Countercyclical and time-varying risk aversion and the equity premium

Jan Antell and Mika Vaihekoski*

September 13, 2017

Abstract

This paper tests the counter-cyclicality of aggregate risk aversion and price of market risk using a novel testing approach introduced in Antell and Vaihekoski (2015) for conditional asset pricing models. Cohen et al. (2015) report experimental evidence that the risk aversion is countercyclical, although empirical support from financial studies is at best inconclusive. This paper applies the new testing approach for the Merton (1973, 1980) model. The testable implications link the realized equity premium to, among others, changes in conditional variance, its long-term persistence, and changes in the time-varying risk aversion. Empirically, the testing is conducted using monthly US stock market data from 1928 to 2013, and GARCH models to estimate time-varying variance. Various methods to model economic expectations are compared. Unlike the traditional estimation approach, the results from the new estimation approach give clear support for time-varying and countercyclical risk aversion.

Keywords: conditional asset pricing, equity premium, GARCH, risk aversion

JEL classification: F3, G12, G15.

*In alphabetical order. Jan Antell is with the Hanken School of Economics, Department of Finance and Statistics. Address: P.O. Box 479, 00101 Helsingfors, Finland. E-mail: jan.antell@hanken.fi. Mika Vaihekoski is with the Turku School of Economics, University of Turku. Address: Turku School of Economics, Department of Accounting and Finance, 2004 University of Turku, Finland. E-mail: mika.vaihekoski@utu.fi. This research was finished, in part, while Vaihekoski was visiting Scancor Unit at Stanford University. We are grateful for the helpful comments from conference participants at the 23rd Annual Conference of the Multinational Finance Society in Stockholm, Sweden and and 2017 FMA European Conference in Prague, Portugal as well as seminar participants at Stockholm Business School, Turku School of Economics and Hanken School of Economics.
1. INTRODUCTION

The Merton (1973 and 1980) model suggests a positive price of market risk parameter, lambda, which links the equity market premium and the variance of the market together, conditional on available information. Under certain assumptions (e.g., power utility function), it can be shown that this price of risk parameter equals to or is a function of the aggregate relative risk aversion measure (Guo and Whitelaw, 2006). A number of empirical financial studies have shown that the risk aversion is time-varying (see, e.g., Smith and Whitelaw, 2009; Guo, Wang and Yang, 2013; Guiso et al., 2013; Cohen et al., 2015). However, opposite results have also been presented (see, e.g., Andersen et al., 2008). In addition, many studies have found it difficult to find significant and positive relationship between realized variance and returns. For example, Yu and Yuan (2011) find that the relationship is positive in low-sentiment periods but not in high-sentiment periods and that this result is strikingly robust to different methods. However, if sentiment-defined regimes are not considered, they find mixed results, the lambda being either negative or positive depending on the method, and smaller in magnitude than during low-sentiment periods. Feunou et al. (2014) review the literature and note that the empirical support for the relationship is “remarkable uneven”.

The financial theory suggests that the price of market risk and aggregate risk aversion should behave in a countercyclical way, that is, during a recession risk aversion should be higher (Campbell and Cochrane, 1999; Guo et al., 2013). Cohen et al. (2015) note that the evidence for this is scarce. For example Yu and Yuan (2011) note that it is very difficult for their empirical results to fit in the existing hypotheses with either constant or time-varying risk aversion. Cohen et al. (2015) note that this is due to “the host of factors that simultaneously change during financial cycles”. The risk aversion and the volatility both exhibit strong countercyclical patterns. Thus it is hard to separate the effects from each other in empirical tests.

This paper argues that the earlier mixed results for pricing market risk as well as for the time-varying price of risk are in part fact due to the weaknesses in the empirical testing approach used to test conditional asset pricing models; using realized returns or an auxiliary expectations model as a proxy for expected returns in tests of conditional asset pricing models is problematic. The testing approach in Antell and Vaihkoski (2015) avoids these problems. They derive a testable model for realized returns conditional on the tested asset pricing model being true – a reverse testing approach, if you will.\footnote{The model is based on Campbell and Hentschel (1992) model for realized returns.} Obviously, if the model is not the one used by the investors, the realized returns would reflect this and the model
would be rejected. They test the approach on the Merton (1973 and 1980) model and find strong support for a positive risk-return relationship on the US market using a constant price of market risk parametrization of the model. Here we take the analysis one step further and derive an empirically testable model in which the relative risk aversion is allowed to be time-varying and we test whether it behaves in a counter-cyclical fashion.

Empirically, we compare the new approach to estimating the price of market risk with the traditional estimation approach. We test the models first using monthly US stock market data from 1928 to 2013. The variance process is modelled using (asymmetric) GARCH models. The price of risk is modelled linear on conditioning variables that are related to economic expectations. A number of ways to model the economic expectations are applied in the empirical analysis. The results show strong and consistent support for the time-varying price of risk and risk aversion using the new estimation approach. Risk aversion is also found to behave in a countercyclical way. The results also clearly show that the traditional testing approach cannot explain the return variance relationship even when one allows for time-varying price of risk.

The remainder of the paper is organized as follows. Section 2 presents the theoretical background, derives a model for the realized returns conditional on the Merton (1973, 1980) model for the return-variance relationship, and discusses the empirical research methodology as well as econometric issues. Section 3 introduces the data used in this paper. Section 4 shows the empirical results. Section 5 presents the conclusions and offers suggestions for further research.

2. THEORETICAL BACKGROUND

2.1. Risk-return models

Merton (1973, 1980) showed that the model for the excess return on the market portfolio, i.e. equity premium, can be written as

\[
E[r_{m,t+1}^e|\Omega_t] = \lambda_{m,t+1}Var(r_{m,t+1}|\Omega_t),
\]

(1)

where \(E[r_{m,t+1}^e|\Omega_t]\) is the expected excess return on the market portfolio and \(Var(r_{m,t+1}|\Omega_t)\) is the variance of the excess market return, both conditional on investors’ information set \(\Omega_t\) available at time \(t\). Lambda, \(\lambda_{m,t+1}\), is a time-varying measure commonly labeled as the conditional price of market risk, also conditional on \(\Omega_t\). Under certain assumptions, it can also be seen as the aggregate relative risk aversion measure, \(-U_{ww}^\prime \cdot W \cdot (U_w^\prime)^{-1}\), where \(U\) is a utility function for the representative investor, \(W\) is wealth, and \(U^\prime\) represents partial
derivatives of the utility function. Equation (1) basically shows that investors must be compensated by a higher expected return if the conditional variance or the price of market risk (risk aversion) increases.

The theoretical model (1) has a number of empirical implications that can be tested. To do this, one must provide empirical proxies for expected returns, lambda, and conditional variance. Typically, the time-varying price of market risk has been modelled as linear on conditioning information variables as

\[ \lambda_{m,t+1} = \lambda_0 + \lambda_1 Z_t, \]  

(2)

where \( \lambda_0 \) is a constant, \( \lambda_1 \) is a \( K \times 1 \) vector of parameters to be estimated, and \( Z_t \) is a \( K \times 1 \) vector of conditional instrument variables available at time \( t \). Some studies have also modeled the price of risk using an exponential model to force positivity (see, e.g., Bekaert and Harvey, 1995; De Santis and Gérard, 1997; Carrieri, Errunza, and Hogan, 2007).

Earlier empirical studies have almost always been conducted as if expected returns can be proxied with the realized returns. An alternative is to use a linear expectations model for the realized returns; one of the first ones to use this approach were Gibbons and Ferson (1985). Given an estimate for the variance and for the price of market risk, one typically proceeds to estimate equation (2) using the following linear model:

\[ r_{e,m,t+1} = \mu + \lambda_{m,t+1} \sigma^2_{m,t+1} + \varepsilon_{m,t+1}, \]  

(3)

where \( r_{e,m,t+1} \) is the realized excess market return from time \( t \) to \( t+1 \), \( \mu \) is a constant expected to be zero if excess returns are used and the asset pricing model is valid, \( \lambda_{m,t+1} \) is the time-varying price of market risk, and \( \sigma^2_{m,t+1} \) is the conditional variance for the period from \( t \) to \( t+1 \), given the information available at time \( t \). We refer to using this equation as the traditional approach to estimating lambda.

