Incorporating the Dynamics of Leverage into Default Prediction

Gunter Löffler, Alina Maurer*

Abstract

A firm's current leverage ratio is one of the core characteristics of credit quality used in statistical default prediction models. Based on the capital structure literature which shows that leverage is mean reverting to a target leverage we forecast future leverage ratios and include them in the set of default risk drivers. The analysis is done with a discrete duration model. Out-of-sample analysis of default events two to five years ahead reveals that the discriminating power of the duration model increases substantially when leverage forecasts are included. We further document that credit ratings contain information beyond the one contained in standard variables but that this information is unrelated to forecasts of leverage ratios.

Field: Financial Economics

Keywords: Default prediction, discrete duration model, leverage targeting, mean reversion, credit rating

JEL-Classification: G32, G33

Acknowledgments: This research was supported by the Deutsche Forschungsgemeinschaft (DFG) through a grant and trough the SFB 649 "Economic Risk". We thank Moody's Investors Service for providing rating and default data.

* Corresponding author: Alina Maurer, Institute of Finance, University of Ulm, 89069 Ulm, Germany. Phone: 0049-731-5023591, email: alina.maurer@uni-ulm.de

1 Introduction

Many observations suggest that firms follow leverage targets. A survey of chief financial officers conducted by Graham and Harvey (2001) reveals that the majority of firms has a target debt-equity ratio. Among others, Fama and French (2002), Leary and Roberts (2005) and Flannery and Rangan (2006) document that empirical leverage ratios are mean reverting. Estimation is usually based on a partial adjustment model in which the optimal capital structure depends on a set of firm characteristics. The results differ regarding the speed of adjustment and the relative importance of targeting behavior, but the finding that firms actively rebalance capital structure in order to close the gap between the current and the targeted leverage appears robust. The results are consistent with the trade-off theory of capital structure in which firms balance the costs and benefits of debt to determine an optimal leverage ratio. Rival theories include the pecking order theory (Myers (1984)) and the market timing theory (Baker and Wurgler (2002)).

If firms follow leverage targets then future leverage ratios should be predictable. In this paper, we examine whether such predictability can be exploited to improve the accuracy of statistical default prediction models. In such models, future defaults are explained with a set of variables including accounting ratios and stock market information. A variable that measures a firm's current leverage is usually included and found to have significant explanatory power.

Our analysis uses two approaches. In a reduced-form approach, we simply extend default risk drives by firm characteristics used to model targeted leverage ratios. In a two-step approach, forecasts for future leverage ratios are derived from a partial adjustment equation in the first step. We then extend a standard default prediction model by a new variable that incorporates these forecasts.

For the two-step approach we find that the accuracy of default prediction can increase significantly when modeling the dynamics of leverage. The added value from incorporating leverage dynamics increases if we extend the length of the default prediction horizon from one to five years. This is consistent with the empirical finding that it takes several years to close the gap between the current and the target leverage ratio.[1] Our results should help improve default prediction models, which are essential ingredients to risk management and pricing in many financial institutions. The results also add a new argument to the discussion of the empirical relevance of leverage-targeting behavior. It would be difficult to maintain the interpretation that targeting behavior is relevant if it were irrelevant in a setting such as default prediction where leverage plays a crucial role.

While statistical default prediction models are commonly used in practice, capital market participants also rely heavily on credit ratings issued by rating agencies such as Fitch, Moody's or Standard & Poor's. According to rating agencies, these ratings have a long-term horizon and are not only based on a firm's current situation but also on projections of its future situation. According to Moody's, for instance, "credit ratings are ordinal measures of through-the-cycle expected loss. As such, while they are certainly

[1] The half-life of deviations from optimal capital structure is five years in Fama and French (2002) and two years in Flannery and Rangan (2006).

based on the current financial strength of the issuer, they incorporate expectations of future performance as well" (Metz and Cantor (2006)). It therefore seems interesting to examine whether rating analysts incorporate information about the dynamics of leverage ratios. If we include the rating as a predictive variable in our default prediction model we find that it is significant but does not affect the significance of the variable that captures leverage dynamics. Thus, while ratings contain valuable information for default prediction, this information is largely unrelated to the leverage predictions that can be derived from the targeting behavior of firms.

Our research combines approaches from two areas: the capital structure literature introduced above and the literature on default prediction. In the latter, a firm's forward default probability within a given time horizon conditional on survival until the beginning of that horizon is modeled as a function of firm specific covariates such as leverage, profitability or liquidity. Some studies additionally include industry or economy-wide factors. Most recent contributions employ discrete duration models or, equivalently, multi-period logit models. They include Shumway (2001), Chava and Jarrow (2004), Hillegeist et al. (2004), Beaver et al. (2004) and Campbell et al. (2006). Often, these studies choose a one-year horizon for default prediction but longer horizons are also used, e.g. in Campbell et al. (2006).

The two papers which are closest to ours are Duffie et al. (2006) and Lo et al. (2008). Both model the dynamics of leverage through time series models inspired by the pricing model of Collin-Dufresne and Goldstein (2001). Model parameters such as target leverage or speed of adjustment are either estimated within a first-order vector autoregressive time series model (Duffie et al. (2006)) or directly taken from the capital structure literature (Lo et al. (2008)). Both studies conclude that modeling the dynamics of leverage increases prediction accuracy. The key difference to our approach is that we model target leverage through a set of firm-specific variables as suggested by the literature on capital structure.

The rest of the paper is organized as follows. In Section 2 we outline the statistical methodology for default analysis as well as the partial adjustment model. We also describe our approach for incorporating leverage dynamics into default prediction. Section 3 describes the data. Econometric issues relevant for the partial adjustment model are discussed in Section 4. The main results are reported in Section 5. Section 6 examines whether credit ratings incorporate information on mean reverting leverage ratios. Section 7 concludes.

2 Methodology

We begin with an overview on the discrete duration model for default analysis. This is followed by a short introduction to the partial adjustment model and its uses for forecasting leverage ratios.

2.1 Default Analysis

2.1.1 Single-Period Default Prediction

We use a discrete duration model. At time t, a firm's conditional probability of default within the next year given the firm's current condition and survival until the beginning of year t is expressed by a discrete hazard function $h(t)$:

$$h(t) = P(Y_{t,t+1} = 1 | Y_{t-1,t} = 0, X_t),$$

where the vector X_t captures the firm's current condition and the binary variable $Y_{t,t+1}$ indicates default in the time interval $(t, t+1]$, i.e.

$$Y_{t,t+1} = \begin{cases} 1 \text{ if default occurs in } (t, t+1] \\ 0 \text{ else.} \end{cases}$$

It has been pointed out by Shumway (2001) and also shown by Allison (1982) and Jenkins (1995) that a discrete duration model is equivalent to a multi-period logit model estimated on a panel data set if the hazard function is specified as

$$h(t) = \frac{1}{1 + \exp(-\alpha(t) - \beta' X_t)}.$$

Consequently, the maximum likelihood function is identical to the one of a binary logistic regression and coefficient estimates can easily be obtained using standard statistical software. In our analysis, we adopt the above specification of the hazard rate, since it has widely been used in the present literature. Among others it has been applied by Shumway (2001), Chava and Jarrow (2004), Hillegeist et al. (2004) and Campbell et al. (2006).

A question that needs to be addressed here is the specification of the baseline hazard function $\alpha(t)$. In a qualitative response model the probability of default is modeled as

$$P(Y_{t,t+1} = 1 | Y_{t-1,t} = 0, X_t) = \frac{1}{1 + \exp(-\alpha - \beta' X_t)}. \tag{1}$$

This implies that the baseline hazard function is constant over time ($\alpha(t) = \alpha$, $\forall t$) which is equivalent to the hypothesis of duration independence in the data. In other words, risk of default is assumed as being independent to the length of time a firm did exist (survive) before default analysis. Yet, if there is duration dependence in the data, the baseline hazard rate should be modeled by an appropriate function in t or, as suggested by Beck et al. (1998), by including temporal dummy variables α_t that index the length of a spell for each firm.[2] Beck et al. (1998) indicate that adding temporal dummy variables when

[2] Shumway (2001) controls for duration dependence by specifying $\alpha(t) = \ln(age)$, with *age* being the number of years a firm has been listed on the particular stock exchange, but finds this variable insignificant. Hillegeist et al. (2004) approximate $\alpha(t)$ by the mean rate of defaults twelve months prior to t. This however only captures economy wide influences, and does not control for duration dependence.

they are not required might introduce unnecessary multicollinearity and hence lead to incorrectly estimated standard errors. A likelihood ratio test is applicable for the null hypothesis of duration independence (i.e. $\alpha_t = 0$, $\forall t$). If the null is rejected, all temporal dummy variables are to be included in the hazard rate model. For the sample used in our analysis, the null hypotheses cannot be rejected. We therefore specify the probability of default as in equation (1). Detailed results are presented in Section 5.

2.1.2 Multi-Period Default Prediction

So far we have only considered default prediction one period ahead providing an answer to the question "What is the probability of a firm to default within the next year if we observe today's covariates?". The methodology illustrated in the previous section is not restricted to single-period default prediction. For default events multiple periods ahead we consider what we will call *discrete multi-period default probability*, meaning that the focus at time t is on the estimation of default probabilities within a one-period time frame k periods from t. In other words, the question addressed here is "What is the probability of default between year $t+k$ and $t+k+1$ given today's covariates?". Again this question is only reasonable for individuals that did not default prior to $t+k$. Discrete multi-period default probabilities are essentially lagged single-period default probabilities, i.e.

$$\begin{aligned} h(t+k) &= P(Y_{t+k,t+k+1} = 1 | Y_{t+k-1,t+k} = 0, X_t) \\ &= \frac{1}{1+\exp(-\alpha_k - \beta_k' X_t)}. \end{aligned} \tag{2}$$

Hereby we explicitly allow the coefficients β and α to vary with the prediction horizon.

From another point of view default prediction multiple periods ahead is a cumulative one. In such a setting the length of one period is enlarged. An analyst might prefer this perception if she is interested in the probability that a firm will default within the next k years, i.e. in the time span $(t, t+k]$, conditioning only on X_t and survival until year t:

$$P(Y_{t,t+k} = 1 | Y_{t-1,t} = 0, X_t).$$

We refer to this measure as *cumulative multi-period default probability.* As pointed out by Campbell et al. (2006) this distribution is no longer logit if we assume that discrete default probabilities follow a logistic distribution. By the Bayesian rule a cumulative multi-period default probability can be deduced from discrete default probabilities via:

$$P(Y_{t,t+k} = 1 | Y_{t-1,t} = 0, X_t) = 1 - \prod_{j=1}^{k-1} (1 - h(t+j)).$$

In the present paper, however, we restrict the analysis to discrete default probabilities.

