DISCUSSION PAPER SERIES

No. 9698

REGIME SWITCHES IN THE RISK-RETURN TRADE-OFF

Eric Ghysels, Pierre Guérin and Massimiliano Marcellino

FINANCIAL ECONOMICS and INTERNATIONAL MACROECONOMICS

ISSN 0265-8003

REGIME SWITCHES IN THE RISK-RETURN TRADE-OFF

Eric Ghysels, University of North Carolina and CEPR
Pierre Guérin, Bank of Canada
Massimiliano Marcellino, Bocconi University and CEPR

Discussion Paper No. 9698
October 2013

Centre for Economic Policy Research
77 Bastwick Street, London EC1V 3PZ, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 9698

October 2013

ABSTRACT

Regime Switches in the Risk-Return Trade-off*

This paper deals with the estimation of the risk-return trade-off. We use a MIDAS model for the conditional variance and allow for possible switches in the risk-return relation through a Markov-switching specification. We find strong evidence for regime changes in the risk-return relation. This finding is robust to a large range of specifications. In the first regime characterized by low ex-post returns and high volatility, the risk-return relation is reversed, whereas the intuitive positive risk-return trade-off holds in the second regime. The first regime is interpreted as a "flight-to-quality" regime.

JEL Classification: G10 and G12
Keywords: conditional variance, Markov-switching, MIDAS and risk-return trade-off

Eric Ghysels
Department of Economics
University of North Carolina
Gardner Hall CB 3305
Chapel Hill
NC 27599-3305
USA

Email: eghysels@unc.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=135789

Pierre Guérin
Bank of Canada
234 Wellington Street
Ottawa, ON K1A 0G9
CANADA

Email: pguerin@bank-banque-canada.ca

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=173118

Massimiliano Marcellino
Bocconi University
Department of Economics
Via Roentgen 1
20136 Milan
ITALY

Email:
massimiliano.marcellino@unibocconi.it

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=139608

*We would like to thank an anonymous referee for helpful comments on a previous version of this paper. Corresponding author: Pierre Guérin, Bank of Canada, 234 Wellington Street, Ottawa, ON K1A 0G9. The first author acknowledges support of a Marie Curie FP7-PEOPLE-2010-IIF grant. The views expressed in this paper are those of the authors. No responsibility for them should be attributed to the Bank of Canada.

Submitted 08 October 2013

1 Introduction

The ICAPM of Merton (1973) states that the expected excess return on the stock market is positively related to its conditional variance:

$$E_t(R_{t+1}) = \mu + \gamma V_t(R_{t+1}), \tag{1}$$

formalizing the intuition that a riskier investment should demand a higher expected return (relative to the risk-free return). However, in the empirical literature there is mixed evidence on whether the coefficient γ is indeed positive and statistically significant. Examples include Ghysels et al. (2005), Guo and Whitelaw (2006) and Ludvigson and Ng (2007), who all find a positive risk-return trade-off.[1] Instead, Glosten et al. (1993), using different GARCH specifications, find a negative relation between risk and return. Similarly, Brandt and Kang (2004) model both the expected returns and conditional variance as latent variables in a multivariate framework and find a negative trade-off.

Omitted variables could play a role to explain these conflicting results. For example, Scruggs (1998) and Guo and Whitelaw (2006)) emphasize the need to include additional variables in the risk-return relation to capture shifts in investment opportunities. Lettau and Ludvigson (2001) suggest using the consumption wealth ratio in the risk-return relation. Ludvigson and Ng (2007) instead include factors summarizing the information from a large set of predictors, and Lettau and Ludvigson (2010) find that a positive risk-return relation is uncovered using lagged mean and lagged volatility as additional predictors.

Another reason for the conflicting results reported in the literature is the way of modeling the conditional variance. Indeed, if one wants to estimate the risk-return trade-off over a long period of time, the conditional variance is not directly observable and must be filtered out from past returns. An attractive approach is the one developed by Ghysels et al. (2005). They introduce a new estimator for the conditional variance - the MIDAS (MIxed DAta Sampling) estimator - where the conditional variance depends on the lagged daily returns aggregated through a parametric weight function. The crucial difference with rolling window estimators of the conditional variance is that the weights on lagged returns are determined endogenously and in a parsimonious way with the MIDAS approach. In this paper, we follow the approach of Ghysels et al. (2005) and use a MIDAS estimator of the conditional variance since it is likely that the MIDAS estimator of the conditional variance can more fully describe the dynamics of market risk. It is also a convenient approach since it permits to easily model the dynamics of the risk-return trade-off at different frequencies.

In this paper, we also consider regime changes in the parameter γ entering before the conditional variance to reflect the possibility of a changing relationship between risk and return.[2] The relation between risk and return should not necessarily be linear. For example,

[1] French et al. (1987) find a strong negative relation between the unpredictable component of volatility and expected returns whereas expected risk premia are positively related to the predictable component of volatility.

[2] While writing the current version of this paper, we became aware of independent and simultaneously written work by Arago et al. (2013) using a similar approach with European data.

Backus and Gregory (1993) and Whitelaw (2000) show that non-linear models are consistent with a general equilibrium approach. Campbell and Cochrane (1999) underline the time-varying nature of risk premia. In particular, Whitelaw (2000) estimates a two-regime Markov-switching model with time-varying transition probabilities that include aggregate consumption as a driving variable for the transition probabilities to account for the changes in investment opportunities. He then finds a non-linear and time-varying relation between expected returns and volatility. Alternatively, Tauchen (2004) criticizes the reduced form nature of the models that estimate the risk-return trade-off. He develops a general equilibrium model where volatility is driven by a two factor structure with a risk premium that is decomposed between risk premia on consumption risk and volatility risk.

More recently, Rossi and Timmermann (2010) proposed new evidence on the risk-return relationship by claiming that the assumption of a linear coefficient entering before the conditional variance is likely to be too restrictive. They use an approach based on boosted regression trees and find evidence for a reversed risk-return relation in periods of high volatility, whereas the relation is positive in periods of low volatility. They also propose to model risk with a new measure, the realized covariance calculated as the product between the changes in the Aruoba et al. (2009) index of business conditions and the stock returns. We follow their approach and include this new measure of risk as a conditioning variable for estimating the risk-return trade-off.

We estimate regime switching risk-return relations using 1-week, 2-week, monthly and quarterly returns ranging from February 1929 to December 2010. Our empirical results can be summarized as follows:

- There is strong evidence for regime changes in the risk-return relation as supported by the test for Markov-switching parameters recently introduced by Carrasco et al. (2013).

- In the first regime characterized by low ex-post returns and high volatility, the risk-return relation is negative, whereas the risk-return relation is positive in the second regime. This is consistent across all frequencies we consider and a wide range of specifications (the inclusion of additional predictors, the use of time-varying transition probabilities, the use of Student-t rather than normal innovations and the use of an Asymmetric MIDAS estimator of the conditional variance).

- The first regime can be interpreted as a "flight-to-quality" regime. This evidence corroborates the findings in Ghysels et al. (2013) who document that the Merton model holds over samples that exclude financial crises, in particular the Great Depression and/or the subprime mortgage financial crisis and the resulting Great Recession. They also report that a simple flight-to-quality indicator, based on the ex post extreme tail events, separates the traditional risk-return relationship from financial crises which amount to fundamental changes in that relationship. In this paper we show that a Markov switching regime model is indeed recovering a similar pattern.

The paper is structured as follows. Section 2 presents the model we use for estimating the risk-return relation. Section 3 details the main results of the paper and a comparison of the estimated conditional variances with GARCH specifications. Section 4 provides a sensitivity analysis across a wide range of models as well as an out-of-sample forecasting exercise. Section 5 concludes.

2 Estimation of the risk-return trade-off with a Markov-switching MIDAS model

If returns are normally distributed, the MIDAS estimation of the risk-return trade-off is such that:

$$R_{t+1} \sim N(\mu + \gamma V_t^{MIDAS}, V_t^{MIDAS}) \tag{2}$$

However, the assumption of a constant parameter γ can be too restrictive and miss changes in investment opportunities due to e.g. changes in the level of market volatility. We therefore propose to model regime changes in the parameter γ through a Markov-switching process that can account for time instability in the risk-return relation. We also consider regime changes in the intercept μ to account for time variation in the mean of the returns. Equation (5) then becomes:

$$R_{t+1} \sim N(\mu(S_{t+1}) + \gamma(S_{t+1})V_t^{MIDAS}, V_t^{MIDAS}) \tag{3}$$

where S_{t+1} is an M-state Markov chain defined by the following constant transition probabilities:

$$p_{ij} = Pr(S_{t+1} = j | S_t = i) \tag{4}$$

$$\sum_{j=1}^{M} p_{ij} = 1 \forall i, j \epsilon \{1, ..., M\} \tag{5}$$

We use a MIDAS estimator for the conditional variance of the stock market since it has already proven to be a useful specification for the estimation of the risk-return trade-off (see e.g., Ghysels et al. (2005)). The MIDAS estimator of the conditional variance is based on the lagged daily returns, which are weighted via a parametric weight function. Two popular choices in the literature are the beta polynomial and the exponential Almon lag weight functions:

$$w(j;\theta) = \frac{(\frac{d}{D})^{\kappa_1} - (1 - \frac{d}{D})^{\kappa_2 - 1}}{\sum_{j=0}^{K} (\frac{j}{D})^{\kappa_1} - (1 - \frac{j}{D})^{\kappa_2 - 1}} \tag{6}$$

$$w(j;\theta) = \frac{exp(\kappa_1 j + \kappa_2 j^2)}{\sum_{j=0}^{K} exp(\kappa_1 j + \kappa_2 j^2)} \tag{7}$$

The above weight functions can take a large variety of shapes depending on the value of the two parameters κ_1 and κ_2. In this paper, we use daily absolute returns rather than squared returns as the use of absolute returns makes the estimated conditional variance less sensitive to outliers. This is relevant as we include periods of high volatility in our estimation sample (1929-2010). In addition, Ghysels et al. (2006) and Forsberg and Ghysels (2007) find that realized power (i.e., the daily sum of the 5-min absolute returns) is the best predictor of future volatility. The MIDAS estimator of the conditional variance is then given by:

$$V_t^{MIDAS} = N \sum_{d=0}^{D} w_j |r_{t-d}| \tag{8}$$

where N is a constant that corresponds to the number of traded days at the frequency of the expected returns to insure that expected returns and conditional variance have the same scale.[3]

The model is estimated by maximum likelihood via the EM algorithm since the EM algorithm performs well for estimating non-linear models (see e.g., Hamilton (1990) and Guérin and Marcellino (2013)).

Several papers estimated Markov-switching models for assessing the risk-return relation. Whitelaw (2000) estimates a Markov-switching model with time-varying transition probabilities with monthly aggregate consumption data and finds a non-linear and time-varying risk-return relation. Mayfield (2004) introduces regime switching in a general equilibrium model where market risk is characterized by periods of high and low volatility, which evolves according to a Markov-switching process. He finds evidence for a shift in the volatility process in 1940 and uncovers a positive risk-return trade-off. Kim et al. (2004) estimate a Markov-switching model for stock returns. They find evidence for a negative and significant volatility feedback effect, which supports a positive risk-return trade-off in normal times.

In particular, in a general equilibrium exchange economy, the sign of the risk-return relation depends on the sign of the correlation in between the marginal rate of substitution (or "stochastic discount factor") and the market return (see e.g. Whitelaw (2000)). Therefore, the parameter $\gamma(S_{t+1})$ entering before the conditional variance in equation (3) is not directly interpretable as the coefficient of relative risk-aversion. Instead, $\gamma(S_{t+1})$ corresponds to the product of the the volatility of the stochastic discount factor and the correlation between the stochastic discount factor and the market return.

[3] $N = \{5, 10, 22, 66\}$ for regressions at 1-week, 2-week, monthly and quarterly horizons.

Table 1: Summary statistics for monthly US excess stock returns

Statistic	1929:02 - 2010:12	1964:02 - 2010:12
Mean	0.399	0.387
Standard deviation	5.581	4.370
Minimum	-29.991	-21.954
Maximum	42.207	15.989
Number of observations	983	563

The last two columns report the sample statistics. Data are the S&P 500 composite portfolio returns obtained from the Global Financial Database website.

