NBER WORKING PAPER SERIES

AGGREGATE IDIOSYNCRATIC VOLATILITY

Geert Bekaert
Robert J. Hodrick
Xiaoyan Zhang

Working Paper 16058
http://www.nber.org/papers/w16058

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2010

We are very grateful to Tim Vogelsang who provided code for the linear time trend test. We have benefited from discussions with and the comments of Lieven Baele, Ian Cooper, Rob Engle, Cheol Eun, Mardi Dungey, Chris Jones, Federico Nardari, Geert Rouwenhorst, Tim Simin and from presentations at the AFA in San Francisco, the EFA in Bergen, Cornell, Rutgers, NYU, the Wharton International Finance Workshop, the Queen's University International Finance Conference, University of Toronto, University of Washington, University of Colorado, Boston College, Purdue, University of Texas at Dallas, Indiana University, University of California at Riverside, University of Houston, and Hong Kong University of Science and Technology. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Aggregate Idiosyncratic Volatility
Geert Bekaert, Robert J. Hodrick, and Xiaoyan Zhang
NBER Working Paper No. 16058
June 2010
JEL No. G12

ABSTRACT

We examine aggregate idiosyncratic volatility in 23 developed equity markets, measured using various methodologies, and we find no evidence of upward trends. Instead, idiosyncratic volatility appears to be well described by a stationary autoregressive process that occasionally switches into a higher-variance regime that has relatively short duration. We also document that idiosyncratic volatility is highly correlated across countries. Finally, we examine the determinants of the time-variation in idiosyncratic volatility. In most specifications, the bulk of idiosyncratic volatility can be explained by a growth opportunity proxy, total (US) market volatility, and in most but not all specifications, the variance premium, a business cycle sensitive risk indicator. Our results have important implications for studies of portfolio diversification, return volatility and contagion.

Geert Bekaert
Graduate School of Business
Columbia University
3022 Broadway, 411 Uris Hall
New York, NY 10027
and NBER
gb241@columbia.edu

Robert J. Hodrick
Graduate School of Business
Columbia University
3022 Broadway
New York, NY 10027
and NBER
rh169@columbia.edu

Xiaoyan Zhang
Johnson Graduate School of Management
366 Sage Hall
Cornell University
Ithaca, NY 14853
xz69@cornell.edu

1. Introduction

Much recent research in finance has focused on idiosyncratic volatility. A rapidly growing literature considers the pricing of idiosyncratic risk[3]. Following the work by Morck, Yeung and Yu (2000), the relative importance of idiosyncratic variance in total variance has been proposed as a measure of market efficiency. The level of idiosyncratic volatility is also an important input in the study of diversification benefits. Here, a growing literature attempts to explain the increase in idiosyncratic volatility first documented by Campbell, Lettau, Malkiel and Xu (2001), CLMX henceforth. Aktas, de Bodt and Cousin (2007) and Kothari and Warner (2004) study how this increase affects the use of one of the most powerful empirical techniques in finance, the event study. Comin and Mulani (2006) examine how and why trends in the macro-economy seem to diverge from the "micro-trend".

Our first contribution is to expand the study of the time-series behavior of aggregate idiosyncratic volatility to international data. This is not only important for purely statistical reasons, but it should also help inform the debate about the determinants of the time-variation in idiosyncratic volatility. Our results are in fact startling: there is no trend in idiosyncratic volatility, not in the U.S. and not in other developed countries.

CLMX's results appear quite robust to alternative methodologies to compute idiosyncratic volatility and to the use of some alternative trend tests. Nevertheless, we show that the implications of the tests are sensitive to the sample period used: ending the sample somewhere around 1997 is key to finding a trend! Of course, when a time series exhibits time trends over part of its sample path, it is likely characterized by near non-stationary-behavior. We show that average idiosyncratic volatility is well described by a relatively stable autoregressive process that occasionally switches into a higher-variance regime that has relatively low duration. We also document a new empirical fact: idiosyncratic volatility is highly correlated across countries, and this correlation has increased over time.

These findings provide a challenge for some of the explanations proposed in the literature for the "trend" in idiosyncratic volatility. Successful determinants must be able to explain the low frequency changes in the idiosyncratic volatility time series data and be correlated across countries. The literature has identified roughly three types of determinants. A first set of articles focuses

[3] See Ang, Hodrick, Xing and Zhang (2006, 2008) and the references therein. We do not address expected return issues here.

on the changing composition of stock market indices. Fink, Fink, Grullon and Weston (2010) ascribe the trend to the increasing propensity of firms to issue public equity at an earlier stage in their life cycle, while Brown and Kapadia (2007) argue that the trending behavior is due to the listings of more riskier firms over the years. The second and largest set of articles, focuses on what we call "corporate variables": firm-specific characteristics that ultimately determine idiosyncratic cash flow variability. These articles include Guo and Savakis (2007) (changes in the investment opportunity set), Cao, Simin and Zhao (2007) (growth options), Comin and Philippon (2005) (research and development spending and access to external financing), and Wei and Zhang (2008) (earnings quality). Gaspar and Massa (2006) and Irvine and Pontiff (2008) point to increasingly competitive product markets as a potential "deeper" explanation of increased idiosyncratic cash flow variability. It is also conceivable that financial development has made stock markets more informative and increased idiosyncratic variability, relative to total market variability (see Chun, Kim, Morck and Yeung 2007). The third set of articles is more "behavioral" in nature and relies on changes in the degree of market inefficiency to generate changes in idiosyncratic variability. Xu and Malkiel (2003) and Bennett, Sias, and Starks (2003) ascribe the rise in idiosyncratic volatility to an increase in institutional ownership, and especially the increased preferences of institutions for small stocks. However, as we do, Brandt, Brav, Graham and Kumar (2010) argue that the increase in idiosyncratic volatility in the 1990s was temporary. They find that the period between 1926 and 1933 exhibited a similar temporary increase in idiosyncratic volatility and ascribe both episodes to "speculative behavior", as evidenced by retail traders in the Internet Bubble.

Our time series characterization of idiosyncratic volatility immediately excludes certain variables as important determinants. For example, because institutional ownership exhibits a clear trend, it cannot fully explain the evidence. However, it is possible that the propensity to issue public equity is not trending upward but also shows regime-switching behavior. The final part of our article runs horse races between the various determinants, in addition to exploring the links between idiosyncratic volatility and market volatility and the business cycle, which have not been studied before. This turns out to be an important omission: together with growth opportunities, market volatility and a cyclical risk aversion indicator appear to drive most of the variation in idiosyncratic volatility, both in the U.S. and internationally.

The rest of the article is organized as follows. Section 2 describes the data. Section 3 contains the main results for trend tests. Section 4 characterizes the time-series properties of aggregate idiosyncratic volatility. Section 5 examines the explanatory power of a large number of potential determinants. In the conclusions, we summarize our findings.

2. Data

2.1 The U.S. Sample

In order to replicate and extend the CLMX study, we first collect daily U.S. stock returns between 1964 and 2008 from CRSP. We calculate excess returns by subtracting the U.S. T-bill rate, which is obtained from the CRSP riskfree file. We calculate the idiosyncratic volatility of a firm's return using two methods. First, we compute the idiosyncratic variance as in CLMX. The model for individual firm j for day t is:

$$R_{j,t} = IND_{J,t} + u_{j,t}^{CLMX}. \tag{1}$$

Here, $IND_{J,t}$ is the return on a corresponding industry portfolio J to which firm j belongs.[4] The firm's idiosyncratic variance is then the variance of the residual $u_{j,t}^{CLMX}$, computed with one month of daily return data. Value-weighting the firm-level idiosyncratic variances produces the CLMX idiosyncratic variance. That is,

$$\sigma^2_{CLMX,m} = \sum_{j=1}^{N} w_{j,m}\ \sigma^2(u_{j,t}^{CLMX}), \tag{2}$$

where day t belongs to month m. Here the weight $w_{j,m}$ is computed using firm j's previous month market capitalization, and N is the number of firms. Implicitly, CLMX assume that systematic risks are captured by the industry return and that firms have unit betas with respect to the industry to which they belong[5].

Bekaert, Hodrick and Zhang (2009), BHZ henceforth, show that the unit beta restrictions in the CLMX approach severely limit the factor model's ability to match stock return comovements. We therefore also consider the Fama-French (1996) model, which fits stock return comovements better:

$$R_{j,t} = b_{0,j,m} + b_{1,j,m}MKT_t + b_{2,j,m}SMB_t + b_{3,j,m}HML_t + u_{j,t}^{FF}, \tag{3}$$

[4] We use 26 industries by merging SIC codes for U.S. firms and FTSE industry codes for foreign firms, as in Bekaert, Hodrick and Zhang (2009).

[5] CLMX also assume that the industry returns have unit betas with respect to the market portfolio, which then leads to a decomposition of total risk into market, industry and idiosyncratic risk, requiring no beta estimation.

where day t belongs to month m. Here, the variable MKT represents the excess return on the market portfolio, SMB is the size factor, and HML is the value factor. This model is more in line with standard methods to correct for systematic risk. Data on the Fama-French factors are obtained from Kenneth French's website. To allow the betas to vary through time, we re-estimate the model every month with daily data. The idiosyncratic variance for firm j is the variance of the residual of the regression, that is, $\sigma^2(u_{j,t}^{FF})$. We again compute the idiosyncratic variance at the country level using value weighting:

$$\sigma_{FF,m}^2 = \sum_{j=1}^{N} w_{j,m}\ \sigma^2(u_{j,t}^{FF}), \tag{4}$$

where day t belongs to month m.

2.2 The Developed Countries Sample

We study daily excess returns for individual firms from 23 developed markets, including the U.S. The sample runs from 1980 to 2008. All returns are U.S. dollar denominated. Our selection of developed countries matches the countries currently in the Morgan Stanley Developed Country Index. Data for the U.S. are from Compustat and CRSP; data for the other countries are from DataStream. In the DataStream data, it is likely that new and small firms are increasingly less under-represented in the sample. This could bias our tests towards finding a trend. We estimate domestic models, such as the CLMX model in equation (1) and the FF model in equation (3), for each developed country, where the industry, size and value factors are constructed in the corresponding national market. In section 3.2, we conduct a robustness check using models which explicitly allow for both global and local factors.

2.3 Summary Statistics

Table 1 presents summary statistics for the time-series of annualized idiosyncratic variances. Panel A focuses on the long U.S. sample where we have 540 monthly observations. The mean of the annualized CLMX idiosyncratic variance is 0.0800 with a time-series standard deviation of 0.0592, and the mean of the annualized Fama-French idiosyncratic variance is 0.0697 with a time-series standard deviation of 0.0484. Hence, the Fama-French risk adjustments lower both the mean and the volatility of the idiosyncratic variance series relative to the CLMX-idiosyncratic variance. The correlation between the two idiosyncratic variance series is nonetheless 98%.

Panel B of Table 1 reports idiosyncratic variance statistics computed for 23 countries, using the CLMX model on the left and the FF model on the right. Among the G7, Japan, the U.S. and Canada have the highest, and Germany and the U.K. the lowest idiosyncratic volatilities. Among the other countries, the idiosyncratic volatility is the highest for Greece at 0.0901 when using the CLMX model, and 0.0798 when using the FF model. The idiosyncratic volatility is the lowest for Switzerland at around 0.03[6].

Panel C of Table 1 presents correlations among the idiosyncratic variances of the G7 countries. No matter which model we use, the idiosyncratic variances are highly correlated across countries. Using Pearson's test, we find that all correlation coefficients are significant at the 5% level. This is an important new fact, as it suggests that there might be a common driving force for idiosyncratic variances across countries.

Figure 1 presents the time-series of the various idiosyncratic variance measures. There are various periods of temporarily higher volatility in the U.S., including 1970, 1974, 1987, a longer-lasting increase in 1998, which seems to reverse after 2003, and the recent crisis period. In other countries, the most obvious high variance periods are again 1998-2001, and the recent crisis period. However, periods of higher volatility are apparent earlier in the sample as well; for instance, around 1987 for a number of countries and in the early 1980s, for France and Italy.

3. Trend Tests

3.1 Main Results

The main result in CLMX is that the average idiosyncratic variance in the U.S. exhibits a positive time trend. To formally test for trends, we use Vogelsang's (1998) linear time trend test as do CLMX. The benchmark model is

$$y_t = b_0 + b_1 t + u_t, \tag{5}$$

where y_t is the variable of interest, and t is a linear time trend. We use the PS1 test in Vogelsang (1998) to test $b_1 = 0$. The conditions on the error terms under which the distributions for the test statistics are derived are quite weak and accommodate most covariance stationary processes

[6] Bartram, Brown, and Stulz (2009) find that a typical U.S. firm has higher idiosyncratic risk than a comparable foreign firm and explore the cross-sectional determinants of this difference.

(such as regime switching models) and even I(1) processes. In all of the ensuing tables, we report the trend coefficient, the t-statistic and the 5% critical value derived in Vogelsang (1998) (for a two-sided test). In addition, Bunzel and Vogelsang (2005) develop a test that retains the good size properties of the PS1 test, but it has better power (both asymptotically and in finite samples). We denote this test with a "dan" subscript, as the test uses a "Daniell kernel" to non-parametrically estimate the error variance needed in the test. In fact, tests based on this kernel maximize power among a wide range of kernels. Vogelsang generously provided us with the code for both the t-ps1 and t-dan tests.

Table 2 reports the trend test results. Panel A presents results using the same U.S. sample as in CLMX, which is 1964-1997. Panel B presents results using the full U.S. sample, 1964-2008. In each panel, we also show the t-dan test based on pre-whitened time-series using an AR(1)-model because Bunzel and Vogelsang (2005) show that pre-whitening improves the finite sample properties of the test. For the sample period 1964-1997, we find a significant trend in the idiosyncratic variance, no matter whether we use the FF or CLMX model. In Panel B, we include 11 more years of data, and the idiosyncratic variance does not display a significant trend for whatever case we consider. Clearly, the trend documented in CLMX is time-period dependent. Since the pre-whitened and non-pre-whitened results are very similar, we only report the pre-whitened results for the t-dan test in later sections.

Panel C of Table 2 reports trend test results for the 23 developed countries, country by country. We fail to detect a significant positive time trend for all countries, either using the t-dan test or the t-ps1 test, and for whatever risk model is used to compute the idiosyncratic variances. France, Italy, Australia, Belgium, Finland, Greece, New Zealand, Portugal and Spain have negative trend coefficients, which are significantly different from zero for Italy and Portugal. In summary, positive trending behavior is simply not visible in idiosyncratic volatility across the developed world.

3.2 Weekly International Data and Alternative Risk Specification

So far, we focused on daily data across different countries, using domestic risk models (CLMX and FF) to arrive at idiosyncratic shocks. One potential drawback of the two models is that they may not adequately capture global risks. In this subsection, we consider an alternative model using both global and local factors. Since the global factors are constructed with data from different countries, and due to the well-known non-synchronous trading problem, we estimate this model using weekly data.