2.2 Mean Reverting Leverage Ratios

Empirical literature on capital structure has widely analyzed the pecking order theory versus the alternative trade-off model. Various studies provide evidence for the hypothesis that leverage is mean reverting, cf. Shyam-Sunder and Myers (1999), Howakimian et al. (2001), Fama and French (2002) and most recently Flannery and Rangan (2006). In this context the dynamics of leverage is described by a partial adjustment model:

$$L_{t+1} - L_t = a_0 + a_1(L^*_{t+1} - L_t) + \epsilon_t, \tag{3}$$

where L_t denotes the leverage ratio observed at time t and L^*_{t+1} is an optimal and hence targeted leverage ratio for the next period. According to the partial adjustment hypothesis, each year a firm closes a fraction of a_1 of the gap between its actual and desired future leverage level. The targeted leverage ratio is the optimal value that sets off the benefits of interest tax shields against the costs of financial distress. Following Flannery and Rangan (2006) the adjustment speed factor a_1 is modeled constant over time and individuals, representing an average firm's adjustment speed.[3]

The major difficulty inherent in the partial adjustment model is the fact that L^* cannot be observed at any time and the researcher faces the question of finding an adequate approximate value for this variable.[4] Howakimian et al. (2001) and Fama and French (2002) model the targeted leverage ratio as a linear combination of observable covariates, i.e.

$$L^*_{t+1} = b_0 + b_1' C_t + \epsilon_t, \tag{4}$$

with C_t being a vector of firm specific characteristics observable at time t. With an appropriate approximate value for L^*_{t+1} at hand, estimation of future leverage ratios becomes feasible after rearranging equation (3) to

$$L_{t+1} = a_0 + a_1 L^*_{t+1} + (1 - a_1)L_t + \epsilon_t. \tag{5}$$

The estimate $\hat{L}_{t+1}$ can then be incorporated into the default prediction process; we will expose this strategy in more detail in Section 2.3.

2.2.1 Forecasting Leverage Ratios One Period Ahead

Forecasts about future leverage ratios can be determined in a two-step procedure (Fama and French (2002)). In a first step an estimate of the optimal leverage ratio is obtained from regression (4), where for estimation purposes L^*_{t+1} is approximated by the future leverage ratio:

$$L_{t+1} = b_0 + b_1' C_t + \epsilon_t.$$

The optimal leverage level in equation (5) can then be replaced by the predicted value $\hat{L}_{t+1}$. Alternatively, as has been described in Flannery and Rangan (2006), a simultaneous estimation of equations (4) and (5) is feasible. Replacing the targeted leverage ratio

[3] We well consider modifications of this assumption in Section 5.3.

[4] A rather simple proxy is the historic mean leverage ratio for a firm or within the respective industry, as used by Shyam-Sunder and Myers (1999).

in (5) by the assumed relation from (4) results in

$$\begin{aligned} L_{t+1} &= a_0 + a_1 b_0 + a_1 b_1' C_t + (1 - a_1) L_t + \epsilon_t \\ &= c_0 + c_1' C_t + c_2 L_t + \epsilon_t, \end{aligned} \tag{6}$$

with $c_0 = a_0 + a_1 b_0, c_1' = a_1 b_1'$ and $c_2 = (1 - a_1)$. In unreported results we find that there are only minor differences in estimation outcomes from the simultaneous or the two-step approach. We prefer the simultaneous estimation due to the intuition that results are less influenced by estimation errors and hence would yield more accurate forecasts.

2.2.2 Forecasting Leverage Ratios Multiple Periods Ahead

From the partial adjustment model it is straightforward to obtain estimates for leverage ratios one period ahead, as is shown in equation (6). In order to extend the approach to further horizons, we modify the partial adjustment model with respect to the length of the time period. At time t, we consider the adjustment to the desired leverage ratio in k years, where the target is set to be the optimal leverage ratio k years ahead given today's firm characteristics. More precisely, an estimate of the targeted leverage ratio at time $t + k$ can be found by regressing the respective observed future values on the vector of explanatory variables, i.e.

$$L_{t+k} = b_0 + b_1' C_t + \epsilon_t.$$

For each horizon k, future leverage ratios can then be estimated through:

$$L_{t+k} = a_0 + a_1 \hat{L}_{t+k} + (1 - a_1) L_t + \epsilon_t,$$

or simultaneously through:

$$L_{t+k} = c_0 + c_1' C_t + c_2 L_t + \epsilon_t. \tag{7}$$

2.3 Incorporating Leverage Dynamics Into Default Prediction

A central question of this paper is whether and to what extent results on the dynamics of capital structure are of benefit in default analysis, in particular for an assessment of long-term credit worthiness. Leverage is one of the key determinants of default risk (cf. Beaver et al. (2004)), and predictive knowledge about leverage should help improve default prediction.

Future leverage ratios can be estimated based on the present leverage ratio L_t and a set of firm specific covariates C_t (cf. equation (7)). This implies that the information on dynamics of leverage can be incorporated into default prediction by extending the vector of covariates to

$$U_t = (X_t, C_t, L_t).$$

We will call this the *reduced-form approach*. An alternative is to first predict future leverage ratios $\hat{L}_{t+k}$ and then include those in the set of explanatory variables in a *two-step approach*.

For each prediction horizon k the reduced-form approach would yield:

$$\begin{aligned} &P(Y_{t+k,t+k+1} = 1 | Y_{t+k-1,t+k} = 0, X_t, C_t, L_t) \\ &= \frac{1}{1 + \exp(-\alpha_k - \beta_k' X_t - c_0 - c_1' C_t - c_2 L_t)} \\ &= \frac{1}{1 + \exp(-\tilde{\alpha}_k - \tilde{\beta}_k' U_t)}, \end{aligned}$$

with $\tilde{\alpha}_k$ and $\tilde{\beta}_k$ being the respective combined coefficients.

Within the two-step approach, we first estimate future leverage ratios via relation (7).[5] In a second step, the base regression model (2) is augmented by including the following explanatory variable to the vector of covariates:

$$d\hat{L}_{t+k} = (\hat{L}_{t+k} - L_t). \tag{8}$$

The new variable $d\hat{L}_{t+k}$ captures the dynamics of leverage since it predicts the change from the most recently observed to the estimated future leverage ratio.[6] In this sense, we will refer to $d\hat{L}_{t+k}$ as to the *dynamic variable*. For estimation at $t + k$ the vector of covariates is defined as:

$$Z_{t+k} = (X_t, d\hat{L}_{t+k}).$$

The two-step approach is a special case of the reduced-form approach. It imposes the restriction that variables in the covariates vector C_t are weighted according to $b_1' C_t$ (cf. equation (4)) as well as the restriction that predicted leverage ratios are bound by 0 and 1. Imposing restrictions in a regression model automatically leads to a deterioration of the in-sample fit. The out-of-sample performance, however, could be improved through the two-step approach if the restrictions turn out to be correct and thus increase the precision of parameter estimates. We therefore examine both approaches.

The two models also differ in ease of interpretation. In the two-step approach the effect of leverage dynamics on credit worthiness is straight-forward: a larger positive difference $d\hat{L}_{t+k}$ comes along with a higher future leverage ratio and should therefore increase the probability of default. Consequently, we expect the regression coefficient of the dynamic variable to have a positive sign. In the reduced-form approach, by contrast, we have to interpret a set of coefficients and it would be difficult to assess the extent to which estimated coefficients are in line with the partial adjustment hypothesis that is the basis of our work.

In the two-step approach an additional comment needs to be made on the definition of the covariates vector Z_{t+k} if the horizon considered is larger than two years. For default in the time interval $(t + 1, t + 2]$ solely the estimated value $d\hat{L}_{t+1}$ is available

[5] Leverage ratios are theoretically bound to the unit interval $[0, 1]$. For this reason, we set all forecasted values that are negative or larger than 1 to this theoretical bounds.

[6] One could also add $\hat{L}_{t+k}$, the prediction of the leverage ratio. The difference $d\hat{L}_{t+k}$ is advantageous as it contains the same information as $\hat{L}_{t+k}$ and is more easily accessible in terms of interpretation. Besides, the estimate $\hat{L}_{t+k}$ is highly correlated with the current leverage ratio and by considering the difference instead of the absolute value we reduce the problem of multicollinearity.

as explanatory variable. This is not the case for longer horizons. E.g., for defaults within $(t+2, t+3]$ both $d\hat{L}_{t+1}$ and $d\hat{L}_{t+2}$ are at hand. In general, all estimated values with forecasting horizon shorter than the targeted horizon are conceivable. Yet, we do not consider this amplification. Estimated leverage ratios at consecutive points in time are highly correlated, which again introduces the problem of multicollinearity. Thus, for default prediction in $(t+k,\ t+k+1]$ the vector of firm specific covariates is only extended by $d\hat{L}_{t+k}$.

3 Data and Definition of Variables

We perform our analysis on a dataset constructed by merging three sources: firm specific accounting variables from the annual COMPUSTAT Database; market variables from the Center for Research in Security Prices (CRSP) as well as credit rating and default information from Moody's Investor Service.

Moody's database covers ratings and defaults for corporate bond issuers. The default indicator in our analysis is based on Moody's definition of default that includes missed or delayed payments, bankruptcy and distressed exchange, see also Hamilton et al. (2006). The state of default is not absorbing; a firm can recover and undergo multiple default events. We include observations for firms with multiple default events only if the credit rating prior to a recurrent default has been upgraded to at least "B2".

We restrict our analysis to a set of US companies excluding financial firms (CRSP SIC 6000-6999) and regulated utilities (CRSP SIC 4000-4999) whose capital decisions are subject to regulatory influence and hence might violate assumptions made in the capital structure model. The final dataset contains 17580 firm-years with complete information, 1694 distinct firms with an average of 10.4 observation years per firm. We consider a time range of 32 years, from 1974 to 2005. We begin the analysis as late as 1974 since there are very few default events prior to this year. During the whole observation period 274 default events were recorded, with the majority of defaults occurring in the years between 1997 and 2002.

3.1 Covariates For Default Analysis

In the literature there is no mutual consent on the right choice of covariates to be used for the estimation of default probabilities. Altman (1968) suggests a set of accounting ratios that is modified by Zmijewski (1984) and Ohlson (1980), among others. More recent research relies on these basic definitions of accounting ratios introducing slight modifications, cf. Campbell et al. (2006). Furthermore, Shumway (2001) shows that market variables measuring excess return and the stock's volatility have additional predicting power and can improve prediction accuracy substantially when added to accounting ratios. This result is confirmed by Chava and Jarrow (2004).

