUMP* weeks than for *ALL OTHER* weeks. The middle set of columns presents analogous results for the market index. In this panel, *CRASH* (or *JUMP*) weeks refer to any week in which any firm in the sample crashes (or jumps). Even though *CRASH* and *JUMP* weeks are categorized using residuals from an expanded index model, the mean and median weekly return to the market index is generally a bit higher in *JUMP* weeks than in *ALL OTHER* weeks in the sample. This is not the case for *CRASH* weeks. Nevertheless, the variance of the market index is not higher in *CRASH* and *JUMP* weeks. Finally, the last set of columns report statistics averaging across industries. If any firm in an industry crashes (or jumps) in a given week, that is defined as a crash (or a jump) week for the industry. Industry results indicate slightly lower (higher) mean and median returns in *CRASH* (*JUMP*) weeks, but equivalent variance.

[8] Given our research question, we obviously do not expect the *Firm-Specific Weekly Returns* to be normally distributed. We merely use the 0.1% cutoff of the normal distribution as a convenient way to obtain reasonable benchmarks for extreme events, *CRASH*es and *JUMP*s.

Figures 2a and 2b highlight the frequency of *CRASH*es and *JUMP*s across the sample years. We include the value-weighted market index in Figures 2a and 2b to demonstrate that the frequency of these extreme outcomes is in fact independent of market movements, which is not surprising given that we employ residuals from an index model to define *CRASH*es and *JUMP*s. The frequency of *CRASH*es is significantly higher in the second half of the sample period than in the first half, whereas the frequency of *JUMP*s is largely unchanged over the time period.

3.5 Control variables

Following prior literature, our control variables include: *SIZE,* defined as the log of the market value of equity at the beginning of the fiscal year; *M to B,* defined as the ratio of the market value of equity to the book value of equity measured at the beginning of the fiscal year; *LEV,* defined as the book value of all liabilities scaled by total assets, again measured at the beginning of the fiscal year; and finally, contemporaneous *ROE*, defined as income before extraordinary items divided by the book value of equity.

We include the variance of the Fama/French weekly industry index (*VAR_INDRET*) as an additional control variable in regressions with *IDIOSYN* as the dependent variable because higher industry variance will increase systematic risk, and hence the R-square of Eq. (4). Finally, following Jin and Myers (2006), we also include the *KURTOSIS* and the *SKEWNESS* of the *Firm-Specific Weekly Returns* as control variables in some of our *IDIOSYN* regressions.

3.6 Descriptive statistics and correlation matrix

Table 3, Panel A, presents descriptive statistics for our key variables of interest; Table 3, Panel B, presents the correlation matrix. Pearson correlations appear above the diagonal, and Spearman correlations below. *OPAQUE* is negatively correlated with R-square, [9] and positively correlated with tail events, as measured by skewness, kurtosis,

---

[9] The negative correlation between *R*-square and *OPAQUE* appears to contradict one of our main hypotheses. However, after controlling for other characteristics, in particular size, the correlation is strongly negative. For example, we show below (Table 5) that once firms are stratified by size, *R*-square increases with *OPAQUE* for *each* size group. The multiple regressions in Table 6 also confirm the positive relation between *R*-square and opacity, after controlling for firm size. The univariate correlation (that does not control for size) appears to be negative because large firms tend to have higher *R*-squares and to be less opaque, as documented in Table 5, Panels A and C.

crash probability, or jump probability. Not surprisingly, the various measures of fat tails, i.e., kurtosis, crash risk, and jump probability, are all positively correlated. Size is negatively correlated with opacity and jump probability and positively correlated with R-square and crash probability. As expected, R-square is negatively correlated with kurtosis, skewness, crash risk, and jump probability in this univariate analysis.

3.7 Opacity and Information Hoarding

Before turning to our main results, which connect our measure of opacity to R-square and crash risk, we pause to first corroborate our interpretation of our opacity measure. We wish to show that firms with high levels of *OPAQUE* are in fact less transparent and in particular, are more likely to withhold negative news from investors.[10] To do so, we examine the relation between *OPAQUE* and the frequency of financial statement restatements. We obtain a sample of such restatements from Cheng and Farber (2008), who in turn developed their sample from 919 restatements of public companies between January 1, 1997 and June 30, 2002 identified in a U.S. General Accounting Office (GAO) Report.[11] The GAO study focuses only on restatements connected to "accounting irregularities," i.e., "an instance in which a company restates its financial statements because they were not fairly presented in accordance with generally accepted accounting principles (GAAP). This would include material errors and fraud." Thus, these are fairly extreme and relatively uncommon events that indicate less-than-forthright accounting practice.

We are able to match 233 of the 919 restatements with our subsample of 14,670 firm-years with fiscal years 1997 - 2001, and find that the incidence of financial restatements in this subsample is in fact related to our opacity measure. We divide the subsample into three groups based on firm-year opacity level. Table 4, Panel A, shows that the rate of restatements increases with opacity. Only 1.2% of the firm-years in the low-opacity group experience a restatement compared to nearly 2% of the firm-years falling in the high-opacity group. This difference is statistically significant at a 0.0038 level. Moreover, differences in restatement rates become more extreme when we cut the data more

[10] Discretionary accruals have been widely used elsewhere as a proxy for earnings management. For a review, see Dechow and Schrand (2004), Chapter 4.

[11] U.S. General Accounting Office, *Financial Statement Restatements: Trends, Market Impacts, Regulatory Responses and Remaining Challenges*, October 2002, www.gao.gov/new.items/d03138.pdf.

finely. For example, if we split our sample into opacity deciles, the incidence of restatements in the highest decile is triple that in the lowest, 3% versus 1%.[12]

A more dramatic view of the association between opacity and restatements comes from Panel B of Table 4, which compares mean and median values of *OPAQUE* in the group of firm-years with restatements to those in the no-restatement group. The restatement group demonstrates significantly higher levels of opacity. The difference in mean values of *OPAQUE* is .0989 (= 0.3634 – 0.2645). Since this is the mean value of a 3-year moving sum, it corresponds to a difference of .033 in the mean annual absolute discretionary accrual. Recall that discretionary accruals are expressed as a fraction of firm assets; therefore, this value corresponds to a swing in measured annual ROA of fully 3.3%.

We conclude that restatement experience corroborates our interpretation of our opacity measure. Firms with higher measures of *OPAQUE* more frequently are required to amend the information they have provided to investors.

4. Regression analysis

4.1 R-square

We begin with results for *IDIOSYN*, which is our measure of the firm-specific information arriving to the securities market. Although several earlier papers have explored the relationship between opacity and R-square, our measure of opacity differs from previous work as it is based on characteristics of the individual firm, specifically on its earnings management. We wish to confirm that our measure of opacity is related to the level of firm-specific information reflected in stock returns.

In Table 5, we find break points that sort our sample into five size quintiles and into three opacity groups (1 = low; 3 = high). With a sample of about 40,800 firm-years,

---

[12] One potential issue in this analysis is matching a restatement event to the time period in which the accounting irregularities actually occurred. Existing research demonstrates that on average firms can manipulate accruals for three years before being discovered and having to restate their financial statements (see e.g., Dechow, Sloan and Sweeney, 1996). Examining the 16 detailed case studies contained in the GAO report confirms this assessment. In 13 out of the 16 cases, (at least) the three prior years' financial statements had to be restated. Because *OPAQUE* is a 3-year moving sum of the absolute value of discretionary accruals, it measures firm behavior in an appropriate window prior to any restatement. To the extent that a 3-year window is not appropriate for some firms, that would only tend to weaken the association we find between *OPAQUE* and restatements.

each size quintile contains a bit more than 8,100 firm-years while each opacity group contains about 13,600 firm-years. Panel A reports the number of firms falling into each size-opacity cell. The average value for market capitalization within each size quintile is presented on the left margin. Consistent with the negative correlation between size and opacity reported in Table 3, we observe that for the small-firm quintiles (1 and 2), the number of observations increases as we move across rows, indicating that small firms are more likely to be characterized by high opacity. Symmetrically, for the larger-firm quintiles (4 and 5), the number of firms decreases across each row.

Panel B presents the average value of opacity for each cell in the 5×3 grid. Moving across rows, it is apparent that the sample exhibits considerable variability in opacity even within each size bracket.

Panel C shows how R-square varies with opacity within various size groups. We know from earlier work (e.g., Roll, 1988) that R-square varies with size, so it seems crucial to condition on at least this variable. Because size and opacity are also correlated, we wish to confirm that any association between opacity and R-square is not spurious, i.e., that opacity *independently* affects R-square after controlling for firm size. Comparing results across each row, we observe that higher opacity is in fact associated with higher R-square. The difference in R-square is small, but small changes in R-square across differing information regimes are consistent with prior research.[13] The entries on the right margin show significance levels for the test of the hypothesis that the R-squares in the low and high opacity groups are identical. For each size quintile the hypothesis is easily rejected. Thus, opacity does appear to have an *independent* impact on R-square.

Table 6 presents regression analysis for *IDIOSYN* that allows us to control for a range of other variables in addition to firm size. The variance of the industry return is mechanically related to R-square: higher industry volatility increases "explained" risk and therefore R-square, which reduces *IDIOSYN*.[14] We also include size, the market-to-book

[13] In Roll's (1988) study, for example, the difference between average R-square across days in which a firm is or is not mentioned in the financial press is also less than .02.

