



Department of Economics

Essays in Applied Time Series Econometrics

Pierre Guérin

Thesis submitted for assessment with a view to obtaining the degree of
Doctor of Economics of the European University Institute

Florence, September, 2011

EUROPEAN UNIVERSITY INSTITUTE
Department of Economics

Essays in Applied Time Series Econometrics

Pierre Guérin

Thesis submitted for assessment with a view to obtaining the degree of
Doctor of Economics of the European University Institute

Jury Members:

Prof. Massimiliano Marcellino, EUI, Supervisor
Prof. Helmut Lütkepohl, EUI
Prof. Monica Billio, Università Ca' Foscari di Venezia
Prof. Eric Ghysels, University of North Carolina, Chapel Hill

© 2011, Pierre Guérin

No part of this thesis may be copied, reproduced or
transmitted without prior permission of the author

Acknowledgements

First and foremost, I would like to thank my supervisor, Massimiliano Marcellino, who has provided me with support and guidance in numerous ways over the last four years. I would also like to express my gratitude to my second supervisor, Helmut Lutkepohl, for his useful comments and suggestions on the different versions of my thesis chapters. I am also grateful to Monica Billio and Eric Ghysels for accepting to be members of my thesis committee.

This thesis also benefited from comments of seminar participants. In particular, I am especially thankful to my fellow students from the EUI Time Series working group.

The French ministries of Foreign Affairs and Education provided me with financial support for this thesis, which is gratefully acknowledged. In addition, I also benefited from the hospitality of the European Central Bank during my PhD internship in the summer of 2010. Also, I am grateful to the staff members of the Economics Department for providing students with a very good environment to work.

Finally, I also owe a special thanks to my family and friends from the EUI and beyond that contributed to make my life in Florence more pleasant and gave me support and encouragement.

Contents

Acknowledgements	i
List of Figures	v
List of Tables	viii
Introduction	1
1 Predicting Business Cycle Regimes Using Markov-switching Models with Time-Varying Transition Probabilities	3
1.1 Introduction	4
1.2 Modelling the asymmetries of the business-cycle	6
1.2.1 Markov-switching models with time-varying transition probabilities (TVTP)	6
1.2.2 Comparison with other modelling approaches	7
1.2.3 Estimation and model selection	8
1.2.4 Selecting the transition variable	10
1.3 In sample results	12
1.4 Out-of-sample results	15
1.5 Case study: an application of the model	21
1.6 Conclusions	24
1.7 Appendix Chapter 1	26
2 Markov-switching MIDAS models	33
2.1 Introduction	34
2.2 Markov-switching MIDAS	36
2.2.1 MIDAS approach	36
2.2.1.1 Basic MIDAS	36

2.2.1.2	Autoregressive MIDAS	36
2.2.2	Markov-switching MIDAS	37
2.2.2.1	The model	37
2.2.2.2	Estimation and model selection	38
2.3	Monte Carlo experiments	39
2.3.1	In-sample estimates	39
2.3.2	Forecasting exercise	41
2.4	An application to the prediction of quarterly GDP	44
2.4.1	Prediction of the US GDP	44
2.4.1.1	In-sample results	44
2.4.1.2	Design of the real-time forecasting exercise	47
2.4.1.3	Out-of-sample results	50
2.4.2	Prediction of the UK GDP	54
2.4.2.1	In-sample results	54
2.4.2.2	Out-of-sample results	58
2.5	Conclusions	59
2.6	Appendix Chapter 2	63
2.6.1	Estimation algorithm	63
2.6.2	Additional Tables	64
3	Trend-cycle decomposition of output and euro area inflation forecasts: a real-time approach based on model combination	71
3.1	Introduction	72
3.2	Trend-cycle decomposition of output	74
3.2.1	Extension to regime changes in the slope of the trend	74
3.2.2	Extension to use auxiliary information	76
3.2.3	Extension to incorporate a time-varying Phillips curve	76
3.3	In-sample estimates for the euro area	78
3.3.1	Time-varying Phillips curve	79
3.3.2	Univariate and bivariate trend-cycle decomposition of euro area real GDP	80
3.3.3	Comparison of the estimated output gaps with the model-averaged measures	83
3.4	Do real-time estimates of the output gap improve inflation forecasts?	85
3.4.1	Real-time estimates of the output gap	85

3.4.2	Inflation forecasts	86
3.5	Conclusions	91
3.6	Appendix Chapter 3	93
3.6.1	State-space representation of the original Clark (1987) model of trend-cycle decomposition of output	93
3.6.2	Kalman filter, Hamilton filter and Kim and Nelson filtering procedure	93
3.6.3	Clark and Mc Craken tests for comparing forecasting performance	94
3.6.4	Additional Tables and Figures	96
4	Switches in the risk-return trade-off	105
4.1	Introduction	106
4.2	The risk-return relation with regime switching	108
4.3	Data and empirical results	110
4.3.1	Data	110
4.3.2	MIDAS and GARCH estimates of the risk-return relation	111
4.3.3	MIDAS estimates of the regime switching risk-return relation	116
4.3.4	Testing for Markov-switching	124
4.4	Sensitivity analysis	125
4.4.1	Additional predictors in the risk-return relation	125
4.4.2	Controlling for asymmetries in stock returns	128
4.4.3	The risk-return trade-off with Student-t innovations	129
4.4.4	Time-varying transition probabilities	134
4.5	Conclusions	136
4.6	Appendix Chapter 4	137
4.6.1	The Carrasco et al. (2009) test for Markov switching parameters	137
4.6.2	Additional robustness checks	140
	Bibliography	143

List of Figures

1.1	INDUSTRIAL PRODUCTION, SLOPE OF THE YIELD CURVE AND US RECES- SIONS 1970:01-2008:11	5
1.2	IN-SAMPLE PROBABILITY OF RECESSION, MS MODEL WITH TVTP AND FTP	18
1.3	ONE-MONTH-AHEAD PREDICTED PROBABILITY OF RECESSION, MS MODEL WITH TVTP AND FTP	22
1.4	THREE-MONTH-AHEAD PREDICTED PROBABILITY OF RECESSION, MS MODEL WITH TVTP AND FTP	23
1.5	ONE-MONTH-AHEAD PREDICTED PROBABILITY OF RECESSION, REAL-TIME DATA	25
1.6	THREE-MONTH-AHEAD MONTE CARLO PREDICTED PROBABILITY OF RECES- SION	26
2.1	MS-MIDAS, QUARTERLY GDP AND MONTHLY SLOPE OF THE YIELD CURVE, 1959:Q1-2009:Q4	47
2.2	WEIGHTS OF THE AR-MIDAS (LHS) AND MSHAR(3)-MIDAS (RHS), EXPONEN- TIAL LAG POLYNOMIAL	48
2.3	ESTIMATED PROBABILITY OF RECESSION, REAL-TIME DATA, 1997:Q4-2010:Q2	56
2.4	MSIAR(2)-MIDAS QUARTERLY GDP AND MONTHLY SLOPE OF THE YIELD CURVE, 1976:Q1-2010:Q1	57
2.5	ESTIMATED PROBABILITY OF RECESSION, REAL-TIME DATA, 2003:Q4-2010:Q1	63
3.1	MODEL DECOMPOSITION OF EURO AREA INFLATION (ANNUAL GROWTH, %)	81
3.2	SMOOTHED PROBABILITY OF BEING IN THE FIRST REGIME	82
3.3	ESTIMATES OF THE EURO AREA OUTPUT GAP DERIVED FROM THE ESTI- MATED UNOBSERVED COMPONENTS MODELS	84
3.4	MODEL AVERAGED MEASURES OF THE NINE OUTPUT GAP ESTIMATES . . .	85

3.5	ESTIMATES OF OUTPUT GAP IN REAL TIME: RANGE OF REVISIONS, (MAX-MIN)	87
3.6	EURO AREA OUTPUT GAP, REAL-TIME DATA, MODEL-AVERAGED MEASURES WITH EQUAL WEIGHTS, 1974Q1-2010Q4	96
3.7	EURO AREA OUTPUT GAP, REAL-TIME DATA, MODEL-AVERAGED MEASURES COMPUTED AS THE MEDIAN OF THE INDIVIDUAL OUTPUT GAPS, 1974Q1-2010Q4	97
3.8	EURO AREA OUTPUT GAP, REAL-TIME DATA, MODEL-AVERAGED MEASURES WITH TIME-VARYING WEIGHTS, 1974Q1-2010Q4	98
3.9	ACTUAL INFLATION AND INFLATION EXPECTATIONS FROM UC AND MUC MODELS	99
4.1	MONTHLY EXCESS STOCK RETURNS 1929:02-2010:12	111
4.2	MIDAS AND REALIZED VARIANCES 1929:02-2010:12	114
4.3	ABSGARCH AND REALIZED VARIANCES 1929:03-2010:12	116
4.4	MIDAS AND REALIZED VARIANCES 1929:02-2010:12	117
4.5	WEIGHTS FOR THE MIDAS ESTIMATOR OF THE CONDITIONAL VARIANCE . .	118
4.6	MONTHLY RETURNS AND PROBABILITY OF BEING IN THE FIRST REGIME 1929:02-2010:12	121
4.7	SCATTERPLOTS OF THE MONTHLY VARIANCES 1929:02-2010:12	123
4.8	WEIGHTS FOR THE ASYMIDAS ESTIMATOR OF THE CONDITIONAL VARIANCE	129

List of Tables

1.1	List of transition variables	11
1.2	Estimation results, Variable by themselves, In-sample estimates	16
1.3	Estimation results, Variables in combination with $SPREAD_{t-k}$, In-sample estimates	17
1.4	RMSE, Variables by themselves, Out-of-sample results	20
1.5	RMSE, Variables in combination with $SPREAD_{t-1}$, Out-of-sample results .	21
1.6	Data appendix	27
1.7	Estimation results, Variable by themselves, In-sample estimates	28
1.8	Estimation results, Variable X_{t-k} in combination with $SPREAD_{t-k}$, In-sample estimates	29
1.9	RMSE, Variables by themselves, Out-of-sample results	30
1.10	RMSE, Variables in combination with $SPREAD_{t-1}$, Out-of-sample results .	31
2.1	Monte Carlo Results, In-Sample Estimates	42
2.2	Monte Carlo Results: Forecasting exercise	45
2.3	Forecasting scheme for $Q1$	49
2.4	Relative Mean Squared Forecast error for forecasting US GDP growth 1998:Q1-2009:Q4	52
2.5	Quadratic Probability Score for forecasting US business cycle regimes 1998:Q1-2009:Q4	53
2.6	Log Probability Score for forecasting US business cycle regimes 1998:Q1-2009:Q4	55
2.7	Relative Mean Squared Forecast error for forecasting UK GDP growth 2004:Q1-2010:Q1	60
2.8	Quadratic Probability Score for forecasting UK business cycle regimes 2004:Q1-2010:Q1	61
2.9	Log Probability Score for forecasting UK business cycle regimes 2004:Q1-2010:Q1	62

2.10	In-sample results, Quarterly US GDP growth rate 1959:Q1-2009:Q4	65
2.11	In-sample results, Quarterly UK GDP growth rate 1975:Q1-2010:Q1	66
2.12	US Estimated Quarterly Probability of Recession updated on a monthly basis	67
2.13	UK Estimated Quarterly Probability of Recession updated on a monthly basis	68
2.14	Classification of the models	69
3.1	Forecasting comparison exercise: results for inflation (HICP excluding energy)	90
3.2	Time-varying Phillips curve	100
3.3	Univariate model	101
3.4	Bivariate model (GDP and rate of capacity utilization)	102
3.5	Bivariate model (GDP and inflation)	103
3.6	Forecasting comparison exercise: results for the change in inflation (HICP excluding energy)	104
4.1	Summary statistics for monthly US excess stock returns	111
4.2	Linear risk return relation: $R_{t+1} \sim N(\mu + \gamma V_t^{MIDAS}, V_t^{MIDAS})$	113
4.3	Monthly GARCH estimates of the risk-return relation	115
4.4	Regime-switching risk-return relation: $R_{t+1} \sim N(\mu + \gamma(S_t)V_t^{MIDAS}, V_t^{MIDAS})$	119
4.5	Explaining the regime probabilities $P(S_{t+1})$, Estimation sample: 1964:2010 .	120
4.6	Comparison of the monthly variance processes	122
4.7	Tests of regime switching in the risk-return relation	125
4.8	Regime-switching risk-return relation with additional predictors	127
4.9	Monthly estimates of the risk-return trade-off with Asymmetric MIDAS esti- mators of the conditional variance	130
4.10	Regime-switching risk-return relation with Student-t innovations	132
4.11	Explaining the regime probabilities $P(S_{t+1})$, Estimation sample: 1964:2010 .	133
4.12	Monthly Regime-switching risk-return relation with time-varying transition probabilities	135
4.13	Comparison of linear, 2-regime and 3-regime models for the risk-return trade-off	139
4.14	The MIDAS estimates of the risk-return trade-off with regime-switching, sub- sample analyses	141

Introduction

In the first chapter of this thesis, I estimate Markov-switching models with time-varying transition probabilities to predict the US business cycle regimes. In particular, I evaluate the predictive power of real and financial indicators and find that the slope of the yield curve turns out to be the most reliable indicator for regime predictions. This first chapter paves the way for the next two chapters of this thesis that also use models with Markov-switching for analysing the business cycle.

The second chapter (a joint work with Massimiliano Marcellino) combines the Markov-switching model with the MIXed DATA Sampling (MIDAS) model. This new model uses information from variables sampled at different frequencies. We first show in a Monte-Carlo experiment that our estimation method yields accurate estimates. We then apply this new model to the prediction of both the business cycle regimes and GDP growth for the US and the UK. We find that the use of high frequency information and parameter switching performs better than using each of these two features separately.

In the third chapter (a joint work with Laurent Maurin and Matthias Mohr), we estimate nine different models of the output gap (univariate, multivariate, linear and non-linear) and compute model-averaged estimates of the output gap. We find some evidence for changes in the slope of the trend of the Euro area output for few periods in 1974 and 2009. Moreover, our model-averages measures of the output gap reduce the uncertainty associated with the output gap estimates and soften the impact of data revisions. We then evaluate the forecasting performance of our output gap estimates for inflation and find that the output gap estimates improve on the forecasting performance of standard AR benchmarks for inflation although the inflation forecasts based on the output gap estimates exhibit a poor forecasting performance since 2008.

The last chapter of this thesis (a joint work with Eric Ghysels and Massimiliano Marcellino) is an empirical evaluation of the risk-return relation. We use a MIDAS estimator of the conditional variance and model regime changes in the parameter entering before the conditional variance. We find evidence for a reversed risk-return relation in periods of high volatility, while we uncover the traditional positive risk-return relation in periods of low volatility. In particular, the high volatility regime is interpreted as a flight-to-quality regime. This finding is robust to a large range of specifications.

Chapter 1

Predicting Business Cycle Regimes Using Markov-switching Models with Time-Varying Transition Probabilities

Abstract

This article examines the usefulness of various economic and financial variables for predicting the US business cycle regimes. To this end, I estimate Markov-switching models with time-varying transition probabilities and report both in-sample and out-of-sample results. The out-of-sample analysis produces regime forecasts one, three and six months ahead, which are then compared with the NBER business cycle reference dates. Variables are assessed by themselves and in combination with the slope of the yield curve, which turns out to be the most reliable indicator.

Keywords: Business cycles, Markov-switching models, Regime prediction, Yield curve.

JEL Classification Code: E37, C53.

1.1 Introduction

Burns and Mitchell (1946) established two defining characteristics of the business cycle: comovement among economic variables through the cycle and non linearity in the evolution of the business cycle. To this extent, Markov-switching models (MS models hereafter) can precisely describe the nature of business cycle movements. Indeed, an important appeal of MS models is their ability to account for the accumulating evidence that business cycles are asymmetric.

Predicting economic activity is a tricky task. The main body of the literature has focused on quantitative measures of future economic activity. However, studies about the identification and prediction of the state of the economy have gained attention over the last decade. For example, Estrella and Mishkin (1998) estimate a probit model for the US business cycle regimes, while Birchenhall et al. (1999) use a logit specification. Kauppi and Saikkonen (2008) and Nyberg (2010) introduce dynamics into these models and find that this improves the out-of-sample performance of this class of models. These papers use the widely accepted National Bureau of Economic Research (NBER) datation of expansions and recessions as dependent variable. However, one major drawback to the use of the NBER datation for predicting economic activity is the delay for the publication of the NBER Business Cycle Dating Committee decisions. The announcements of turning points may indeed be published up to twenty months after the turning point has actually occurred. In this respect, MS models have a strong advantage over probit and logit models for forecasting purposes since MS models endogenously determine the probabilities of being in a given state of the economy. More recently, Aruoba et al. (2009) and Camacho and Perez-Quiros (2010) estimate large state-space models to extract a business cycle indicator. This approach is conceptually attractive but it is often computationally cumbersome.

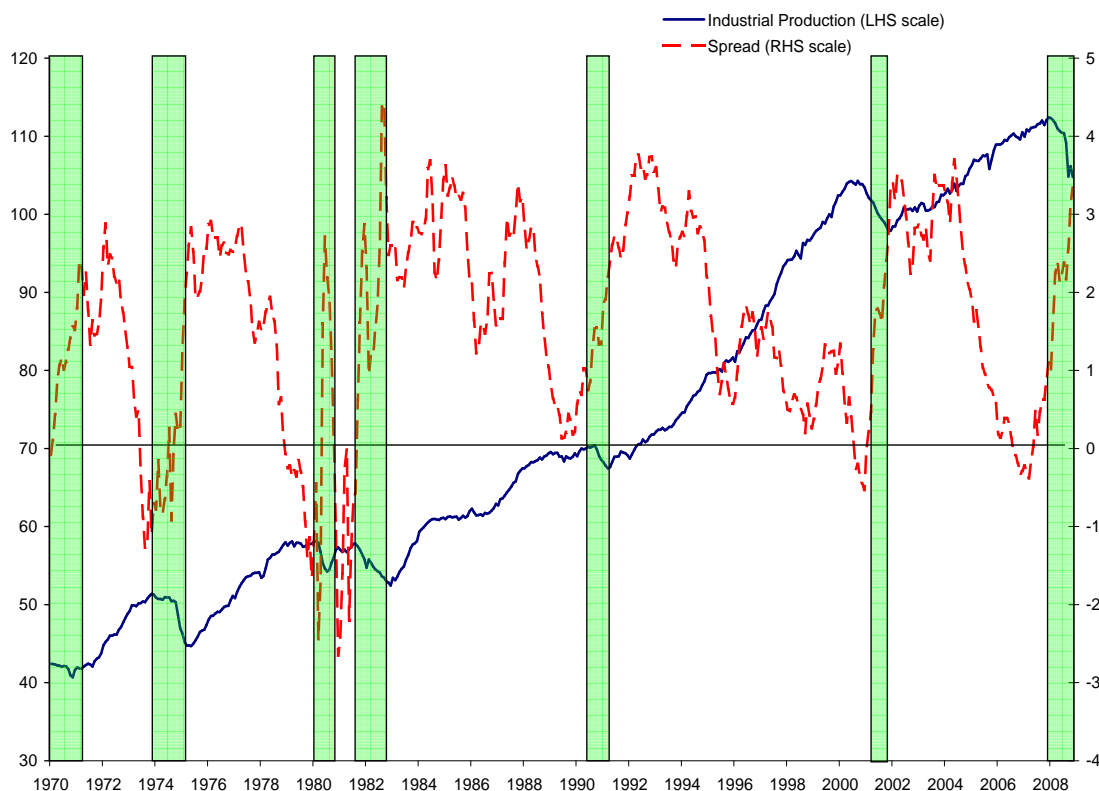
In this paper, I examine the forecasting performance of various economic and financial variables using MS models with time-varying transition probabilities. In particular, the variables with potential interest that are considered are monetary aggregates, interest rates, inflation, exchange rate and an index of leading indicators. I look at the predictive content of these variables by themselves and in combination with the slope of the yield curve, which emerges as the most reliable indicator.

The out-of-sample performance for the economic and financial variables under scrutiny is also evaluated. I produce regime forecasts one, three and six months ahead. I find that stock prices and the Conference Board index of leading indicators perform very well for one-month-ahead forecasts. However, the slope of the yield curve proves to be the best choice for

predictions with one, three and six months horizon. The literature indeed presents evidence for the good performance of the slope of the yield curve as a predictor for US recessions (see e.g. Estrella and Mishkin (1998) and Rudebusch and Williams (2009)). Interestingly, Figure 1.1 shows that the yield curve has become inverted prior to six out of the seven recessions of the post-1970 era. Finally, my results also tend to indicate that combining variables to the slope of the yield curve does not help for out-of-sample predictions.

In the following section, I describe the MS model with time-varying transition probabilities and its advantages over the competing models. I also discuss the strategy to select the MS model that obtains the best fit. In the third section, I look at the in-sample performance of the selected model, while Section 1.4 presents the out-of-sample results. Section 1.5 carries out an application to the 2007-2009 recession using real-time data. Section 1.6 concludes.

Figure 1.1: INDUSTRIAL PRODUCTION, SLOPE OF THE YIELD CURVE AND US RECESSIONS 1970:01-2008:11



1.2 Modelling the asymmetries of the business-cycle

1.2.1 Markov-switching models with time-varying transition probabilities (TVTP)

Markov-switching models have been introduced in Time Series analysis by Hamilton (1989). The basic idea behind this class of models is that the parameters of the underlying D.G.P. depends on the unobservable regime variable S_t , which represents the probability of being in a different state of the world. Krolzig (1997) carefully details the different sorts of Markov-switching models and extends the univariate specification to a multivariate framework. The Hamilton model (1989) of the business cycle is a fourth-order model fitted to the quarterly change in the log of the US real GNP for $t=1951:II$ to $1984:IV$:

$$y_t - \mu(S_t) = \phi_1(y_{t-1} - \mu(S_{t-1})) + \dots + \phi_4(y_{t-4} - \mu(S_{t-4})) + \varepsilon_t \quad (1.1)$$

where the conditional mean switches between two regimes and the variance is constant across the two regimes $\varepsilon_t \sim NID(0, \sigma^2)$.

According to Krolzig notation (1997), this is a Markov-switching model with a switch in the mean: MS(M) model. One could also use Markov-switching model with a switch in the intercept (MS(I) model):

$$y_t = \mu(S_t) + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \varepsilon_t \quad (1.2)$$

If there is a switch in the mean (MS(M) model), there will be an immediate jump of the dependent variable onto its new level, whereas the specification with a switch in the intercept (MS(I) model) implies a gradual adjustment of the dependent variable toward its new level.

Besides, one can also allow the variance of the disturbances to change across regimes. This proves to be useful in light of the debate about the changing behavior of the US monetary policy. Sims and Zha (2006) argue that the main source of fluctuation in the US monetary policy is due to time variation in disturbance variances, while Cogley and Sargent (2005) lean in favor of a changing behavior of the US monetary authorities.

The regime generating process is an ergodic Markov-chain with a finite number of states $S_t = 1, \dots, M$ defined by the following transition probabilities:

$$p_{ij} = \Pr(S_{t+1} = j | S_t = i) \quad (1.3)$$

$$\sum_{j=1}^M p_{ij} = 1 \quad \forall i, j \in (1, \dots, M) \quad (1.4)$$

However, the assumption of constant transition probabilities can be too restrictive. Diebold et al. (1993) and Filardo (1994) use time-varying transition probabilities and model them as logistic functions of economic variables, whereas Kim et al. (2008) model them as probit functions. In contrast, Durland and McCurdy (1994) use duration dependent transition probabilities where the probability of staying in a recession can decrease with the length of the recession.

Following Filardo (1994), I use time-varying transition probabilities (TVTP hereafter) modelled as a logistic function in order to bound the probabilities between 0 and 1. In the two-regime case, information about transition probabilities can be summarized in the matrix of transition probabilities \mathbf{P} . For simplicity of the presentation, I label state 0 as the recession state and state 1 as the boom state.

$$\mathbf{P} = \begin{bmatrix} p_t^{00} = q(z_t) & p_t^{01} = 1 - p(z_t) \\ p_t^{10} = 1 - q(z_t) & p_t^{11} = p(z_t) \end{bmatrix} \quad (1.5)$$

where:

$$q(z_t) = \exp(q_0 + \mathbf{z}_{t-d}\mathbf{q}) / (1 + \exp(q_0 + \mathbf{z}_{t-d}\mathbf{q})) \quad (1.6)$$

and

$$p(z_t) = \exp(p_0 + \mathbf{z}_{t-d}\mathbf{p}) / (1 + \exp(p_0 + \mathbf{z}_{t-d}\mathbf{p})) \quad (1.7)$$

\mathbf{z}_t is a matrix of economic and financial variables, \mathbf{q} is the vector of parameters for the probability of staying in a recession and \mathbf{p} is the vector of parameters for the probability of staying in a boom. The strategy to choose the number of regimes and lags for MS models is detailed in the beginning of section 3.