Although being different in detail, the various accounting ratios can be grouped as proxies for either leverage, profitability, liquidity or coverage; four main characteristics that determine the quality of an obligor. For our analysis, we decided on the definition of each characteristic that shows best explanatory power in terms of significance when compared to alternative proxies.

Leverage is measured as the ratio of total debt to the market value of total assets (L=TD/MTA), with MTA being the sum of total debt and market capitalization, cf. Campbell et al. (2006). Highly leveraged firms are theoretically at higher risk of default. We proxy expected profitability by the ratio of earnings before interest and taxes to total assets (EBIT/TA), as has been suggested by Altman (1968) and Fama and French (2002). In general, more profitable firms are less likely to default. We observe that neither of the alternative measures for liquidity (e.g. working capital to total assets or current assets to current liabilities) is significant. For this reason, we exclude a proxy for liquidity from the analysis. Similar to Blume et al. (1998) we quantify coverage by the ratio of earnings before interest and taxes over interest expenses of the same period (EBIT/XINT). The lower the ratio, the higher is the uncertainty that the company will satisfy its interest expenses and meet its debt obligations.

In addition to the core characteristics of a firm's credit-worthiness we include a measure of firm size along with a proxy for future investment opportunities as explanatory variables. Similar to Shumway (2001), firm size (SIZE) is measured as the natural logarithm of the ratio of a firm's market capitalization to the capitalization of the S&P500. Larger firms are suspected to be less exposed to default. We proxy future investment opportunities by growth in assets (dTA $= \frac{TA_t - TA_{t-1}}{TA_t}$), cf. Fama and French (2002). The corresponding coefficient is presumed to be positive.

As noted previously there is substantial evidence that market variables do improve prediction accuracy. For this reason we enlarge the vector of covariates by the stock return over the previous twelve months and the respective stock return volatility. Each firm's stock return (RET) over the previous twelve months is computed by cumulating monthly returns. Distressed firms are expected to perceive falling stock prices. Such, stock returns are associated to be negatively related to default. We define volatility (VOLA) as the standard deviation of monthly stock returns over the previous twelve months. We expect more volatile stocks to be an indicator for declining credit quality, which should be confirmed by a positive coefficient estimate.

All accounting variables at lagged by six months to ensure availability at the start of default prediction horizons. Furthermore, in order to mitigate the influence of outliers, we winsorize all observed covariates, except SIZE and EBIT/XINT, at the 1% and 99% quantiles. Taking the logarithm when measuring firm size already reduced the impact of outliers. Interest coverage is truncated similar to the procedure suggested by Blume et al. (1998): all observations with negative earnings are set to zero; all observations with negative interest expenses are set to the maximum value of 5; all remaining values are cut of at the maximum value of 5. The procedure results in a range of $[0, 5]$ for transformed coverage values.[7]

A summary of the included variables along with the respective COMPUSTAT items

[7] Blume et al. (1998) point out that coverage is likely to have non-linear effects. Consequently, in a logit model a change in values close to the mean of the distribution would have a larger impact on the dependent variable than the same change in values close to the tails of the distribution.

is provided in Panel A of Table 6 in the Appendix. Table 7 (Panel A) in the Appendix reports the descriptive statistics of the data set.

3.2 Covariates For Leverage Forecasts

In line with the specification for default analysis leverage is measured as a firm's market debt ratio. For the vector C_t of covariates we consider a set of widely used firm characteristics (cf. Rajan and Zingales (1995), Howakimian et al. (2001), Fama and French (2002), Flannery and Rangan (2006)). These include profitability, firm size, non-debt tax shields and tangibility as well as product uniqueness along with growth and investment opportunities.[8]

More profitable firms have less leverage since firms with higher retained earnings are more likely to use equity finance. On the other hand, the inverse relationship between profitability and bankruptcy costs implies better debt conditions. Firm size has a positive relation to leverage. Reasons might be lower default probability, less volatile earnings and better access to the debt market. Non-debt tax shields are proxied by the proportion of depreciation expense relative to total assets (XD/TA). Higher non-debt tax shields most likely increase the preference for equity relative to debt financing. Tangibility is measured by the ratio of fixed to total assets (FA/TA). Firms with a higher proportion of tangible assets operate with higher leverage. Product uniqueness, measured by the proportion of research and development expenses to total assets (R&D/TA), should be negatively related to leverage. We proxy growth opportunities by the market-to-book ratio (MB); firms with high expected future growth tend to have more equity. As in Flannery and Rangan (2006) we also include a dummy variable indicating whether a firm does report R&D expenses. Again growth in assets is used as a proxy for investments opportunities.[9] In addition, following Flannery and Rangan (2006) we include the industry mean of leverage (INDMEAN) in order to control for industry wide characteristics not captured otherwise. For industry definitions we refer to Chava and Jarrow (2004).[10]

A detailed description of covariate definitions is provided in Panel B of Table 6 in the Appendix. Table 7 (Panel B) in the Appendix reports the descriptive statistics of the data set. Accounting variables are lagged by six months to ascertain availability and outliers are removed through winsorization at the 1% and 99% quantiles.

4 Partial Adjustment Model - Econometric Issues

Estimation of the partial adjustment model is associated with three critical questions. Firstly, panel data is in general subject to two sources of dependence: correlation of residuals for a given year across firms and correlation of residuals for a given firm across

[8] For consistency reasons, variables that approximate the same characteristics considered in the default analysis are not redefined.

[9] Fama and French (2002) include growth in assets to capture temporary movements in leverage away from its target rather than as a direct determinant of target leverage.

[10] A finer indexing of industry dummies as in Fama and French (1997) does not change the results significantly.

years. The first source is usually controlled for using time dummies, whereas cross-sectional dependence is dealt with by including dummies for each firm (fixed effects or within groups estimation), c.f. Peterson (2007). Secondly, the set of covariates includes a lagged dependent variable, i.e. the most recently observed leverage ratio L_t. The introduction of a lagged dependent variable comes along with the problem of endogeneity, since L_t will be correlated with the error term. A standard approach in this case is a two-step ordinary least squares regression with an instrumental variable that is highly correlated with the endogenous covariate, but orthogonal to the error term.

If the set of covariates in a panel data regression includes a lagged dependent variable the data structure is referred to as *dynamic panel data*. Several estimation techniques have been suggested for dynamic panel data, in particular for samples where a large number of individuals has been observed for a small number of time periods. A detailed overview can be found in Baltagi (2003). However, the appropriate technique is not always obvious and as pointed out by Bond (2002) the methodology of choice largely depends on the characteristics of the data being analyzed. In this setting Bond (2002) shows that OLS and the fixed effects estimates for the lagged dependent variable are likely to be biased in opposite directions. Consequently, a reliable estimator should lie in (or close) to the implied range. However, in our application we primarily focus on accurate forecasts rather than on correct estimates of coefficients. Within the two-step approach, we use forecasts of future leverage ratios as an independent covariate in the default analysis in order to increase prediction accuracy. Thus, it is obvious that our findings are sensitive to the results from the first-step regression, and poor forecasts should be avoided.

Keeping this in mind, we compare several estimation methods and use the one that performs best with respect to forecasting ability. More precisely, we consider within groups estimation, instrumental variables regression with and without fixed effects and Arellano and Bond's (1991) GMM procedure.[11] For the instrumental variables regression we follow Flannery and Rangan (2006) who show that book debt ratio can serve as an instrument for L_t. Book debt ratio is defined as the ratio of total debt to the book value of total assets (whereas leverage is defined as the ratio of total debt to the market value of total assets). We test the forecasting performance out-of-sample. For this purpose we estimate the coefficients on a sample including the years 1985 to 1995. Future leverage ratios are estimated for years 1996 to 2000 and the forecasts are compared with the observed values. Table 8 in the Appendix quantifies the performance of alternative methods through the root mean squared error (RMSE). Each method is applied to a forecasting period of one to four years. We observe smallest RMSE values for instrumental variables regression without fixed effects. Second best performs the GMM estimation suggested by Arellano and Bond (1991). The latter, however, is computationally very expensive (using STATA's xtabond2 procedure).

In line with this results we use instrumental variables regression without fixed effects for leverage forecasts. All following inference will be based on this choice. Estimates for

[11] The Arellano and Bond's (1991) technique is implemented in STATA's xtabond2 procedure described in Roodman (2006).

the partial adjustment model one to four periods ahead are presented in Table 9 in the Appendix (we do not list coefficient estimates for time dummies). The coefficient for the present leverage ratio quantifies one minus the average fraction of the gap between the targeted and observed future leverage that is closed. For a one-year horizon this is approximately $11\% = (1 - 0.89)$ (Fama and French (2002) report a mean reversion of 7%-10% for dividend payers and 15%-18% for non-payers per year). The estimated adjustment speeds for two to four years are close to the theoretical values given the observed one-year adjustment speed. E.g. the estimated two-year adjustment speed is $(1-0.80) = 0.2$, whereas with a one-year adjustment speed of 0.11 the geometric decline would result in $(1 - (1 - 0.11)^2) = 0.21$.

5 Empirical Results

Our presentation of empirical results is structured in the following way. We first report estimation results from the in-sample estimation for each of the considered models: base, two-step dynamic and reduced-form dynamic. The respective prediction accuracies are compared out-of-sample in Section 5.2.

5.1 In-Sample Estimation

5.1.1 Base Model

Results from the base model when estimating default probabilities one to five years ahead are summarized in Table 1. We report robust standard errors adjusted for clustering on firms and provide Mc-Faddens Pseudo-R^2 as an indicator for the overall fit of the model as well as the area under the Receiver Operating Characteristic (ROC) curve (AUC).[12]

Beforehand we have carried out a test for temporal dependence. The likelihood-ratio test for the significance of temporal dummy variables indicates no temporal dependence in the data. P-values for the likelihood-ratio tests are shown in the last row of Table 1. On account of this result, we omit temporal dummies in all subsequent regressions.