Using realized returns as a proxy for the expected returns has been justified by the rational expectations assumption. However, the rational expectations hypothesis states that investors’ expectations may be wrong in the short run. Since the inherited nature of the conditional models couples with time-varying expectations, we argue that realized returns are a biased proxy for the expected returns for short return measurement intervals, and therefore this approach is problematic for empirical tests of conditional asset pricing models. Antell and Vaihekoski (2015) present a reverse testing approach for estimating conditional asset pricing models which is based on the idea of analyzing realized returns conditional on the asset pricing model being true (c.f. Guo and Whitelaw, 2006). Campbell and Hentschel (1992) show that continuously compounded realized returns, \( r_{t+1} \), for a dividend paying
security can be written as

\[ r_{t+1} \approx k_1 + (1 - \rho) \left[ \sum_{i=1}^{\infty} \rho^i \left( d_{t+1,t+1+i} - \rho^{-1} d_{t,t+i} \right) \right] + \sum_{i=1}^{\infty} \rho^{i-1} \left( r_{t,t+i} - \rho r_{t+1,t+1+i} \right), \tag{4} \]

where \( r_{t,t+i} = E_t[r_{t+i}] \) expresses the continuously compounded required rate of return for the period \( t+i \), conditional on information available at time \( t \). \( d_{t,t+i} = E_t[\ln(D_{t+i})] \) is the expected log dividend at time \( t+i \) \((i > 0)\) conditional on information available at time \( t \). The terms \( r_{t+1,t+1+i} \) and \( d_{t+1,t+1+i} \) are defined in a similar fashion although both are conditional on information available at time \( t+1 \). The parameter \( \rho \) is positive and less than one by definition. Campbell, Lo, and MacKinlay (1997) note that \( \rho \) should be 0.997 for monthly data.

Now taking our candidate asset pricing model (1) and using it to define conditional expected (required) returns for the market portfolio, we can rewrite the last term of equation (4) for the market portfolio as follows

\[ \sum_{i=1}^{\infty} \rho^{i-1} \left( r_{t,t+i} - \rho r_{t+1,t+1+i} \right) \]

\[ = \sum_{i=1}^{\infty} \rho^{i-1} \left( r_{m,t+i} - \rho r_{m+1,t+1+i} \right) + \sum_{i=1}^{\infty} \rho^{i-1} \left( \lambda_{m,t+1} \sigma_{t,t+i}^2 - \rho \lambda_{m,t+2} \sigma_{t+1,t+1+i}^2 \right). \tag{5} \]

Following earlier studies, we assume that lambda is a linear function of conditioning information variables, and use equation (2) for \( \lambda_{m,t+1} \) and \( \lambda_{m,t+2} \). Utilizing these definitions, we can rewrite the second term in equation (5) as follows

\[ \sum_{i=1}^{\infty} \rho^{i-1} \left[ \left( \lambda_0 + \lambda_1 Z_i \right) \sigma_{t,t+i}^2 - \rho \left( \lambda_0 + \lambda_1 Z_{t+1}^i \right) \sigma_{t+1,t+1+i}^2 \right] \]

\[ = \lambda_0 \left[ \sum_{i=1}^{\infty} \rho^{i-1} (\sigma_{t,t+i}^2 - \rho \sigma_{t+1,t+1+i}^2) \right] + \lambda_1 \left[ \sum_{i=1}^{\infty} \rho^{i-1} (\sigma_{t,t+i}^2 Z_i^i - \rho \sigma_{t+1,t+1+i}^2 Z_{t+1}^i) \right]. \tag{6} \]

Using the assumption that the conditional variance is a mean-reverting process (cf., e.g., Engle and Patton, 2001) and that one-step-ahead forecasts can be assessed, we can calculate conditional multistep forecasts \( i \) periods ahead using the conditional variance for the next period. Here, we further assume that the conditional variance for any future period \( i \geq 1 \) can be expressed as a function of the next period’s forecast as follows:
\[
\sigma^2_{t,t+i} = \phi^{i-1}\sigma^2_{t,t+1} + \sigma^2(1 - \phi^{i-1}),
\]

where \(|\phi| < 1\) is a persistence parameter reflecting the speed of convergence of the conditional variance toward the long-term unconditional variance \(\sigma^2\). As a result, equation (4) can be written as

\[
rm_{t+1} \approx k_1 + (1 - \rho)\left[\sum_{i=1}^{\infty} \rho^i(d_{t+1,t+1+i} - \rho^{-1}d_{t,t+i})\right] + \sum_{i=1}^{\infty} \rho^{i-1}(rf_{t+1,t+1+i} - \rho rf_{t+1,t+1+i})
+ \lambda_0 \sum_{i=1}^{\infty} \rho^{i-1} [\phi^{i-1}(\sigma^2_{t,t+1} - \rho \sigma^2_{t+1,t+2}) + \sigma^2(1 - \rho)(1 - \phi^{i-1})]
+ \lambda_1 \sum_{i=1}^{\infty} \rho^{i-1}\{Z'_{t}[\phi^{i-1}\sigma^2_{t,t+1} + \sigma^2(1 - \phi^{i-1})] - \rho Z'_{t+1}[\phi^{i-1}\sigma^2_{t+1,t+2} + \sigma^2(1 - \phi^{i-1})]\}.
\]

Introducing two new variables, \(\varphi_{\Delta \sigma}\) and \(\varphi_{\sigma}\), the equation can be expressed in a more compact form as

\[
rm_{t+1} \approx k_1 + (1 - \rho)\left[\sum_{i=1}^{\infty} \rho^i(d_{t+1,t+1+i} - \rho^{-1}d_{t,t+i})\right] + \sum_{i=1}^{\infty} \rho^{i-1}(rf_{t+1,t+1+i} - \rho rf_{t+1,t+1+i})
+ \lambda_0 \sum_{i=1}^{\infty} \rho^{i-1} [\sigma^2_{t,t+1} - \rho \sigma^2_{t+1,t+2} + \sigma^2\varphi_{\Delta \sigma}]
+ \lambda_1 \sum_{i=1}^{\infty} \rho^{i-1}\{Z'_{t}\sigma^2_{t,t+1} - Z'_{t+1}\rho \sigma^2_{t+1,t+2} + \sigma^2\varphi_{\sigma}\},
\]

where \(\varphi_{\Delta \sigma}\) is equal to \(1/(1 - \rho \phi)\) and \(\varphi_{\sigma}\) equals \((1 - \varphi_{\Delta \sigma} \cdot (1 - \rho))\). Henceforth \(\varphi_{\Delta \sigma}\) and \(\varphi_{\sigma}\) are collectively called sigma multipliers. In theory, if the variance persistence parameter \(\phi\) equals, say, 0.9 and \(\rho\) equals 0.997 for monthly data, \(\varphi_{\Delta \sigma}\) equals \(1/(1 - 0.997 \cdot 0.9) = 9.74\), and \(\varphi_{\sigma} = 0.97\). The parameter \(\varphi_{\Delta \sigma}\) indicates how much changes in the conditional variance over one period are magnified due to the persistence of variance.