3 Data and empirical results

3.1 Data

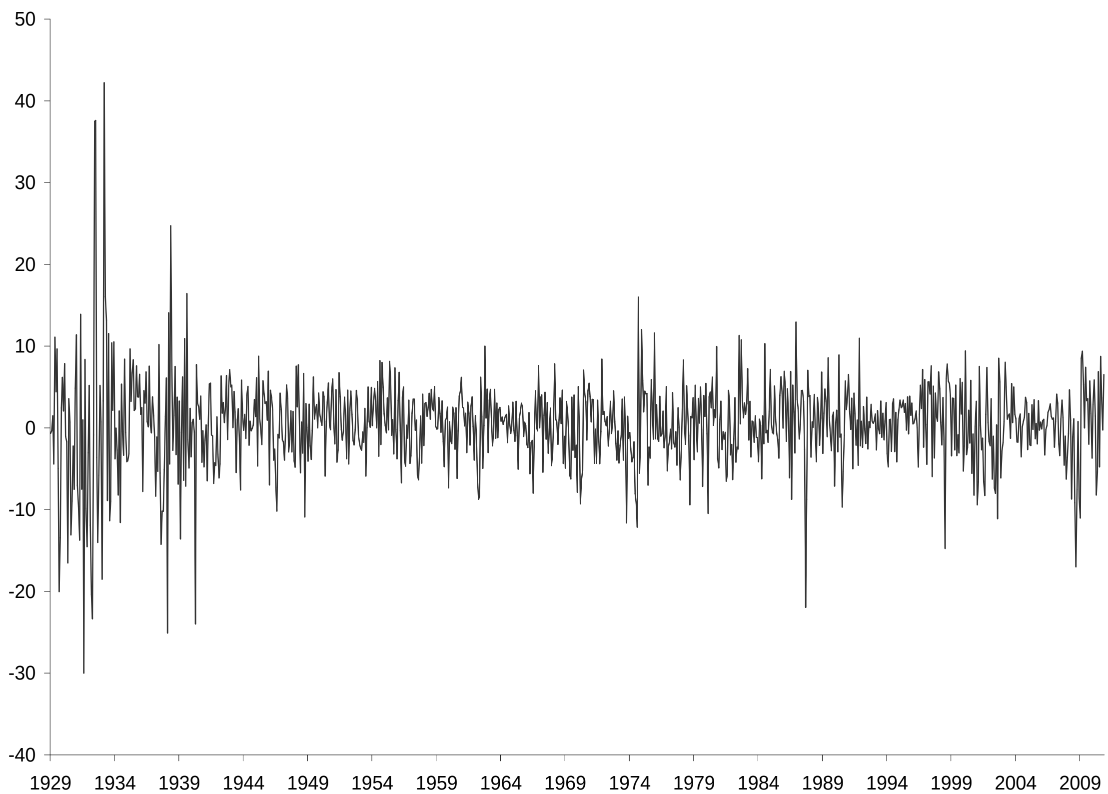
We use the S&P500 composite portfolio index ranging from February 1, 1929 to December 31, 2010 as a proxy for stock returns. The daily returns are taken as 100 times the daily change in the index. The risk-free rate is obtained from the 3-month Treasury bill, which is transformed at the daily frequency by appropriately compounding it. We use excess returns in the empirical analysis of the paper and for brevity we refer to them as returns. The data for stock returns are obtained from the Global Financial Data website. The risk-free rate series from 1929 to 1933 are the "*Yields on Short-Term US Securities Three-Six Month Treasury Notes and Certificates, Three Month Treasury*" from the NBER Macrohistory database. The risk-free rate from 1934 to 2010 is the 3-month Treasury bill taken from the Federal Reserve website.

Table 1 reports summary statistics for monthly excess returns. We consider two estimation samples: from 1929:02 to 2010:12 and from 1964:02 to 2010:12. Following Ghysels et al. (2005), we choose 1964 as the start year for the sub-sample analysis. The average monthly excess return over the full sample sample is 0.399%, which is slightly higher than in the shorter estimation sample 0.387%. The monthly excess returns over the full estimation sample also have higher standard deviation and a larger range than the shorter estimation sample. Figure 1 plots the data.

3.2 MIDAS and GARCH estimates of the risk-return relation

The MIDAS estimator of the conditional variance aggregates past absolute daily returns so that to compute the conditional variance for a given month N, we use daily returns until

Figure 1: MONTHLY EXCESS STOCK RETURNS 1929:02 - 2010:12



the last traded day of month $N-1$. The past daily returns are aggregated with the beta weight function since Ghysels et al. (2006) find that it performs well with S&P500 data.[4] We then regress the returns of month N on the MIDAS estimator of the conditional variance for month N to estimate the risk-return relation in equation (1).

The monthly realized absolute variance is computed from the within-month daily absolute returns:

$$RVAR_{t+1} = \sum_{d=0}^{D} |r_{t+1-d}|$$

where D is the number of traded days in month $t+1$. For brevity, in the sequel, we refer to realized absolute variance simply as realized variance.

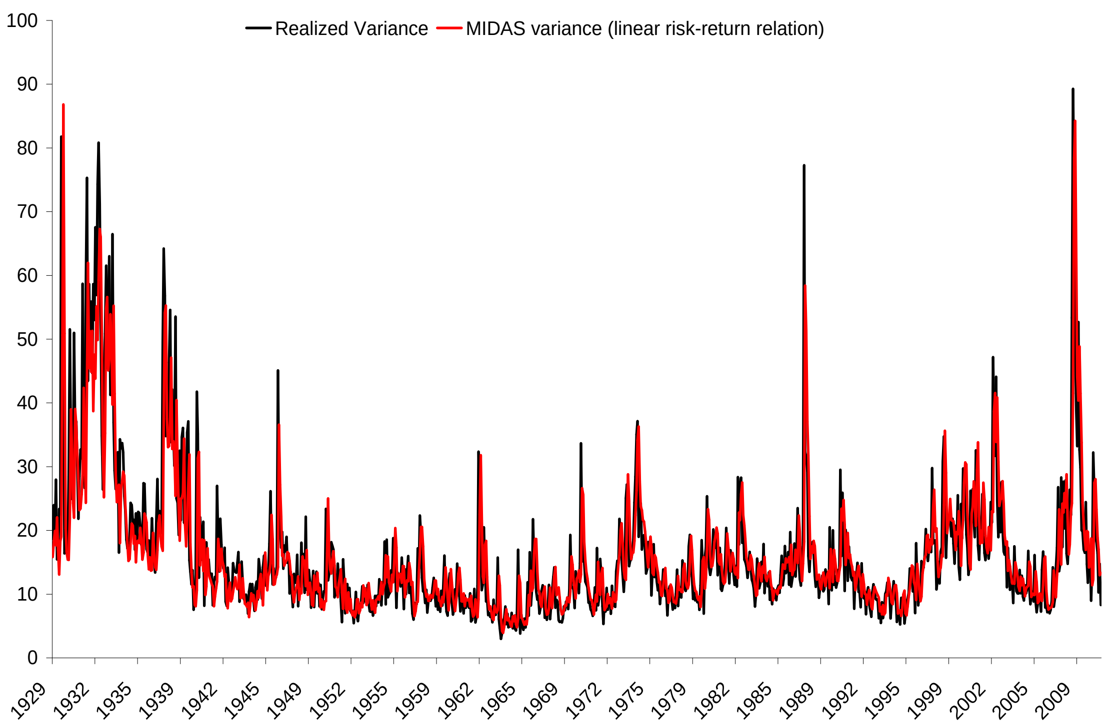
Table 2 reports the empirical results for the linear estimates of the risk-return trade-off using returns R_{t+1} for the LHS of equation (1) ranging from the weekly to the quarterly frequency. The results show a positive relation between expected returns and conditional volatility for both the sub-sample and full sample analyses and across all different frequencies for the expected returns R_{t+1}. However, the coefficient γ entering before the conditional variance is not significant at the 10% level except in the sub-sample analysis

[4]The use of exponential Almon lag weight function yields qualitatively similar results.

at the 2-week horizon. R_R^2 is the coefficient of determination from regressing R_{t+1} on the MIDAS estimator of the conditional variance. The explanatory power for the returns is low and typically increasing at lower frequency. The last column of Table 2 reports the $R_{\sigma^2}^2$s, which are obtained from the regression of the realized variance on the MIDAS estimator of the conditional variance. MIDAS estimators of the conditional variance explain from 48.71% to 58.74% of the realized variance. Besides, the predictive power of the MIDAS estimators is higher at the monthly frequency. Indeed, Figure 2 shows that the monthly MIDAS estimator of the conditional variance tracks very well the monthly realized variance.

These results, however, differ slightly from the findings of Ghysels et al. (2005) since they find a positive and significant risk-return trade-off. We see two reasons for this discrepancy: (i) our MIDAS estimator of the conditional variance is calculated from the absolute returns rather than the squared returns (ii) our estimation sample is longer as it includes the 2007-2009 financial crisis, which is likely to affect significantly the results previously reported in the literature.

Figure 2: MIDAS AND REALIZED VARIANCES 1929:02 - 2010:12



Another way to model the conditional variance is to use GARCH specifications. The GARCH-in-mean specification is another estimate of the risk-return trade-off (see for example French et al. (1987) and Glosten et al. (1993)). It is described by the following equations:

$$R_t = \mu + \gamma V_t^{GARCH} + \epsilon_t \tag{9}$$

Table 2: Linear risk return relation: $R_{t+1} \sim N(\mu + \gamma V_t^{MIDAS}, V_t^{MIDAS})$

	μ $(*10^2)$	γ	$LogL$	R_R^2	$R_{\sigma^2}^2$
Full sample analysis: February 1929 - December 2010					
Quarterly	0.129	0.025	-1227.190	0.94%	57.38%
	[0.030]	[0.273]			
Monthly	0.262	0.009	-2987.476	0.01%	58.74%
	[1.334]	[0.664]			
2-week	0.088	0.013	-5562.106	0.02%	56.26%
	[0.828]	[0.750]			
1-week	0.066	0.007	-9532.267	0.01%	50.56%
	[1.750]	[0.948]			
Sub-sample analysis: February 1964 - December 2010					
Quarterly	-0.327	0.034	-642.927	0.33%	47.52%
	[-0.339]	[1.418]			
Monthly	0.271	0.008	-1593.271	0.10%	54.12%
	[0.709]	[0.285]			
2-week	-0.001	0.026	-2992.930	0.03%	54.08%
	[-0.008]	[2.214]			
1-week	0.014	0.022	-5133.369	0.01%	48.71%
	[0.783]	[1.626]			

The MIDAS estimator of the conditional variance is computed using 120 lags for the daily absolute returns, which are aggregated with the beta polynomial weight function. T-statistics are computed from the inverse of the outer product estimate of the Hessian and are reported in brackets. $LogL$ is the value of the log likelihood function. R_R^2 is the coefficient of determination when regressing the returns on V_t^{MIDAS} and $R_{\sigma^2}^2$ is the coefficient of determination when regressing the realized variance on V_t^{MIDAS}.

$$V_t^{GARCH} = \omega + \alpha\epsilon_{t-1}^2 + \beta V_{t-1}^{GARCH} \quad (10)$$

The absolute GARCH-in-mean (ABSGARCH) specification is instead defined as:

$$(V_t^{ABSGARCH})^{1/2} = \omega + \alpha|\epsilon_{t-1}| + \beta(V_{t-1}^{ABSGARCH})^{1/2} \quad (11)$$

We use both Student-t innovations and Normal innovations and consider two different sample sizes (1929-2010 and 1964-2010). Table 3 presents the results for the monthly GARCH-in-mean and monthly absolute GARCH-in-mean specifications, estimated with quasi-maximum likelihood via the EM algorithm. First note that the use of Student-t innovations rather than Normal innovations increases the log-likelihood by about 20 in the full sample case, which is a significant gain from estimating a single parameter ν. In the shorter sample size, the increase in the log-likelihood is lower (about 10). Second, the estimates for γ - the parameter entering before the conditional variance - are positive in each case. However, it is significant only with the absolute GARCH-in-mean specification with Student-t and Normal innovations in the full sample period 1929-2010. Finally, the coefficients of determination $R^2_{\sigma^2}$s are roughly equivalent to their MIDAS counterparts and the R^2_Rs are higher, especially with Student-t innovations (see Table 2).

Figure 3 plots the ABSGARCH variance with the realized variance. Unlike the MIDAS variance, the ABSGARCH variance has troubles to accomodate the periods of high volatility ranging from 1929 to 1940.

3.3 MIDAS estimates of the regime switching risk-return relation

Table 4 provides the estimates for the regime switching risk-return relation described by equation (3).[5] For the full sample analysis (1929-2010), we find that for regressions at the 1-week, 2-week and monthly horizons, the coefficient γ_1 is negative and significant, while the coefficient in the second regime γ_2 is positive and significant. In both regimes, the coefficients γ_1 and γ_2 tend to be higher in absolute value at higher frequency, which indicates a steeper risk-return relation at higher frequencies. For the sub-sample 1964-2010, we find qualitatively similar results.[6]

An attractive feature of Markov-switching models is their ability to endogenously generate probabilities of being in a given regime. The unconditional probabilities of being in the first regime are low (between 2.49% and 20.09%) and are - as expected - typically higher in the full estimation sample (1929-2010) than in the shorter estimation sample (1964-2010). Besides, in the regime switching case, the coefficients of determination $R^2_{\sigma^2}$s are roughly equivalent to the linear case, and so are the R^2_Rs. The monthly MIDAS conditional variance obtained from the regime switching risk-return relation is very close to the monthly realized variance (see Figure 4).

[5]Note that considering regime changes only in the slope parameter γ yields qualitatively similar results.

[6]Table C1 in the appendix provides additional estimation results with different estimation window sizes. The results reported are consistent with those of Table 4.