[14] Higher market volatility also should increase R-square, but while this factor varies across years, it is identical for all stocks in any year, whereas industry variability exhibits considerable cross-sectional dispersion. Moreover, the set of year-by-year industry volatilities should span much of the time series variation in market index volatility. Nevertheless, we did estimate the Table 6 regressions including market

ratio, leverage, and contemporaneous ROE as additional controls for firm characteristics.[15] Larger firms operating in a wider cross-section of the economy are expected to have higher R-squares. The market-to-book ratio places firms along a growth-versus-value spectrum and thus may be systematically related to R-square. Leverage is also expected to affect R-square through its impact on the sensitivity of firm returns to macroeconomic conditions and because it affects the division of risk between equity and debt holders.

Finally, Jin and Myers include skewness and kurtosis as control variables in their examination of R-square and opacity. Following their lead, we add these controls in Model 2 of Table 6. Neither of these variations has a meaningful impact on the relation between opacity and R-square: The coefficients on the moving sum of absolute discretionary accruals, *OPAQUE*, are −.163 and −.167 in Models 1 and 2 respectively. In both cases, the estimate is highly significant, with *t*-statistics of 8.95 and 9.34.

Consistent with Table 5, firm size is a highly significant predictor of *IDIOSYN*. Roll (1988) also finds that R-square increases with size, especially in simple index-model regressions (see his Figure 2). Market-to-book ratio has a statistically significant but economically trivial impact on *IDIOSYN*, and leverage is not at all significant. Not surprisingly, industry variance is highly significant, with high-volatility industries exhibiting lower values of *IDIOSYN* (higher R-square). Both kurtosis and skewness turn out to be statistically significant, with higher values of each implying higher values of *IDIOSYN*. This seems reasonable, since jump events would tend to weaken the link between firm returns and market returns.

In Model 3 of Table 6, we examine potential non-linearities in the relation between *OPAQUE* and *IDIOSYN*. These could arise from two sources. First, it is possible that earnings management has an impact on information flow only when it is relatively aggressive, possibly at the upper end of the observed range of *OPAQUE*. If so, the relation between *OPAQUE* and *IDIOSYN* would be concave, with a steeper reduction

---

volatility as a right-hand side variable. Inclusion of this variable had negligible effects on any of our results.

[15] In the accounting literature, studies of the impact of discretionary accruals on stock returns typically include controls for contemporaneous firm performance such as ROE. On the other hand, one might prefer to examine stock return characteristics during a given period without allowing for the influence of contemporaneous announcements. We therefore re-estimate all of our regressions without ROE as a control variable; as it turns out, the inclusion of ROE has almost no impact on any of the other regression estimates. Panel B of Table 6 contains these and other robustness checks.

in *IDIOSYN* at higher values of *OPAQUE*. Alternatively, we might observe thresholds beyond which opacity is already so high that additional earnings management doesn't substantially degrade already-impaired information flow; if so, we would find that the reduction in *IDIOSYN* as *OPAQUE* increases tails off, resulting in a convex relation between *OPAQUE* and *IDIOSYN*.

In fact, the empirical relation appears to be convex. In Model 3 of Table 6, we add the square of *OPAQUE* to the list of explanatory variables. Adding this quadratic term causes the coefficient on *OPAQUE* to more than double to −.402, but the coefficient on $OPAQUE^2$ is positive and highly statistically significant (*t*-statistic = 5.35).[16]

Figure 3 plots implied values of R-square using the parameter estimates from Model 3 as *OPAQUE* moves from the bottom to the top of its sample distribution, holding all other variables at their mean values.[17] R-square increases by around 13 percent, from .206 to .233, from the bottom to the top of the distribution of *OPAQUE*. Given the lack of response in R-square to explicit news about the firm documented in Roll (1988), a difference of this magnitude may be all that we can expect to uncover with respect to underlying earnings management. Put slightly differently, the quantitative impact of *OPAQUE* on R-square is as strong as the appearance of explicit firm-specific news in the financial press.

The last column of Table 6 presents estimates of "economic impacts" of the other right-hand side variables using coefficient estimates from Model 3, the most inclusive of the specifications. Each entry in that column is the expected impact on R-square resulting from an increase in the right-hand side variable from the 25th to the 75th percentile of the sample distribution. (The regression actually estimates the impact on *IDIOSYN*; we obtain the impact on R-square by inverting Eq. 5). Not surprisingly given the results in Roll (1988), size clearly dominates, but that is in part due to its tremendous

[16] We experimented with higher order terms for other independent variables, specifically skewness and kurtosis. As size already was a concave (log) transformation of market value, we did not add a non-linear term for that variable. The inclusion or exclusion of these additional quadratic terms affected neither the coefficients nor the significance levels of *OPAQUE* or $OPAQUE^2$ in any meaningful manner. Similarly, we experimented with additional squared terms in the Table 7 regressions below. Again, our inferences on crash risk were unaffected.

[17] We invert Eq. (5) to solve for $R^2 = 1/(\exp(IDIOSYN) + 1)$. While *IDIOSYN* is decreasing and convex in *OPAQUE*, R^2 is increasing and concave. The effect of *OPAQUE* on synchronicity tails off as *OPAQUE* increases.

sample skewness and correspondingly large interquartile range. Also not surprising, the variance of the industry return has a strong positive impact on R-square, and kurtosis has a large negative impact. Of most interest to us, the combined economic impact of *OPAQUE* and $OPAQUE^2$ is .015 − .003 = .012 or 1.2 percent. Recall that in Roll's (1988) study, the difference between average R-square across days in which a firm is or is not mentioned in the financial press is only about two percent; in this light, the impact of *OPAQUE* is not trivial.

4.2 Robustness

Our results are robust to reasonable changes in methodology. Panel B of Table 6 shows the coefficient estimates on *OPAQUE* and $OPAQUE^2$ for a variety of specifications. Line 1 of the Panel repeats the estimates from Model 3 of Table 6. Line 2 shows that the coefficients on *OPAQUE* and $OPAQUE^2$ are barely affected by excluding ROE from the right-hand side. We also experimented with our sample-selection criterion, and found in Line 3 that requiring 51 weeks of returns rather than 26 made virtually no difference in our estimates, presumably because the more stringent criterion actually results in the loss of very few firms from the sample. We tried measuring opacity using single-year absolute values of discretionary accruals rather than a three-year moving sum. The regression coefficients (Line 4) were qualitatively similar, but with somewhat lower significance levels, consistent with the interpretation that this is a noisier measure of earnings manipulation. We re-ran our index model regressions using 2-week rather than one-week holding periods and found that this made little difference to our results (Line 5).

We also focused on the model of discretionary accruals. It is possible that the Jones model fits accruals behavior better for some industries that others. If so, what appears to be high discretionary accruals and therefore earnings management for some firms might simply reflect higher model error for that industry. Therefore, for each industry we rank ordered firms by *OPAQUE*, and assigned each firm its percentile score of *OPAQUE* within its industry. This ordinal ranking automatically adjusts for systematic differences in absolute abnormal accruals across industries. As it turns out, regression coefficients on the ordinal ranking of *OPAQUE* are also negative (Line 6), with still-high levels of statistical significance.[18]

A potential problem with the panel regressions in Table 6 is the possibility of within-firm autocorrelation, which would result in biased standard errors. Therefore, we also estimated each regression equation as a cross-sectional regression for each year of the sample independently and computed autocorrelation-corrected Fama-MacBeth (1973) estimates using a method suggested by Pontiff (1996). Specifically, the time-series of the parameter estimates are regressed on a constant, with regression residuals modeled as a fifth-order autoregressive process. The standard error of the intercept term is used as the corrected standard error. As long as the fifth-order autoregressive process captures all of the serial dependence, these standard errors are not biased by serial or cross-sectional correlation.[19] The by-year Fama-MacBeth results (Line 7) are generally consistent with the results in Panel A. However, both the magnitude and the significance levels of the coefficients are considerably reduced. Our analysis below (Table 10) suggests that there is a structural break in the model after the passage of Sarbanes Oxley in 2002 and its likely impact on accounting practice and transparency. Therefore, in Line 8 we repeat the Fama-MacBeth analysis, but ending the sample in 2002. These results are far more in line with the other variants in the table. There are not enough post Sarbanes-Oxley years to produce a meaningful Fama-MacBeth analysis, but those years contribute considerably to variation in the coefficient estimates in the full sample and therefore reduce the *t*-statistics that appear in Line 7.

4.3 Crash risk

The coefficients on *OPAQUE* in Table 6 suggest that earnings management does in fact impede the flow of firm-specific information to the securities market. We turn next to the question of whether such opacity also predicts crash risk.

Table 7 presents logit analysis supporting the link between opacity and crash risk. Each firm-year is assigned a zero if the firm experiences no crash during the fiscal year,

[18] Given that we now use an ordinal rank, we exclude the square of *OPAQUE* from the right-hand side: squaring the ordinal rank of opacity would be a meaningless transformation.

