Forthcoming, Review of Quantitative Finance and Accounting

Can rating agencies look through the cycle?

Gunter Löffler

Institute of Finance, University of Ulm, Helmholtzstrasse 18, 89069 Ulm, Germany.

Phone: ++49-731-5023597. Fax: ++49-731-5023950. e-mail: gunter.loeffler@uni-ulm.de

Abstract

Rating agencies claim to look through the cycle when assigning corporate credit ratings, which entails that they are able to separate trend components of default risk from transitory ones. To test whether agencies possess this competence, I take market-based estimates of one-year default probabilities of corporate bond issuers and estimate their long-run trend using the Hodrick-Prescott filter, local regression, or centered moving averages. I find that ratings help identify the current split into trend and cycle. In addition, rating stability is similar to the one of hypothetical ratings based on long-term trends. The results are robust to the use of different filter techniques. They are confirmed by a model-free analysis, which shows that ratings predict future changes in market-based default probability estimates. Since the examined trends are forward-looking in the sense that the trend filtering algorithms use future data, agency ratings exhibit important characteristics one would expect from ratings that see through the cycle.

Key words: default risk, credit ratings, through-the-cycle, Hodrick-Prescott filter, Merton model

JEL classification: G20, G33

1 Introduction

A key characteristic of corporate credit ratings produced by the major rating agencies is that they are meant to look through the cycle.[1] The agencies' aim is to base rating assignments on the expected long-term credit quality, and not respond to transitory fluctuations in credit quality (Basel Committee on Banking Supervision 2000; Cantor 2001; Standard and Poor's 2003). In justification of this approach, agencies cite investor preferences for stable ratings (Fons 2002). Such preferences can arise, for example, from rating-based portfolio governance rules which lead to high transaction cost if ratings change frequently.

As stressed by the agencies themselves, looking through the cycle is difficult.[2] This seems to be one of the reasons banks usually do not employ the through-the-cycle approach when assigning internal credit ratings for their borrowers (Treacy and Carey 1998). Academic research, too, is beset by complications in detecting long-term trends. It is difficult to identify and disentangle the transitory and permanent components of stock prices, for example, even when fifty years of data, or more, are available; and the evolution of the literature shows that the choice of appropriate statistical techniques is not obvious.[3]

The ability to rate through the cycle thus seems to be a core competence of rating agencies, even though there has not, until now, been an empirical study that examines whether agencies actually possess this competence. The lack of research is even more surprising since a skeptic might contend that rating through the cycle is nothing more than a marketing story contrived to cover a low-cost policy of reviewing ratings only infrequently. By classifying new information as transitory, agencies could avert frequent criticism for reacting too slowly to new information. The view that major rating agencies are too slow to change their ratings is held by a number of investors (Baker and Mansi 2002), academics (Daníelsson et al. 2001), and competitors (Egan and Jones 2002).

In this paper, I document that agency ratings exhibit key characteristics one would expect from through-the-cycle ratings. Before elaborating on this statement, I clarify the terminology because different terminologies and usages can give rise to confusion. The following terminologies are equivalent for the purpose of this paper:

Time series	=	trend component	+	cycle component
Time series	=	permanent component	+	transitory component

Another word for trend would be permanent component; the analogous term for cycle is transitory. An example is a series composed of a random walk plus an autoregressive process of order 1, AR(1); in this example, the stochastic trend of the random walk constitutes the trend of the aggregate series. How does the distinction between systematic and non-systematic risk components relate to this?

---

[1] Note that this paper deals with corporate credit ratings. Structured finance ratings, which feature prominently in the discussion of the subprime crisis, are assigned by different rating units based on a different rating approach. An assessment of rating quality should therefore be conducted separately for the two types of ratings. The methodology for sovereign ratings is more similar. Due to significant differences in the determinants of corporate and sovereign defaults, however, the conclusions of the paper should not be translated to sovereign ratings.

[2] Cf. Standard and Poor's (2003), pp.41-43.

[3] Cf. Campbell, Lo and MacKinlay (1997), pp. 78-80.

Cycles can be both systematic and non-systematic, the same holds for trends.[4] One might first think of macroeconomic or aggregate stock market cycles. However, the major part of a typical firm's risk is non-systematic (Campbell et al. 2001); consistent with this finding, the default risk series used in this paper are largely driven by non-systematic cycles and trends, too. I therefore refrain from using the distinction in the estimation and analysis of the series.

I start the empirical analysis with a robust, model-free analysis. I examine whether ratings are useful for predicting changes in EDFs, which are market-based estimates of one-year default probabilities. The results – ratings predict changes in EDFs over a horizon of up to 3 years – already indicate that ratings are forward-looking. Then, I set out to provide a more direct test of the agencies' ability to look through the cycle. Although the rating agencies do not give a precise, formal definition of what they aim to do when looking through the cycle, a characterization that captures its essence is to say that ratings are based on the trend component of a firm's short-term default risk (cf. Löffler 2004).[5]

The research strategy is as follows: I take EDFs and estimate their trend component using the Hodrick-Prescott filter (Hodrick and Prescott 1997), centered moving averages, or local regressions (Cleveland 1979). These three filter methods use both past and future data to identify the trend at a given date, thus seeing through the cycle. Irrespective of the filtering method, I find that agency ratings help predict the current state of the EDF trend. Also, the low variability of ratings is largely in line with the low variability of long-term trends. Summing up, ratings are not a crystal ball but they exhibit important characteristics one would expect from ratings that see through the cycle.

The remainder of this paper is structured as follows. The relevant literature is summarized in section 2. After a description of the data in section 3, section 4 examines whether ratings predict future changes in EDFs. Section 5 analyses the relationship between trends, cycles and ratings from different perspectives. Section 6 concludes.

2 Relevant literature

Conceptual aspects of the through-the-cycle methodology have been explored by Carey and Hrycay (2001) and Löffler (2004). Simulations in Löffler (2004) already suggest that key empirical characteristics of ratings, such as their low variability, *could* stem from a through-the-cycle approach. In a similar vein, Altman and Rijken (2004) quantify the *potential* effects of a through-the-cycle policy by benchmarking it against quantitative scoring models. The present paper documents the empirical relevance and validity of the possibility results derived by Löffler (2004) and Altman and Rijken (2004).

---

[4] Let me give an example for each combination: a change in the cyclical component of a firm's default risk would be systematic if it is associated with aggregate changes in profitability over the business cycle; a change in the trend component would be systematic if it is part of a permanent, market-wide increase in leverage. Non-systematic cycles can arise if firms respond to firm-specific shocks by rebalancing their capital structure (cf. Graham and Harvey 2001; Fama and French 2002; Flannery and Rangan 2006; Leary and Roberts 2005). If a firm decides to increase its leverage permanently without the average firm doing this at the same time, there would be a non-systematic change in the firm's trend component.

[5] One could argue that this statement is incomplete because it misses the stress scenario approach described in the literature. I discuss this aspect in section 5.2.

Over long horizons the default prediction power of ratings is higher than the one of common quantitative prediction models (Altman and Rijken 2006; Löffler 2007). Agency ratings therefore appear to provide added value, which is in contrast to evidence on the added value of another source of qualitative information, namely auditors' assessments (Sun 2006; Yip 2006). The fact that the performance of ratings increases with the horizon could reflect an ability to look through the cycle, but it could also be due to backward-looking smoothing. Similarly, findings that rating actions are procyclical (Nickell, Perraudin and Varotto 2000; Bangia et al. 2002; Amato and Furfine 2004) do not establish that agencies do not look or rate through the cycle. Business cycle fluctuations that coincide with permanent changes in credit quality could provide one explanation for such a pattern.[6] An example would be a recession that is sparked by increased international competition, which can permanently lower the profit margins of some firms even after the economy has adjusted to it and overcome the recession.

While rating agencies employ a through-the-cycle architecture, extant quantitative default prediction models can be described as following a point-in-time, or current-condition approach. They estimate the probability of default over a fixed horizon, which is usually one year, based on today's information. There are two main types of default prediction models: (i) statistical, or reduced-form approaches exemplified in the models of Bharath and Shumway (2008), Campbell, Hilscher and Szilagyi (2008) and Trustorff, Konrad and Leker (2011), and (ii) option-theoretic, or structural models based on Merton (1974). In statistical models, historical data are used to estimate the link between determinants of default risk and default probabilities. Typical candidates include accounting ratios such as leverage and profitability but also stock return volatility and stock prices. The second approach is based on the following insight: In a limited liability company, equity holders have a walk-away option, which will be exercised if the value of equity becomes negative. The walk-away option can be priced as a put-option on the firm's assets; the default probability is the probability that equity holders exercise this option. The classical Merton (1974) analysis has been generalized in many directions. For example, Geske (1977) and Geske and Johnson (1984) generalize the simplified structure of debt assumed in Merton (1974); Longstaff and Schwartz (1995) introduce stochastic interest rates; Leland and Toft (1996) endogenize capital structure decisions. Recent studies on the performance of structural default prediction models include Leland (2004), Bharath and Shumway (2008) and Campbell, Hilscher and Szilagyi (2008). A commercial implementation is the Expected Default Frequency (EDF) by Moody's KMV, which is an estimate of a firm's one-year probability of default based on a generalized Merton model. Details on the approach are given in Crosbie and Bohn (2003) and Kealhofer (2003). A key difference between EDFs and models implemented in academic papers is that Moody's KMV calibrates the model-based default probabilities to actual default frequencies.