We calculate firm idiosyncratic volatilities according to a modified Fama-French type model that we call WLFF (for Fama-French model with world and local factors), as in BHZ (2009). The model has six factors, a global market factor ($WMKT$), a global size factor ($WSMB$), a global value factor ($WHML$), a local market factor ($LMKT$), a local size factor ($LSMB$) and a local value factor ($LHML$):

$$\begin{aligned} R_{j,t} &= b_{0,j,s} + b_{1,j,s} WMKT_t + b_{2,j,s} WSMB_t + b_{3,j,s} WHML_t \\ &\quad + b_{4,j,s} LMKT_t + b_{5,j,s} LSMB_t + b_{6,j,s} LHML_t + u_{j,t}^{WLFF}. \end{aligned} \tag{6}$$

where week t belongs to a six-month period s. To allow for time-varying betas, the above model is re-estimated every six months with weekly data. The combination of local and global factors with time-varying betas makes the model flexible enough to fit stock market comovements in an environment where the degree of global market integration may change over time. The local factors are in fact regional factors, where we consider three regions: North America, Europe and the Far East. The global market factor, $WMKT$, is calculated as the demeaned value-weighted sum of returns on all stocks. To calculate $WSMB$, we first compute $SMB(k)$ for each country k, which is the difference between the value-weighted returns of the smallest 30% of firms and the largest 30% of firms within country k. Factor $WSMB$ is the demeaned value weighted sum of individual country $SMB(k)$s. Factor $WHML$ is calculated in a similar manner as the demeaned value weighted sum of individual country $HML(k)$s using high versus low book-to-market values. The local factors ($LMKT, LSMB, LHML$) are all orthogonalized relative to the global factors ($WMKT, WSMB, WHML$). BHZ (2009) show that this model fits the comovements between country-industry portfolios and country-style portfolios very well, and it also captures firm level comovements well.

We calculate the idiosyncratic variance for stock j as the variance of the residual of the regression, that is, $\sigma^2(u_{j,t}^{WLFF})$, and we then aggregate to the country level:

$$\sigma^2_{WLFF,s} = \sum_{j=1}^{N} w_{j,s}\ \sigma^2(u_{j,t}^{WLFF}), \tag{7}$$

where week t belongs to the six-month period s. The weight $w_{j,s}$ is computed from firm j's relative market capitalization at the end of the last six-month period, and N represents the number of firms within one country.

Panel A of Table 3 report trend test results for the 23 developed countries. We fail to detect a significant time trend for any country, using either the t-dan test or the t-ps1 test.

3.3 Equal Weighting

In this subsection, we examine the time-series behavior of equally weighted idiosyncratic variances. The dominance of small firms in the equally weighted idiosyncratic variances may cause the results to change. For example, one of the reasons suggested for the trend in aggregate idiosyncratic variance is that small firms may have sought public funding at an earlier stage in their life cycle than before (see Fink, Fink, Grullon and Weston 2010).

The results are presented in Panels B and C of Table 3, where we focus on the U.S. idiosyncratic variance over 1964 - 2008, computed from daily data. Since the results for the other developed countries are very similar, we do not report those to save space. In Panel B, the time-series mean of the equal-weighted CLMX (FF) idiosyncratic variance is 0.4308 (0.3530), which is much larger than its value-weighted counterpart of 0.0800 (0.0697). Obviously, the returns of smaller firms are much more volatile.

In Panel C, we report the Vogelsang trend test results. Interestingly, the equally weighted idiosyncratic variance time series shows a larger trend coefficient than the value-weighted time series, but the coefficient is now insignificantly different from zero for all cases, even for the 1964-1997 period. This, in fact, confirms the results in CLMX. Equally weighted idiosyncratic variances are too noisy to allow strong statistical inference.

4. Characterizing the Dynamics of Idiosyncratic Volatility

The results in Section 3 strongly reject the presence of a gradual and permanent increase in idiosyncratic variances, as captured by a deterministic time trend. Other forms of non-stationary behavior remain a possibility, however. We first examine the presence of stochastic trends. Using the Dickey and Fuller (1979) and Phillips and Perron (1988) tests, we invariably reject the null of a unit root, consistent with the evidence in Guo and Savickas (2008). We also examine models with structural breaks, adopting the methodology in Bai and Perron (1998). For all countries, we identify a relatively large number of breaks, with the break dates highly correlated across countries. In particular, the tests consistently reveal the end of 1997/1998 and 2001/2002 as break dates, thus selecting a temporary period of higher idiosyncratic volatility associated with what many economists have called the Internet or Tech Bubble. Generally, the "break tests" identify periods of temporary higher volatility that may occur more than once during the sample period. A better model to capture such behavior is a regime-switching model.

4.1 Country Specific Regime Switching Model

4.1.1 The Model

Let y_t represent the original idiosyncratic variance. Following Hamilton (1994), we allow y_t to follow an AR(1) model where all parameters can take on one of two values, depending on the realization of a discrete regime variable, s_t. The regime variable follows a Markov Chain with constant transition probabilities. Let the current regime be indexed by i.

$$y_t = (1 - b_i)\mu_i + b_i y_{t-1} + \sigma_i e_t,\ i \in \{1, 2\}. \tag{8}$$

with e_t is $N(0, 1)$. In estimation, we force regime 1 (2) to be the lower (higher) idiosyncratic variance regime, and the mean levels of idiosyncratic variances in both regimes to be non-negative, that is, we constrain $\mu_2 > \mu_1 > 0$.

The transition probability matrix, Φ, is 2x2, where each probability represents $P[s_t = i|s_{t-1} = j]$, with $i, j \in \{1, 2\}$:

$$\Phi = \begin{pmatrix} p11 & 1 - p11 \\ 1 - p22 & p22 \end{pmatrix}, \tag{9}$$

The model is parsimonious, featuring only 8 parameters, $\{\mu_1, \mu_2, b_1, b_2, \sigma_1, \sigma_2, p11, p22\}$.

4.1.2 Estimation Results

In Panel A of Table 4, we report the estimation results for both σ^2_{CLMX} and σ^2_{FF} for the long U.S. sample. The standard errors are computed using the robust White (1980) covariance matrix. The annualized idiosyncratic variance level for regime 1, μ_1, is 0.062 for σ^2_{CLMX}, and 0.055 for σ^2_{FF}, but the level increases dramatically for regime 2, with μ_2 equal to 0.181 for σ^2_{CLMX} and 0.155 for σ^2_{FF}. Using a Wald test, the level differences between the two regimes are highly statistically significant. It seems likely that a regime with high mean volatility also has high innovation volatility, and that is indeed what we find. Regime 2 has much higher volatility than regime 1, as σ_1 is 0.011 but σ_2 is 0.082 for σ^2_{CLMX}, with similar results when we use σ^2_{FF}. It is also typical for a high-variance regime to show more mean-reverting behavior, and we also find this to be the case for the point estimates for both σ^2_{CLMX} and σ^2_{FF}. The difference between the two autocorrelation coefficients is statistically significant.

Figure 2 presents time-series of the smoothed probabilities of being in regime 2. The smoothed probability of being in regime 2 at time t is computed using information from the whole time-series up to time T, that is, $P[s_t = 2|y_T, ..., y_1]$. As can be expected from the parameter estimates, the high-variance regime is a short-lived regime. However, it does occur several times over the sample period with some consistency over the two risk models. High variance episodes that occur in both cases include 1970, 1974, 1987, 1996, 1998-2002 and 2007-2008. If we define y_t to be in regime 2 if the probability of being in regime 2 is higher than 0.5, and vice versa for regime 1, then there are 13(11) regime switches over the 45 year sample for σ^2_{CLMX} (σ^2_{FF}), and y_t is in regime 2 14% of the time. On average, regime 2 lasts about 10 to 11 months[7].

It is not difficult to give some economic content to the regimes. The shaded areas in Figure 2 are NBER recessions. Clearly, the high-level idiosyncratic variance regimes mostly coincide with periods of recessions, although recessions are neither necessary nor sufficient to have a high volatility regime. It is well-known that in recession periods market volatility tends to be high as well (see Schwert, 1989). We also find that our high idiosyncratic variance regimes coincide with market volatility being about twice as high as in normal regimes. We come back to this finding in a later section. The link between high idiosyncratic variance regimes and recessions appears stronger for σ^2_{CLMX} than for σ^2_{FF}.

We report the results for the shorter sample period of G7 countries in Panel B of Table 4. The levels of idiosyncratic variances differ across countries, but not dramatically. In the low variance regime, the means vary between 0.033 (Germany) and 0.072 (Japan). For the high variance regime, the means vary between 0.121 (Germany) and 0.198 (U.S.). The persistence parameters are mostly lower in regime 2, but not significantly in Japan, the U.K. and the U.S. The corresponding time-series of smoothed probabilities in regime 2 (high volatility regime) using the CLMX model are

[7]When estimating a RS model for the equally weighted aggregate idiosyncratic variance, the regime identification is mostly similar to the value-weighted case, but with stronger persistence in the low-variance regime. The equally weighted idiosyncratic variances move into regime 2 from 1998 to 2001, similar to what happens in Figure 2, where we use value-weighting. However, the equally weighted idiosyncratic variances stay in regime 2 much longer than their value-weighted counterparts, returning only to regime 1 during the second half of 2003. Other shifts into regime 2 occur in 1970, 1974, 1987, 1988, and the early 1990s.

presented in Figure 3. The results using the FF model are very similar, and we do not report them to save space. The idiosyncratic variances of all 6 countries are mostly in regime 2 around 1997 to 2002 and back again in the recent crisis period. They are also likely to be in regime 2 around the 1987 crash. Most countries experience additional transitions into the higher variance regime.

4.1.3 Regime Switches and the CLMX Results

These results help us interpret CLMX's trend findings. Essentially, the idiosyncratic variance process does not exhibit a trend but exhibits covariance stationary behavior with regime switching. Of course, trend tests, despite having good finite sample properties, may perform much worse in an environment where we start the sample in a low level regime and end the sample in a high level regime. That is exactly what happened in the case of CLMX's analysis.

The CLMX sample started during a "normal" idiosyncratic variance regime in the 1960's and ended in 1997. While 1997 itself is not classified as a high variance regime, it is in the middle of a period with frequent shifts into the high variance regime. As Figure 2 shows, the probability that σ^2_{CLMX} is in the high-level, high-variance regime increased briefly around October 1987[8], increases slightly several times in the following years, before increasing substantially but briefly in June and July of 1996. In April 1998, a longer lasting high variance regime starts. Conditioning on such a sample selection, a trend test may be more likely to reject than the asymptotic size of the test indicates.

To see the effect more concretely, Figure 4 shows the values of the t-dan test recursively, starting the sample in 1964:01 but varying the end point between 1970:04 and the end of the sample (2008:12). The date 1970:04 is not chosen arbitrarily, it is in the first high variance regime selected by the regime switching model, and it is striking that the trend test would have rejected when dates around that time were chosen as sample end points. The trigger date for finding an upward trend was the crash of October 1987, and an upward trend would have been found all the way until 2000:4. Using the less powerful t-ps1 test, actually employed by CLMX, this period of "false rejections" would have lasted about two years less long. The "fake trend" experiment has two important implications: first, the level of idiosyncratic volatility has been high over the 1990s; second, if the time-series starts in a low volatility period, and ends in a high volatility period, the trend test tends

[8] Schwert (1990) shows how stock market volatility, during and after the crash, was very unusual but returned to normal levels relatively quickly.

to be significant, even though the time-series follows an overall stationary process. The recent data reinforce the regime-switching nature of the idiosyncratic variance process. By 2004, the level of idiosyncratic variance had dropped back to pre 1970 levels, only to rise starkly (and most likely temporarily) in the current crisis that started in 2008.

The regime switching process identifies periods of unusually high and volatile idiosyncratic volatility that are not likely associated with gradual fundamental increases in idiosyncratic volatility, but may nonetheless bias trend statistics. It is therefore also of interest to check whether there is a trend in idiosyncratic volatility during the "normal" periods, that is, excluding the high idiosyncratic volatility periods identified by the regime-switching model. When we do so, we still fail to find evidence for a trend, but in the case of the t-ps1 test, the t-stat is relatively high for the CLMX risk model. Consequently, there is some very weak evidence of long-run increases in idiosyncratic volatility.

4.1.4 Specification Tests

It is important to verify that a RS model indeed fits the data well, and that it fits the data better than simpler alternative models. We conduct a number of specification tests on the residuals of various RS models we estimated. We also report tests for two alternative benchmark models: an AR(1) model in levels with Gaussian shocks, and an AR(1) model with a GARCH(1,1) volatility process. Our tests examine 6 moment conditions: the mean and one auto-correlation of the residuals; the variance and one auto-correlation of the squared residuals; and the third and fourth order moments. Appendix A details the tests. While we use asymptotic critical values, it is quite likely that our tests over-reject in small samples. This is particularly true if the data are actually generated from a non-linear RS model (see e.g. Baele, Bekaert and Inghelbrecht (2010)).

Table 5 reports the results. The three panels investigate, respectively, the mean and variance specification; the higher moments conditions; and finally in Panel C, a joint test. For the RS models, we use smoothed ex-post probabilities to infer residuals and model moments. We present both the long sample for U.S. for the two risk models and the short sample for all countries. Focusing on the joint test first, the regime switching model clearly outperforms the two other models. Over the 9 tests, there is not a single 1% rejection, and three 5% rejections (for Canada, Germany and the UK). The AR model on the other hand is always rejected at the 1% level, whereas the GARCH model features only two cases for which it is not rejected at the 5% level (short sample US, and long sample US when the FF model is used). The sources of the rejections differ across countries, and in some cases the joint test simply adds power to the two sub-tests.

In unreported work, we also apply the three models to the logarithm of the variance. Such a model keeps the variance everywhere positive and the non-linear transformation may sufficiently reduce the outliers in the data to make the idiosyncratic variance process more amenable to linear modeling. However, none of the models performs better in logarithm forms, so that we restrict further analysis to the untransformed variances.

4.2 Commonality in Idiosyncratic Volatilities

One new empirical fact that we uncovered deserves further scrutiny. Idiosyncratic volatility has a large common component across countries. This fact may have implications for the analysis of issues such as international diversification and contagion. It should also be explored in theoretical and empirical work that examines why we have different idiosyncratic volatility regimes.

Table 6 provides more information about this phenomenon. We report the correlations of the aggregate idiosyncratic variances in the G7 countries with respect to the aggregate U.S. idiosyncratic variance. Because a missing common risk factor is a potential explanation of this phenomenon, we show results for our two risk models and one additional method to compute idiosyncratic volatility. In particular, we also use weekly returns over six month intervals and the international WLFF factor model introduced above. BHZ (2009) show that the idiosyncratic return correlations of country portfolios, computed using the WLFF model, are essentially zero.

Across the panels, the correlations with the U.S. vary between 0.20 (Italy) and 0.81 (UK) for the monthly time-series, and between 0.15 (Italy) and 0.87 (Canada) for the half-year time-series. Table 6 also shows these correlations over the first and second halves of the sample. The increase in the correlation with the U.S. over time is remarkable.

As a simpler summary statistic, we also compute the equally weighted correlation between the idiosyncratic variances of the G7 countries. Using the weekly WLFF model to compute idiosyncratic variances, it is 57% over the whole sample; 24% over the 1980-1994 period but 75% over the 1995-2008 period. The magnitudes are qualitatively robust to the use of other methods to obtain idiosyncratic volatilities, or to the use of daily data, instead of weekly data. Hence, while our article explained away an existing puzzle in the literature about idiosyncratic volatility, we may have well introduced a novel one.

It is possible that the phenomenon is also related to the regime switching behavior of idiosyncratic variances. In the last two columns of the table, we report the bivariate correlations with the US, conditioning on the idiosyncratic variance in the US being either in the low level/low variance or high level/high variance regime. Perhaps not surprisingly, we find that the correlations are generally higher in the high-level/high variance regime[9]. In the remainder of the article, we focus on understanding more fundamental determinants of idiosyncratic variances, and investigate whether they can capture the RS behavior and commonality just documented.

5. Determinants of Idiosyncratic Volatility Dynamics

Idiosyncratic variances follow a stationary autoregressive process, characterized by relatively low frequency changes in regime, when they temporarily become higher, more variable and more mean-reverting. These patterns are apparent in all countries. Moreover, there is a strong common component in idiosyncratic variances across countries that has increased in importance over time. These facts have significant implications for the rapidly growing literature trying to explain the time variation in idiosyncratic volatility. In particular, it does not bode well for studies that have focused on trying to fit trends in the series. In this section, we attempt to determine which prevailing explanations best fit the time series movements of idiosyncratic variances in the U.S. and in other countries. Because we have more data available in the US, we start there in section 5.1. Table 7 lists all the independent variables we use in the analysis and the acronyms we assign to them.