For the one-year prediction horizon coefficient estimates for all covariates, but EBIT/TA, carry the expected sign and show high significance. EBIT/TA has the right sign, but is only marginally significant with a p-value of 0.082. For further horizons this covariate remains insignificant. Other accounting ratios (L, EBIT/XINT and dTA) keep their predictive power with longer horizons. Among the market variables only SIZE exhibits long term explanatory power. Neither past returns, nor volatility can explain defaults in the remote future. We find a rather rapid decline in the overall fit of the model, comparing the relatively high Pseudo-R^2 value for a one-year horizon (Pseudo-$R^2 = 0.4012$) to the very small overall fit for the five-year horizon (Pseudo-$R^2 = 0.0718$). A similar

[12] The area under the ROC curve measures the discriminating power of a model; a value of AUC close to the maximum of 1.0 corresponds to a perfect discrimination between defaulters and non-defaulters, whereas a random decision would lead to a value of about 0.5. Another widely used measure for prediction accuracy is the accuracy ratio (AR), which is the area under the CAP curve. However, AR is a simple linear transformation of AUC: $AR = 2AUC - 1$.

	Prediction Horizon in Years				
	1	2	3	4	5
L	4.89***	2.97***	2.41***	1.49**	0.99*
	(0.53)	(0.41)	(0.39)	(0.46)	(0.48)
EBIT/TA	-1.85	0.95	-0.24	1.86	2.23
	(1.06)	(1.16)	(1.29)	(1.48)	(1.68)
EBIT/XINT	-0.30**	-0.33***	-0.19**	-0.29**	-0.31***
	(0.10)	(0.08)	(0.07)	(0.09)	(0.09)
SIZE	-0.18***	-0.20***	-0.19***	-0.22***	-0.21***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)
dTA	1.09**	1.13***	1.12**	1.17*	1.37**
	(0.39)	(0.34)	(0.40)	(0.5)	(0.45)
RET	-1.67***	-0.65***	-0.29	-0.24	-0.11
	(0.31)	(0.18)	(0.16)	(0.19)	(0.20)
VOLA	4.70***	4.39***	3.42**	0.66	0.33
	(1.01)	(1.11)	(1.07)	(1.36)	(1.39)
Constant	-9.05***	-7.53***	-7.08***	-6.47***	-6.12***
	(0.54)	(0.49)	(0.44)	(0.50)	(0.53)
Firm-years	17580	15569	14614	13721	13232
Pseudo-R^2	0.4012	0.2238	0.1398	0.0965	0.0718
AUC	0.9465	0.8809	0.8259	0.7856	0.7501
P-value	0.9309	0.8851	0.7358	0.3884	0.3030

Table 1: Base model: coefficient estimates for default probabilities one to five years ahead. Standard errors are in parentheses. Stars indicate a significance level: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. P-values quantify the significance of temporal dummies.

trend is observed for prediction accuracy, the area under the ROC curve reduces from $\text{AUC}_1 = 0.9465$ to $\text{AUC}_5 = 0.7501$.

5.1.2 Two-step Dynamic Model

Table 2 presents coefficient estimates in the two-step dynamic model for default probabilities two to five years ahead. For any of the considered horizons, the dynamic variable has the right sign and exhibits high statistical significance; all p-values are smaller than 0.001.

The covariates EBIT/XINT, RET, VOLA and SIZE remain unaffected. There are only minor changes in magnitude and negligible changes in significance for these variables after inclusion of $d\hat{L}$. The proxy for profitability remains insignificant. Only the covariate dTA is affected. The coefficient estimate shows a loss of explanatory power and changes the sign when combined with leverage forecasts. This might be due to the fact that dTA is part of the covariates vector in the first step regression (i.e. C_t) that leads to the leverage forecasts.

We further observe that adding the dynamic variable increases the overall fit of the base model. For all horizons, values for the Pseudo-R^2 and AUC of the two-step dynamic model exceed the respective values of the base model. The best relative gain appears for a prediction horizon of four years, with an increase from $\text{AUC}_{\text{Base}} = 0.7856$ to $\text{AUC}_{\text{Two-Step}} = 0.8181$.

5.1.3 Reduced-form Dynamic Model

Results for the reduced-form dynamic model are reported in Table 3. For the sake of brevity we do not list coefficient estimates for time dummies. The model shows the best overall fit and discriminatory power which is what we expect because it imposes fewer restrictions than the other models. The improvement is pronounced. It is largest at the five-year horizon. Discriminatory power increases from $\text{AUC}_{\text{Base}} = 0.7501$ and $\text{AUC}_{\text{Two-Step}} = 0.7775$ to $\text{AUC}_{\text{Reduced-Form}} = 0.8366$.

The last row of Table 3 reports p-values for the null hypothesis: $C_t = 0$, i.e. covariates from the partial adjustment model have no explanatory power.[13] Applying the likelihood ratio test the null hypothesis is rejected for all horizons which means that in combination elements of the covariates vector C_t do contribute to default prediction. Yet, when inspecting each covariate separately we find no significant coefficients (except for R&D/TA). The low precision of the estimates indicates that imposing the restriction $L_{t+k} = b_0 + b_1' C_t$ (as is done in the two-step dynamic model) might improve out-of-sample performance.

[13] We only consider covariates that are not part of the base model.

	Prediction Horizon in Years			
	2	3	4	5
L	4.33***	5.09***	4.78***	4.13***
	(0.49)	(0.53)	(0.66)	(0.75)
$d\hat{L}_1$	12.66***			
	(2.24)			
$d\hat{L}_2$		13.38***		
		(1.68)		
$d\hat{L}_3$			11.82***	
			(1.58)	
$d\hat{L}_4$				9.00***
				(1.61)
EBIT/TA	0.58	-1.57	0.74	1.61
	(1.18)	(1.34)	(1.61)	(1.82)
EBIT/XINT	-0.33***	-0.18*	-0.27**	-0.28**
	(0.08)	(0.08)	(0.09)	(0.09)
SIZE	-0.14**	-0.11*	-0.14**	-0.16**
	(0.05)	(0.05)	(0.05)	(0.06)
dTA	0.27	-0.16	-0.09	0.48
	(0.36)	(0.45)	(0.53)	(0.49)
RET	-0.65***	-0.31	-0.27	-0.13
	(0.19)	(0.17)	(0.21)	(0.21)
VOLA	4.52***	4.16***	1.90	1.89
	(1.14)	(1.10)	(1.41)	(1.40)
Constant	-7.52***	-7.37***	-7.05***	-6.94***
	(0.48)	(0.45)	(0.52)	(0.58)
Firm-years	15569	14614	13721	13232
Pseudo-R^2	0.2361	0.1699	0.1279	0.0933
AUC	0.8885	0.851	0.8181	0.7775

Table 2: Two-step dynamic model: coefficient estimates for default probabilities two to five years ahead. Standard errors are in parentheses. Stars indicate a significance level: $^{*}p < 0.05;^{**} p < 0.01;^{***} p < 0.001$.

	Prediction Horizon in Years			
	2	3	4	5
L	2.89***	2.60***	1.56**	1.37*
	(0.49)	(0.47)	(0.56)	(0.56)
EBIT/TA	0.65	-0.69	2.34	2.67
	(1.31)	(1.42)	(1.80)	(1.87)
EBIT/XINT	-0.3***	-0.14	-0.25*	-0.26**
	(0.08)	(0.08)	(0.10)	(0.10)
SIZE	-0.20***	-0.17***	-0.17**	-0.2**
	(0.05)	(0.05)	(0.06)	(0.06)
dTA	0.83*	0.70	0.85	1.18*
	(0.35)	(0.42)	(0.51)	(0.46)
RET	-0.53**	-0.30	-0.48*	-0.13
	0.20)	(0.18)	(0.22)	(0.23)
VOLA	5.39***	5.13***	4.05*	3.93*
	(1.32)	(1.31)	(1.72)	(1.68)
XD/TA	2.73	-1.71	1.66	4.09
	(3.51)	(3.86)	(4.41)	(4.10)
FA/TA	0.16	-0.06	-0.51	-0.39
	(0.43)	(0.41)	(0.44)	(0.44)
R&D/TA	-7.98*	-6.54	-9.41*	-7.88*
	(3.59)	(3.63)	(4.05)	(3.7)
R&D_Dummy	-0.14	-0.22	-0.28	-0.42*
	(0.16)	(0.17)	(0.17)	(0.18)
MB	-0.10	-0.24	-0.32	-0.03
	(0.25)	(0.19)	(0.23)	(0.19)
INDMEAN	0.41	0.29	0.13	-0.88
	(1.19)	(1.31)	(1.38)	(1.58)
Constant	-9.08***	-7.33***	-6.75***	-7.57***
	(1.24)	(0.94)	(1.05)	(1.35)
Firm-years	15569	14614	13721	13232
Pseudo-R^2	0.2709	0.2049	0.1702	0.1453
AUC	0.9050	0.8771	0.8556	0.8366
P-Value	<0.001	<0.001	<0.001	<0.001

Table 3: Reduced-form dynamic model: coefficient estimates for default probabilities two to five years ahead. Standard errors are in parentheses. Stars indicate a significance level: $^{*}p < 0.05;^{**} p < 0.01;^{***} p < 0.001$. P-values quantify the simultaneous significance of covariates from the partial adjustment equation.

5.2 Out-of-Sample Analysis Of Prediction Accuracies

In-sample results presented in the previous section suggest that additional information on future capital structure might actually lead to more accurate default prediction. At this point an out-of-sample validation is needed to see the impact of the dynamic variable in a realistic setting, where an analyst aims to predict default probabilities at a certain point in time only having information up to this point.

We conduct the out-of-sample validation as follows. Given a certain validation year t_0, we first estimate coefficients on a sample consisting of observations of the previous n years.[14] This step is repeated ten times. For each iteration we shift the validation year one year back, resulting in default probability estimates for ten validation years. The choice of the starting validation year t_0 depends on the prediction horizon. It is the most recent year with a positive number of defaults, e.g. for horizon $k = 2$ we set $t_0 = 2002$, for horizon $k = 3$ we set $t_0 = 2001$ and so on. The performance of the competing models is again compared by means of the AUC. In order to have a sufficient number of observations, the ROC curve is computed on the set of pooled probability estimates for all validation years. Additionally, we apply the algorithm suggested by DeLong et al. (1988) in order to test the null hypothesis of equality of areas under two ROC curves.

We vary the number of years used to estimate coefficients from $n = 8$ to $n = 13$ years.[15] The variable $d\hat{L}$ in the dynamic model is obtained in a way which is consistent with the out-of-sample methodology. The partial adjustment regression is estimated only with data from the years that make up the estimation sample; to compute the mean industry leverage ratio, which is an explanatory variable in this regression, we also use only observations from the estimation sample.