1.2.2 Comparison with other modelling approaches

Estrella and Mishkin (1998) estimate a probit model to predict the occurrence of US recessions using a wide range of financial and non-financial variables. They find that the slope of the yield curve exhibits the best out-of-sample performance and performs better itself than in conjunction with other variables. Kauppi and Saikkonen (2008) extend the static specification of Estrella and Mishkin (1998) by adding lagged values of the dependent variable and the variables included in the information set. On the basis of an out-of-sample

exercise, Nyberg (2010) shows that this "dynamic probit" model yields better results than the static model for predicting recessions. Birchenhall et al. (1999) use a logistic model for the identification and prediction of US recessions. They find that a binary model with a logistic specification achieves better results than an MS model with TVTP, using leading indicators as explanatory variables. However, their specification relies heavily on leading indicators, which are often subject to data revision, which makes real-time inference problematic.

In addition, these models take the NBER reference dates of turning points as dependent variable, which can be problematic since the NBER Business Cycle Dating Committee can wait a year or more to confirm that a turning point has occurred¹. In contrast to binary choices models, which take the NBER classification as dependent variable, MS models determine endogenously the business cycle regimes experienced by the US.

Strengthening the case for the yield curve as a predictor of US recessions, Galvao (2006) develops a model to predict recessions that accounts for both non-linearity and structural breaks. This so-called Structural Break Threshold VAR with the slope of the yield curve as a leading indicator performs well in an out-of-sample experiment. Dueker (2005) instead proposes an innovative approach combining information from discrete variables in vector autoregressions.

Recently, Aruoba et al. (2009) and Camacho and Perez-Quiros (2010) use large state-space models to extract business cycle indicators. Both papers incorporate data sampled at different frequencies. However, these models are often difficult to estimate. As a consequence, it is often not clear whether this increases in complexity yields better forecasting results than simpler models (see e.g. Giannone et al. (2008)).

To sum up, the MS model with TVTP is attractive since it does not rely on the NBER chronology. In addition, it is rather simple to estimate and it allows one to check whether a given variable has asymmetric effects, i.e. helps to predict that the economy enters into a recession, a boom or in both phases of the business cycle.

1.2.3 Estimation and model selection

The estimation of MS models is carried out by maximizing the log-likelihood function with the EM algorithm. The log-likelihood function is given by:

$$\log L = \sum_{t=1}^T \ln f(y_t | \Omega_{t-1}) \quad (1.8)$$

¹See the NBER website for the dating of the US recession and the release dates: <http://www.nber.org/cycles/cyclesmain.html>

where $f(y_t|\Omega_{t-1})$ is the conditional density of y_t given Ω_{t-1} , which is the information available up to time $t-1$. Since the states S_t are not observable, the determination of the log likelihood is not straightforward. This can be done using the filtering procedure proposed by Hamilton (1989) with slight changes to take into account the time variability of the transition probabilities as in Diebold et al. (1993) ².

Choosing the number of regimes for Markov-switching models is a tricky business. Indeed, the econometrician has to deal with two problems when estimating MS models. First, since some parameters are not identified under the null hypothesis, one has to use nonstandard tests to determine the number of regimes. This comes from the fact that the asymptotic distribution of standard tests are not standard and usually depends on the covariance of chi-square processes. Second, under the null hypothesis, scores are identically equal to zero.

Solutions to these problems have been suggested in the literature. Hansen (1992) considers the likelihood function as a function of unknown parameters and uses empirical process to bound the asymptotic distribution of a standardized likelihood ratio test statistic. Garcia (1998) points out that the test is computationally expensive if the grid search over the parameter space is extensive. In addition, the Hansen's test might be too conservative since it does not provide a critical value but only a lower bound for the likelihood ratio test statistic. Altissimo and Corradi (2002) derive asymptotic bounds for LM and Wald tests and show that it chooses in a robust way between a linear and a threshold model. However, Carrasco (2002) shows that tests for structural changes have no power if the data are generated by a MS or a Threshold model. Smith et al. (2006) propose a new information criterion, the Markov Switching Criterion, based on the Kullback-Leibler divergence between the true and candidate model to select the number of states and variables simultaneously for MS models. However, this information criterion does not apply to mixed regression models where only some parameters of the model change across regimes. Finally, Carrasco et al. (2009) introduce a new test for Markov switching parameters that only requires to estimate the model under the null hypothesis of constant parameters. However, the test does not allow to discriminate between models with 2-regime and 3-regime.

Since I am primarily interested in predicting business-cycle regimes, I compute the Mean Squared Error (MSE) using the difference between the NBER classification and the estimated filtered probabilities to select the number of regimes and lags in line with Birchenhall et al. (1999). In doing so, I only concentrate on MS models with constant transition probabilities (FTP hereafter). The MSE is defined as follows:

²Parameter estimation for all MS models is carried out using the optimization package Optmum of GAUSS 7.0. I extend Bellone (2005) toolbox MSVARlib 2.0 to allow for the estimation of MS models with TVTP.

$$MSE = \frac{1}{T} \sum_{t=1}^T (NBER_t - P(S_t = 0))^2 \quad (1.9)$$

where $NBER_t$ is a dummy variable that takes on a value of 1 if the US economy is in recession according to the NBER and 0 otherwise. $P(S_t = 0)$ is the estimated probability of recession.

1.2.4 Selecting the transition variable

First, I use the industrial production index as measure of economic activity since it is available on a monthly basis and often used for business cycle analysis.

Table 1.1 lists the driving variables z_t used for the time-varying transition probabilities. A thorough description of the data and their sources is reported in Table 1.6 in the appendix. I first include a set of monetary policy instruments. I also consider financial variables as possible driving forces of the business cycle. Estrella and Mishkin (1998) find that financial variables are useful predictors of the real activity: the slope of the yield curve exhibits the best predictive power of US recessions beyond a one-quarter horizon, while share prices prove to be useful at a one-to-three quarter horizon. In addition, Andreou et al. (2010) extract factors from daily financial series and use MIDAS regressions to show that financial variables are useful for predicting quarterly real GDP and inflation. I also include an index of oil prices as possible transition variable in order to account for the changes in commodity prices. Hamilton (2009) points out the crucial role of oil prices in the triggering of US recessions.

Following Ravn and Sola (1999), I look at the significance of the coefficients entering into the logistic function. An interesting issue is to check whether a given variable z_t has asymmetric effects, i.e. helps to predict that the economy enters into a recession, a boom or in both phases of the business cycle. However, one should be cautious while analyzing the significance of the parameters contained in the \mathbf{p} and \mathbf{q} vectors, this should not be directly interpreted as if there was a causality link between the z_t variable and the y_t variable. Finding that a variable z_t enters significantly into the logistic function means that this z_t variable has some predictive power on y_t . Only under the condition that the z_t variable is exogenous, one can say that variations in this variable have driven the y_t variable into recessions or expansions. For example, changes in the short-term interest rate might reflect the reaction of the central bank to deteriorating economic conditions. In such a situation, monetary policy can be seen as endogenous. This can explain the so-called "price puzzle" where a

Table 1.1: List of transition variables

Monetary Policy Instruments
Federal Funds (FF)
Monetary Aggregate M1 (M1)
Monetary Aggregate M2 (M2)
Monetary Aggregate M3 (M3)
Real Monetary Aggregate M1 (RM1)
Real Monetary Aggregate M2 (RM2)
Real Monetary Aggregate M3 (RM3)
Financial variables
Real Effective Exchange Rate (REER)
Stock Prices (SP)
3-month Treasury bill (T-bill)
10-year Treasury bond (T-bond)
Term premium (SPREAD)
Others
Index of Leading Indicators (CLI)
Consumer Price Index (CPI)
Oil Price (OIL)

This table lists the variables entering into the time-varying transition probabilities. The labels attached to each variable are in parentheses. A full description of the data is reported in Table 1.6 in the appendix.

contraction in the monetary policy leads to an increase in the price level. Christiano et al. (1996) show that including an index of commodity prices can resolve the "price puzzle".

1.3 In sample results

The general strategy of the in-sample analysis is the following:

- Estimate MS models with fixed transition probabilities (FTP) considering 2 to 4 regimes, 0 to 12 autoregressive lags and a switch or not in the variance of the residuals. For each model, compute the Mean Squared Error (MSE) defined in equation 1.9. Select the number of regimes, lags and structure of the disturbance variance on the basis of the MSE. The selected numbers of regimes, lags and form of disturbance variance will be used for all MS models with TVTP.
- Estimate an MS model with TVTP using each series of Table 1.1 in turn for $1 \leq d \leq 12$. Compute the LR statistic for testing the FTP model against the TVTP model. Calculate the SIC for selecting the best delay d for the transition variable. Delete the least significant variable in either p_t^{00} and p_t^{11} . If necessary, estimate the reduced model and compute the SIC for selecting the delay that has the most important predictive power.
- Finally, since the slope of the yield curve produces consistent results across all horizons, I estimate again MS models with TVTP using each series of Table 1.1 in turn in combination with the variable Spread. Again, I compute the SIC for selecting the best delay d in the transition function and delete the least significant variable in either p_t^{00} and p_t^{11} . I repeat this process until all explanatory variables are significant at the 10% level.

The Schwarz Information Criterion is defined as follows:

$$SIC = -2\text{Log}L + n * \log(T) \quad (1.10)$$

where n is the number of parameters estimated and T is the number of observations. Under the null hypothesis of the Likelihood Ratio test for checking the validity of the TVTP model against the FTP model, the FTP model is not accepted if $2[\log L_{TVTP} - \log L_{FTP}]$ exceeds $\chi_{c,\alpha}^2$ where c is the number of constraints and α the level of significance of the test.

Note that I do not extend the number of variables in the transition function beyond the two-variable case since I encounter convergence problems for the estimation of model with more than two variables in the transition function.

The model that gets the best fit in terms of MSE is a model with no lag, two regimes and a switch in the variance of the shocks. Choosing a model with no lag could be problematic if one wants to forecast future values of the index of industrial production. However, since I am primarily interested in modelling business-cycle regimes and predicting turning points, I concentrate my analysis on the estimated probabilities of being in a given state and leave aside the estimated measures of the index of industrial production. In this respect, the MSE criteria computed as the squared difference between the estimated filtered probabilities and the NBER classification appears to be the relevant goodness-of-fit measure for regime prediction purpose. Besides, unlike most of the models that include autoregressive lags, the model with no lag has a negative coefficient for the mean in the recession state, which is crucial for business cycle analysis as emphasized by Smith and Summers (2004). Note that the variance of the residuals is always higher in the recession state than in the boom state (see Table 1.2). This means that shocks hitting the US economy are higher during recessions than during expansions.

In-sample results are obtained by estimating equation 1.2 over the entire sample period (i.e. for $t=1970:1$ to $2008:11$) except for the variables OIL and REER for which the estimation period starts respectively for $t=1975:1$ and for $t=1971:1$ since observations for earlier periods are not available for these variables. I report two types of results: the coefficients and their associated standard errors and the Schwarz Information Criterion (SIC) based on the log-likelihood. Table 1.2 and Table 1.3 contain the variables that perform best for the in-sample analysis. The full in-sample results are reported in Table 1.8 and Table 1.9 in the appendix.

Table 1.2 examines the predictive power of each variable in turn. First, the variable that gets the best result in terms of SIC is the composite leading index from the Conference Board³, which yields its highest predictive power one month ahead. These results are consistent with Estrella and Mishkin (1998), who show that the predictive power of leading indicators tends to be short. However, one major drawback to the use of leading indicators for real-time forecasting is that the components and the weights of the indicator are often changed to improve the ex-post performance as pointed out by Diebold and Rudebusch (1989). In addition, using a real-time dataset, Koenig and Emery (1994) find that composite

³The composite leading index is computed by the Conference Board and includes among other variables: the spread between the 10-year Treasury bonds and the Federal Funds, the Federal Funds, money supply M2 and an index of stock prices. See http://www.conference-board.org/pdf_free/economics/bci/begweek.pdf for a full list of the variables included in the index.

leading indices have a limited predictive power for detecting cyclical turning points.

Second, stock prices - measured by the S&P500 index - are strong predictors in the very short run (i.e. one month ahead) but have a lower predictive content than the composite leading index in terms of SIC. Interestingly, I find that stock prices are relevant only during expansions. Stock prices are expected to be meaningful for predicting real economic activity since stock prices are usually explained by the expected stream of future dividends, which are often viewed as being correlated with future activity. Hamilton and Lin (1996) estimate a bivariate MS model and provide evidence for the links between stock prices and economic activity.

Third, the slope of the yield curve produces consistent results across all horizons. Unlike the Federal Funds, it appears to be significant only during expansions. When the yield curve flattens due e.g. to an increase in the short-term interest rate, real activity is expected to slow down. I indeed find an expected positive sign for p_1 . The reason for the good predictive performance of the yield curve is that it generally reflects the stance of the current monetary policy, which influences real activity. In addition, the yield curve also reflects expectations about monetary policy and inflation, which plays a substantial role for predicting economic activity. Estrella and Mishkin (1997) and Dueker (1997) carefully detail the importance of the yield curve as a predictor for the economic activity and provide empirical evidence for its predictive power.

Fourth, the price of oil helps to predict the evolution of the business cycle but only during expansions, it yields its highest predictive power at a two-month horizon. In particular, the results indicate that an increase in oil price decreases significantly the probability of staying in a boom.

Table 1.7 in the appendix reports results for the remaining variables. First, nominal monetary aggregates have a poor predicting power for the real economic activity. However, real monetary aggregates obtain better results especially for M3, which yields its highest predictive content two months ahead. The better fit from the real monetary aggregates (as compared to nominal monetary aggregates) can be attributed to the good fit of the inflation rate. Second, the Federal Funds - which are usually considered as the main instrument for monetary policy - seem to affect most real activity at a six-month horizon but only during recessions. Third, the predictive power of the real effective exchange rate is also tested. This measure includes the market exchange rates but also variations in relative price level, it can therefore be used as a proxy for the degree of competitiveness of a given country. The relevance of exchange rate as a mechanism driving business cycle has been widely identified in the literature about contagion. I find that a real depreciation of the US dollar increases

significantly the probability of staying in a boom. However, the LR test indicates that I cannot reject the validity of the FTP model against the TVTP model for the model including REER as a transition variable.

Table 1.3 reports the results when the slope of the yield curve is associated with one of the other variables of Table 1.1 in turn. The results of the single-variable analysis are generally confirmed. Stock prices and composite leading index have a good predictive power. Similarly, the price of oil proves to be relevant for predicting the economic activity when it is associated with the slope of the yield curve. The model that performs best in terms of SIC is the one with the yield curve at two different horizons in the logistic function (the spread lagged one and four periods).

Table 1.8 in the appendix reports results for the remaining variables associated with the slope of the yield curve. Combining the slope of the yield curve with monetary aggregates improves the in-sample fit relatively to models that only include monetary aggregates. I find the same pattern for the models with FF and REER. Note that the coefficients p_2 for the models with RM2, RM3 and REER and the coefficient q_2 for the model with FF are not significant at the 10% level, which implies that combining these variables with the slope of the yield curve is not useful. However, I do not further reduce these models since they will be used in the next section devoted to the out-of-sample results.

Figure 1.2 plots the predicted probabilities of being in a recession for the model with Spread lagged one and four months as a transition variable and for the MS model with FTP. The latter model gives false signals for recessions in the 1970s, in 1998 and in 2005 and miss the 1990/91 and 2001 recessions whereas the former model provides an accurate description of the recessions experienced by the US in the post-1970 era although the 1990/91 recession and the recession that started in December 2007 are detected with a certain lag.

1.4 Out-of-sample results

The methodology for the recursive out-of-sample analysis is the following. First, I estimate the parameters of the model using data ranging from 1970:1 to 1997:12 in order to capture enough recessions to get accurate parameter estimates. I then compute one, three and six periods ahead forecasts for 1998:1, 1998:3 and 1998:6. The model is then re-estimated using data from 1970:1 to 1998:1, and a one, three and six periods ahead forecasts are computed for 1998:2, 1998:4 and 1998:7. The process is repeated until the end of the sample 2008:11 is reached. This means that I compute 132 out-of-sample forecasts for each of the

Table 1.2: Estimation results, Variable by themselves, In-sample estimates

	FTP	$SPREAD_{t-10}$	SP_{t-1}	CLI_{t-1}	CPI_{t-2}	OIL_{t-2}^1
μ_0	-0.333 (0.254)	-0.590*** (0.175)	-0.441* (0.244)	-0.645*** (0.155)	-0.569* (0.314)	-0.347 (0.216)
μ_1	0.293*** (0.038)	0.318*** (0.033)	0.308*** (0.032)	0.334*** (0.033)	0.301*** (0.033)	0.311*** (0.039)
σ_0	1.665*** (0.399)	1.314*** (0.303)	1.520*** (0.319)	1.128*** (0.238)	1.614*** (0.424)	1.173*** (0.301)
σ_1	0.271*** (0.036)	0.314*** (0.025)	0.286*** (0.036)	0.314*** (0.025)	0.302*** (0.038)	0.255*** (0.034)
q_0	0.835*** (0.065)	0.861*** (0.059)	0.797*** (0.065)	3.994** (1.639)	0.829*** (0.077)	0.873*** (0.055)
q_1	-	-	-	-2.872** (1.409)	-	-
p_0	0.964*** (0.016)	2.452*** (0.611)	4.799*** (1.136)	4.861*** (0.874)	6.096*** (1.451)	4.079*** (0.754)
p_1	-	2.284** (1.139)	0.541*** (0.177)	2.474*** (0.815)	-4.242*** (1.557)	-0.133** (0.058)
SIC	946.732	931.457	937.218	921.479	934.993	777.314
LogL	-465.358	-456.385 ²	-459.266 ²	-450.062 ²	-458.154 ²	-379.523 ²

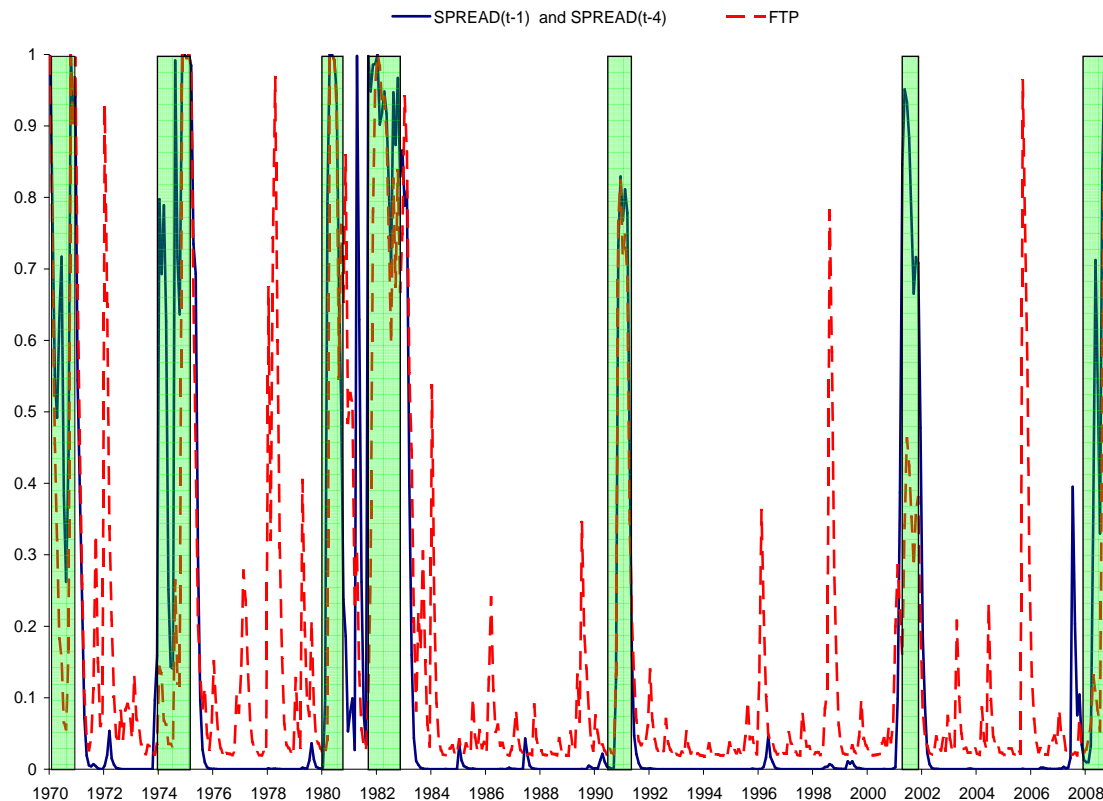
***, ** and * indicate significance at the 1%, 5% and 10% level. Standard errors are in parentheses. ¹ means that the estimation period for the model with oil price starts in 1975:1. ² means that a standard likelihood ratio test can reject the null hypothesis of constant transition probabilities at the 5% level.

Table 1.3: Estimation results, Variables in combination with $SPREAD_{t-k}$, In-sample estimates

	FTP	$SPREAD_{t-1}$ & $SPREAD_{t-4}$	SP_{t-1}	CLI_{t-2}	CPI_{t-9}	OIL_{t-3}^1
μ_0	-0.333 (0.254)	-0.669*** (0.158)	-0.368** (0.173)	-0.420*** (0.113)	-0.586*** (0.175)	-0.563*** (0.184)
μ_1	0.293*** (0.038)	0.330*** (0.030)	0.318*** (0.031)	0.382*** (0.031)	0.315*** (0.035)	0.337*** (0.033)
σ_0	1.665*** (0.399)	1.144*** (0.235)	1.407*** (0.262)	0.926*** (0.144)	1.346*** (0.331)	0.864*** (0.219)
σ_1	0.271*** (0.036)	0.318*** (0.024)	0.269*** (0.030)	0.283*** (0.027)	0.314*** (0.025)	0.279*** (0.026)
q_0	0.835*** (0.065)	0.854*** (0.052)	0.805*** (0.061)	3.984** (1.712)	0.859*** (0.062)	0.856*** (0.055)
q_1	-	-	-	-	-	-
q_2	-	-	-	4.253* (2.431)	-	-
p_0	0.964*** (0.016)	6.673** (2.736)	3.702*** (1.057)	0.247 (0.588)	2.996** (1.159)	3.269*** (0.906)
p_1	-	-7.186** (3.387)	0.588* (0.233)	-2.399** (1.119)	1.971* (1.210)	1.334** (0.631)
p_2	-	9.911** (4.431)	0.812** (0.421)	2.541*** (0.835)	-1.021* (1.671)	-0.177** (0.081)
SIC	946.732	927.762	935.160	926.399	933.742	771.471
LogL	-465.358	-453.204 ²	-456.903 ²	-451.188 ²	-456.194 ²	-375.297 ²

***, ** and * indicate significance at the 1%, 5% and 10% level. Standard errors are in parentheses. ¹ means that the estimation period for the model with oil price starts in 1975:1. ² means that a standard likelihood ratio test can reject the null hypothesis of constant transition probabilities at the 5% level.

Figure 1.2: IN-SAMPLE PROBABILITY OF RECESSION, MS MODEL WITH TVTP AND FTP



forecasting horizons (1, 3 and 6-month ahead)⁴. I evaluate the forecasting performance of each model on the basis of the Relative Mean Squared Error, which is the ratio between the MSE of an MS model with TVTP using a given indicator and the MSE of the benchmark model (FTP model).

Computing forecasts for MS models with TVTP is not straightforward since one needs to forecast the variables entering into the transition function for the forecasts beyond a one-period horizon. I decide to compute "iterated" forecasts for the variables entering into the transition function. The good performance of iterated forecasts with respect to direct forecasts for time-series model is detailed in Marcellino et al. (2006). Iterated forecasts require first to estimate an AR model, then iterating using the parameters of the estimated AR model to obtain the forecasts. I select the "best" AR model for the transition variables on the basis of the SIC computed for the entire sample period.

⁴To evaluate the 3-month and 6-month-ahead forecasts, I lengthen appropriately the evaluation sample.

Once again, I concentrate my analysis on the prediction of business cycle regimes (i.e. recession or expansion) rather than forecasting quantitative measures of the index of industrial production. This means that I forecast the probabilities of being in a given regime. The j -month ahead predicted filtered probabilities are computed recursively as follows:

$$\mathbf{FP}_{t+j} = \mathbf{P}_{t+j} \mathbf{FP}_{t+j-1} \quad (1.11)$$

where \mathbf{FP}_{t+j} is the (2x1) vector of filtered probabilities in period $t + j$ and \mathbf{P}_{t+j} is the (2x2) matrix of transition probabilities in period $t + j$.

Granger and Terasvirta (1993) point out that forecasts for \mathbf{P}_{t+j} (for $j > 1$) can be biased. The basic idea is that in general the linear conditional expectation operator cannot be exchanged with the non-linear operator, which is a linear combination of logistic functions in our case. In other words, the expected value of a nonlinear function is not equal to the function evaluated at the expected value of its argument. In the next section, I come back to this issue and compute Monte Carlo forecasts, which allows me to potentially correct this bias.

In this section, I report the MSE for $h = 1$, $h = 3$ and $h = 6$. I use the MS model with constant transition probabilities as a benchmark for assessing the forecasting performance of the variables under scrutiny. Note that for obtaining the out-of-sample forecasts, I use the selected models obtained from the in-sample analysis (i.e. models with 2 regimes).