We distinguish three different types of variables: variables affecting changes in the index composition, "corporate" variables correlated with cash flow volatility, and finally, business cycle variables and market wide volatility, a category new to this literature. The first three sub-sections in 5.1 discuss these three groups of variables in more detail. Section 5.1.4 runs a statistical horse race to determine which variables best capture the time series variation in aggregate US idiosyncratic volatility. This analysis employs two alternative model reduction techniques, which we describe below. Section 5.1.5 assesses whether accounting for these determinants leads to residuals that are well-behaved and no longer exhibit regime switching behavior. In Panel B of Table 7, we list the

[9] In a previous version of the article, we actually estimated a joint regime-switching model over the G7 countries, where the US regime variable functions as the standard regime variable and the regime variables in other countries depend on the U.S. variable. Results are available upon request. The joint model shows that when the U.S. is in the high level/high variance regime, the other G7 countries are more likely to be in this regime as well, and vice versa.

variables that are available internationally, which is a subset of the variables available for the US. Section 5.2 then conducts an analysis of the determinants of aggregate idiosyncratic volatility in the G7 countries. Given the more limited nature of our data, this analysis should be viewed as a preliminary first look at the data.

5.1 Analysis of U.S. Data

Most of the literature has focused on U.S. data. Our general approach is to regress the idiosyncratic variance time series on a set of explanatory variables, mostly constructed exactly as in the extant literature. The reported standard errors are heteroskedasticity consistent and always allow for 12 Newey-West (1987) lags.

5.1.1 Index Composition and Behavioral Variables

One possible explanation offered in the literature for a potential increase in idiosyncratic variances over time is that the composition of the index has changed towards younger, more volatile firms. Fink, Fink, Grullon and Weston (2010) show that the age of the typical firm at its IPO date has fallen dramatically from nearly 40 years old in the early 1960s to less than 5 years old by the late 1990s. Since younger firms tend to be more volatile, this systematic decline in the average age of IPOs, combined with the increasing number of firms going public over the last 30 years, may have caused a significant increase in idiosyncratic risk. Brown and Kapadia (2007) also ascribe increases in firm-specific risk to new listings by riskier companies (although not necessarily solely related to age), whereas Xu and Malkiel (2003) argue that the increase can be partly attributed to the increasing prominence of the NASDAQ market. We proxy for the age effect using the "pyoung" variable, the percentage of market cap of firms which are less than 10 years old since foundation.

A related possibility is that small capitalization stocks, which tend to have higher idiosyncratic volatilities, have become relatively more important (see Bennett, Sias and Starks, 2003). It is possible such a trend is a fundamental response to markets becoming more efficient over time, making it possible for smaller firms to list and be priced efficiently. Bennett, Sias and Starks (2003) explicitly ascribe the trend to institutional investors becoming more actively interested in small capitalization stocks over time, which could increase trading in these stocks and make these markets more liquid and consequently more efficient, thereby providing higher valuations. Xu and Malkiel (2003) also argue that an increase in institutional ownership is associated with higher idiosyncratic volatility. To measure the effect of the relative importance of small capitalization firms on idiosyncratic variability, we use the variable "psmall", the proportion of market capitalization, represented by the smallest 25% of all listed firms.

Such an explanation is hard to distinguish from certain "behavioral explanations". Brandt, Brav, Graham and Kumar (2010) ascribe episodic shifts in idiosyncratic volatility to speculative behavior. While difficult to measure, they claim that the 1990s episode of high and increasing idiosyncratic volatility is concentrated primarily in firms with low stock prices and limited institutional ownership. Their explanation is hence quite different, and almost contradictory to the arguments made by Bennett, Sias and Starks (2003) and Xu and Malkiel (2003), but it nonetheless also gives a primary role to small stocks. We use two variables to imperfectly measure the "speculative trading" channel. The variable "plow" measures the market cap percentage of firms with a stock price lower than \$5 and the variable "lowto" represents the average turnover of firms with a stock price lower than \$5.

More generally, retail investors can potentially act as noise traders and increase trading volume and volatility (see also Foucault, Sraer and Thesmar 2008). We also use a general measure of turnover, computed as total dollar volume over total market capitalization. A positive effect of turnover on volatility could also reflect increased turnover indicating a more developed, more efficient stock market, which in turn may be associated with higher idiosyncratic variability, see Jin and Myers (2006).

The first column on the left hand side of Table 8 runs a regression of our CLMX measure of idiosyncratic variances onto the 5 variables described above. The results are surprising. The fraction of young firms is positively associated with aggregate idiosyncratic variability, but the effect is not statistically significant. Both the proportions of small stocks and low priced stocks are negatively associated with aggregate idiosyncratic risk; with the effects significant at 10% level. The turnover in low priced stocks is negatively associated with idiosyncratic risk, yet overall turnover is positively associated with idiosyncratic risk. Neither effect is significant. In summary, the variables that provide some marginal explanatory power have the wrong sign. The R^2 of the regression is 35%.

Introducing institutional ownership would help to distinguish the Bennett, Sias and Starks (2003) story from Brandt et al.'s (2010). Unfortunately, the fraction of shares owned by institutional investors is only available from 1981 onwards, eliminating 17 years from the sample period. Over this shorter sample period, institutional ownership is univariately positively related with idiosyncratic variability, but the coefficient is insignificantly different from zero. When we run a regression including our five other variables, the coefficient on institutional ownership becomes significantly negative, which is not consistent with the Bennett, Sias and Starks hypothesis. The results are somewhat hard to interpret, because institutional ownership is quite highly correlated with the 5 variables included in our analysis. In fact, these five variables explain 77% of the variation in institutional ownership. The fact that institutional ownership shows a clear upward trend also implies that it cannot really be a major factor driving the time-series variation in idiosyncratic variability.

5.1.2. Corporate variables

Another part of the literature essentially argues that the movements in idiosyncratic variances reflect fundamental (idiosyncratic) cash flow variability. The various articles differ in how they measure this fundamental variability, and how they interpret their results. Details about how to construct those variables can be found in Appendix B.

Wei and Zhang (2006) claim that the upward trend in average stock return volatility is fully accounted for by a downward trend in return-on-equity, indicating poorer earnings quality, and a upward trend in the volatility of return-on-equity. To mimic their results, we create three empirical measures, the variable "vwroe" is the value-weighted average of the firm level return on equity; the variable "vwvroe" is the value-weighted firm level time-series variance of the return on equity (computed using the past 12 quarters of data), and the variable "cvroe" is the cross-sectional variance of the return on equity at each point in time.

Irvine and Pontiff (2008) attribute the increases in idiosyncratic return volatility to an increase in the idiosyncratic volatility of fundamental cash flows. We mimic their procedure to construct idiosyncratic cash flow volatility. First, they use a pooled AR(3) model for firms' earnings per share to create earnings innovations. Then, they take the cross-sectional variance of these innovations. This is the variable "veps". Both Irvine and Pontiff (2008) and Gaspar and Massa (2006) ascribe increases in fundamental idiosyncratic variability to more intense economy wide product competition. To proxy for "competition", we use a measure of industry turnover, denoted "indto". We first compute the percentage of market cap of firms entering and exiting the same industry at the industry level each month. Then this percentage is assigned to individual firms in the various industries. The variable "indto" is the value-weighted average of firm level industry turnover.

Cao, Simin and Zhao (2007) show that both the level and variance of corporate growth options are significantly related to idiosyncratic volatility. We therefore use two variables: the value-weighted firm level "maba" (market value of assets over book value of assets), as a proxy for growth options, and "vmaba", the value-weighted variance of the firm level's maba, computed using data over the past three years. Finally, following the spirit of "cvroe", we also compute "cvmaba", the cross-sectional variance of maba at each point in time.

Finally, Chun, Kim, Morck and Yeung (2008) argue that a more intensive use of information technology and faster production growth created a wave of "creative destruction", leading to higher idiosyncratic volatility. Chun et al. (2008) and Comin and Philippon (2008) therefore link idiosyncratic volatility to research intensity and spending. Following Comin and Mulani (2006), for each firm, we first compute its fiscal year R&D expenditure divided by the quarter's total revenue. We then construct two R&D related variables: the variable "rd" is the value-weighted average of the scaled R&D expenditure, and the variable "cvrd" is the cross-sectional variance of firm level scaled R&D expenditure.

We use the variables from the extant literature described above in our empirical analysis, without any modification. We do want to point out that, in our opinion, to explain aggregate idiosyncratic variability, the measures should likely also be "idiosyncratic," which is really only true for the measures created by Pontiff and Irvine (2008), who essentially apply CLMX's methodology to earnings per share innovations.

The second column of Table 8 reports results for a regression of our CLMX measure of idiosyncratic variances onto the 10 corporate variables we described above. The total R^2 is 56%, so these variables explain more of the time-series variation in aggregate idiosyncratic variances than the index variables did. We expect positive coefficients for all of these variables. Of course, many of them are highly correlated, causing some multicollinearity. Of the return on equity variables, only the level is significant. Both earnings per share variability and industry turnover are significant. In addition, the growth option variable, maba, and both research spending variables are significant, but the volatility of R&D has a surprisingly negative effect on idiosyncratic volatility.

5.1.3 Business Cycles and Market Wide Volatility

In this section, we examine a number of potential determinants that the extant literature has not yet considered, namely business cycle variables and aggregate market volatility. There are a variety of channels through which business cycle variables could affect aggregate idiosyncratic variability. A first channel is simply that recessions are associated with increases in macro-economic uncertainty, which in turn drives up both systematic and idiosyncratic risk. In principle, the corporate variables we have used so far should pick up this effect, but it is possible they do not, or do so imperfectly.

Another possibility is that there is discount rate volatility that somehow was missed in the systematic factor measurement and causes idiosyncratic volatility fluctuations that even may be correlated across countries. This could be a missing risk factor or it could be because the functional form of systematic risk measurement is incorrect (e.g. the true factor model is really non-linear). For example, if aggregate stock return predictability reflects variation in discount rates, the evidence in Henkel, Martin and Nardari (2008) suggests it is concentrated in certain periods, particularly recessions.

To analyze business cycle effects, we use 6 variables. The first variable is the so-called variance premium, denoted by "mvp", which is the difference between the square of the VIX index, an option based measure of the expected volatility in the stock market, and the actual physical expected variance of stock returns. We take this measure from Baele, Bekaert and Inghelbrecht (2009), who show that it is an important determinant of stock market volatility. Ang, Hodrick, Xing and Zhang (2006) argue that the conditional variance of the market should be a priced cross-sectional risk factor. Bollerslev, Tauchen and Zhou (2009) show that mvp is an important predictor of stock returns, and hence constitutes an aggregate discount rate factor. Theoretically, a variance premium can arise through stochastic risk aversion and nonlinearities in fundamentals (see Bekaert and Engstrom (2009), and Drechschler and Yaron (2008)) or through Knightian uncertainty (see Drechschler (2009)). Next, the regime switching model unearthed an important link between market volatility ("mkttv") and idiosyncratic volatility. We use this variable as the second business cycle variable in the regression. Note that idiosyncratic and aggregate volatility need not be automatically correlated, as long as the index is sufficiently diversified[10]. We also include a growth in industrial

[10] We actually checked the source of correlations between aggregate total and aggregate idiosyncratic variance for the G7 countries, by splitting up the aggregate total variance in two components, a variance component (which should

production variable, computed as industrial production minus its own two year moving average, which we denote by "dip". The variables "term" and "def" are market-driven indicators of business cycle conditions, representing the term spread and default spread respectively. There is an old literature suggesting default spreads predict economic activity (See e.g. Harvey (1988)); more recently the links between default spreads and future economic activity have been explored, by, e.g. Mueller (2009) and Gilchrist, Yankov and Zakrejsek (2009). Finally, we also use a survey measure of consumer confidence, from the conference board, denoted "confi".

When we regress aggregate idiosyncratic variability onto these 6 variables, we explain more than 57% of the total variation. Thus, business cycle variables are slightly more important than cash flow variables, and much more important than compositional variables. The significant variables include the variance premium, the market variance, industrial production (albeit marginally), and the default spread. However, the sign of the default spread is surprisingly negative.

5.1.4 What drives idiosyncratic variability?

We now run horse races between the various determinants. Unfortunately, a regression model using all variables together would be plagued by extreme multi-collinearity and would include many useless and insignificant variables. We use two methodologies to pare down the regression model: subgroup regression and a stepwise regression approach motivated by the work of Hendry and Krolzig (2001).

Subgroup Regressions We first take all the 10% significant variables from the regressions using the three different groups and run a joint regression. This regression has 13 regressors. We eliminate from this regression variables yielding coefficients that are not significant at the 10% level and re-run the regression once more. The final result is reported in the fourth section of Table 8, Panel A. We find that with just 7 independent variables, we explain 80% of the variation in the aggregate idiosyncratic variance. No compositional variables survive, but four cash flow variables and three business cycle variables do survive. The cash flow variables are industry turnover and the growth option variable, the variables stressed by, respectively, Irvine and Pontiff (2008) and Cao, Simin and Zhao (2009); and the research and development spending variables stressed by Chun et al. (2000)

converge to zero in well-diversified indices as the number of firms get large) and a systematic component (which only depends on covariances). We find that the bulk of the correlation is accounted for by the systematic component, with the lowest proportions being 88% for Germany and Italy.

and Comin and Philippon (2005). The business cycle variables are the total market variance, the growth in industrial production and the default spread, with the default spread now having the correct sign, indicating higher idiosyncratic variability when credit conditions are bad.

To gauge the relative importance of the various variables in explaining the time variation in idiosyncratic variances, the last column reports a simple covariance decomposition of the fitted value of the regression. Let independent variable x_{it} have a regression coefficient of $\widehat{b_i}$, and denote the fitted value of the regression by $\widehat{y_t}$. Then, for each variable, we report the sample analogue of the ratio $\frac{cov(\widehat{b_i}x_{it},\widehat{y_t})}{var(\widehat{y_t})}$. These ratios add up to one by construction. Clearly, the most important variables are the growth option variable, maba, and the market wide volatility. Research and development expenditure also explains a non-negligible part of the variation of aggregate idiosyncratic variability.

When applied to the idiosyncratic variances, computed using the FF model, the final "subgroup" model is similar but includes four more variables: vwroe, veps, cvmaba and mvp. Qualitatively, the results for the FF model are largely similar, with maba, mkttv and rd accounting for most of the variation in idiosyncratic volatility (not reported).

Hendry Regressions Recent research on model reduction techniques by Hendry and Krolzig (2001), among many others, suggests that starting from the most general model may yield better specified parsimonious models. While not literally applying the PCGets ("general-to-specific") system, proposed and commercialized by Hendry and his associates, our second model reduction technique is quite close in spirit to it. We first run a regression using all possible regressors (21 in total). We then verify the joint significance of all the variables that are not significant at the 10% level. The joint test uses also a 10% significance level. If the joint test fails to reject that a set of variables is jointly insignificant, we eliminate these variables from the regression and then run one final regression with the remaining variables. However, if the set of variables is jointly significant, we increase the significance level by 5% for both the individual and joint tests. The results are reported in Panel B of Table 8. We end up using a 10% significance level.