Table 3: Monthly GARCH estimates of the risk-return relation

	Model	μ (x10^2)	γ	ω (x10^4)	α	β	ν	R^2_R	$R^2_{\sigma^2}$	LogL
Student-t innovations										
1929-2010	GARCH-in-mean	0.478	0.009	1.199	0.809	0.152	7.598	0.29%	47.27%	-2889.709
		[2.207]	[0.933]	[3.344]	[28.104]	[5.374]	[4.464]			
	ABSGARCH-in-mean	0.331	0.019	0.410	0.775	0.188	6.862	0.10%	58.07%	-2895.954
		[1.562]	[2.206]	[5.083]	[28.212]	[7.082]	[5.012]			
1964-2010	GARCH-in-mean	0.124	0.028	0.843	0.848	0.114	8.012	0.23%	30.95%	-1587.517
		[0.397]	[1.588]	[2.334]	[25.023]	[3.577]	[3.293]			
	ABSGARCH-in-mean	0.342	0.020	0.380	0.808	0.137	8.355	0.61%	53.89%	-1587.601
		[0.974]	[0.988]	[3.769]	[20.945]	[4.093]	[3.225]			
Normal innovations										
1929-2010	GARCH-in-mean	0.334	0.012	1.024	0.813	0.159	-	0.24%	47.13%	-2907.405
		[1.580]	[1.325]	[4.109]	[40.627]	[6.876]				
	ABSGARCH-in-mean	0.147	0.028	0.400	0.773	0.195	-	0.02%	58.46%	-2920.288
		[0.807]	[4.306]	[7.143]	[42.595]	[9.693]				
1964-2010	GARCH-in-mean	0.154	0.020	0.640	0.864	0.112	-	0.06%	28.18%	-1596.769
		[0.576]	[1.280]	[2.546]	[34.206]	[4.050]				
	ABSGARCH-in-mean	0.412	0.012	0.349	0.821	0.129	-	0.32%	51.93%	-1596.442
		[63.787]	[1.280]	[5.027]	[28.391]	[4.441]				

In the estimation, we impose constraints on the parameters ω, α and β to ensure that the conditional variance is positive. T-statistics are calculated from the inverse of the outer product estimate of the Hessian and are reported in brackets. R^2_R is the coefficient of determination when regressing the returns on the estimated GARCH variance and $R^2_{\sigma^2}$ is the coefficient of determination when regressing the realized variance on the estimated GARCH variance. LogL is the value of the log-likelihood function.

Figure 3: ABSGARCH AND REALIZED VARIANCES 1929:03 - 2010:12

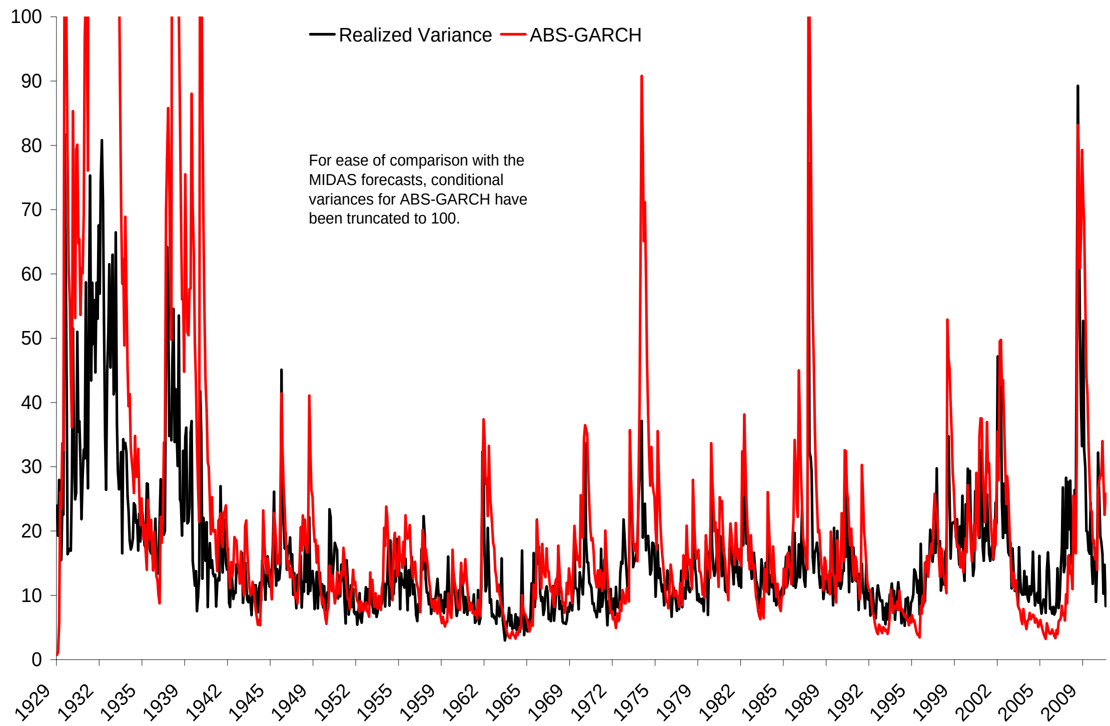
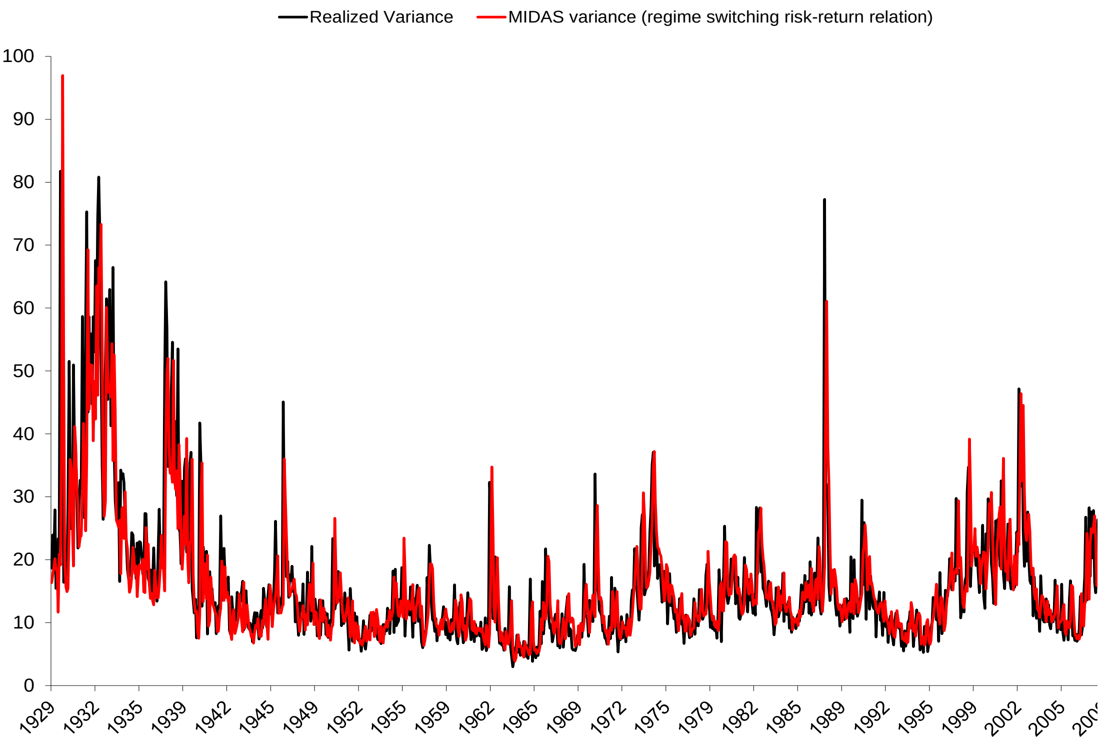


Figure 5 plots the weights attached to the lagged daily absolute returns at different frequencies for the regime-switching risk-return relation. For the 1-week and 2-week horizons, the weight function has a decreasing shape, whereas the weight function has a hump shape at the monthly and quarterly horizons. In all cases, the weights are negligible after 80 traded days, which emphasizes the importance of including more than a month of daily returns for measuring the conditional variance and the relevance of the MIDAS approach.

Figure 6 shows the estimated probability of being in the first regime (dotted line) and the actual returns (solid line), the probability is high in periods of high volatility and low returns. In particular, it peaks at one in all periods of financial turmoil.

To further understand the regime probabilities, we first run OLS regressions for the smoothed probabilities of the first regime on the slope of the yield curve, the expected returns, the changes in volatility and we control for business cycle conditions by including the Aruoba et al. (2009) index of business cycle conditions in the regression. Second, we use the same set of explanatory variables but instead run logistic regressions using as a dependent variable a dummy variable that takes on a value of 1 if the smoothed probability of being in regime 1 is higher than 0.5 and 0 otherwise. The results are reported in Table 5. First, expected returns always affect negatively and significantly the regime probabilities. Second, an increase in volatility is positively related to the regime probabilities. Third, the

Figure 4: MIDAS AND REALIZED VARIANCES 1929:02 - 2010:12



slope of the yield curve affects negatively and significantly the regime probabilities, except at the 2-week and quarterly horizons for the OLS regressions where the coefficient on the slope of the yield curve is not significant at the 10% level. This means that when the slope of the yield curve becomes less steep (resulting from a flight-to-quality episode for example) the probability of the first regime increases. This holds even when controlling for business cycle conditions as defined by the Aruoba et al. (2009) index of business cycle conditions.

Therefore, in the first regime - characterized by high volatility and low ex-post returns - we find that there is a reversed risk-return relation with a low premium for volatility. By contrast, in the second regime, a positive and significant risk-return relation holds. In addition, the first regime can be interpreted as a flight-to-quality regime since the slope of the yield curve appears to be negatively related to the regime probabilities of the first regime. As noted earlier, this evidence corroborates the findings in Ghysels et al. (2013) who estimate the risk-return relationship using a simple flight-to-quality indicator.

We now compare the different estimated variance processes in Table 6. Panel A reports the means, variances and goodness-of-fit measures for the MIDAS (for both linear and non-linear cases) and ABSGARCH conditional variances using the realized variance as a benchmark. The goodness-of-fit measure is computed as one minus the sum of the absolute differences between the estimated conditional variance and the realized variance divided by

Table 4: Regime-switching risk-return relation: $R_{t+1} \sim N(\mu(S_{t+1}) + \gamma(S_{t+1})V_t^{MIDAS}, V_t^{MIDAS})$

	p_{11}	p_{22}	μ_1 $(*10^2)$	μ_2 $(*10^2)$	γ_1	γ_2	$LogL$	R_R^2	$R_{\sigma^2}^2$	$P(S_t = 1)$
Full sample analysis: February 1929 - December 2010										
Quarterly	0.516	0.907	-13.106	-1.845	0.006	0.135	-1174.979	1.99%	52.76%	16.13%
	[4.475]	[32.219]	[-4.143]	[-1.738]	[0.159]	[4.377]				
Monthly	0.255	0.934	-2.610	0.158	-0.346	0.066	-2915.114	0.02%	54.17%	8.11%
	[3.352]	[43.215]	[-1.548]	[0.517]	[-3.223]	[2.827]				
2-week	0.269	0.938	-2.228	-0.137	-0.459	0.122	-5407.312	0.01%	56.38%	7.86%
	[4.464]	[72.979]	[-2.810]	[-0.886]	[-6.281]	[4.993]				
1-week	0.315	0.914	-0.387	-0.088	-0.747	0.174	-9271.982	0.01%	50.79%	11.11%
	[6.454]	[37.292]	[-0.843]	[-1.009]	[-9.878]	[4.016]				
Sub-sample analysis: February 1964 - December 2010										
Quarterly	0.647	0.911	-5.256	-2.342	-0.035	0.138	-632.893	0.30%	50.48%	20.09 %
	[3.402]	[13.357]	[-0.829]	[-1.496]	[-0.319]	[3.444]				
Monthly	0.202	0.944	-1.551	0.208	-0.303	0.043	-1583.303	0.02%	53.23%	6.54%
	[1.059]	[22.766]	[-0.530]	[0.446]	[-1.360]	[1.157]				
2-week	0.104	0.977	-1.716	-0.039	-0.789	0.061	-2951.106	0.04%	53.75%	2.49%
	[1.039]	[99.000]	[-1.092]	[-0.441]	[-3.072]	[3.322]				
1-week	0.273	0.961	-0.238	-0.033	-0.924	0.094	-5070.936	0.01%	48.51%	5.08%
	[3.005]	[50.523]	[-0.212]	[-0.487]	[-4.557]	[4.675]				

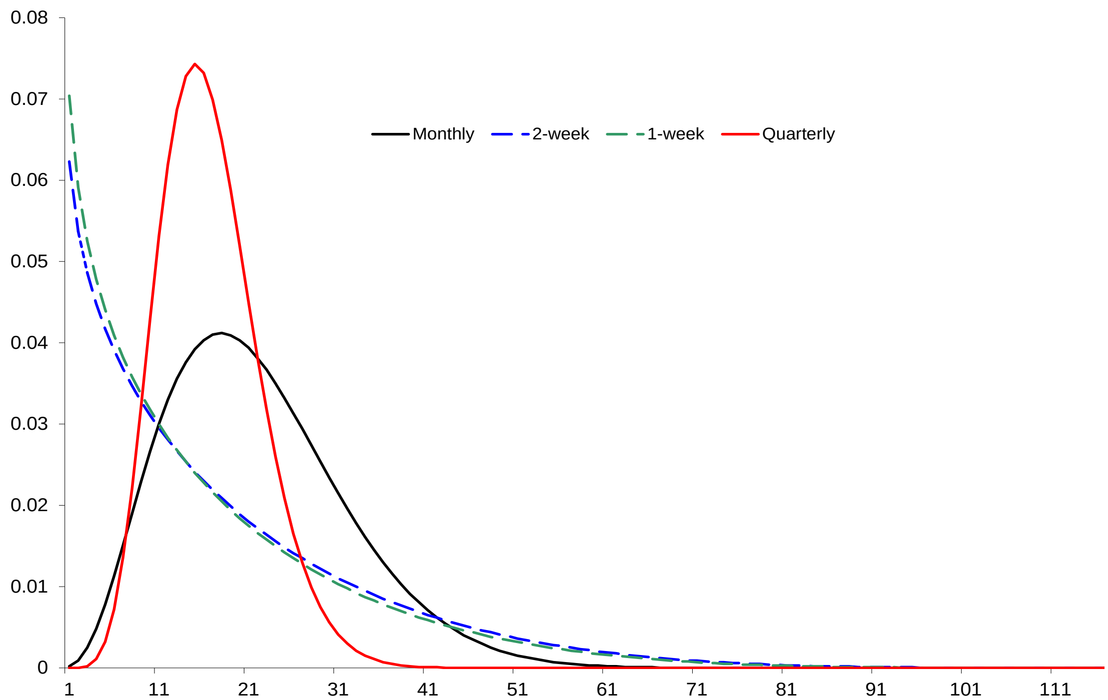
The MIDAS estimator of the conditional variance is calculated using 120 lags for the daily absolute returns, which are aggregated with the beta polynomial weight function. T-statistics are calculated from the inverse of the outer product estimate of the Hessian and are reported in brackets. $LogL$ is the value of the log likelihood function. R_R^2 is the coefficient of determination when regressing the returns on V_t^{MIDAS} and $R_{\sigma^2}^2$ is the coefficient of determination when regressing the realized variance on V_t^{MIDAS}. p_{11} and p_{22} are the transition probabilities of staying in the first and second regime, respectively. $P(S_t = 1)$ is the unconditional probability of being in the first regime.