In line with the in-sample results, Table 1.4 shows that share prices outperform the benchmark model for 1- and 3-month-ahead predictions whereas the composite leading index is better than the benchmark model only for 1-month-ahead predictions. The slope of the yield curve outperforms the benchmark model beyond a 1-month horizon, while the price of oil outperforms the benchmark model for 6-month-ahead predictions. Note that the CPI loses its predictive power, while M2 achieves very good forecasting results since it outperforms the benchmark model for the three horizons of predictions.

Table 1.9 in the appendix shows that the remaining variables have a poor fit as they do not outperform the benchmark model apart from the model with M3 that outperforms the benchmark model for one and three months ahead predictions.

Table 1.5 reports the out-of-sample forecasting performance of selected variables in combination with the slope of the yield curve. In line with Estrella and Mishkin (1998), I find that combining variables to the slope of the yield curve tend to worsen out-of-sample predictions. This finding differs from the in-sample results. In particular, most of the models do

Table 1.4: RMSE, Variables by themselves, Out-of-sample results

	$h = 1$	$h = 3$	$h = 6$
$SPREAD_{t-1}$	1.02	0.99	0.98
SP_{t-1}	0.96	0.98	1.07
CLI_{t-1}	0.97	1.03	1.13
OIL_{t-1}	1.06	1.01	0.93
$M2_{t-1}$	0.95	0.95	0.98

This table reports the relative mean squared forecast error (RMSE). The benchmark model is a model with constant transition probabilities.

not perform better than the benchmark model (see Table 1.10 in the appendix). However, I find that using the slope of the yield curve lagged one and four months produces excellent results since this model largely outperforms the benchmark model.

Figure 1.3 and Figure 1.4 plot the 1-month and 3-month-ahead forecasts of the filtered probabilities at a one month and three months horizon for the FTP model and the model including the slope of the yield curve lagged one and four periods. I decide to consider that the economy is in recession if the probability of being in a recession is higher than .5. The FTP model yields contrasted results since it detects with a certain lag the 2001 recession and the recession that began in December 2007. In addition, the FTP model gives a false signal of recession in October and November 2005.

Conversely the model with the slope of the yield curve is more convincing since the 2001 recession is declared on time for both one and three months ahead forecasts although it is identified to last a bit longer than the NBER datation. Concerning the recession that has started in December 2007, Figure 3 shows that a first recession signal is given in June 2008 (using information available up to May 2008). However, the probability of recession falls below .5 for July and August 2008 before increasing again in September 2008. Figure 1.4 reports the predicted probabilities of recession three months ahead, it does not give further insight for an early warning of this recession. In the next section, I implement the out-of-sample exercise with a real-time dataset.

Table 1.5: RMSE, Variables in combination with $SPREAD_{t-1}$, Out-of-sample results

	$h = 1$	$h = 3$	$h = 6$
$SPREAD_{t-4}$	0.70	0.77	0.92
SP_{t-1}	1.01	0.97	1.02
CLI_{t-1}	1.22	1.20	1.21
OIL_{t-1}	1.74	1.54	1.19
$M2_{t-1}$	1.02	1.00	0.99

This table reports the relative mean squared forecast error (RMSE). The benchmark model is a model with constant transition probabilities.

1.5 Case study: an application of the model

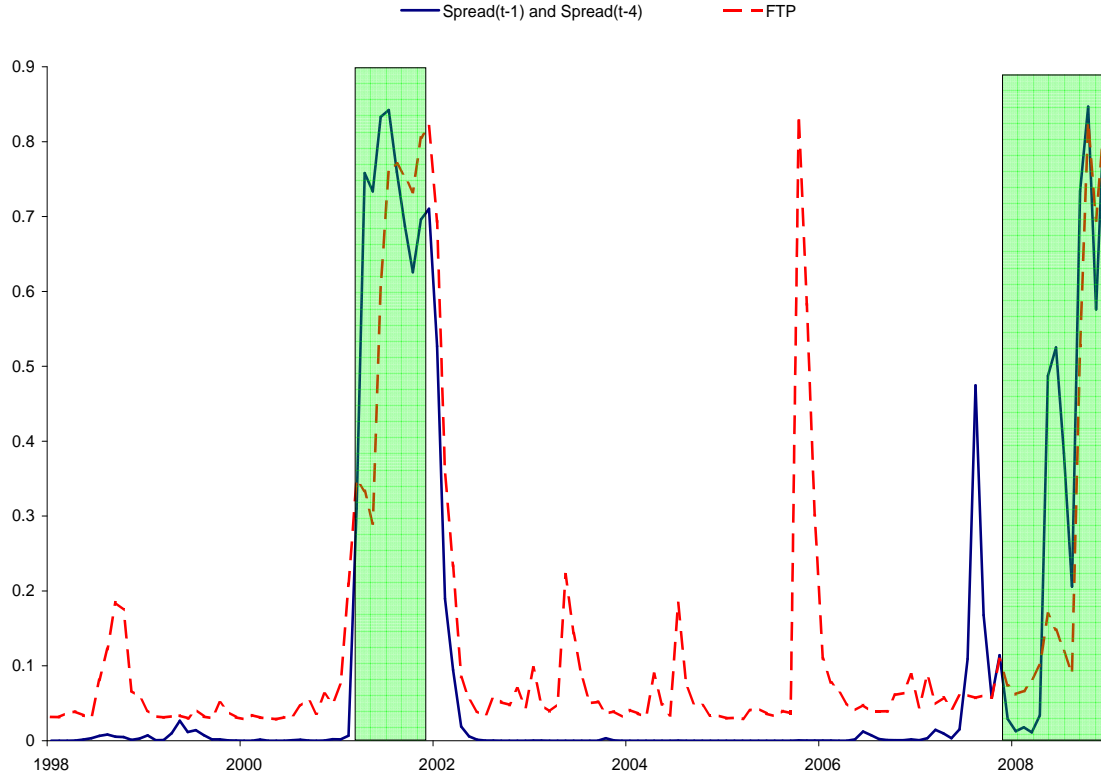
In this section, I look at the performance of the MS model with TVTP to predict the US recession that began in December 2007 according to the NBER. It is found that it is possible to predict a recession in the summer of 2008, while the NBER declared a peak had occurred in December 2007 using information available up to December 2008. Unlike the NBER Dating Committee, I only use current and lagged information. In addition, I use real-time data so that I obtain real-time forecasts, which are the relevant ones for policy makers and market participants.

The real-time data for the index of industrial production is obtained from the Federal Reserve Bank of St. Louis⁵. When I have two sets of data for the same month, I always choose the vintage data with the latest release date. I use real-time data from 2005:10 to 2009:7. I then run an out-of-sample exercise in the same way than in section 1.4. This means that I initially estimate the parameters of the model using data from 1970:1 to 2005:10, then I increase the final date of the estimation by a month until I reach the end of the sample, which is 2009:7 in this section. From $t=2005:10$ to $t=2009:7$, I compute one month and three months ahead forecasts.

For one-step-ahead forecasts, there is no bias arising from the naïve forecasts since the values of the variables entering into the transition function are known. However, for forecasts

⁵The data are drawn from the ALFRED database available at: <http://alfred.stlouisfed.org/>

Figure 1.3: ONE-MONTH-AHEAD PREDICTED PROBABILITY OF RECESSION, MS MODEL WITH TVTP AND FTP



with horizon $h > 1$, the strategy for computing Monte Carlo forecasts is closely linked to the methodology proposed by Koop et al. (1996) for computing impulse responses in non linear models. The algorithm I use is the following:

- *Step 1:* Assume that the transition variable z_t follows an AR(p) process:

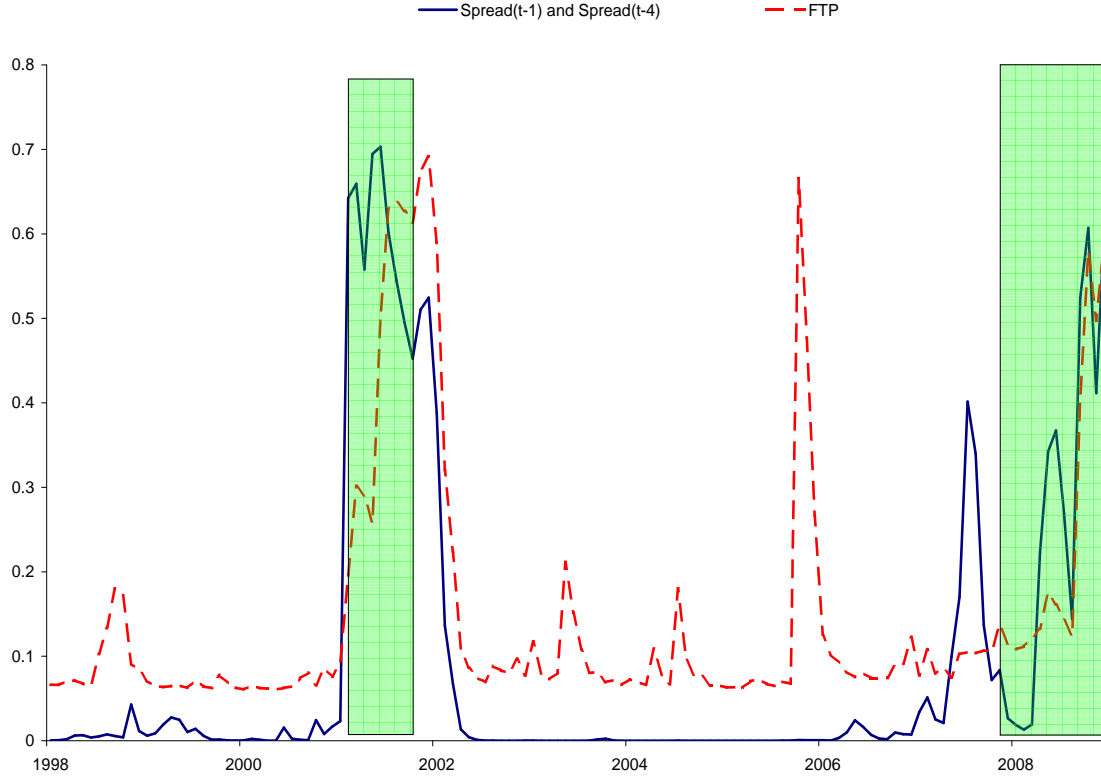
$$z_t = \alpha_0 + \alpha_1 z_{t-1} + \dots + \alpha_p z_{t-p} + e_t = g(z_t, e_t) \quad (1.12)$$

where e_t is taken to be i.i.d. with zero mean and distribution D . The transition probabilities h-period ahead (for $h > 1$) p_{T+h}^{ij} are given by:

$$p_{T+h}^{ij} = L\{g(z_{T+h-1}, e_{T+h-1})\} \quad (1.13)$$

where $L\{.\}$ is the logit transformation.

Figure 1.4: THREE-MONTH-AHEAD PREDICTED PROBABILITY OF RECESSION, MS MODEL WITH TVTP AND FTP



- *Step 2*: For a given forecast horizon h , randomly sample $(h - 1)xN$ values of the innovation e_t ⁶, where N is the number of Monte Carlo replications.
- *Step 3*: Use the first $(h - 1)$ random shocks to compute the forecasts for the transition probabilities $p_{T+h}^{ij(1)}$.
- *Step 4*: Repeat Step 2 N times and form the average:

$$p_{T+h}^{ij} = \frac{1}{N} \sum_{n=1}^N p_{T+h}^{ij(n)} \quad (1.14)$$

If N is large enough, by the Law of Large Numbers the average of the Monte Carlo replication will converge to the exact conditional expectation.

⁶The normal distribution is used in the application.

- *Step 5:* Repeat Steps 2 to 4 a sufficient number of times R to get accurate estimation of the forecasts. Again, as the number of repetition increases pointwise convergence will be guaranteed by the Law of Large Numbers.

Figure 1.5 shows the one-month-ahead predicted probabilities of recession for the model that include the slope of the yield curve lagged one and four months into the transition function. I do not use the above algorithm for getting these forecasts since for one month ahead predictions the value of the transition variable is known. Figure 1.5 only differs from Figure 1.4 in the sense that real-time data are used. Note that there is a first jump in the probabilities of recession in August 2007. However, these probabilities decrease below .5 (our cut-off value that determines whether the economy is in recession or not) until June 2008. Indeed, there is a peak in the probability of recession in June 2008 based on predictive information from May 2008. However, the probabilities of recession for July and August 2008 decrease below .5 before increasing again above .5 in September 2008. Interestingly, this model indicates that the recession that began in December 2007 according to the NBER reached its end in August 2009 ⁷.

Figure 1.6 plots the three-month-ahead probability of recession for the model with the slope of the yield curve lagged one and four months. For these forecasts the above algorithm is used with $N = 100$ and $R = 100$. The probability looks very similar though there are shifted downward and to the right so that the model indicates that the recession will be ending in October 2009 using information available until July 2009.

Overall, the model using the spread at two different horizons does a fairly good for tracking the US business cycle regimes since it is able to anticipate the NBER Business Cycle Dating Committee announcements.

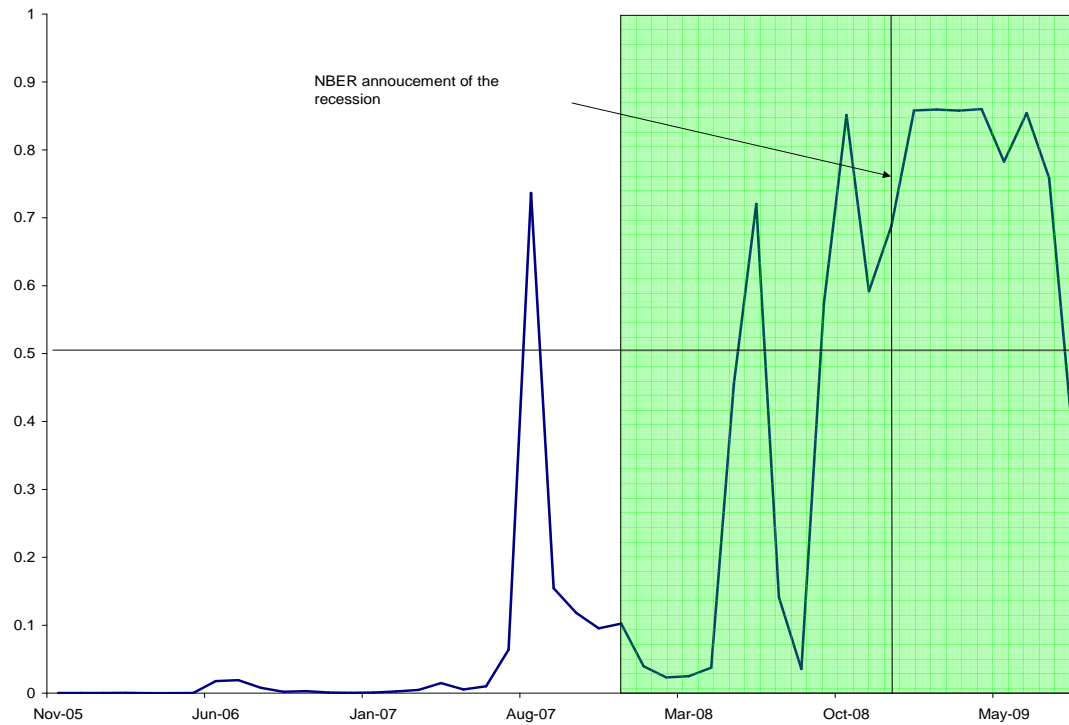
1.6 Conclusions

This paper presents a consistent framework for model selection and parameter estimation of Markov-switching models with time-varying transition probabilities. This class of model proves to be successful for the identification and prediction of the US business cycle regimes as defined by the NBER (i.e. expansion and contraction). Besides, it significantly outperforms the Markov-switching model with constant transition probabilities.

In line with previous studies, it is found that stock indices and index of leading indicators perform well for predictions with one month horizon, while the slope of the yield curve turns

⁷On September 20, 2010, the NBER announced that the last recession came to an end in June 2009. This announcement was not available at the time the paper was written.

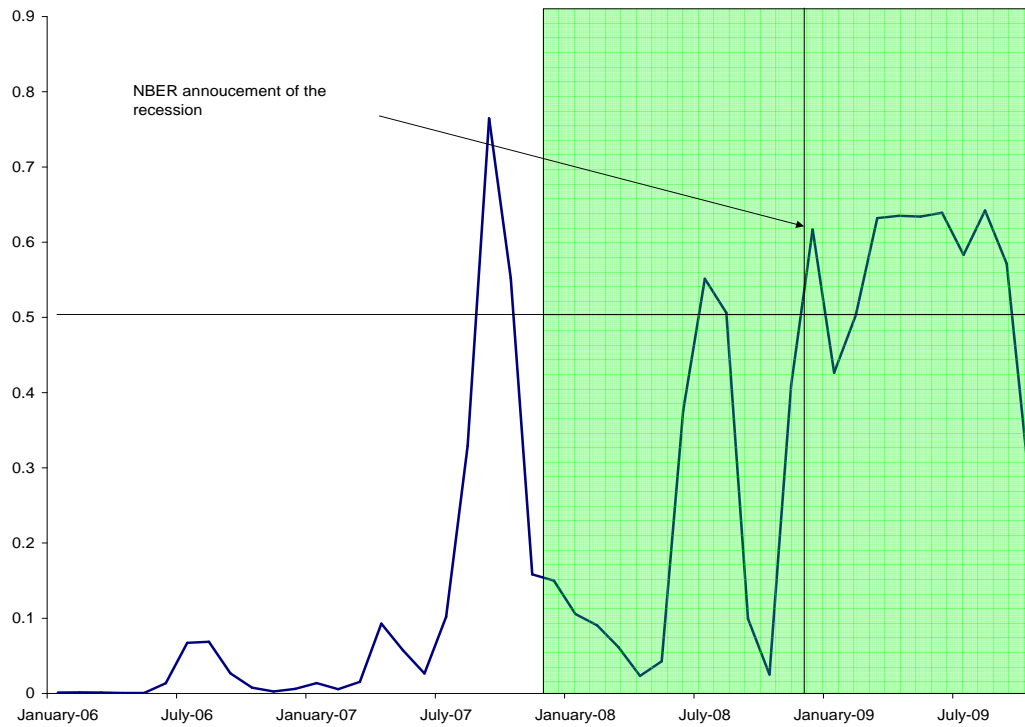
Figure 1.5: ONE-MONTH-AHEAD PREDICTED PROBABILITY OF RECESSION, REAL-TIME DATA



out to be the most useful indicator for out-of-sample predictions since it performs better by itself than in combination with other variables.

The one month-ahead forecasts using the slope of the yield curve at two different horizons as a transition variable is successful for out-of-sample predictions. In particular, using a real-time dataset, the recession that began in the US in December 2007 is identified before the announcement of the NBER Business Cycle Dating Committee.

Figure 1.6: THREE-MONTH-AHEAD MONTE CARLO PREDICTED PROBABILITY OF RECESSION



1.7 Appendix Chapter 1

Table 1.6: Data appendix

Series	Transformation	Sample Period	Description	Sources
IP	$100\Delta Ln$	1970:1-2008:11	Industrial Production	Federal Reserve Database
FF	$\Delta(Mean)$	1970:1-2008:11	Federal Funds	Federal Reserve Database
M1	$100\Delta Ln$	1970:1-2008:11	Monetary Aggregate M1	OECD Database
M2	$100\Delta Ln$	1970:1-2008:11	Monetary Aggregate M2	OECD Database
M3	$100\Delta Ln$	1970:1-2008:11	Monetary Aggregate M3	OECD Database
Real M1	$100\Delta Ln$	1970:1-2008:11	M1 deflated by CPI	OECD Database
Real M2	$100\Delta Ln$	1970:1-2008:11	M2 deflated by CPI	OECD Database
Real M3	$100\Delta Ln$	1970:1-2008:11	M3 deflated by CPI	OECD Database
REER	$100\Delta Ln$	1971:1-2008:11	Real Effective exchange rate	OECD Database
SP	$100\Delta Ln$	1970:1-2008:11	S&P500	OECD Database
T-bill	$\Delta(Mean)$	1970:1-2008:11	3-month Treasury bill	Federal Reserve Database
T-bond	$\Delta(Mean)$	1970:1-2008:11	10-year Treasury bond	Federal Reserve Database
SPREAD	Level	1970:1-2008:11	Spread between T-bond and T-bill	Federal Reserve Database
CPI	$100\Delta Ln$	1970:1-2008:11	Consumer Price Index	OECD Database
CLI	$100\Delta Ln$	1970:1-2008:11	Index of Leading Indicators	Conference Board
OIL	$100\Delta Ln$	1975:1-2008:11	US Crude oil imported acquisition cost by refiners	Energy Information Administration

The first column indicates the series name, the second column the transformation applied to the series, the third column is the sample period, the fourth column shortly describes the series and the last column indicates the data sources. $100\Delta Ln$ means that the series are taken as 100 times its log difference, Level indicates that the series is taken in level, $\Delta(Mean)$ indicates that the series is demeaned.

Table 1.7: Estimation results, Variable by themselves, In-sample estimates

	FF_{t-6}	$M1_{t-5}$	$M2_{t-12}$	$M3_{t-12}$	$RM1_{t-1}$	$RM2_{t-3}$	$RM3_{t-2}$	$REER_{t-7}^1$	$Tbill_{t-11}$
μ_0	-0.276 (0.214)	-0.456** (0.229)	-0.224 (0.217)	-0.187 (0.218)	-0.583*** (0.165)	-0.719*** (0.223)	-0.703*** (0.225)	-0.403 (0.277)	0.154 (0.189)
μ_1	0.283*** (0.032)	0.312*** (0.044)	0.272*** (0.034)	0.266*** (0.035)	0.341*** (0.033)	0.295*** (0.033)	0.306*** (0.031)	0.310*** (0.037)	0.277*** (0.034)
σ_0	1.745*** (0.366)	1.486*** (0.469)	1.831*** (0.406)	1.849*** (0.404)	1.123*** (0.244)	1.527*** (0.396)	1.465*** (0.348)	1.592*** (0.405)	1.629*** (0.326)
σ_1	0.263*** (0.029)	0.284*** (0.035)	0.259*** (0.028)	0.256*** (0.029)	0.303*** (0.026)	0.335*** (0.027)	0.319*** (0.028)	0.272*** (0.037)	0.245*** (0.031)
q_0	0.684 (0.636)	2.196* (1.196)	0.796*** (0.069)	0.791*** (0.069)	0.879*** (0.047)	0.854*** (0.063)	0.849*** (0.067)	0.826*** (0.067)	0.950*** (0.514)
q_1	0.279** (0.133)	-2.113* (1.350)	-	-	-	-	-	-	0.348* (0.170)
p_0	0.954*** (0.019)	0.972*** (0.014)	4.399*** (0.745)	4.432*** (0.752)	4.809*** (0.967)	7.361*** (3.749)	5.232*** (1.230)	3.822*** (0.689)	0.949** (0.022)
p_1	-	-	-1.884** (0.762)	-1.934** (0.752)	2.629** (1.035)	9.198* (5.344)	5.248*** (1.822)	-0.849** (0.385)	-
SIC	940.289	944.289	943.409	943.310	941.024	936.064	932.754	903.661	941.366
LogL	-460.802 ²	-462.802 ²	-462.362 ²	-462.313 ²	-461.169 ²	-458.689 ²	-457.034 ²	-442.527	-461.340 ²

***, ** and * indicate significance at the 1%, 5% and 10% level. For each variable, I select the lag that obtains the best fit in terms of SIC. ¹ means that the estimation period for the model with REER starts in 1971:1. ² means that a standard likelihood ratio test can reject the null hypothesis of constant transition probabilities at the 5% level.