The model is less parsimonious than the one previously considered, as it retains 14 variables, all significant at the 1% level except for industry turnover, which is significant at 5% level. After eliminating the useless variables, the R^2 remains unchanged at 86%, and the coefficients of the retained variables remain similar to what they were in the full model. Interestingly, the signs for the compositional variables "psmall" (the relative importance of small firms) and "pyoung" (the relative importance of young firms) are now as expected. However, psmall contributes a negative 13% to the explained variation, whereas psmall's contribution is about 12%. Together, they explain nothing. General turnover does come in significantly and explains about 10% of the total explained variation. The four cash flow variables retained in Panel A survive here too with about the same economic and statistical significance. Two additional cash flow variables are retained as well (earnings variability, veps, and the cross-sectional variance of the return on equity, cvroe). While veps explains 7% of the variation in idiosyncratic variances, cvroe's contribution is a negative 11%. Maba (growth options) remains very important with a 44% contribution to the overall variance of the fitted value. As to the business cycle variables, the variance premium, the term spread and the confidence index are the additional business variables in the final model. The term spread and confidence index are not economically important, but the variance premium accounts for 10% of the explained variation. Its coefficient is in line with expectations. The default spread is no longer significant, but industrial production and the total market variance still are. They have similar coefficients as they do in the model reported in Panel A, and similar economic significance as well, with the market variance now contributing 32% of the total explained variation.

Our results shed new light on the debate regarding the determinants of the time-variation in US aggregate idiosyncratic volatility. We find that compositional and behavioral variables are relatively unimportant, failing to survive in our first multivariate model, and barely accounting for 10% of the total in the second model. Cash flow variables account for 55-60% of the explained variation, leaving a significant part of the variation to business cycle variables. The addition of these latter variables, not examined before, helps increase the explained variation to over 80%.

The Role of Business Cycle Variables Why are these business cycle variables so important? There are two main possibilities. First, the business cycle variables may reflect cash flow variability not accounted for by our cash flow variables. Second, they may reflect discount rate variation not accounted for in our factor model. Let's start with the cash flow channel. One possibility here is related to the counter-cyclical nature of idiosyncratic volatility that we uncovered analyzing the regime switching behavior of aggregate idiosyncratic volatility. When regressing the aggregate idiosyncratic variance time series onto a NBER recession indicator, we obtain a highly significant positive coefficient. It is possible that true cash flow variability is counter-cyclical, but that measurement error in our variables implies that business cycle variables capture this countercyclicality better. The evidence seems largely inconsistent with this interpretation, as most of our cash flow variables either show no significant relation with the NBER recession dummy variable, or a significantly negative relation (e.g. all return on equity variables).

In a more direct analysis, we orthogonalize the cash flow variables with respect to all of our business cycle variables and repeat the regression of Table 8, Panel A, regressing idiosyncratic variability on "pure" cash flow variables. The R^2 drops from 56% to 25%, but the coefficients on some of the most important variables, including maba, hardly change. The coefficients on the return on equity variables and veps do become significantly smaller in absolute magnitude. When we reverse the exercise the R^2 of the business cycle variables also gets cut in half, but importantly the coefficient on the total market variance hardly changes. These results (available upon request) suggest that some of the explanatory power of the business cycle variables may run through cash flow variability and is related to the corporate variables we use, but a significant part is totally independent of it.

One obvious candidate for the independent explanatory power of business cycle variables is the presence of discount rate variation not accounted for in our factor model. To assess the validity of this interpretation, we replace our business cycle variables by a risk premium proxy extracted from these variables. Specifically, we run a regression of market excess returns at time $t+1$ onto the 6 variables at time t. The fitted value of this regression is an estimate of the risk premium on the market, which naturally varies through time. We then repeat our explanatory analysis of aggregate idiosyncratic variance, but we replace the business cycle variables by this risk premium proxy. Consequently, in this regression the business cycle variables only enter to the extent they can predict market excess returns. Despite using only one business cycle variable, the explanatory power of the regression, pared down using the Hendry approach, only drops from 86% to 69%. The risk premium variable is highly significant, and accounts for 26% of the explained variation. This suggests that more than half of the explanatory power of the business cycle variables is related to these variables capturing discount rate variation not accounted for by standard risk models. The current literature on return predictability suggests that most of the predictable variation is concentrated in recession periods (see e.g. Henkel, Martin and Nardari (2009), and it is this time-variation in discount rates that standard models of risk may not quite capture, leading to common risk factors contaminating estimates of idiosyncratic variance.

Our analysis is related but more comprehensive than the recent work by Zhang (2009). Zhang also runs a horse race between explanatory variables for aggregate US idiosyncratic variability, but the set of variables he uses is much more limited. He uses the return on equity and "maba", but he uses one variable we do not use: institutional ownership (see our discussion above and note that this variable is highly correlated with our "dto" variable). We do not use institutional ownership because our sample starts in 1964 and the ownership data are only available from 1980 onwards. Zhang (2009) finds the "fundamentals" variables to be the more robust determinants of idiosyncratic variability. He also notes that there is a trend upward in idiosyncratic volatility from 1980 till 2000, and a trend downward after 2000. In his empirical analysis, he allows for different coefficients in the two periods and finds some evidence in favor of coefficient changes. We examine the possibility of shifts in the relationship between idiosyncratic variability and its determinants in more detail in the next sub-section.

5.1.5 Regression Fit

In this section, we assess how well the regression model fits the time series dynamics of the aggregate idiosyncratic time series. We start by repeating the specification tests we applied to the residuals of various statistical models in Section 3. When we consider all 6 moments, reported on the first line of Table 9, Panel A, the specification test rejects both regression models at the 1% level. From results not reported, we find that the rejection is mainly due to the autocorrelations in the residuals rather than skewness or kurtosis. Consequently, the residuals still exhibit serial correlation and some non-normal behavior.

To explore this further, we estimate a regime-switching model for the regression residuals, using the exact same specification as in Section 3. However, because the residuals ought to have mean zero, we identify the two regimes by their variability rather than their mean. In Panel B of Table 9, we report parameter estimates as well as the results of a Wald test for equal means in the last row. If the mean level of residuals is systematically different across regimes, the model has failed to really capture the RS behavior. From the Wald test in the last row, the two means are not significantly different. We also report the autocorrelation coefficients, which should really be close to zero. Consistent with the results in Panel A, there seems to be plenty of autocorrelation in the regression residuals. There is, of course, still some regime switching behavior left in the variance, and the variances in two regimes are significantly different. Panel A also shows the specification tests applied to the residuals of the RS model. Using ex post probabilities, we fail to reject at the

5% level.

Finally, we investigate whether the coefficients in the regression models vary with the regime. We therefore let each coefficient (including the intercept) in the two final models depend on a regime 2 dummy variable, which takes the value of 1 if the smoothed probability of being in regime 2 is higher than 0.5 and 0 otherwise. We report the results in Panel C of Table 9. Across the two regressions, roughly half of the dummy coefficients is significantly different from zero. When significant, the coefficients mostly become larger in magnitude in regime 2. This is true for the two most important determinants, namely maba and aggregate market volatility. Overall, allowing for this non-linearity makes the fit nearly perfect (an R2 of 94%) in the case of the Hendry model. It is conceivable that this non-linear dependence reflects a "crisis" effect, where in time of turmoil all volatility measures increase dramatically. We further reflect on this in the conclusions.

5.2 International Analysis

Our international data are much more limited than the data we have for the US. First, we do not construct compositional variables. The data vendor, DataStream, gradually increased its coverage of international firms over time, which makes the time-series of compositional variables difficult to interpret. Fortunately, the analysis on the U.S. seems to suggest these variables are far less important than cash flow and business cycle variables. In Table 6, Panel B, we show the data we have available for the international countries. Unfortunately, many of our variables are only available at the annual frequency. With such limited data, the best we can do is run a panel analysis.

We create country specific variables for the fundamentals and the business cycle variables, except for the variance premium where we simply use the US values as an indicator of global risk appetite. We consider two different models for the international analysis.

In a first model, we simply take the same model as we applied to the US, with all explanatory variables country-specific, but coefficients pooled across the G7 countries. The panel model uses country dummies and clusters the standard errors on year, so that correlations between countries are taken into account. The assumption of pooled coefficients is restrictive, but with the sample only starting in 1983, we have only 26 time series points so imposing such restrictions is necessary. Table 10 reports the results from the two analyses, sub-group regressions and the Hendry model reduction technique. For the sub-group analysis, we end up with 7 significant variables, indto and the three maba variables, among the fundamentals; and market volatility, the term spread and the US variance premium, among the business cycle variables. Market wide variability is now the most important determinant accounting for 31% of the predictable variation. Together, the business cycle variables account for about 55% of the predictable variation, the corporate variables for about 45% with maba still being the most important cash flow variable. This decomposition reverses the relative importance of the corporate versus business cycle variables relative to the US results, but it is at the same time rather similar. A surprising result is the importance of the US market variance premium as a determinant of time series movements in international idiosyncratic variances. Note that this decomposition excludes the effect of the country dummies. The country dummies by themselves account for about 14% of the 69% R^2 of the model.

When we consider the Hendry model, we retain 9 variables in the final model, all of which are significant at the 5% level. The business cycle variables are the same as in the sub-group model. For the corporate variables, vwroe and veps now also enter significantly, but "vwroe" enters with the wrong sign. In the decomposition, it is again "maba", the growth option proxy, consistent with the US results, that is by far the most important cash flow variable. Among the business cycle variables, the decomposition again reveals that about 30% of the total explained variation is accounted for by the total market variance, and about 20% by the variance premium. The split between corporate variables and business cycle variables is now about 50-50.

One interesting question to be addressed, is how much of the strong international commonality in idiosyncratic variances these models can explain. Table 11 provides the answer. We first report the correlation for the original raw data and then for the residuals of the two regression models we just discussed. With few exceptions, the correlations drop rather substantially, often becoming negative. While the correlations do not seem negligible in many cases, they are statistically much closer to zero than the original, raw correlations. Of the 21 correlations, 16 were statistically significantly different from zero originally. Using regression residuals, only 9(5) significant correlations remain when we use the subgroup (Hendry) model[11]. We also report the correlations of model-implied idiosyncratic volatility across countries in the last two panels of Table 11. It is not surprising that the model-implied idiosyncratic volatilities are all highly and significantly correlated.

[11] It is difficult to dismiss the possibility of a missing common factor. In that scenario, country residuals should still show significant correlations. In the CLMX model these residuals are by construction zero. When we use the FF model, the average correlation among the residuals is 23% , but when we employ the WLFF model, the average correlation becomes 12%. This indicates that the domestic risk models do omit important systematic variation.

Finally, within the context of this model, we ask formally whether the US idiosyncratic volatility process is different from that of the other countries. To do so, we take the model in Panel B of Table 10, and re-estimate it allowing for all coefficients to be of the form $b_0 + b_1 D_{US}$, where the D_{US} dummy variable is 1 for US data and zero otherwise. The joint test for 9 dummy coefficients to be zero fails to reject at the 5% level, but would reject at the 10% level. Because there are only two significant dummy coefficients, and the regression clearly suffers from multi-collinearity; we do not discuss these results any further.

The second model we consider recognizes the strong correlation between idiosyncratic volatilities across countries, and investigates whether country-specific determinants of idiosyncratic variability are still important once we control for a "U.S. factor" in idiosyncratic variances. The model is as follows:

$$\sigma_{i,t}^2 = \beta_i \sigma_{US,t}^2 + \gamma'(z_{i,t} - z_{US,t}) + e_{i,t}.$$

That is, we allot each country a beta relative to the US, and then see if differences in the usual explanatory variables, $z_{i,t} - z_{US,t}$, further explain time-variation in the idiosyncratic variance. Table 12 reports the results. In the first regression, when we only include the US idiosyncratic variance, all betas with respect to the US's variance are highly statistically significant, ranging from 0.306 for U.K. to 0.610 for Japan. The high beta for Japan may be surprising, but that Canada also has a high beta makes sense. In regressions 2 through 4, we add the explanatory variables, showing that about half of them remain statistically significant in the presence of the U.S. factor. In the final regression, we apply the Hendry procedure to pare down the regression. The explanatory variables surviving remain similar to what we found in the first model[12]. The corporate variables surviving are veps, industry turnover, but only one of the maba (growth options) variables survives, namely vmaba. While the variables have the right sign, their explanatory power has become very limited, compared to the first model. We still find the total market volatility, the term spread and the variance premium to come in significantly, but their contribution to the explained variance has also been significantly reduced, remaining only economically significant for the variance premium. In other words, the joint comovement with the US captures most of the explained variance and the economic importance of market volatility in explaining idiosyncratic volatility seems to be primarily U.S. driven. Of course, such a conclusion may change if we had better international data, but it

[12] This would not be surprising if the betas were all close to one, as then the second model is implied by the first model we estimated.

again confirms the importance of the common component in idiosyncratic variances.

6. Conclusions

This article first documents a simple fact: there is no upward trend in idiosyncratic volatility anywhere in the developed world. Instead, we find that idiosyncratic volatility is well described by a stationary mean-reverting process with occasional shifts to a higher-mean, higher-variance regime. Given the claim to the contrary for the U.S. in the influential CLMX article, a substantial literature has attempted to explain trending behavior in idiosyncratic volatilities.

Such explanations include the increasing propensity of firms to issue public equity at an earlier stage in their life cycle, and more volatile cash flows / fundamentals. In this article, we ran a comprehensive horse race using the variables proposed in the literature regarding index composition and cash flow variability, but adding business cycle variables and market wide variability to the mix. We find that the cash flow variables (especially a growth option proxy, market to book value of assets), various business cycle variables and market wide volatility are the most important determinants of the time variation in US aggregate idiosyncratic variability. However, a linear regression model does not eliminate the regime switching characteristics of the idiosyncratic variability, and we find a significant regime dependence of the regression coefficients.

One potential explanation is that in times of crisis, all risk variables increase disproportionately in ways that are hard to capture by simple linear models. To provide some initial exploration, define a crisis or bear market to be a market return two standard deviations below the mean for the US sample series over 1980-2008 (to be consistent with our international sample). Aggregate idiosyncratic volatility is almost 13% higher in crises than the sample average, just as aggregate market volatility is also considerably higher in bear markets. Using the US-based definition of a crisis to investigate idiosyncratic variability in other countries, we find that idiosyncratic variability is uniformly higher in these crisis periods than in normal periods, typically by a large margin. On average, the average idiosyncratic variability over the G7 countries is 8.9% higher than the sample mean in crises. Note that these US crises also represent local crises, as the mean return is -12.1% over the G7 countries. In a nice analogy with findings regarding international return correlations (see Longin and Solnik, 2001; Ang and Bekaert, 2002), idiosyncratic variances are also much more highly correlated across countries during crises. In fact, the difference between normal and bear market correlations is much larger than it is for actual returns. For actual returns, the G7 correlation is on average 52%, and the bear market correlation is 60.9%; for idiosyncratic volatilities, the average correlation across G7 countries is 56%, but the bear market correlation is 80.3%. While extreme movements in discount rates may be part of the story here, a full explanation of this phenomenon is beyond the scope of the article[13].

Consequently, the crisis interpretation may also partially explain another major new finding in this article: idiosyncratic variability is highly correlated across countries, and this correlation has increased over time. It is higher when the US idiosyncratic variability and market wide variability are high. Preliminary work with a linear model for annual data also detected some significant explanatory power for cash flow variables, the business cycle and market wide volatility, and the model did succeed in significantly reducing the correlation across countries, suggesting part of the comovement may have a fundamental explanation.

[13]We performed some preliminary work with a regime switching model for the US long sample accommodating three regimes. The third regime captures periods of extreme high idiosyncratic volatility, and such periods, apart from a short period during the "Tech bubble", mostly coincide with periods of market stress and low stock returns, such as the October 1987 crisis, the bear market in 1998-2002, and the recent crisis period in 2008.

References

Aktas, N., E. De Bodt, and J. Cousin, 2007, Assessing the Power and the Size of the Event Study Method through the Decades, working paper.