In the out-of-sample analysis we compare prediction accuracies from the two dynamic models to those of the base model. Tables 10 and 11 in the Appendix report the respective results in Panels A to F. Each panel represents an estimation sample; the validation samples remain unchanged. We compute the relative increase in AUC values for both dynamic models compared to the base model and report p-values for the null hypothesis of equality of areas under two ROC curves. Additionally, a more concise presentation of results is provided for each model separately, in Figure 1. We plot the AUC value against the number of years in the estimation sample. For the dynamic models each value is represented by stars, a circle or a cross. The symbols show the decision for the null hypotheses. If there is a relative increase in the AUC value when compared to the base model and the null hypothesis is rejected, stars indicate the significance level. If the null hypothesis cannot be rejected at a significance level of 5%, the plot displays a circle. A cross is used if we find a significant decrease in the AUC value.

The out-of-sample analysis exhibits two interesting facts. Firstly, we observe that the absolute as well as the relative out-of-sample performance does depend on the number

[14] More accurately, this technique should be referred to as out-of time validation.

[15] For a fixed sample, the number of years available for estimation declines with the prediction horizon. If the history covers n years, only $n - (k - 1)$ years can be used for prediction or forecast k years ahead.

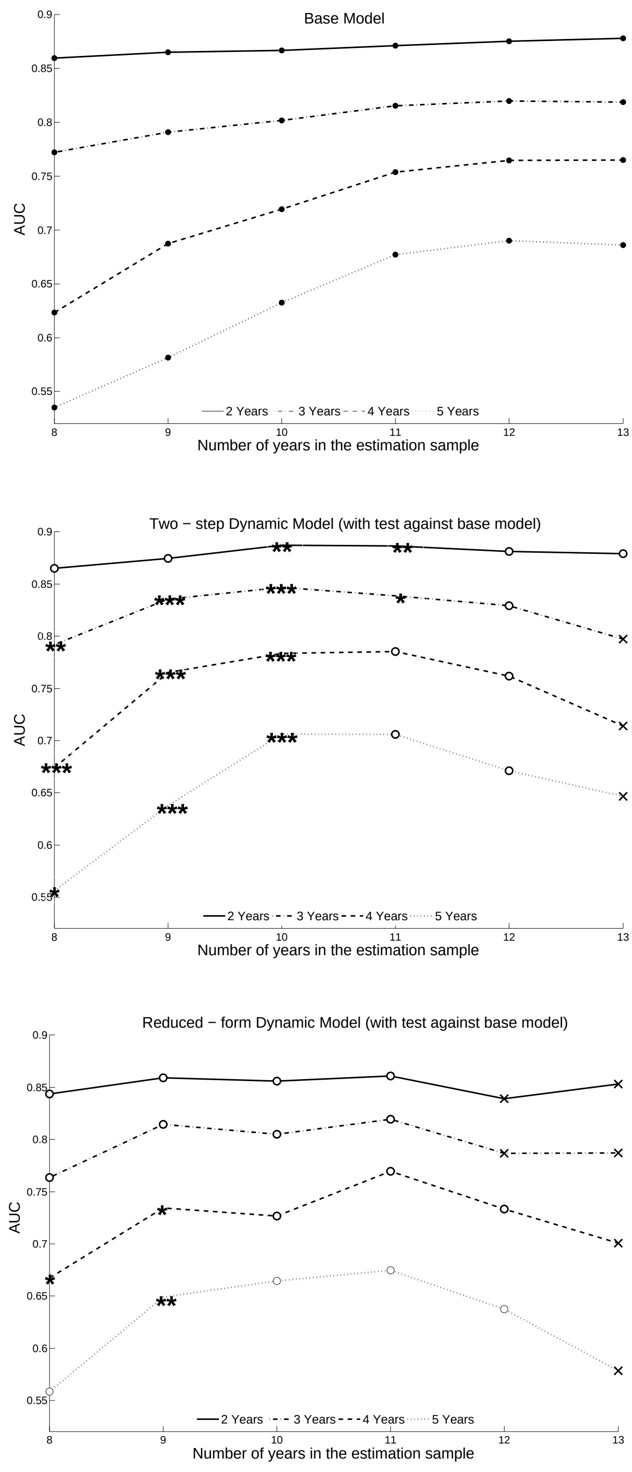

Figure 1: Out-of-sample prediction accuracies: prediction horizon of 2 to 5 years. Number of years in the estimation samples ranges from 8 to 13. Stars indicate a significance level: $^{*}p < 0.05$;$^{**}p < 0.01$;$^{***}p < 0.001$, circles indicate no significance for the null hypotheses of equality of areas under two ROC curves, crosses show significant inferiority.

of years included in the estimation sample. For the two-step dynamic model a consistent significant increase only shows for estimation samples that cover less then eleven years. Secondly, we see that the increase in prediction accuracy for the reduced-form dynamic model is much lower than that for the two-step dynamic model when compared to the base model. Dynamics of leverage cannot be captured sufficiently by simply adding the explanatory variables to the set of default risk drivers. The two-step procedure is clearly more favorable even though there might be uncertainties about the right estimation techniques for the partial adjustment model.

A separate inspection of the models reveals that in the base model AUC values increase with the number of years covered by the estimation sample, whereas in the two-step dynamic model AUC values first increase, reach a maximum for a history of ten years and then decline (Figure 1, second subfigure). This effect appears for all prediction horizons considered, but is more pronounced for default predictions four and five years ahead. In the reduced-form dynamic model, a similar trend shows for the four and five-year horizons.

How can these findings be explained? First, we will try to understand the trend exhibited for the base model. Basically, we see that prediction accuracy increases if the estimation sample becomes larger. In larger samples coefficient estimates are less exposed to overfitting and thus will lead to more accurate predictions. Thus, the more years are on hand for estimation, the better will the base model perform.

The poor performance of the base model for small estimation samples comes along with a significant increase of prediction accuracy after adding the dynamic variable. However, unlike to what we would expect, the dynamic model becomes less superior the larger we choose the estimation sample until finally no improvement is noticeable. We notice best improvement, both in terms of absolute increase in the AUC value and statistical significance, for the nine and ten-year estimation samples. Why does this effect grow weaker? The partial adjustment model assumes that the adjustment speed factor as well as the linkage between firm characteristics and target leverage are constant over time. In-sample, departures from this assumption are mitigated by time dummies. Out-of-sample, they are not, and will therefore lead to a reduction in the precision of leverage forecasts. A plot of the root mean squared errors for leverage forecasts presented in Figure 5 in the Appendix supports this conjecture. We see a trend that is in line with the pattern of the AUC values. The root mean squared errors are u-shaped with a minimum for an estimation sample of ten years.

A convenient feature of the two-step procedure is that it separates default prediction from leverage forecasts. Hence we can test the above assumption by combining different estimation sample histories. More precisely, for histories up to 10 years, estimation samples for leverage forecasts and default predictions cover the same number of years. For longer histories, estimation samples for leverage forecasts are fixed to 10 years and the number of years used for default prediction is varied from 11 to 13 years. We use a history of at maximum 10 years in the first step of the dynamic model since here we have observed the highest increase in the AUC. Results are shown in Figure 2 and Table 12 in the Appendix. The trend observed above is no longer present; prediction accuracy does not decline for larger estimation samples. Moreover, we even observe significantly

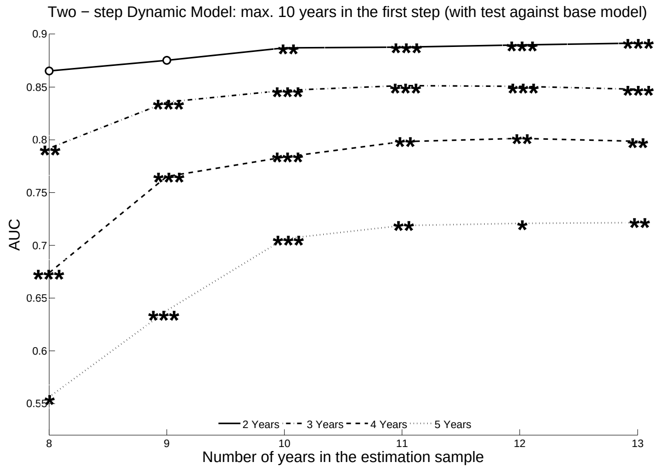

Figure 2: Out-of-sample prediction accuracies: prediction horizon of 2 to 5 years. Number of years in the estimation samples for default analysis ranges from 8 to 13. Leverage forecasts are based on estimation samples covering at most 10 years. Stars indicate a significance level: $^{*}p < 0.05;^{**}p < 0.01;^{***}p < 0.001$ and circles indicate no significance for the null hypotheses of equality of areas under two ROC curves.

higher AUC values for the two-step dynamic model, where in previous computations the two-step dynamic model has been inferior. The increase is moderate for a two-year prediction horizon, ranging from 1.5% to 1.9%. Defaults three to five years ahead can be predicted with much higher increase in accuracy; here the values improve by 3.6% to 6.1%.

Figure 3 exemplarily compares the AUC values given a prediction horizon of four years. The base and the reduced-form dynamic model show lowest values for estimation samples with a history of less respectively larger than 12 years. Best performance is observed for the two-step dynamic model, restricted to a ten-year estimation sample in the first step.

In essence, the rather poor out-of-sample performance of the reduced-form dynamic model indicates that a simultaneous estimation of leverage dynamics and default probabilities cannot capture the dynamics in leverage sufficiently. We observe a larger effect for the two-step approach. An inclusion of the dynamic variable does improve prediction accuracy significantly if out-of-sample forecasts of leverage ratios are based on estimation samples covering at maximum the past ten years, even if a longer data history is available. In particular, the two-step dynamic model is of benefit when historical data available for default analysis is scarce.

5.3 Variations of the Two-Step Dynamic Model

So far we have estimated future leverage ratios under the assumption of a homogeneous adjustment behavior for all firms. The coefficient a_1 in equation (3) is the average ad-

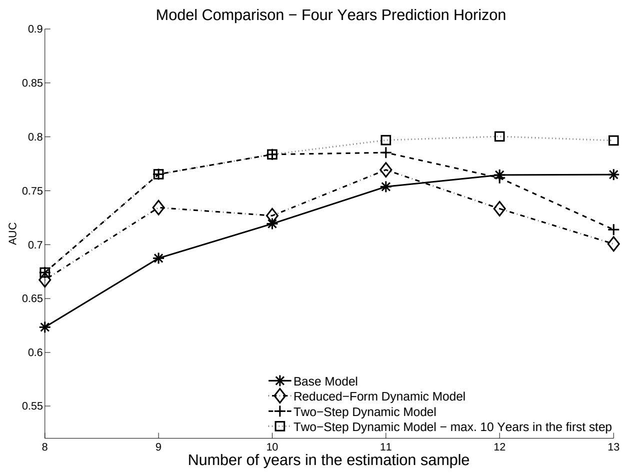

Figure 3: Out-of-sample prediction accuracies: comparison of the considered models at a prediction horizon of four years.

justment rate for a firm operating between 1974 and 2005. The simplifying assumption of a constant adjustment rate has widely been used in the literature on capital structure. Only a few studies model the adjustment rate as firm specific and time varying (Banerjee et. al (2000), Roberts (2002)).[16] The analysis conducted by Roberts (2002) concludes that a firm's adjustment speed depends on its industry and financial health. An immediate implication hereof is that accuracy of leverage forecasts (and hence the accuracy of estimated default probabilities) might increase if the dynamic model allows for differences in industry and financial health.