Table 1.8: Estimation results, Variable X_{t-k} in combination with $SPREAD_{t-k}$, In-sample estimates

X_{t-k}	FF_{t-10}	$M1_{t-12}$	$M2_{t-5}$	$M3_{t-4}$	$RM1_{t-5}$	$RM2_{t-10}$	$RM3_{t-10}$	$REER^1_{t-10}$
μ_0	-0.586*** (0.178)	-0.599*** (0.192)	-0.498*** (0.147)	-0.591*** (0.152)	-0.519*** (0.153)	-0.588*** (0.176)	-0.595*** (0.181)	-0.644*** (0.203)
μ_1	0.316*** (0.033)	0.299*** (0.031)	0.346*** (0.031)	0.351*** (0.032)	0.324*** (0.024)	0.316*** (0.034)	0.313*** (0.035)	0.322*** (0.034)
σ_0	1.342*** (0.319)	1.537*** (0.347)	1.106*** (0.202)	1.052*** (0.209)	1.326*** (0.292)	1.331*** (0.318)	1.367*** (0.349)	1.278*** (0.325)
σ_1	0.313*** (0.025)	0.319*** (0.024)	0.297*** (0.024)	0.300*** (0.024)	0.295*** (0.026)	0.314*** (0.025)	0.314*** (0.025)	0.314*** (0.025)
q_0	1.574* (0.763)	0.847*** (0.070)	0.870*** (0.056)	0.852*** (0.051)	1.021 (0.725)	0.860*** (0.025)	0.856*** (0.063)	0.875*** (0.058)
q_1	-	-	-	-	1.117* (0.637)	-	-	-
q_2	0.058 (0.132)	-	-	-	-3.461** (1.519)	-	-	-
p_0	2.408*** (0.609)	3.258*** (0.951)	-1.653 (1.883)	0.257 (1.166)	3.229*** (0.968)	2.453*** (0.629)	2.542*** (0.692)	2.685*** (0.670)
p_1	2.159* (1.198)	3.256** (1.305)	2.420** (1.191)	1.492*** (0.539)	1.347** (0.589)	2.070* (1.237)	1.924* (1.186)	2.139* (1.127)
p_2	-	-1.842** (0.907)	10.628* (6.466)	5.676* (3.043)	1.942* (1.040)	0.606 (1.749)	0.969 (1.590)	-0.011 (0.118)
SIC	933.920	932.722	931.935	932.358	934.842	933.987	933.777	894.714
LogL	-456.283 ²	-455.684 ²	-455.290 ²	-455.502 ²	-454.074 ²	-456.316 ²	-456.211 ²	-436.725 ²

***, ** and * indicate significance at the 1%, 5% and 10% level. For each variable, I select the lag that obtains the best fit in terms of SIC. Standard errors are in parentheses. ¹ means that the estimation period for the model with REER starts in 1971:1. ² means that a standard likelihood ratio test can reject the null hypothesis of constant transition probabilities at the 5% level.

Table 1.9: RMSE, Variables by themselves, Out-of-sample results

	$h = 1$	$h = 3$	$h = 6$
FF_{t-1}	1.48	1.33	1.14
$M1_{t-1}$	1.14	1.02	1.04
$M2_{t-1}$	0.95	0.95	0.98
$M3_{t-1}$	0.96	0.98	1.00
$RM1_{t-1}$	1.05	1.04	1.06
$RM2_{t-1}$	1.15	1.15	1.14
$RM3_{t-1}$	1.20	1.19	1.17
CPI_{t-1}	1.17	1.16	1.14
$REER_{t-1}$	2.99	2.54	1.96
$Tbill_{t-1}$	1.52	1.34	1.15
$Tbond_{t-1}$	1.32	1.28	1.22

This table reports the relative mean squared forecast error (RMSE). The benchmark model is a model with constant transition probabilities.

Table 1.10: RMSE, Variables in combination with $SPREAD_{t-1}$, Out-of-sample results

	$h = 1$	$h = 3$	$h = 6$
FF_{t-1}	1.21	1.25	1.13
$M1_{t-1}$	1.03	0.99	0.98
$M2_{t-1}$	1.02	1.00	0.99
$M3_{t-1}$	1.23	1.13	1.22
$RM1_{t-1}$	2.46	1.67	1.24
$RM2_{t-1}$	1.34	1.17	1.16
$RM3_{t-1}$	0.98	1.02	1.05
CPI_{t-1}	0.94	0.98	1.03
$REER_{t-1}$	1.01	1.01	1.01

This table reports the relative mean squared forecast error (RMSE). The benchmark model is a model with constant transition probabilities.

Chapter 2

Markov-switching MIDAS models

Abstract This paper introduces a new regression model - Markov-switching mixed data sampling (MS-MIDAS) - that incorporates regime changes in the parameters of the mixed data sampling (MIDAS) models and allows for the use of mixed-frequency data in Markov-switching models. After a discussion of estimation and inference for MS-MIDAS, and a small sample simulation based evaluation, the MS-MIDAS model is applied to the prediction of the US and UK economic activity, in terms both of quantitative forecasts of the aggregate economic activity and of the prediction of the business cycle regimes. Both simulation and empirical results indicate that MS-MIDAS is a very useful specification.

Keywords: Business cycle, Mixed-frequency data, Non-linear models, Forecasting, Nowcasting.

JEL Classification Code: C22, C53, E37.

⁰This is a joint work with Massimiliano Marcellino

2.1 Introduction

The econometrician often faces a dilemma when observations are sampled at different frequencies. One solution consists in estimating the model at the lowest frequency, temporally aggregating the high-frequency data. However, this solution is not fully satisfactory since important information can be discarded in the aggregation process. A second solution is to temporally disaggregate (interpolate) the low frequency variables. However, there is no agreement on the proper interpolation method, and the resulting high frequency variables would be affected by measurement error.

The third option is represented by regression models that combine variables sampled at different frequencies. They are particularly attractive since they can use the information of high-frequency variables to explain variables sampled at a lower frequency without any prior aggregation or interpolation. In this context, the MIDAS (Mixed Data Sampling) model of Ghysels et al. (2004) and Ghysels et al. (2007) has recently gained considerable attention. A crucial feature of this class of models is the parsimonious way of including explanatory variables through a weighting function, which can take various shapes depending on the value of its parameters.

MIDAS models have been applied for predicting both macroeconomic and financial variables. Ghysels et al. (2006) use the MIDAS framework to predict the volatility of equity returns, while Clements and Galvao (2008) and Clements and Galvao (2009) successfully apply MIDAS models to the prediction of quarterly US GDP growth using monthly indicators as high frequency variables. Andreou et al. (2010) exploit the informational content of daily financial variables to predict quarterly GDP and inflation in the US. In particular, they extend the standard MIDAS model to include factors in a dynamic framework, along the lines of Marcellino and Schumacher (2010).

MIDAS models are generally used as single-equation models where the dynamics of the indicator is not modelled. By contrast, system-based models such as the mixed-frequency VAR (MF-VAR) explicitly model the dynamics of the indicator. Kuzin et al. (2011) compare the forecasting performance of MIDAS and MF-VAR models for the prediction of the quarterly GDP growth in the Euro area. They find that MIDAS models outperform MF-VAR for short horizons (up to five months), while MF-VAR tend to perform better for longer horizons. A similar comparison is provided in Bai et al. (2010).

An issue that has attracted so far limited attention in the MIDAS literature is the stability of the relationship between the high and low frequency variables. Time-variation in MIDAS models has been only introduced by Galvao (2009) via a smooth transition function governing

the change in some parameters of the model. This Smooth Transition MIDAS is applied to the prediction of quarterly US GDP using weekly and daily financial variables.

In this paper, we propose an alternative way to allow for time-variation in the MIDAS model, introducing the Markov-switching MIDAS (MS-MIDAS) model. Regime changes may result from asymmetries in the process of the mean or variance. From an economic point of view, the predicting ability of the higher frequency variables could change across regimes following, e.g., changes in market conditions or business cycle phases. For example, the slope of the yield curve is often considered as a strong predictor of US recessions, an inverted yield curve signaling a forthcoming recession. However, Galbraith and Tkacz (2000) argue that the predictive power of the slope of the yield curve is limited in normal times. Therefore, it could be important to permit time-variation in the predictive ability of the high-frequency data. Indeed, our empirical applications show that in general the predictions from MS-MIDAS models are more accurate than those from simple MIDAS models.

An additional attractive feature of Markov-switching models is the possibility of estimating and predicting the probabilities of being in a given regime. The literature (e.g., Estrella and Mishkin (1998), Birchenhall et al. (1999)) often uses binary response models to predict the state of the economy using the NBER dating of expansions and contractions as a dependent variable. However, this method can be problematic since the announcements of turning points may be published up to twenty months after the turning point has actually occurred. Our MS-MIDAS model instead allows for real time evaluation and forecasting of the probability of being in a given regime.

Finally, MS-MIDAS is also a convenient approach to allow for the use of mixed frequency information in standard Markov-switching models. Hamilton (2010) pointed out the importance of using models with mixed frequency data for predicting recessions in real time. In our applications, the forecasting performance of standard MS models is indeed improved by the use of higher frequency information.

The paper is organized as follows. Section 2.2 reviews the MIDAS approach, introduces the MS-MIDAS, and discusses the estimation method. Section 2.3 presents Monte-Carlo simulations to assess the accuracy of the proposed estimation method in finite samples and its forecasting accuracy. Section 2.4 discusses an empirical application to the prediction of quarterly GDP growth and business cycle turning points in the US and the UK. Both empirical applications use financial variables as indicators. Section 2.5 concludes.

2.2 Markov-switching MIDAS

2.2.1 MIDAS approach

2.2.1.1 Basic MIDAS

The MIDAS approach of Ghysels et al. (2004) and Ghysels et al. (2007) involves the regression of variables sampled at different frequencies. Following the notation of Clements and Galvao (2008), and assuming that the model is specified for h -step ahead forecasting, the basic univariate MIDAS model is given by:

$$y_t = \beta_0 + \beta_1 B(L^{1/m}; \theta) x_{t-h}^{(m)} + \epsilon_t \quad (2.1)$$

where $B(L^{1/m}; \theta) = \sum_{j=1}^K b(j; \theta) L^{(j-1)/m}$ and $L^{s/m} x_{t-1}^{(m)} = x_{t-1-s/m}^{(m)}$. Note that t refers to the time unit of the dependent variable y_t and m to the time unit of the higher frequency variables $x_{t-h}^{(m)}$.

The forecasts of the MIDAS regression are computed directly so that no forecasts for the explanatory variables are required. However, unlike iterated forecasts, direct forecasts require to re-estimate the model when the forecasting horizon changes, see Chevillon and Hendry (2005) and Marcellino et al. (2006) for a comparison of the relative merits of iterated and direct forecasts.

The crucial difference between MIDAS and Autoregressive Distributed Lag models is that the content of the higher frequency variable is exploited in a parsimonious way through the polynomial $b(j; \theta)$, which allows to have a rich variety of shapes with a limited number of parameters. Ghysels et al. (2007) detail various specifications for the polynomial of lagged coefficients $b(j; \theta)$. A popular choice for the weighting scheme is the exponential Almon lag:

$$b(j; \theta) = \frac{\exp(\theta_1 j + \dots + \theta_Q j^Q)}{\sum_{j=1}^K \exp(\theta_1 j + \dots + \theta_Q j^Q)} \quad (2.2)$$

Note that the weighting function of the exponential Almon lag implies that the weights are always positive. In the empirical applications, we employ the exponential Almon lag scheme with two parameters $\theta = \{\theta_1, \theta_2\}$.

2.2.1.2 Autoregressive MIDAS

Introducing an autoregressive lag in the MIDAS specification is not straightforward as pointed out by Clements and Galvao (2008), who show that a seasonal response of y to x can

appear. However, this can be done without generating any seasonal patterns if autoregressive dynamics is introduced through a common factor, so that equation 2.1 becomes:

$$y_t = \beta_0 + \lambda y_{t-d} + \beta_1 B(L^{1/m}; \theta)(1 - \lambda L^d)x_{t-h}^{(m)} + \epsilon_t \quad (2.3)$$

2.2.2 Markov-switching MIDAS

2.2.2.1 The model

The basic idea behind Markov-switching models is that the parameters of the underlying data generating process (DGP) depend on an unobservable discrete variable S_t , which represents the probability of being in a different state of the world (see Hamilton (1989)). The basic version of the Markov-switching MIDAS (MS-MIDAS) regression model we propose is:

$$y_t = \beta_0(S_t) + \beta_1(S_t)B(L^{1/m}; \theta)x_{t-h}^{(m)} + \epsilon_t(S_t) \quad (2.4)$$

where $\epsilon_t|S_t \sim NID(0, \sigma^2(S_t))$.

The MS-MIDAS that includes autoregressive dynamics is instead defined as:

$$y_t = \beta_0(S_t) + \lambda y_{t-d} + \beta_1(S_t)B(L^{1/m}; \theta)(1 - \lambda L^d)x_{t-h}^{(m)} + \epsilon_t(S_t) \quad (2.5)$$

The regime generating process is an ergodic Markov-chain with a finite number of states $S_t = \{1, \dots, M\}$ defined by the following transition probabilities:

$$p_{ij} = Pr(S_{t+1} = j | S_t = i) \quad (2.6)$$

$$\sum_{j=1}^M p_{ij} = 1 \forall i, j \in \{1, \dots, M\} \quad (2.7)$$

Here the transition probabilities are constant. This assumption has been originally relaxed by Filardo (1994), who used time-varying transition probabilities modelled as a logistic function, while Kim et al. (2008) model them as a probit function. However, we stick to the assumption of constant transition probabilities to keep the model tractable.

The parameters that can switch are the intercept of the equation, β_0 , the parameter entering before the weighting scheme, β_1 , and the variance of the disturbances, σ^2 . Changes in the intercept β_0 are important since they are one of the most common sources of forecast failure, see e.g. Clements and Hendry (1999). The switch in the parameter β_1 allows the predictive ability of the higher frequency variable to change across the different states of the

world ¹. Besides, we also allow the variance of the disturbances σ^2 to change across regimes. This proves to be useful not only for modelling financial variables but also for applications with macroeconomic variables.

Another attractive feature of the Markov-switching models is that they allow the estimation of the probabilities of being in a given regime. This is relevant, for example, when one wants to predict business cycle regimes. Indeed, studies about the identification and prediction of the state of the economy have gained attention over the last decade (see e.g. Estrella and Mishkin (1998), Berge and Jorda (2009), Stock and Watson (2010), and the literature review in Marcellino (2006)).

2.2.2.2 Estimation and model selection

In the literature, MIDAS models are usually estimated by nonlinear least squares (NLS). However, for implementing the filtering procedure described in Hamilton (1989), we estimate the MS-MIDAS via (pseudo) maximum likelihood. We thus need to make a normality assumption about the distribution of the disturbances, which is not required with the NLS estimation. We aim at maximizing the log-likelihood function given by:

$$L = \sum_{t=1}^T \ln f(y_t | \Omega_{t-1}) \quad (2.8)$$

where $f(y_t | \Omega_{t-1})$ is the conditional density of y_t given the information available up to time $t - 1$, Ω_{t-1} . Note that $f(y_t | \Omega_{t-1})$ can be rewritten as:

$$f(y_t | \Omega_{t-1}) = \sum_{j=1}^M P(S_t = j | \Omega_{t-1}) f(y_t | S_t = j, \Omega_{t-1}) \quad (2.9)$$

The computations are carried out with the optimization package OPTMUM of GAUSS 7.0 using the BFGS algorithm. Appendix 2.6.1 provides more details about the estimation method we use.

Choosing the number of regimes for Markov-switching models is a tricky problem. Indeed, the econometrician has to deal with two problems: first, some parameters are not identified under the null hypothesis and, second, the scores are identically equal to zero under the null. Hansen (1992) considers the likelihood function as a function of unknown parameters and

¹Galvao (2009) proposed a regression model (STMIDAS) that captures changes in β_1 with a smooth transition function. This so-called STMIDAS model performs well for the prediction of the US GDP using financial variables as high frequency data.

uses empirical processes to bound the asymptotic distribution of a standardized likelihood ratio test statistic. Garcia (1998) pointed out that the test is computationally expensive if the number of parameters and regimes is high. Carrasco et al. (2009) recently proposed a new method for testing the constancy of parameters in Markov-switching models. Their procedure is attractive since it only requires to estimate the model under the null hypothesis of constant parameters. However, this testing procedure does not allow one to discriminate between Markov-switching models with different number of regimes since the parameters must be constant under the null hypothesis.

Psaradakis and Spagnolo (2006) study the performance of information criteria based on the optimization of complexity-penalized likelihood for model selection. They find that the AIC, SIC and HQ criteria perform well for selecting the correct number of regimes and lags as long as the sample size and the parameter changes are large enough. Smith et al. (2006) propose a new information criterion for selecting simultaneously the number of variables and lags of the Markov-switching models. However, both studies run their analysis with models where all parameters switch across regimes, which might not always be desirable. For example, in equations 2.4 and 2.5, we do not consider switches in the vector of parameters θ since we encountered serious convergence problems in the empirical applications due to the relatively small size of our sample ($T=200$). However, with larger sample sizes, the MS-MIDAS model could easily accommodate changes in the θ vector. In addition, Driffill et al. (2009) show that a careful study of the parameters that can switch is crucial for forecasting accurately bond prices with the CIR model for the term structure.

In the empirical part, we will follow Psaradakis and Spagnolo (2006) and use the Schwarz information criterion for selecting the number of regimes and deciding whether the variance of the disturbances should also change across regimes. We will also report results for different parameterizations of the Markov-switching models.

2.3 Monte Carlo experiments

2.3.1 In-sample estimates

The first purpose of the Monte Carlo experiments is to assess the accuracy of the maximum likelihood estimation procedure we propose for the MS-MIDAS model. The DGP used in the Monte Carlo experiments is the MSHAR(2)-MIDAS model defined by equations 2.5 to 2.7, i.e. it is a model with two regimes and switches in the intercept β_0 , in the parameter entering before the weighting function β_1 and in the variance σ^2 , since models with two

regimes are often used in the literature. We consider two sample sizes for the simulated series $T = 200$ and $T = 500$. The matrix of explanatory variables includes a constant and the process for $x_t^{(m)}$ is an AR(1) with a large autoregressive coefficient (0.95) and a small drift (0.025). We are primarily interested in the predictive content of monthly variables for forecasting quarterly variables, so we set $K = 3$ and $K = 13$. We use the following true parameter values:

$$(\beta_{0,1}, \beta_{0,2}) = (-1, 1), (\beta_{1,1}, \beta_{1,2}) = (0.6, 0.2), (\sigma_1, \sigma_2) = (1, 0.67) \quad (2.10)$$

$$(\theta_1, \theta_2) = (2 * 10^{-1}, -3 * 10^{-2}) \quad (2.11)$$

These parameter values are similar to those used in Kim et al. (2008) and closely match the in-sample parameter estimates of our empirical application for the UK (see Table 2.11 in the appendix). The transition probabilities are first set such that both regimes are equally persistent ($p_{11} = 0.95$, $p_{22} = 0.95$). We also consider another set of transition probabilities: $p_{11} = 0.85$ and $p_{22} = 0.95$. Indeed, with these transition probabilities, if one thinks of y_t as quarterly observations, the duration of the first regime (6.67 quarters) is lower than the duration of the second regime (20 quarters), which roughly corresponds to the average duration of recessions and expansions experienced by the US and the UK.

We first simulate a Markov chain with two regimes using one of the two sets of transition probabilities. The dependent variable y_t is then constructed depending on the outcome of the simulated Markov chain using the above parameter values and the simulated series for x_t . The first 100 data points are discarded to eliminate start-up effects². We repeat the estimation 1000 times and report the means of the maximum likelihood point estimates³. In addition, we report the standard deviations of the point estimates from the true parameter values.

We do not show the point estimates for θ_1 and θ_2 but rather the approximation error computed as the sum of the squared error between the estimated and the true weighting function, normalized by the squared weights of the true weighting function. We proceed this way since it is the shape of the weighting function which is important rather than the point

²Discarding more than 100 initial observations leads to identical results.

³Note that we do not initialize the algorithm with the true parameter values. Instead, we use the same rule of thumb than in the empirical applications for the initialization of the parameters (i.e. we run OLS regressions on sub-samples after sorting the x_t variable with respect to the dependent variable y_t).

estimates for θ_1 and θ_2 . The approximation error is defined by:

$$\frac{\sum_{j=1}^{M=K} [b(j, \hat{\theta}) - b(j, \theta)]^2}{\sum_{j=1}^{M=K} b(j, \theta)^2} \quad (2.12)$$

Table 2.1 shows that the parameter estimates for the intercepts $\beta_{0,1}$ and $\beta_{0,2}$, the autoregressive parameter λ and the transition probabilities p_{11} and p_{22} are very close to their true values. The estimates for $\beta_{1,1}$ and $\beta_{1,2}$ - the parameters entering before the weighting function - are slightly downward biased. The standard deviations for the estimates are lower when the sample size is large, as expected. Similarly, the shape of the weighting function is better approximated for $T = 500$ and the average R^2 is higher.

Overall, the Monte Carlo experiments suggest that maximum likelihood estimation of this specification of the MSHAR(2)-MIDAS provides accurate estimates of the model parameters, including the transition probabilities.

2.3.2 Forecasting exercise

We carry out another Monte Carlo experiment to assess the forecasting accuracy of the MS-MIDAS model. To this end, we generate data from the D.G.P. used in the previous subsection with the parameter values defined in equations 2.10 and 2.11 with a sample size of $T=200$ and $T=500$ using 13 lags for the high frequency indicator x_t^m . The sample size T is split between an estimation sample and an evaluation sample. We choose three different sizes H for the evaluation sample, $H = \{20, 50, 100\}$. We then run the following out-of-sample forecasting experiment: we use the first $T-H$ observations and compute one-step ahead forecasts ⁴. We recursively expand the estimation sample until we reach the end of the sample T so that we compute H forecasts. The design of this forecasting experiment is very close to the empirical application we run later in the paper.

We use seven different models to compute the forecasts: the MSHAR(2)-MIDAS model (i.e. the true model), the MSH(2)-MIDAS model (i.e. the true model without an autoregressive lag), a standard MIDAS and AR-MIDAS models as defined in equations 2.1 and 2.3 respectively. We also consider an AR(1) model and a standard Markov-switching model with two regimes, a switch in the intercept and in the variance of the disturbances and one autoregressive lag (i.e. MSIHAR(2) model). Finally, we also show the results for an

⁴Note that we use a different estimation sample for each H in order to consider the common trade-off in empirical analysis between a longer estimation sample or a longer evaluation sample. Our results suggest that, as long as the estimation sample remains long enough, a longer evaluation sample is to be preferred to a longer estimation sample.

Table 2.1: Monte Carlo Results, In-Sample Estimates

		$p_{11} = 0.95$	$p_{11} = 0.95$	$\beta_{0,1} = -1$	$\beta_{0,2} = 1$	$\beta_{1,1} = 0.6$	$\beta_{1,2} = 0.2$	$\lambda = 0.2$	R^2	Approx. error
<hr/>										
K=3										
	T=200	0.936 (0.039)	0.946 (0.032)	-1.043 (0.224)	1.042 (0.140)	0.504 (0.103)	0.170 (0.056)	0.165 (0.074)	0.769	0.262
	T=500	0.944 (0.018)	0.949 (0.016)	-1.042 (0.116)	1.028 (0.074)	0.532 (0.045)	0.181 (0.023)	0.174 (0.042)	0.795	0.157
<hr/>										
K=13										
	T=200	0.935 (0.040)	0.945 (0.028)	-1.047 (0.242)	1.041 (0.137)	0.501 (0.104)	0.174 (0.046)	0.163 (0.074)	0.768	0.256
	T=500	0.944 (0.019)	0.949 (0.016)	-1.036 (0.117)	1.037 (0.077)	0.532 (0.044)	0.181 (0.022)	0.172 (0.044)	0.793	0.174
<hr/>										
		$p_{11} = 0.85$	$p_{11} = 0.95$	$\beta_{0,1} = -1$	$\beta_{0,2} = 1$	$\beta_{1,1} = 0.6$	$\beta_{1,2} = 0.2$	$\lambda = 0.2$	R^2	Approx. error
<hr/>										
K=3										
	T=200	0.808 (0.114)	0.951 (0.025)	-1.015 (0.455)	1.030 (0.116)	0.455 (0.211)	0.176 (0.033)	0.178 (0.069)	0.734	0.237
	T=500	0.832 (0.053)	0.952 (0.014)	-1.025 (0.185)	1.026 (0.065)	0.506 (0.081)	0.180 (0.017)	0.181 (0.040)	0.769	0.133
<hr/>										
K=13										
	T=200	0.816 (0.097)	0.950 (0.028)	-0.993 (0.423)	1.035 (0.121)	0.451 (0.193)	0.175 (0.032)	0.176 (0.070)	0.730	0.234
	T=500	0.833 (0.051)	0.951 (0.014)	-1.032 (0.191)	1.022 (0.065)	0.507 (0.076)	0.181 (0.017)	0.185 (0.041)	0.772	0.149
<hr/>										

This table reports the average of the 1000 point estimates of the Monte Carlo experiments. The last column reports the average approximation error for the weighting scheme as defined by equation 2.12.

Standard deviations of the 1000 point estimates are reported in brackets.

MSIHAR(2)-MIDAS model (i.e. a model with a constant β_1). We always use the true number of lags ($K = 13$) for the high frequency variable x_t^m when the models have mixed-frequency data.

We repeat the forecasting experiment N times for each evaluation sample and report in Table 2.2 the average of the mean square forecast error over the number of replications for the seven models under consideration. We also report the Quadratic Probability Score (QPS) and Log Probability Score (LPS) for the models with Markov-switching features in order to check how well these models can predict the true regimes. Here, the MSFE is a criterion that allows us to assess the quantitative forecasting abilities of the model under scrutiny, whereas QPS and LPS criteria evaluate their qualitative forecasting abilities, i.e. to what extent the true regimes are predicted.