The structure imposed in the implementation of Merton models can help to find the correct functional relationships and to limit estimation error present in statistical models. Inevitably, however, model assumptions will not be literally correct. Ultimately, it is therefore an empirical question how the forecasts from a given structural model compare to alternatives. EDFs have been shown to provide powerful default predictions relative to ratings and the Z-Score, a standard

[6] "In any case, purely cyclical factors are difficult to differentiate from coincident secular changes in industry fundamentals, such as the emergence of new competitors, changes in technology, or shifts in customer preferences" (Standard and Poor's 2003, p. 42).

statistical measure of default risk (Bohn, Arora and Korablev 2005). Bharath and Shumway (2008) and Campbell, Hilscher and Szilagyi (2008) conclude that the Merton model underlying the EDF is inferior to reduced-form statistical models. However, they do not test EDFs produced by Moody's KMV but standard Merton model implementations. The latter have been shown to underperform the former by Bohn, Arora and Korablev (2005). For the purpose of the present paper, what is needed are reliable measures of short-term default risk. Based on the evidence summarized above, state-of-the-art statistical models and EDFs both seem to serve the purpose well. For reasons of data availability, I conduct the analysis with Moody's KMV EDFs.

In the analysis of the next sections, the time series of a company's EDF will be detrended in order to separate permanent and transitory components. The literature on detrending is wide. Several contributions indicate problems with the Hodrick-Prescott filter, which is widely used in empirical macroeconomics. The studies document that detrending results are sensitive to the choice of filter, and that the Hodrick-Prescott filter can identify spurious trends (Harvey and Jaeger 1993; King and Rebelo 1993; Cogley and Nason 1995). Ehlgen (1998) and Pedersen (2001) give new support to the use of the Hodrick-Prescott filter by showing that even optimal (in the mean-squared error sense) filters can give rise to distortions, and that the Hodrick-Prescott filter is less distorting than other filters.

For the purpose of the present paper, the issue of statistically spurious trends does not seem critical, for the following reasons: (i) the reported results are robust to the use of different filter techniques; (ii) statistical tests typically have low power in identifying long-term cyclical patterns, while rating agencies might have used information beyond the time series data examined by the statistician. Rating agencies could be able to exploit predictability in circumstances where statistical tests do not find significant cyclical patterns and thus classify filtered trends as spurious. In such a situation, forward-looking trends also serve their purpose for the present paper. A forward-looking trend that is classified as spurious will nevertheless have the property that the difference between the trend and the current value of the series predicts subsequent movements of the series. In other words, it is not evident that a rating that looks through a possibly spurious cycle is less valuable to rating users than a rating that is assigned for a firm whose cycle withstands statistical tests.

Kisgen (2006) finds that firms' capital structure decisions are affected by their desire to meet rating targets. This may raise concerns about the interpretation of the results from the present paper. Firms that change their capital structure to maintain a rating tend to confirm the current rating assessment. If ratings are found to predict future default probabilities, some part should perhaps be credited to this self-fulfilling prophecy character of ratings rather than to a genuine forecasting ability. A look at the empirical findings indicates that this part is not likely to dominate. In Kisgen (2006, Table IX), ratings and contemporaneous control variables explain less than 3% of the variance of one-year capital structure changes. If there were a 100% correlation between capital structure changes and default probabilities, the fit from those regressions could thus explain 3% of the variation in default probabilities. In the present paper (Table 2), however, ratings and control variables explain 7% of the variance of one-year changes in default probabilities. The part that can be attributed to self-fulfilling prophecies is likely to be significantly less than 3/7 because (i) the correlation between capital structure and default risk is not 100%; (ii) Kisgen uses contemporaneous control variables which I do not; (iii) some of the capital structure changes explained in the Kisgen regressions are motivated by a desire to change the current rating, not to maintain it.

3 Data

The analyses use monthly data on Moody's long-term senior ratings and Moody's KMV Expected Default Frequencies (EDFs) for US and non-US corporate bond issuers. Issuers contained in the database are made up of the intersection of issuers which have a Moody's rating and traded equity. The data covers the period from 1980 to 2005, but because Moody's refined its rating system in 1982, I start the regression analyses in December 1982. Ratings enter the analysis as cardinal numbers from 1 (Aaa) to 21 (C). Empirical default rates rise exponentially with deteriorating grades (cf. Hamilton 2004, Exhibit 31) which is why I use logarithmic EDFs for the entire analysis. (When mentioning specific EDF values, the simple probabilities will nevertheless be used.)
Table 1 shows descriptive statistics on EDFs and ratings. Skewness and kurtosis of the two variables are fairly close. The simple conversion of ratings to numbers from 1 to 21 has been used by Warga and Welch (1993), Stohs and Mauer (1996) and Kisgen (2006). In the present paper, EDF information is explained through ratings. Therefore, the conversion is adequate if the relationship between ratings and log EDFs is (approximately) linear. To check the appropriateness of this assumption, I consider the following alternatives for modeling the relationship between log EDFs and ratings:

(i) Use a fifth order polynomial in the simple rating conversion. This is a data-driven approach to capture nonlinearities.
(ii) As in Stohs and Mauer (1996), two dummy variables are included along with the simple rating conversion: A dummy that is one if the rating is better than A1, and a dummy that is one if the rating is worse than B3. This is a parsimonious way of allowing for nonlinearities in the tails of the rating distribution.
(iii) A conversion of ratings to default probabilities. Specifically, I take Moody's idealized default probabilities from Yoshizawa (2003). As Yoshizawa does not specify default probabilities for rating grades Ca and C, I estimate them through a quadratic extrapolation. This leads to default probabilities of 59% and 79% for rating grades Ca and C, respectively.

Results reported in Table 1 show that the default probability conversion leads to a worse fit of EDFs (as measured by the adjusted R^2) than the simple numerical representation. Allowing nonlinearities in the latter does not greatly improve the fit. With a fifth-order polynomial, the R^2 rises only from 44.4% to 46.6%, and the predicted fit is non-monotonic for high EDF levels, which indicates overfitting. Adding the two dummy variables also increases the fit but the gain in the R^2 is only 1.2 percentage points. The simple conversion therefore appears to work relatively well. Nevertheless, I have also run the analyses of the next two sections based on approaches (ii) and (iii), respectively. Selected results for these alternative approaches will be presented.

4 Do rating agencies predict future changes in default probabilities?

Consider a process x_{t+1} that is made up of a random walk $z_{t+1} = z_t + v_{t+1}$ plus an AR(1) that follows $y_{t+1}=\rho y_t + u_{t+1}$, $0 < \rho < 1$. With such a process, knowing the current z_t and x_t helps explain future changes in x:

$$\begin{aligned}
&x_{t+1} - x_t = y_{t+1} + z_{t+1} - y_t - z_t = y_{t+1} - y_t + z_{t+1} - z_t = y_{t+1} - y_t + v_{t+1} \\
&= \rho y_t - y_t + u_{t+1} + v_{t+1} \\
&= (\rho - 1) y_t + \varepsilon_{t+1} = (\rho - 1)(x_t - z_t) + \varepsilon_{t+1} \\
&= (\rho - 1) x_t - (\rho - 1) z_t + \varepsilon_{t+1}
\end{aligned} \tag{1}$$

with $\varepsilon_{t+1} = u_{t+1} + v_{t+1}$. In the regression $x_{t+1} - x_t = \gamma_0 + \gamma_1 x_t + \gamma_2 z_t + \varepsilon_{t+1}$, we would therefore have $\gamma_1 = -\gamma_2 = (\rho - 1)$. If the z_t is only observed with error, the estimated γ coefficients will be biased away from $(\rho - 1)$, but the x_t and z_t will still explain future changes in x. The example illustrates that the ability to look through the cycle—i.e. the ability to identify the z —implies the ability to forecast the original series. This motivates a first, model-free test of whether rating agencies can look through the cycle, namely, check whether they can predict changes in future default probabilities. The correspondences to the example from above are that the EDF corresponds to x, and that the rating contains information about z.