Within the two-step dynamic model one can adjust for differences in industry and financial health when running the first step regression on industry or rating category subsamples.[17] We use the rating category as a proxy for financial health. In order to ensure a sufficient sample size in each rating category, we aggregate Moody's rating classes into five categories (high = Aaa to Aa3, medium = A1 to Baa3, speculative = Ba1 to B3, poor = Caa1 to C, and unrated). Industry subgroups are defined as in Chava and Jarrow (2004), resulting in a total of eight industry classes.

For the in-sample results (not tabulated), we find that neither the significance nor the magnitude of coefficient estimates in the default regression is affected when the first-step regression is run on industry or rating category subsamples; values of the Pseudo-R^2 and the AUC increase moderately. Out-of-sample prediction accuracy increases only if the number of years used in the estimation sample is sufficiently large. Figure 4

[16] In our analysis the out-of-sample results presented in Section 5.2 also suggest that the conjecture of a constant adjustment speed factor is inadequate.

[17] Alternatively one could include a proxy for industry characteristics (e.g. mean industry leverage ratio) and financial health (e.g. present rating category) as independent variable in the partial adjustment model. Our specification of the optimal leverage ratio does control for industry characteristics, yet the proxy INDMEAN shows no explanatory power.

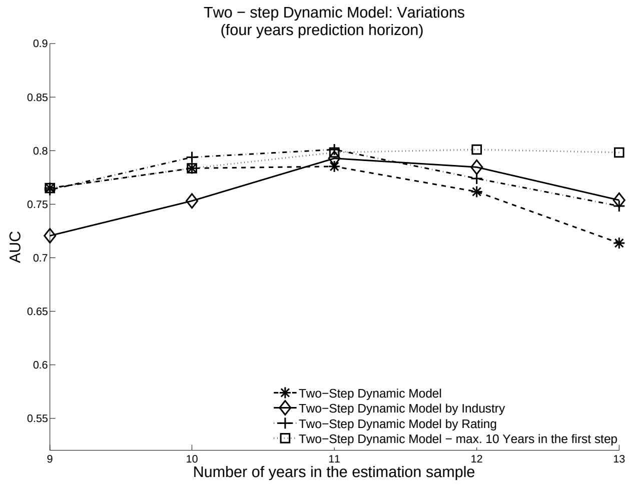

Figure 4: Out-of-sample prediction accuracies: Comparison of out-of-sample AUC values for variations of the two-step dynamic model at a prediction horizon of four years.

exemplarily compares the out-of sample performance for the modifications of the two-step dynamic model for a prediction horizon of four years.[18] Forecasting leverage based on rating subsamples does consistently increase the AUC; using industry subsamples leads to improvements only if at least 12 years are included in the estimation sample. The improvement from using subsamples is relatively small compared to using the "optimal" estimation sample in the first step regression. This can be seen in Figure 4 where for comparison we also show AUC values for the two-step dynamic model in which we fix the estimation sample in the first step to at maximum ten years.

Overall, the results show that refinements of the partial adjustment regression could help to further improve default prediction. Improvements are not clear-cut though, suggesting that the two-step dynamic model with its rather simplistic assumption of homogeneous adjustment behavior is a good starting point for practical purposes.

6 Do Credit Ratings Capture the Dynamics of Leverage?

In the preceding section we have shown that additional information on the dynamics of capital structure can significantly improve prediction accuracy of long-term default probabilities. Credit ratings assigned by rating agencies such as Moody's or Standard & Poor's also focus on the long term quality of an obligor's debt. Fons et al. (2002) state on Moody's ratings: "[. . .] credit ratings powerfully discriminate among relative long-term risks. They target multiple horizons, rather than a single, defined investment horizon.". In other words, credit ratings are meant to be forward-looking with respect to credit quality and the question arises whether such a forward-looking ability covers

[18] A similar pattern is observed for all prediction horizons.

leverage dynamics. If it does, adding credit rating information to the default regression in the two-step dynamic model would make the dynamic leverage variable redundant.

We address this question by considering the following covariates in the default prediction model:

(i) $Z_{t+k} = (X_t, d\hat{L}_{t+k})$, i.e. two-step the dynamic model,

(ii) $V_t = (X_t, RAT_t)$, i.e. the base model plus ratings,

(iii) $W_{t+k} = (X_t, d\hat{L}_{t+k}, RAT_t)$, i.e. the two-step dynamic model plus ratings,

where RAT_t denotes the credit rating assigned by Moody's at time t. We convert each of the 21 rating categories to a numerical variable; 1 corresponds to rating class Aaa, 2 to rating class Aa1, ..., 21 corresponds to rating class C. A higher rating should therefore be associated with a higher probability of default.

We first report in-sample coefficient estimates for discrete default probabilities two to five years ahead when both the dynamic variable and credit rating are included as covariates. The respective values are presented in Table 4. The number of complete observations available for estimation declines after including RAT, due to lacking rating information for some firm-years. When we compare coefficient estimates for the dynamic model with and without inclusion of rating information (i.e. figures from Tables 2 and 4), we find high additional explanatory power for the rating variable, which is also mirrored in the increasing Pseudo-R^2 values. Further we can observe that the inclusion of RAT has almost no effect on the significance and magnitude of the dynamic variable. This implies that dynamics in leverage are not captured by credit ratings.

Mutual significance of leverage and rating information can further be validated by a comparison of the respective out-of-sample ROC areas. More precisely, we estimate AUCs after default regression with covariates Z_{t+k}, V_t and W_{t+k} and contrast all pairs. As has been observed in Section 5.2 out-of-sample forecasts on capital structure are best when based on an estimation sample including at most the preceding ten years. We follow this rule for the out-of-sample comparison of prediction accuracies implied by the models in question. Results are presented in Panels A to C of Table 5. In Panel A we compare ROC areas for the pair (Z_{t+k}, V_t), in Panel B for the pair (W_{t+k}, V_t) and in Panel C for the pair (W_{t+k}, Z_{t+k}).

We find the implication from coefficient estimates in Table 4 confirmed in all three Panels. More precisely, Panel A shows that $d\hat{L}$ is marginally superior compared to RAT. Panel B suggests that adding the dynamic variable to the base model enlarged by rating information significantly increases prediction accuracy (up to 9.5% for a five years prediction horizon). In Panel C, we see that adding the rating information to the dynamic model can also increase the AUC significantly, albeit only for horizons lager than three years.

Summarizing, both coefficient estimates and the out-of-sample analysis of ROC curves suggest that dynamics in leverage and rating are complementary in the sense that both do carry exclusive information relevant for long term credit quality. Neither one can be replaced by the other without loss of prediction accuracy.

	Prediction Horizon in Years			
	2	3	4	5
L	4.11***	4.78***	5.18***	2.76***
	(0.55)	(0.56)	(0.68)	(0.72)
$d\hat{L}_1$	12.62***			
	(2.36)			
$d\hat{L}_2$		13.54***		
		(1.71)		
$d\hat{L}_3$			13.54***	
			(1.57)	
$d\hat{L}_4$				4.93***
				(1.35)
RAT	0.12**	0.19***	0.22***	0.13***
	(0.04)	(0.04)	(0.04)	(0.04)
EBIT/TA	-0.21	-2.02	-1.57	1.21
	(1.22)	(1.33)	(1.55)	(1.76)
EBIT/XINT	-0.25**	-0.07	-0.06	-0.20*
	(0.09)	(0.08)	(0.09)	(0.1)
SIZE	-0.10	0.02	0.03	-0.06
	(0.06)	(0.07)	(0.07)	(0.08)
dTA	0.48	-0.10	-0.32	0.79
	(0.37)	(0.45)	(0.51)	(0.48)
RET	-0.71***	-0.44**	-0.41	-0.35
	(0.19)	(0.17)	(0.21)	(0.21)
VOLA	3.53**	3.23**	0.15	-0.32
	(1.23)	(1.16)	(1.55)	(1.45)
Constant	-8.38***	-8.28***	-8.2***	-7.06***
	(0.54)	(0.51)	(0.61)	(0.62)
Firm-years	13450	13225	13025	12854
Pseudo-R^2	0.2528	0.1915	0.1591	0.0933
AUC	0.8952	0.8583	0.8124	0.7584

Table 4: Dynamic model enlarged by credit rating: coefficient estimates for default probabilities two to five years ahead. Standard errors are in parentheses. Stars indicate a significance level: $^{*}p < 0.05$;$^{**}p < 0.01$;$^{***}p < 0.001$.

	Prediction Horizon in Years			
	2	3	4	5
Panel A				
AUC Dynamic	0.8919	0.8513	0.7925	0.7245
AUC Rating	0.881	0.8256	0.7557	0.6903
Change in %	1.2	3.1	4.9	5.0
P-value	0.0835	0.0103	0.0451	0.0498
Panel B				
AUC Dynamic and Rating	0.8944	0.8577	0.8124	0.7561
AUC Rating	0.881	0.8256	0.7557	0.6903
Change in %	1.5	3.9	7.5	9.5
P-value	0.006	<0.001	<0.001	<0.001
Panel C				
AUC Dynamic and Rating	0.8944	0.8577	0.8124	0.7561
AUC Dynamic	0.8919	0.8513	0.7925	0.7245
Change in %	0.3	0.8	2.5	4.4
P-value	0.2098	0.1626	0.0144	0.0029
Firm-years	5627	5323	5054	4790

Table 5: Dynamic Variable vs. Credit Rating: out-of-sample prediction accuracies. The p-value refers to the null hypotheses of equality of areas under two ROC curves. The estimation sample covers a history of 10 years.