Table 5: Explaining the regime probabilities $P(S_{t+1})$

	Slope of the yield curve$_{t+1}$	ΔV_{t+1}^{MIDAS}	R_{t+1}	ADS_{t+1}
Panel A: OLS regression				
1-week	-0.006***	0.016***	-0.039***	-0.009**
2-week	-0.003	0.007***	-0.021***	-0.004
Monthly	-0.007**	0.001	-0.023***	-0.011**
Quarterly	-0.010	0.001	-0.024***	-0.098***
Panel B: Logistic regression				
1-week	-0.372***	0.331	-1.450***	0.286
2-week	-1.163***	0.019	-1.378***	1.625***
Monthly	-0.591*	0.010	-1.086***	0.229
Quarterly	-0.375*	0.004	-0.287***	-1.308***

Panel A reports the results of OLS regressions of the estimated smoothed probability of being in the first regime $P(S_{t+1})$ on the level of the slope of the yield curve, the changes in the MIDAS estimator of the conditional variance ΔV_{t+1}^{MIDAS}, the expected returns R_{t+1} and the level of the ADS index of business cycle conditions ADS_{t+1}. Panel B reports results of logistic regressions using a dummy variable as a dependent variable and the same set of explanatory variables. The dummy variable takes on a value of 1 if the smoothed probability of being in regime 1 is higher than 0.5 and 0 otherwise. The slope of the yield curve is defined as the difference between the yields on a 10-year Treasury bond and the yields on a 3-month Treasury bill. *, **, *** indicate significance at the 10% level, 5% level and 1% level, respectively. We only use the sub-sample 1964-2010 since we do not have data for the ADS index and the weekly slope of the yield curve for the whole full sample.

Figure 5: WEIGTHS FOR THE MIDAS ESTIMATOR OF THE CONDITIONAL VARIANCE (REGIME-SWITCHING RISK-RETURN RELATION) AT DIFFERENT FREQUENCIES 1929:02 - 2010:12



the sum of the realized variance. The means and the variances of the MIDAS estimators of the conditional variances are close but slightly below the mean and the variance of the realized variance. The mean and variance of the ABSGARCH variance are instead strongly higher than the mean and variance of the realized variance. The goodness-of-fit measure is higher for the MIDAS estimators of the conditional variance than the ABSGARCH variance. This is particularly acute in the full sample case, which is expected since the ABSGARCH variance has troubles to accommodate the high volatility episodes of the late 1920's and 1930's.

Panel B of Table 6 reports the cross-correlation matrix for the MIDAS (for both the linear and non-linear cases), the ABSGARCH conditional variances and the realized variance. The MIDAS conditional variance in the linear case exhibits the highest correlation with the realized variance for both samples. Not surprisingly, the MIDAS conditional variances in the linear and non-linear cases are very highly correlated. The ABSGARCH conditional variance is the second best correlated with the realized variance although they have smaller goodness-of-fit values than the MIDAS conditional variances (see last column of Panel A).

Figure 7 provides further insights about the variance processes under scrutiny. Panels A, B and C plot the MIDAS conditional variances (both in the linear and non-linear

Table 6: Comparison of the monthly variance processes

Panel A: Summary Statistics

Full sample analysis: February 1929 - December 2010

Estimator	Mean ($x10^4$)	Variance ($x10^8$)	Goodness-of-fit
Realized	16.376	133.302	-
MIDAS (linear)	16.170	104.876	0.722
MIDAS (MS)	16.180	109.682	0.706
ABSGARCH	27.836	2271.446	0.138

Sub-sample analysis: February 1964 - December 2010

Estimator	Mean ($x10^4$)	Variance ($x10^8$)	Goodness-of-fit
Realized	14.583	73.508	-
MIDAS (linear)	15.198	72.466	0.734
MIDAS (MS)	15.221	70.093	0.734
ABSGARCH	15.346	132.171	0.647

Panel B: Correlations

Full sample analysis: February 1929 - December 2010

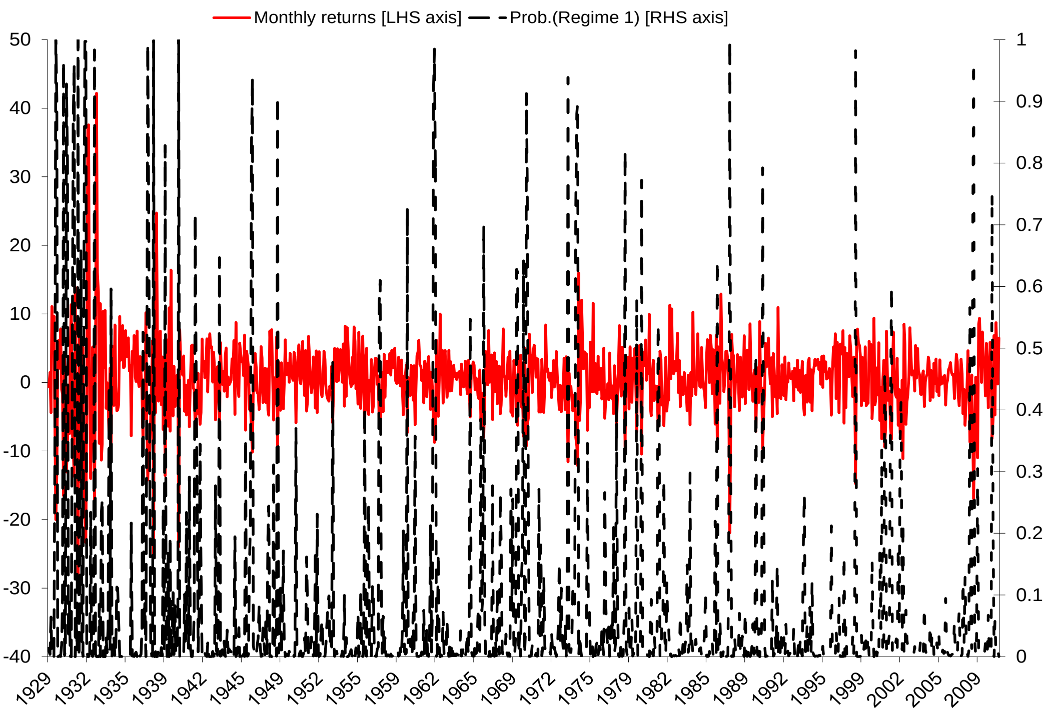
	Realized	MIDAS (linear)	MIDAS (MS)	ABSGARCH
Realized	1	-	-	-
MIDAS (linear)	0.766	1	-	-
MIDAS (MS)	0.736	0.988	1	-
ABSGARCH	0.759	0.715	0.706	1

Sub-sample analysis: February 1964 - December 2010

	Realized	MIDAS (linear)	MIDAS (MS)	ABSGARCH
Realized	1	-	-	-
MIDAS (linear)	0.736	1	-	-
MIDAS (MS)	0.730	0.996	1	-
ABSGARCH	0.734	0.762	0.770	1

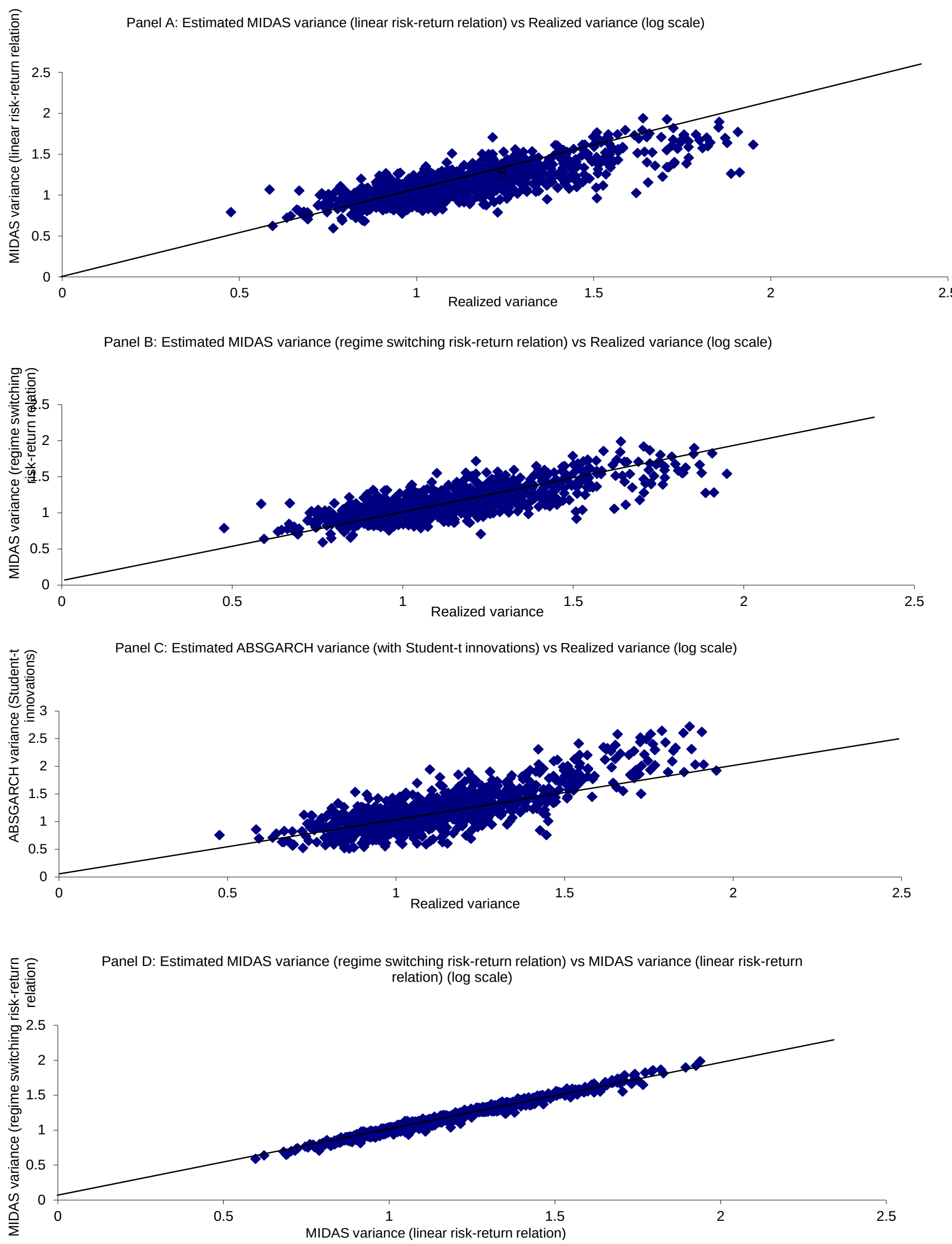
Panel A reports summary statistics for the MIDAS estimated conditional variances, the realized variance and the ABSGARCH conditional variances with Student-t innovations. The goodness-of-fit measure is calculated as one minus the sum of absolute differences between the estimated variance process and the realized variance divided by the sum of realized variance. Panel B reports a cross-correlation matrix for the different variance processes under scrutiny.

Figure 6: MONTHLY RETURNS AND PROBABILITIES OF BEING IN THE FIRST REGIME 1929:02-2010:12



cases) and the ABSGARCH variance against the realized variance with a 45° line, which indicates a perfect fit with the realized variance. The MIDAS variances show no clear sign of asymmetry (panels A and B), whereas the estimated ABSGARCH variance (Panel C) shows that the ABSGARCH variance tends to overestimate the realized variance. Finally, Panel D of Figure 7 plots the MIDAS variance in the regime switching case against the MIDAS variance in the linear case: this shows that the MIDAS variances are very close to each other.

Figure 7: scatterplots of the monthly variances 1929:02 - 2010:12



3.4 Testing for Markov-switching

Testing for parameter changes in Markov-switching models is a difficult issue since under the null hypothesis of constant parameters (i) the transition probabilities are not identified and (ii) the scores of the log-likelihood are identically equal to zero. Hansen (1992) and Garcia (1998) proposed tests for Markov-switching but these tests require to estimate the model under the alternative hypothesis and are often computationally very expensive. Recently, Carrasco et al. (2013) have introduced a new test for Markov-switching parameters that only requires to estimate the model under the null hypothesis of constant parameters. Appendix A details the Carrasco et al. (2013) test for Markov-switching parameters. Table 7 reports their test statistics for regressions at 1-week, 2-week, monthly and quarterly horizons and the corresponding 5% bootstrapped critical values.