Now, applying certain simplifying assumptions\(^2\) on expected dividends and term structure, equation (9) can be expressed in a simplified form as follows

\(^2\) The term structure is assumed to be flat. Investors are assumed to revise their views on dividend growth rate from time \(t\) to \(t + 1\), but the difference between the new and the old view converges geometrically to zero for period \(t + i\) when \(i \to \infty\).
\[
\begin{align*}
\bar{r}_{m,t+1} & \approx k_2 + (g_{t+1,t+2} - g_{t,t+2}) \cdot \varphi_d + (r_{ft} - r_{ft+1}) \cdot \varphi_{rf} \\
& + \lambda_0 \left( (\sigma^2_{t,t+1} - \sigma^2_{t+1,t+2}) \cdot \varphi_{\Delta \sigma} + \sigma^2 \cdot \varphi_{\sigma} \right) \\
& + \lambda_1 \left( (Z'_t \sigma^2_{t,t+1} - Z'_{t+1} \sigma^2_{t+1,t+2}) \cdot \varphi_{\Delta \sigma} + \sigma^2 \cdot (Z'_t - \rho Z'_{t+1}) \cdot (1 - \rho)^{-1} \cdot \varphi_{\sigma} \right),
\end{align*}
\]

where \( g_{t+1,t+2} \) and \( g_{t,t+2} \) are conditional dividend growth rates. Parameters \( \varphi_{rf} \) and \( \varphi_d \) measure the impact of the changes in the risk-free rate and investors’ expectations for the dividend growth rate, respectively. Both of them are by definition positive. The constant \( k_2 \) can be either positive or negative.

Analyzing equation (10) reveals that realized returns are higher if investors’ conditional expectations of the long-term dividend growth rate increase from period \( t \) to \( t+1 \), ceteris paribus. The same is true if the risk-free interest rate decreases. Assuming that the asset pricing model is correct, a decrease in conditional volatility also leads to higher realized returns. If there is a positive relationship between the conditioning information variable and the risk aversion (i.e., \( \lambda_1 \) is positive), an increase in the value of the information variable should lead to lower realized return. It is easy to see that this is the case. Thus all implications of the model are in line with the intuition. Equation (10) also allows us to separate the effect of changes in the variance and in the relative risk aversion.

2.2. Empirical estimation

The main objective of this paper is to test whether the price of market risk, or lambda, is time-varying and whether it behaves in a countercyclical fashion. In addition, we also want to compare the estimate of lambda from the traditional approach with the estimate from the new estimation approach. To estimate lambdas, we need an empirical proxy for the conditional variance. Antell and Vaihekoski (2015) compared GARCH, MIDAS, and the VIX-index as estimates of the variance, but they concluded that, overall, the choice of the variance estimation method does not have a major effect on the empirical performance of the models. Hence, here we utilize GARCH models in the estimation.

To estimate the traditional lambda, we combine equations (2) and (3) and a GARCH in mean model. In practice, if a GARCH(1,1) model is selected for the variance process, the variance of the excess market returns is given by the following equation:

\[
\sigma^2_{m,t+1} = \omega + \alpha \varepsilon^2_{m,t} + \beta \sigma^2_{m,t},
\]  

where the parameters \( \omega \), \( \alpha \) and \( \beta \) relate to the GARCH(1,1) variance specification. The
estimation can proceed in one phase.

To estimate the lambda using the new approach, we estimate model (10) in two phases for two reasons. First, the reverse testing specification uses the change in the conditional variance in the mean equation rather than the variance per se. Second, unless separated, the variance specification affects the mean equation (through $\phi \Delta \sigma$ and $\phi \sigma$) which in turn affects the variance estimation.\(^3\)

In the first phase, we estimate the conditional variance process for the market using the variance equation (11). Given the conditional variance series, we calculate the speed of convergence in the variance process ($\phi$), unconditional variance ($\sigma^2$), and the ensuing sigma multipliers. In the second step we run a linear regression model where we have operationalized equation (10) in excess-return form as follows

$$r_{m,t+1} = b_1 + b_2 (g_{t+1,t+2} - g_{t,t+2}) + b_3 (r_{ft} - r_{ft+1})$$
$$+ b_4 \left[ (\sigma_{t,t+1}^2 - \sigma_{t+1,t+2}^2) \cdot \varphi_{\Delta \sigma} + \sigma^2 \cdot \varphi_{\sigma} \right]$$
$$+ b_5 \left[ (Z_i' \sigma_{t,t+1}^2 - \rho Z_{t-1} \sigma_{t+1,t+2}^2) \cdot \varphi_{\Delta \sigma} + \sigma^2 \cdot (Z_i' - \rho Z_{t-1}') \cdot (1 - \rho)^{-1} \cdot \varphi_{\sigma} + u_{m,t+1} \right], \quad (12)$$

where $b_1$ to $b_5$ are the coefficients to be estimated. All coefficients except $b_5$ are expected to be positive given their definitions. $b_4$ is our estimate for the unconditional, long-term mean lambda given that the information variables have been demeaned. The parameters in vector $b_5$ show time-variation in lambda. The risk-free rate at time $t$ is given by $r_{ft}$. The term $\sigma_{t,t+1}^2$ is the variance of the continuously compounded excess market return from time $t$ to $t+1$, conditional on information available at time $t$. Variables are defined similarly for time $t+1$. Other variables are as defined earlier.

As our main interest is to get an estimate for the price of market risk and to assess whether it behaves in a counter-cyclical fashion, we begin our estimation with the following baseline version of the model. It is shortened under the assumption that changes in both the interest rate level and the dividend growth rates are of lesser importance.

$$r_{m,t+1} = b_1 + b_2 \left[ (\sigma_{t,t+1}^2 - \sigma_{t+1,t+2}^2) \cdot \varphi_{\Delta \sigma} + \sigma^2 \cdot \varphi_{\sigma} \right]$$
$$+ b_3 \left[ (Z_i' \sigma_{t,t+1}^2 - \rho Z_{t-1} \sigma_{t+1,t+2}^2) \cdot \varphi_{\Delta \sigma} + \sigma^2 \cdot (Z_i' - \rho Z_{t-1}') \cdot (1 - \rho)^{-1} \cdot \varphi_{\sigma} + u_{m,t+1} \right], \quad (13)$$

where $b_1$ is also expected to account for the mean effect from the components excluded

\(^3\) Rossi and Timmermann (2015) also utilize a two-step estimation strategy. When estimating the conditional variance series, they utilize an EGARCHX model with the residuals conditional on the same X variables (tbill-rate and dp-ratio).
from the model and $b_2$ is our estimate for the lambda. The parameter $b_3$ accounts for the time-variation in lambda.

In practice, in the first phase, we estimate the GARCH model for the excess market returns using only a constant in the mean equation. In the second step, we estimate equation (14) or (15) using estimates for conditional variance as well as $\varphi_{\Delta \sigma}$, $\varphi_{\sigma}$, and the unconditional variance, $\sigma^2$, from the first step. To estimate $\varphi_{\Delta \sigma}$ and $\varphi_{\sigma}$, we use definitions (??) and (??). They require an estimate for the speed of conditional variance returning to its long-term mean, i.e., the $\phi$ parameter and the dividend-to-price-related $\rho$ parameter. The latter can be easily calculated from the data, but the former utilizes the results from the model for the conditional variance. Assuming that the conditional variance follows a GARCH(1,1) process, we can write the $i$-step ahead forecasts for the conditional variances as a combination of the next period’s conditional variance and a long-term (unconditional) level, i.e.,

$$
\sigma^2_{t,t+1+i} = (\alpha + \beta)^i \sigma^2_{t,t+1} + \omega \frac{1 - (\alpha + \beta)^i}{1 - \alpha - \beta},
$$

(14)

where $\alpha$, $\beta$, and $\omega$ are the GARCH parameters. Now, assuming that the GARCH parameters remain constant, our variance convergence speed parameter $\phi$ is the sum of $\alpha$ and $\beta$. The unconditional variance can be estimated as $\omega/(1 - \alpha - \beta)$.

### 2.3. Econometric issues

To estimate the conditional variance series, we utilize quasi-maximum likelihood approach. In practice we utilize the likelihood function for the selected distribution given the mean equation (a constant) and the variance process. When utilizing conditional normal as the error distribution, we utilize heteroskedasticity-consistent standard errors using the methods described by Bollerslev and Wooldridge (1992). Both Marquard and Berndt-Hall-Hall-Hausman (BHHH) algorithms are used for the optimization. Conditional variance series is constructed using the estimated parameter values.