7 Summary and Concluding Remarks

Empirical corporate finance research documents that firms actively rebalance their capital structure towards a target leverage ratio. The main question addressed in our paper is whether this insight can be used to improve the prediction of corporate defaults. We use two related econometric specifications. In both cases, a standard discrete hazard rate model (our base model) is enriched by variables that capture the dynamics of leverage. In the reduced-form approach, we just add the variables that the capital structure literature uses to model targeted leverage ratios. In the two-step approach, we augment the base model by the predicted change in the leverage ratio; the prediction is obtained from a first-step regression along the lines of partial adjustment model used in the capital structure literature. The reduced-form approach nests the two-step approach.

In-sample, both approaches lead to significant improvements over the base model. The reduced-form model performs best because it imposes the fewest restrictions. To judge the usefulness of a default prediction model, however, it is crucial to conduct an out-of-sample analysis. For the reduced-form dynamic model, we find no material increase in predictive accuracy as measured by the area under the ROC curve and hence conclude that dynamics in leverage cannot be captured by solely adding covariates from the partial adjustment model to the set of risk drivers. For the two-step dynamic model, by contrast, a significant increase in prediction accuracy is observed. The restrictions imposed through the partial adjustment model increase the precision of coefficient estimates and thereby lead to better out-of-sample performance. The magnitude of this improvement varies with the length of the sample used for estimating the partial adjustment equation. We attribute this effect to non-stationarities in the speed of adjustment and the linkages between firm characteristics and leverage targets. Specifically, the performance of the dynamic model deteriorates if the estimation sample is longer than 10 years.

Standard default prediction models do not use information contained in credit ratings but many market participants do. Given that rating agencies claim their ratings to be forward looking we further ask whether credit ratings contain information about leverage dynamics. We address this question by adding credit ratings to our default prediction models. Including credit ratings improves predictive power but the contribution of the leverage forecast is not affected. While rating analysts do appear to have a forward-looking ability, this ability is unrelated to the predictability of leverage ratios that has been documented in the corporate finance literature.

The paper could inspire other extensions of existing default prediction models. Until now, specification of such models is largely based on a search for covariates that perform better than the ones previously considered. The final regression equation is reduced-form in the sense that it imposes no restrictions on variables (except for the commonly made linearity restriction). Our analysis has shown that default prediction models can benefit from imposing restrictions that are derived from other research areas, e.g. capital structure. Finally, the work contributes to the corporate finance literature since it shows that it is important to model the leverage targeting behavior of firms. If we had obtained that leverage forecasts from a partial adjustment equation are useless in default prediction, this would have cast doubts on the empirical relevance of the trade-off theory.

APPENDIX

Panel A - Default Prediction		
Covariate	Source	CRSP/COMPUSTAT Item
Total Debt/Market Value of Total Assets (L=TD/MTA)	CS;CRSP	(CS9+CS34)/(CS9+CS34+MCAP)
EBIT/Total Assets (EBIT/TA)	CS	(CS18+CS15+CS16)/CS6
EBIT/Interest Expense (EBIT/XINT)	CS	(CS18+CS15+CS16)/CS15
ln(MCAP/Cap. of S&P500) (SIZE)	CRSP	
Growth in Assets ($dTA = \frac{TA_t - TA_{t-1}}{TA_t}$)	CS	CS6
Return over the previous 12 months (RET)	CRSP	
Volatility of Returns over the previous 12 months (VOLA)	CRSP	
Panel B - Leverage Forecast		
Covariate	Source	CRSP/COMPUSTAT Item
Total Debt/Market Value of Total Assets (L=TD/MTA)	CS;CRSP	(CS9+CS34)/(CS9+CS34+MCAP)
EBIT/Total Assets (EBIT/TA)	CS	(CS18+CS15+CS16)/CS6
ln(MCAP/Cap. of S&P500) (SIZE)	CRSP	
Growth in Assets ($dTA = \frac{TA_t - TA_{t-1}}{TA_t}$)	CS	CS6
Depreciation Expense/Total Assets (XD/TA)	CS	CS14/CS6
Fixed Assets/Total Assets (FA/TA)	CS	CS8/CS6
R&D Expenses/Total Assets (R&D/TA)	CS	CS46/CS6
R&D Dummy	CS	CS46
Market-to-Book Ratio (MB)	CS;CRSP	(CS9+CS34+MCAP)/CS6
Mean Industry Leverage (INDMEAN)	CS;CRSP	(CS9+CS34)/(CS9+CS34+MCAP)

Table 6: Definition of Covariates: Panel A gives an overview of covariates used for default prediction, Panel B lists variables that explain leverage levels. The CRSP item "market capitalization" is defined as MCAP = price per share*shares outstanding.

Panel A - Default Prediction					
Covariate	Obs.	Mean	Std. Dev.	Min.	Max.
L = TD/MTA	17580	0.34	0.23	0	0.92
EBIT/TA	17580	0.09	0.09	-0.26	0.31
EBIT/XINT	17580	3.23	1.84	0.00	5.00
RET	17580	0.15	0.44	-0.73	1.89
VOLA	17580	0.11	0.06	0.04	0.34
dTA	17580	0.06	0.17	-0.62	0.56
SIZE	17580	-8.42	1.90	-15.81	-2.81
Panel B - Leverage Forecast					
Covariate	Obs.	Mean	Std. Dev.	Min.	Max.
L = TD/MTA	15770	0.33	0.22	0.00	0.92
EBIT/TA	15770	0.10	0.09	-0.26	0.31
MB	15770	1.26	0.89	0.37	5.7
XD/TA	15770	0.05	0.02	0.01	0.16
FA/TA	15770	0.35	0.20	0.02	0.89
R&D Dummy	15770	0.38	0.49	0	1
R&D/TA	15770	0.02	0.03	0.00	0.18
INDMEAN	15770	0.34	0.06	0.19	0.72
SIZE	15770	-8.32	1.86	-14.88	-2.81
dTA	15770	0.06	0.16	-0.62	0.56

Table 7: Descriptive statistics: Panel A summarizes variables used in default prediction, Panel B describes the set of covariates explaining leverage levels.

	Forecast Horizon in Years			
	1	2	3	4
Within Regression	0.162	0.244	0.277	0.271
IV Regression with Fixed Effects	0.159	0.242	0.276	0.271
IV Regression without Fixed Effects	0.123	0.161	0.177	0.197
Arellano and Bond's (1991) GMM Procedure	0.126	0.163	0.187	0.205

Table 8: Leverage Forecasts: out-of-sample root mean squared errors for a variation of estimation techniques applicable to the partial adjustment model.

	Forecast Horizon in Years			
	1	2	3	4
L	0.8927***	0.8037***	0.7337***	0.6656***
	(0.0085)	(0.015)	(0.0208)	(0.0261)
EBIT/TA	0.0358*	0.0805***	0.1057***	0.1155**
	(0.0152)	(0.0221)	(0.0286)	(0.0351)
SIZE	-0.0023**	-0.0027*	-0.0031	-0.0022
	(0.0007)	(0.0014)	(0.002)	(0.0025)
dTA	0.0462***	0.0644***	0.0731***	0.0722***
	(0.0071)	(0.0097)	(0.0123)	(0.0138)
XD/TA	-0.0719	-0.1908*	-0.1713	-0.1722
	(0.0491)	(0.087)	(0.1231)	(0.1595)
FA/TA	0.0215***	0.0391***	0.041**	0.0494*
	(0.0059)	(0.0107)	(0.0151)	(0.0195)
R&D Dummy	-0.0007	-0.0009	-0.0003	0.0014
	(0.002)	(0.0036)	(0.0051)	(0.0065)
R&D/TA	-0.1123***	-0.1906***	-0.291***	-0.382***
	(0.0266)	(0.049)	(0.0676)	(0.0839)
MB	-0.0063***	-0.0096***	-0.0118***	-0.0137***
	(0.0014)	(0.0023)	(0.003)	(0.0038)
INDMEAN	0.031	0.0274	-0.0006	-0.0294
	(0.0194)	(0.0354)	(0.0482)	(0.0582)
Constant	0.0061	0.0260	0.1081***	0.1022***
	(0.0108)	(0.0187)	(0.0249)	(0.0295)
Firm-years	14829	13293	11932	11046

Table 9: Partial adjustment model: IV Regression without fixed effects; lagged leverage ratio is instrumented by book debt ratio. Standard errors are in parentheses. Stars indicate a significance level: $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$. Leverage ratio forecasts one to four years ahead are used in the second step of long term default analysis.

	Prediction Horizon in Years			
	2	3	4	5
Panel A Years in the estimation sample: 8				
AUC Two-Step Dyn. Model	0.8651	0.792	0.674	0.5564
AUC Base Model	0.8596	0.7723	0.6232	0.5353
Change in %	0.6	2.6	8.2	3.9
P-value	0.2795	0.0035	<0.001	0.0143
Panel B Years in the estimation sample: 9				
AUC Two-Step Dyn. Model	0.8747	0.8356	0.7652	0.638
AUC Base Model	0.8652	0.7908	0.6874	0.5814
Change in %	1.2	5.7	11.3	9.7
P-value	0.0535	<0.001	<0.001	<0.001
Panel C Years in the estimation sample: 10				
AUC Two-Step Dyn. Model	0.8871	0.8468	0.7836	0.7063
AUC Base Model	0.8669	0.8019	0.7195	0.6327
Change in %	2.3	5.6	8.9	11.6
P-value	0.0019	<0.001	<0.001	<0.001
Panel D Years in the estimation sample: 11				
AUC Two-Step Dyn. Model	0.8866	0.8388	0.7853	0.7061
AUC Base Model	0.8712	0.8154	0.7537	0.6773
Change in %	1.8	2.9	4.2	4.3
P-value	0.0012	0.0254	0.0801	0.1377
Panel E Years in the estimation sample: 12				
AUC Two-Step Dyn. Model	0.8811	0.8292	0.7618	0.6713
AUC Base Model	0.8753	0.8198	0.7647	0.6900
Change in %	0.7	1.1	-0.4	-2.7
P-value	0.3279	0.2462	0.8408	0.2455
Panel F Years in the estimation sample: 13				
AUC Two-Step Dyn. Model	0.8790	0.7974	0.7139	0.6466
AUC Base Model	0.8781	0.8187	0.765	0.6861
Change in %	0.1	-2.6	-6.7	-5.8
P-value	0.8214	0.0079	<0.001	0.0025
Firm-years	6861	6592	6342	6086

Table 10: Two-step dynamic vs. base model: out-of-sample prediction accuracy for varying estimation samples. Number of years in the estimation samples ranges from 8 to 13. Change in % shows the increase in accuracy for the two-step dynamic model. P-values refer to the null hypotheses of equality of areas under two ROC curves.