Note that the QPS is bounded between 0 and 2 and the range of LPS is 0 to ∞ . LPS penalizes large forecast errors more than QPS. LPS and QPS are computed as follows:

$$QPS = \frac{2}{H} \sum_{t=1}^H (P(S_{t+1} = 1) - S_{t+1})^2 \quad (2.13)$$

$$LPS = -\frac{1}{H} \sum_{t=1}^H (1 - S_{t+1}) \log(1 - P(S_{t+1} = 1)) + S_{t+1} \log(P(S_{t+1} = 1)) \quad (2.14)$$

where S_{t+1} is a dummy variable that takes on a value of 1 if the true regime is the first regime and $P(S_{t+1} = 1)$ is the predicted probability of being in the first regime in period $t+1$.

For $T = 200$, Table 2.2 shows that the true model (i.e. the MSHAR(2)-MIDAS) gets the best results for the MSFE, QPS and LPS when the size of the evaluation sample is $H = 100$. The MSH(2)-MIDAS model obtains the best results for $H = 50$ in terms of QPS and LPS, while the AR(1) model gets the lowest MSFE. For $H = 20$, the true model yields the best performance in terms of LPS and MSFE but it is slightly outperformed by the MSIHAR model according to the QPS criterion.

For $T = 500$ and $H = 100$, the true model obtains the best performance for both discrete and continuous forecasts. For $H = 50$, the true model obtains the best results in terms of MSFE, whereas it is outperformed by the MSH(2)-MIDAS model in terms of QPS and LPS criteria. Finally, for $H = 20$, the MSIHAR(2) model obtains the best regime forecasts but it is outperformed by the MSIHAR(2)-MIDAS according to the MSFE criterion.

Overall, the true MS-MIDAS model is either ranked first or exhibits a performance very close to the best model for both discrete and continuous forecasts. Besides, given the DGP

we use, the simple MIDAS model has a poor forecasting performance as compared to the AR-MIDAS model. Finally, the MS and the AR(1) models yield rather inaccurate continuous forecasts.

2.4 An application to the prediction of quarterly GDP

2.4.1 Prediction of the US GDP

2.4.1.1 In-sample results

We analyze quarterly data for the US GDP, taken from the real-time dataset of the Philadelphia Federal Reserve⁵, which originates from the work of Croushore and Stark (2001). Quarterly vintages reflect the information available in the middle month of each quarter. The dependent variable is taken as 100 times the quarterly change in the log of the US real GDP from $t=1959:Q1$ to $2009:Q4$. For the in-sample analysis, we use the 2010:Q1 vintage.

We first consider the slope of the yield curve as high frequency indicator since its predictive power for GDP growth has been widely documented (Estrella and Hardouvelis (1991), Galvao (2006), Rudebusch and Williams (2009)). We use the difference between the 10-year Treasury bond and the 3-month Treasury-bill as a proxy for the slope of the yield curve. We also consider stock returns as a monthly indicator for forecasting quarterly aggregate economic activity. Stock returns are taken as 100 times the monthly change in the log of the S&P500 index. We finally consider the Federal Funds as a monthly indicator to take into account the stance of the monetary policy, which is often considered as an important determinant of economic activity. We take the first difference for both the slope of the yield curve and the Federal Funds since we achieve better forecasting results with this transformation. The data for the 10-year Treasury bond yields, the 3-month Treasury bill and the Federal Funds are taken from the Federal Reserve website, while the data for the S&P500 index are downloaded from Yahoo Finance.

For selecting the number of regimes and whether there is also switching in the variance of the disturbances, we use the SIC with a maximum number of regimes of $M = 3$. For $M = \{2, 3\}$, we then estimate a model with or without a switch in the variance and in the parameter β_1 . Whatever indicator we use, the model that gets the best fit has three regimes and switches in the intercept β_0 , in the parameter entering before the weighting function β_1 and in the variance σ^2 .

⁵The real-time vintage quarterly data for the US GDP are available at <http://www.philadelphiafed.org/research-and-data/real-time-center/real-time-data/data-files/ROUTPUT/>

Table 2.2: Monte Carlo Results: Forecasting exercise

Number of out-sample forecasts H :	20			50			100		
<hr/>									
<i>First estimation</i>									
<i>sample size</i>	QPS	LPS	MSFE	QPS	LPS	MSFE	QPS	LPS	MSFE
MSHAR(2)-MIDAS	0.398	0.609	1.432	0.300	0.447	1.782	0.295	0.464	1.207
MSIHAR(2)	0.392	0.642	1.709	0.370	0.584	1.617	0.872	1.690	1.301
MSH(2)-MIDAS	0.412	0.635	1.468	0.258	0.398	1.860	0.304	0.469	1.226
MSIHAR(2)-MIDAS	0.444	0.683	1.646	0.356	0.561	1.671	0.821	1.629	1.237
AR-MIDAS	-	-	1.563	-	-	1.722	-	-	1.423
MIDAS	-	-	2.510	-	-	2.432	-	-	1.738
AR(1)	-	-	1.596	-	-	1.563	-	-	1.456
<i>Second estimation</i>									
<i>sample size</i>									
MSHAR(2)-MIDAS	0.456	0.653	1.385	0.238	0.404	1.487	0.387	0.581	1.121
MSIHAR(2)	0.358	0.543	1.787	0.365	0.547	2.034	0.628	0.861	1.622
MSH(2)-MIDAS	0.524	0.736	1.442	0.216	0.374	1.519	0.403	0.605	1.183
MSIHAR(2)-MIDAS	0.361	0.546	1.326	0.326	0.498	1.625	0.536	0.758	1.342
AR-MIDAS	-	-	1.410	-	-	1.711	-	-	1.469
MIDAS	-	-	1.893	-	-	2.202	-	-	1.549
AR(1)	-	-	1.832	-	-	1.931	-	-	1.668

This table reports the average QPS, LPS and MSFE over 200 Monte Carlo replications. In the first estimation sample, the initial estimation sample size is $T - H$ where $T = 200$. In the second estimation sample, the initial estimation sample size is $T - H$ where $T = 500$. Both estimation samples are recursively expanded until the end of the sample is reached. Entries in bold outline the model with the lowest QPS, LPS or MSFE for each combination of the evaluation sample H and sample size T . The true model is the MSHAR(2)-MIDAS model. A classification of the models is provided in Table 2.14.

Table 2.10 in the appendix reports the in-sample results for each indicator for the models with three regimes, a switch in the variance of the disturbances, in the intercept and with or without a switch in β_1 . Note that the intercept in the first regime $\beta_{0,1}$ is negative, whereas the intercept in the second regime $\beta_{0,2}$ is positive, but smaller than the intercept in the third regime $\beta_{0,3}$. Therefore, the first regime can be interpreted as the recessionary regime, the second regime is instead the low but positive growth regime, while the third regime is the strong growth regime. As expected, the variance in the second regime is always the lowest among the three regimes. Moreover, there are noticeable differences across regimes in the coefficient β_1 , which measures the impact of the monthly indicators on quarterly GDP growth, while β_1 remains statistically significant in most cases. These results highlight the importance of allowing for parameter changes in MIDAS models, but also the relevance of including high frequency information in MS models.

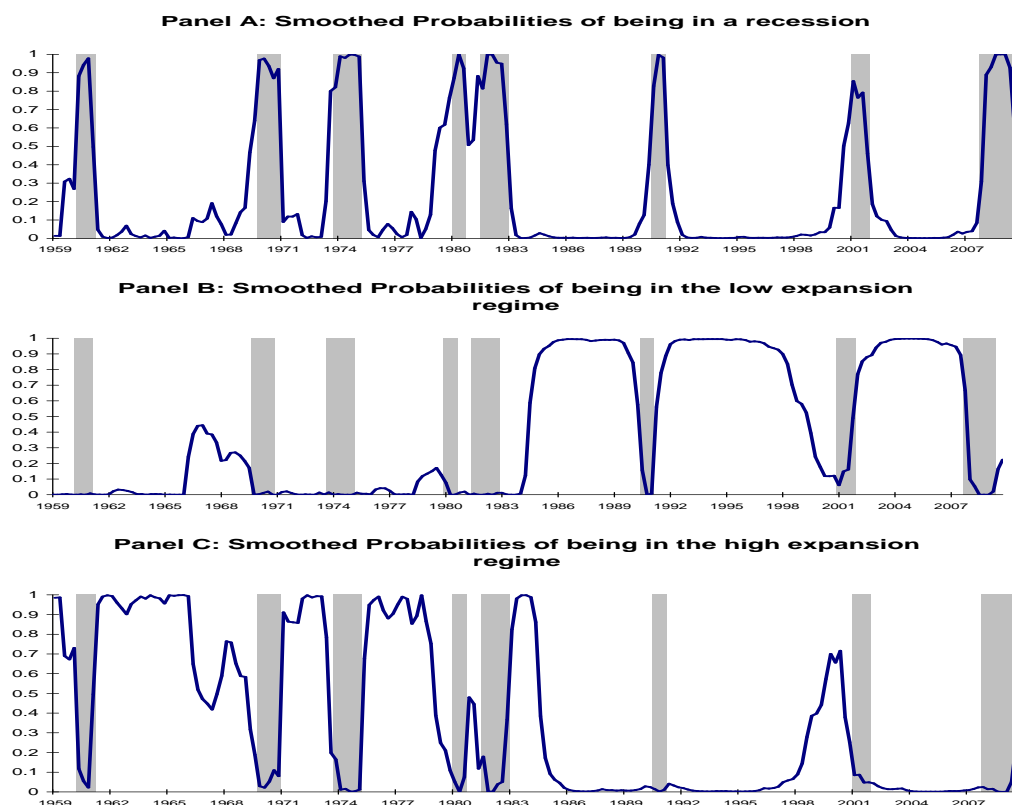
Figure 2.1 depicts the estimated smoothed probabilities resulting from the MSIHAR(3)-MIDAS model with the slope of the yield curve as a monthly indicator. The shadow areas represent the recessions identified by the National Bureau of Economic Research (NBER). First, one can see that the estimated probabilities of recession match quite well the actual recessions, including the recession that started in December 2007 (panel A). Interestingly, the probability of recession falls in the third quarter of 2009, which confirms the NBER dating of the end of the last recession. The first (moderate) expansion regime - depicted in panel B - is predominant in the post-1984 era and is characterized by a much lower variance than the second (stronger) expansion regime reported in Panel C. This finding is in line with the great moderation phenomenon and supports the McConnell and Perez-Quiros (2000) dating of the break in volatility experienced by the US. Panel C reports the estimated probabilities of being in a high growth regime, this regime is predominant in the 1960s and 1970s and is shortly resurgent in the late 1990s, reflecting the high growth experienced by the US thanks to the technology boom.

Consequently, the MS-MIDAS model with three regimes seems to be a proper specification for describing quarterly US GDP, and its forecasting performance will be assessed in the next subsection.

To conclude, Figure 2.2 plots the estimated weights corresponding to the three monthly indicators for the AR-MIDAS and MSHAR(3)-MIDAS models ⁶. The figure illustrates the variety of weights that can be attached to the indicators thanks to the MIDAS specification.

⁶The MSHAR(3)-MIDAS model is a model with three regimes, an autoregressive parameter and switches in β_0 , β_1 and in the variance σ^2 . Table 2.14, which is the last one in the Appendix, reports the labels we used for each model.

Figure 2.1: MS-MIDAS, QUARTERLY GDP AND MONTHLY SLOPE OF THE YIELD CURVE, 1959:Q1-2009:Q4

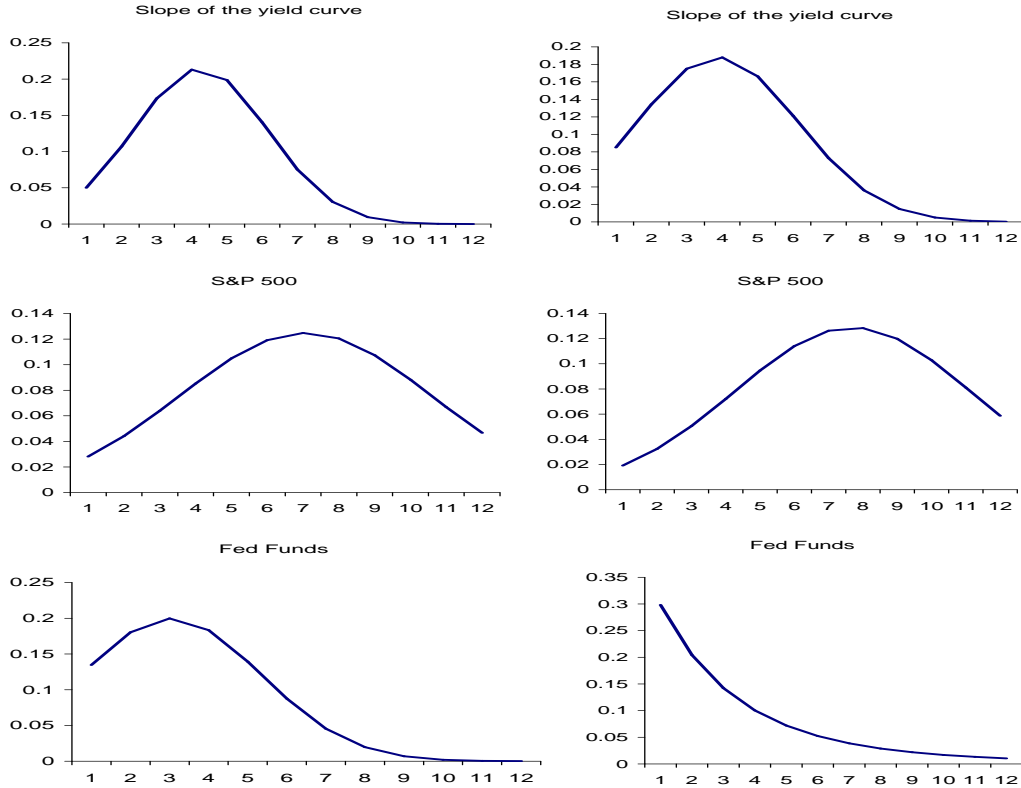


2.4.1.2 Design of the real-time forecasting exercise

The sample is split into an estimation sample and an evaluation sample. The evaluation sample consists of quarterly GDP growth in the quarters 1998:Q1 to 2009:Q4. For each of these quarters, we generate forecasts with horizons $h = \{0, 1/3, 2/3, 1, 4/3, 5/3, 2\}$. The initial estimation sample goes from 1959:Q1 to 1997:Q4 and is recursively expanded over time until 2009:Q2 ⁷. The design of the exercise is similar to the one described in section 3.1.1 of Clements and Galvao (2008). We denote $y_{\tau,\nu}$ as output growth in period τ released in the vintage ν data set. We aim at forecasting final estimates of the output growth $y_{t,T}$ as defined in the latest vintage available to us $T = 2010:Q1$. Note that for GDP, the vintage

⁷The last observation used in the estimation sample is 2009:Q2 since we need the actual values of GDP for the next two quarters to compute the MSFE. Therefore, there are 47 forecasts computed for each forecast horizon h .

Figure 2.2: WEIGHTS OF THE AR-MIDAS (LHS) AND MSHAR(3)-MIDAS (RHS), EXPONENTIAL LAG POLYNOMIAL



data set released in quarter $t + 1$ contains data up to quarter t , and quarterly vintages reflect information available in the middle month of each quarter. We use financial variables as higher frequency variables, which are available without any delays and are not subject to data revisions.

A few additional comments are required. First, forecasts for the regime probabilities k quarters ahead are computed recursively as:

$$P(S_{t+k} = j) = \sum_{i=1}^M p_{ij} P(S_{t+k-1} = i) \quad (2.15)$$

Note that the predicted probabilities only depend on the transition probabilities and on the filtered probabilities.

Second, forecasts with an horizon $h = 0$ (i.e. nowcasts) imply that we want to forecast

output growth for the current quarter knowing the values of the monthly indicators for all months of the current quarter. The nowcasts are computed as follows: we first regress $y_{t|t+1}$ on $B(L^{(1/3)}; \theta)x_{t|t+1}$ and $y_{t-1|t+1}$, where $y_{t|t+1} = [y_{1|t+1}, y_{2|t+1}, \dots, y_{t-1|t+1}, y_{t|t+1}]$ and $x_{t|t+1} = [x_{1|t+1}, \dots, x_{t-1|t+1}, x_{t|t+1}]$. We then use these estimates, the forecasts for the regime probabilities $P(S_{t+1} = j|x_t)$, $y_{t|t+1}$ and $x_{t+1|t+1}$ to compute the forecasts $\hat{y}_{t+1|t+1}$.

Forecasts with an horizon $h = 1/3$ imply that we only know the values for the first two months of the monthly indicator. To obtain these forecasts, we first regress $y_{t|t+1}$ on $B(L^{(1/3)}; \theta)x_{t-1/3|t+1}$ and $y_{t-1|t+1}$, where $x_{t-1/3|t+1} = [x_{1-1/3|t+1}, \dots, x_{t-4/3|t+1}, x_{t-1/3|t+1}]$. We then use these estimates, the forecasts for the regime probabilities $P(S_{t+1} = j|x_t)$, $y_{t|t+1}$ and $x_{t+2/3|t+1}$ to obtain forecasts for y_{t+1} , which is conditioned on $x_{t+2/3|t+1}$ and $y_{t|t+1}$. Forecasts with an horizon $h = 4/3$ are generated from a regression of $y_{t|t+1}$ on $B(L^{(1/3)}; \theta)x_{t-4/3|t+1}$ and $y_{t-2|t+1}$.

Similarly, forecasts with an horizon $h = 2/3$ imply that we only know the values for the first month of the monthly indicator. To obtain these forecasts, we first regress $y_{t|t+1}$ on $B(L^{(1/3)}; \theta)x_{t-2/3|t+1}$ and $y_{t-1|t+1}$ where $x_{t-2/3|t+1} = [x_{1-2/3|t+1}, \dots, x_{t-5/3|t+1}, x_{t-2/3|t+1}]$. We then use these estimates, the forecasts for the regime probabilities $P(S_{t+1} = j|x_t)$ and $x_{t+1/3|t+1}$ to obtain forecasts for y_{t+1} , which is conditioned on $x_{t+1/3|t+1}$. Forecasts with an horizon $h = 5/3$ are generated from a regression of $y_{t|t+1}$ on $B(L^{(1/3)}; \theta)x_{t-5/3|t+1}$ and $y_{t-2|t+1}$.

Finally, forecasts with an horizon $h = 1$ are computed from the regression of $y_{t|t+1}$ on $B(L^{(1/3)}; \theta)x_{t-1|t+1}$ and $y_{t-1|t+1}$, while forecasts with an horizon $h = 2$ come from the regression of $y_{t|t+1}$ on $B(L^{(1/3)}; \theta)x_{t-2|t+1}$ and $y_{t-2|t+1}$. Hence, for example, forecasts for the first quarter Q1 of a given year are generated as described in Table 2.3.

Table 2.3: Forecasting scheme for Q1

Forecast horizon h	0	1/3	2/3	1	4/3	5/3	2
Data up to month	<i>March_t</i>	<i>Feb_t</i>	<i>Jan_t</i>	<i>Dec_{t-1}</i>	<i>Nov_{t-1}</i>	<i>Oct_{t-1}</i>	<i>Sept_{t-1}</i>

2.4.1.3 Out-of-sample results

Table 2.4 reports the relative mean square forecast errors (MSFE) for seven different models for different forecast horizons using an AR(1) model as a benchmark⁸. The MIDAS model is the standard MIDAS as defined in equation 2.1. The AR-MIDAS is the model defined in equation 2.3. The MSIH(3)-MIDAS is a model with three regimes, a switch in the intercept and in the variance of the shocks. The MSIHAR(3)-MIDAS is an MSIH(3)-MIDAS model with an autoregressive lag introduced through a common factor as described in Section 2.2. The MSH(3)-MIDAS and MSHAR(3)-MIDAS are similar to the MSIH(3)-MIDAS and MSIHAR(3)-MIDAS apart from the fact that they also include a switch in the parameter β_1 . We also report results for a standard Markov-switching model with three regimes, one autoregressive lag, a switch in the intercept and in the variance (MSIHAR(3) model). The number of lags included in the weighting function is selected using the SIC.

Table 2.4 reports the out-of-sample forecasting results for the period 1998:Q1 to 2009:Q4. Note first that the AR-MIDAS always outperforms the MIDAS with the Federal Funds and slope of the yield curve as a monthly indicator. When using stock prices, AR-MIDAS and MIDAS yield comparable forecasting performance. In the Markov-switching case, including an autoregressive lag seems to be of less importance. Second, the S&P500 index is the best indicator among the three variables considered and it also largely outperforms the AR(1) model. For forecast horizons $h = \{0, 1/3, 2/3, 1\}$, the MSIHAR-MIDAS model with stock prices turns out to be the best model for predicting quarterly GDP growth across all models under consideration. Third, the slope of the yield curve exhibits a poor forecasting performance as compared to the AR(1) model: this is a disappointing result but it is in line with the findings of Galvao (2009). Fourth, within the class of models that use the Federal Funds as a monthly indicator, the AR-MIDAS yields a better forecasting performance than models with Markov-switching features for $h = \{0, 2/3, 1\}$, while the MSHAR(3)-MIDAS model is the best forecasting model for $h = \{1/3, 4/3, 2\}$. Finally, the standard Markov-switching model is slightly better than the AR(1) for two-quarter ahead predictions but slightly worse for one-quarter ahead predictions. It is beaten by several MS-MIDAS

⁸We do not report tests of equal forecast accuracy since it is not straightforward to implement them in the context of nested MIDAS models with real-time data. Indeed it is uncertain whether the test proposed by Clark and McCracken (2009a) for nested models with real-time data can be applied in the context of MIDAS models. Furthermore, the test for nested models by Clark and McCracken (2005) is computationally very expensive since Monte Carlo simulations should be undertaken for each model and forecast horizon. In addition, the usual approach of adopting the Giacomini and White (2006) test combined with rolling estimation is not suited in our context, since we want to use all the available sample in each point in time to improve inference on the regimes. Hence we leave the issue of testing equal forecast accuracy for future research.

specifications, which confirms the usefulness of introducing higher frequency information into MS models.

In addition to predicting quarterly GDP growth, Markov-switching MIDAS models can endogenously generate probabilities of being in a given regime. Table 2.5 and Table 2.6 below provide the quadratic probability score (QPS) and the log probability score (LPS) as defined in equations 2.13 and 2.14. We use the classification of the economic activity from the NBER so that S_t is a dummy variable that takes on a value of 1 if the economy is in recession in quarter t according to the NBER, while $P(S_t = 1)$ is the probability of being in the recession regime in period t . Forecasts with an horizon $h = \{0, 1/3, 2/3, 1\}$ predict the regime of the economy one quarter ahead, while forecasts with an horizon $h = \{4/3, 5/3, 2\}$ predict the state of the economy two quarters ahead.

In contrast to forecasting the level of GDP growth, the slope of the yield curve and the Federal Funds tend to better predict the state of the economy than stock prices. This is in line with the results from binary recession models that emphasize the predictive power of the slope of the yield curve (see e.g. Estrella and Mishkin (1998)). Moreover, using information from the monthly indicators produces better regime forecasts than a pure MS model for quarterly GDP.

Additional evidence on the predictive ability of the Markov-switching MIDAS specification is presented in Figure 2.3, where we report the nowcasted probability of being in a recession with the slope of the yield curve as a monthly indicator using the MSH(3)-MIDAS model with $h = 0$. These probabilities are generated from the recursive exercise and correspond to the filtered probabilities for the last observation T , where T is recursively expanded over time from $t=1997:Q4$ to $2009:Q4$.

Figure 2.3 shows that there is a first signal of recession in the second quarter of 2001 using information up to August 2001. The probability of recession then rises above .90 in the third and fourth quarter of 2001. Interestingly, the probability of recession stays above .35 until the second quarter of 2003 and only fall below .10 in the third quarter of 2003. This illustrates the slow economic recovery that followed the 2001 recession. Figure 2.3 also shows that there is a first peak in the probability of recession in the fourth quarter of 2007 using information up to February 2008. The probability of recession then jumps above .85 from the end of the first quarter of 2008 until the second quarter of 2009 and it starts declining in the third quarter of 2009. This confirms the NBER dating of the end of the last recession in June 2009. Note that our model gives the first signal of recession well before the announcement of the recession by the NBER that occurred in December 2008.