In addition, we construct a measure of economic expectations (economic cycles). To do this we estimate the probability of recession using a probit-model with a set of selected information (forecasting) variables. Our approach is similar to Kauppi and Saikkonen (2008); in other words, the full sample is utilized in the estimation instead of a rolling estimation which produces a slight forward looking bias (see their paper for the discussion). Dynamic and non-dynamic estimation approaches are used. The extracted recession probability estimates are used as one of the information variables.

The two-step estimation procedure to estimate the full model (12) and the baseline model (13) raises the question of whether there might be biases in our second-step estimator for the
lambda because the independent variable is subject to an errors-in-the-variables problem. Following earlier studies, we argue that the potential measurement error in the variance decreases due to the long sample period (cf. Shanken, 1992) and that, as a result, the lambda estimates are not biased systematically. For example, Hedegaard and Hodrick (2014) use a four-step procedure in a multivariate setup.

When estimating the linear regression model for equations (12) and (13), we also test for autocorrelation in the residuals of our model. Because there are indications of first-order autocorrelation, we use the Newey-West (1987) adjustment for autocorrelation and heteroskedasticity with automatically selected fixed bandwidth and Bartlett kernel when calculating the standard errors for the parameters using the OLS. Thus, all reported $t$-values for equations (12) and (13) are calculated with the adjustment.

3. DATA

3.1. Main variables

We utilize monthly excess equity market returns for the US market to test the models. The sample period is from January 1928 to December 2013, i.e., 1,032 months of data. Month-end CRSP value-weighted total returns are used as a proxy for the market returns. The risk-free rate of return is calculated for any month $t+1$ as the one-month holding period return on US Treasury bills closest to one month at the end of month $t$. These data are also from the CRSP. The excess return is obtained as the difference between the market return and the risk-free rate. Continuously compounded returns in decimal format are used throughout this study.

For the full model, we also need a measure for the change in the risk-free interest rate level. Here, we proxy the risk-free interest rate level with the long-term US government bond yield taken from the Ibbotson SBBI (2014). In addition, we need a measure for the change in the conditional dividend growth rates. For this end, we assume that the past annual growth rate can be used as a proxy. To calculate the realized dividend growth rate, we first calculate monthly time series for the dividends paid (USD) in the past twelve months. This is done by multiplying the CRSP price index a year ago with the difference between the total return and price index percentage returns in the twelve-month period. In the second step, we calculate monthly series for the logarithmic annual change in the paid dividends.

---

4 This is based on the idea that we can calculate a proxy for the dividends by solving $D_t$ from the following equation $R_t - R^x_t = (\frac{P_t + D_t}{P_{t-12}} - 1) - (\frac{P_t}{P_{t-12}} - 1)$, where on the left hand side we have the difference between realized percentage annual returns for the total return and the price index.
and use it as a proxy for the expected future growth rate of dividends. Finally, we take the first difference of these rates and insert them into our model.\textsuperscript{5}

3.2. \textit{Conditioning information variables for time-varying relative risk aversion}

Several different variables have been used to model the potential time-variation in the relative risk aversion or the price of market risk. Studies on an individual level have linked several characteristics to risk aversion. For example, Poterba (2001) studies the claims that risk tolerance declines as household ages. Here, however, the main interest is on aggregate effect on the expected equity premium. The effect has to be exogenous to the market uncertainty, i.e., go over and above the effect caused by the time-varying market risk level. By the very nature of the relative risk aversion, we also argue that the potential and hopefully visible time-variation in the risk-aversion is rather slow-moving compared to the variance. Hence, we expect the risk aversion to behave fairly smoothly.

A number of different variables have been suggested and their effect on the (aggregate) risk aversion have been studied in the literature. Yee (2006) studied link between savings rate, earnings quality and the risk aversion. Other studies have studied the effect of e.g. war (Gilson et al., 2015), financial crisis (Guiso et al., 2013), liquidity (Gibson and Mougeot, 2002), catastrophe risk (Barro, 2006), uncertainty about government policy (Pastor and Veronesi, 2012), and aggregate level of education (Outreville, 2015). De Santis and Gérard (1998) utilize financial variables – such as the dividend yield in excess of the short-term interest rate, the change in the U.S. term premium, the change in the one-month Eurodollar deposit rate, and the U.S. default premium – to model the price of market risk. From the very definition of the relative risk aversion, one can also argue that for some forms of utility functions, we should also see a relationship between relative risk aversion and the overall wealth level.

Here we concentrate on economic conditions and its effect on the risk aversion. Cohn et al. (2015) find that finance professionals are willing to take considerably less risk in a financial bust as compared to a boom, i.e., that risk aversion seems to be countercyclical. As a result, we test a hypothesis that bad economic conditions increase risk aversion and vice versa. However, it is natural to assume that for investors the current economic condition is not as important as the expectations regarding the development of the economy in the future.

\textsuperscript{5}In other words, the difference in conditional growth rates, $g_{t+1,t+2} - g_{t,t+2}$, is proxied with the first difference in realized annual dividend growth rates, $(d_{t+1} - d_{t+1-12}) - (d_{t} - d_{t-12})$. 

10
Now, it is an empirical and practical question how to model economic expectations. We follow three different approaches. The first approach allows the risk aversion to vary directly with the selected conditioning variables that can be argued to be related to the economic conditions. The second approach uses the actual knowledge of the recession to measure its effect on the risk aversion. The third approach uses forecasting variables to generate a prediction of the economic conditions, which, in turn, is used to model the risk aversion. In the first approach, the estimated system is open and the focus is on the economic activity in general terms, whereas the latter approaches focus directly on modelling the economic situation or investors’ expectations regarding it.

In the first approach the price of risk (risk aversion) is modelled linear on variables that have been found to predict the economic activity. To keep the system estimable, only a few variables are chosen. The first variable \(dIP\) is continuously compounded monthly change in the US industrial production. The second variable \(BaaAaa\) is the U.S. default premium measured as the difference between Baa and Aaa rated bond yields. The third variable is the long-term term premium \(LTS\), a measure of the long end yield curve, namely the difference between 10 year government bond yield and 3 month T-bill rate. Our fourth variable is a measure of the short-term term structure \(STS\) measured as the difference between 3 and 1 month T-bill rates. All series are downloaded from FRED with the exception of 10 year government bond yield which is taken from Ibbotson (2014).

We expect an increase in the \(dIP\) to lower the risk aversion since higher production signals improvement in the economy. Similarly an increase in the \(LTS\) should be inversely related to the risk aversion as positive term structure has been found to be positively related to future economic activity (c.f., Estrella and Mishkin, 1998). The converse is true for the \(BaaAaa\) variable as an increase in the corporate credit risk premium typically forecasts economic downturn. The expected effect of the \(STS\) variables on the risk aversion is more ambiguous to pinpoint as the values are partly affected by the decisions made by the Fed and partly by the short-term market anticipations.

Our second approach uses an ex post recession indicator \(DREC\) for the US where a value of one indicates a recession period and a value of zero is an expansion period. The indicator is based on the announcements made by the Business Cycle Dating Committee of the National Bureau of Economic Research. Due to the way the indicator is calculated, the values are announced approximately twelve months after the fact. As a result, the potential

---

6 The NBER provides monthly values that are based on average of daily data for the whole sample period as well as daily values from 1986 forward. We have augmented the former time series with the month-end values from the latter time series in the estimation.

7 The NBER does not define a recession in terms of two consecutive quarters of decline in real GDP. Rather, a recession is defined as a significant decline in economic activity spread across the economy, lasting
effect suffers from a look-ahead bias, but we do not consider it a major issue as investors’
can recognize a recession in a timely manner even if it is officially announced afterwards.