[19] An analysis of the significance of the autoregressive parameters indicates that a fifth-order process is well specified. Detailed results for the by-year regressions are available upon request.

and a one if there is at least one week during the fiscal year in which the stock price crashes. In light of the strong evidence of non-linearity in the Table 6 regressions, we include both *OPAQUE* and *OPAQUE*^{2} on the right-hand side. The "marginal impact" of the explanatory variables is presented in the last column of Table 7. Analogous to Table 6, it is the increase in the probability of a crash in any year corresponding to a shift from the 25th to the 75th percentile of the distribution of each right-hand side variable, while holding all other explanatory variables at their mean value.

Not surprisingly, *ROE* is highly significant with a negative coefficient, implying that firms with good operating performance have lower crash risk during the year. Size is associated with greater crash risk and has the largest marginal impact of any of the explanatory variables. As we discuss in detail below, this positive association is largely a function of how we define firm-specific crashes. The market-to-book ratio is statistically significant in predicting crash risk in Table 7, but its marginal impact is small. Leverage is associated with lower crash risk. One would think the partial impact of leverage would be to increase crash risk; this contrary finding most likely reflects endogeneity in firms' capital structure choices, specifically, that more stable, less crash-prone firms are more willing or able to establish higher levels of indebtedness.

Turning now to our primary variables of interest, the coefficient on *OPAQUE* is positive and significant at better than a .01% level while the coefficient on *OPAQUE*^{2} is negative and also significant at better than a .01% level. The concavity of the relationship in this table is consistent with the convexity of the relation between *IDIOSYN* and *OPAQUE* found in Table 6. In both cases, the impact of *OPAQUE* tails off at higher values of earnings management. To examine the marginal impact for changes in earnings management we must combine the impacts of *OPAQUE* and *OPAQUE*^{2} because both variables will move in lockstep as earnings management increases. The combined marginal impact for the two right-hand side variables is .0156.[20]

[20] Because crash probability is non-linear in *OPAQUE* and *OPAQUE*^{2}, one cannot simply add the marginal impacts of the two variables (as we did in Table 6 where *IDIOSYN is* a linear function of the explanatory variables). The combined impact on probability is not the sum of the individual impacts that would be obtained by mechanically adding up the impacts of each. We compute the combined impact by setting *both* *OPAQUE* and *OPAQUE*^{2} to their 25th and 75th percentile values respectively, and comparing crash probabilities at the two percentile values.

Of course, given the concavity of the regression relationship in Table 7 with respect to *OPAQUE*, the incremental impact of an increase in *OPAQUE* steadily diminishes as its value increases. Figure 4 presents crash probabilities computed using mean values of *OPAQUE* for each of the ten portfolios formed by ranking on *OPAQUE*, with all other right-hand side variables set equal to their sample means. The sensitivity of crash risk to *OPAQUE* substantially declines at higher values of *OPAQUE*. The change in the implied probability of crashes across the bottom to the top decile is 2.7 percent.

To summarize, the impact of *OPAQUE* on crash risk is statistically highly significant and consistent with the Jin-Myers (2006) model. The quantitative or economic import of the variable is more difficult to assess. On the one hand, the 2.7 percentage point increase in crash probability that corresponds to the difference between the bottom and top decile of *OPAQUE* is certainly enough to warrant attention in any application focused on tail risk. On the other hand, the unconditional probability of observing a crash during a full year is about 17.1% (Table 2). So our earnings management-based measure of opacity roughly accounts for 2.7/17.1 = 15.8% of the variation in crash risk. On balance, while it appears that the majority of crashes are independent of accounting practice, knowing that a firm is an active earnings manager increases the conditional probability of a crash by an economically as well as statistically significant amount.

We performed similar robustness checks for the determinants of crash risk as for R-square, and obtain similar results. Panel B of Table 7 shows that the results are robust to the exclusion of ROE on the right-hand side (Line 2), requiring 51 weeks of annual data for each firm (Line 3), to the use of a single-year absolute value of discretionary accruals to measure opacity (Line 4), and to the use of 2-week versus 1-week holding periods (Line 5). Using the firm's percentile score for *OPAQUE* in its industry does not affect the qualitative conclusion that opacity increases crash risk, although statistical significance is reduced (Line 6). Finally, as in Table 6, Fama-MacBeth coefficients are a bit lower than the panel results, but the relation is actually stronger when we restrict the sample to the pre-Sarbanes-Oxley years.

For a different look at the magnitude and determinants of crash risk, Table 8, Panel A, presents crash risk for the same size-opacity sort presented in Table 5.

Interestingly, there is no pattern moving across the rows in the two small-firm quintiles, but a strong relation for the three larger ones. For the three largest quintiles the difference in crash probability between the high- and low-opacity groups increases from 3.4, to 4.1, to an astonishing 6.0 percentage points. These are statistically and certainly economically significant differences. Moving down each column the patterns are not as clear, with one exception. In each case, the lowest crash probability is associated with the smallest quintile. This seems counter to casual observation, but is in fact consistent with the positive coefficient on size in Table 7, suggesting that large firms appear to be more crash-prone than smaller ones.

This surprising result is an artifact of our definition of a crash. We define a crash as a tail event of sufficient magnitude to fall in the lower 0.1% of the normal distribution. This is roughly a three-sigma event. But large firms have lower standard deviations of returns than smaller ones (the Pearson correlation between size and standard deviation is −.465, significant at better than .01 percent level in Table 3). Therefore, the absolute magnitude of a return needed to qualify as a crash is smaller for larger firms. This is evident in Panel B of Table 8, which presents the mean crash return for each size-opacity cell, and in the margins, the mean crash return for each size quintile and opacity group. Moving down the columns, a clear pattern is evident: raw crash returns are smaller in absolute value for larger firms. The mean crash return is −0.25 for the smallest size quintile and −0.17 for the largest size quintile. Thus, using our definition of a crash, the higher crash propensity of larger firms does not imply necessarily wider swings in returns.

Given this pattern, there is an obvious temptation to define a crash event uniformly across firms, for example, as a 3-sigma event based on the average residual standard deviation from the entire sample of firm years. But any such uniform definition would *by construction* result in high-volatility firms being identified as high-crash firms. In general, to evaluate the "jumpiness" of any stochastic process, one needs to evaluate movements relative to the standard deviation *of that particular process*. In our context, crash episodes for each firm must be defined relative to the return volatility of that particular firm. By their nature, fat tails can be neither assessed nor identified without reference to a benchmark determined by the volatility of the specific underlying process.

While size and standard deviation are negatively correlated, *OPAQUE* and the volatility of weekly returns are positively correlated (the Pearson correlation between opacity and standard deviation is .358, significant at better than a .01 percent level). Therefore, the threshold return that qualifies as a crash is *larger* in absolute value for firms with greater earnings management. This is also apparent in Panel B of Table 8. Within each size quintile, the absolute value of the mean one-week crash return increases monotonically with opacity. Averaging across size quintiles, the mean crash return for the lowest level of opacity is –0.19 and –0.27 for the highest level of opacity.

4.4 Positive jump risk

Table 7 indicates that earnings management or the opacity of firms' financial reports predicts crashes. These results immediately raise another question: Does opacity predict fat tails generally, or only one-sided exposure to crashes? We define a positive jump symmetrically to a crash, that is, as a 3.09-standard deviation positive value of ln(1 + residual weekly return), and repeat the logit analysis for jump probability.

The simple sort in Table 8, Panel C shows no evidence of a relation between jump probability and opacity for any of the size quintiles. Table 9 presents logit analysis of jump probability similar to the results in Table 7 for crash risk. The important result in Table 9 is the striking difference between the power of *OPAQUE* to predict crashes (in Table 7) versus its lack of significance as a predictor of positive jumps. The Chi-square statistics for both *OPAQUE* and $OPAQUE^2$ in Table 9 are exceedingly low; none of these variables achieves significance at even a 10% confidence level in either of the specifications. Thus, in agreement with the Jin-Myers (2006) model, opacity appears to predict only negative crashes in stock prices.

The control variables in Table 9 are largely mirror images of their values in Table 7, with positive coefficients in Table 7 becoming negative coefficients in Table 9. This is the pattern that one would expect. The interesting exception to this rule is the coefficient on *ROE*, which is also negative (but not even close to statistically significant) in the positive jump regressions. The negative coefficient for contemporaneous *ROE* in both tables is supportive of a broad interpretation of the Jin-Myers (2006) and Kirschenheiter and Melumad (2002) models: In both models, a firm receiving bad news will try to hide the negative news until it is no longer feasible to do so, which ultimately results in a crash; thus opaque firms with poor performance are more subject to crashes. This explains why the coefficient on *ROE* would be negative in the Table 7 crash regressions. However, firms experiencing good performance have no incentives to hide such news. Because they would not be expected to stockpile positive information, their stock returns ought to reflect good news more steadily over time. Therefore, the impact of *ROE* on jump probability need not be symmetric to its impact on crash risk. Better performance encourages more revelation and therefore more continuous information release.

4.5 Sarbanes-Oxley and transparency

We have used the moving sum of absolute discretionary accruals as a proxy for firm opacity, and have documented that this variable predicts both R-square and crash risk. But earnings management may wax and wane over time, thus providing a natural experiment to further test the impact of opacity on stock price dynamics. Specifically, consider the potential impact of the Sarbanes-Oxley Act (SOX), passed into law in 2002. Other papers have noted that the passage of the Act, which substantially increased the penalties for earnings manipulation, materially reduced the incidence of accounting-based earnings management [Cohen, et al. (2007); Graham, et al. (2005)]. Presumably then, the Act would also have reduced the power of discretionary accruals to predict the properties of stock returns.