As additional explanatory variable, I include the lagged change in EDFs. Including the lagged change is a simple way of capturing time series dependencies; we would expect the rating to add something to it. Specifically, I run the following regressions of changes in EDFs on ratings, EDFs, and EDF momentum:

$$EDF_{i,t+a} - EDF_{it} = \beta_0 + \beta_1 Rating_{it} + \beta_2 EDF_{it} + \beta_3 (EDF_{it} - EDF_{i,t-a}) + u_{it} \tag{2}$$

$$EDF_{i,t+a} - EDF_{i,t+a-1} = \beta_0 + \beta_1 Rating_{it} + \beta_2 EDF_{it} + \beta_3 (EDF_{it} - EDF_{i,t-a}) + u_{it} \tag{3}$$

An assumption implicit in these regressions is that the relation between the EDF trend (z_t) and ratings is linear; it is motivated by the analysis of the previous section, which has found that the relation between ratings and EDFs is approximately linear.

In the first regression, the dependent variable is the cumulative EDF change over a months. In the second it is the one-month change which is a months ahead. To make the statistical inference robust, I use in (2) only non-overlapping forecast horizons starting in December; forecast horizons in (3) are non-overlapping by construction. I also de-mean all variables by their cross-sectional means in order to reduce cross-sectional dependence and to bring the forecasting test closer to the meaning of a rating.[7] Ratings are assigned on a relative scale and meant to be stable on average (cf. Cantor 2001). They are not intended to vary with aggregate default risk, so they should not be expected to predict the average trend in EDFs. Finally, I correct the standard errors for clustering within a given forecast horizon

[7] If the stock market declines, for example, the EDFs will on average go up, resulting in the error terms being correlated across firms. Running the regressions without de-meaning does not change the conclusions. Note that de-meaning also purges the data of the effects of aggregate stock market fluctuations.

as well as within issuers.[8] An open question is how to deal with firms that default within the prediction horizon. I examine the following two approaches: (i) use the EDF recorded in the database or set it to the maximum of 20% used by Moody's KMV in case there is no EDF in the database (ii) set the EDF to 100% in case of a default.

Table 2 shows the results for approach (i). Ratings significantly predict EDFs over up to 36 months. The positive coefficient is consistent with the example in equation (1), as rating information corresponds to z, and $-(\rho - 1) > 0$ for $\rho<1$. As for answering the question of how far agencies look into the future it is important to note that this observation holds for cumulative changes as well as for one-month changes. (One might well be able to predict 36-month cumulative changes just on the basis of predicting the next 6 months.) The predictive quality seems to fade, but relatively slowly. For the sake of brevity, I do not report results for approach (ii) because coefficients and t-statistics for the rating variable are consistently higher, strengthening the conclusions drawn from approach (i). The negative coefficient on EDFs is consistent with the analysis in equation (1), as EDFs correspond to x, and $(\rho - 1) < 0$ for $\rho<1$. The coefficient on the twelve-month EDF difference varies with the horizon. It indicates that there is short-run momentum but long-run mean reversion in EDFs, which is consistent with the behavior of stock returns. In order to check whether results are robust to the use of the simple rating conversion (Aaa=1, C=21) I run the analysis with the default probability conversion described in the previous section. For the regressions from Panel A of Table 2, results are reported in Appendix A. The predictive power of ratings is marginally smaller; apart from that, there are no major differences. Further analysis shows that adding the dummy variables for very good grades (better than A1) as well as for very bad grades (worse than B3) improves the fit, but only slightly; the maximum increase in the R^2 is 0.5 percentage points.

A possible concern is that the results could be influenced by Moody's acquisition of KMV in April 2002. Though Moody'sKMV was run as a separate business segment throughout the sample period of this paper, there might have been some alignment of information usage. In a sensitivity analysis, I limit the data used in the regressions to observations before January 2002.[9] As is evident from Appendix B, which reports selected results, the predictive power in the pre-2002 period does not differ noticeably from the entire sample period. In another variation reported in Appendix B, I replace Moody's ratings by long-term credit ratings from Standard and Poor's.[10] Again, there is almost no change in the results, suggesting that the results do not depend on the choice of rating agency. I also re-run Table 2 using only Moody's ratings for non-US issuers. Again, there are no conspicuous changes in coefficients.

How do the results compare to findings that rating changes can be predicted by market-based information (e.g. by a measure that is close to EDFs as in Delianedis and Geske, 2003)? Moody's claims that it deliberately slows down the adjustment of ratings to new information in order to increase rating stability (Cantor 2001). This can lead to a lagged response to public information and will tend to reduce any predictive content that the ratings have, but it need not reduce the latter to zero.

[8] Firm-specific clusters also take account of firm-specific serial correlation. The estimator is implemented using the procedure discussed and provided by Petersen (2009).

[9] For example, the last 36 month forecast horizon entering the analysis is from December 1998 to December 2001.

[10] Data is from S&P Compustat, for US firms only.

Summing up, the regressions establish that ratings meet an important requirement that should be met if they look through-the-cycle. The fact that their predictive ability extends over three years corroborates the interpretation. Since a cycle can cover several years, the interpretation would be questionable if we had found that the rating agency predicts EDFs only over a few months. Of course, the regressions provide an indirect test and do not formally prove that ratings actually look through the cycle.[11] In the next section, I will use filter methods to extract trends and cycles from the EDF series. Relative to the reduced-form regression approach of this section, filtering requires assumptions on filter techniques and also leads to a loss of observations, because short series or series ending in default cannot be filtered in a meaningful way. However, it will allow more specific tests of the agencies' ability to look through the cycle.

5 An analysis of ratings and EDF trends and cycles

5.1 Trend extraction techniques and trend properties

I use three methods for identifying a long-term EDF trend from the time series of EDFs: a simple moving average, local regression, and the Hodrick-Prescott filter (Hodrick and Prescott 1997). My choice of techniques and parameters is meant to document robustness. I implement two Hodrick-Prescott filters that differ greatly in the intensity of smoothing; then I select a moving average (local regression) that is close to the Hodrick-Prescott trend with low (high) smoothing intensity. Hence, the results can show robustness with respect to smoothing intensity, as well as with respect to the technique used to achieve a given smoothing intensity.
Applied to a series of EDFs for issuer *i*, the Hodrick-Prescott filter splits it into a trend-component *HPTREND* and a cyclical component *HPCYCLE*:

$$EDF_{it} = HPTREND_{it} + HPCYCLE_{it} \tag{4}$$

The split is determined by minimizing (separately for each issuer *i*, which has data from $t_0(i)$ to $T(i)$):

$$\sum_{t=t_0(i)}^{T(i)} (HPCYCLE_{it})^2 + \lambda \sum_{t=t_0(i)+2}^{T(i)} \left[(HPTREND_{it} - HPTREND_{i,t-1}) - (HPTREND_{i,t-1} - HPTREND_{i,t-2}) \right]^2 \tag{5}$$

where λ is a smoothing constant. In the minimization of (5), λ determines the weight deviations from the trend receive relative to variation in the trend. The larger λ is, the smoother is the estimated trend. Results are reported for two alternative choices, λ=10,000 and λ=500,000. For quarterly (monthly) macroeconomic data, a choice commonly made in the literature is λ=1600

[11] It cannot be ruled that the influence of the rating variable is affected by endogeneity. For example, if ratings and EDFs are two different measures of the underlying credit quality, errors in measuring this credit quality could lead to endogeneity. This would not invalidate the conclusion that ratings are useful for predicting EDFs. It could even be linked with a through-the-cycle interpretation in the sense that ratings look through error cycles in EDFs. The standard econometric response – instrumental variables – does not appear feasible due to the lack of suitable variables.

(λ=14,400). One reason for choosing more extreme smoothing constants than are typically seen in the literature is to demonstrate robustness. A further justification for the large value of λ=500,000 is that EDFs, driven by stock prices, are more volatile than the macroeconomic data on which the Hodrick-Prescott filter is commonly applied.[12]

Smoothing EDFs by local regression involves regressing EDFs on time separately for each issuer-month. The subset of observations that is used for a local regression extends over a bandwidth that has to be specified by the researcher. It is centered except for the end points, where uncentered subsets are used. I choose to report results for a bandwidth of eight years (96 months). For firm *i* with observations from $t = t_0(i),\ldots T(i)$, the smoothed value at time τ is the prediction for $t = \tau$ from the regression

$$EDF_{it} = b_0 + b_1 t + u_{it}, \qquad t = \max(\tau - 48, t_0(i)),\ldots,\ \min(\tau + 48, T(i)) \qquad (6)$$

Finally, I report results for a centered 37-month moving average computed over months -18 to +18. In contrast to the local regression described above, the moving average is computed only if observations over the entire 37-month interval are available.

The three filter techniques are all forward-looking but differ in the way future data is used. The Hodrick-Prescott filter uses the entire series; the local regression is run over a fixed interval that is truncated at the end points; the moving average is computed over a fixed interval that is not truncated. Unreported analyses show that varying the parameters does not lead to qualitatively different or unexpected results. With the Hodrick-Prescott filter, for example, results for λ=50,000 lie between those for λ=10,000 and λ=500,000.