There is overwhelming evidence for regime changes in the risk-return relation since the null hypothesis is rejected at the 5% level in all cases. Note that the test statistics are higher for the full sample estimates (1929-2010) than in the shorter sample (1964-2010). This is expected since the full sample contains periods of higher volatility and is thus more prone to exhibit non-linear behavior. Besides, the test-statistics are higher with higher frequency data for both samples, which indicates that the evidence for regime switching is stronger at higher frequencies.

Note that the above test requires the parameters to be constant under the null so that we cannot test a 3-regime model against a 2-regime model. We nevertheless report in appendix B goodness-of-fit measures for these two models and the linear model. First, the linear model is always outperformed in terms of SIC by the Markov-switching models. Second, for the subsample period 1964-2010, the 2-regime model is preferred at the quarterly and monthly horizons since it obtains the lowest SIC for these regressions, whereas the 3-regime model gets the lowest SIC at the 1-week and 2-week horizons. Third, the 3-regime model always obtains the lowest SIC for the full sample estimates. However, the three regime switching parameters $\gamma(S_{t+1})$ are not all significant at the 10% level at the monthly and quarterly horizons. In addition, the SIC tends to overestimate the true number of regimes (see e.g., Smith et al. (2006)), particularly when parameter changes are small.

Finally, we also consider models with switches in all parameters of the model (that is, μ, γ and the MIDAS parameters κ_1 and κ_2). In this way, the weight function also changes across regimes. The SICs for these models are reported in the fifth column of Table B in the appendix, which shows that these models are always outperformed by the regime-switching models with constant parameters κ_1 and κ_2 (except for the sub-sample analysis at the monthly horizon).

We therefore decide to keep the model with two regimes and regime changes in the parameters μ and γ in the subsequent analysis.

Table 7: Tests of regime switching in the risk-return relation

		Carrasco et al. test statistic	5% Bootstrapped critical values
1929-2010	Quarterly	9.265	2.724
	Monthly	14.126	3.443
	2-week	21.456	4.522
	1-week	54.707	5.605
1964-2010	Quarterly	4.358	2.445
	Monthly	3.987	3.060
	2-week	6.044	3.807
	1-week	17.121	4.397

This table shows the Carrasco et al. (2013) test statistics and the corresponding 5% bootstrapped critical values. Under the null hypothesis, there is no regime switching in the risk-return relation. The bootstrapped critical values are based on 1000 Monte Carlo repetitions. Appendix A details the test.

4 Sensitivity analysis

4.1 Additional predictors in the risk-return relation

The lack of conditioning variables is often cited as a source of misspecification for the estimates of the risk-return trade-off (see e.g., the literature review in Lettau and Ludvigson (2010)). Guo and Whitelaw (2006) use two additional predictors: the consumption-wealth ratio from Lettau and Ludvigson (2001) and the stochastically detrended risk-free rate to approximate the hedge component of Merton (1973)'s model. Ludvigson and Ng (2007) use factors extracted from a large macroeconomic and financial database to enlarge the information set. Both studies conclude that including additional predictors allows to uncover a positive risk-return trade-off.

Table 8 presents the results when we include as additional predictors the lagged returns R_t, the slope of the yield curve $Slope_t$, the dividend-price ratio $(D/P)_t$ and the realized covariance Cov_t in the risk-return relation. The realized covariance measure is computed as the product between the daily changes in the Aruoba et al. (2009) index of business cycle conditions and the expected returns. Rossi and Timmermann (2010) show that the changes in the ADS index are highly correlated with the changes in consumption, the realized covariance can then be seen as an approximation for the time-varying risk premium on consumption that is likely to be important for the estimation of the risk-return trade-off as emphasized by Tauchen (2004). More generally, it can be seen as a way of controlling for business cycle conditions. Monthly realized covariance is calculated as follows:

$$Cov_t = \sum_{i=1}^{N} \Delta ADS_{i,t} * R_{i,t}$$

where $\Delta ADS_{i,t}$ is the daily change in the ADS index on day i of month t and $R_{i,t}$ is the corresponding stock return.

The slope of the yield curve is taken as the difference between the 10-year Treasury bond and the 3-month Treasury bill. The dividend-price ratio is the difference between the log of dividends and the log of prices, where dividends are 12-month moving sums of dividends. The data for the 10-year Treasury bond and the dividend-price ratio are from Robert Schiller's website.

Note that, unlike a large part of the literature, we consider returns sampled from the weekly to the quarterly frequency to describe more precisely the dynamics of the risk-return trade-off. The results suggest the following.[7] First, across all frequencies we consider, the risk-return relation is reversed in the first regime, while it is positive in the second regime. Second, the risk-return relation is typically steeper at higher frequencies since the coefficients entering before the conditional variance are higher in absolute value at higher frequencies. Third, expected returns, the dividend-price ratio and the slope of yield curve enter positively and significantly in the risk-return relation at the quarterly horizon. Overall, the results do not differ much from Table 4, suggesting that the detected regime switching risk-return relation is robust to the inclusion of additional predictors.

[7] Note that the full sample analysis does not include Cov_t as an additional predictor since the ADS index of business cycle conditions is not available before 1960. Likewise, we do not include the slope of the yield curve and the dividend-price ratio at the 1-week and 2-week horizons due to data availability.

Table 8: Regime-switching risk-return relation with additional predictors

	p_{11}	p_{22}	μ_1 $(*10^2)$	μ_2 $(*10^2)$	γ_1	γ_2	R_t	$(D/P)_t$	$Slope_t$	Cov_t	$LogL$	R^2_R	$R^2_{\sigma^2}$	$P(S_t = 1)$
Full sample analysis: February 1929 - December 2010														
Quarterly	0.372	0.911	-12.576	-0.888	-0.036	0.119	0.113	6.148	0.971	-	-1164.396	2.42%	50.97%	12.47%
	[3.573]	[35.049]	[-3.368]	[-0.689]	[-0.812]	[4.908]	[2.068]	[2.889]	[2.491]					
Monthly	0.276	0.934	-1.737	0.871	-0.369	0.060	-0.023	1.990	0.120	-	-2908.059	0.03%	52.78%	8.40%
	[3.631]	[45.627]	[-0.982]	[1.743]	[-3.351]	[2.550]	[-0.469]	[2.771]	[0.674]					
2-week	0.266	0.938	-2.238	-0.138	-0.459	0.122	0.002	-	-	-	-5407.307	0.01%	56.38%	7.83%
	[4.335]	[73.921]	[-2.958]	[-0.985]	[-6.343]	[5.299]	[0.214]							
1-week	0.398	0.919	-0.442	-0.079	-0.725	0.181	-0.074	-	-	-	-9263.007	0.01%	50.79%	11.82%
	[7.178]	[58.010]	[-1.128]	[-1.127]	[-9.931]	[6.625]	[-4.261]							
Sub-sample analysis: February 1964 - December 2010														
Quarterly	0.001[a]	0.912	12.408	-0.206	-0.567	0.073	0.177	4.580	0.677	2.848	-624.871	0.34%	48.89%	8.12%
		[19.675]	[1.426]	[-0.167]	[-2.506]	[2.673]	[2.305]	[1.690]	[1.826]	[1.532]				
Monthly	0.324	0.934	-1.399	0.695	-0.223	0.056	-0.060	1.557	0.136	1.025	-1578.731	0.04%	52.55%	8.94%
	[1.297]	[19.154]	[-0.530]	[1.034]	[-1.156]	[1.178]	[-1.003]	[1.477]	[0.970]	[0.899]				
2-week	0.094	0.977	-1.792	-0.046	-0.779	0.061	0.011	-	-	0.245	-2950.941	0.04%	53.82%	2.48%
	[0.932]	[98.397]	[-1.082]	[-0.929]	[-2.959]	[3.772]	[0.252]			[0.478]				
1-week	0.393	0.953	-0.359	-0.016	-0.790	0.108	-0.086	-	-	0.823	-5063.848	0.01%	48.54%	7.13%
	[4.859]	[61.597]	[-0.603]	[-0.537]	[-4.302]	[5.085]	[-3.541]			[1.474]				

The MIDAS estimator of the conditional variance is calculated using 120 lags for the daily absolute returns, which are aggregated with the beta polynomial weight function. T-statistics are computed from the inverse of the outer product estimate of the Hessian and are reported in brackets. $LogL$ is the value of the log likelihood function. R^2_R is the coefficient of determination when regressing the returns on V_t^{MIDAS} and $R^2_{\sigma^2}$ is the coefficient of determination when regressing the realized variance on V_t^{MIDAS}. p_{11} and p_{22} are the transition probabilities of staying in the first and second regime, respectively. $P(S_t = 1)$ is the unconditional probability of being in the high volatility regime. The additional predictive variables are the lagged returns (R_t), the dividend-price ratio ($(D/P)_t$), the slope of the yield curve ($Slope_t$) and the realized covariance (Cov_t). [a] In that case, the transition probability for regime 1 hit the lower bound that was imposed to respect the properties of a Markov chain.

4.2 Controlling for asymmetries in stock returns

Modelling asymmetries in the process for conditional variance is potentially important since one can expect different responses of the conditional variance following negative or positive shocks. For example, Glosten et al. (1993) find that the sign of the risk-return trade-off becomes negative when allowing for a different effect of positive and negative returns on the conditional variance. Ghysels et al. (2005) instead introduce the asymmetric MIDAS estimator of the conditional variance, which gives different weights to the lagged returns depending on whether they are positive or negative. They find that negative returns have a stronger effect on the conditional variance upon impact but this effect dies away quickly, whereas positive returns have a smaller effect upon impact but are more persistent.

The asymmetric MIDAS estimator of the conditional variance is given by:

$$V_t^{ASYMIDAS} = N[\phi \sum_{d=0}^{\infty} w_d(\kappa_1^-, \kappa_2^-) 1_{t-d}^- |r_{t-d}| + (2-\phi) \sum_{d=0}^{\infty} w_d(\kappa_1^+, \kappa_2^+) 1_{t-d}^+ |r_{t-d}|] \tag{12}$$

where 1_{t-d}^- is the indicator function for $\{r_{t-d} < 0\}$ and 1_{t-d}^+ is the indicator function for $\{r_{t-d} \geq 0\}$.

Table 9 reports the results when estimating a linear and regime switching risk-return relation at the monthly frequency with an Asymmetric MIDAS estimator of the conditional variance. First, the results are broadly consistent with Table 4. In the linear case, the coefficients γ entering before the conditional variance are not significant at the 10% level for both the full sample and sub-sample analyses. In the regime-switching case, the risk-return relation is reversed in the first regime, while the traditional positive risk-return trade-off holds in the second regime. Besides, the coefficient ϕ - that governs the weights allocated to the negative returns - is higher than 1 in all cases, which suggests that negative returns have a stronger impact on the conditional variance than positive returns. In addition, the asymmetric MIDAS estimator cannot be rejected by a standard likelihood ratio test of no asymmetries (i.e. $\kappa_1^+ = \kappa_1^-$, $\kappa_2^+ = \kappa_2^-$, $\phi = 1$) except for the regime switching risk-return relation in the full sample case where the p-value exceeds .05.

Figure 8 plots the weights attached to the positive and negative returns, the overall asymmetric weights and the symmetric weights for a regime switching risk-return relation. The positive weights have a bell shape with a maximum effect on the conditional variance after about 20 traded days. The negative returns have a maximum effect on the conditional variance upon impact and the effect dies away after 80 traded days.[8] Overall, the symmetric and asymmetric weights are relatively close from each other.

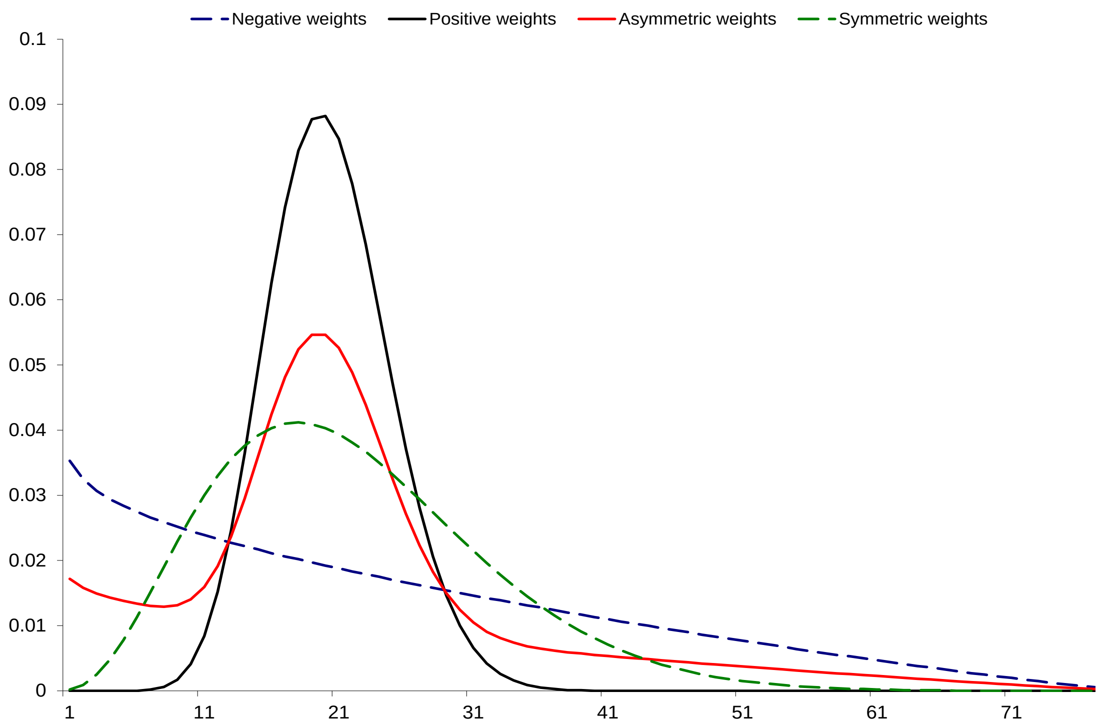
[8]We find the same shapes for the weight functions when we use Student-t rather than Normal innovations and the exponential Almon lag weigth function rather than the beta polynomial weight function for aggregating the lagged daily absolute returns. We use 80 daily lagged returns for estimating the Asymmetric MIDAS estimator of the conditional variance since we encountered convergence problems of the algorithm when we included more than 80 daily lagged returns.