One may view discretionary accruals derived from residuals in the modified Jones model as the sum of model error plus intentional accruals management. If SOX has in fact been successful in reducing earnings management, then in the post-SOX period residuals from the modified Jones model would be composed more of model error and less of intentional accruals management. It follows that the predictive power of *OPAQUE* for both *IDIOSYN* and crash risk would dissipate in the post-SOX environment. Tables 10 and 11 test this hypothesis.

In Table 10, Models 1 and 2, we introduce two new explanatory variables: a SOX dummy (equal to zero prior to 2002, and equal to one in 2002 and beyond), and an interaction term equal to the product of *OPAQUE* and the SOX dummy. As in Table 6, Models 1 and 2 present specifications with and without kurtosis and skewness as control variables. In both cases, the coefficient on the SOX dummy is negative (−.309 to −.334), and highly significant (*t*-statistics above 25). The negative coefficient on SOX indicates that the baseline level of *IDIOSYN* has shifted down, or equivalently, that typical R-squares have increased since 2002. It is hard to assign causation here, as there have been several other changes that may have affected information flows from firms to investors, most notably Regulation Fair Disclosure.

Of greater interest is the coefficient on the product of *OPAQUE* and the SOX dummy. This coefficient can be interpreted as the change in the coefficient on *OPAQUE* in the post-SOX period. In both specifications, while the coefficient on *OPAQUE* is negative (−.189 in both models), the coefficient on the interaction term is positive and of slightly greater magnitude (either .233 or .234). The coefficients on both *OPAQUE* and the interaction term are highly significant, with *t*-statistics all greater than 6. In Model 3, we also add $OPAQUE^2$ and $OPAQUE^2$ times the SOX dummy. As we found for *OPAQUE*, the coefficient on $OPAQUE^2$ times the SOX dummy has the opposite sign and is slightly greater in magnitude than that on $OPAQUE^2$, suggesting that the impact of earnings management ends after 2002.

Therefore, it seems that the relationship between abnormal accruals and *IDIOSYN* largely disappears in the post-SOX period. This is consistent with the hypothesis that in a period with greater monitoring and scrutiny of accounting practice, earnings management has substantially subsided, and residuals from the modified Jones model no longer proxy for the opacity of firms' financial reports.

Table 11 repeats our logit crash analysis allowing for different impacts of opacity in the pre- and post-SOX periods. The results are highly consistent with Tables 7 and 10. *OPAQUE* and $OPAQUE^2$ predict crashes with high significance in the pre-SOX period, but while the coefficients on the post-SOX dummy interaction variables are estimated with far less precision, they are all of opposite sign and comparable magnitude as the pre-SOX values, implying that opacity does not predict crashes in the post-SOX period. Table 12 examines jump propensities in the pre- and post-SOX periods. Consistent with Table 9, *OPAQUE* is not statistically significant in predicting jumps in either the pre- or the post-SOX periods.

5. Conclusions

We confirm that our firm-specific measure of opacity, the three-year moving sum of the absolute value of discretionary accruals, is a reliable predictor of R-square. We also examine the more sparsely considered relation between opacity and crash risk. We find that our measure of earnings management reliably predicts such risk. In contrast, positive jumps in stock returns are essentially unrelated to our measure of opacity. Reasonable variation in this measure of opacity changes either R-square or crash risk by about 15% of their unconditional or baseline levels, but the effect on crash risk is substantially more pronounced for larger firms. While opacity, or at least our proxy for it, is clearly part of a larger story in explaining stock price dynamics, it does in fact appear to be an economically meaningful part of that story, and the statistical relation between it and both R-square and crash risk is strong. These results are new and consistent with a theoretical model of crash risk based on information management: Firms manage earnings to shelter bad information up to a point, but once some tipping point is crossed, the information comes out in one fell swoop, which results in a price crash. However, this phenomenon appears to have dissipated in the post-SOX years.

References

Ashbaugh, Hollis, Joachim Gassen, and Ryan Lafond, 2005, Does Stock Price Synchronicity Reflect Information Or Noise? The International Evidence, mimeo

Barberis, Nicholas, Andrei Shleifer, and Jeffrey Wurgler, 2005, Comovement, Journal of Financial Economics 75, 283-317.

Bates, D., 2001, Post-'87 Crash Fears in the S&P 500 Futures Options Market, Journal of Financial Economics 94, 181-238.

Bekaert, Geert, and Guojun Wu, 2000, Asymmetric Volatility and Risk in Equity Markets, Review of Financial Studies 13, 1-42.

Bergstresser, D., and T. Philippon, 2006. CEO incentives and earnings management. Journal of Financial Economics 80, 511-529.

Bergstresser, D., M. Desai and J. Rauh, 2006. Earnings manipulation, pension assumptions, and managerial investment decisions. Quarterly Journal of Economics 121, 157-195.

Burgstahler, David and Michael Eames, 2006, "Management of Earnings and Analysts' Forecasts to Achieve Zero and Small Positive Earnings Surprises" . Journal of Business Finance & Accounting, Vol. 33, No. 5-6, pp. 633-652

Campbell, J. Y., and L. Hentschel, 1992. No News is Good News: An Asymmetric Model of Changing Volatility in Stock Returns, Journal of Financial Economics 31, 281-318.

Cheng, Qiang and Farber David, 2008. Earnings restatements, changes in CEO compensation, and firm performance, forthcoming Accounting Review.

Cohen, D., A. Dey, and T. Lys, 2007, Real and Accrual-Based Earnings Management in the Pre- and Post-Sarbanes-Oxley Periods, forthcoming in the Accounting Review.

Collins, D.W. and S.P. Kothari, 1989, An analysis of intertemporal and cross-sectional determinants of earnings response coefficients, Journal of Accounting and Economics 11, 143-181.

Cornett, M.M., A.J. Marcus, and H. Tehranian, 2008. Corporate governance and pay-for-performance: The impact of earnings management," Journal of Financial Economics 87, 357-373.

Dechow, P., R. Sloan, A. Sweeney, 1995, Detecting Earnings Management. Accounting Review 70, 193-225.

Dechow, P., R. Sloan, A. Sweeney, 1996, Causes and Consequences of Earnings Manipulation: An Analysis of Firms Subject to Enforcement Actions by the SEC. Contemporary Accounting Research 13, 1-36.

Dechow, P. and C.M. Schrand, 2004, Earnings Quality. Research Foundation of CFA Institute monograph.

Dimson, Elroy, 1979, Risk Measurement when Shares are Subject to Infrequent Trading, Journal of Financial Economics 7, 197-227.

Dumas, B., J. Fleming, and R. E. Whaley, 1998, Implied Volatility Functions: Empirical Tests, Journal of Finance 53, 2059-2106.

Durnev, A., R. Morck, B. Yeung, & P. Zarowin, 2003, Does Greater Firm-Specific Return Variation Mean More Or Less Informed Stock Pricing? Journal of Accounting Research 41, 797-836.

Fama, E. and J. MacBeth, 1973. Risk, return, and equilibrium: Empirical tests. Journal of Political Economy (May-June) 81, 607-636.

Ferreira, Miguel A., and Paul A. Laux, 2007, Corporate Governance, Idiosyncratic Risk, and Information Flow, Journal of Finance 62, 951-989.

French, Kenneth R., G. W. Schwert, and R. F. Stambaugh, 1987, Expected Stock Returns and Volatility, Journal of Financial Economics 19, 3-29.

Graham, J., C. Harvey, and S. Rajgopal, 2005, The Economic Implications of Corporate Financial Reporting, Journal of Accounting and Economics 40, 3-73.

Hong, H., and J. Stein, 2003, Differences of Opinion, Short-Sales Constraints, and Market Crashes, Review of Financial Studies 16, 487-525.

Jin, Li and S. C. Myers, 2006. R^2 Around the World: New Theory and New Tests. Journal of Financial Economics 79, 257-292.

Jones, J., 1991, Earnings management during import relief investigations. Journal of Accounting Research 29: 193-228.

Kinney, William, David Burgstahler, and Roger Martin, 2002, Earnings Surprise “Materiality” as Measured by Stock Returns, Journal of Accounting Research 40, 1297-1329.

Kirschenheiter, Michael, and Nahum D. Melumad, 2002, Can “Big Bath” and Earnings Smoothing Co-Exist as Equilibrium Financial Reporting Strategies? Journal of Accounting Research 40, 761-796.

Klein, A., 2002. Audit committee, board of director characteristics, and earnings management, Journal of Accounting and Economics 33, 375-400.

Kothari, S.P., Susan Shu, and Peter Wysocki, 2008, Do Managers Withhold Bad News? forthcoming *Journal of Accounting Research*.

Li, K., R. Morck,, F. Yang, B. Yeung, 2004, Firm-specific variation and openness in emerging markets, Review of Economics and Statistics 86, 658-669.

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

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 58, 215-260.

Piotroski, Joseph D., and Darren T. Roulstone, 2005, Do Insider Trades Reflect both Contrarian Beliefs and Superior Knowledge about Future Cash Flow Realizations? Journal of Accounting and Economics 39, 55-81.