Several sample selection decisions are related to the identification of trends. Default is a special situation in which the normal trend and cycle is left; I thus split the time series on the occasion of default. Observations before the default month are treated as a separate series; observations after a default are discarded until the firm emerges from default. After emergence, observations are again treated as a separate series.[13] If there are missing observations, I treat stretches of data interrupted by missing observations as separate series.

Regardless of the filter methodology, the empirical analyses of the later sections use only observations where EDFs for the prior and future 18 months are available. The Hodrick-Prescott filter and the local regression smoother can be computed for all observations including the first and the last, but the closer one is to the end of a series, the less future information is available for smoothing. This diminishes the forward-looking character of the trend estimate, which is important for the purpose of this study. Requiring 18 months of future observations is meant to ensure a forward-looking property that is sufficient in the sense that it is not easy to achieve by an analyst. Finally, I disregard series shorter than 48 months in order to reduce situations in which the filters yield implausible results because of a lack of cyclical fluctuations.[14] Numbers for the selection process are as follows: starting with 4,244 companies, split-ups due to missing observations and defaults

---

[12] Pedersen (2001) also recommends λ >100,000 for monthly data.

[13] Conclusions do not change if defaults are ignored and trends are computed over all observations of a company. I define emergence from default as an upgrade to B3 or better.

[14] Consider a Hodrick-Prescott trend computed over one sinusoidal cycle. Rather than being horizontal, it is a downward sloping line. This trend line is as smooth as a horizontal line, but the squared deviations from the sinusoid are smaller. Many short series are similar in that they contain just one peak and one trough.

increase the number of series by 1,907. The requirement that series length should be at least 48 months reduces the number of series to 2,643. The mean length of the series in this sample is 120 months (median 98).

To get an indication of the smoothing properties of the methods described above, I calculate for each series the variance of the trend in EDF and the variance of the EDF itself. The lower the ratio of these two variances, the more variability has been removed by the filter. As reported in Table 3 average ratios range from 32% to 50% and show that the Hodrick-Prescott trend with λ=10,000 is similar to the 37-month moving average, while the Hodrick-Prescott trend with λ=500,000 is similar to the trend obtained through local regression.[15] Table 3 also reports R^2s from regressions of the form $x_{it} = b_0 + b_1\bar{x}_t + u_{it}$, i.e. a variable is regressed on its cross-sectional means. The lower the R^2 from such a regression the lower the importance of systematic, marketwide components. The R^2s for ratings, EDFs and EDF trends are all below 10%; for EDF cycles they stay below 25%. This means that the data are predominantly driven by non-systematic trends and cycles. Thus, the identification of firm-specific trends and cycles is essential to a successful through-the-cycle rating.

Figure 1 shows the time series of EDFs, ratings and the two Hodrick-Prescott trends (λ=10,000 or λ=500,000) for four firms. In the first chart, credit quality exhibits cyclical fluctuations around a slowly-moving trend. The second chart shows that an EDF trend can remain flat for years despite large fluctuations in EDFs (here between 0.25% and 3.5%). The final two charts are for two defaulting firms, Fruit of the Loom and Enron; they end in the month before default. The charts illustrate that the perfect-foresight character of the Hodrick-Prescott trend is reduced towards the end of the series (which justifies excluding the first and last 18 months from the later analysis).

Table 4 complements the visual analysis through correlations of EDFs, ratings and filtered EDFs; Pearson correlations (below diagonal) and Spearman rank correlations (above diagonal) exhibit very similar patterns. Panel A contains the correlation computed across all observations. Consistent with the agencies' practice of following a through-the-cycle approach, ratings display larger correlations with the filtered EDF trends than with original EDFs. The correlation becomes greater with increasing smoothing intensity. When moving from pooled correlation to within-time correlation (Panel B), which is correlation computed after subtracting the variables' cross-sectional means, the correlation between ratings and EDF based variables increase. This is what we expect because ratings give a relative ordering and are not intended to capture aggregate fluctuations in default risk. For the same reason, correlations are generally lower if correlation is computed after subtracting the series-specific means (panel C) as this removes much of the relative differences in default risk from the data. In both panels B and C, EDF trends continue to be more strongly correlated with ratings than EDFs. I again check whether results are robust to the default probability transformation of ratings described in section 3. As is evident from Appendix A, in which the pooled correlations for that approach are shown, correlation patterns are the same as in Table 4.

[15] Standard deviations of variance ratios are high because they reach high values (>1) in some cases. The reason is that the sample selection excludes observations at the start and end of the series. Though excluded from the final analysis they are used in the computation of trends and their variability. If the variability of EDFs in the middle of the series is lower than at the start and end one can observe low EDF variability but high trend variability.

To better gauge the magnitude of the differences in correlation coefficients, I conduct bivariate linear regressions of the type $Rating_{it} = a + b \cdot Predictive\_Variable_{it} + u_{it}$, where the predictive variable is either the EDF or one of the estimated trends analyzed in Table 4.[16] Table 5 gives information on absolute prediction errors $|\hat{u}_{it}|$. If $|\hat{u}_{it}|$ is smaller than 0.5, for example, a rating derived from OLS (if the prediction is rounded to the nearest integer) would be identical to the actual rating. When the predictive variable is the EDF, the prediction error is in 62.4% of all cases smaller or equal to two 2.5 rating grades. If the same is done with the HP-trend with λ=500,000, 69.3% of all predictions are at most 2.5 grades away from the actual rating. Such differences appear to be significant.

5.2 Empirical analysis of ratings, trends and cycles

5.2.1 Do ratings predict trends in EDFs?

I start with the following regression based on three-month differences:

$$TREND_{i,t+3} - TREND_{it} = \beta_0 + \beta_1 (EDF_{i,t+3} - EDF_{it}) + \beta_2 (Rating_{i,t+3} - Rating_{it}) + u_{it} \quad (7)$$

where the dependent variable TREND is the trend in EDFs determined through one of the smoothing methods. The coefficient β_2 shows whether ratings help identify the trend in EDFs after controlling for the EDF; other control variables can and will be added. Since the dependent variable of (7) is not the future realization of some variable, the regression is not a classical forecasting analysis. However, it examines the forecasting ability of ratings because the left-hand side variable is based on forward-looking trends and hence contains information about the future. Note that I use three-month differences instead of levels to avoid possible concerns about non-stationarities.[17] Six-month or twelve-month differences do not lead to qualitative changes in results.

To elucidate the nature and the value of the agency's predictive ability, it seems important to include additional control variables. Viewed from a particular date, the trend is obtained by smoothing over the past and the future. If rating agencies simply smoothed over the past, they could help explain the current state of the trend without being able to \`foresee' future developments. In addition, the predictive quality of rating agencies could result from the exploitation of trends or trend reversals in the EDF series. Especially in the case of short-term trends, this could easily be replicated with a statistical model.

In an extended regression, I thus control for backward-looking smoothing and short term trends or reversals. Specifically, I include rolling estimates of the trend, which use only information available at time *t*, denoted by $TREND_t|t^-$, as well as

[16] I also produced predictions based on ordered logit or ordered probit regressions. Results are very similar.

[17] Even though ratings and EDFs are bounded between 1 and 21 and 0.02% and 20%, respectively, one cannot rule out that trending behavior within the bounds affects the distribution of t-statistics. Regressions using levels are available upon request; they strengthen the conclusions drawn from the analysis using differences.

the twelve-month difference in EDFs.[18] Following the procedure from section 4, I de-mean all regression variables by their cross-sectional means as this brings the analysis in line with the relative nature of ratings. To obtain robust standard errors, I use only non-overlapping quarterly data and allow for clustering within quarters as well as companies.

For the Hodrick-Prescott filter, Kaiser and Maravall (1999) and Meyer and Winker (2005) document the possibility of spurious cross-correlations induced by filtering. Applied to this paper, it means that finding an influence of ratings on trends after controlling for EDFs could be spurious. To take possible effects of spurious correlations into account, I use a bootstrap study to simulate critical *t*-values. The bootstrap is structured as follows:

(1) For each series *i* contained in the analysis, randomly select (with replacement) another series out of those which include the time span covered by series *i*. Replace the rating values of series *i* by the rating values of the randomly picked series *j*.
(2) Run the regression whose *t*-statistic is to be simulated and record the *t*-statistic associated with the rating variable.
(3) Repeat (1) to (2) 1,000 times.

I report the 97.5% percentile of the simulated t-statistics as the simulated critical value.