Table 9: Monthly estimates of the risk-return trade-off with Asymmetric MIDAS estimators of the conditional variance

	p_{11}	p_{22}	μ_1 $(*10^2)$	μ_2 $(*10^2)$	γ_1	γ_2	ϕ	$LogL$	$LRtest$	R^2_R	$R^2_{\sigma^2}$	$P(S_t = 1)$
Asymmetric MIDAS conditional variance (regime-switching risk-return relation)												
1929:02 - 2010:12	0.258	0.890	-1.939	0.204	-0.275	0.082	1.514	-2911.398	7.432	0.02%	59.19%	12.89%
	[3.153]	[21.354]	[1.580]	[0.766]	[-3.954]	[3.099]	[7.575]		[0.059]			
1964:02 - 2010:12	0.032	0.882	-4.667	0.930	-0.036	0.009	1.592	-1590.444	15.718	0.05%	49.30%	10.84%
	[0.166]	[18.604]	[-5.064]	[3.132]	[-0.371]	[0.374]	[11.520]		[0.001]			
Asymmetric MIDAS conditional variance (linear risk-return relation)												
1929:02 - 2010:12	-	-	0.357	-	0.003	-	1.242	-2977.774	19.404	0.02%	58.09%	-
			[1.593]		[0.173]		[7.557]		$[2*10^{-4}]$			
1964:02 - 2010:12	-	-	0.267	-	0.008	-	1.690	-1585.467	15.608	0.03%	55.07%	-
			[0.907]		[0.357]		[14.148]		[0.001]			

The Asymmetric MIDAS estimator of the conditional variance is computed following equation (12). The estimators of the conditional variance are computed using 80 lags for the daily absolute returns, which are aggregated with the beta polynomial weight function. T-statistics are computed from the inverse of the outer product estimate of the Hessian and are reported in brackets. $LogL$ is the value of the log likelihood function. $LRtest$ reports the value of a likelihood ratio test statistic. Under the null hypothesis of no asymmetric effects for the MIDAS conditional variance $\kappa_1^+=\kappa_1^-$, $\kappa_2^+=\kappa_2^-$ and $\phi = 1$, p-values for the LR tests are reported in brackets. R^2_R is the coefficient of determination when regressing the returns on V_t^{MIDAS} and $R^2_{\sigma^2}$ is the coefficient of determination when regressing the realized variance on V_t^{MIDAS}. p_{11} and p_{22} are the transition probabilities of staying in the first and second regime, respectively. $P(S_t = 1)$ is the unconditional probability of being in the high volatility regime.

Figure 8: WEIGHTS FOR THE ASYMIDAS ESTIMATOR OF THE CONDITIONAL VARIANCE (REGIME SWITCHING RISK-RETURN RELATION) 1929:02-2010:12



4.3 The risk-return trade-off with Student-t innovations

As an additional robustness check, we use a Student-t rather than a Normal distribution for the innovations since the Student-t distributions can better account for outliers that are present in stock returns than the Normal distribution. The log-likelihood function is then written as:

$$\mathcal{L}_{\mathcal{T}}(\theta) = \sum_{t=1}^{T} l_t(\theta) \tag{13}$$

where:

$$l_{t+1}(\theta) = ln\Gamma(\frac{1+\nu}{2}) - ln\Gamma(\frac{\nu}{2}) - 0.5ln(\pi(\nu-2)) - 0.5ln(V_t^{MIDAS}) - \frac{(\nu+1)}{2} ln(1 + \frac{\epsilon_{t+1}(S_{t+1})^2}{(\nu-2)V_t^{MIDAS}})$$

and,

$$\epsilon_{t+1}(S_{t+1}) = R_{t+1} - \mu(S_{t+1}) - \gamma(S_{t+1})V_t^{MIDAS}$$

$\Gamma(.)$ is the Gamma function, ν are the degrees of freedom for the Student-t innovations and θ is the vector of parameters to be estimated. The maximum likelihood estimates $\theta^{\hat{MLE}}$ are obtained with the EM algorithm and are reported in Table 10.

First, the coefficient γ_1 is always negative, whereas the coefficient γ_2 is always positive in the second regime. Both coefficients are significant across all frequencies we consider (except for γ_1 in the sub-sample analysis at the quarterly horizon). This is in line with the results reported in Table 4. However, in absolute terms, the coefficient γ_1 is smaller than in Table 4 (except at the quarterly frequency). This is not surprising since the use of Student-t innovations - unlike Normal innovations - makes the estimates less sensitive to outliers. As a result, the first regime now captures periods with less volatile and less negative returns. This translates into higher unconditional probabilities of being in the first regime. Conversely, the coefficients γ_2 are typically higher than in Table 4 as the second regime captures fewer episodes of negative returns and moderate volatility, which are now mostly associated with the first regime.

The $R^2_{\sigma^2}$'s are comparable to those reported in Table 4, expect for regressions at the monthly and quarterly frequencies where the coefficients of determination for the realized variance $R^2_{\sigma^2}$ are higher for the full sample (1929-2010) estimates.

Table 11 reports the results when regressing the smoothed probabilities of being in the first regime on the slope of the yield curve, the expected returns, the changes in volatility and the Aruoba et al. (2009) index of business cycle conditions. We also report in Table 11 results for logistic regressions using as a dependent variable a dummy variable that takes on a value of 1 if the smoothed probability of being in regime 1 is higher than 0.5 and zero otherwise. First, the coefficients for the slope of the yield curve are negative (except at the 1-week horizon for OLS regressions and 1-week and monthly horizons for logistic regressions). Second, the changes in volatility affects positively the regime probabilities (except at the 1-week horizon). Third, the coefficients on expected returns are negative and strongly significant, which is consistent with the results reported in Table 5. The coefficient on the ADS index of business cycle conditions is negative and significant (except at the 2-week horizon in the case of logistic regressions). Overall, the results are broadly consistent with those presented in Table 5 in that the first regime tends to be characterized by a flattening of the yield curve, a weakening of economic activity as well as an increase in volatility owing to negative returns.

4.4 Time-varying transition probabilities

In this sub-section, we consider the use of time-varying transition probabilities since (i) we have provided evidence that some variables can explain the pattern of the probability of being in a given regime; (ii) it can help us to better understand the regime probabilities; (iii) it could improve the fit with respect to Markov-switching models with constant transition probabilities. Filardo (1994) relaxed the assumption of constant transition probabilities and use logistic functions to bound the transition probabilities between 0 and 1. The transition probability matrix P is then given by:

$$P = \begin{bmatrix} p_t^{11} = q(z_t) & p_t^{12} = 1 - p(z_t) \\ p_t^{21} = 1 - q(z_t) & p_t^{22} = p(z_t) \end{bmatrix}$$

Table 10: Regime-switching risk-return relation with Student-t innovations

	p_{11}	p_{22}	μ_1 $(*10^2)$	μ_2 $(*10^2)$	γ_1	γ_2	ν	$LogL$	R^2_R	$R^2_{\sigma^2}$	$P(S_t=1)$
Full sample analysis: February 1929 - December 2010											
Quarterly	0.509	0.782	-1.732	0.985	-0.104	0.088	4.189	-1159.082	0.23%	63.01%	30.72%
	[8.428]	[14.177]	[-1.170]	[0.913]	[-4.023]	[3.621]	[5.967]				
Monthly	0.380	0.703	-0.260	0.603	-0.196	0.109	5.022	-2882.261	0.01%	62.66%	32.40%
	[5.781]	[10.961]	[-0.517]	[1.597]	[-5.406]	[4.166]	[6.605]				
2-week	0.383	0.825	-0.105	0.152	-0.388	0.136	4.966	-5331.816	0.01%	56.74%	22.09%
	[5.967]	[25.759]	[-0.459]	[0.779]	[-5.940]	[3.580]	[9.006]				
1-week	0.407	0.660	0.172	-0.053	-0.444	0.311	5.332	-9167.628	0.01%	50.80%	36.47%
	[12.610]	[14.774]	[1.718]	[-2.148]	[-10.570]	[15.799]	[11.241]				
Sub-sample analysis: February 1964 - December 2010											
Quarterly	0.563	0.857	-2.526	-1.955	-0.079	0.142	4.568	-630.823	0.36%	51.01%	24.63%
	[3.255]	[10.310]	[-0.710]	[-0.664]	[-0.962]	[2.105]	[2.193]				
Monthly	0.388	0.785	-0.374	0.050	-0.162	0.107	6.766	-1579.143	0.05%	55.12%	26.02%
	[2.545]	[3.979]	[-0.264]	[0.187]	[-2.866]	[3.045]	[1.814]				
2-week	0.403	0.813	-0.650	0.012	-0.234	0.150	4.061	-2931.188	0.07%	53.68%	23.87%
	[5.590]	[15.490]	[-1.174]	[0.258]	[-2.892]	[6.144]	[7.176]				
1-week	0.382	0.486	0.374	-0.245	-0.360	0.350	7.344	-5056.935	0.01%	47.95%	45.37%
	[3.952]	[24.558]	[1.657]	[-1.670]	[-4.677]	[10.769]	[4.361]				

The MIDAS estimator of the conditional variance is computed using 120 lags for the daily absolute returns, which are aggregated with the beta polynomial weight function. T-statistics are computed from the inverse of the outer product estimate of the Hessian and are reported in brackets. $LogL$ is the value of the Log-likelihood function. R^2_R is the coefficient of determination when regressing the returns on V_t^{MIDAS} and $R^2_{\sigma^2}$ is the coefficient of determination when regressing the realized variance on V_t^{MIDAS}. p_{11} and p_{22} are the transition probabilities of staying in the first and second regime, respectively. $P(S_t = 1)$ is the unconditional probability of being in the high volatility regime.

Table 11: Explaining the regime probabilities $P(S_{t+1})$

	Slope of the yield curve$_{t+1}$	ΔV_{t+1}^{MIDAS}	R_{t+1}	ADS_{t+1}
Panel A: OLS regression				
1-week	0.002*	-0.019***	-0.094***	-0.003*
2-week	-0.001	0.004	-0.065***	-0.023***
Monthly	-0.004	0.002*	-0.045***	-0.022***
Quarterly	-0.021*	0.001*	-0.028***	-0.056***
Panel B: Logistic regression				
1-week	0.017	-0.516*	-6.636***	-0.322**
2-week	-0.143	0.089	-2.703***	-0.081
Monthly	0.013	0.050	-2.250***	-0.667*
Quarterly	-0.367	0.059*	-0.727**	-0.735*

Panel A reports the results of OLS regressions of the estimated smoothed probability of being in the first regime $P(S_{t+1})$ on the level of the slope of the yield curve, the changes in the MIDAS estimator of the conditional variance ΔV_{t+1}^{MIDAS}, the expected returns R_{t+1} and the level of the ADS index of business cycle conditions ADS_{t+1}. Panel B reports results of logistic regressions using a dummy variable as a dependent variable and the same set of explanatory variables. The dummy variable takes on a value of 1 if the smoothed probability of being in regime 1 is higher than 0.5 and 0 otherwise. The slope of the yield curve is defined as the difference between the yields on a 10-year Treasury bond and the yields on a 3-month Treasury bill. *, **, *** indicate significance at the 10% level, 5% level and 1% level, respectively. We only use the sub-sample 1964-2010 since we do not have data for the ADS index and the weekly slope of the yield curve for the whole full sample.

where:

$$q(z_t) = \frac{exp(\theta_1 + \theta_2 z_t)}{1 + exp(\theta_1 + \theta_2 z_t)}$$

and:

$$p(z_t) = \frac{exp(\theta_3 + \theta_4 z_t)}{1 + exp(\theta_3 + \theta_4 z_t)}$$

We use alternatively the slope of the yield curve ($Slope_{t+1}$), the dividend-price ratio ($(D/P)_{t+1}$), the lagged returns (R_t) and the realized covariance measure (Cov_{t+1}) calculated as the product between the changes in the ADS index and the returns as driving variables for the transition probabilities. All regressions are sampled at the monthly frequency.

Table 12 displays the results. First, the coefficients γ_1 and γ_2 are close to the estimates reported in Table 4: across all indicators, the risk-return relation is negative in the first regime, while it is positive in the second regime. None of the indicators enters significantly for explaining the transition probabilities of the first regime, whereas all indicators enter significantly at the 5% level for explaining the transition probabilities of the second regime (except for the slope of the yield curve for the full sample analysis and the dividend-price ratio for the sub-sample analysis).