A recession indicator is only a crude measure of the economic cycles and it may not
indicate the time-varying expectations of the investors. Hence, we also want to create another
variable that might capture investors’ expectations in a better fashion. Our third approach
does exactly this as we use a probit model to generate a prediction of the economic condition.
For this purpose, we use the following dynamic prediction model for the recessions:

\[ y_t = \delta y_{t-1} + x_{t-1}' \beta + \epsilon_t, \]

where \( y_t \) is the (lagged) value of \( D_{REC} \), and \( x_{t-1}' \) is a vector of conditioning prediction
variables. A non-dynamic version excludes the first term from the right-hand side. If we
define \( \Phi(\pi_t) = \Phi(\hat{\delta} y_{t-1} + x_{t-1}' \hat{\beta}) \), where \( \Phi(.) \) is the cumulative normal function, then \( \Phi(\pi_t) \)
is the cumulative probability of recession at time \( t \).

In practice, we utilize the same four variables as before as prediction variables in the
empirical analysis. The probit analysis allows us to compute the fitted probabilities for
recession as follows: \( p_t = 1 - \Phi(-\hat{\delta} y_{t-1} - x_{t-1}' \hat{\beta}) \). This variable \( (PROB_{REC}) \) is used
to test whether risk aversion increases when the probability of recession increases. We do
the analysis in-sample similar to Kauppi and Saikkonen (2008). First we use a dynamic
version of the model in which investors’ know the prevailing state. Then we utilize a non-
dynamic prediction model in which investors do not recognize the recession after the fact.
In both cases, we assume as Kauppi and Saikkonen, that even though data on recessions are
announced with delay, we assume that investors have some understanding of the prevailing
recession in real time.

### 3.3. Descriptive analysis

Table 1 provides descriptive statistics for the asset returns (Panel A) and the conditional
variables (Panel B) in this study. The sample period is from January 1928 to December
2013. The arithmetic mean monthly risk premium is 0.477 per cent per month (or 5.72% per
annum), with a volatility of 5.44 per cent per month (18.84% p.a.). The average dividend
growth rate is 4.49%. Government bond yields have been, on average, 5.09%.

Almost all of the time series are non-normal according to the Jarque-Bera (1987) test for
normality. The monthly risk premia are negatively skewed and show kurtosis. As expected,
the monthly risk premium series show fairly low, albeit significant, positive first-order auto-
more than a few months, normally visible in real GDP, real income, employment, industrial production, and
wholesale-retail sales. In practice we have used the indicator series DREC from the FRED.
correlation. The dividend growth rate shows high autocorrelation (0.895) as expected due to overlapping dividend observations used to calculate the growth rate, as do the government bond yield series (0.980).

Before using the information variables in the estimation, we test whether they are stationary using the augmented Dickey-Fuller test as there are some evidence that non-stationary information variables might bias the results in tests of lambda pricing model (see, e.g., Antell and Vaihekoski, 2012). To test this we run an augmented Dickey-Fuller test with a constant and four lags in the test equation. MacKinnon (1996) one-sided p-values are reported in Table 1. The results show all of the series clearly reject the null hypothesis of unit root (non-stationarity).

Finally, to study whether the selected information variables show multicollinearity and whether some of them might be redundant in the estimation, we calculate their cross-correlations (available upon request). They are all quite low – the highest cross-correlation is between $BaaAaa$ and $LTS$, 0.32, but can still be considered reasonable for the analysis at hand.

3.4. Probit estimation

Table 2 presents estimation results for the probit model (equation (15)). All prediction variables are lagged by one month. This means that when we model recession indicator variable for month $t + 1$, we utilize information that is available at the end of month $t$, which – in some cases – include the difference of values at the end of month $t$ and $t - 1$. We report the parameter estimates first for the non-dynamic model which does not utilize the past information on the recession and then for the dynamic model. In addition to the parameter estimates, an adjusted McFadden R-squared is reported for the models.

The results show, as expected, that the dynamic version is much better at capturing the recession periods, as the knowledge of past state of the economy is a natural predictor for the next period. Higher credit yield spread and decreasing industrial production are positively related to the recession probability with the non-dynamic version of the model. Steep long-end yield curve is negatively related to future recessions in both versions of the model. Figure 1 shows the development of the fitted probabilities from both estimations.

Although we utilize the probabilities from both dynamic and non-dynamic version of the Probit model, our main interest is with the latter version. Dynamic model clearly uses information available after the fact and thus it is unlikely to give false signals which in reality are natural. In other words, it is more likely that investors might anticipate a recession which did not materialize, yet affecting investor behavior.
4. EMPIRICAL RESULTS

4.1. Preliminary analysis

A simple test to see if the price of market risk is time-varying is to estimate the lambda parameter using a rolling sample window. We do this first with a sixty year estimation window (720 observations) utilizing both the traditional and new estimation approaches. For simplicity, we use the GARCH(1,1)-M model with normally distributed errors when we estimate equation (2). A similar variance process is used when the lambda is estimated using the new estimation approach of Antell and Vaihekoski (2015). The time-series evolution of the lambda can be seen in Figure 2.

It is clear from Figure 2 that lambda is time-varying regardless of the estimation approach. In addition, the Second World War seems to have an expected impact on the price of risk reflecting investors’ higher risk aversion (c.f., Gilson et al., 2015). The rolling estimation approach, however, can only be applied after the fact and as such it does not utilize the latest information to the fullest. Hence, in order to get more detailed view on the behavior of the price of risk and how the economic cycles affect it, we proceed to estimating the conditional model.

4.2. Time-varying price of risk and the traditional estimation approach

We begin our analysis by studying the price of market risk using the traditional estimation approach. Following earlier research we model conditional variance with the generalized autoregressive conditional heteroskedasticity approach. To induce the time-variability into lambda, we combine equations (2) and (3) and utilize the variance estimate from the GARCH model. In effect, we estimate a GARCH in mean model with time-varying coefficient for the explanatory variable (variance) in the mean equation. As the GARCH-type estimation is always sensitive to the starting values, we take them from the model in question estimated with a constant price of risk parameter (lambda). Initial values for the conditioning variables are set to zero.

We estimate the model first using the standard GARCH(1,1) model for the variance and then using the GJR-GARCH(1,1) model by Glosten, Jaganathan, and Runkle (1993) for the variance. It is better at capturing asymmetry in the variance process (cf., e.g., Bekaert and Wu, 2000; Cappiello et al., 2006). The reported results are from the latter model. For simplicity, variance estimations are done initially under the assumption of gaussian errors for models I-IV, but for the model V, we have introduced $t$-distributed errors.

Table 3 presents the results for different model specifications, all estimated using the
quasi-maximum likelihood. Conditioning variables have been demeaned in the estimation with the exception of the recession dummy. The results for the variance process’ parameters are not reported, but in all cases they are significant and as such the conditional volatility is asymmetric with a positive response to negative shocks. We also report an adjusted McFadden $R^2$-squared for the models.

Model I provides the starting values for the other models as well as an estimate of the constant lambda. The estimate for constant lambda (0.224) is not significant ($t$-value 0.255). The explanatory power of the model is also low. Model II allows lambda to vary freely with the selected four information variables. As the information variables have been demeaned, we can interpret the constant as an estimate for the long-term unconditional lambda. The results are disappointing – none of the conditioning variables is significant.

To test the main hypothesis in this paper, we proceed to test model III where the price of risk varies with the recession indicator, $D_{REC}$. Lambda is still not significant and, if anything, the relationship seems to be opposite to economic theory – recession seems to induce lower risk aversion, although the relationship is not significant. This, of course, could be due to the fact that the recession variable is only known after the fact and it may not really capture investors’ expectations regarding economic conditions. Thus we proceed to estimate model IV with the time-varying probability, $Pr_{REC}$, for the recession. It is estimated separately using recession probabilities from a dynamic (model IV) and a non-dynamic (model Va) probit models.