	Prediction Horizon in Years			
	2	3	4	5
Panel A Years in the estimation sample: 8				
AUC Reduced-Form Dyn. Model	0.8436	0.7637	0.6672	0.5586
AUC Base Model	0.8596	0.7723	0.6232	0.5353
Change in %	-1.9	-1.1	7.1	4.4
P-value	0.0786	0.6461	0.0245	0.3137
Panel B Years in the estimation sample: 9				
AUC Reduced-Form Dyn. Model	0.8591	0.8146	0.7345	0.6491
AUC Base Model	0.8652	0.7908	0.6874	0.5814
Change in %	-0.7	3.0	6.9	11.6
P-value	0.4360	0.1152	0.0469	0.0068
Panel C Years in the estimation sample: 10				
AUC Reduced-Form Dyn. Model	0.8559	0.8051	0.7267	0.6645
AUC Base Model	0.8669	0.8019	0.7195	0.6327
Change in %	-1.3	0.4	1.0	5.0
P-value	0.1995	0.8448	0.7573	0.2893
Panel D Years in the estimation sample: 11				
AUC Reduced-Form Dyn. Model	0.8608	0.8195	0.7694	0.6748
AUC Base Model	0.8712	0.8154	0.7537	0.6773
Change in %	-1.2	0.5	2.1	-0.4
P-value	0.1607	0.7681	0.4938	0.9298
Panel E Years in the estimation sample: 12				
AUC Reduced-Form Dyn. Model	0.8389	0.7869	0.7333	0.6376
AUC Base Model	0.8753	0.8198	0.7647	0.6900
Change in %	-4.2	-4.0	-4.1	-7.6
P-value	0.0015	0.0125	0.1700	0.0592
Panel F Years in the estimation sample: 13				
AUC Reduced-Form Dyn. Model	0.8531	0.7870	0.7007	0.5786
AUC Base Model	0.8781	0.8187	0.765	0.6861
Change in %	-2.8	-3.9	-8.4	-15.7
P-value	0.0049	0.0026	<0.001	<0.001
Firm-years	6861	6592	6342	6086

Table 11: Reduced-form dynamic vs. base model: out-of-sample prediction accuracy for varying estimation samples. Number of years in the estimation samples ranges from 8 to 13. Change in % shows the increase in accuracy for the reduced-form dynamic model. P-values refer to the null hypotheses of equality of areas under two ROC curves.

	Prediction Horizon in Years			
	2	3	4	5
Panel A				
Years in the estimation sample: 10 (leverage), 11 (default)				
AUC Two-Step Dyn. Model	0.8877	0.8514	0.7983	0.7187
AUC Base Model	0.8712	0.8154	0.7537	0.6773
Change in %	1.9	4.4	5.9	6.1
P-value	<0.001	<0.001	0.0031	0.0093
Panel B				
Years in the estimation sample: 10 (leverage), 12 (default)				
AUC Two-Step Dyn. Model	0.8899	0.8504	0.8013	0.7207
AUC Base Model	0.8753	0.8198	0.7647	0.6900
Change in %	1.7	3.7	4.8	4.4
P-value	<0.001	<0.001	0.0060	0.0280
Panel C				
Years in the estimation sample: 10 (leverage), 13 (default)				
AUC Two-Step Dyn. Model	0.8916	0.8478	0.7985	0.7215
AUC Base Model	0.8781	0.8187	0.7650	0.6861
Change in %	1.5	3.6	4.4	5.2
P-value	<0.001	<0.001	0.0060	0.0011
Firm-years	6861	6592	6342	6086

Table 12: Two-step dynamic vs. base model: out-of-sample prediction accuracy for varying estimation samples. Leverage forecasts are based on a 10 years history. The estimation sample for default analysis covers 11 to 13 years. Change in % shows the increase in AUC for the two-step dynamic model. P-values refer to the null hypotheses of equality of areas under two ROC curves.

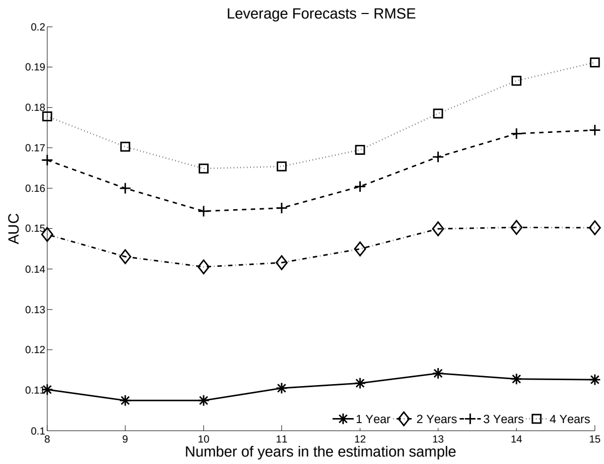

Figure 5: Leverage forecasts 1 to 4 years ahead: out-of-sample root mean squared errors for varying estimation samples. (IV Regression without fixed effects.)

References

Allison, P., 1982, *Discrete-time Methods for the Analysis of Event Historie.* Sociological Methodology, 61-99.

Altman, E.I., 1968, *Financial ratios, discriminant analysis and the prediction of corporate bankruptcy.* The Journal of Finance, 23, 589-609.

Arellano, M., Bond, S., 1991. *Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations.* Review of Economic Studies 58: 277-97.

Baker, M. and Wurgler, J., 2002, *Market Timing and Capital Structure.*, The Journal of Finance, 57(1), 132.

Baltagi, B.H., 2003, *Econometric analysis of panel data.* John Wiley & Sons, Ltd.

Banerjee, S., Heshmati, A. and Whilborg, C., 2000, *The Dynamics of Capital Structure* SSE/EFI Working Paper Series in Economics and Finance, No. 333

Beaver, W.H., McNichols, M.F. and Rhie, J.-W., 2004, *Have Financial Statements Become Less Informative? Evidence from the Ability of Financial Ratios to Predict Bankruptcy.* Working Paper.

Beck, N., Katz, J.N. and Tucker, R., 1998, *Taking Time Seriously: Time-Series-Cross-Section Analysis with a Binary Dependent Variable.* American Journal of Political Science, 42, 1260-1288.

Blume M.E., Lim F., MacKinley A.C., 1998, *The declining credit quality of U.S. corporate debt: myth or reality?* The Journal of Finance, 53, 1389-1413.

Bond S., 2002, *Dynamic panel data models: a guide to micro data and practice.* The Institute for Fiscal Studies, Department of Economics, UCL, cemmap working paper CWP09/02.

Campbell, J.Y., Hilscher, J. and Szilagyi, J., 2006, *In Search of Distress Risks.* Harvard Institute of Economic Research Discussion Paper No. 2081.

Chava, S., Jarrow R.A., 2004, *Bankruptcy Prediction with Industry Effects.* Review of Finance, 8, 537-569.

Collin-Dufresne, P., Goldstein, R., 2001, *Do credit spreads reflect stationary leverage ratios?* Journal of Finance 56, 1929-1957.

Duffie, D., Saita, L., Wang, K., 2006, *Multi-period corporate default prediction with stochastic covariates.* Journal of Financial Economics, Forthcoming.

DeLong, E.R., DeLong, D.M., Clarke-Pearson, D.L., 1988, *Comparing the areas under two or more correlated receiver operating curves: A nonparametric approach.* Biometrics, 44, 837-845.

Fama, E.F., French, K.R., 1997, *Industry costs of equity.* Journal of Financial Economics 43, 153-193.

Fama, E.F., French, K.R., 2002, *Testing tradeoff and pecking order predictions about dividends and debt.* Review of Financial Studies 15, 1-33.

Flannery, M.J., Rangan, K.P., 2006, *Partial adjustment toward target capital structure.* Journal of Financial Economics 79, 469-506.

Fons, J.S., Cantor, R., Mahoney C., 2002, *Understanding Moody's corporate bond ratings and rating process.* Moody's Investor Service, Special Comment.

Graham, J., Harvey, C., 2001, *The theory and practice of corporate finance: Evidence from the field.* Journal of Financial Economics 60, 187-243.

Hamilton, D.T., Varma, P., Ou, S., and Cantor, R., 2006, *Default and Recovery Rates of Corporate Bond Issuers, 1920-2005.* Moody's Investor Service, Special Comment.

Hillegeist, S.A., Keating, E.K., Cram, D.P. and Lundstedt, K.G., 2004, *Assessing the Probability of Bankruptcy.* Review of Accounting Studies 9, 5-34.

Howakimian, A. Opler, T. and Titman, S., 2001, *The Debt-Equity Choice.* Journal of Financial and Quantitative Analysis 36, 3247-3259.

Jenkins, S.P., 1995, *Easy Estimation Methods for Discrete-Time Duration Models.* Oxford bulletin of Economics and Statistic 57.

Leary, M., Roberts, M., 2005, *Do firms rebalance their capital structure?* Journal of Finance 60, 2575-2619.

Lo, C.F., Wong, T.C., Hui, C.H. and Huang, M.X., 2008, *Assessing Credit Risk of Companies with Mean-Reverting Leverage Ratios.* HKIMR Working Paper No. 4/2008

Myers, S. C., 1984, *The Capital Structure Puzzle.*, The Journal of Finance, 39(3), 575-592.

Metz, A., Cantor, R., 2006, *Moody's Credit Rating Prediction Model.* Moody's Investor Service, Special Comment.

Ohlson, J.S., 1980, *Financial Ratios and the probabilistic prediction of bankruptcy.* Journal of Accounting Research, 19, 109-131.

Petersen M.W., 2007, *Estimating standard errors in finance panel data sets: Comparing approaches.* Kellogg Finance Dept. Working Paper No. 329.

Rajan, R.G., Zingales L., 1995, *What do we know about capital structure? Some evidence from international data.* The Journal of Finance 50, 1421-1458.

Roberts, Michael R., 2002, *The Dynamics of Capital Structure: An Empirical Analysis of a Partially Observable System.*, Duke University, Fuqua School of Business Working Paper.

Roodman, D., 2006, *How to do xtabond2: An introduction to "Difference" and "System" GMM in Stata.* Center for Global Development, Working Paper No. 103.

Shumway, T., 2001, *Forecasting bankruptcy more accurately: A simple hazard rate model.* Journal of Business 74, 101-124.

Shyam-Sunder, L., Myers, S., 1999, *Testing static tradeoff against pecking order models of capital structure.* Journal of Financial Economics 51, 469-506.

Zmijewski, M.E., 1984, *Methodological issues related to the estimation of financial distress prediction models*, Journal of Accounting Research, 22, 59-82.