Table 12 also reports a likelihood ratio test for testing the statistical significance of the time-varying transition probabilities. Under the null hypothesis of constant transition probabilities: $\theta_2 = \theta_4 = 0$. The null hypothesis of no time variation in the transition probabilities cannot be rejected at the 5% level when using the slope of the yield curve (full sample analysis) and the dividend-price ratio (sub-sample analysis). This provides mixed evidence for the use of time-varying transition probabilities for estimating the risk-return trade-off with regime switching, but overall confirms the robustness of the results we have obtained.

4.5 Out-of-sample forecasting exercise

In this section, we look at the forecasting performance of the MIDAS estimators for forecasting realized volatility. We use the MIDAS estimators from both the linear and regime-switching risk return relation and we take as benchmark a standard AR(1) model for realized variance following Ludvigson and Ng (2007). Unlike Welch and Goyal (2008) and Campbell and Thompson (2008) - who study the prediction of excess returns - we concentrate our analysis on the prediction of realized variance since the MIDAS approach is primarily designed for modeling the conditional variance.

The design of the out-of-sample forecasting exercise is the following. The first estimation sample goes from February 1929 to December 1969 so that we first forecast the realized volatility for January 1970. We then recursively expand the sample size until we reach the end of the sample December 2010. Therefore, the evaluation sample goes from January 1970 to December 2010. We concentrate our analysis on one-step-ahead forecasts. We forecast

Table 12: Monthly Regime-switching risk-return relation with time-varying transition probabilities

	μ_1 $(*10^2)$	μ_2 $(*10^2)$	γ_1	γ_2	θ_1	θ_2	θ_3	θ_4	$LogL$	$LRtest$	R^2_R	$R^2_{\sigma^2}$
Full sample analysis: February 1929 - December 2010												
$Slope_{t+1}$	-2.143	0.233	-0.374	0.062	-1.233	0.089	2.227	0.273	-2913.387	3.454	0.03%	54.03%
	[-0.786]	[0.709]	[-2.601]	[2.804]	[-0.841]	[0.086]	[5.327]	[1.392]		[0.178]		
$(D/P)_{t+1}$	-3.296	0.095	-0.310	0.069	-0.923	0.990	1.579	-3.278	-2909.457	11.314	0.02%	54.45%
	[-2.052]	[0.437]	[-3.116]	[3.750]	[-1.439]	[0.374]	[3.932]	[-2.857]		[0.004]		
R_t	-2.421	0.226	-0.378	0.060	-2.096	-0.219	2.658	0.895	-2910.712	8.804	0.04%	51.19%
	[-1.650]	[0.760]	[-4.512]	[3.054]	[-1.800]	[-0.543]	[7.092]	[3.013]		[0.012]		
Sub-sample analysis: February 1964 - December 2010												
$Slope_{t+1}$	-2.514	0.237	-0.249	0.042	-1.873	0.399	2.046	0.743	-1579.038	8.530	0.05%	53.37%
	[-0.462]	[0.408]	[-0.633]	[0.913]	[-1.268]	[0.380]	[3.653]	[2.522]		[0.012]		
$(D/P)_{t+1}$	-0.854	0.307	-0.405	0.032	-0.018	5.521	1.843	-2.504	-1581.855	2.896	0.05%	52.59%
	[-0.215]	[0.675]	[-1.348]	[0.954]	[-0.181]	[1.207]	[1.637]	[-0.949]		[0.235]		
R_t	-2.925	-0.011	-0.364	0.050	-70.819	100.379	3.511	1.145	-1580.194	6.218	0.13%	40.19%
	[-1.088]	[-0.262]	[-1.791]	[4.018]	[-0.331]	[0.331]	[6.161]	[2.877]		[0.045]		
Cov_{t+1}	-2.052	0.162	-0.349	0.037	-196.454	-39.320	3.678	1.299	-1577.947	10.712	0.06%	52.86%
	[-0.810]	[0.408]	[-2.319]	[1.307]	[-0.077]	[-0.076]	[6.057]	[2.583]		[0.005]		

The MIDAS estimator of the conditional variance is computed using 120 lags for the daily absolute returns, which are aggregated with the beta polynomial weight function. T-statistics are computed from the inverse of the outer product estimate of the Hessian and are reported in brackets. $LogL$ is the value of the log-likelihood function. $LRtest$ reports the value of a likelihood ratio test statistic. Under the null hypothesis of no time variation in the transition probabilities $\theta_2 = \theta_4 = 0$, p-values for the LR tests are reported in brackets. R^2_R is the coefficient of determination when regressing the returns on V_t^{MIDAS} and $R^2_{\sigma^2}$ is the coefficient of determination when regressing the realized variance on V_t^{MIDAS}.

1-week, 2-week, 3-week and monthly realized volatility, and compute relative mean squared forecast error (RMSE) and relative mean absolute forecast error (RMAE):

$$RMSE = \frac{\sum_{t=1}^{T}(V_{t+1|t}^{MIDAS} - RVAR_{t+1})^2}{\sum_{t=1}^{T}(V_{t+1|t}^{AR(1)} - RVAR_{t+1})^2}$$

$$RMAE = \frac{\sum_{t=1}^{T}|V_{t+1|t}^{MIDAS} - RVAR_{t+1}|}{\sum_{t=1}^{T}|V_{t+1|t}^{AR(1)} - RVAR_{t+1}|}$$

where $V_{t+1|t}^{MIDAS}$ is the one-step-ahead MIDAS forecast of the realized variance $RVAR_{t+1}$, and $V_{t+1|t}^{AR(1)}$ is the one-step-ahead forecast of the realized variance $RVAR_{t+1}$ from an AR(1) model. Table 13 presents the results. For monthly forecasts, the AR(1) model outperforms both MIDAS conditional forecasts. At the monthly horizon, the MIDAS forecasts from the linear risk-return relation are better than the MIDAS forecasts obtained from the regime switching risk-return relation. However, at the 1-week horizon, MIDAS forecasts are better than the forecasts from the AR(1) model. The MIDAS forecasts from the regime-switching risk-return relation are (slightly) better than the ones from the linear risk-return relation. This reasonably confirms the in-sample evidence since we found more evidence for regime switching at the 1-week frequency than at the monthly frequency (see Table 7).

Table 13: Forecasting Realized Volatility: one-step-ahead forecast

	Model	RMSE	RMAE
1-week	MIDAS (linear)	0.920	0.915
	MIDAS (MS)	0.915	0.913
2-week	MIDAS (linear)	1.078	1.015
	MIDAS (MS)	1.072	1.013
3-week	MIDAS (linear)	1.196	1.043
	MIDAS (MS)	1.185	1.039
Monthly	MIDAS (linear)	1.178	1.097
	MIDAS (MS)	1.292	1.128

This table reports the relative mean squared forecast error (RMSE) and the relative mean absolute forecast error (RMAE) for forecasting one-step-ahead realized volatility. The two competing models - MIDAS (linear) and MIDAS (MS) - are two MIDAS estimators of the conditional variance: one is estimated from a linear risk-return relation and the other one is estimated from a regime-switching risk-return relation. The benchmark model is a standard AR(1) model for realized volatility. The first estimation sample goes from February 1929 to December 1969 and is recursively expanded until we reach the end of the sample December 2010.

5 Conclusions

This paper provides evidence for time instability in the risk-return relation. We allow for regime changes in the risk-return relation through regime switching in the parameter entering before the conditional variance as well as the intercept of the risk-return relation. The conditional variance is modeled with a MIDAS estimator, which is less prone to misspecifications than GARCH models. We consider as dependent variable the US excess stock returns ranging from the weekly to the quarterly frequency and use two different estimation samples: (i) from February 1929 to December 2010 and (ii) from February 1964 to December 2010. We find strong statistical evidence for regime changes in the risk-return relation using the test recently introduced by Carrasco et al. (2013) for Markov-switching parameters.

In the first regime, we find that the risk-return relation is reversed. Conversely, in the second regime, we uncover the traditional positive risk-return relation. The regime probabilities for the first regime are associated with a decline in stock returns, an increase in volatility as well as a flattening of the yield curve, which is concomitant with flight-to-quality episodes. Our findings help to understand why the literature has reported conflicting results and are qualitatively close to the recent contribution of Rossi and Timmermann (2010). Our results are also robust to a wide range of modifications: (i) the inclusion of additional predictors, (ii) the use of Student-t rather than Normal innovations, (iii) the use of time-varying rather than constant transition probabilities, (iv) an asymmetric MIDAS estimator of the conditional variance.

One possible avenue for further research on this topic would be to study the dynamics of the risk-return trade-off using intra-daily returns. This could be done along the lines of the work by Rosenberg and Engle (2002) and Bakshi et al. (2003).

References

Arago, V., Floros, C., and Salvador, E. (2013). Re-examining the risk-return relationship in Europe: linear or non-linear trade-off? *mimeo.*

Aruoba, S. B., Diebold, F. X., and Scotti, C. (2009). Real-Time Measurement of Business Conditions. *Journal of Business & Economic Statistics*, 27(4):417–427.

Backus, D. K. and Gregory, A. W. (1993). Theoretical Relations between Risk Premiums and Conditional Variances. *Journal of Business & Economic Statistics*, 11(2):177–85.

Bakshi, G., Kapadia, N., and Madan, D. (2003). Stock Return Characteristics, Skew Laws, and the Differential Pricing of Individual Equity Options. *Review of Financial Studies*, 16(1):101–143.

Brandt, M. W. and Kang, Q. (2004). On the relationship between the conditional mean and volatility of stock returns: A latent VAR approach. *Journal of Financial Economics*, 72(2):217–257.

Campbell, J. Y. and Cochrane, J. (1999). Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior. *Journal of Political Economy*, 107(2):205–251.

Campbell, J. Y. and Thompson, S. B. (2008). Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average? *Review of Financial Studies*, 21(4):1509–1531.

Carrasco, M., Hu, L., and Ploberger, W. (2013). Optimal Test for Markov Switching Parameters. *mimeo.*

Filardo, A. J. (1994). Business-Cycle Phases and Their Transitional Dynamics. *Journal of Business & Economic Statistics*, 12(3):299–308.

Forsberg, L. and Ghysels, E. (2007). Why do absolute returns predict volatility so well? *Journal of Financial Econometrics*, 5(1):31–67.

French, K. R., Schwert, G. W., and Stambaugh, R. F. (1987). Expected stock returns and volatility. *Journal of Financial Economics*, 19(1):3–29.

Garcia, R. (1998). Asymptotic Null Distribution of the Likelihood Ratio Test in Markov Switching Models. *International Economic Review*, 39(3):763–88.

Ghysels, E., Plazzi, A., and Valkanov, R. (2013). The risk-return relationship and financial crises. Discussion Paper, UNC, University of Lugano and UCSD.

Ghysels, E., Santa-Clara, P., and Valkanov, R. (2005). There is a risk-return trade-off after all. *Journal of Financial Economics*, 76(3):509–548.

Ghysels, E., Santa-Clara, P., and Valkanov, R. (2006). Predicting volatility: getting the most out of return data sampled at different frequencies. *Journal of Econometrics*, 131(1-2):59–95.

Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993). On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *Journal of Finance*, 48(5):1779–1801.

Guérin, P. and Marcellino, M. (2013). Markov-switching MIDAS models. *Journal of Business and Economic Statistics*, 31(1):45–56.

Guo, H. and Whitelaw, R. F. (2006). Uncovering the Risk-Return Relation in the Stock Market. *Journal of Finance*, 61(3):1433–1463.

Hamilton, J. D. (1990). Analysis of time series subject to changes in regime. *Journal of Econometrics*, 45(1-2):39–70.

Hansen, B. E. (1992). The Likelihood Ratio Test under Nonstandard Conditions: Testing the Markov Switching Model of GNP. *Journal of Applied Econometrics*, 7(S):S61–82.

Kim, C.-J., Morley, J. C., and Nelson, C. R. (2004). Is There a Positive Relationship between Stock Market Volatility and the Equity Premium? *Journal of Money, Credit and Banking*, 36(3):339–60.

Lettau, M. and Ludvigson, S. (2001). Resurrecting the (C)CAPM: A Cross-Sectional Test When Risk Premia Are Time-Varying. *Journal of Political Economy*, 109(6):1238–1287.

Lettau, M. and Ludvigson, S. (2010). Measuring and Modeling Variation in the Risk-Return Trade-off. *Handbook of Financial Econometrics*, pages 617–690.

Ludvigson, S. C. and Ng, S. (2007). The empirical risk-return relation: A factor analysis approach. *Journal of Financial Economics*, 83(1):171–222.

Mayfield, E. (2004). Estimating the market risk premium. *Journal of Financial Economics*, 73(3):465–496.

Merton, R. C. (1973). An Intertemporal Capital Asset Pricing Model. *Econometrica*, 41(5):867–87.

Rosenberg, J. V. and Engle, R. F. (2002). Empirical pricing kernels. *Journal of Financial Economics*, 64(3):341–372.

Rossi, A. and Timmermann, A. (2010). What is the Shape of the Risk-Return Relation? *mimeo.*

Scruggs, J. T. (1998). Resolving the Puzzling Intertemporal Relation between the Market Risk Premium and Conditional Market Variance: A Two-Factor Approach. *Journal of Finance*, 53(2):575–603.

Smith, A., Naik, P. A., and Tsai, C.-L. (2006). Markov-switching model selection using Kullback-Leibler divergence. *Journal of Econometrics*, 134(2):553–577.