Pontiff, Jeffrey, 1996, Costly Arbitrage: Evidence from Closed-End Funds. Quarterly Journal of Economics 111, 1135-1151.

Roll, Richard, 1988, R^2, Journal of Finance 43, 541-566.

Romer, David, 1993, Rational Asset-Price Movements without News, American Economic Review 83, 1112-1130.

Sloan, Richard, 1996, Do Stock Prices Fully Reflect Information in Accruals and Cash Flows about Future Earnings? Accounting Review 71, 289-315.

Skinner, Douglas, and Richard Sloan, 2002, Earnings surprises, growth expectations, and stock returns or don't let an earnings torpedo sink your portfolio. Review of Accounting Studies 7 (2-3), 289-312

Teoh, Siew Hong, Ivo Welch and T. J. Wong, 1998a, Earnings management and the underperformance of seasoned equity offerings, Journal of Financial Economics 50, 63-99.

Teoh, Siew Hong, Ivo Welch and T. J. Wong, 1998b, Earnings management and the long-term market performance of initial public offerings, Journal of Finance 53, 1935-1974.

West, Kenneth D., 1988, Bubbles, Fads, and Stock Price Volatility Tests: A Partial Evaluation, Journal of Finance 43, 639-656.

Wie, Steven, and Chu Zhang, 2006, Why did Individual Stocks Become More Volatile? Journal of Business 79, 259-292.

Appendix A: Variable Definitions

OPAQUE is the prior three years' moving sum of the absolute value of discretionary accruals. Specifically,

$$OPAQUE = \text{AbsV}(DiscAcc_{t\text{-}1}) + \text{AbsV}(DiscAcc_{t\text{-}2}) + \text{AbsV}(DiscAcc_{t\text{-}3})$$

where $DiscAcc_t$ is measured using the Modified Jones Model.

Firm-Specific Weekly Return is equal to ln(1+ residual), where the residual is $\varepsilon_{j,t}$ from the expanded index-model regression:

$$r_{j,t} = \alpha_j + \beta_{1,j}r_{m,t-1} + \beta_{2,j}r_{i,t-1} + \beta_{3,j}r_{m,t} + \beta_{4,j}r_{i,t} + \beta_{5,j}r_{m,t+1} + \beta_{6,j}r_{i,t+1} + \varepsilon_{j,t}$$

IDIOSYN is a measure of the firm-specific information arriving to the securities market based on $(1 - R^2)$ from the modified index model regression. Specifically,

$$IDIOSYN = \ln\left(\frac{1 - R^2}{R^2}\right)$$

Std Dev ln(1 + residual) is the standard deviation of the *Firm-Specific Weekly Return* over the fiscal year.

Kurtosis is the kurtosis of the *Firm-Specific Weekly Return* over the fiscal year.

Skewness is the skewness of the *Firm-Specific Weekly Return* over the fiscal year.

CRASH is an indicator variable equal to one if within its fiscal year a firm experiences one or more *Firm-Specific Weekly Returns* falling 3.09 or more standard deviations below the mean *Firm-Specific Weekly Return* for its fiscal year; zero otherwise.

JUMP is an indicator variable equal to one if within its fiscal year a firm experience one or more *Firm-Specific Weekly Returns* rising 3.09 or more standard deviations above the mean *Firm-Specific Weekly Return* for its fiscal year; zero otherwise.

SIZE is the natural log of the market value of equity at the beginning of the fiscal year.

M to B is the ratio of the market value of equity to the book value of equity measured at the beginning of the fiscal year.

LEV is the book value of all liabilities scaled by total assets, measured at the beginning of the fiscal year.

ROE is contemporaneous return on equity defined as income before extraordinary items divided by the book value of equity.

VAR(Industry Index) is the variance of the weekly returns of the Fama / French industry index during the firm's fiscal year.

SOX is an indicator variable equal to one in sample years 2002 and beyond; zero otherwise.

Table 1: Sample Development, Industry Membership, and Fiscal Years of Sample

Panel A: Sample Development

	# firm years
All Compustat firm fiscal years 1991 through 2005	109,026
Excluding firm fiscal years:	
with incomplete stock return data	22,777
financial services and utilities	28,051
with insufficient fin'l data to calc. control variables	7,413
with insufficient fin'l data to calc. three lags of discretionary accruals	9,903
Final Sample	**40,882**

Table 1: Sample Development, Industry Membership and Fiscal Years of Sample

Panel B: Fama French Industries*

Industry	# firm years	Industry	# firm years
Aero	242	Hshld	922
Agric	181	LabEq	1,188
Autos	882	Mach	1,870
Beer	181	Meals	910
BldMt	1,066	MedEq	1,610
Books	517	Mines	223
Boxes	173	Oil	1,981
BusSv	2,260	Paper	867
Chems	1,003	PerSv	482
Chips	3,019	Retail	2,682
Clths	792	Rubbr	509
Cnstr	514	Ships	107
Coal	37	Smoke	53
Drugs	2,370	Soda	110
ElcEq	846	Softw	2,851
FabPr	247	Steel	912
Food	936	Telcm	1,246
Fun	681	Toys	433
Gold	253	Trans	1,295
Guns	80	Txtls	320
Hardw	1,390	Whlsl	1,889
Hlth	752		**40,882**

* Source: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/det_49_ind_port.html

Table 1: Sample Development, Industry Membership and Fiscal Years of Sample

Panel C: Observations in each Fiscal Year

Fiscal Year	Number of Obs.
1991	2,022
1992	2,295
1993	2,441
1994	2,557
1995	2,680
1996	2,872
1997	3,023
1998	2,976
1999	2,858
2000	3,053
2001	2,754
2002	2,726
2003	2,842
2004	3,105
2005	2,678
	40,882

Table 2: Crashes and Jumps – Frequency and Weekly Returns

Panel A: Crash and Jump Frequency

Panel A reports the frequency of firm-specific crashes and jumps for 40,882 firm-years in the sample period 1991-2005. Crashes and jumps are defined based on residuals from an expanded index model regression with market and industry returns as explanatory variables. Weekly firm-specific residual returns that are 3.09 standard deviations below or above the mean for the firm's fiscal year are categorized as crashes or jumps.

Frequency of *Crashes* in the firm year			Frequency of *Jumps* in the firm year		
	# of Obs.	% of sample		# of Obs.	% of sample
0	33,932	83.0%	0	31,853	77.9%
1	6,766	16.6%	1	8,767	21.4%
2	184	0.5%	2	261	0.6%
3	0	0.0%	3	1	0.0%
	======			======	
	40,882			40,882	

See **Appendix A** for all variable definitions.

Table 2: Crashes and Jumps: Frequency and Weekly Returns

Panel B: Return in Crash and Jump Weeks vs. All Other Weeks

Panel B reports the mean, median, and variance of raw weekly returns for *CRASH* weeks, *JUMP* weeks, and *ALL OTHER* weeks. The first set of columns presents statistics for individual firms. The middle set of columns presents analogous results for the market index. In this panel, *CRASH* (or *JUMP*) weeks refer to any week in which any firm in the sample crashes (or jumps). The last set of columns report statistics averaging across industries. If any firm in an industry crashes (or jumps) in a given week, that is defined as a crash (or a jump) week for the industry. Weekly returns for 40,882 firm-years in the sample period 1991-2005.

		Firm Returns			**Market Index**			**Industry Index**		
	# of obs.	mean	median	variance	mean	median	variance	mean	median	variance
***CRASH* Weeks**	7,134	-0.2274	-0.2014	0.0147	0.0033	0.0039	0.0004	0.0020	0.0020	0.0010
***JUMP* Weeks**	9,292	0.3327	0.2623	0.1007	0.0055	0.0066	0.0003	0.0084	0.0071	0.0008
***ALL OTHER* Weeks**	2,101,147	0.0026	0.0000	0.0054	0.0023	0.0036	0.0005	0.0025	0.0029	0.0010

Table 3: Key Variables of Interest

Panel A: Descriptive Statistics

40,882 firm years in the sample period 1991-2005. Variable definitions appear in Appendix A.

	Q1	Mean	Median	Q3	Std Dev
Market Value (lagged)	63.2	3,157.8	258.0	1,191.5	14,300.5
Opaque	0.102	0.243	0.174	0.302	0.222
R^2	0.128	0.250	0.209	0.331	0.164
Std Dev[ln(1 + residual)]	0.036	0.058	0.052	0.074	0.030
Kurtosis	0.152	1.672	0.878	2.146	2.730
Skewness	-0.258	0.113	0.134	0.523	0.776
ROE	-0.006	-0.011	0.087	0.155	0.426
Market-to-book (lagged)	1.240	3.077	2.028	3.453	3.434
Leverage (lagged)	0.308	0.475	0.484	0.634	0.213
Var(Industry Index)	0.0004	0.0010	0.0007	0.0012	0.0010

Table 3: Key Variables of Interest (continued)

Panel B: Correlation Matrix (Pearson correlations are above the diagonal; Spearman correlations below the diagonal.)
Correlations computed from 40,882 firm-years in the sample period 1991-2005. *p*-values appear below correlations. See Appendix A for variable definitions.