Regression results are reported in Table 6. Simulated critical *t*-statistics (5% significance) for the rating variable are close to the expected value of 1.96, indicating that the problem of spurious trends is not important in the data set here. Ratings contribute significantly to the explanation of EDF trend, even after controlling for backward-looking trends. The only exception is the extended regression based on the moving average trend, where the rating variable loses significance upon inclusion of the backward-looking trend, which is here the uncentered moving average computed over the preceding 18 months. This, however, should not come as a surprise. Viewed in *t*, the change in the 37-months moving average from *t* to *t*+3 is completely determined by months *t*-18 to *t*-16 and months *t*+19 to *t*+21. The case is different with the other smoothing methods, because their change is determined by the interplay of past, current, and future observations. Therefore, changes in the backward-looking moving averages do a relatively good job of explaining the subsequent change of the centered moving average, and leave less room for other variables, including ratings. In Table 6, this is evident in the relatively high coefficient and *t*-statistic of the backward-looking moving average trend. Results are again robust to the way in which rating information is coded. When ratings are coded based on a default probability transformation (see Appendix A), the resulting coefficients on the rating variable are very similar to the ones in Table 6.

The evidence thus suggests that ratings look through the cycle as they contain information that is valuable for identifying the trend. Whether the predictive quality is economically significant is difficult to judge because there is no straightforward benchmark. The trends computed here are based on perfect foresight, so we cannot expect agencies to predict them perfectly or to a large extent. If a rating changes by three grades, e.g. from Aa2 to A2, the predicted change in the log HPI trend with λ=10.000 is 6.9% (= coefficient × ΔRating =

[18] If the trend is computed with the Hodrick-Prescott Filter, for example, $TREND_t|t^-$ is obtained by applying the Hodrick-Prescott filter (2) to observations $t_0(i),\ldots t$ instead of $t_0(i),\ldots,T(i)$.

0.023 × 3). This is not large but note that the correlation analysis from Table 4 has shown that most of the information contained in ratings is about relative ordering. When we examine three-month differences as we do here in order to increase statistical robustness, the bulk of this information is not used. If regressions are run on levels rather than on three-month differences, ratings therefore have a stronger impact on ratings. Unreported analyses show that in such regressions, a two standard-deviation difference in ratings leads to a difference in the EDF trend of up to 81%, which is sizeable.[19]

5.2.2 Stability of ratings, EDFs and EDF trends

One striking empirical characteristic of agency ratings is their considerable stability. On a one year time horizon, typically 80% to 90% of ratings remain stable, compared to 40% to 50% for categories based on EDFs (Kealhofer 2003) or other short-term forecasts of default probabilities (Carey and Hrycay 2001). To compute stability measures for the present data I choose fixed EDF cutoffs to assign grades based on EDFs or EDF trends. They are set in line with average default rates of letter-only rating classes (cf. Hamilton 2004). For example, EDFs larger than 0.02% (the minimum value of EDFs) and smaller or equal to 0.06% are assigned to one grade; stability of this grade is then compared to the stability of Moody's category Aa. The same ranges are used to classify EDF trends into grades.

Results are shown in Table 7. The reported stability is the number of observations which have grade *i* at the start and the end of a calendar year divided by the number of all observations which have grade *i* at the start of the year *and* which have EDF trends available at the end of the year or defaulted within the year. Availability of EDF trends is subject to the criteria defined in section 5.1. For ratings and EDFs, the figures correspond to the stylized facts cited above. The interesting finding is that stability of grades based on the Hodrick-Prescott trend with λ=500,000 or the local regression is close to the stability of ratings. From Table 4, these two trends have the highest correlation with ratings, suggesting that they provide a good approximation of what rating agencies have in mind when rating through the cycle. The results empirically support the findings of Löffler (2004), who performed simulations based on plausible parameter choices for the credit quality process and concluded that the through-the-cycle approach might explain a large part of the stability of agency ratings.

5.2.3 Can ratings be described through stress scenarios?

Carey and Hrycay (2001) and Löffler (2004) associate through-the-cycle ratings with a stress scenario approach. A stress scenario is an unfavorable cyclical deviation from the long-term trend that occurs with a low probability. The probability of large cyclical deviations can be approximated based on the standard deviation of the cyclical component. The rating can therefore be described as $TREND_{it} + m \cdot \sigma(CYCLE_i)$, where TREND is an EDF trend, σ(CYCLE) is the

[19] The regressions are of the form $TREND_{it} = \beta_0 + \beta_1 EDF_{it} + \beta_2 Rating_{it} + u_{it}$.

standard deviation of CYCLE = EDF – TREND, and *m* is a multiplier related to the extremeness of the stress scenario. If ratings follow EDFs for firms that are already in stress, this leads to the following definition of a stress scenario measure:

$$\max\left[TREND_{it} + m\sigma\left(CYCLE_i\right),\ EDF_{it}\right],\quad m > 0. \tag{8}$$

Alternatively, I neglect the move towards EDFs for firms that are in stress and examine

$$TREND_{it} + m\sigma\left(CYCLE_i\right),\quad m > 0. \tag{9}$$

If these two definitions provide a good description of the agencies' rating method they should have a higher correlation with ratings than the underlying trend. For *m*=1.64 and *m*=2.33, corresponding to (nominal) stress scenario probabilities of 5% and 1%, respectively, Table 8 shows that this is not clearly visible in the data. Often, correlations with ratings decrease when moving from the EDF trend to the stress scenario measures defined above. If they rise, the increase is small, especially when compared to the increase in correlation when moving from EDFs to EDF trends (cf. Table 4).

Possibly, the analysis fails to capture relevant aspects of the scenario approach. For several reasons, however, it seems plausible that the role of stress scenarios is indeed limited. First, rating agencies do not ascribe particular importance to the stress scenario concept. In recent years, Moody's has published several statements meant to clarify its rating policy (e.g. Cantor 2001; or Fons 2002). The analysis of stress scenarios is either not mentioned, or it is described as just one element of the rating process: "Fundamental credit analysis (…) seeks to predict the credit performance of bonds, other financial instruments, or firms across a range of plausible economic scenarios, some of which will include credit stress" (Fons 2002, p. 5). Second, Löffler (2004, fn. 9) reports for his simulations that the key results do not depend on whether ratings are defined with or without stress scenarios. In the present data, finally, the standard deviation of cyclical components varies so little across firms that it would not strongly change rating assignments even if rating agencies put more weight on it. The standard deviation of $\sigma(CYCLE_i)$ averages 0.15 across the four different filters, while the average standard deviation of the variable TREND is 1.43.

6 Summary and conclusion

Using a broad data set, I have shown that agency ratings contribute to the identification of long-term trends in market-based estimates of short-term default probabilities. The long-term trends have been identified by Hodrick-Prescott filtering, local regression or by taking centered moving averages. A significant part of this contribution is predictive in the sense that it adds information to statistical filters that are merely backward-looking. The paper thus lends empirical support to the agencies' claim that they follow a through-the-cycle rating concept, and that they have expertise in applying this concept. It also shows that the relatively large stability of ratings, which is often interpreted as an indicator of inefficient response to new information, could be due to the through-the-cycle methodology.

When assessing rating quality, an important aspect is the performance of ratings in the prediction of defaults. Prior literature has already documented that ratings provide valuable information, especially over horizons longer than one year (Altman and Rijken 2006; Löffler 2007). One could surmise that this is simply due to backward-looking smoothing, which could easily be reproduced through statistical models. The results of this paper indicate, however, that the informational value of ratings rests on a forward-looking ability.

While these findings support the usefulness of ratings, investors should be aware that, by construction, through-the-cycle ratings may underperform other predictors when it comes to short-term default prediction (cf. Löffler 2004). For the same reason, the results of this paper are consistent with evidence that ratings underperform relative to alternative measures of default risk or are slow to adjust their ratings. The short-term underperformance and the long-term information content of through-the-cycle ratings are two sides of the same coin. When assessing rating quality it is therefore essential to distinguish between the question of how well rating agencies fulfill their own objectives (e.g. looking through the cycle), and how well their objectives are aligned with those of rating users (e.g. short-term risk management vs. long-term investment). In the practical usage of ratings, statistical adjustments of rating information could help to overcome possible misalignments of objectives. For example, investors focusing on short-term risk could employ a method similar to Butera and Faff (2006) to convert through-the-cycle ratings into short-term ratings.

Acknowledgments

I am grateful to Moody's Investors Service for providing data used in this paper. I wish to thank the anonymous referee as well as seminar participants at the annual meeting of the German Finance Association, the GRETA credit conference, the Federal Reserve Board, and the universities of Frankfurt and Mannheim for their many helpful comments. Parts of the paper were written while the author was enjoying the hospitality of the College of Management at NC State.