Ang, A. and G. Bekaert, 2002, International asset allocation with regime shifts, Review of Financial Studies, 15(4), 1137-1187.

Ang, A., R. J. Hodrick, Y. Xing and X. Zhang, 2006, The Cross-Section of Volatility and Expected Returns, Journal of Finance, 61, 259-299.

Ang, A., R. J. Hodrick, Y. Xing and X. Zhang, 2009, High Idiosyncratic Volatility and Low Returns: International and Further U.S. Evidence, Journal of Financial Economics, 91, 2-23.

Baele, L., Bekaert, G., and K. Inghelbrecht, 2010, The Determinants of Stock and Bond Return Comovements, Review of Financial Studies, forthcoming.

Bai, J. and P. Perron, 1998, Estimating and Testing Linear Models with Multiple Structural Changes, Econometrica, 66, 47-78.

Bartram, S., G. Brown and R. Stulz, 2009, Why Do Foreign Firms Have Less Idiosyncratic Risk than U.S. Firms? Working paper.

Bekaert, G., and E. Engstrom, 2009, Asset Return Dynamics under Bad Environment-Good Environment Fundamentals, working paper.

Bekaert, G., R. J. Hodrick and X. Zhang, 2009, International Stock Return Comovements, Journal of Finance, 64, 2591-2626.

Bennett, J. A., R. W. Sias, and L. T. Starks, 2003, Greener Pastures and the Impact of Dynamic Institutional Preferences, Review of Financial Studies, 16, 1203-1238.

Bollerslev, T., Tauchen, G. E., and Zhou, H., 2009, Expected Stock Returns and Variance Risk Premia, Review of Financial Studies, 22, 4463-4492.

Brandt, M. W., A. Brav, J. Graham, and A. Kumar, 2010, The Idiosyncratic Volatility Puzzle: Time Trend or Speculative Episodes? Review of Financial Studies, 23, 863-899.

Brown, G., and N. Kapadia, 2007, Firm-Specific Risk and Equity Market Development, Journal of Financial Economics, 84, 358-388.

Bunzel, H. and T. J. Vogelsang, 2005, Powerful Trend Function Tests That Are Robust to Strong Serial Correlation, with an Application to the Prebisch-Singer Hypothesis, Journal of Business and Economic Statistics, 381-394.

Campbell, J. Y., M. Lettau, B. G. Malkiel and Y. Xu, 2001, Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. Journal of Finance, 56, 1-43.

Cao, C., T. Simin and J. Zhao, 2007, Can Growth Options Explain the Trend in Idiosyncratic Risk? Review of Financial Studies, 2599-2633.

Chun, H., J. Kim, R. Morck and B. Yeung, 2007, Creative Destruction and Firm-Specific Performance Heterogeneity, Journal of Financial Economics, 109-135.

Comin, D. and S. Mulani, 2006, Diverging Trends in Macro and Micro Volatility: Facts, Review of Economics and Statistics, 374-383.

Comin, D. and T. Philippon, 2005, The Rise in Firm-Level Volatility: Causes and Consequences, NBER Macroeconomics Annual 2005, 20, NBER, MA, in (eds.) M. Gertler and K. Rogoff.

Dickey, D. and W. Fuller, 1979, Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association 74, 427-431.

Drechsler, I. 2009, Uncertainty, Time-Varying Fear, and Asset Prices, working paper

Drechsler, I., and A. Yaron, 2008, What's Vol Got to Do With It, working paper.

Fama, E. F. and K. R. French, 1996, Multifactor explanations of asset pricing anomalies, Journal of Finance 51, 55-84.

Ferreira, M. and P. Gama, 2005, Have world, country and industry risks changed over time? An investigation of the developed stock market volatility, Journal of Financial and Quantitative Analysis, 40, 195-222.

Fink, J., K. Fink, G. Grullon, and J. Weston, 2010, What Drove the Increase in Idiosyncratic Volatility during the Internet Boom?, forthcoming, Journal of Financial and Quantitative Analysis.

Foucault, T., D. Sraer, and D. Thesmar, 2008, "Individual Investors and Volatility", Working paper.

Gaspar, J. and M. Massa, 2006, Idiosyncratic Volatility and Product Market Competition, Journal of Business, 3125-3152.

Gilchrist, S., V. Yankov and E. Zakrajsek, 2009, Credit Market Shocks and Economic Flucutations: Evidence from Corporate Bond and Stock Markets, working paper.

Guo, H., and R. Savickas, 2008, Average Idiosyncratic Volatility in G7 Countries, Review of Financial Studies, 21, 1259-1296.

Hamilton, 1994, Time-Series Analysis, Prentice Hall.

Harvey, C., 1988, The Real Term Structure and Consumption Growth, Journal of Financial Economics 22, 305-334.

Henkel, S., J. C. Martin and F. Nardari, 2009, Time-varying Short Horizon Return Predictability, working paper.

Hendry, D. F., and H.-M. Krolzig, 2001, Automatic Econometric Model Selection, London.

Irvine, P. J., and J. Pontiff, 2008, Idiosyncratic Return Volatility, Cash Flows, and Product Market Competition, Review of Financial Studies, 1149-1177.

Jin, L. and S. C. Myers, 2006, "R2 Around the World: New Theory and New Tests," Journal of Financial Economics, 71, 257-292.

Kothari, S. P. and Warner, J. B., 2004, The Econometrics of Event Studies, unpublished working paper.

Longin, F. and B. Solnik, 2001, Extreme Correlation of International Equity Markets, Journal of Finance, 56, 649-676.

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

Mueller, P., 2009, Credit Spreads and Real Activity, working paper.

Phillips, P. and P. Perron, 1988. Testing for a Unit Root in Time Series Regression, Biometrika 75, 335-346.

Schwert, G. W., 1989, Why does Stock Market Volatility Change Over Time, Journal of Finance, 44 (5), 1115-1153.

Schwert, G. W., 1990, Stock Volatility and the Crash of '87, Review of Financial Studies, 3 (1990) 77-102.

Timmermann, A., 2000, Moments of Markov Switching Models, Journal of Econometrics, 96, 75-111.

Vogelsang, T., 1998, Trend function hypothesis testing in the presence of serial correlation, Econometrica 66, 123-148.

Wei, S. X., and C. Zhang, 2006, Why Did Individual Stocks Become More Volatile?, Journal of Business, 79, 259-292.

White, H., 1980, A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity, Econometrica 48, 817-838.

Xu, Y., and B. G. Malkiel, 2003, Investigating the Behavior of Idiosyncratic Volatility , Journal of Business, 76 (4), 613-644

Zhang, C., 2009, A Re-examination of the Causes of Time-varying Stock Return Volatilities, forthcoming in Journal of Financial and Quantitative Analysis.

Appendix

A. Residual specification tests

We apply the specification test to three models:

$$AR(1) \quad : \quad y_t = b_0 + b_1 y_{t-1} + e_t^{AR}, \tag{10}$$

$$GARCH(1,1) \quad : \quad y_t = b_0 + b_1 y_{t-1} + e_t^{GARCH}, \tag{11}$$

$$h_t \quad = \quad \sigma^2_{t|t-1}(e_t^{GARCH}) = w + c_0 h_{t-1} + c_1 (e_{t-1}^{GARCH})^2, \tag{12}$$

$$RS \quad : \quad y_t = (1 - b_i)\mu_i + b_i y_{t-1} + \sigma_i e_t^{RS}, i = 1, 2. \tag{13}$$

In the estimation, all error terms are assumed to be normally distributed. We examine specification tests for the residuals, e_t, which should have the following first order moment conditions:

$$E(e_t) \quad = \quad 0, \tag{14}$$

$$E(e_t e_{t-1}) \quad = \quad 0. \tag{15}$$

Because e_t is forced to have mean zero in the autoregressive specification, but may not have zero mean in other specifications, we work with demeaned residuals. For second order moments, we have

$$E(e_t^2) - \sigma^2_{t|t-1}(e_t) = 0,$$

where $\sigma^2_{t|t-1} = var(e_t)$ for the AR(1) model, and $\sigma^2_{t|t-1} = h_t$ for the GARCH(1,1) model. Moreover, the serial correlation of the squared residuals also ought to be zero, for which we use

$$E[(e_t^2 - \sigma^2_{t|t-1})(e_{t-1}^2 - \sigma^2_{t-1|t-2})] = 0, \tag{16}$$

Finally, we test the correct specification of the higher order moments for the residuals. For skewness, we have

$$E[e_t^3 - (\sigma^2_{t|t-1})^{3/2}] = 0; \tag{17}$$

for kurtosis, we have

$$E[e_t^4 - [\sigma^2_{t|t-1}]^2] = 0. \tag{18}$$

The calculations are more complicated for the RS model. We start by computing the residuals conditioning on the $t-1$ information in an obvious manner:

$$e_t^{RS} = y_t - E(y_t|t-1) = y_t - p_1[(1-b_1)\mu_1 + b_1 y_{t-1}] - p_2[(1-b_2)\mu_2 + b_2 y_{t-1}], \tag{19}$$

where p_1 denotes the probability of being in regime 1, and p_2 the probability of being in regime 2. We compute residuals using ex-post (smoothed) probabilities.

To shorten future formulas, we define

$$\begin{aligned}
y_{t|t-1} &= E(y_t|t-1) \\
&= p_1[(1-b_1)\mu_1 + b_1 y_{t-1}] + p_2[(1-b_2)\mu_2 + b_2 y_{t-1}], \\
k_1 &= (1-b_1)\mu_1 + b_1 y_{t-1} - y_{t|t-1}, \\
k_2 &= (1-b_2)\mu_2 + b_2 y_{t-1} - y_{t|t-1},
\end{aligned}$$

The conditional variance is as follows,

$$\sigma^2_{t|t-1} = p_1\sigma_1^2 + p_2\sigma_2^2 + p_1 k_1^2 + p_2 k_2^2. \tag{20}$$

Now we can compute the moment conditions (14) through (16) as before.

The formulas for the unscaled skewness and kurtosis are in Timmermann (2000) and become, for our model:

$$\begin{aligned}
skew_{t|t-1} &= [p_1(3\sigma_1^2 k_1 + k_1^3) + p_2(3\sigma_2^2 k_2 + k_2^3)], \qquad (21)\\
kurt_{t|t-1} &= [p_1(3\sigma_1^4 + k_1^4 + 6\sigma_1^2 k_1^2) + p_2(3\sigma_2^4 + k_2^4 + 6\sigma_2^2 k_2^2)] - 3(\sigma^2_{t|t-1})^2. \qquad (22)
\end{aligned}$$

Consequently, the last two moment conditions are,

$$\begin{aligned}
E[(e_t^{RS})^3 - skew_{t|t-1}] &= 0, \qquad (23)\\
E[(e_t^{RS})^4 - kurt_{t|t-1}] &= 0. \qquad (24)
\end{aligned}$$

To test all moment conditions jointly, we always use a Newey-West (1987) covariance matrix with 12 lags.

B. Accounting Data Details

In this appendix, we describe how we construct the accounting data variables as in Table 7.

All return on equity (ROE) related variables are computed as in Wei and Zhang (2006), where the variable "vwroe" is the value weighted average of firm level return on equity; the variable "vwvroe" is the value weighted average of the 12-quarter time-series variance of firm level return on equity, and the variable "cvroe" is the cross-sectional variance of the firm level return on equity.

Irvine and Pontiff (2008) focus on competition measures. We follow their procedure and compute "veps" as the cross-sectional variance of shocks to earnings per share (EPS). The shocks to EPS are computed using a pooled auto-regressive regression of year-to-year changes in quarterly EPS. To be more specific, the dependent variable is the annual difference in earnings per share, $EPS(t) - EPS(t-4)$, at the firm level, where t is current quarter, and the independent variables are $EPS(t-1) - EPS(t-5)$, $EPS(t-2) - EPS(t-6)$, and $EPS(t-3) - EPS(t-7)$. This regression attemps to adjust for seasonality in the EPS data. By computing the cross-sectional variance, this approach implicitly adjusts for the market average level of shocks to EPS, in the same spirit of "cvroe". We also compute industry turnover as the cross-sectional average at firm level for industry entries and exits each month.

Cao, Simin and Zhao (2007) consider growth options as an explanation. The most successful variable in their paper is "maba", the value weighted average of firm level market assets over book assets. We also compute "vmaba" as the value weighted average of the 12-quarter time-series variance of firm level market assets over book assets. Following the same reasoning as for "cvroe" and "veps", we also compute "cvmaba" as the cross-sectional variance of firm level market assets over book assets.

For the R&D expenditure variables, quarterly data on R&D is not reported by the majority of firms in US. So we rely on annual data on R&D. Following Comin and Mulani (2006), for each quarter, we take the corresponding fiscal year R&D, and then divide by the quarter's total revenues (sales). We also compute the cross-sectional dispersion in R&D (denoted cvrd) across firms each quarter, and it has a correlation of 80% with R&D expenditures.

Notice that all U.S. accounting data are from Compustat, and thus they are quarterly data on a firm-by-firm basis. However, because firms have different fiscal year end's, the data are spread out over the year. To ensure that each month represents the full sample of firms, we follow the procedure in Irvine and Pontiff (2008) and for each month average the accounting measures of that month and the previous two months. We apply the same methodology to all quarterly accounting data.

For the international data, we compute ROE, maba, and competition related variables as we do for U.S. firms. The variable "vwroe" is the value weighted average of the annual firm level return on equity; the variable "vwvroe" is the value weighted average of the 3-year time-series variance of the annual firm level return on equity, and one variable "cvroe" is the cross-sectional variance of the firm level return on equity each year. We compute "veps" as the cross-sectional variance of shocks to annual earnings per share, where the shocks are estimated using a pooled regression within each country. To be more specific for the pooled regression, the dependent variable becomes $EPS(t) - EPS(t-1)$, where t is current year, and the independent variable is $EPS(t-1) - EPS(t-2)$. The variable "maba" is the value weighted average of the annual firm level market assets over book assets. We also compute "vmaba" as the value weighted average of the 3-year time-series variance of annual firm level market assets over book assets, and "cvmaba" as the cross-sectional variance of annual firm level market assets over book assets.

Table 1. Idiosyncratic variance summary statistics

Panel A provides summary statistics for the U.S. sample of January 1964 to December 2008. Panel B reports summary statistics for the developed countries sample of January 1980 to December 2008. Panel C presents correlations between G7 idiosyncratic variances. We use bold font if the correlation is significantly different from zero at the 5% level. The U.S. return data are obtained from CRSP, and the return data for other countries are obtained from DataStream. All the returns are denominated in U.S. dollars. The variables σ^2_{CLMX} and σ^2_{FF} are the aggregate firm level idiosyncratic variances, as defined in equations (2) and (4), respectively. All variance time-series statistics are annualized.