	Size	Opaque	R^2	Std Dev ln(1+resid)	Kurtosis	Skewness	Crashes	Jumps	ROE	Market to Book	Leverage	Variance of Indus. Index
Size		-0.215	0.580	-0.465	-0.083	-0.210	0.076	-0.171	0.182	0.245	0.142	0.048
(lagged)		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
Opaque	-0.282		-0.074	0.358	0.035	0.030	0.004	0.034	-0.216	0.182	-0.113	0.129
	<.0001		<.0001	<.0001	<.0001	<.0001	0.4048	<.0001	<.0001	<.0001	<.0001	<.0001
R^2	0.548	-0.100		-0.305	-0.179	-0.117	-0.024	-0.157	0.096	0.151	0.072	0.176
	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
Std Dev	-0.511	0.407	-0.311		0.236	0.038	0.082	0.131	-0.388	0.116	-0.129	0.309
ln(1+resid)	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
Kurtosis	-0.129	0.070	-0.215	0.224		-0.098	0.430	0.373	-0.064	-0.008	-0.027	0.002
	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	0.1074	<.0001	0.6691
Skewness	-0.239	0.048	-0.163	0.077	0.086		-0.597	0.556	-0.001	-0.073	0.014	-0.032
	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	0.8881	<.0001	0.0063	<.0001
Crashes	0.088	0.010	-0.015	0.082	0.486	-0.565		-0.149	-0.023	0.034	-0.021	0.021
	<.0001	0.0398	0.0023	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001
Jumps	-0.175	0.042	-0.169	0.119	0.535	0.592	-0.149		-0.034	-0.052	-0.003	-0.025
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	0.5510	<.0001
ROE	0.281	-0.157	0.148	-0.374	-0.100	-0.025	-0.030	-0.048		-0.062	0.034	-0.099
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001
M/B Ratio	0.422	0.097	0.242	-0.030	-0.051	-0.131	0.068	-0.097	0.294		0.059	0.090
(lagged)	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001
Leverage	0.136	-0.142	0.061	-0.157	-0.031	0.008	-0.022	-0.004	0.127	-0.020		-0.090
(lagged)	<.0001	<.0001	<.0001	<.0001	<.0001	0.0991	<.0001	0.4767	<.0001	<.0001		<.0001
Variance of	0.045	0.170	0.177	0.309	0.012	-0.036	0.022	-0.024	-0.092	0.027	-0.075	
Industry Index	<.0001	<.0001	<.0001	<.0001	0.0162	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	

Table 4: Opacity and Financial Restatements

This table documents the relation between opacity and the incidence of restatements of financial statements. Panel A divides firm-years into three groups based on the value of *OPAQUE* and presents the rate of restatements for each group. Panel B compares the levels of *OPAQUE* for firm-years with and without restatements.

Panel A: Incidence of restatements by opacity measure

Opacity group	**Percent of observations with restatements**
Lowest	1.20%
Middle	1.63%
Highest	1.92%
Chi-square for equality across groups	8.35 (p = .0154)
Chi-square for equality between high and low opacity groups	8.37 (p = .0038)

Panel B: Levels of opacity for restatement and no-restatement firm-years.

	Observations	**Mean *OPAQUE***	**Median *OPAQUE***
No-restatements	14,437	.2645	.1862
Restatements	233	.3634	.2183
t-statistic for equality across groups		2.60	3.75
p-value		.0098	<.0001

Table 5: Firm characteristics in groups sorted by size and opacity

We sort our sample into five size quintiles and into three opacity groups (1 = low; 3 = high). Break points for the two sorts are estimated independently, so the number of observations in each cell differs. Panel A presents the number of observations in each cell; Panel B demonstrates that opacity varies substantially even within size groups. Panel C presents the average R^2 for firm-years in each cell, and a test of the hypothesis that R^2 is equal in the high versus low opacity group within each size quintile. 40,882 firm-years in the sample period 1991-2005.

Panel A: Observations in each size/opacity group

	Market cap	Opacity group			
Size	**($M)**	**1**	**2**	**3**	**Total**
1	39	1,568	2,701	3,902	8,171
2	123	2,002	2,772	3,406	8,180
3	331	2,536	2,778	2,865	8,179
4	994	3,249	2,855	2,076	8,180
5	14,318	4,267	2,526	1,379	8,172
Total		13,622	13,632	13,628	

Panel B: Level of opacity for each size/opacity group

	Opacity group		
Size	**1**	**2**	**3**
1	0.085	0.184	0.482
2	0.082	0.181	0.488
3	0.082	0.178	0.477
4	0.080	0.176	0.448
5	0.074	0.175	0.433

Panel C: R-square for each size/opacity group

	Opacity group			
Size	**1**	**2**	**3**	***p*-value: (3) – (1)**
1	0.147	0.150	0.160	<.0001
2	0.177	0.181	0.189	<.0001
3	0.224	0.225	0.233	0.0058
4	0.278	0.284	0.292	0.0011
5	0.395	0.403	0.417	0.0003

Table 6: Relation between *IDIOSYN* and Opacity

Panel A: OLS regressions of *IDIOSYN* as a function of opacity and control variables. *IDIOSYN* measures firm-specific information. $IDIOSYN = \ln[(1 - R^2)/R^2]$. 40,882 firm-years in the sample period 1991-2005. *t*-statistics appear below regression coefficients. Economic impact is the expected change in firm R-square resulting from an increase in each right-hand side variable from the 25th to the 75th percentile of the sample distribution. It is derived by calculating the impact on *IDIOSYN* and inverting to estimate the impact on R-square. See **Appendix A** for all variable definitions.

	Model 1	**Model 2**	**Model 3**	**Economic Impact (Model 3)**
Intercept	2.870	2.740	2.782	
	190.22	177.52	160.66	
OPAQUE	-0.163	-0.167	-0.402	.015
	-8.95	-9.34	-8.49	
Var(Industry Index)	-125.228	-124.669	-123.064	.019
	-33.49	-33.89	-33.35	
Size (lagged)	-0.254	-0.247	-0.249	.157
	-127.44	-122.82	-121.89	
Market-to-book (lagged)	0.003	0.003	0.003	−.001
	2.51	2.43	2.67	
Leverage (lagged)	0.001	0.006	0.003	.000
	0.07	0.35	0.15	
ROE	-0.022	-0.008	-0.006	.000
	-2.43	-0.94	-0.67	
Skewness		0.030	0.030	−.004
		6.17	6.06	
Kurtosis		0.051	0.051	−.018
		37.02	37.14	
$OPAQUE^2$			0.228	−.003
			5.35	
R-square (adj)	0.335	0.357	0.357	

Panel B: Robustness checks: coefficients on *OPAQUE* and *OPAQUE*^{2} for variant specifications of Model 3 from Panel A. *t*-statistics appear below regression coefficients.

Model variation	Coefficient on *OPAQUE* *t*-statistic	Coefficient on *OPAQUE*^{2} *t*-statistic
1. Baseline regression (specification 3 of Panel A)	−.402 −8.49	.228 5.35
2. Excluding ROE from set of explanatory variables	−.401 −8.48	.230 5.39
3. Requiring 51 weeks of data per year	−.391 −8.04	.223 5.02
4. Measuring opacity using single-year discretionary accruals	−.462 −5.66	.554 3.44
5. Using biweekly rather than weekly returns in index model	−.431 −7.73	.231 4.58
6. Using percentile rank of *OPAQUE* instead of its numerical value	−.099 −7.33	N/A
7. By-year (Fama-Macbeth) regressions, full sample	−.189 −1.72	.115 1.53
8. By-year (Fama-Macbeth) regressions, excluding post-2002 data	−.333 −3.31	.211 3.16

Table 7: Using Opacity to Predict Crash Risk

Panel A: Logit regressions to explain crash risk. The dependent variable, *CRASH* is an indicator variable equal to one if within its fiscal year a firm experiences one or more *Firm-Specific Weekly Returns* falling 3.09 or more standard deviations below the mean *Firm-Specific Weekly Return* for its fiscal year; zero otherwise. 40,882 firm-years in the sample period 1991-2005. See **Appendix A** for all variable definitions.

	Coefficient	Chi-sqr.	*p*-value	Marginal impact
Intercept	-2.232	1364.49	<.0001	
OPAQUE	0.855	25.61	<.0001	
$OPAQUE^2$	-0.729	22.11	<.0001	0.0156
ROE_t	-0.213	51.73	<.0001	-0.0047
$SIZE_{t-1}$	0.118	278.63	<.0001	0.0477
M to B_{t-1}	0.007	2.85	0.091	0.0020
LEV_{t-1}	-0.399	39.38	<.0001	-0.0181
Chi-square	369.53	<.0001		
# of observations				
Crash	6,950			
No Crash	33,932			

Panel B: Robustness checks: coefficients on *OPAQUE* and *OPAQUE*^{2} for variant specifications of the Panel A regression. Significance levels appear below regression coefficients.