References

1. Altman EI, Rijken HA (2004) How rating agencies achieve rating stability. J Bank Financ 28: 2679-2714
2. Altman EI, Rijken HA (2006) A point-in-time perspective on through-the-cycle ratings. Financ Analysts J 62: 54-70
3. Amato JD, Furfine CH (2004) Are credit ratings procyclical? J Bank Financ 28: 2641-2677
4. Baker HK, Mansi SA (2002) Assessing credit rating agencies by bond issuers and institutional investors. J Bus Finance Account 29: 1367-1368
5. Bangia A, Diebold FX, Kronimus A, Schagen C, Schuermann T (2002) Ratings migration and the business cycle, with application to credit portfolio stress testing. J Bank Financ 26:445-474
6. Basel Committee on Banking Supervision (2000) Credit ratings and complementary sources of credit quality information. Basel
7. Bharath ST, Shumway T (2008) Forecasting default with the Merton distance to default model. Rev Financ Stud 21:1339-1369
8. Bohn JR, Arora N, Korablev I (2005) Power and level validation of the EDF credit measure in the U.S. market. Working Paper, Moodys KMV
9. Butera G, Faff R (2006) An integrated multi-model credit rating system for private firms. Rev Quant Finan Acc 27:311–340
10. Campbell JY, Hilscher J, Szilagyi J (2008) In search of distress risk. J Finance 63:2899-2939
11. Campbell JY, Lettau M, Malkiel BG and Xu Y (2001) Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. J Finance 56:1-43
12. Campbell JY, Lo A, MacKinlay C (1997) The econometrics of financial markets. Princeton University Press, Princeton, NJ
13. Cantor R (2001) Moody's investors service response to the consultative paper issued by the Basel Committee on Banking Supervision and its implications for the rating agency industry. J Bank Financ 25:171-186
14. Carey M, Hrycay M (2001) Parameterizing credit risk models with rating data. J Bank Financ 25:171-270
15. Cleveland WS (1979) Robust locally weighted regression and smoothing scatterplots. J Am Stat Assoc 74:829-837
16. Cogley T, Nason JM (1995) Effects of the Hodrick-Prescott on trend and difference stationary time series. Implications for business cycle research. J Econ Dynamics Control 19:253-278
17. Crosbie P, Bohn J (2003). Modeling Default Risk, Moody's KMV Technical Paper
18. Daníelsson J, Embrechts P, Goodhart C, Keating C, Muennich F, Renault O, Shin HS (2001) An academic response to Basel II. Working paper, London School of Economics
19. Delianedis G, Geske R (2003) Credit risk and risk neutral default probabilities: Information about rating migrations and defaults, EFA 2003 Annual Conference Paper No. 962
20. Egan SJ, Jones WB (2002) Statement of Egan Jones on credit rating agencies. Securities and Exchange Commission. http://www.sec.gov/news/extra/credrate/ eganjones2.htm
21. Ehlgen J (1998) Distortionary effects of the optimal Hodrick-Prescott filter. Econ Lett 61:345-349
22. Fama EF, French KR (2002) Testing tradeoff and pecking order predictions about dividends and debt. Rev Financ Stud 15:1-33
23. Flannery MJ, Rangan KP (2006) Partial adjustment toward target capital structure. J Financ Econ 79:469-506
24. Fons J (2002) Understanding Moody's corporate bond ratings and rating process. Special Comment, Moody's Investors Service
25. Geske R (1977) The valuation of corporate liabilities as compound options. J Financ Quant Analysis 12: 541-552
26. Geske R, Johnson HE (1984) The valuation of corporate liabilities as compound options: A correction. J Financ Quant Analysis 19:231-232
27. Graham JR, Harvey CR (2001) The theory and practice of corporate finance: Evidence from the field. J Financ Econ 60:187-243
28. Hamilton DT (2004) Default & recovery rates of corporate bond issuers. Moody's Investors Service

29. Harvey AC, Jaeger A (1993) Detrending, stylized facts, and the business cycle. J Appl Econometrics 8:231-247
30. Hodrick RJ, Prescott EC (1997) Postwar U.S. business cycles: an empirical investigation. J Money Credit Bank 29:1-16
31. Kaiser R, Maravall A (1999) Estimation of the business cycle: a modified Hodrick-Prescott filter. Span Econ Rev 1:175-206
32. Kealhofer S (2003) Quantifying credit risk I: Default prediction. Financ Analysts J 59:30-44.
33. King RG, Rebelo ST (1993) Low frequency filtering and real business cycles. J Econ Dynamics Control 17:207-231
34. Kisgen D (2006) Credit ratings and capital structure. J Finance 61:1035-1072
35. Leary MT, Roberts M (2005) Do firms rebalance their capital structure? J Finance 60:2575-2619
36. Leland HE (2004) Predictions of default probabilities in structural models of debt. J Invest Manag 2:5-20
37. Leland HE, Toft KB (1996) Optimal capital structure, endogenous bankruptcy, and the term structure of credit spreads, J Finance 51:987-1019
38. Löffler G (2004) An anatomy of rating through the cycle. J Bank Financ 28:695-720
39. Löffler G (2007) The complementary nature of ratings and market-based measures of default risk. J Fixed Income 17:38-47
40. Longstaff FA, Schwartz ES (1995) A simple approach to valuing risky debt. J Finance 50: 789-821
41. Merton RC (1974) On the pricing of corporate debt: The risk structure of interest rates. J Finance 29:449-470
42. Meyer M, Winker P (2005) Using HP filtered data for econometric analysis: some evidence from Monte Carlo simulations. Allg Stat Arch 89:303-320
43. Nickell P, Perraudin W, Varotto S (2000) Stability of rating transitions. J Bank Financ 24:203-227
44. Pedersen TM (2001) The Hodrick-Prescott filter, the Slutzky effect, and the distortionary effect of filters. J Econ Dynamics Control 25:1081-1101
45. Petersen MA (2009) Estimating standard errors in finance panel data sets: comparing approaches. Rev Financ Stud 22:435-480
46. Standard & Poor's (2003) Corporate ratings criteria. Standard & Poor's
47. Stohs MH, Mauer D (1996) The determinants of corporate debt maturity structure. J Bus 69:279-313
48. Sun L (2007) A re-evaluation of auditors' opinions versus statistical models in bankruptcy prediction. Rev Quant Finan Acc 28:55–78
49. Treacy W, Carey M (1998) Credit risk rating at large US banks. Federal Reserve Bulletin, 898-921
50. Trustorff JH, Konrad PM, Leker J (2011) Credit risk prediction using support vector machines. Rev Quant Finan Acc 36:565–581
51. Warga A, Welch I (1993) Bondholder losses in leveraged buyouts. Rev Financ Stud 6:959-982
52. Yip A (2006) Business failure prediction: A case-based reasoning approach. Rev Pac Basin Financ Mark Policies 9, 491–508
53. Yoshizawa Y (2003) Moody's approach to rating synthetic CDOs. Special comment, Moody's Investors Service