Panel A. U.S. sample, 1964 – 2008

	σ^2_{CLMX}		σ^2_{FF}	
N	Mean	Std	Mean	Std
540	0.0800	0.0592	0.0697	0.0484

Panel B. Developed countries sample, 1980 – 2008

		σ^2_{CLMX}		σ^2_{FF}	
	N	Mean	Std	Mean	Std
CANADA	342	0.0880	0.0476	0.0844	0.0433
FRANCE	342	0.0692	0.0377	0.0696	0.0386
GERMANY	342	0.0537	0.0655	0.0492	0.0426
ITALY	342	0.0758	0.0536	0.0727	0.0485
JAPAN	342	0.0912	0.0487	0.0815	0.0426
U.K.	342	0.0529	0.0429	0.0550	0.0459
U.S.	342	0.0931	0.0661	0.0814	0.0544
AUSTRALIA	342	0.0745	0.0482	0.0712	0.0455
AUSTRIA	342	0.0413	0.0503	0.0433	0.0422
BELGIUM	342	0.0487	0.0584	0.0459	0.0367
DENMARK	342	0.0473	0.0297	0.0523	0.0365
FINLAND	288	0.0547	0.0527	0.0711	0.0529
GREECE	251	0.0901	0.0701	0.0798	0.0480
HK	342	0.0792	0.0564	0.0710	0.0454
IRELAND	342	0.0474	0.0598	0.0683	0.0683
NETHERLANDS	342	0.0292	0.0303	0.0369	0.0322
NEW ZEALAND	275	0.0404	0.0311	0.0518	0.0252
NORWAY	342	0.0800	0.0526	0.0859	0.0537
PORTUGAL	251	0.0677	0.0958	0.0597	0.0385
SINGAPORE	342	0.0656	0.0556	0.0591	0.0388
SPAIN	275	0.0435	0.0406	0.0457	0.0361
SWEDEN	342	0.0568	0.0425	0.0664	0.0403
SWITZERLAND	342	0.0312	0.0292	0.0326	0.0262

Panel C. Correlations between the idiosyncratic variances of the G7 countries, 1980 – 2008

	σ^2_{CLMX}					
	Canada	France	Germany	Italy	Japan	U.K.
France	**56%**					
Germany	**62%**	**57%**				
Italy	**31%**	**51%**	**20%**			
Japan	**56%**	**54%**	**57%**	**23%**		
U.K.	**74%**	**68%**	**81%**	**31%**	**72%**	
U.S.	**75%**	**65%**	**68%**	**20%**	**70%**	**80%**

	σ^2_{FF}					
	Canada	France	Germany	Italy	Japan	U.K.
France	**63%**					
Germany	**77%**	**67%**				
Italy	**32%**	**53%**	**19%**			
Japan	**65%**	**62%**	**71%**	**31%**		
U.K.	**68%**	**62%**	**74%**	**33%**	**70%**	
U.S.	**76%**	**71%**	**81%**	**27%**	**72%**	**71%**

Table 2. Trend tests

Panel A and Panel B reports trend test results for the U.S. idiosyncratic variance time-series, and Panel C reports trend test results for the idiosyncratic variance time-series of all developed countries. All panels use Vogelsang's (1998) t-PS1 test and Bunzel and Vogelsang's (2008) t-dan test. The 5% critical value (two sided) for t-dan is 2.052, and for t-ps1 is 2.152. We report both pre-whitened results using AR (1) and non-pre-whitened results for the t-dan test in Panel A and B, and for Panel C, we only use the pre-whitened results. Variables σ^2_{CLMX} and σ^2_{FF} are the aggregate firm level idiosyncratic variances, as defined in equations (2) and (4), respectively. All variance time-series statistics are annualized. All coefficients are multiplied by 100.

Panel A. Idiosyncratic variances over 1964-1997, daily data

	Pre-whitened		Not pre-whitened			
	b-dan	t-dan	b-dan	t-dan	b-ps1	t-ps1
σ^2_{CLMX}	0.011	4.84	0.011	4.97	0.011	3.89
σ^2_{FF}	0.009	4.35	0.009	4.72	0.009	3.36

Panel B. Idiosyncratic variances over 1964-2008, daily data

	Pre-whitened		Not pre-whitened			
	b-dan	t-dan	b-dan	t-dan	b-ps1	t-ps1
σ^2_{CLMX}	0.015	0.95	0.015	1.04	0.016	1.35
σ^2_{FF}	0.013	0.76	0.013	0.87	0.014	1.15

Panel C. Idiosyncratic variances over 1980-2008, daily data

	σ^2_{CLMX}				σ^2_{FF}			
	b-dan	t-dan	b-ps1	t-ps1	b-dan	t-dan	b-ps1	t-ps1
CANADA	0.059	0.35	0.097	0.48	0.086	0.37	0.120	0.50
FRANCE	-0.090	-0.28	-0.004	-0.01	-0.051	-0.16	0.030	0.09
GERMANY	0.290	0.48	0.346	0.73	0.326	0.24	0.393	0.51
ITALY	-0.414	-2.62	-0.375	-1.86	-0.351	-2.29	-0.305	-1.56
JAPAN	0.016	0.07	0.050	0.23	0.027	0.08	0.046	0.16
U.K.	0.070	0.06	0.053	0.08	0.075	0.09	0.064	0.10
U.S.	0.053	0.03	0.141	0.15	0.072	0.03	0.166	0.15
AUSTRALIA	-0.055	-0.27	-0.125	-0.81	-0.051	-0.19	-0.115	-0.61
AUSTRIA	0.496	0.93	0.543	0.82	0.531	1.35	0.587	1.18
BELGIUM	-0.052	-0.06	-0.058	-0.13	-0.061	-0.09	-0.083	-0.19
DENMARK	0.065	0.37	0.126	0.50	0.501	0.30	0.600	0.45
FINLAND	-0.514	-0.21	-0.584	-0.41	-0.593	-0.34	-0.718	-0.99
GREECE	-0.323	-1.57	-0.377	-1.77	-0.230	-1.07	-0.282	-1.42
HK	0.030	0.14	0.001	0.00	0.029	0.11	-0.005	-0.02
IRELAND	0.077	0.08	-0.008	-0.02	0.055	0.04	-0.022	-0.03
NETHERLANDS	0.198	0.31	0.253	0.45	0.195	0.16	0.235	0.28
NEW ZEALAND	-0.019	-0.04	-0.089	-0.20	-0.145	-0.14	-0.243	-0.31
NORWAY	0.017	0.06	0.060	0.18	0.020	0.07	0.055	0.21
PORTUGAL	-0.651	-4.33	-0.773	-4.93	-0.485	-1.30	-0.610	-2.27
SINGAPORE	0.049	0.42	0.002	0.01	0.238	2.00	0.259	1.63
SPAIN	-0.298	-1.38	-0.350	-1.61	-0.259	-0.38	-0.262	-0.49
SWEDEN	7.009	0.00	8.203	0.01	0.050	0.13	0.103	0.29
SWITZERLAND	0.120	0.41	0.159	0.51	0.174	0.53	0.223	0.71

Table 3. Robustness checks: global model idiosyncratic volatilities and equal weighted idiosyncratic variances

Panel A reports trend test results for 23 countries idiosyncratic volatilities time series, using the Vogelsang (1998) t-PS1 test and the Bunzel and Vogelsang (2008) t-dan test. The 5% critical value (two sided) for t-dan is 2.052, and for t-ps1 is 2.152. The variable σ^2_{WLFF} is the aggregate firm level idiosyncratic variances, as defined in equation (6). In Panel B, we report summary statistics for the U.S. sample, and the sample period is January 1964 to December 2008. In Panel C, we report trend test results for U.S. idiosyncratic variance time-series, using the trend tests described above. We compute idiosyncratic variances using the CLMX and FF model. In both Panels A and C, we use a pre-whitened model for the t-dan test. Coefficients in Panel A and C are multiplied by 100. All variance time-series statistics are annualized.

Panel A. Trend test for WLFF model idiosyncratic volatilities

	σ^2_{WLFF}			
	b-dan	t-dan	b-ps1	t-ps1
CANADA	0.218	0.07	0.234	0.16
FRANCE	-1.495	-0.18	-1.294	-0.44
GERMANY	1.019	0.01	1.089	0.10
ITALY	-2.566	-0.91	-2.212	-1.08
JAPAN	0.234	0.02	0.223	0.07
U.K.	-0.154	-0.01	-0.120	-0.04
U.S.	0.003	0.00	0.255	0.03
AUSTRALIA	-0.732	-0.03	-1.089	-0.36
AUSTRIA	1.533	0.47	1.756	0.66
BELGIUM	-0.705	0.00	-0.866	-0.06
DENMARK	-0.693	-0.04	-0.682	-0.16
FINLAND	-1.951	-0.01	-1.909	-0.07
GREECE	-1.877	-0.05	-1.673	-0.19
HK	-0.669	-0.36	-0.778	-0.55
IRELAND	0.180	0.01	-0.337	-0.06
NETHERLANDS	0.261	0.01	0.295	0.06
NEW ZEALAND	-1.683	-0.33	-1.798	-0.82
NORWAY	-0.805	-0.59	-0.975	-1.33
PORTUGAL	-3.383	-0.12	-4.344	-0.77
SINGAPORE	0.049	0.01	-0.001	0.00
SPAIN	-2.733	-0.16	-3.007	-0.71
SWEDEN	-0.748	-0.40	-0.514	-0.35
SWITZERLAND	0.526	0.04	0.759	0.19

Panel B. Summary statistics for the U.S.

	σ^2_{CLMX}		σ^2_{FF}	
N	Mean	Std	Mean	Std
540	0.4308	0.3036	0.3530	0.2436

Panel C. Trend test for the U.S.

	1964-1997				1964-2008			
	b-dan	t-dan	b-ps1	t-ps1	b-dan	t-dan	b-ps1	t-ps1
σ^2_{CLMX}	0.154	1.20	0.135	0.87	0.109	0.23	0.142	0.71
σ^2_{FF}	0.132	0.52	0.116	0.59	0.088	0.17	0.118	0.60

Table 4. Regime switching model estimation results.

This table reports the regime switching model results for the idiosyncratic variance time-series computed using daily data, where the model is described as follows:

$y_t = (1-b_i)\mu_i + b_i y_{t-1} + \sigma_i e_t, i = 1, 2.$

Transition probability matrix $\Phi = \begin{bmatrix} p11 & 1-p11 \\ 1-p22 & p22 \end{bmatrix}$.

The transition probability parameters, p11 and p22 are constrained to be in (0,1) during estimation. We also re-parameterize to ensure $0 < \mu_1 < \mu_2$. The left half panel reports results for the U.S. The sample period is 1964-2008. The variables σ^2_{CLMX} and σ^2_{FF} are the aggregate firm level idiosyncratic variances, as defined in equations (2) and (4), respectively. The right half panel reports results for the aggregate idiosyncratic variance time-series of the G7 countries. The sample period becomes 1980-2008. All variance time-series statistics are annualized.

	long sample: 1964-2008				short sample: 1980-2008						
	U.S.				CA	FR	GE	IT	JP	U.K.	U.S.
	σ^2_{CLMX}		σ^2_{FF}		σ^2_{CLMX}	σ^2_{CLMX}	σ^2_{CLMX}	σ^2_{CLMX}	σ^2_{CLMX}	σ^2_{CLMX}	σ^2_{CLMX}
	coef.	std.	coef.	std.	coef.	coef.	coef.	coef.	coef.	coef.	coef.
μ_1	0.062	0.003	0.055	0.003	0.071	0.051	0.033	0.049	0.072	0.036	0.070
μ_2	0.181	0.023	0.155	0.020	0.160	0.125	0.121	0.139	0.156	0.122	0.198
b_1	0.823	0.028	0.813	0.027	0.604	0.628	0.761	0.681	0.639	0.735	0.686
b_2	0.585	0.090	0.677	0.080	0.226	0.273	0.316	0.431	0.564	0.584	0.512
σ_1	0.011	0.001	0.010	0.001	0.013	0.012	0.008	0.016	0.018	0.009	0.013
σ_2	0.082	0.007	0.057	0.005	0.073	0.048	0.113	0.061	0.055	0.058	0.086
p_{11}	0.981	0.008	0.987	0.007	0.934	0.904	0.933	0.927	0.952	0.940	0.984
p_{22}	0.900	0.061	0.935	0.044	0.672	0.593	0.750	0.804	0.830	0.716	0.934

Table 5. Regime switching model specification tests

AR stands for a first-order autoregressive model with homoskedastic errors. GARCH is the AR model with the variance of the error term following a GARCH(1,1) process. RS stands for the regime switching model discussed in the text. The moment conditions for RS models are computed following Timmermann (2000). We use smoothed ex-post regime probabilities to compute moments. In Panel A, we use 4 moments: mean, variance, and first order autocorrelations for both. In Panel B, we consider 2 moments: skewness and kurtosis. In Panel C, we combine the 6 moments in Panels A and B. To compute the p-values of the Wald tests, we always use 12 Newey-West lags to adjust for serial correlation.

Panel A. Mean, variance and auto-correlations

	US long CLMX Wald	p	US long FF Wald	p	CA CLMX Wald	p	FR CLMX Wald	p	GE CLMX Wald	p	IT CLMX Wald	p	JP CLMX Wald	p	UK CLMX Wald	p	US CLMX Wald	p
AR	14.902	0.5%	24.566	0.0%	15.425	0.4%	6.432	16.9%	23.428	0.0%	13.834	0.8%	10.323	3.5%	7.353	11.8%	11.250	2.4%
GARCH	9.317	5.4%	5.961	20.2%	14.643	0.6%	9.738	4.5%	2.382	66.6%	14.158	0.7%	11.100	2.6%	5.319	25.6%	5.078	27.9%
RS	7.037	13.4%	7.088	13.1%	4.639	32.6%	1.251	87.0%	9.526	4.9%	5.381	25.0%	0.849	93.2%	0.456	97.8%	3.336	50.3%

Panel B. Skewness and kurtosis

	US long CLMX Wald	p	US long FF Wald	p	CA CLMX Wald	p	FR CLMX Wald	p	GE CLMX Wald	p	IT CLMX Wald	p	JP CLMX Wald	p	UK CLMX Wald	p	US CLMX Wald	p
AR	3.709	15.7%	3.898	14.2%	6.458	4.0%	10.823	0.5%	5.699	5.8%	9.280	1.0%	7.159	2.8%	6.089	4.8%	4.259	11.9%
GARCH	3.493	17.4%	3.795	15.0%	5.297	7.1%	4.706	9.5%	1.110	57.4%	6.247	4.4%	9.602	0.8%	2.914	23.3%	3.203	20.2%
RS	2.432	29.6%	4.656	9.8%	5.362	6.9%	2.763	25.1%	9.487	0.9%	5.509	6.4%	7.216	2.7%	1.610	44.7%	2.375	30.5%

Panel C. All

	US long CLMX Wald	p	US long FF Wald	p	CA CLMX Wald	p	FR CLMX Wald	p	GE CLMX Wald	p	IT CLMX Wald	p	JP CLMX Wald	p	UK CLMX Wald	p	US CLMX Wald	p
AR	25.948	0.0%	32.537	0.0%	24.609	0.0%	24.428	0.0%	26.516	0.0%	22.814	0.1%	21.562	0.2%	25.122	0.0%	21.869	0.1%
GARCH	18.937	0.4%	10.005	12.4%	24.211	0.1%	18.603	0.5%	19.906	0.3%	22.117	0.1%	17.508	0.8%	13.356	3.8%	12.501	5.2%
RS	9.543	14.5%	9.279	15.9%	12.771	4.7%	3.563	73.6%	14.650	2.3%	10.668	9.9%	11.310	7.9%	13.106	4.1%	6.530	36.7%

Table 6. The common component in idiosyncratic variances across countries

For each panel, the table reports the correlations of G7 countries' idiosyncratic variance with the U.S. idiosyncratic variance. The different panels use different models to compute idiosyncratic variances. In Panels A and B, we also report the average correlations in the two regimes identified for the U.S., using smoothed regime probabilities in regime 2.