However, again the results do not provide any support for time-variability of the lambda. This result is clearly counterintuitive and against our hypothesis that the price of market risk (relative risk aversion) should be higher when investors are faced with bad economic conditions or they anticipate them in the future. Overall, the explanatory power of the traditional model is low, with an adjusted $R^2$ typically very close to zero.

There are several reasons for the estimation not finding a significant relationship between expected returns and variance. One potential explanation could be that the model used is too complex for the estimation. As a robustness check, we estimate the models without the asymmetry in the variance process. The results are basically the same, although lambda is somewhat higher but still insignificant. Another explanation is that the conditional return is not normally distributed. Therefore, we re-estimate models III, IV and Va assuming a conditional $t$-distribution instead of the normal distribution. The results (provided for model Va in Table 3 as model Vb) do not provide support for the traditional model and time-varying lambda, although the degrees of freedom for the $t$-distribution is estimated to be 7.012 with a $t$-value of 6.314, meaning that the tails of the distribution are fatter than is commensurate with the normal distribution. However, the explanatory power of the
model does not materially increase, and the price of risk parameter remains insignificant. In fact, the lambdas are even lower than before. It does not seem unfair to conclude that the traditional approach, when estimated with the commonly used GARCH-in-mean approach, does not seem able to find a statistically significant (positive) relationship between variance and return even when one allows for time-varying price of market risk.

4.3. New approach and the baseline model

Next, we turn to the new, reverse estimation approach. We first estimate our baseline model utilizing the same GARCH processes as with the traditional approach and again with conditional normality and \( t \)-distribution assumed. The estimation is conducted in two steps. First, we estimate the parameters for the (GJR-)GARCH process with a constant in the mean equation. We then use the results to calculate the variance persistence parameter \( \phi \), unconditional variance, and the sigma-multipliers as indicated by equations (??) and (??). Note that when we are utilizing the GJR-GARCH specification, the variance persistence parameter is the sum of \( \alpha \), \( \beta \), and half the asymmetry parameter, \( \gamma \). The unconditional variance can be stated as \( \omega/(1 - \alpha - \beta - \gamma/2) \). In the second stage, we run a linear regression model according to equation (15) and again with the information variables in excess of their means. The results are reported in Table 4. The reported \( t \)-values use calculated using the Newey-West (1987) adjustment on standard errors for autocorrelation and heteroscedasticity.

In line with the traditional approach, our results show that the variance process parameters (not reported) are significant in all cases. However, in contrast to the results shown in Table 3, the lambda estimates are significant in all of the estimated cases. The explanatory power of the models are also considerably higher than they were for the corresponding traditional models. More importantly, though, lambda is significant and time-varying. As before, we begin the analysis by allowing our four conditioning variables to have an impact on lambda without any restrictions (model II). The results show that higher credit risk spreads (\( t \)-value 2.083) and steeper short end of the yield curve (\( t \)-value 2.361) lead to higher risk aversion. In addition, we find a negative relationship, although statistically weaker (\( t \)-value 1.834), between the change in the industrial production and the risk aversion. In all cases, the sign of the effect is as implied by the economic theory.

When we use an explicit measure for the recession, \( D_{REC} \), we can clearly see that the risk aversion is significantly and positively related to a recession (model III). The impact is not major (0.029, \( t \)-value 2.179) in economic terms, but it confirms our theoretical expectation. When we use a richer variable to model expectations on economic conditions, the results show that there is clearly a positive relationship between the probability of a recession
and risk aversion (Models IV and Va with dynamic and non-dynamic estimates for the recession, respectively). Finally, we re-estimate Model Vb with the variance process estimated under the assumption of $t$-distribution. The results basically confirm the earlier results, although interestingly the explanatory power of the model increases clearly indicating that the normal distribution is too restrictive for the stock returns.

4.4. **Full model**

Finally, we test our full model as given by equation (12). We estimate the model as before using the OLS with conditional variance estimates from the first-pass estimation. The conditional variance process is based on the GJR-GARCH model, first under normal distribution and then assuming $t$-distribution (Model Vb). The results are reported in Table 5.

The results show that the long-term (unconditional) lambda is statistically significant as before. Our second and third explanatory variables, changes in the expected dividend growth rate and the change in the risk free rate, are also found statistically significant and positive as implied by our model. The explanatory power of the full model can be considered to be high (adjusted $R^2$s range from 51.3% to 56.3%), and in fact, it is always clearly higher than for the corresponding baseline model.

High explanatory power and high $t$-values for the $\Delta g$ variable raise the question whether we are explaining our realized returns with realized contemporary dividends creating an endogeneity/simultaneity problem. This is, however, not the case, even though our proxy for the change in the dividend growth rates does utilize contemporary dividends when it is calculated. Using the growth rate of a variable to explain the level of the variable, even if contemporary, is not a statistical issue if both variables are stationary and if there is no loop of causality between the explanatory and dependent variable. Obviously higher returns do not cause higher dividends.

Although the recession dummy is not statistically significantly when the full model is estimated, the results give strong support for a higher risk aversion when investors’ expectations regarding the economic condition worsen when measured with the fitted probabilities from the Probit estimations (models IV, Va, and Vb).

4.5. **Additional considerations**

Analysing the economic significance of the time-variation in the price of risk shows that it is rather limited. For example, the parameter values for Model Vb in Table 5 suggest that lambda varies from 1.251 to 1.296 (range of 3.53%) depending on the value for $Pr.REC_t$. This is surprisingly small compared to the variation shown in Figure 2. In practice, the sug-
gested variation implies that at the average market conditions with the conditional variance matching its time series average (here 0.003046 per month), the expected risk premium varies from 0.38% to 0.39% per month (4.573% – 4.737% p.a.). Small range of variation raises the question whether the conditional variance and change in the conditional dividend growth rates variable are correlated with the recession probability variable. In other words, a higher likelihood of recession could be reflected also as lower dividend growth expectations. Similarly, conditional variance might also reflect higher probability of a recession which leaves no room for the time-variation in the lambda.

To study this possibility, we calculate the correlation coefficients between these three variables. The results show that the conditional variance is highly correlated with the recession probability (0.56), but this is not the case with the change in the dividend growth rates variable – correlation is only 0.01. This leaves us to believe that the information in the selected $Pr_{REC_t}$ variable is already reflected to some degree in the market as higher level of market uncertainty. As a result, investors do not seem to change their risk aversion as much as a priori expected.

It is also interesting to study the model using totally different exogenous variables. There is plenty of evidence that the person’s risk aversion could be affected by one’s physical safety concerns. At the aggregate level, a war would be a natural cause for investors to increase their risk aversion. To test this hypothesis we create an indicator for the Second World War (i.e. December 1939 – August 1945) and re-estimate our model V in Table 5 using non-dynamic probability of a recession and the war indicator as two conditioning variables.

The results show that the long-term unconditional lambda is still highly significant (1.248, $t$-value 6.976) and higher when the likelihood of recession increases (0.047, $t$-value 2.00). Interestingly we find the risk aversion to be clearly higher during the war time (0.091, $t$-value 6.00). The increase is economically meaningful, driving the risk aversion up by more than 7.3% resulting for 0.255% higher risk premium per annum if measured using the average conditional variance during the same period (somewhat surprisingly low, only 0.002906 per month).

5. SUMMARY AND CONCLUSIONS

In this paper, we study the relationship between the equity market risk premium and variance using the asset pricing model of Merton (1973, 1980). Our main interest is to study whether the price of market risk, or under certain assumptions, relative risk aversion, is time-varying and whether it behaves in a counter-cyclical fashion. For the empirical estimation of the model, we use two different approaches. The first approach is based on the traditional
estimation approach which uses realized returns are a proxy for the expected returns. The second approach uses the testing approach developed in Antell and Vaihekoski (2015). It is based on the idea of studying realized returns conditional on the tested asset pricing model being true.