A crucial point of the MS-MIDAS specification is that the quarterly probabilities of

Table 2.4: Relative Mean Squared Forecast error for forecasting US GDP growth 1998:Q1-2009:Q4

		Forecast horizon (h)						
	Model	0	1/3	2/3	1	4/3	5/3	2
Slope of the yield curve	MSIH(3)-MIDAS	1.080	1.061	1.066	1.066	1.022	1.002	1.009
	MSIHAR(3)-MIDAS	0.960	1.019	1.038	0.964	1.032	1.049	1.041
	MSH(3)-MIDAS	1.080	1.043	1.059	1.025	1.067	1.073	1.049
	MSHAR(3)-MIDAS	1.073	1.073	1.044	0.992	0.996	1.019	1.015
	AR-MIDAS	1.037	1.022	1.022	1.010	0.995	0.997	0.997
	MIDAS	1.242	1.231	1.231	1.252	1.065	1.069	1.276
S&P 500	MSIH(3)-MIDAS	0.691	0.739	0.712	0.765	0.740	0.784	0.785
	MSIHAR(3)-MIDAS	0.680	0.690	0.639	0.726	0.775	0.776	0.792
	MSH(3)-MIDAS	0.913	0.876	0.881	1.078	0.872	0.902	0.861
	MSHAR(3)-MIDAS	0.871	1.160	0.932	1.128	0.873	0.885	0.845
	AR-MIDAS	0.769	0.726	0.715	0.746	0.715	0.754	0.779
	MIDAS	0.762	0.728	0.713	0.729	0.684	0.727	0.767
Fed Funds	MSIH(3)-MIDAS	0.909	0.996	0.983	1.005	1.074	1.086	1.284
	MSIHAR(3)-MIDAS	0.944	0.952	0.995	1.044	1.009	1.046	1.185
	MSH(3)-MIDAS	1.041	0.916	0.963	0.997	1.126	1.110	1.119
	MSHAR(3)-MIDAS	0.974	0.886	0.983	0.971	0.966	1.172	1.009
	AR-MIDAS	0.874	0.942	0.945	0.945	1.081	1.124	1.278
	MIDAS	1.041	1.078	1.171	1.174	1.237	1.237	1.442
	MSIHAR(3)	-	-	-	1.071	-	-	0.984

Real-time data set. Relative Mean Squared Forecast Error for US output growth in the quarters 1998:Q1-2009:Q4. Benchmark: AR(1) model. Recursive forecasting scheme. Entries in bold outline the model with the lowest MSFE for each indicator and forecast horizon. A classification of the models is reported in Table 2.14.

Table 2.5: Quadratic Probability Score for forecasting US business cycle regimes 1998:Q1-2009:Q4

		Forecast horizon (h)						
	Model	0	1/3	2/3	1	4/3	5/3	2
Slope of the yield curve	MSIH(3)-MIDAS	0.391	0.395	0.395	0.395	0.436	0.448	0.469
	MSIHAR(3)-MIDAS	0.384	0.374	0.374	0.383	0.433	0.481	0.468
	MSH(3)-MIDAS	0.270	0.296	0.297	0.407	0.485	0.461	0.402
	MSHAR(3)-MIDAS	0.270	0.287	0.306	0.370	0.448	0.494	0.491
S&P 500	MSIH(3)-MIDAS	0.399	0.449	0.464	0.439	0.465	0.437	0.457
	MSIHAR(3)-MIDAS	0.434	0.436	0.443	0.408	0.520	0.519	0.509
	MSH(3)-MIDAS	0.442	0.414	0.382	0.412	0.473	0.495	0.517
	MSHAR(3)-MIDAS	0.391	0.287	0.410	0.353	0.456	0.466	0.635
Fed Funds	MSIH(3)-MIDAS	0.391	0.375	0.388	0.391	0.385	0.390	0.463
	MSIHAR(3)-MIDAS	0.394	0.401	0.400	0.385	0.459	0.462	0.435
	MSH(3)-MIDAS	0.415	0.369	0.374	0.337	0.483	0.389	0.495
	MSHAR(3)-MIDAS	0.423	0.403	0.436	0.411	0.489	0.502	0.445
	MSIHAR(3)	-	-	-	0.357	-	-	0.452

Entries in bold outline the model with the lowest QPS for each indicator and forecast horizon. QPS is computed as follows:

$$QPS = \frac{2}{F} \sum_{t=1}^T (P(S_{t+h} = 1) - NBER_{t+h})^2$$

where F is the number of forecasts, $P(S_{t+h})$ are the predicted regime probabilities of being in the first regime and $NBER_{t+h}$ is a dummy variable that takes on a value of 1 if the US economy is in recession in quarter $t+h$ according to the NBER. For $h=\{0,1/3,2/3,1\}$, we predict business cycle regimes one quarter ahead, whereas for $h=\{4/3,5/3,2\}$ we predict business cycle regimes two quarters ahead. A classification of the models is reported in Table 2.14.

recession can be updated on a monthly basis (i.e. at the frequency of the x_t^m variable). This makes this class of models very attractive for real-time estimation of business cycle conditions. Indeed, Table 2.12 in the appendix reports the nowcasted probability of recession for the MSH(3)-MIDAS model with the slope of the yield curve as an indicator for three different forecast horizons $h=\{0,1/3,2/3\}$. The table confirms that using the slope of the yield curve as a monthly indicator provides strong calls of recession in the first quarter of 2008 since the probability of recession gradually increased in the first quarter of 2008 to reach .87 in March 2008 (using information available up to May 2008). Table 2.12 also shows that the probability of recession is decreasing in the third quarter of 2009 in line with the NBER datation of the end of the last recession. However, the probability of recession remains fairly high in the third and fourth quarters of 2009, reflecting the moderate growth path experienced by the US.

In summary, Markov-switching MIDAS models not only generate good forecasting results for the level of GDP growth, but they also provide relevant information about the state of the economy. The combination of high frequency information and parameter switching performs better than using each of these two features separately, as in standard MIDAS and MS models, respectively.

2.4.2 Prediction of the UK GDP

2.4.2.1 In-sample results

The data for the UK GDP are taken from the Bank of England Real-Time Database.⁹ We retain only the vintages corresponding to the first estimates of GDP. This database is updated every year following the publication of the ONS Blue Book. For the in-sample analysis, the dependent variable is taken as 100 times the quarterly change in the log of the UK real GDP from $t=1975:Q1$ to $2010:Q1$. We consider comparable predictors as in the application for the US GDP: the slope of the yield curve, the Financial Times All Shares Index and the Bank of England base rate. The slope of the yield curve is taken as the difference between a bond with a 10-year maturity and a bond with a 1.5-year maturity. We applied the same data transformation as in the US case. The data for the FT All Shares Index and the Bank of England base rate are taken from Datastream, while the data for the UK yield curve are taken from the Bank of England database.

We select a model with two regimes and no switch in the variance of the error term since this model matches well the business cycle regimes experienced by the UK (see Figure 2.4).

⁹The Bank of England Real-Time Database is available at: <http://www.bankofengland.co.uk/statistics/gdpdatabase/>

Table 2.6: Log Probability Score for forecasting US business cycle regimes 1998:Q1-2009:Q4

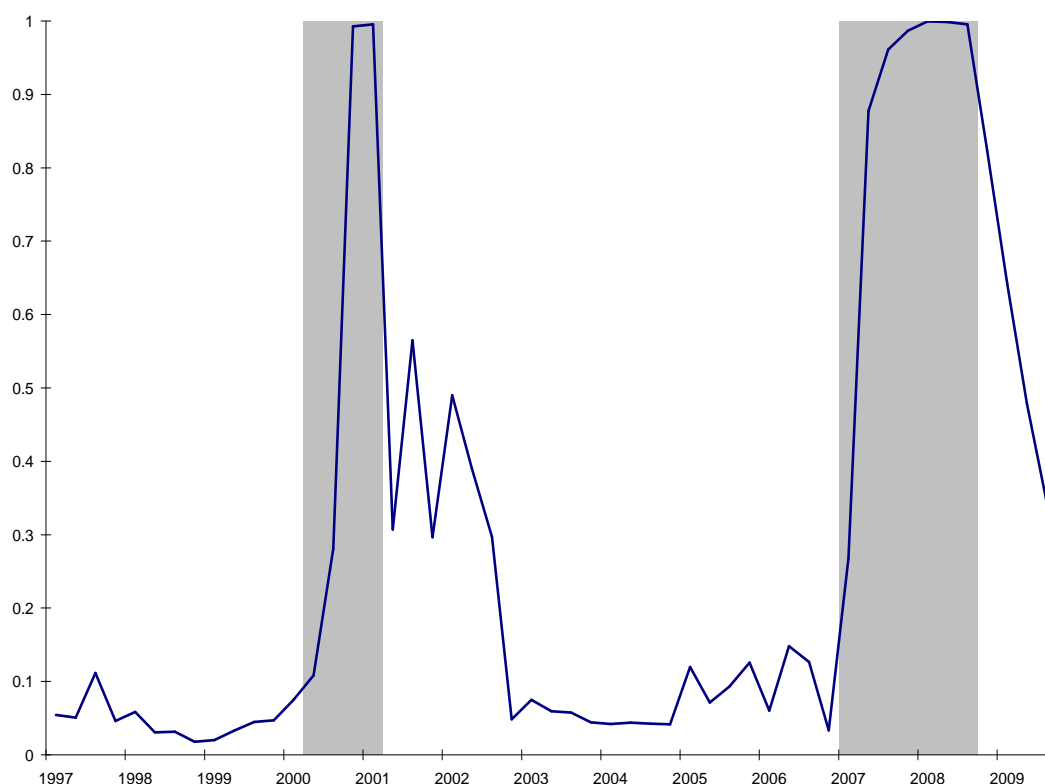
		Forecast horizon (h)						
	Model	0	1/3	2/3	1	4/3	5/3	2
Slope of the yield curve	MSIH(3)-MIDAS	0.739	0.749	0.747	0.748	0.738	0.745	0.775
	MSIHAR(3)-MIDAS	0.708	0.680	0.681	0.693	0.899	0.991	0.963
	MSH(3)-MIDAS	0.437	0.473	0.465	0.690	0.901	0.746	0.604
	MSHAR(3)-MIDAS	0.438	0.454	0.474	0.807	0.841	0.941	0.905
S&P 500	MSIH(3)-MIDAS	0.781	0.924	1.050	0.914	0.838	0.697	0.736
	MSIHAR(3)-MIDAS	0.853	0.856	0.891	0.815	0.981	0.985	0.920
	MSH(3)-MIDAS	0.707	0.756	0.633	0.823	0.723	0.889	1.052
	MSHAR(3)-MIDAS	0.741	0.442	0.714	0.558	0.919	0.923	1.127
Fed Funds	MSIH(3)-MIDAS	0.734	0.621	0.722	0.722	0.581	0.601	0.671
	MSIHAR(3)-MIDAS	0.805	0.736	0.734	0.693	0.959	0.961	0.661
	MSH(3)-MIDAS	0.924	0.616	0.677	0.518	0.768	0.585	0.822
	MSHAR(3)-MIDAS	0.785	0.760	0.870	0.748	0.850	0.850	0.904
	MSIHAR(3)	-	-	-	0.594	-	-	0.875

Entries in bold outline the model with the lowest LPS for each indicator and forecast horizon. LPS is computed as follows:

$$LPS = -\frac{1}{F} \sum_{t=1}^T (1 - NBER_{t+h}) \log(1 - P(S_{t+h} = 1)) + NBER_{t+h} \log(P(S_{t+h} = 1))$$

where F is the number of forecasts, $P(S_{t+h})$ are the predicted regime probabilities of being in the first regime and $NBER_{t+h}$ is a dummy variable that takes on a value of 1 if the US economy is in recession in quarter $t+h$ according to the NBER. For $h=\{0,1/3,2/3,1\}$, we predict business cycle regimes one quarter ahead, whereas for $h=\{4/3,5/3,2\}$ we predict business cycle regimes two quarters ahead. A classification of the models is reported in Table 2.14.

Figure 2.3: ESTIMATED PROBABILITY OF RECESSION, REAL-TIME DATA, 1997:Q4-2010:Q2



Note: MSH(3)-MIDAS model with the monthly slope of the yield curve, forecast horizon $h = 0$

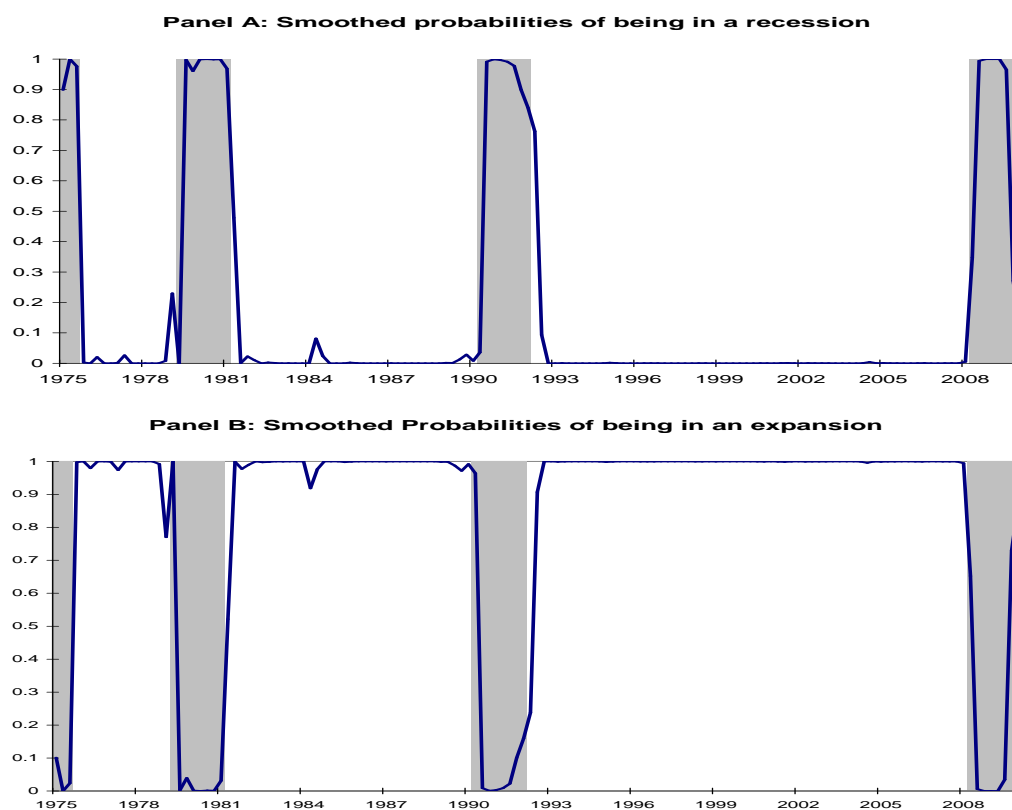
Information criteria (SIC and HQ) selected a model with three regimes. However, very few observations were associated with the third regime so that we decided to keep the model with two regimes for the sake of parsimony and for ease of information.

Table 2.11 in the appendix presents the in-sample results for each indicator for the MSI(2)-MIDAS and MS(2)-MIDAS models. The intercept $\beta_{0,1}$ in the first regime is always negative, while the intercept in the second regime $\beta_{0,2}$ is always positive. Both coefficients are highly significant in all cases. The coefficient β_1 is significant in most of the cases, which emphasizes the importance of including variables sampled at a monthly frequency for predicting quarterly GDP.

Figure 2.4 reports the estimated smoothed probabilities. The shadow areas are the

recessions identified by the ECRI ¹⁰. As mentioned above, the four recessions experienced by the UK are all very well matched by the model with two regimes and no switch in the variance of the error term. We therefore use this class of models in the out-of-sample forecasting exercise. The different MS-MIDAS specifications are recursively estimated in each forecasting period and are used to predict not only the regime but also the level of GDP growth, so that the influence of the full sample regime fitting based specification is very small.

Figure 2.4: MSIAR(2)-MIDAS QUARTERLY GDP AND MONTHLY SLOPE OF THE YIELD CURVE, 1976:Q1-2010:Q1



¹⁰The ECRI business cycle chronology is available at: <http://www.businesscycle.com/resources/cycles/>. The ECRI provides a business cycle chronology at the monthly frequency. We transformed it into quarterly frequency by considering that the UK economy is in recession in quarter t if the ECRI indicates so for at least one of the months of quarter t .

2.4.2.2 Out-of-sample results

The design of the real-time forecasting exercise is identical to the one described in section 2.4.1.2. The actual values for GDP are taken from the last vintage of data available to us T=2010:Q2. Table 2.7 reports the Mean Square Forecast Error relative to an AR(1) model. The main results are the following.

First, note that the Markov-switching MIDAS models always outperform the MIDAS and AR-MIDAS models across all indicators for one-step ahead predictions of GDP, confirming the importance of allowing for time variation in the MIDAS regression for nowcasting and short-term forecasting. Second, each of the three indicators in the MS-MIDAS specification yield relevant information since they produce better forecasting results than the AR(1) model. Third, the AR-MIDAS and MIDAS models always obtain the best performance for two-step ahead predictions. Fourth, unlike for the US, share prices do not clearly outperform the slope of the yield curve and the short-term interest rate. Fifth, the AR-MIDAS model always outperforms the standard MIDAS with the slope of the yield curve and the short-term interest rate, while the standard MIDAS model performs better than the AR-MIDAS with share prices as an indicator for two-step ahead predictions. Finally, the standard Markov-switching model gets better forecasts than the AR(1) model for one-step ahead predictions but worse for two-step ahead predictions. At each horizon the MS model is beaten by at least one MS-MIDAS specification, confirming the importance of introducing higher frequency information in the MS model.

Table 2.8 and Table 2.9 show the QPS and LPS criteria that allow us to assess the regime prediction ability of the models under scrutiny. First, the slope of the yield curve and the BoE base rate perform better than share prices for regime prediction: this confirms what we have found in the empirical application for the US. Besides, the standard Markov-switching model exhibits a good performance by itself, suggesting that the use of mixed frequency data does not improve regime prediction as much as in the case for the US. However, while the standard MS model can be only implemented at the quarterly level, the MS-MIDAS specifications allow for a timely monthly update of the forecasts.

Figure 2.5 shows the estimated nowcasted probability of being in a recession for the model with the slope of the yield curve. These probabilities come from the forecasting exercise and correspond to the estimated filtered probabilities of being in a recession for the last observation T, where T is recursively expanded over time from t=2003:Q4 to 2010:Q1. The ECRI business cycle chronology indicates that the last recession started in May 2008. The chart shows that there is a peak in the probability of recession for the third quarter of 2008 and it indicates that the probability of recession sharply declined in the first quarter

of 2010 ¹¹.

Table 2.13 in the appendix provides further insight on the ability of the MS-MIDAS models to detect recessions in real time, and illustrates how the probability of recession can be updated on a monthly basis. This table shows that there is signal of recession in July 2008 using data for GDP from the October 2008 vintage as the probability of recession amounts to .68 before rising to 1 in the last quarter of 2008. The probabilities of recession are equal or close to 1 in 2009, but decline significantly in the first quarter of 2010. Indeed, the probability of recession in March 2010 is .39 suggesting that the UK recession that started in May 2008 according to the ECRI and two months later according to us, came to an end in the first quarter of 2010.

2.5 Conclusions

Mixed data sampling (MIDAS) models are attracting considerable attention in the literature for their ability to combine in a rather simple regression framework variables sampled at different frequencies. Time-varying parameter models, with changes both in the conditional mean and in the variance, are also more and more used in applied macroeconomics. In this paper we combine these two strands of literature, and introduce the Markov-switching (MS-)MIDAS model, which allows for time-variation in the parameters of MIDAS models, and for the use of high frequency information in standard MS models.

The MS-MIDAS model can be estimated by maximum likelihood, and Monte Carlo experiments indicate that the resulting estimates are rather accurate. Information criteria can then be used for the selection of the number of lags and regimes. Two empirical applications to nowcasting and forecasting quarterly GDP growth for the US and the UK using monthly financial indicators confirm the good performance of the MS-MIDAS model. It can also rather accurately predict changes in regimes.

Due to its generality and ease of implementation, we believe that the MS-MIDAS model can provide a convenient specification for a large class of empirical applications in applied macroeconomics and finance.

¹¹At the time we have completed this draft, the ECRI has not announced yet the end of the last recession for the UK.

Table 2.7: Relative Mean Squared Forecast error for forecasting UK GDP growth 2004:Q1-2010:Q1

		Forecast horizon (h)						
	Model	0	1/3	2/3	1	4/3	5/3	2
Slope of the yield curve	MSI(2)-MIDAS	0.621	0.714	0.617	0.693	1.088	1.151	1.130
	MSIAR(2)-MIDAS	0.684	0.684	0.681	0.869	1.069	1.063	1.167
	MS(2)-MIDAS	0.801	0.896	0.819	0.844	1.070	1.168	1.164
	MSAR(2)-MIDAS	0.698	0.686	0.684	0.802	1.097	1.139	1.133
	AR-MIDAS	0.857	0.917	0.880	0.939	0.954	0.953	0.975
	MIDAS	0.906	0.969	0.897	1.017	1.115	1.174	1.165
Share prices	MSI(2)-MIDAS	0.697	0.699	0.861	0.704	1.019	1.009	1.006
	MSIAR(2)-MIDAS	0.814	0.810	0.810	0.806	1.070	0.998	0.994
	MS(2)-MIDAS	0.861	0.791	1.118	1.201	1.113	1.129	1.129
	MSAR(2)-MIDAS	0.750	0.801	0.916	0.766	1.190	1.041	1.040
	AR-MIDAS	0.868	0.868	0.850	0.933	0.958	0.939	0.938
	MIDAS	0.957	0.897	0.899	0.887	0.956	0.925	0.921
BoE base rate	MSI(2)-MIDAS	0.633	0.685	0.828	0.752	1.099	1.095	1.137
	MSIAR(2)-MIDAS	0.746	0.788	0.833	0.859	1.088	1.070	1.105
	MS(2)-MIDAS	0.637	0.767	0.698	0.698	0.925	0.925	1.149
	MSAR(2)-MIDAS	0.792	0.757	0.833	0.843	0.917	0.920	1.113
	AR-MIDAS	0.953	0.957	0.954	1.078	0.862	0.903	1.079
	MIDAS	1.034	1.028	1.032	1.089	1.131	1.132	1.280
	MSIAR(2)	-	-	-	0.866	-	-	1.082

Real-time data set. Relative Mean Squared Forecast Error for output growth in the quarters 2004:Q1-2010:Q1. Benchmark: AR(1) model. Recursive forecasting scheme. Entries in bold outline the model with the lowest MSFE for each indicator and forecast horizon. A classification of the models is reported in Table 2.14.

Table 2.8: Quadratic Probability Score for forecasting UK business cycle regimes 2004:Q1-2010:Q1

		Forecast horizon (h)						
	Model	0	1/3	2/3	1	4/3	5/3	2
Slope of the yield curve	MSI(2)-MIDAS	0.204	0.215	0.203	0.201	0.434	0.379	0.345
	MSIAR(2)-MIDAS	0.174	0.174	0.174	0.184	0.403	0.329	0.447
	MS(2)-MIDAS	0.227	0.232	0.237	0.234	0.354	0.383	0.323
	MSAR(2)-MIDAS	0.182	0.173	0.173	0.178	0.356	0.341	0.391
Share prices	MSI(2)-MIDAS	0.263	0.259	0.244	0.237	0.320	0.333	0.345
	MSIAR(2)-MIDAS	0.206	0.206	0.189	0.192	0.398	0.411	0.410
	MS(2)-MIDAS	0.260	0.196	0.321	0.180	0.383	0.388	0.389
	MSAR(2)-MIDAS	0.367	0.469	0.326	0.432	0.488	0.474	0.474
BoE base rate	MSI(2)-MIDAS	0.214	0.215	0.213	0.187	0.332	0.339	0.328
	MSIAR(2)-MIDAS	0.176	0.177	0.176	0.170	0.371	0.337	0.287
	MS(2)-MIDAS	0.154	0.209	0.175	0.175	0.239	0.239	0.335
	MSAR(2)-MIDAS	0.267	0.174	0.167	0.175	0.316	0.250	0.346
	MSIAR(2)	-	-	-	0.175	-	-	0.315

Entries in bold outline the model with the lowest QPS for each indicator and forecast horizon. QPS is computed as follows:

$$QPS = \frac{2}{F} \sum_{t=1}^T (P(S_{t+h} = 1) - ECRI_{t+h})^2$$

where F is the number of forecasts, $P(S_{t+h})$ are the predicted regime probabilities of being in the first regime and $ECRI_{t+h}$ is a dummy variable that takes on a value of 1 if the UK economy is in recession in quarter $t+h$ according to the ECRI. For $h = \{0, 1/3, 2/3, 1\}$, we predict business cycle regimes one quarter ahead, whereas for $h = \{4/3, 5/3, 2\}$ we predict business cycle regimes two quarters ahead. A classification of the models is reported in Table 2.14.