Tauchen, G. (2004). Stochastic Volatility in General Equilibrium. *mimeo.*

Welch, I. and Goyal, A. (2008). A Comprehensive Look at The Empirical Performance of Equity Premium Prediction. *Review of Financial Studies*, 21(4):1455–1508.

Whitelaw, R. F. (2000). Stock Market Risk and Return: An Equilibrium Approach. *Review of Financial Studies*, 13(3):521–47.

Appendix A: We describe here the Carrasco et al. (2013) test for Markov switching parameters.

Denote $l_{t,\theta}^{(1)}$ and $l_{t,\theta}^{(2)}$ the first and second derivatives of the log-likelihood function with respect to the regime switching parameters θ (where $\theta = (\mu, \gamma)$).[9]

Due to the presence of nuisance parameters β that are not identified under the null hypothesis of no Markov-switching, the Carrasco et al. (2013) test statistic for Markov-switching parameters TS can be constructed as a sup-type test, that is:

$$\sup TS = \sup \frac{1}{2}\left(\max\left(0, \frac{\Gamma_T}{\sqrt{\hat{\epsilon}^{*\prime}\hat{\epsilon}^{*}}}\right)\right)^2$$

where

$$\Gamma_T = \frac{1}{2\sqrt{T}}\sum_{t=1}^{T}\gamma_t(\beta) \quad ,$$

$$\gamma_t(\beta) = tr\left(\left(l_{t,\theta}^{(2)} + l_{t,\theta}^{(1)} l_{t,\theta}^{(1)\prime}\right)E[\eta_t\eta_t^{\prime}]\right) + 2\sum_{s<t} tr\left(l_{t,\theta}^{(1)} l_{s,\theta}^{(1)\prime} E[\eta_t\eta_t^{\prime}]\right)$$

$$\hat{\epsilon}^{*} = \frac{\hat{\epsilon}}{\sqrt{T}}$$

and $\hat{\epsilon}$ is the vector of residuals from the OLS regression of $\frac{1}{2}\gamma_t(\beta)$ on the entire vector of derivatives and η_t is the latent variable.

We find the maximum value of TS using a fixed range of values for $\rho \in [-0.98, 0.98]$ with increment 0.01.

We compute critical values with bootstrapping techniques. We first generate M data series using the maximum likelihood estimates as true parameter values such that:

$$y_t^{(m)} \sim N(\hat{\mu} + \hat{\gamma}V_t^{MIDAS}, V_t^{MIDAS})$$

where m is the m^{th} sample. We then estimate each of the M samples with maximum likelihood and compute the test statistic by maximizing $TS^{(m)}$ over a fixed range of values for $\rho \in [-0.98, 0.98]$. The 5% bootstrapped critical value is then calculated as the 95^{th} percentile of the distribution of the M test statistics $TS^{(m)}$.

[9] Note that here we kept the MIDAS parameters κ_1 and κ_2 constant since the first derivatives with respect to these parameters is often zero, which is problematic when we regress $\gamma_t(\beta)$ on the vector of derivatives.

Table **Appendix B:** Comparison of linear, 2-regime and 3-regime models for the risk-return trade-off

		$M = 1$	$M = 2$	$M = 2$ switch in κ_1 and κ_2	$M = 3$	at least one $\gamma(S_{t+1})$ is not significant when $M = 3$
				1929-2010		
Quarterly	LogL	-1227.190	-1174.979	-1178.063	-1151.743	YES
	SIC	2464.439	2370.074	2381.272	**2338.690**	
Monthly	LogL	-2987.476	-2915.114	-2926.598	-2887.108	YES
	SIC	5986.922	5854.169	5883.121	**5816.111**	
2-week	LogL	-5562.106	-5407.312	-5421.151	-5362.068	NO
	SIC	11137.531	10841.261	10875.599	**10770.749**	
1-week	LogL	-9532.267	-9271.982	-9304.984	-9160.942	NO
	SIC	19079.056	18573.009	18646.277	**18372.714**	
				1964-2010		
Quarterly	LogL	-642.927	-632.893	-633.033	-627.604	YES
	SIC	1294.942	**1283.960**	1288.785	1287.013	
Monthly	LogL	-1593.271	-1583.303	-1580.036	-1579.456	YES
	SIC	3197.544	3188.610	**3187.578**	3197.419	
2-week	LogL	-2992.930	-2951.106	-2951.495	-2939.819	NO
	SIC	5998.212	5926.914	5933.868	**5922.866**	
1-week	LogL	-5133.369	-5070.936	-5084.024	-5042.182	NO
	SIC	10280.293	10168.9819	10201.936	**10131.809**	

LogL is the value of the log-likelihood function, SIC is the Schwarz Information Criterion. The fifth column reports the LogL and SIC for the models with switches in μ, γ and the MIDAS parameters κ_1 and κ_2 so that the weight function aggregating the lagged daily returns also changes across regime. The last column indicates whether at least one parameter $\gamma(S_{t+1})$ entering before the conditional variance is not significant at the 10% confidence level when a 3-regime model is estimated. Entries in bold outline the model with the lowest SIC for each regression.

Appendix C: Additional robustness checks

We report below additional estimation results of the risk-return trade-off with MIDAS conditional variance and regime switching risk-return relation:

- We stop the estimation in December 2000 for both the full sample and sub-sample analyses following Ghysels et al. (2005) and Mayfield (2004) so that we do not include the 2007-2009 financial crisis in the estimation sample.

- We consider estimates of the risk-return trade-off at the weekly frequency for two short estimation samples 2001-2010 and 2007-2010.

- We use as a proxy for stock returns CRSP data rather than the S&P 500 composite portfolio.

- We use a model with a NBER dummy variable entering before the estimate of the conditional variance. The NBER dummy variable takes a value of 1 if the US economy is in recession and a value of 0 is the US economy is in expansion according to the NBER business cycle dating committee.

- We use a credit spread (defined as the difference between the yields on the Moody's Corporate bond (all industries - BAA) and the yields on the 10-year US Treasury bond) instead of the slope of the yield curve as an additional predictor in the risk-return relation.

- We use the realized variance instead of a MIDAS estimator for the conditional variance.

First, the results shown in Panel A of Table C1 are consistent with the results reported previously so that the choice of the estimation window does not seem to drive our results. Second, using CRSP value-weighted portfolio as a proxy for stock market returns yield comparable results to those obtained using the S&P500 composite portfolio index. Third, estimating the risk-return relation with a NBER dummy variable entering before the estimate of the conditional variance to take into account the fluctuations of the business cycle also yields an inverted risk return relation during US recessions, while the risk-return relation remains positive during US expansions. Fourth, using the lagged realized variance as a proxy for the conditional variance (instead of a MIDAS estimator) does not affect qualitatively the results. Finally, using the credit spread as an additional predictor in the risk-return relation yields similar results than when using the slope of the yield curve.

Table C1: The MIDAS estimates of the risk-return trade-off with regime switching, robustness checks

		p_{11}	p_{22}	μ_1 $(*10^2)$	μ_2 $(*10^2)$	γ_1	γ_2	$LogL$	R^2_R	$R^2_{\sigma^2}$	$P(S_t = 1)$
Panel A: Regime-switching risk-return relation: $R_{t+1} \sim N(\mu(S_{t+1}) + \gamma(S_{t+1})V_t^{MIDAS}, V_t^{MIDAS})$											
1929:02-2000:12	Monthly	0.257	0.926	-2.184	0.034	-0.373	0.087	-2569.827	0.19%	53.54%	9.09%
		[3.455]	[38.901]	[-1.415]	[0.264]	[-3.578]	[5.379]				
1964:02-2000:12	Monthly	0.249	0.961	0.525	-0.571	-0.624	0.108	-1238.063	2.64%	27.35%	4.98%
		[1.762]	[54.051]	[0.088]	[-1.239]	[-1.538]	[3.086]				
2001:02-2010:12	Weekly	0.423	0.934	2.927	-0.183	-1.234	0.104	-1135.468	0.01%	61.94%	10.30%
		[2.998]	[27.615]	[2.812]	[-0.963]	[-4.912]	[2.056]				
2007:02-2010:12	Weekly	0.349	0.842	-1.431	0.283	-0.393	0.114	-497.281	0.18%	63.05%	19.54%
		[2.752]	[13.058]	[-1.588]	[0.779]	[-2.621]	[1.400]				
Panel B: Regime-switching risk-return relation: $R_{t+1} \sim N(\mu(S_{t+1}) + \gamma(S_{t+1})V_t^{MIDAS}, V_t^{MIDAS})$ *with CRSP data*											
1929:02-2000:12	Monthly	0.274	0.914	-2.393	0.378	-0.323	0.104	-2581.244	0.22%	55.67%	10.63 %
		[3.618]	[32.338]	[-1.358]	[0.954]	[-2.695]	[3.124]				
1964:02-2000:12	Monthly	0.046	0.966	14.440	0.027	-2.079	0.092	-1253.604	2.04%	27.06%	3.49%
		[0.399]	[74.226]	[2.180]	[0.229]	[-3.759]	[5.403]				
Panel C: risk-return relation with a NBER dummy variable											
1929:02-2010:12	Monthly	-	-	0.190	-0.410	-0.045	0.080	-2974.156	0.01%	58.70%	-
				[0.297]	[-1.380]	[-1.498]	[3.496]				
1964:02-2010:12	Monthly	-	-	-0.076	-1.693	0.050	0.041	-1588.072	0.04%	52.30%	-
				[-0.119]	[-1.165]	[1.046]	[0.599]				

The MIDAS estimator of the conditional variance is computed using 120 lags for the daily absolute returns, which are aggregated with the beta polynomial weight function. T-statistics are computed from the inverse of the outer product estimate of the Hessian and are reported in brackets. $LogL$ is the value of the log likelihood function. R^2_R is the coefficient of determination when regressing the returns on the V_t^{MIDAS} and $R^2_{\sigma^2}$ is the coefficient of determination when regressing the realized variance on V_t^{MIDAS}. p_{11} and p_{22} are the transition probabilities of staying in the first and second regime, respectively. $P(S_t = 1)$ is the unconditional probability of being in the high volatility regime. Panel A reports estimation results using the S&P500 composite portfolio, whereas Panel B reports estimation results using the CRSP value weighted portfolio. Note that we do not have data for the CRSP value weighted portfolio for the 2001-2010 period. Panel C reports results when using a NBER dummy variable entering before the conditional variance instead of a regime switching parameter.

Table C2: The MIDAS estimates of the risk-return trade-off with regime switching, robustness checks

	p_{11}	p_{22}	μ_1 $(*10^2)$	μ_2 $(*10^2)$	γ_1	γ_2	R_t	$(D/P)_t$	$Credit_t$	Cov_t	$LogL$	R^2_R	$R^2_{\sigma^2}$	$P(S_t=1)$
Full sample analysis: February 1929 - December 2010 - Using lagged realized variance instead of a MIDAS estimator for the conditional variance														
Quarterly	0.635	0.905	-11.687	-3.786	0.024	0.192	-	-	-	-	-1187.359	0.27%	56.54%	20.73%
	[6.288]	[24.020]	[-2.076]	[-2.880]	[0.450]	[4.652]								
Monthly	0.225	0.919	-3.083	0.282	-0.275	0.064	-	-	-	-	-2922.956	0.01%	58.75%	9.46%
	[2.597]	[30.470]	[-1.884]	[1.045]	[-2.755]	[2.700]								
Full sample analysis: February 1929 - December 2010 - Using the credit default spread instead of the slope of the yield curve														
Quarterly	0.388	0.902	-11.396	-0.233	-0.057	0.102	0.098	7.101	1.188	-	-1165.614	2.35%	51.15%	13.77%
	[3.352]	[32.080]	[-3.106]	[-0.212]	[-1.147]	[3.598]	[1.678]	[3.292]	[1.743]					
Monthly	0.271	0.932	-1.417	0.942	-0.412	0.039	-0.025	2.229	0.295	-	-2907.666	0.04%	51.52%	8.56%
	[3.606]	[47.060]	[-0.907]	[2.076]	[-4.050]	[1.418]	[-0.709]	[2.907]	[1.283]	-				
Sub-sample analysis: February 1964 - December 2010 - Using the credit default spread instead of the slope of the yield curve														
Quarterly	$0.001^{(a)}$	0.913	12.399	0.427	-0.569	0.066	0.177	4.902	0.447	2.776	-626.305	0.33%	49.57%	8.04%
		[19.535]	[1.371]	[0.194]	[-2.390]	[1.703]	[2.119]	[1.434]	[0.479]	[1.492]				
Monthly	0.296	0.933	-1.340	0.695	-0.257	0.035	-0.057	1.642	-0.057	0.973	-1578.787	0.05%	52.88%	8.68%
	[1.393]	[21.093]	[-0.531]	[1.047]	[-1.359]	[0.716]	[-0.996]	[1.573]	[-0.996]	[0.891]				

The MIDAS estimator of the conditional variance is calculated using 120 lags for the daily absolute returns, which are aggregated with the beta polynomial weight function. T-statistics are calculated from the inverse of the outer product estimate of the Hessian and are reported in brackets. $LogL$ is the value of the log likelihood function. R^2_R is the coefficient of determination when regressing the returns on V_t^{MIDAS} and $R^2_{\sigma^2}$ is the coefficient of determination when regressing the realized variance on V_t^{MIDAS}. p_{11} and p_{22} are the transition probabilities of staying in the first and second regime, respectively. $P(S_t=1)$ is the unconditional probability of being in the first regime. [(a)] In that case, the transition probability for regime 1 hit the lower bound that was imposed to respect the properties of a Markov chain.