Tests are conducted using data for the equity premium on the US stock market from 1928 to 2013. In the empirical tests, we utilize different ways to estimate economic expectations that are related to cyclicality in the economy. For robustness, we also utilize different specifications for the variance process as well as for the return distribution. The traditional estimation approach cannot find a positive risk-return relationship during the sample period. Furthermore, it does not support the hypothesis of time-varying risk aversion nor its counter-cyclical behavior.

The results from the new estimation approach, on the other hand, give support for the positive risk-return relationship as well as for the time-varying and counter-cyclical risk aversion. However, the results indicate that the variation in the risk aversion could be smaller than thought a priori, at least when the outlook for an economic recession is used as the driving variable for the risk aversion. On the other hand, the results provide support for higher risk aversion during the Second World War, although the aggregate effect is likely to be smaller than previously anticipated.
REFERENCES


Fig. 1. Fitted probabilities for a recession from non-dynamic (1a) and dynamic (1b) probit models estimated with four conditioning information variables. The dynamic model uses previous realizations in the model. Monthly US data from January 1928 to December 2013 are used in the estimation. Recession periods as indicated by NBER are marked with grey area.
Fig. 2. Price of risk estimates from the traditional as well as the new estimation approach using a sixty-year rolling window. In both cases, variance has been modelled as a GARCH(1,1) process with normal distribution. Dates on the x-axis indicate the beginning of the sample period.
Table 1: **Descriptive statistics.** Descriptive statistics for the main variables of interest (Panel A) and information variables (Panel B). All variables are for the USA. The main pricing variables are excess market returns, dividend growth rate as well as long-term government bond yield. Monthly data are used from January 1928 to December 2013 (1,032 observations). All returns, yields, and growth rates are continuously compounded, in percentage form. Dividend growth rate per annum is the continuously compounded growth of the dividends paid during the past twelve months compared to dividends paid a year ago. Long-term government bond yield per annum is taken from the SBBI book. The conditional information variables are: difference between the continuously compounded yield on Baa and Aaa rated corporate bonds (BaaAaa), 10 year government bond yield minus 3 month T-Bill rate (LTS), a measure of short-term yield curve (3 month minus 1 month TB-rate, STS), and monthly change in industrial production (dIP). Normality refers to the p-value for the Jarque-Bera (1987) test for normal distribution. ADF is MacKinnon (1996) one-sided p-value for the Augmented Dickey-Fuller test statistic for non-stationarity. Autocorrelation coefficients significantly (5%) different from zero are marked with an asterisk.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. dev.</th>
<th>Skewness</th>
<th>Excess Kurtosis</th>
<th>J-B (p-val)</th>
<th>Autocorrelation</th>
<th>ADF</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Main variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equity premium $R_m - R_f$</td>
<td>0.477</td>
<td>5.44</td>
<td>-0.53</td>
<td>6.51</td>
<td>&lt; 0.001</td>
<td>0.110*</td>
<td>-0.011</td>
</tr>
<tr>
<td>Market return squared $R_m^2$</td>
<td>30.055</td>
<td>85.08</td>
<td>8.29</td>
<td>84.70</td>
<td>&lt; 0.001</td>
<td>0.230*</td>
<td>0.169*</td>
</tr>
<tr>
<td>Dividend growth rate p.a.</td>
<td>4.488</td>
<td>16.33</td>
<td>-0.51</td>
<td>1.32</td>
<td>&lt; 0.001</td>
<td>0.895*</td>
<td>0.768*</td>
</tr>
<tr>
<td>Long-term government bond yield p.a.</td>
<td>5.092</td>
<td>2.64</td>
<td>0.96</td>
<td>0.30</td>
<td>&lt; 0.001</td>
<td>0.996*</td>
<td>0.991*</td>
</tr>
<tr>
<td><strong>Panel B: Information variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baa-Aaa (BaaAaa)</td>
<td>1.067</td>
<td>0.66</td>
<td>2.42</td>
<td>8.24</td>
<td>&lt; 0.001</td>
<td>0.975*</td>
<td>0.939*</td>
</tr>
<tr>
<td>Term structure, long (LTS)</td>
<td>1.603</td>
<td>1.28</td>
<td>-0.24</td>
<td>-0.10</td>
<td>0.005</td>
<td>0.961*</td>
<td>0.920*</td>
</tr>
<tr>
<td>Term structure, short (STS)</td>
<td>0.205</td>
<td>30.19</td>
<td>2.56</td>
<td>11.54</td>
<td>&lt; 0.001</td>
<td>0.518*</td>
<td>0.424*</td>
</tr>
<tr>
<td>Industrial production growth (dIP)</td>
<td>0.268</td>
<td>1.81</td>
<td>0.35</td>
<td>14.61</td>
<td>&lt; 0.001</td>
<td>0.533*</td>
<td>0.245*</td>
</tr>
</tbody>
</table>
Table 2: Recession model estimation results. Quasi-maximum likelihood estimates for the probit model under different model specifications are provided. In-sample estimation is conducted using the full sample from January 1928 to December 2013 (1,032 observations). In the estimation, the recession indicator variable ($D_{REC}$) is regressed against four lagged information variable. The information variables are: difference between the continuously compounded yield on Baa and Aaa rated corporate bonds ($BaaAaa$), a measure of short end of the yield curve (3 month minus 1 month TB-rate, $STS$) as well as the long end of the yield curve (10 year government bond yield minus 3 month T-Bill rate, $LTS$), and monthly change in industrial production ($dIP$). Non-dynamic and dynamic versions of the model are estimated. In the latter model, the lagged realized value for the dependent variable appears as an additional explanatory variable. McFadden $R^2$ are reported for the model. $t$-values are provided below the parameter values in parentheses. Coefficients significantly (10%, 5% or 1%) different from zero are marked with one, two, or three asterisks, respectively.

<table>
<thead>
<tr>
<th></th>
<th>Constant</th>
<th>$y_{t-1}$</th>
<th>$BaaAaa_{t-1}$</th>
<th>$LTS_{t-1}$</th>
<th>$STS_{t-1}$</th>
<th>$dIP_{t-1}$</th>
<th>McFadden $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-dynamic model</td>
<td>-1.478***</td>
<td>69.075***</td>
<td>-18.041***</td>
<td>0.063</td>
<td>-35.157***</td>
<td>0.258</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-12.807)</td>
<td>(-7.859)</td>
<td>(-4.087)</td>
<td>(0.397)</td>
<td>(-9.850)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dynamic model</td>
<td>-1.909***</td>
<td>3.594***</td>
<td>19.417</td>
<td>-29.994***</td>
<td>-0.247</td>
<td>-5.326</td>
<td>0.769</td>
</tr>
<tr>
<td></td>
<td>(-10.952)</td>
<td>(16.031)</td>
<td>(1.441)</td>
<td>(-3.801)</td>
<td>(-0.980)</td>
<td>(-0.939)</td>
<td></td>
</tr>
</tbody>
</table>
Table 3: **Baseline model and the traditional estimation approach.** Quasi-maximum likelihood estimates for the unconditional (Model I) and conditional price of market risk are reported using the traditional estimation approach under different model specifications. The conditional price of market risk (relative risk aversion) is modelled linear on the variables indicated in below. Variable $Pr.REC_t$ is the in-sample probability of a recession from dynamic (model IVa) and non-dynamic (models Va and Vb) Probit models, respectively. GJR-GARCH with normal distribution is used to estimate the variance process except for the model Vb where $t$-distribution has been used. Excess US continuously compounded returns (CRSP total return index) from January 1928 to December 2013 (1,032 observations) are used in the estimation. Conditioning variables have been demeaned (except the recession indicator). Adjusted $R^2$ is calculated using the likelihood values for the full and the restricted model with adjustment for the number of parameters. Coefficients significantly (10%, 5% or 1%) different from zero are marked with one, two, or three asterisks, respectively.