Table 2.9: Log Probability Score for forecasting UK business cycle regimes 2004:Q1-2010:Q1

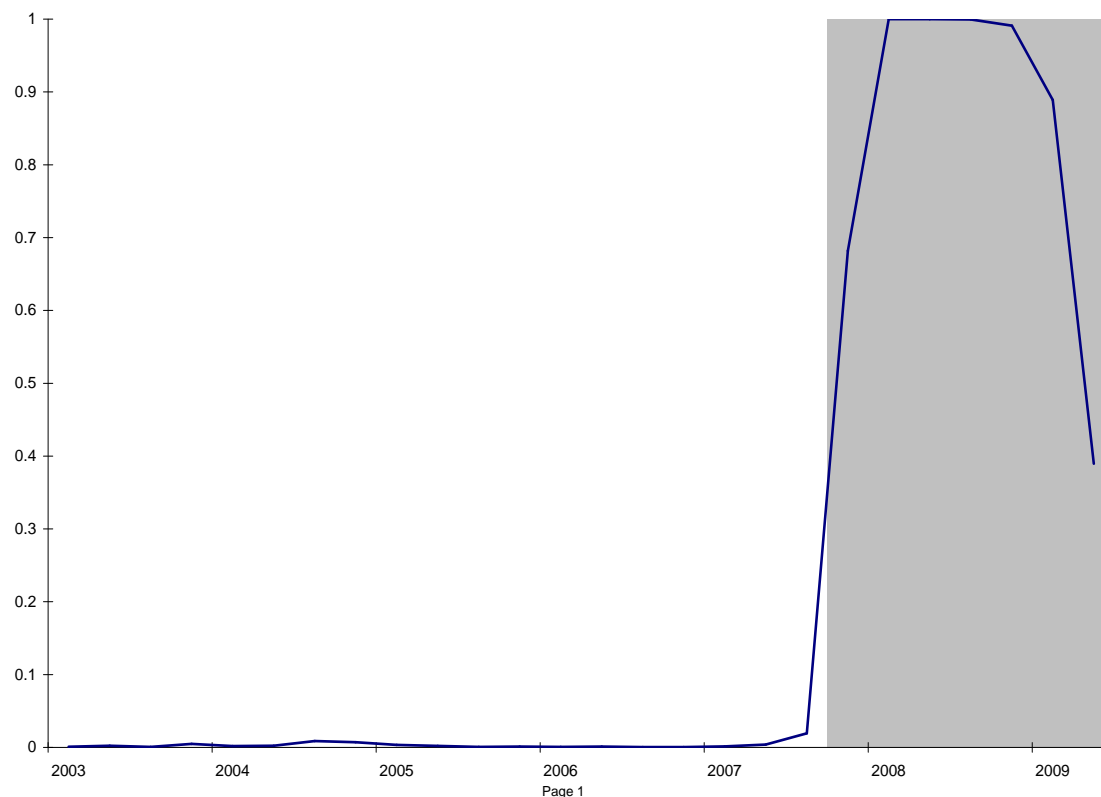
		Forecast horizon (h)						
	Model	0	1/3	2/3	1	4/3	5/3	2
Slope of the yield curve	MSI(2)-MIDAS	0.370	0.383	0.368	0.367	0.656	0.578	0.544
	MSIAR(2)-MIDAS	0.345	0.345	0.344	0.353	0.621	0.532	0.653
	MS(2)-MIDAS	0.397	0.404	0.388	0.387	0.555	0.558	0.511
	MSAR(2)-MIDAS	0.348	0.342	0.342	0.346	0.553	0.521	0.593
Share prices	MSI(2)-MIDAS	0.442	0.437	0.414	0.409	0.554	0.568	0.579
	MSIAR(2)-MIDAS	0.369	0.367	0.353	0.357	0.683	0.700	0.697
	MS(2)-MIDAS	0.586	0.351	0.650	0.335	0.677	0.663	0.664
	MSAR(2)-MIDAS	0.659	0.855	0.589	0.768	0.885	0.812	0.812
BoE base rate	MSI(2)-MIDAS	0.381	0.379	0.365	0.343	0.532	0.545	0.533
	MSIAR(2)-MIDAS	0.341	0.338	0.338	0.334	0.607	0.542	0.485
	MS(2)-MIDAS	0.275	0.372	0.346	0.346	0.406	0.406	0.536
	MSAR(2)-MIDAS	0.524	0.333	0.335	0.330	0.520	0.419	0.549
	MSIAR(2)	-	-	-	0.340	-	-	0.517

Entries in bold outline the model with the lowest LPS for each indicator and forecast horizon. LPS is computed as follows:

$$LPS = -\frac{1}{F} \sum_{t=1}^T (1 - ECRI_{t+h}) \log(1 - P(S_{t+h} = 1)) + ECRI_{t+h} \log(P(S_{t+h} = 1))$$

where F is the number of forecasts, $P(S_{t+h})$ are the predicted regime probabilities of being in the first regime and $ECRI_{t+h}$ is a dummy variable that takes on a value of 1 if the UK economy is in recession in quarter $t+h$ according to the ECRI. For $h = \{0, 1/3, 2/3, 1\}$, we predict business cycle regimes one quarter ahead, whereas for $h = \{4/3, 5/3, 2\}$ we predict business cycle regimes two quarters ahead. A classification of the models is reported in Table 2.14.

Figure 2.5: ESTIMATED PROBABILITY OF RECESSION, REAL-TIME DATA, 2003:Q4-2010:Q1



Note: MSIAR(2)-MIDAS model with the monthly slope of the yield curve, forecast horizon $h = 0$

2.6 Appendix Chapter 2

2.6.1 Estimation algorithm

All the models are estimated by maximum likelihood. The computations are carried out with the optimization library OPTMUM of Gauss 7.0. selecting the BFGS algorithm.

The algorithm we use is described by the following steps:

Denote ω the parameters of the models to be estimated.

- STEP 1: Give initial values to all parameters of the model ω^0 .
- STEP 2: If there is regime switching, implement the Hamilton (1989) filtering proce-

dure using in the first iteration ω^0 and in the following iterations ω^j . We thus obtain an estimate of the filtered probabilities - if there is regime switching - and the value of the log-likelihood function.

- STEP 3: Maximize the log-likelihood function to obtain an updated version of the parameters ω^j
- STEP 4: Iterate over STEP 2 and STEP 3 until the algorithm has converged.

Hamilton (1994) pointed out that this algorithm is a special case of the EM algorithm: the expectation (E) step is step 2 and the maximization (M) step is step 3. Note that the expectation step aims at the formulation of guesses about the latent variables given the data and the initial or updated values of the parameters, while the maximization step yields the values of the parameters that maximize the log-likelihood function over the iterations.

2.6.2 Additional Tables

Table 2.10: In-sample results, Quarterly US GDP growth rate 1959:Q1-2009:Q4

	Slope of the yield curve	S&P 500	Fed Funds	Slope of the yield curve	S&P 500	Fed Funds
$\beta_{0,1}$	-0.273 [0.207]	-0.024 [0.206]	-0.193 [0.205]	0.009 [0.143]	-0.030 [0.206]	-0.063 [0.206]
$\beta_{0,2}$	0.778*** [0.053]	0.718*** [0.058]	0.801*** [0.052]	0.799*** [0.052]	0.781*** [0.052]	0.797*** [0.052]
$\beta_{0,3}$	1.366*** [0.140]	1.271*** [0.136]	1.360*** [0.149]	1.352*** [0.154]	1.363*** [0.194]	1.410*** [0.130]
$\beta_{1,1}$	-0.303** [0.145]	0.098*** [0.026]	0.363** [0.159]	-1.323*** [0.329]	0.121*** [0.039]	0.638*** [0.208]
$\beta_{1,2}$	- -	- -	- -	0.030 [0.225]	0.032** [0.017]	0.513*** [0.147]
$\beta_{1,3}$	- -	- -	- -	0.215 [0.464]	0.068 [0.047]	-0.919*** [0.243]
σ_1	0.651*** [0.187]	0.647*** [0.215]	0.618*** [0.230]	0.592*** [0.154]	0.578*** [0.212]	0.560*** [0.155]
σ_2	0.214*** [0.036]	0.197*** [0.038]	0.213*** [0.036]	0.205*** [0.036]	0.198*** [0.036]	0.207*** [0.033]
σ_3	0.633*** [0.126]	0.626*** [0.131]	0.653*** [0.134]	0.668*** [0.148]	0.626*** [0.165]	0.680*** [0.128]
$P(S_t = 1)$	0.191	0.237	0.209	0.239	0.247	0.189
SIC	491.473	479.164	490.113	487.364	484.587	485.800

The first three columns report in-sample results for the MSIH(3)-MIDAS model (i.e. the model with no switch in β_1), while the last three columns report in-sample results for the the MSH(3)-MIDAS model(i.e. the model with a switch in β_1). ***, ** and * indicate significance at 1%, 5% and 10%. Standard deviations are reported in brackets. SIC is the Schwarz Information Criterion.

Table 2.11: In-sample results, Quarterly UK GDP growth rate 1975:Q1-2010:Q1

	Slope of the yield curve	Share prices	BoE base rate	Slope of the yield curve	Share prices	BoE base rate
$\beta_{0,1}$	-0.724*** [0.214]	-0.825* [0.378]	-0.675*** [0.173]	-0.847*** [0.177]	-0.782*** [0.262]	-0.971*** [0.185]
$\beta_{0,2}$	0.746*** [0.065]	0.704*** [0.071]	0.747*** [0.062]	0.763*** [0.061]	0.664*** [0.068]	0.755*** [0.059]
$\beta_{1,1}$	-0.925* [0.485]	0.055** [0.023]	-0.538 [0.188]	1.236*** [0.421]	0.301** [0.118]	-1.462*** [0.405]
$\beta_{1,2}$	- -	- -	- -	0.388 [0.238]	0.097** [0.030]	-0.069 [0.131]
σ	0.393*** [0.054]	0.352*** [0.081]	0.394*** [0.052]	0.391*** [0.050]	0.364*** [0.056]	0.376*** [0.048]
$P(S_t = 1)$	0.166	0.141	0.197	0.195	0.142	0.160
SIC	325.058	320.831	326.327	327.942	321.760	324.844

The first three columns report in-sample results for the MSI(2)-MIDAS model (i.e. the model with no switch in β_1), while the last three columns report in-sample results for the the MS(2)-MIDAS model(i.e. the model with a switch in β_1). ***, ** and * indicate significance at 1%, 5% and 10%. Standard deviations are reported in brackets. SIC is the Schwarz Information Criterion. QPS is the quadratic probability score using the ECRI business cycle datation and the estimated probabilities of being in the first regime.

Table 2.12: US Estimated Quarterly Probability of Recession updated on a monthly basis

Quarter	Month	$P(S_t = 1)$	$NBER_t$	Quarter	Month	$P(S_t = 1)$	$NBER_t$
Q1 2007	January	0.146	0	Q1 2009	January	0.995	1
	February	0.137	0		February	1.000	1
	March	0.148	0		March	0.999	1
Q2 2007	April	0.128	0	Q2 2009	April	0.994	1
	May	0.123	0		May	0.987	1
	June	0.127	0		June	0.995	1
Q3 2007	July	0.053	0	Q3 2009	July	0.829	0
	August	0.036	0		August	0.822	0
	September	0.033	0		September	0.824	0
Q4 2007	October	0.283	0	Q4 2009	October	0.663	0
	November	0.241	0		November	0.639	0
	December	0.268	1		December	0.642	0
Q1 2008	January	0.636	1	Q1 2010	January	0.486	0
	February	0.742	1		February	0.468	0
	March	0.878	1		March	0.479	0
Q2 2008	April	0.931	1	Q2 2010	April	0.339	0
	May	0.940	1		May	0.343	0
	June	0.961	1		June	0.345	0
Q3 2008	July	0.985	1				
	August	0.983	1				
	September	0.987	1				
Q4 2008	October	1.000	1				
	November	1.000	1				
	December	1.000	1				

$P(S_t = 1)$ are the estimated probabilities of being in the first regime from the MSH(3)-MIDAS model with the monthly slope of the yield curve. $NBER_t$ is a dummy variable that takes on a value of 1 if the economy is in recession and 0 otherwise. For the months of March, June, September and December, the probabilities are obtained from the model with a forecast horizon $h = 0$. For the months of February, May, August and November, the probabilities are obtained from the model with a forecast horizon $h = 1/3$. For the months of January, April, July and October, the probabilities are obtained from the model with a forecast horizon $h = 2/3$.

Table 2.13: UK Estimated Quarterly Probability of Recession updated on a monthly basis

Quarter	Month	$P(S_t = 1)$	$ECRI_T$	Quarter	Month	$P(S_t = 1)$	$ECRI_T$
Q1 2007	January	0.001	0	Q1 2009	January	1.000	1
	February	0.001	0		February	1.000	1
	March	0.001	0		March	1.000	1
Q2 2007	April	0.001	0	Q2 2009	April	1.000	1
	May	0.001	0		May	0.999	1
	June	0.001	0		June	1.000	1
Q3 2007	July	0.000	0	Q3 2009	July	0.993	1
	August	0.000	0		August	0.991	1
	September	0.000	0		September	0.991	1
Q4 2007	October	0.001	0	Q4 2009	October	0.913	1
	November	0.001	0		November	0.892	1
	December	0.001	0		December	0.889	1
Q1 2008	January	0.004	0	Q1 2010	January	0.523	1
	February	0.004	0		February	0.437	1
	March	0.004	0		March	0.390	1
Q2 2008	April	0.019	0				
	May	0.019	1				
	June	0.019	1				
Q3 2008	July	0.683	1				
	August	0.691	1				
	September	0.683	1				
Q4 2008	October	1.000	1				
	November	1.000	1				
	December	1.000	1				

$P(S_t = 1)$ are the estimated probabilities of being in the first regime from the MSIHAR(2)-MIDAS model with the monthly slope of the yield curve. $ECRI_t$ is a dummy variable that takes on a value of 1 if the economy is in recession and 0 otherwise. Note that at the time we have written this paper, the ECRI has not announced yet the end of the last recession. For the months of March, June, September and December, the probabilities are obtained from the model with a forecast horizon $h = 0$. For the months of February, May, August and November, the probabilities are obtained from the model with a forecast horizon $h = 1/3$. For the months of January, April, July and October, the probabilities are obtained from the model with a forecast horizon $h = 2/3$.

Table 2.14: Classification of the models

The general Markov-switching MIDAS model with M regimes we consider is:

$$y_t = \beta_0(S_t) + \beta_1(S_t)B(L^{1/m}; \theta)x_{t-h}^{(m)} + \epsilon_t(S_t)$$

where $\epsilon_t|S_t \sim NID(0, \sigma^2(S_t))$.

Model	regime changes in	AR component
MSI(M)-MIDAS	β_0	NO
MS(M)-MIDAS	β_0 and β_1	NO
MSIH(M)-MIDAS	β_0 and σ^2	NO
MSH(M)-MIDAS	β_0, β_1 and σ^2	NO
MSIAR(M)-MIDAS	β_0	YES
MSAR(M)-MIDAS	β_0 and β_1	YES
MSIHAR(M)-MIDAS	β_0 and σ^2	YES
MSHAR(M)-MIDAS	β_0, β_1 and σ^2	YES
MSIAR(M)	β_0	YES
MSIHAR(M)	β_0 and σ^2	YES

The suffix "H" refers to models with a switch in the variance of the shocks. The suffix "I" refers to models with a switch in the intercept β_0 . The suffix "AR" means that we include an AR component in the model through a common factor to avoid a seasonal response of y to x as it is described in equation 2.5. The last two rows of the table show the labels we use for the standard Markov-switching models.

Chapter 3

Trend-cycle decomposition of output and euro area inflation forecasts: a real-time approach based on model combination

Abstract

To cope with both model uncertainty and parameter instability that are inherent to trend-cycle decomposition models of GDP, we construct model-averaged measures of the output gap (both ex-post and real-time estimates). We first estimate nine models of trend-cycle decomposition of euro area GDP, both univariate and multivariate, some of them allowing for changes in the slope of trend GDP and/or its error variance using Markov-switching specifications, or including a Phillips curve. We then pool the estimates using three weighting schemes. We finally run a forecasting experiment to evaluate the predictive power of the output gap for inflation in the euro area. Our measures help forecasting inflation over most of our evaluation sample (2001-2010) but fail dramatically over the last recession.

Keywords: Trend-cycle decomposition, Phillips curve, Unobserved components, Kalman filter, Markov-switching.

JEL Classification Code: C53, E37.

⁰This is a joint work with Laurent Maurin and Matthias Mohr.

3.1 Introduction

The estimation of potential output is of primary importance for policy makers since it represents the maximum level of output which is not associated with inflationary pressures. The output gap - i.e. the difference between the actual level of output and the potential output - conveniently summarizes the transitory state of the economy by determining whether the economy operates below or above its sustainable level.

Unobserved components (UC) models are often used to measure potential output since they are specifically designed to deal with latent (i.e. unobserved) variables. Univariate trend-cycle decomposition model of real GDP can be traced back to Watson (1986) and Clark (1987). These studies, which focus on the US economy, find that the cyclical part of output closely matches the US recessions identified by the NBER. Indeed, they allocate most of the variation of output to the cycle and leave the trend mostly unchanged over time. Conversely, the Beveridge and Nelson (1981) (BN) decomposition of GDP attributes most of its variability to its trend, whereas its cyclical component remains small, noisy and does not match the NBER business cycle dating of the economic activity.

Morley et al. (2003) explain the discrepancy between the BN and UC decompositions by the fact that it is usually assumed in the literature that there is no correlation between the shocks to the trend and the cycle. The authors find that relaxing this restriction makes the UC decomposition of GDP identical to the BN decomposition. Moreover, they report a negative and significant correlation between the shocks to the trend and to the cycle.

Conversely, Perron and Wada (2009) emphasize the importance of allowing for a change in the slope of the trend. They model the shocks to the trend and cycle as a mixture of two normal distributions that permits to capture endogenously changes in the slope of trend GDP. In doing so, they identify a structural break in the slope of the trend of US real GDP around 1973:Q1 and obtain a cycle component of GDP that is consistent with the NBER dating of the economic activity. In this paper, we extend this approach and propose to capture changes in the slope of trend GDP with regime switches in its slope and its shocks variance.

The Markov-switching model of Hamilton (1989) is appealing since it makes the probability of parameter changes dependent on past realizations, whereas assuming that the errors of the state follow a mixture of normal distributions (i.e. the approach followed by Perron and Wada (2009)) implies that the probabilities that the errors are drawn from one regime to the other are independent from past realizations. In this respect, adopting a Markov-switching specification implies that, unlike Perron and Wada (2009), we allow for a change in trend

growth to last several quarters while remaining short lived, and to happen more than once.

To cope with model uncertainty inherent to trend-cycle decomposition models of GDP, we also incorporate additional information to improve the estimation of the output gap. First, we consider the use of an auxiliary indicator - the rate of capacity utilization - to help identifying the transitory component of GDP. Given the high correlation between this indicator and the business cycle component of economic activity, we can expect that this improves the estimation of the output gap. Second, we add a Phillips curve to the model of trend-cycle decomposition of GDP. The use of a Phillips curve for estimating the output gap has been first advocated by Kuttner (1994), who appends a Phillips curve to the univariate trend-cycle decomposition of GDP and finds that this bivariate model helps to better estimate the output gap.

The estimation of the output gap is characterized by both model uncertainty and parameter instability. Model uncertainty means that model selection is a tricky issue since we do not observe the true level of the output gap, while parameter instability refers to the idea that parameter estimates can be sensitive to the estimation window chosen. As a consequence, the output gap estimates are surrounded by a large uncertainty. One solution consists in reporting predictive densities of the output gap (see e.g. Garratt et al. (2009)). Another solution is to compute model-averaged measures of the output gap in order to reduce model uncertainty (see e.g. Morley and Piger (2009)).

Another issue pointed out by Orphanides and Van Norden (2002) is the unreliability of the estimates of output gap in real-time. However, Marcellino and Musso (2010) find that the use of real-time data is less problematic to estimate the euro area output gap.

In this paper, we estimate nine models of the euro area output gap: linear, non-linear, univariate and bivariate models. We then report model-averaged measures of the output gap with their single model counterparts and show that the differences across estimates are sizeable. We find some evidence of regime changes in the slope of the trend of the euro area GDP for few periods around 1974 and since 2009. We then run a pseudo out-of-sample forecasting experiment to forecast the level and the change in inflation using both ex-post and real-time estimates of the output gap. We find that our output gap measures help forecasting inflation over most of the sample but fail dramatically since the last recession. This may be due to a strong anchoring of inflation expectations since we find that a large part of inflation appears to be driven by its forward-looking component.

The paper is organized as follows. Section 3.2 presents the univariate and multivariate models of trend/cycle decomposition of GDP with and without regime switching. Section 3.3 discusses the estimation method and reports the empirical results for the euro area.

In this section, we also discuss the estimation of a univariate time-varying Phillips curve. The estimation of the output gap in real-time and its forecasting performance for predicting inflation in the euro area is analysed in Section 3.4. Section 3.5 concludes.

3.2 Trend-cycle decomposition of output

In this section, we present the models used to decompose GDP in between trend and cycle. We start with the univariate model, discuss the inclusion of Markov-switching parameters and then present the bivariate models.

Watson (1986) provides a starting point to decompose the level of output y_t into a trend n_t and a cycle z_t :

$$y_t = n_t + z_t \quad (3.1)$$

The trend n_t is modeled as a random walk with drift and the cyclical component z_t is modeled as an AR(2) process:

$$n_t = \mu + n_{t-1} + \epsilon_t^n \quad (3.2)$$

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \epsilon_t^z \quad (3.3)$$

The disturbances ϵ_t^n and ϵ_t^z are assumed to be normally distributed, i.i.d, with mean-zero and are not correlated. The trend component n_t is interpreted as the level of potential output, while the cycle z_t is interpreted as the output gap. This model is relatively standard in the literature and can be cast in state-space form, with a state vector of dimension 3 (see Appendix 3.6.1 for the measurement and state equations).

3.2.1 Extension to regime changes in the slope of the trend

To extend the standard model, we consider regime changes in the intercept of the trend component μ and in the variance of the shock affecting it, ϵ_t^n , using regime switches governed by a Markov chain. This allows trend growth to be regime dependent. The general Markov-switching model we consider is:

$$y_t = n_t + z_t \quad (3.4)$$

$$n_t = \mu(S_t) + n_{t-1} + \epsilon_t^n(S_t) \quad (3.5)$$

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \epsilon_t^z \quad (3.6)$$

where $\epsilon_t^n | S_t \sim NID(0, \sigma_n^2(S_t))$ and $\epsilon_t^z \sim NID(0, \sigma_z^2)$

The regime generating process is an ergodic Markov chain with a finite number of states $S_t = \{1, \dots, M\}$ defined by the following transition probabilities:¹

$$p_{ij} = Pr(S_{t+1} = j | S_t = i) \quad (3.7)$$

$$\sum_{j=1}^M p_{ij} = 1 \forall i, j \in \{1, \dots, M\} \quad (3.8)$$

Regime changes in the intercept μ of the trend component can occur due to a decline in productivity due to unemployment hysteresis or stronger scrapping of capital during recessions associated with a restructuring of the economy. Similarly, changes in the variance of shocks to trend GDP can be attributed to stronger shocks affecting the economy during recessions. For example Cogley and Sargent (2005), Sims and Zha (2006) and Fernández-Villaverde et al. (2010) emphasize the importance of allowing the variance of the shocks to vary. The prior view is that low growth is associated with large negative shocks. In the set of models estimated below, we consider both changes in the slope together with changes in the variance of the shocks.

As the level of potential output n_t and the output gap z_t are not observed, the model has to be cast in state-space form before being estimated with the Kalman filter. The inclusion of regime changes in some parameters of the model complicates the estimation since there is an additional latent variable S_t in the model. However, Kim and Nelson (1999b, chapter 5) show how to estimate state-space models with regime switching, i.e. how to combine the Kalman and Hamilton filters in a tractable way. Further details about the estimation are provided in Section 3.3.1., while Appendix 3.6.2 reports the equations for the Kalman and Hamilton filters.

It is important to note that we only include regime changes in some parameters of the trend equation of GDP since we want to capture possible changes in the level of potential output. Conversely, Kim and Nelson (1999a) include regime switches in the intercept of the cycle equation of GDP and Sinclair (2009) extends their specification by allowing for a correlation between the errors in the trend and the cycle.

In the empirical application, the linear model given by equations 3.1 to 3.3 is labeled as MODEL UC-1, the Markov-switching model with only a switch in the intercept of the trend component of GDP is labeled as MODEL UC-2 and the Markov-switching model with a switch in both the drift of the trend component of GDP and its shock variance is labeled as MODEL UC-3.

¹See Hamilton (1989) for more details.

3.2.2 Extension to use auxiliary information

Beside the three univariate models described above, we also consider the use of an auxiliary indicator to better estimate the output gap. In the empirical application, the indicator is the rate of capacity utilization which is often used as a proxy for the cyclical component of GDP. Indeed, if one considers the output gap as the transitory component of GDP, appending an indicator well correlated to the economic activity should provide relevant information for estimating the output gap.

The measurement and transition equations for the bivariate model of GDP and the auxiliary indicator are then respectively given by:

$$\begin{bmatrix} y_t \\ aux_t \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & \alpha_1 & \alpha_2 \end{bmatrix} \begin{bmatrix} n_t \\ z_t \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} 0 \\ \epsilon_t^{aux} \end{bmatrix} \quad (3.9)$$

$$\begin{bmatrix} n_t \\ z_t \\ z_{t-1} \end{bmatrix} = \begin{bmatrix} \mu(S_t) \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & \phi_1 & \phi_2 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} n_{t-1} \\ z_{t-1} \\ z_{t-2} \end{bmatrix} + \begin{bmatrix} \epsilon_t^n(S_t) \\ \epsilon_t^z \\ 0 \end{bmatrix} \quad (3.10)$$

where $\epsilon_t^{aux} \sim NID(0, \sigma_{aux}^2)$, $\epsilon_t^n(S_t) \sim NID(0, \sigma_n^2(S_t))$, $\epsilon_t^z \sim NID(0, \sigma_z^2)$ and $\epsilon_t^\kappa \sim NID(0, \sigma_\kappa^2)$.