Model	Coefficient on *OPAQUE* significance level	Coefficient on *OPAQUE*^{2} significance level
1. Baseline regression (as reported in Panel A)	.855 <.0001	−.729 <.0001
2. Excluding ROE from set of explanatory variables	.882 <.0001	−.682 <.0001
3. Requiring 51 weeks of data per year	.942 <.0001	−.802 <.0001
4. Measuring opacity using single-year discretionary accruals	.800 .0067	−.802 <.0001
5. Using biweekly rather than weekly returns in index model	.804 .0067	−.773 .0047
6. Using percentile rank of OPAQUE instead of its numerical value	.083 .088	N/A
7. By-year (Fama-Macbeth) regressions	.762 <.0001	−.702 <.0001
8. By-year (Fama-Macbeth) regressions, excluding post-2002 data	.932 <.0001	−.851 <.0001

Table 8: Crash and Jump Probabilities for firms sorted on size and opacity

We sort our sample into five size quintiles and into three opacity groups (1 = low; 3 = high). Break points for the two sorts are estimated independently, so the number of observations in each cell differs. *CRASH* is an indicator variable equal to one if within its fiscal year a firm experiences one or more *Firm-Specific Weekly Returns* falling 3.09 or more standard deviations below the mean *Firm-Specific Weekly Return* for its fiscal year; zero otherwise. *JUMP* is an indicator variable equal to one if within its fiscal year a firm experiences one or more *Firm-Specific Weekly Returns* falling 3.09 or more standard deviations above the mean *Firm-Specific Weekly Return* for its fiscal year; zero otherwise. 40,882 firm-years in the sample period 1991-2005.

Panel A: Crash probability for each size/opacity group

	Opacity Group			Difference	*p*-value
Size	**1**	**2**	**3**	**(3) – (1)**	**(3) – (1)**
1	0.117	0.108	0.108	−.009	0.362
2	0.151	0.153	0.158	.007	0.499
3	0.174	0.184	0.208	.034	0.001
4	0.202	0.221	0.243	.041	<.0001
5	0.161	0.181	0.221	.060	<.0001

Panel B: Mean of crash returns (raw one-week stock returns) for each size/opacity group

	Opacity Group			mean
Size	**1**	**2**	**3**	
1	-0.22	-0.24	-0.26	-0.25
2	-0.23	-0.25	-0.28	-0.26
3	-0.21	-0.23	-0.29	-0.25
4	-0.19	-0.22	-0.26	-0.22
5	-0.15	-0.18	-0.22	-0.17
mean	-0.19	-0.22	-0.27	

Panel C: Jump probability for each size/opacity group

	Opacity Group			Difference	*p*-value
Size	**1**	**2**	**3**	**(3) – (1)**	**(3) – (1)**
1	0.337	0.358	0.340	.003	0.813
2	0.255	0.257	0.247	−.008	0.521
3	0.208	0.208	0.198	−.010	0.348
4	0.174	0.168	0.161	−.013	0.227
5	0.132	0.143	0.123	−.009	0.367

Table 9: Using Opacity to Predict Positive Jumps

Logit regressions to explain jump probability. The dependent variable, *JUMP* is an indicator variable equal to one if within its fiscal year a firm experiences one or more *Firm-Specific Weekly Returns* falling 3.09 or more standard deviations above the mean *Firm-Specific Weekly Return* for its fiscal year; zero otherwise. 40,882 firm-years in the sample period 1991-2005. See **Appendix A** for all variable definitions.

	Coefficient	Chi-sqr.	*p*-value	Marginal impact
Intercept	-0.167	9.39	0.002	
OPAQUE	-0.164	1.15	0.283	
$OPAQUE^2$	0.172	1.59	0.208	-0.0031
ROE_t	-0.030	1.12	0.290	-0.0007
$SIZE_{t-1}$	-0.214	962.59	<.0001	-0.0881
M to B_{t-1}	-0.011	7.06	0.008	-0.0033
LEV_{t-1}	0.221	14.86	0.000	0.0100
Chi-square	1270.48	<.0001		
# of observations				
Jump	9,029			
No Jump	31,853			

Table 10: Relation between *IDIOSYN* and Opacity Pre- and Post-SOX

OLS regressions of *IDIOSYN* as a function of opacity and control variables. *IDIOSYN* measures firm-specific information. SOX is an indicator variable equal to one in 2002 and beyond and zero otherwise. 40,882 firm-years in the sample period 1991-2005. *t*-statistics appear below regression coefficients. See **Appendix A** for all variable definitions.

	Model 1	**Model 2**	**Model 3**
Intercept	2.916	2.782	2.833
	191.14	179.12	159.77
SOX	-0.309	-0.334	-0.399
	-25.06	-27.58	-23.03
OPAQUE	-0.189	-0.189	-0.512
	-8.67	-8.81	-9.20
SOX*OPAQUE	0.233	0.234	0.723
	6.54	6.70	7.46
Var(industry return)	-135.30	-135.76	-134.40
	-36.42	-37.24	-36.76
Size (lagged)	-0.243	-0.234	-0.236
	-120.63	-115.78	-114.81
Market-to-book (lagged)	0.001	0.001	0.001
	1.02	0.71	0.89
Leverage (lagged)	-0.037	-0.035	-0.036
	-2.05	-1.97	-2.04
ROE	-0.030	-0.016	-0.014
	-3.32	-1.77	-1.53
Skewness		0.028	0.028
		5.88	5.76
Kurtosis		0.054	0.054
		39.91	39.99
$OPAQUE^2$			0.322
			6.31
SOX* $OPAQUE^2$			-0.485
			-5.49
R-square (adj)	0.349	0.374	0.374

.

Table 11: Using Opacity to Predict Crash Risk, Pre- and Post-SOX

Logit regressions to explain crash risk. The dependent variable, *CRASH* is an indicator variable equal to one if within its fiscal year a firm experiences one or more *Firm-Specific Weekly Returns* falling 3.09 or more standard deviations below the mean *Firm-Specific Weekly Return* for its fiscal year; zero otherwise. SOX is an indicator variable equal to one in 2002 and beyond and zero otherwise. 40,882 firm-years in the sample period 1991-2005. See **Appendix A** for all variable definitions.

	Coefficient	Chi-sqr.	*p*-value	Marginal impact
Intercept	-2.285	1284.14	<.0001	
SOX	0.389	41.61	<.0001	0.0570
OPAQUE	0.972	21.35	<.0001	
$OPAQUE^2$	-0.893	20.14	<.0001	0.0173
SOX*OPAQUE	-0.622	3.37	0.067	
$SOX*OPAQUE^2$	0.565	3.23	0.072	-0.0102
ROE_t	-0.209	49.08	<.0001	-0.0046
$SIZE_{t-1}$	0.105	209.86	<.0001	0.0421
M to B_{t-1}	0.010	6.11	0.013	0.0030
LEV_{t-1}	-0.357	31.51	<.0001	-0.0162
Chi-square	477.47	<.0001		
# of observations				
Crash	6,950			
No Crash	33,932			

Table 12: Using Opacity to Predict Positive Jumps, Pre- and Post- SOX

Logit regressions to explain jump probability. The dependent variable, *JUMP* is an indicator variable equal to one if within its fiscal year a firm experiences one or more *Firm-Specific Weekly Returns* falling 3.09 or more standard deviations above the mean *Firm-Specific Weekly Return* for its fiscal year; zero otherwise. SOX is an indicator variable equal to one in 2002 and beyond and zero otherwise. 40,882 firm-years in the sample period 1991-2005. See **Appendix A** for all variable definitions.

	Coefficient	Chi-square	*p*-value	Marginal impact
Intercept	-0.169	8.87	0.003	
SOX	0.196	11.20	0.001	0.0570
OPAQUE	-0.266	2.13	0.145	
$OPAQUE^2$	0.239	2.08	0.149	-0.0056
SOX*OPAQUE	0.108	0.12	0.734	
$SOX*OPAQUE^2$	-0.116	0.16	0.686	0.0023
ROE_t	-0.026	0.8687	0.3513	-0.0046
$SIZE_{t-1}$	-0.224	1012.58	<.0001	0.0421
M to B_{t-1}	-0.009	4.81	0.028	0.0030
LEV_{t-1}	0.248	18.7109	<.0001	-0.0162
Chi-square	1327.45	<.0001		
# of observations				
Jump	9,029			
No Jump	31,853			

Figure 1a: Mean Absolute Value of Discretionary Accruals and Mean Weekly Return on the CRSP value-weighted market index