<table>
<thead>
<tr>
<th></th>
<th>Model I</th>
<th>Model II</th>
<th>Model III</th>
<th>Model IV</th>
<th>Model Va</th>
<th>Model Vb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant alpha</td>
<td>0.006**</td>
<td>0.006**</td>
<td>0.006</td>
<td>0.006**</td>
<td>0.005**</td>
<td>0.008***</td>
</tr>
<tr>
<td></td>
<td>(2.550)</td>
<td>(2.180)</td>
<td>(1.188)</td>
<td>(2.552)</td>
<td>(1.998)</td>
<td>(3.192)</td>
</tr>
<tr>
<td>$\lambda_0$</td>
<td>0.224</td>
<td>0.224</td>
<td>0.224</td>
<td>0.227</td>
<td>0.287</td>
<td>0.081</td>
</tr>
<tr>
<td></td>
<td>(0.255)</td>
<td>(0.187)</td>
<td>(0.165)</td>
<td>(0.250)</td>
<td>(0.226)</td>
<td>(0.069)</td>
</tr>
<tr>
<td>$BaaAaa_{t-1}$</td>
<td>-0.002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-0.001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$LTS_{t-1}$</td>
<td>0.023</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$STS_{t-1}$</td>
<td>1.830</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.622)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$dIP_{t-1}$</td>
<td>0.005</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.002)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D_{REC_t}$</td>
<td>-1.327</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-1.016)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Pr.REC_t$</td>
<td></td>
<td>-1.434</td>
<td>-0.194</td>
<td>0.096</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-0.731)</td>
<td>(-0.101)</td>
<td>(0.051)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adj. $R^2$</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
</tr>
</tbody>
</table>
Table 4: **Baseline model and the new estimation approach.** OLS estimates for the unconditional (Model I) and conditional price of market risk as are reported using the new estimation approach under different model specifications. The conditional price of market risk (relative risk aversion) is modelled linear on the variables indicated in below. Variable `$Pr_{REC}$` is the in-sample probability of a recession from dynamic (model IVa) and non-dynamic (models Va and Vb) Probit models, respectively. GJR-GARCH with normal distribution is used to estimate the variance process except for the model Vb where $t$-distribution has been used. Excess US continuously compounded returns (CRSP total return index) from January 1928 to December 2013 (1,032 observations) are used in the estimation. Conditioning variables have been demeaned (except the recession indicator). Newey-West (1987) standard errors that are robust to heteroskedasticity and autocorrelation are used in the estimation. $t$-values are provided below parameter values in parentheses. Coefficients significantly (10%, 5% or 1%) different from zero are marked with one, two, or three asterisks, respectively.

<table>
<thead>
<tr>
<th></th>
<th>Model I</th>
<th>Model II</th>
<th>Model III</th>
<th>Model IV</th>
<th>Model Va</th>
<th>Model Vb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant alpha</td>
<td>0.002</td>
<td>0.002*</td>
<td>0.002</td>
<td>0.002</td>
<td>0.002</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>(1.485)</td>
<td>(1.761)</td>
<td>(1.566)</td>
<td>(1.602)</td>
<td>(1.562)</td>
<td>(0.687)</td>
</tr>
<tr>
<td>$\lambda_0$</td>
<td>0.890***</td>
<td>0.790***</td>
<td>0.863***</td>
<td>0.861***</td>
<td>0.876***</td>
<td>1.791***</td>
</tr>
<tr>
<td></td>
<td>(8.389)</td>
<td>(5.587)</td>
<td>(7.962)</td>
<td>(7.686)</td>
<td>(7.942)</td>
<td>(7.810)</td>
</tr>
<tr>
<td>$BaaAaa_{t-1}$</td>
<td></td>
<td>4.265**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.083)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$LTS_{t-1}$</td>
<td>0.058</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.070)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$STS_{t-1}$</td>
<td>0.014**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.361)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$dIP_{t-1}$</td>
<td>−0.412*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(−1.834)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D_{REC_t}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.029**</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(2.179)</td>
<td></td>
</tr>
<tr>
<td>$Pr_{REC_t}$</td>
<td></td>
<td></td>
<td></td>
<td>0.055**</td>
<td>0.074**</td>
<td>0.074**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(2.040)</td>
<td>(2.251)</td>
<td>(2.050)</td>
</tr>
<tr>
<td>Adj. $R^2$</td>
<td>0.246</td>
<td>0.290</td>
<td>0.251</td>
<td>0.263</td>
<td>0.275</td>
<td>0.378</td>
</tr>
</tbody>
</table>


Table 5: **Full model.** The OLS estimates for the unconditional (Model I) and conditional price of market risk as reported using the new estimation approach under different model specifications. GJR-GARCH with normal distribution is used to estimate the variance process. The conditional price of market risk (relative risk aversion) is modelled linear on the variables indicated in below. Variable $PrREC$ is the in-sample probability of a recession from dynamic (model IV) and non-dynamic (models Va and Vb) Probit models, respectively. GJR-GARCH with normal distribution is used to estimate the variance process except for the model Vb where $t$-distribution has been used. Excess US continuously compounded returns (CRSP total return index) from January 1928 to December 2013 (1,032 observations) are used in the estimation. Conditioning variables have been demeaned (except the recession indicator). $\Delta g$ is the first difference in the expected dividend growth rate per annum. $\Delta rf$ is the first difference in the long-term US government bond yield. The Newey-West (1987) adjustment has been used to calculate the standard errors. $t$-values are provided in parentheses. Coefficients significantly (10%, 5% or 1%) different from zero are marked with one, two, or three asterisks, respectively.

<table>
<thead>
<tr>
<th></th>
<th>Model I</th>
<th>Model II</th>
<th>Model III</th>
<th>Model IV</th>
<th>Model Va</th>
<th>Model Vb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant alpha</td>
<td>0.003**</td>
<td>0.003**</td>
<td>0.003**</td>
<td>0.003**</td>
<td>0.003**</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>(2.148)</td>
<td>(2.137)</td>
<td>(2.181)</td>
<td>(2.141)</td>
<td>(2.194)</td>
<td>(1.512)</td>
</tr>
<tr>
<td>$\lambda_0$</td>
<td>0.604***</td>
<td>0.619***</td>
<td>0.595***</td>
<td>0.591***</td>
<td>0.605***</td>
<td>1.259***</td>
</tr>
<tr>
<td>$BaaAaa_{t-1}$</td>
<td>0.951</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.978)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$LTS_{t-1}$</td>
<td>0.070</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.131)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$STS_{t-1}$</td>
<td>0.007</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.281)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$dIP_{t-1}$</td>
<td>-0.180</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-1.227)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$DREC_t$</td>
<td></td>
<td>0.011</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.121)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$PrREC_t$</td>
<td></td>
<td></td>
<td></td>
<td>0.034***</td>
<td>0.041**</td>
<td>0.045*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(2.764)</td>
<td>(2.048)</td>
<td>(1.867)</td>
</tr>
<tr>
<td>$\Delta g$</td>
<td>0.389***</td>
<td>0.375***</td>
<td>0.387***</td>
<td>0.386***</td>
<td>0.375***</td>
<td>0.336***</td>
</tr>
<tr>
<td>$\Delta rf$</td>
<td>1.275***</td>
<td>1.268**</td>
<td>1.248**</td>
<td>1.235**</td>
<td>1.530***</td>
<td>1.423***</td>
</tr>
<tr>
<td></td>
<td>(2.554)</td>
<td>(2.580)</td>
<td>(2.507)</td>
<td>(2.482)</td>
<td>(3.189)</td>
<td>(2.977)</td>
</tr>
<tr>
<td>Adj. $R^2$</td>
<td>0.513</td>
<td>0.517</td>
<td>0.513</td>
<td>0.519</td>
<td>0.521</td>
<td>0.563</td>
</tr>
</tbody>
</table>