Panel A. Daily CLMX model

σ^2_{CLMX}	Correlation with U.S.			Correlation with U.S.	
	1980-2008	1980-1994	1995-2008	Regime 1	Regime 2
Canada	80%	63%	86%	53%	82%
France	66%	45%	74%	51%	68%
Germany	77%	49%	79%	37%	57%
Italy	20%	14%	47%	25%	60%
Japan	66%	23%	76%	28%	79%
U.K.	81%	73%	81%	57%	70%
U.S.	100%	100%	100%	100%	100%

Panel B. Daily FF model

σ^2_{FF}	Correlation with U.S.			Correlation with U.S.	
	1980-2008	1980-1994	1995-2008	Regime 1	Regime 2
Canada	82%	64%	84%	56%	68%
France	73%	34%	83%	41%	76%
Germany	82%	40%	82%	27%	73%
Italy	34%	26%	57%	28%	58%
Japan	69%	18%	77%	24%	77%
U.K.	84%	65%	86%	17%	73%
U.S.	100%	100%	100%	100%	100%

Panel C. Weekly WLFF model

σ^2_{WLFF}	Correlation with U.S.		
	1980-2008	1980-1994	1995-2008
Canada	82%	80%	86%
France	45%	32%	82%
Germany	69%	6%	71%
Italy	15%	35%	56%
Japan	73%	-7%	85%
U.K.	87%	52%	92%
U.S.	100%	100%	100%

Table 7. Explanatory variables

Panel A. For the U.S. analysis

variable	description
I. Index composition/behavioral variables	
pyoung	the % of market cap of firms less than 10 years old since foundation
psmall	the % of market cap of firms smaller than 25% of all firms listed
plow	the % of market cap of firms with share price lower than $5
lowto	the average volume turnover for firms with share price lower than $5
dto	aggregate dollar volume over aggregate market capitalization
II. Corporate variables	
vwroe	the value weighted average of firm level return on equity
vwvroe	the value weighted average of 12-quarter time-series variance of firm level return on equity
cvroe	the cross-sectional variance of firm level return on equity
veps	the cross-sectional variance of shocks to earnings per share
indto	the average industry turnover
maba	the value weighted average of firm level market assets over book assets
vmaba	the value weighted average of 12-quarter time-series variance of firm level market assets over book assets
cvmaba	the cross-sectional variance of firm level market assets over book assets
rd	the value weighted average of firm level R&D expenditure scaled by sales
cvrd	the cross-sectional variance of firm level R&D expenditure scaled by sales
III. Business cycle variables	
dip	the first order difference in industrial production
confi	the conference board’s index of consumer confidence
def	the yield spread between BAA and AAA corporate bonds
term	the yield spread between 10 year and 1 year government bond
mvp	the market variance premium
mkttv	the market index realized variance

Panel B. For the international analysis

variable	description
I. Corporate variables	
vwroe	the value weighted average of firm level return on equity
vwvroe	the value weighted average of 12-quarter time-series variance of firm level return on equity
cvroe	the cross-sectional variance of firm level return on equity
veps	the cross-sectional variance of shocks to earnings per share
indto	the average industry turnover
maba	the value weighted average of firm level market assets over book assets
vmaba	the value weighted average of 12-quarter time-series variance of firm level market assets over book assets
cvmaba	the cross-sectional variance of firm level market assets over book assets
II. Business cycle variables	
mkttv	the market index realized variance
usmvp	the U.S. market value premium
dgdp	the first order difference in each country's annual GDP
def	the yield spread between each country's corporate debt and government bonds
term	the yield spread between each country's long term and short term government bonds

Table 8. What drives U.S. idiosyncratic volatility?

OLS regressions of aggregate idiosyncratic variances in the U.S. over 1964-2008, computed using the CLMX model, on various determinants, labeled on the left. In Panel A, we show 4 regressions, one for each group of variables, and a final one based on a paring down technique picking significant variables from the previous regressions, discussed in the text. In Panel B, we show the regression on all variables simultaneously and a regression reduced by the general-to-specific paring down technique, described in the text. In the last row of Panel B, we also report a joint Wald test of all variables dropped from regression I to II are significantly different from zero. All p-values are based on a standard error, using 12 Newey-West lags. The last column reports the covariance decomposition described in the text.

Panel A. Subgroup regressions

	I. Behavioral and compositional		II. Corporate cash flow		III. Business cycle variables		IV. Significant variables from I-III		
	coef.	p-value	coef.	p-value	coef.	p-value	coef.	p-value	Cov decomp
pyoung	0.783	15.4%							
psmall	-12.380	8.4%							
plow	-3.903	7.2%							
lowto	-0.007	75.0%							
dto	0.043	21.6%							
vwroe			1.342	2.0%					
vwvroe			-5.053	38.0%					
cvroe			-0.564	31.1%					
veps			0.014	2.8%					
indto			0.006	1.0%			0.006	0.3%	2%
maba			0.083	1.3%			0.091	0.0%	42%
vmaba			-0.006	55.9%					
cvmaba			0.001	7.1%					
rd			0.251	2.9%			0.140	3.4%	26%
cvrd			-0.008	0.2%			-0.006	0.0%	-10%
mvp					1.405	0.1%			
mkttv					0.697	0.0%	0.726	0.0%	40%
dip					-0.664	5.3%	-0.842	0.8%	1%
def					-0.025	2.1%	0.019	0.1%	-2%
term					-0.004	17.6%			
confi					0.0002	44.6%			
Adj. R2	35%		56%		57%		80%		

Panel B. Hendry regression

	I. All variables		II. Significant variables from I		
	coef.	p-value	coef.	p-value	Cov decomp
pyoung	0.776	0.0%	0.879	0.0%	12%
psmall	11.063	0.7%	13.467	0.0%	-13%
plow	-0.339	80.4%			
lowto	-0.001	80.3%			
dto	0.023	0.3%	0.024	0.9%	10%
vwroe	0.092	82.5%			
vwvroe	1.286	74.1%			
cvroe	-0.847	0.8%	-0.952	0.0%	-11%
veps	0.009	0.1%	0.007	0.3%	7%
indto	0.003	0.7%	0.004	1.3%	1%
maba	0.094	0.0%	0.101	0.0%	44%
vmaba	-0.008	15.0%			
cvmaba	0.000	53.6%			
rd	0.172	0.1%	0.134	0.0%	23%
cvrd	-0.006	0.0%	-0.006	0.0%	-8%
mvp	0.535	0.0%	0.542	0.0%	10%
mkttv	0.642	0.0%	0.633	0.0%	32%
dip	-0.513	0.3%	-0.510	0.5%	1%
def	0.004	49.9%			
term	-0.010	0.0%	-0.011	0.0%	2%
confi	-0.001	0.0%	-0.001	0.0%	-9%
Adj. R2	86%		86%		
Wald test for eliminated variables from regression I to II					
	p-value	68.0%			

Table 9. Analyzing Model Residuals

Panel A reports specification tests of the regression model residuals and residuals for a RS model on the residuals. The residuals are computed using the subgroup model and Hendry model. We conduct specification tests on the RS model residuals using both ex ante and ex post probabilities, denoted RS ante and RS post, respectively. The moment conditions include: mean, variance, autocorrelation of first order for both mean and variance, skewness and kurtosis. To compute the p-value of the Wald tests, we always use 12 Newey-West lags to adjust for serial correlation. The parameters of the RS model are reported in Panel B. We estimate the RS model as in Table 4, except we re-parameterize to ensure $0 < \sigma_1 < \sigma_2$. The last row in Panel B reports the Wald test of $\mu_1 = \mu_2$. In Panel C, we re-estimate the subgroup and Hendry models, by allowing all coefficients to be linear functions of a dummy variable in the form of $b = b_0 + b_1 dummy$. We refer to b_0 as the "constant" coefficient, and b_1 as the "dummy" coefficient. The dummy variable takes a value of 1 if the CLMX smoothed probability in regime 2 is higher than 0.5.

Panel A. Specification test on the residuals and RS residuals of the residuals

Models	Subgroup Model		Hendry Model	
	Wald	p-value	Wald	p-value
residuals	25.71	0.0%	22.26	0.1%
RS post	9.34	15.5%	7.21	30.2%

Panel B. Regime switching model for the regression residuals

	Subgroup model		Hendry model	
	coef.	std.	coef.	std.
μ_1	-0.002	0.003	-0.001	0.002
μ_2	0.009	0.011	0.006	0.008
b_1	0.705	0.042	0.632	0.044
b_2	0.413	0.121	0.237	0.136
σ_1	0.013	0.001	0.012	0.001
σ_2	0.049	0.005	0.044	0.005
p_{11}	0.985	0.009	0.982	0.011
p_{22}	0.929	0.062	0.903	0.097
P($\mu_1 = \mu_2$)	34.9%		42.4%	

Panel C. Allowing for Regime 2 Dummy in Regression Coefficients

	Subgroup Model				Hendry Model			
	b_0		b_1		b_0		b_1	
	coef.	p-value	coef.	p-value	coef.	p-value	coef.	p-value
pyoung					0.028	10.5%	0.383	41.4%
psmall					0.200	6.3%	27.402	2.8%
plow								
lowto								
dto					-0.016	0.4%	0.057	0.0%
vwroe								
vwvroe								
cvroe					0.390	0.2%	0.754	43.1%
veps					0.000	73.4%	0.016	0.2%
indto	0.003	2.3%	0.003	48.6%	0.002	0.3%	-0.004	47.2%
maba	0.066	0.0%	0.067	0.7%	0.058	0.0%	0.087	0.2%
vmaba								
cvmaba								
rd	0.002	96.9%	0.170	1.7%	0.047	31.0%	0.116	3.8%
cvrd	-0.003	0.6%	-0.003	20.0%	-0.002	3.4%	-0.011	0.0%
mvp					0.479	0.0%	-0.295	29.9%
mkttv	0.526	0.0%	0.674	0.0%	0.514	0.0%	0.664	0.0%
dip	-0.189	20.3%	-0.642	50.4%	-0.070	55.9%	-0.780	16.4%
def	0.011	0.5%	0.039	12.7%				
term					-0.005	0.0%	-0.015	0.5%
confi					-0.001	0.0%	-0.001	4.4%
Adj. R2	89%				94%			

Table 10. Idiosyncratic volatility across G7 countries?

OLS regressions of aggregate idiosyncratic variances in the G7 countries over 1983-2008, computed using the CLMX model, on various determinants, labeled on the left. The annual data time series for idiosyncratic variance are average over monthly observations in the year. More details about data are in Appendix B. In Panel A, we show 3 regressions, one for each group of variables, and a final one based on a paring down technique picking significant variables from the previous regressions, discussed in the text. In Panel B, we show the regression on all variables simultaneously and a regression reduced by general-to-specific paring down technique, described in the text. In the last 4 row of Panel B, we also report joint Wald tests of whether all variables dropped step by step from regression I to II are significantly different from zero. All regressions include country dummies. All p-values are based on a standard error using 12 Newey-West lags and they are adjusted by clustering on years. All regressions include country dummies. The last column reports the covariance decomposition described in the text.

Panel A. Subgroup regressions

	I. Corporate cash flow		II. Business cycle variables		III. Significant variables		
	coef.	p-value	coef.	p-value	coef.	p-value	Cov decomp
vwroe	-0.129	29.9%					
vwvroe	0.129	91.8%					
cvroe	0.232	23.4%					
veps	0.042	15.3%					
indto	0.133	0.7%			0.062	3.3%	0.3%
maba	0.024	0.0%			0.022	0.0%	25.2%
vmaba	0.004	0.0%			0.002	3.6%	7.2%
cvmaba	0.003	8.8%			0.003	0.1%	11.3%
mkttv			0.455	0.0%	0.459	0.0%	31.1%
dgdp			-0.035	22.5%			
def			-0.001	70.2%			
term			-0.004	9.4%	-0.002	6.1%	0.4%
usmvp			1.180	0.1%	0.916	0.2%	24.4%
Adj. R2	41%		56%		69%		
Adj. R2 (w/o country dummies)	27%		41%		55%		

Panel B. Hendry regressions

	I. all variables		II. significant variables		
	coef.	p-value	coef.	p-value	Cov decomp
vwroe	-0.142	3.8%	-0.146	2.6%	7.1%
vwvroe	-0.498	43.6%			
cvroe	0.119	24.5%			
veps	0.039	6.6%	0.050	0.5%	0.1%
indto	0.081	0.8%	0.083	0.6%	0.3%
maba	0.024	0.0%	0.024	0.0%	24.6%
vmaba	0.002	2.3%	0.002	2.1%	6.7%
cvmaba	0.003	0.1%	0.003	0.1%	10.1%
mkttv	0.457	0.0%	0.475	0.0%	29.9%
dgdp	-0.027	29.2%			
def	0.000	91.2%			
term	-0.003	2.2%	-0.002	3.5%	0.4%
usmvp	0.859	0.2%	0.832	0.5%	20.6%
Adj. R2	0.71%		71%		
R2 (w/o country dummies)	59%		59%		
Wald test for eliminating vwvroe, cvroe, dgdp, and def at 10%					
	p-value	0.678			

Table 11. Correlations for the annual idiosyncratic volatility data

This table reports correlations for annual data of aggregate idiosyncratic variance time-series. Panel A reports correlation coefficients for the original annual idiosyncratic variance data. Panels B and C report correlation coefficients for the residuals from the final regression in the subgroup model and Hendry model, respectively. The bold font indicates that the correlation is significant at the 5% level using a Pearson test.

Panel A. Correlation for original annual idiosyncratic variance

	CANADA	FRANCE	GERMANY	ITALY	JAPAN	UK
FRANCE	**77%**					
GERMANY	**85%**	**67%**				
ITALY	14%	**55%**	2%			
JAPAN	**75%**	**78%**	**74%**	29%		
UK	**87%**	**74%**	**90%**	19%	**86%**	
US	**92%**	**81%**	**82%**	21%	**84%**	**89%**

Panel B. Correlation for regression residuals from the subgroup model

	CANADA	FRANCE	GERMANY	ITALY	JAPAN	UK
FRANCE	**65%**					
GERMANY	35%	-26%				
ITALY	20%	**71%**	**-52%**			
JAPAN	**49%**	**42%**	**44%**	16%		
UK	**67%**	32%	**53%**	-8%	**76%**	
US	34%	28%	27%	26%	16%	15%

Panel C. Correlation for regression residuals from the Hendry model

	CANADA	FRANCE	GERMANY	ITALY	JAPAN	UK
FRANCE	38%					
GERMANY	**52%**	-22%				
ITALY	3%	**66%**	-37%			
JAPAN	34%	37%	**47%**	22%		
UK	**71%**	33%	**54%**	2%	**63%**	
US	30%	27%	37%	20%	13%	12%

Panel D. Correlation for regression implied idiosyncratic volatility from the subgroup model

	CANADA	FRANCE	GERMANY	ITALY	JAPAN	UK
FRANCE	**94%**					
GERMANY	**90%**	**95%**				
ITALY	**79%**	**83%**	**85%**			
JAPAN	**86%**	**87%**	**92%**	**78%**		
UK	**93%**	**93%**	**93%**	**81%**	**85%**	

US	**74%**	**74%**	**73%**	**56%**	**69%**	**61%**

Panel C. Correlation for regression implied idiosyncratic volatility from the Hendry model

	CANADA	FRANCE	GERMANY	ITALY	JAPAN	UK
FRANCE	**89%**					
GERMANY	**84%**	**90%**				
ITALY	**72%**	**71%**	**70%**			
JAPAN	**85%**	**83%**	**86%**	**72%**		
UK	**90%**	**86%**	**87%**	**69%**	**76%**	
US	**77%**	**73%**	**68%**	**45%**	**71%**	**60%**