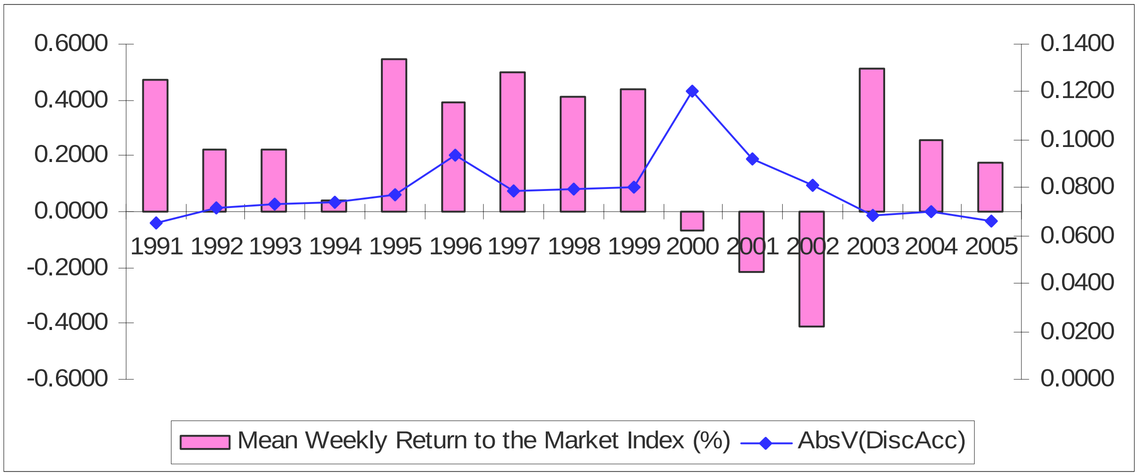

Figure 1b: OPAQUE and Mean Weekly Return on the CRSP value-weighted market index

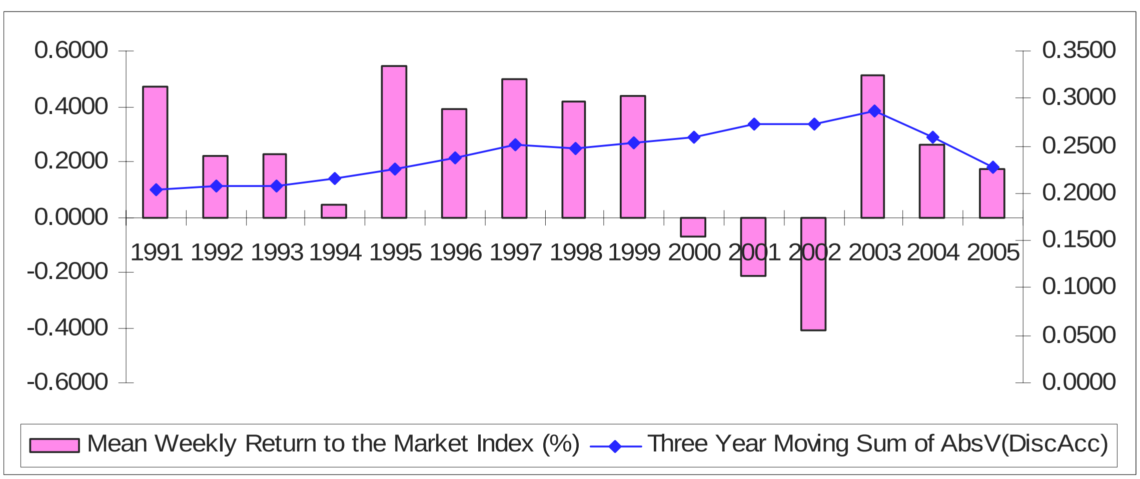

Figure 2a: Crash Frequency and the Mean Weekly Return on the CRSP value-weighted market index

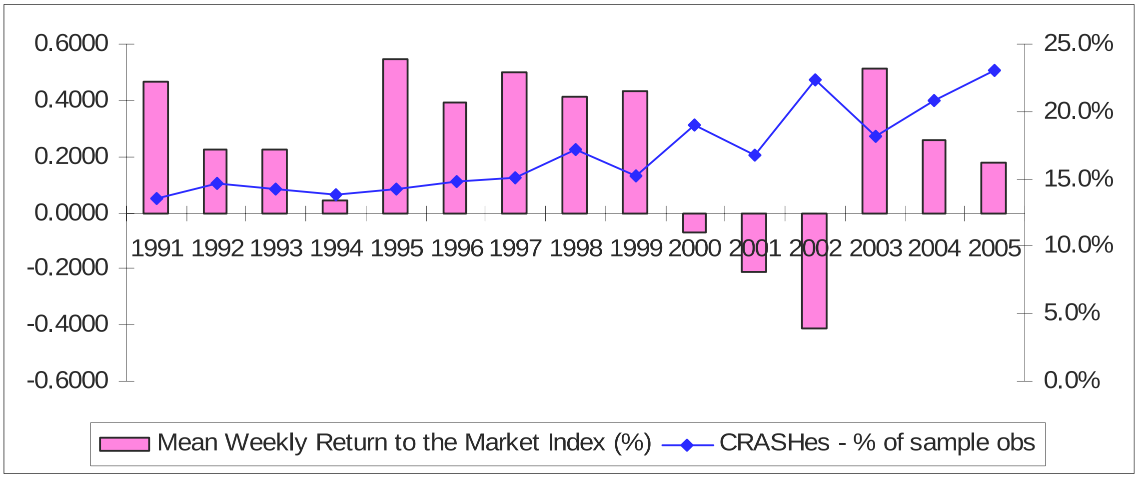

Figure 2b: Jump Frequency and the Mean Weekly Return on the CRSP value-weighted market index

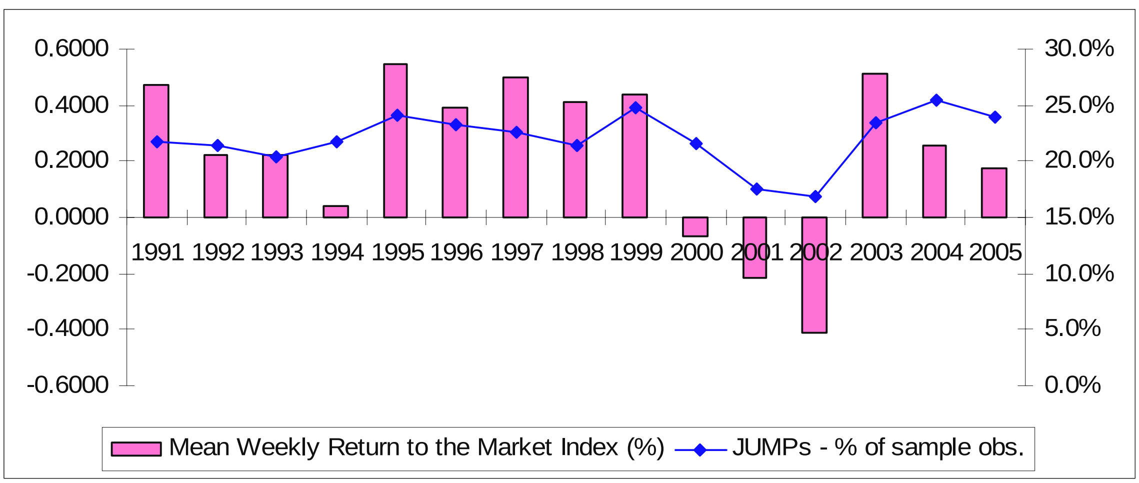

Figure 3: R-square as a function of opacity

The horizontal axis value for point on the graph is the average value of *OPAQUE* for decile portfolios formed by ranking on *OPAQUE*. The vertical axis value is the inferred value of R-square from regression Model 5 of Table 6, with all other right-hand side variables set equal to their sample average values.

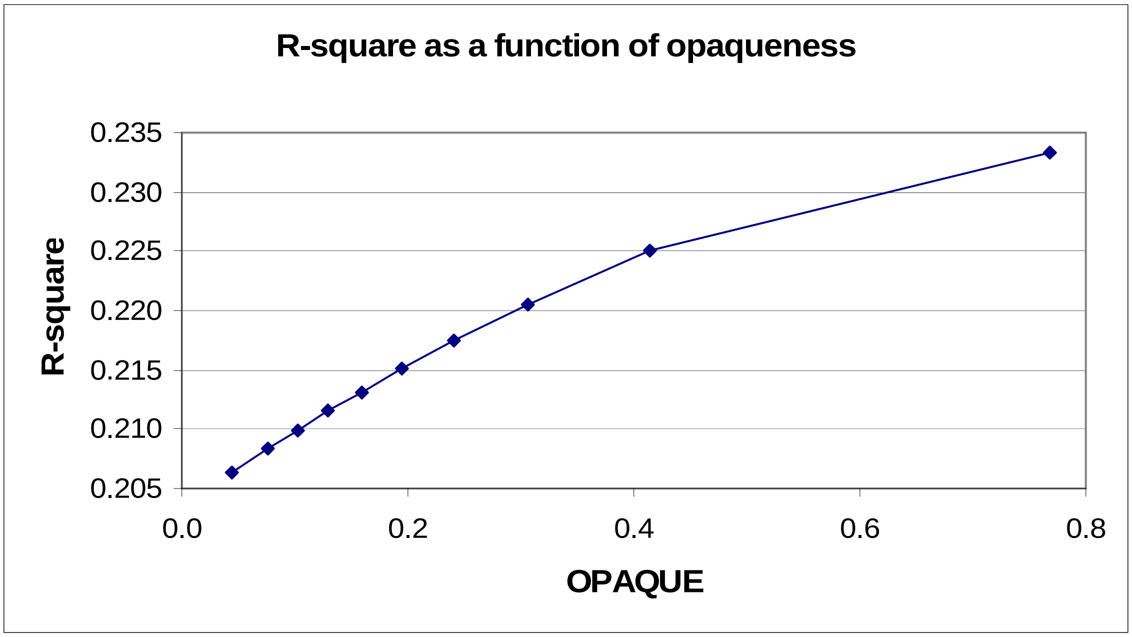

Figure 4: Annual crash probability as a function of opacity

The horizontal axis value for point on the graph is the average value of *OPAQUE* for decile portfolios formed by ranking on *OPAQUE*. The vertical axis value is the implied probability of a crash during the year for each value of *OPAQUE* based on Model 1 of Table 7, with all other right-hand side variables set equal to their sample average values.