Figure 1: EDFs, EDF trends (HP=Hodrick-Prescott) and ratings for four companies

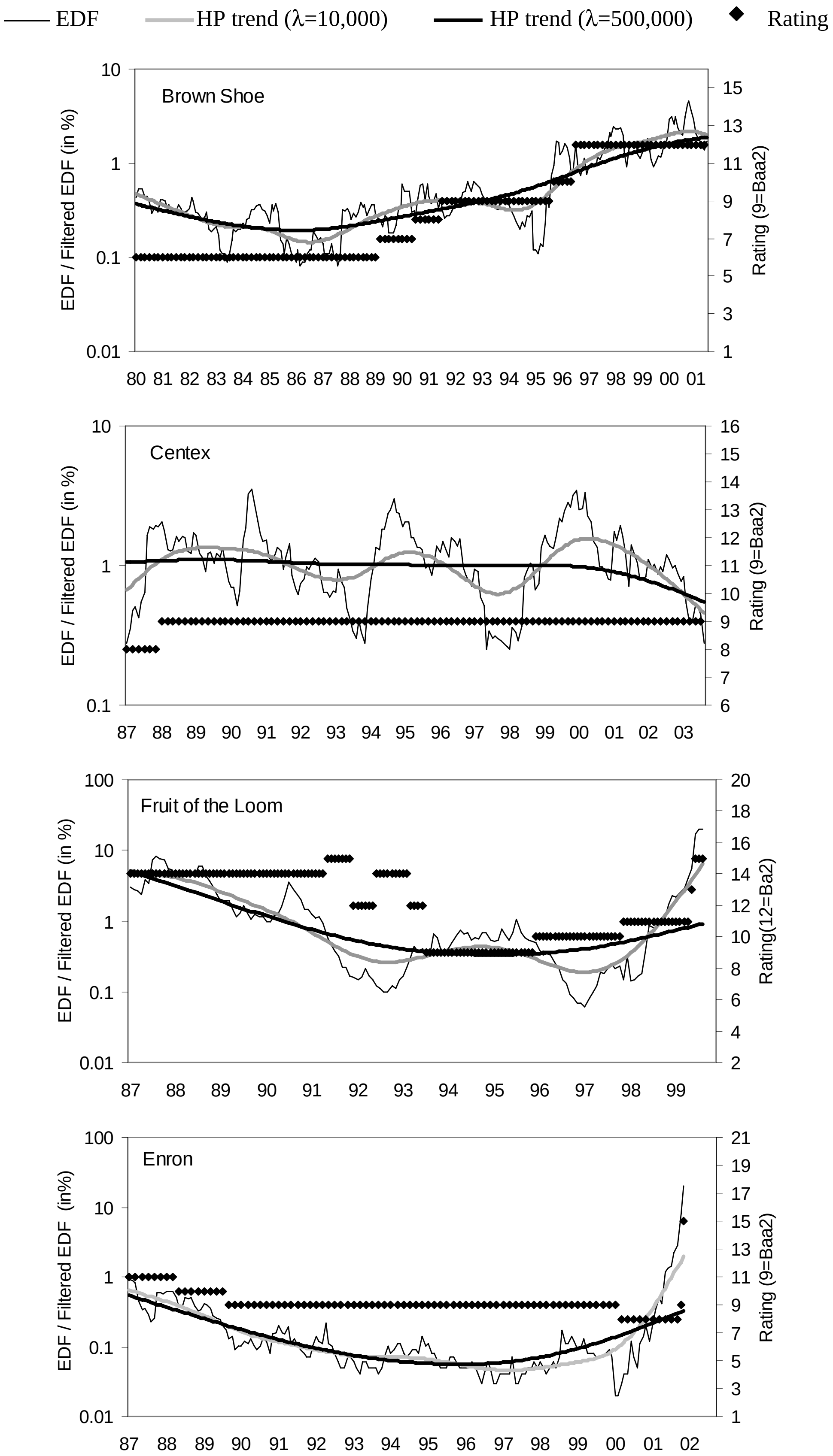

Table 1: Descriptive statistics for ratings and EDFs

Panel A: Univariate statistics

	Mean	Median	St.dev.	Skewness	Kurtosis
EDF	-0.96	-1.05	1.70	0.28	2.62
Rating	9.39	9.00	4.00	0.13	2.27

Panel B: Adjusted R^2 from a regression of EDFs on

	Adj. R^2
Rating	0.444
Rating, $Rating^2$, $Rating^3$, $Rating^4$, $Rating^5$	0.466
Rating, $I_{Rating<5}$, $I_{Rating>16}$	0.456
PD(Rating)	0.412

Notes: The sample includes all observations not in default. The number of observations is 356,436 firm-months. EDF is Moody's KMV expected default frequency (in logs), Rating is Moody's rating (Aaa=1, C=21), PD(Rating) is Moody's rating represented by the logarithms of idealized one-year default probabilities stated in Yoshizawa (2003). $I_{Rating<5}$ is one for ratings better than A1, and zero otherwise; $I_{Rating>16}$ is one for ratings worse than B3, and zero otherwise.

Table 2: Do ratings predict changes in EDFs? – regressions of EDF-changes on rating and EDF information

	Prediction horizon *a*				
	6 months	12 months	18 months	24 months	36 months
Panel A: Explaining cumulative changes in EDFs ($EDF_{i,t+a} - EDF_{it}$)					
Rating	0.033	0.061	0.072	0.083	0.109
	(6.27)	(6.57)	(10.53)	(6.29)	(5.89)
EDF	-0.109	-0.208	-0.246	-0.270	-0.317
	(-7.82)	(-7.56)	(-12.05)	(-6.97)	(-6.89)
EDF_t-EDF_{t-a}	0.055	0.084	-0.007	-0.078	-0.186
	(3.60)	(3.26)	(-0.24)	(-1.44)	(-7.22)
Adj. R^2	0.040	0.072	0.081	0.106	0.158
NOB	50835	22173	12385	8000	3906
Panel B: Explaining future one-month changes in EDFs ($EDF_{i,t+a} - EDF_{i,t+a-1}$)					
Rating	0.005	0.003	0.003	0.002	0.001
	(5.80)	(8.90)	(7.21)	(4.83)	(1.93)
EDF	-0.017	-0.013	-0.011	-0.007	-0.004
	(-7.74)	(-13.30)	(-10.39)	(-6.41)	(-3.10)
EDF_t-EDF_{t-a}	0.010	-0.001	-0.005	-0.005	-0.002
	(4.42)	(-0.86)	(-4.11)	(-4.18)	(-1.55)
Adj. R^2	0.006	0.004	0.004	0.002	0.001
NOB	304655	261141	222849	190669	140035

Notes: The results are based on regressions of the form

Panel A: $$EDF_{i,t+a} - EDF_{it} = \beta_0 + \beta_1\, Rating_{it} + \beta_2\, EDF_{it} + \beta_3\,(EDF_{it} - EDF_{i,t-a}) + u_{it}$$

Panel B: $$EDF_{i,t+a} - EDF_{i,t+a-1} = \beta_0 + \beta_1\, Rating_{it} + \beta_2\, EDF_{it} + \beta_3\,(EDF_{it} - EDF_{i,t-a}) + u_{it}$$

where Rating is Moody’s rating (Aaa=1, C=21) and EDF is Moody’s KMV expected default frequency (in logs). Regression use only observations which are non-overlapping in the dependent variable. T-values (in parentheses) are corrected for heteroscedasticity and clustering within companies as well as observations belonging to the same time period [t,t+a].

Table 3: Characteristics of EDF trends used in the empirical analysis: Smoothing intensity and role of systematic components

Panel A: Variance[$TREND_i$] / variance[EDF_i]			
	Mean	Median	Standard deviation
Hodrick-Prescott trend (λ=10,000)	0.503	0.483	0.701
Hodrick-Prescott trend (λ=500,000)	0.321	0.193	0.779
Moving average trend	0.500	0.449	1.127
Local regression trend	0.323	0.192	0.803
Panel B: R^2 from pooled regression of a variable on its monthly cross-sectional means			
	R^2 for raw series	R^2 for trend	R^2 for cycle=EDF – Trend
Rating	0.018		
EDF	0.097		
Hodrick-Prescott (λ=10,000) on EDF		0.087	0.160
Hodrick-Prescott (λ=500,000) on EDF		0.041	0.219
Moving average of EDF		0.079	0.172
Local regression on EDF		0.037	0.220

Notes: The lower the variance ratios in Panel A, the higher the smoothing intensity. The lower the R^2 in Panel B, the lower the role of systematic components. EDF is Moody's KMV expected default frequency (in logs), the moving average is computed over 37 months; the local regression with an eight-year bandwidth. The number of series *i* used in Panel A is 2643, the number of observations used in Panel B is 215,326.

Table 4: Correlations between ratings, EDFs and EDF trends

	Rating	EDF	HP TREND (λ=10,000)	HP TREND (λ=500,000)	MA TREND	LR TREND
Panel A (pooled observations)						
Rating	1	0.649	0.684	0.716	0.690	0.713
EDF	0.673	1	0.965	0.908	0.958	0.893
HP TREND (λ=10,000)	0.709	0.967	1	0.963	0.999	0.949
HP TREND (λ=500,000)	0.740	0.916	0.968	1	0.970	0.998
MA TREND	0.714	0.960	0.999	0.974	1	0.957
LR TREND	0.737	0.902	0.956	0.997	0.964	1
Panel B (within-time)						
Rating	1	0.672	0.709	0.735	0.713	0.732
EDF	0.695	1	0.967	0.920	0.961	0.907
HP TREND (λ=10,000)	0.731	0.969	1	0.972	0.999	0.961
HP TREND (λ=500,000)	0.755	0.927	0.975	1	0.977	0.997
MA TREND	0.735	0.963	0.999	0.980	1	0.967
LR TREND	0.752	0.916	0.966	0.998	0.971	1
Panel C (within-series)						
Rating	1	0.205	0.231	0.237	0.231	0.231
EDF	0.279	1	0.873	0.699	0.849	0.642
HP TREND (λ=10,000)	0.311	0.891	1	0.864	0.995	0.806
HP TREND (λ=500,000)	0.318	0.730	0.880	1	0.875	0.984
MA TREND	0.310	0.871	0.996	0.892	1	0.818
LR TREND	0.311	0.682	0.835	0.989	0.848	1

Notes: Rating is Moody's rating (Aaa=1, C=21), EDF is Moody's KMV expected default frequency (in logs), HP TREND is the Hodrick-Prescott trend in EDFs with smoothing parameter λ; MA TREND denotes the centered moving average of EDFs computed over 37 months; LR TREND is obtained through local regression with an eight-year bandwidth. Entries below the diagonal are Pearson correlation coefficients, entries above are Spearman rank correlation coefficients. The number of observations is 215,326.