Table 12. Idiosyncratic volatility across G7 countries: beta model

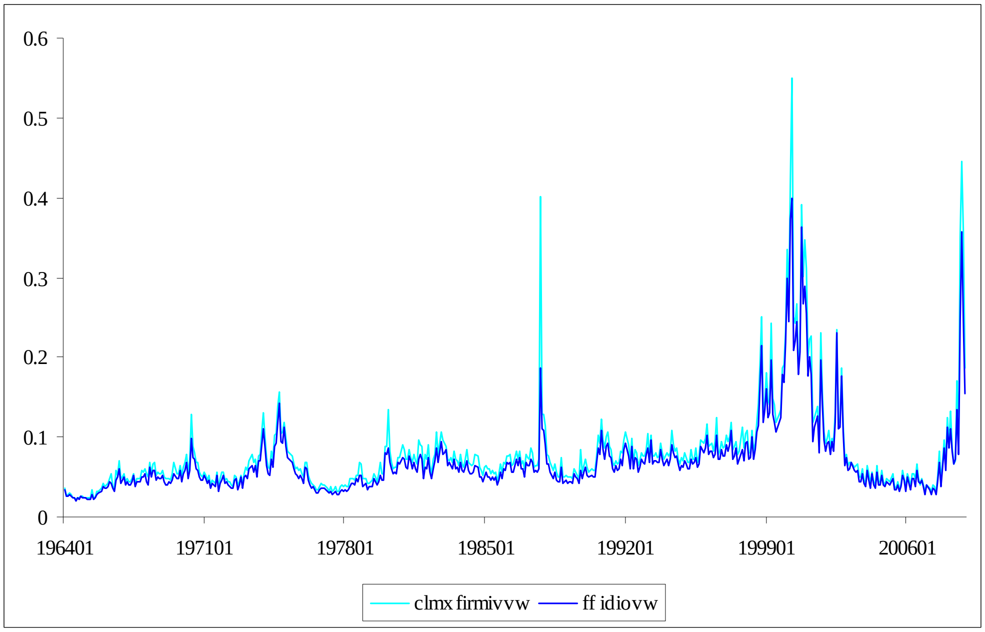
OLS regressions of aggregate idiosyncratic variances in the G7 countries over 1983-2008, computed using the CLMX model, on various determinants, labeled on the left. The annual data time series for idiosyncratic variance are average over monthly observations in the year. More details about the data are in Appendix B. We show 5 regressions: one for each group of variables, one for all regressors, starting with the betas with respect to the U.S. variance, a final subgroup one based on a paring down technique picking significant variables from the full regression. All p-values are based on a standard error using 12 Newey-West lags and adjusted for clustering on years. The last column for the fifth regression reports the covariance decomposition described in the text.

		only US idio		all corp + US idio		all cycle + US idio		all variables + US idio		Hendry variables + US idio		
		coef.	p-value	coef.	p-value	coef.	p-value	coef.	p-value	coef.	p-value	cov decomp
	vwroe			-0.105	7.1%			-0.079	12.4%			
	vwvroe			0.862	23.2%			0.886	21.6%			
	cvroe			-0.046	51.4%			0.064	33.9%			
	veps			0.077	0.2%			0.064	0.6%	0.078	0.0%	1.5%
	indto			0.080	3.6%			0.059	3.0%	0.045	5.4%	1.4%
	maba			0.005	12.0%			0.001	78.0%			
	vmaba			0.001	1.4%			0.002	0.0%	0.001	0.0%	-14.1%
	cvmaba			0.000	92.6%			0.000	94.5%			
	mkttv					0.538	0.0%	0.505	0.0%	0.514	0.0%	-0.1%
	dgdp					0.018	15.6%	0.012	35.3%			
	def					0.001	32.8%	0.001	25.1%			
	term					-0.002	0.2%	-0.002	1.3%	-0.002	1.6%	0.4%
	usmvp					0.510	0.1%	0.679	0.0%	0.703	0.0%	18.6%
dca	usidio	0.549	0.0%	0.681	0.0%	0.575	0.0%	0.628	0.0%	0.612	0.0%	34.4%
dfr	usidio	0.375	0.0%	0.539	0.0%	0.399	0.0%	0.478	0.0%	0.446	0.0%	7.0%
dge	usidio	0.339	0.0%	0.472	0.0%	0.343	0.0%	0.407	0.0%	0.371	0.0%	2.7%
dit	usidio	0.318	0.0%	0.476	0.0%	0.283	0.0%	0.405	0.0%	0.331	0.0%	0.8%
djp	usidio	0.610	0.0%	0.809	0.0%	0.562	0.0%	0.728	0.0%	0.673	0.0%	47.5%
duk	usidio	0.306	0.0%	0.534	0.0%	0.310	0.0%	0.442	0.0%	0.374	0.0%	-0.1%
adj. R2		59.6%		64.9%		67.5%		71.1%		71.5%		

Figure 1. Idiosyncratic variances over time

In Panel A, we plot the time-series idiosyncratic variance for the U.S. sample. The sample period is January 1964 to December 2008. In Panels B and C, we plot the time-series idiosyncratic variances for G7 countries. The aggregate idiosyncratic variance measures using CLMX and FF are defined in equations (2) and (4), respectively. The U.S. return data are obtained from CRSP, and the return data for other countries are obtained from DataStream. All the returns are denominated in U.S. dollars. All variance time-series statistics are annualized.

Panel A. U.S. (daily data)



Panel B. G7 (daily data, CLMX)

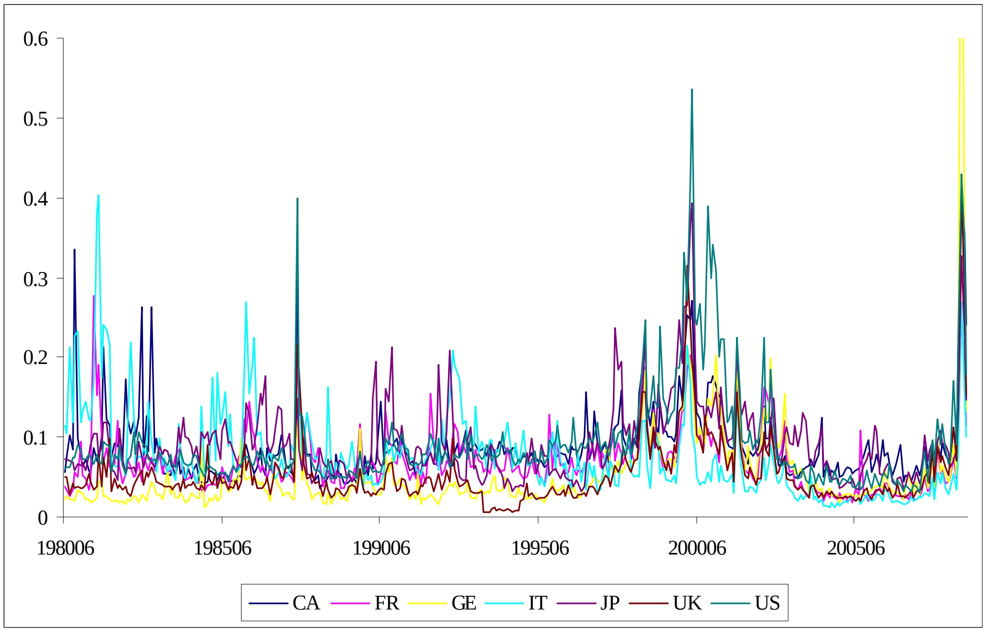
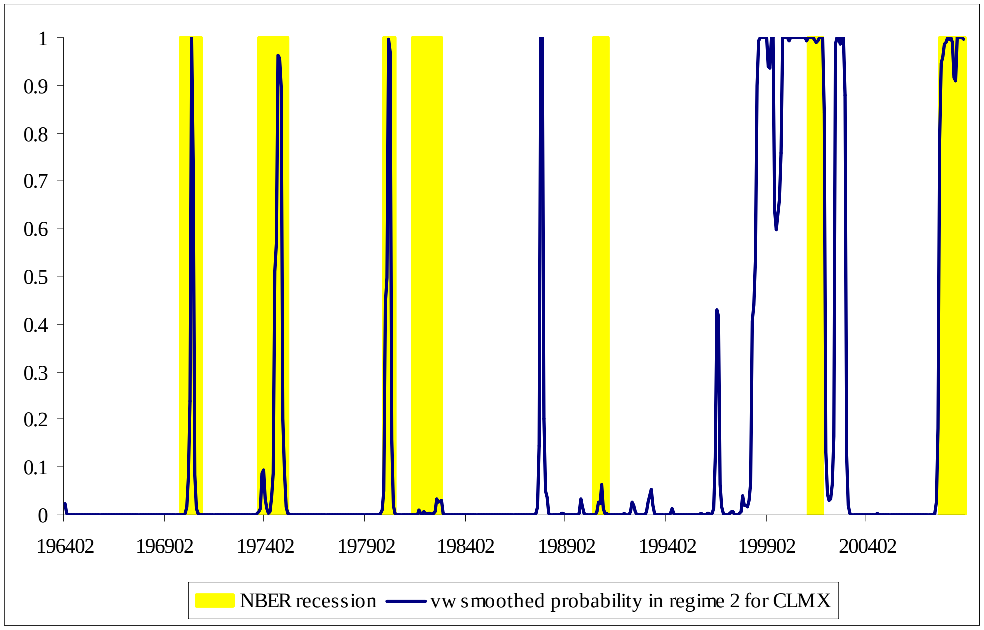


Figure 2. Regime probabilities for U.S. idiosyncratic variances
This figure reports the smoothed probability of being in regime 2 for the U.S., using a regime switching model defined in equations (6) and (7). The model is estimated over sample period 1964 – 2008. The variables σ^2_{CLMX} and σ^2_{FF} are the aggregate firm level idiosyncratic variances, as defined in equations (2) and (4), respectively.

Panel A. σ^2_{CLMX}



Panel B. σ^2_{FF}

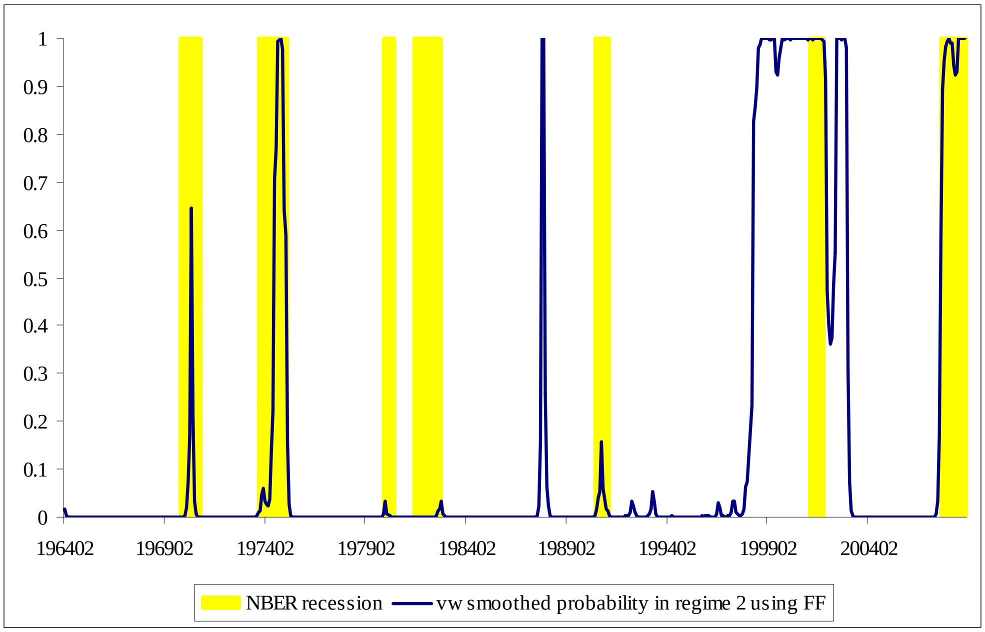


Figure 3. Regime probabilities for G7 countries

This figure reports the smoothed probability of being in regime 2 for the G7 countries other than the U.S., using a regime switching model defined in equations (6) and (7). The model is estimated over sample period 1980 – 2008, country by country. The variable σ^2_{CLMX} is the aggregate firm level idiosyncratic variance, as defined in equation (2), estimated using daily data.

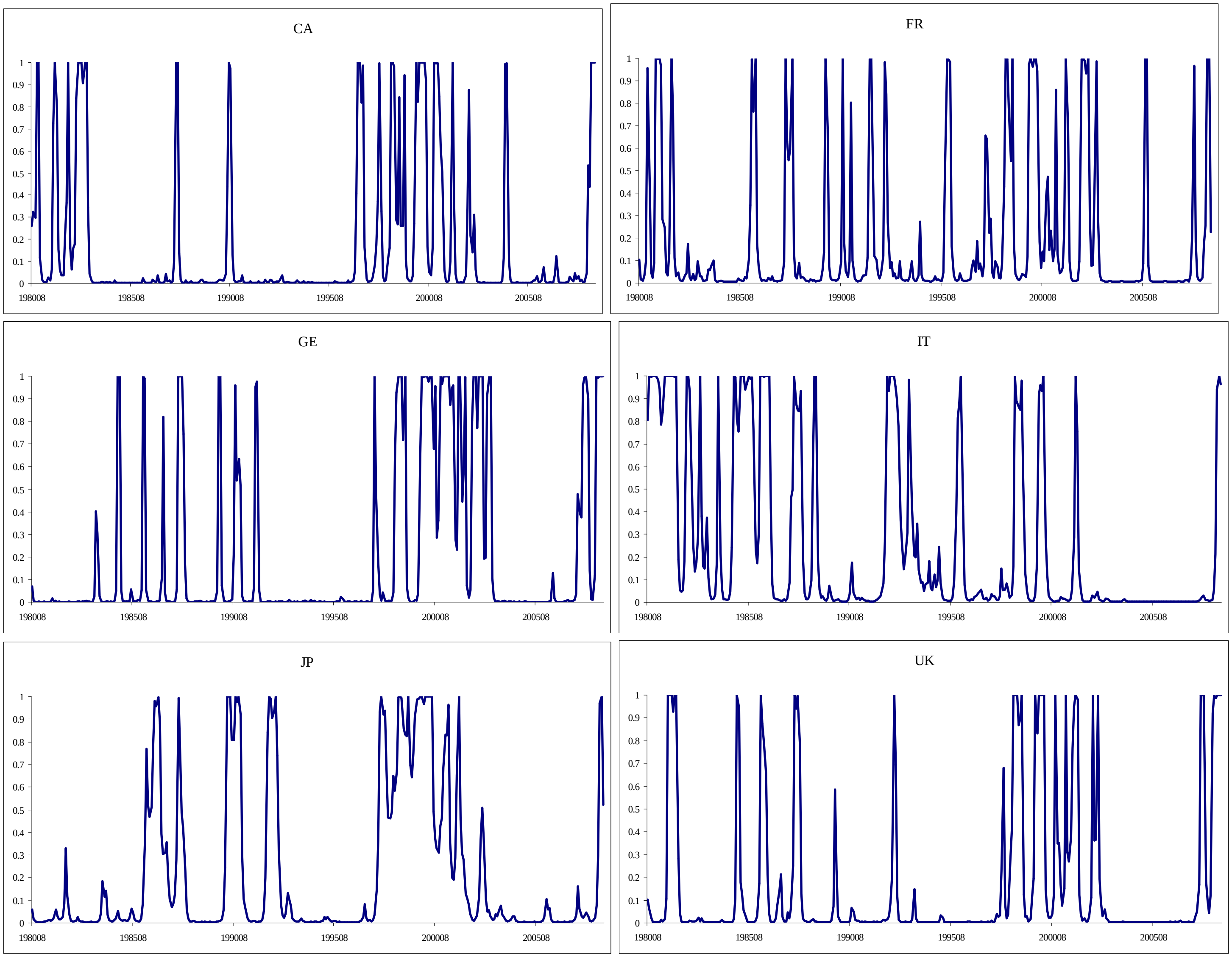


Figure 4. Recursive trend tests for the U.S.

This figure reports the t-dan test statistics for the U.S., which is estimated for a sample period starting in 1964:01 and ending between 1970:04 (the first high variance regime) and 2008:12. The variable σ^2_{CLMX} is the annualized aggregate firm level idiosyncratic variance computed using daily data, as defined in equation (2). The horizontal line at 2.05 represents the critical value for the trend test (t-dan test) to be significant at 5%.