Table 5: Error distribution in bivariate rating prediction

	Predictive variable				
Absolute prediction error	EDF	HP TREND (λ=10,000)	HP TREND (λ=500,000)	MA TREND	LR TREND
≤ 0.5	13.7%	14.7%	16.2%	14.9%	16.1%
≤ 1.5	40.4%	42.6%	45.6%	43.1%	45.6%
≤ 2.5	62.4%	65.6%	69.3%	66.2%	69.2%
≤ 3.5	79.5%	82.2%	84.3%	82.5%	84.0%

Notes: The Table summarizes the distribution of absolute regression errors $|\hat{u}_{it}|$ when Moody's rating (Aaa=1, C=21) is predicted in a bivariate regression of the form: $\text{Rating}_{it} = a + b \cdot \text{Predictive\_Variable}_{it} + u_{it}$. EDF is Moody's KMV expected default frequency (in logs), HP TREND is the Hodrick-Prescott trend in EDFs with smoothing parameter λ; MA TREND denotes the centered moving average of EDFs computed over 37 months; LR TREND is obtained through local regression with an eight-year bandwidth.

Table 6: Do rating changes explain changes in EDF trends?

	Dependent variable:							
	ΔHP TREND (λ=10,000)		ΔHP TREND (λ=500,000)		ΔMA TREND		ΔLR TREND	
ΔEDF_t	0.087	0.035	0.027	0.002	0.063	0.029	0.022	0.010
	(32.96)	(14.10)	(18.50)	(1.12)	(21.78)	(12.26)	(14.15)	(5.63)
$\Delta Rating_t$	0.023	0.004	0.014	0.005	0.023	0.000	0.014	0.007
	(15.76)	(5.15)	(14.14)	(8.51)	(15.62)	(-0.04)	(11.97)	(10.14)
simulated t (5%)	*2.09*	*2.05*	*1.95*	*1.94*	*2.03*	*2.06*	*2.00*	*1.96*
$\Delta TREND_t \mid t^-$		0.215		0.193		0.387		0.076
		(21.93)		(19.19)		(33.18)		(6.34)
EDF_t - EDF_{t-12}		0.026		0.000		0.008		0.010
		(11.25)		(0.51)		(4.47)		(6.97)
Adj. R^2	0.14	0.49	0.05	0.27	0.08	0.46	0.03	0.13
NOB	69729	69729	69729	69729	67588	67588	69729	69729

Notes: Δ denotes 3-month differences. Rating is Moody's rating (Aaa=1, C=21), EDF is Moody's KMV expected default frequency (in logs), HP TREND is the Hodrick-Prescott trend in EDFs with smoothing parameter λ; MA TREND denotes the centered moving average of EDFs computed over 37 months; LR TREND is obtained through local regression with an eight-year bandwidth. TREND | t^- is computed with the same method as the dependent variable, but uses only information up until the current date t. Regressions use only non-overlapping observations. T-values (in parentheses) are corrected for heteroscedasticity and clustering within companies as well as within time. Simulated critical *t*-statistics for the rating variable are obtained through a bootstrap in which rating series are reshuffled across observations.

Table 7: One-year stability of ratings and EDF-based categories

Rating / EDF-Range(%)	Rating	EDF	HP TREND (λ=10,000)	HP TREND (λ=500,000)	MA TREND	LR TREND
Aa / 0.02-0.06	0.89	0.36	0.73	0.87	0.75	0.86
A / 0.06-0.12	0.92	0.32	0.51	0.75	0.53	0.73
Baa / 0.12-0.5	0.90	0.59	0.77	0.89	0.79	0.89
Ba / 0.5-2.5	0.84	0.60	0.75	0.85	0.77	0.84
B / 2.5-15	0.83	0.52	0.73	0.80	0.80	0.80

Notes: The table shows the fraction of firms whose rating grade stayed constant over one year. Rating is Moody's rating, EDF is Moody's KMV expected default frequency, HP TREND is the Hodrick-Prescott trend in EDFs with smoothing parameter λ; MA TREND denotes the centered moving average of EDFs computed over 37 months; LR TREND is obtained through local regression with an eight-year bandwidth.

Table 8: Correlations of ratings with EDF trends and stress scenario measures of default risk

Correlation of ratings with	Trend method			
	HP (λ=10,000)	HP (λ=500,000)	MA	LR
TREND	0.709	0.740	0.714	0.737
max(TREND + 1.64 σ(CYCLE), EDF)	0.713	0.732	0.716	0.725
max(TREND + 2.33 σ(CYCLE), EDF)	0.713	0.720	0.714	0.712
TREND + 1.64 σ(CYCLE)	0.714	0.732	0.716	0.725
TREND + 2.33 σ(CYCLE)	0.713	0.744	0.718	0.741

Notes: Rating is Moody's rating (Aaa=1, C=21), EDF is Moody's KMV expected default frequency (in logs), HP denotes the Hodrick-Prescott filter with smoothing parameter λ; MA a centered moving average of EDFs computed over 37 months; LR trend is obtained through local regression with an eight-year bandwidth. CYCLE is EDF – TREND. The number of observations is 215,326.

Appendix A: Selected results when rating information is captured through the log of the associated idealized default probability

Panel 1: Results from regressions as in Table 2, Panel A

	Prediction horizon *a*				
	6 months	12 months	18 months	24 months	36 months
Panel A: Explaining cumulative changes in EDFs ($EDF_{i,t+a} - EDF_{it}$)					
Rating	0.041	0.076	0.086	0.100	0.125
	(5.96)	(6.16)	(9.13)	(5.53)	(4.79)
EDF	-0.102	-0.194	-0.226	-0.248	-0.280
	(-7.65)	(-7.29)	(-11.38)	(-6.51)	(-5.87)
EDF_t-EDF_{t-a}	0.052	0.078	-0.014	-0.085	-0.196
	(3.40)	(3.02)	(-0.52)	(-1.56)	(-7.10)
Adj. R^2	0.038	0.068	0.076	0.100	0.149
NOB	50835	22173	12385	8000	3906

Panel 2: Results from correlation analysis as in Table 4, Panel A

	Rating	EDF	HP TREND (λ=10,000)	HP TREND (λ=500,000)	MA TREND	LR TREND
Panel A (pooled observations)						
Rating	1	0.649	0.684	0.716	0.690	0.713
EDF	0.653	1	0.965	0.908	0.958	0.893
HP TREND (λ=10,000)	0.689	0.967	1	0.963	0.999	0.949
HP TREND (λ=500,000)	0.720	0.916	0.968	1	0.970	0.998
MA TREND	0.694	0.960	0.999	0.974	1	0.957
LR TREND	0.716	0.902	0.956	0.997	0.964	1

Panel 3: Coefficients on $\Delta Rating_t$ *from regression analysis as in Table 6*

$\Delta Rating_t$	0.030	0.006	0.018	0.006	0.029	0.000	0.018	0.009
	(13.73)	(4.75)	(12.76)	(7.73)	(13.49)	(0.09)	(11.01)	(9.19)

Appendix B: Sensitivity analysis of the regressions in Table 2

	Prediction horizon *a*				
	6 months	12 months	18 months	24 months	36 months
Data up to 2001, coefficient on rating when explaining cumulative EDF changes ($EDF_{i,t+a} - EDF_{it}$),					
Rating	0.038	0.069	0.076	0.089	0.116
	(6.40)	(6.76)	(9.68)	(6.82)	(5.17)
Data up to 2001, coefficient on rating when explaining one-month EDF changes ($EDF_{i,t+a} - EDF_{i,t+a-1}$)					
Rating	0.006	0.004	0.003	0.002	0.001
	(5.48)	(8.81)	(7.31)	(4.71)	(1.51)
S&P ratings, coefficient on rating when explaining cumulative EDF changes ($EDF_{i,t+a} - EDF_{it}$),					
Rating	0.031	0.059	0.070	0.076	0.087
	(8.94)	(8.52)	(7.52)	(5.75)	(4.10)
S&P ratings, coefficient on rating when explaining one-month EDF changes ($EDF_{i,t+a} - EDF_{i,t+a-1}$)					
Rating	0.005	0.003	0.003	0.002	0.001
	(7.41)	(7.86)	(6.76)	(4.29)	(2.50)
Moody's non-US ratings, coefficient on rating when explaining cumulative EDF changes ($EDF_{i,t+a} - EDF_{it}$),					
Rating	0.036	0.064	0.075	0.096	0.134
	(4.94)	(5.17)	(10.56)	(8.10)	(6.76)
Moody's non-US ratings, coefficient on rating when explaining one-month EDF changes ($EDF_{i,t+a} - EDF_{i,t+a-1}$)					
Rating	0.005	0.004	0.003	0.002	0.0004
	(4.56)	(7.96)	(5.69)	(3.50)	(0.53)

Notes: The table reports β_1 coefficients from regressions of the form

$$EDF_{i,t+a} - EDF_{it} = \beta_0 + \beta_1\, Rating_{it} + \beta_2\, EDF_{it} + \beta_3\,(EDF_{it} - EDF_{i,t-a}) + u_{it}$$

$$EDF_{i,t+a} - EDF_{i,t+a-1} = \beta_0 + \beta_1\, Rating_{it} + \beta_2\, EDF_{it} + \beta_3\,(EDF_{it} - EDF_{i,t-a}) + u_{it}$$