We consider linear bivariate models (labeled as MODEL MUC-1 (auxiliary)). For the non-linear bivariate models we estimate, we include regime changes in the parameters of the trend equation of GDP in the same way we did in the univariate modeling: (i) switch in the slope of the trend only (labeled as MODEL MUC-2 (auxiliary)) or (ii) switch in both the slope of the trend and its shock variance (labeled as MODEL MUC-3 (auxiliary)).

3.2.3 Extension to incorporate a time-varying Phillips curve

We follow Kuttner (1994) and add an equation for inflation along with the trend-cycle decomposition of GDP. We can indeed expect gains by adding an inflation equation to our model since - in theory - inflation is linked to the level of the output gap. Although it is indeed sometimes found that inflation can help to estimate the transitory component of output (see e.g. Kuttner (1994) and Proietti et al. (2007)), there is no clear agreement in the literature. For instance, based on US data, Orphanides and Van Norden (2002) find that multivariate models do not outperform their univariate counterparts.

An additional problem with the Phillips curve specification relates to the well known

fact that over the forty years covered in our empirical analysis, the inflation regime has changed. To account for this, we use a time-varying version of the Phillips curve, which is then incorporated in the model of trend-cycle decomposition of GDP. We consider a time-varying Phillips-curve of the form:

$$\pi_t = \kappa_t + \sum_{j=1}^J \lambda_{\pi,j} \pi_{t-j} + \sum_{j=0}^J \lambda_{z,j} z_{t-j} + \sum_{j=1}^J \lambda_{EXR,j} EXR_{t-j} + \sum_{j=1}^J \lambda_{OIL,j} OIL_{t-j} + \epsilon_t^\pi \quad (3.11)$$

$$\kappa_t = \kappa_{t-1} + \epsilon_t^\kappa \quad (3.12)$$

where $\epsilon_t^\pi \sim NID(0, \sigma_\pi^2)$, $\epsilon_t^\kappa \sim NID(0, \sigma_\kappa^2)$ and π_t , z_t , EXR_t and OIL_t are the inflation rate, the cyclical component of output ², the nominal effective exchange rate and the price of oil respectively. The intercept κ_t is modeled as a random walk without drift in order to capture changes in the trend of inflation and can be interpreted as the level of inflation expectations. The other parameters of the model (λ 's, σ_κ^2 and σ_π^2) are kept constant.

Again, as the parameter κ is not constant over time, equations 3.11 and 3.12 have to be estimated via maximum likelihood using the Kalman filter. The state-space representation of this model is given by:

$$\pi_t = \kappa_t + \lambda x_t + \epsilon_t^\pi \quad (3.13)$$

$$\kappa_t = \kappa_{t-1} + \epsilon_t^\kappa \quad (3.14)$$

where x_t is a matrix of observables and λ its corresponding vector of coefficients.

The measurement and transition equations for the bivariate model of GDP and inflation are instead respectively given by:

$$\begin{bmatrix} y_t \\ \pi_t \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & \lambda_z & 0 & 1 \end{bmatrix} \begin{bmatrix} n_t \\ z_t \\ z_{t-1} \\ \kappa_t \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} 0 \\ \mathbf{x}_t \end{bmatrix} + \begin{bmatrix} 0 \\ \epsilon_t^\pi \end{bmatrix} \quad (3.15)$$

²We use here the HP filtered cycle as a proxy for the cyclical component of output.

where \mathbf{x}_t is a matrix of explanatory variables and λ its corresponding vector of coefficients.

$$\begin{bmatrix} n_t \\ z_t \\ z_{t-1} \\ \kappa_t \end{bmatrix} = \begin{bmatrix} \mu(S_t) \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \phi_1 & \phi_2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} n_{t-1} \\ z_{t-1} \\ z_{t-2} \\ \kappa_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_t^n(S_t) \\ \epsilon_t^z \\ 0 \\ \epsilon_t^\kappa \end{bmatrix} \quad (3.16)$$

where $\epsilon_t^{aux} \sim NID(0, \sigma_{aux}^2)$, $\epsilon_t^n(S_t) \sim NID(0, \sigma_n^2(S_t))$, $\epsilon_t^z \sim NID(0, \sigma_z^2)$ and $\epsilon_t^\kappa \sim NID(0, \sigma_\kappa^2)$.

We consider linear bivariate models labeled as MODEL MUC-1 (inflation). For the non-linear bivariate models we estimate, we include regime changes in the parameters of the trend equation of GDP in the same way we did in the univariate modeling: (i) switch in the slope of the trend only (labeled as MODEL MUC-2 (inflation)) or (ii) switch in both the slope of the trend and its shock variance (labeled as MODEL MUC-3 (inflation)).

3.3 In-sample estimates for the euro area

The estimation of the nine models described above is carried out using quarterly data for the euro area as a whole over the period 1970Q1-2010Q4 (i.e. 164 observations). Real GDP is taken from Eurostat and backcasted with the AWM database before 1995Q1. The auxiliary indicator is the rate of capacity utilization published by the European Commission. It is available for the euro area since 1985Q1 and backcasted with country data before. It is demeaned prior to the estimation. Regarding the variables entering the Phillips curve, the harmonised index of consumer prices (HICP) is taken from Eurostat, the oil price in US dollars from TWI, while the euro US dollar exchange rate and the euro nominal effective exchange rate against its 16 main competitors are taken from BIS data.

All models are estimated with maximum likelihood. The computations are carried out with the optimization library OPTMUM of GAUSS 9.0.0. selecting the BFGS algorithm. Denoting ω the parameters of the model to be estimated, the algorithm we use is described by the following steps:

- STEP 1: Give initial values to all parameters of the model ω^0 and to the expectation of the state vector and its variance.
- STEP 2: If there is regime switching in at least one parameter of the model, implement the filtering procedure of Kim and Nelson (1999b) for state-space models with regime

switching using in the first iteration ω^0 and in the following iterations ω^j . If there is no regime switching, one needs to implement the standard Kalman filter using in the first iteration the initial values for the state vector and its variance, and in the next iterations their updated versions. At the end of Step 2, we thus obtain an estimate of the filtered probabilities (if there is regime switching), the state vector and the log-likelihood function.

- STEP 3: Maximize the log-likelihood function to obtain an updated version of the parameters ω^j .
- STEP 4: Iterate over STEPS 2 to 3 until the algorithm has converged.

Hamilton (1994) pointed out that this algorithm is a special case of the EM algorithm: the expectation (E) step is step 2 and the maximization (M) step is step 3. Note that the expectation step aims at formulating guesses about the latent variables (i.e. the unobserved components and the regime probabilities) given the data and the initial or updated values of the parameters, while the maximization step yields the values of the parameters that maximize the log-likelihood over the iterations.

3.3.1 Time-varying Phillips curve

We first estimate a univariate time-varying Phillips curve without a trend-cycle decomposition model of GDP since model selection would raise difficulties in the multivariate framework.³ The specification of the Phillips curve that best fits the data is chosen in a univariate context, using a basic HP filter as a measure of the output gap, before being included and re-estimated jointly with the output gap in the multivariate framework.

We estimate equations 3.11 and 3.12 by maximizing the log-likelihood function via the EM algorithm as described in the previous subsection. Inflation is 100 times the quarterly change in consumer prices (HICP) and the output gap is the cycle extracted from the Hodrick-Prescott filter. The exchange rate and oil prices in euro terms are 100 times the quarterly change of their logarithm. For selecting the right number of lags in equation 3.11, we proceed sequentially: we first estimate a model with four lags for each of the explanatory variables and delete the least significant variables until all coefficients are significant at least at the 10% level.

³We experimented problems of convergence of our algorithm when we carried out model selection within the bivariate framework with regime switching.

Table 3.2 in the Appendix reports the maximum likelihood parameter estimates and their standard errors. First, applying the above criterion to determine the lags on each explanatory variable, we select a model with no lagged inflation. We see two explanations for this result: (i) the time-varying parameter can capture part of the significance of lagged inflation (ii) confirmation of the purely forward looking New Keynesian Phillips Curve, which states that current inflation only depends on expected inflation and current marginal cost.⁴ Besides, this result is in line with Hondroyannis et al. (2009), who use time-varying parameter models on data for Germany, France, Italy and the United Kingdom and also favor specifications that exclude lagged inflation. Second, the coefficient entering before the output gap is highly significant and positive: a one percentage point increase in the output gap pushes up inflation by 0.15%. This is the lower bound of the estimates reported in the literature. Third, the coefficients on lagged exchange rate and oil price, taken in euros, are both significant and have the expected signs. An appreciation of the euro has a negative impact on inflation, with a 10% appreciation diminishing inflation by about 0.2%. Finally, an increase in oil price has a positive impact on inflation.

Figure 3.1 shows the time-varying parameter of the Phillips curve, which is interpreted as the level of inflation expectations with actual inflation and the difference between actual and expected inflation. For example, over the most recent period, oil price, the exchange rate and the output gap are estimated to have contributed to annual inflation by almost 2 p.p. at the end of 2008 and around -1 p.p. at the end of 2009.

3.3.2 Univariate and bivariate trend-cycle decomposition of euro area real GDP

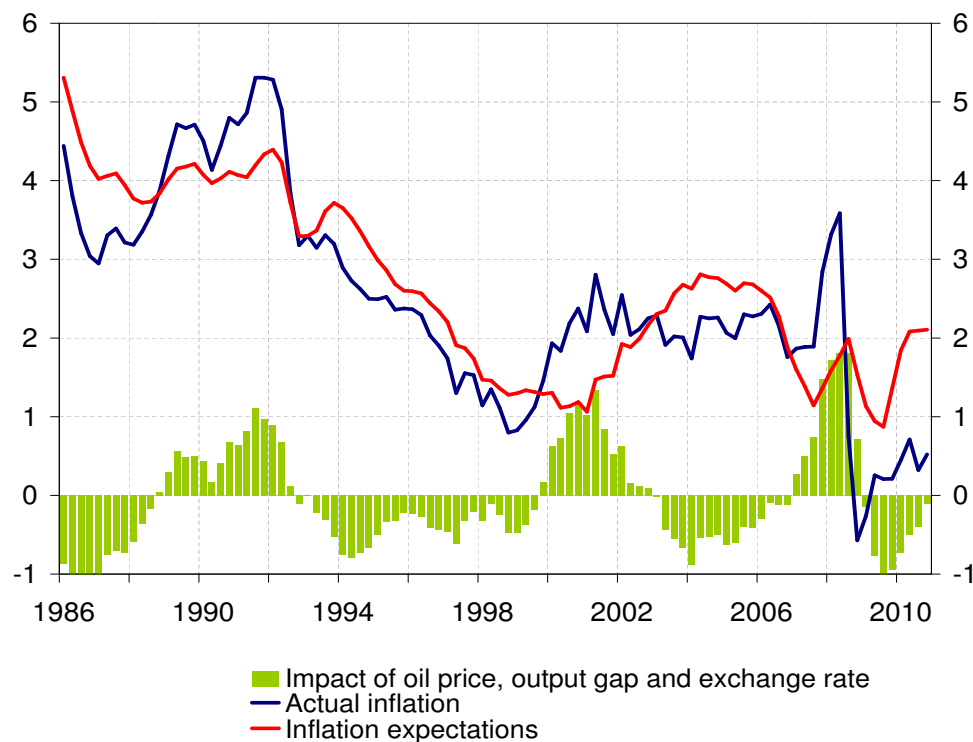
We estimate the univariate and bivariate trend-cycle decomposition models of the euro area GDP described in Section 3.2: the three univariate models and the six bivariate models.⁵ Tables 3.3, 3.4 and 3.5 in the Appendix report the maximum likelihood estimates.

Table 3.3 in the Appendix shows the results for the univariate models of the trend-cycle decomposition of GDP. First, the regime switching intercepts are highly significant in the two regimes. In addition, both regime switching models increase the log likelihood by about 15

⁴In the estimation of the New Keynesian Phillips curve, current marginal costs are often approximated by the output gap (see e.g. Rudd and Whelan (2007)). However, some argue that unit labor costs should be used as the driving variable in the Phillips curve (see e.g. Gali and Gertler (1999)). This debate is beyond the scope of this paper.

⁵In all the models, stationarity constraints on the parameters ϕ_1 and ϕ_2 and positive definiteness constraints on the variance parameters of the innovations were imposed. Standard deviations were computed from the inverse of the outer product estimate of the Hessian.

Figure 3.1: MODEL DECOMPOSITION OF EURO AREA INFLATION (ANNUAL GROWTH, %)



Note: The moving constant corresponds to the time-varying parameter of the univariate Phillips curve (see equation 3.11). The difference reflect the cumulated impact of exchange rate, oil price and output gap.

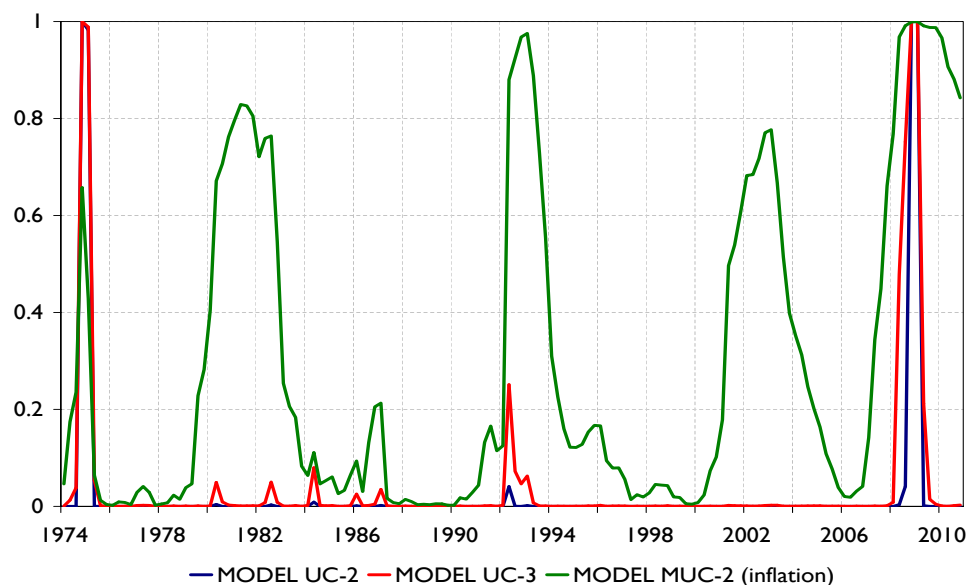
with respect to the linear model.⁶ This points out the relevance of parameter switching in the trend equation of GDP and provides evidence for possible decreases in trend output growth during recessions. Indeed, Figure 3.2 shows that the probability for a negative intercept for potential output peaks for few periods in 1974 and 2009.

Table 3.4 in the Appendix reports the results for the models using the demeaned rate of capacity utilization as an auxiliary indicator to better estimate the output gap. The coefficients for the auxiliary indicator are highly significant, which shows its relevance to estimate the output gap.

The results for the bivariate models using inflation as an extra variable in the system are reported in Table 3.5 of the Appendix, while the resulting inflation expectations are

⁶However, the improvement in the log-likelihood cannot be tested since a standard likelihood ratio test cannot be implemented since (i) the transition probabilities are not identified and (ii) the scores of the log likelihood are identically equal to zero under the null hypothesis of no regime switching.

Figure 3.2: SMOOTHED PROBABILITY OF BEING IN THE FIRST REGIME



Note: MODEL UC-2 is the model with a switch in the slope of the trend only. MODEL UC-3 is the model with in the slope of the trend and its error variance. MODEL MUC-2 (inflation) is a bivariate model with an equation for inflation and a switch in the slope of the trend of GDP only.

represented in Figure 3.9. In the equation for inflation, we include one lag for the output gap, the exchange rate and the oil price following the results obtained in the previous subsection. As a robustness check, we also include one lag for inflation even if it is not significant in the univariate analysis (the state-space representation of the model is given by equations 3.15 and 3.16). The parameter estimates for the coefficients of the exchange rate, oil price and output gap are similar across all specifications and consistent with the results obtained in the univariate analysis (see Table 3.2 in the Appendix). The coefficient for lagged inflation is not significant at the 5% level, in line with the results obtained in the univariate analysis (except for the model MUC-3 (inflation)).

Figure 3.2 shows the probability of low trend GDP growth. A high probability of this regime is associated with all the recession episodes recorded in the euro area over the estimation period: 1982-1984, 1992-1993 and 2008-2009. However, there is less evidence for regime switching for the models using inflation since the increase in the log-likelihood for the regime switching models is modest with respect to the linear model (see the last row of Table 3.5). Figure 3.1 plots the actual level of inflation and the inflation expectations (i.e. the

time-varying parameters from the bivariate models MUC-1 (inflation), MUC-2 (inflation), MUC-3 (inflation) and the univariate time-varying Phillips curve).

3.3.3 Comparison of the estimated output gaps with the model-averaged measures

Figure 3.3 shows the output gaps estimated for the nine models under scrutiny. There are important differences between the estimates of the output gap, with a range max-min between the estimates of the output gap reaching high levels during the two important recessions identified in the sample: 4% in the beginning of 1992 and 5% in the beginning of 2010.

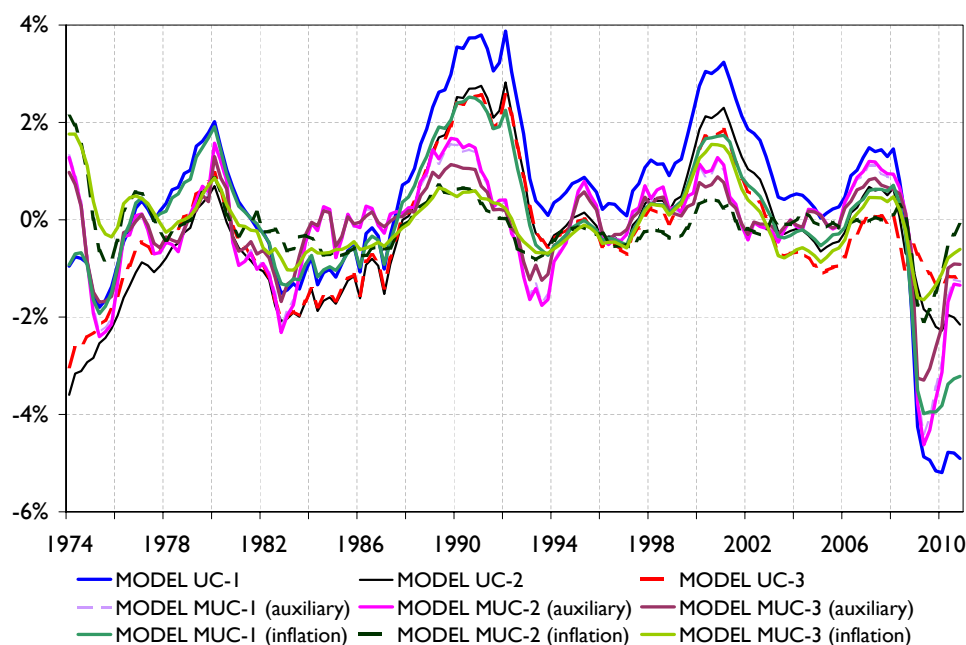
The output gaps estimated from the univariate models differ depending on whether there is regime switching or not. In particular, the output gap estimated from a linear univariate model captures well the expansions and recessions experienced by the euro area. However, the univariate regime switching models estimate a smaller negative output gap for the last recession, which suggests that the last recession also affected the level of potential output. The models with inflation tend to yield smoother estimates of the output gap and therefore allocate more variation to the trend of output. Conversely, the models with the rate of capacity utilization as an auxiliary indicator are very close to each other. They closely match the evolution of the euro area economic activity, and therefore allocate little variation to the level of potential output.

In the forecasting literature, it is often found that combining forecasts from different models allows to improve the forecasts from individual models (see e.g. Drechsel and Maurin (2011)). This is particularly relevant for estimating the output gap since the model estimates are characterized by both model uncertainty and parameter instability. Therefore, we compute three different model-averaged measures of the output gap: (i) one measure obtained as the simple arithmetic average over each of the nine models, labeled EST. 1, (ii) one measure obtained as the median estimate over each of the nine models, labeled EST. 2, (iii) the last measure takes into account the uncertainty in the estimation of the output gap and is labeled EST. 3. In particular, the latter measure gives higher weights $w_t(l)$ to the models with smaller variances attached to the estimated output gaps:

$$w_t(l) = \frac{[V_t(z_t^{(l)})]^{-1}}{\sum_{l=1}^L [V_t(z_t^{(l)})]^{-1}} \quad (3.17)$$

where $w_t(l)$ are the weights given to model l at time t , $z_t^{(l)}$ is the output gap from model

Figure 3.3: ESTIMATES OF THE EURO AREA OUTPUT GAP DERIVED FROM THE ESTIMATED UNOBSERVED COMPONENTS MODELS



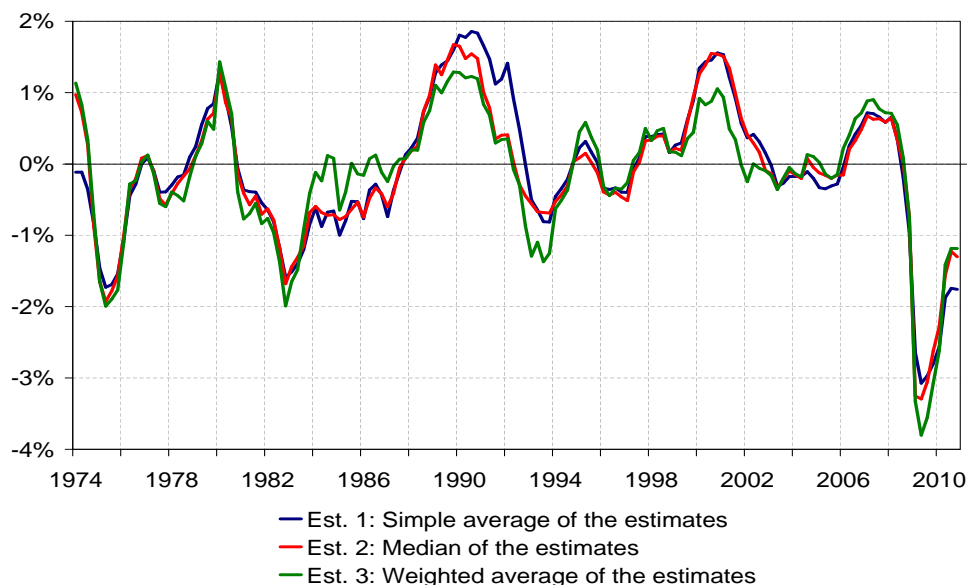
Note: MODEL UC-1 is the linear univariate trend-cycle decomposition of GDP, MODEL UC-2 is the univariate trend-cycle decomposition of GDP with a switch in the slope of the trend only and MODEL UC-3 is the univariate trend-cycle decomposition of GDP with a switch in both the slope of the trend of GDP and its error variance. MODEL MUC-1 (auxiliary) is the linear bivariate model with the trend-cycle decomposition of GDP and an equation for capacity utilization, MODEL MUC-2 (auxiliary) is the bivariate model with capacity utilization and a switch in the slope of the trend of GDP and MODEL MUC-3 (auxiliary) is the bivariate model with capacity utilization and a switch in the slope of the trend of GDP and its error variance. MODEL MUC-1 (inflation) is the linear bivariate model with the trend-cycle decomposition of GDP and an equation for inflation, MODEL MUC-2 (inflation) is the bivariate model with inflation and a switch in the slope of the trend of GDP and MODEL MUC-3 (inflation) is the bivariate model with inflation and a switch in the slope of the trend of GDP and its error variance.

l at time t and $V_t(z_t^{(l)})$ its corresponding variance estimated from the Kalman filter. In this way, the weights are time-varying, positive and sum to one.

Figure 3.4 plots the three model-averaged measures. The model-averaged measure with time-varying weights (labelled as "EST. 3") is more cyclical since it gives more weights to the models using the demeaned rate of capacity utilization, which yield more precise

estimates of the output gap (i.e. with a smaller variance). In particular, focussing on the most recent period, the amplitude of the model-averaged output gap estimates is largely reduced compared to the one of the initial nine estimates (from between -0.5 and -4.8 p.p. to between -1.2 p.p. and -1.8 p.p. at the end of 2010). The three model-averaged measures are overall fairly close unlike the estimates from the individual models, which shows the relevance of combining individual model estimates to obtain more reliable estimates of the output gap.

Figure 3.4: MODEL AVERAGED MEASURES OF THE NINE OUTPUT GAP ESTIMATES



3.4 Do real-time estimates of the output gap improve inflation forecasts?

3.4.1 Real-time estimates of the output gap

In this section, we compute the real-time estimates of the output gap measures estimated previously since it has long been advocated that there are severe differences between the real-time estimates of the output gap and their final vintages counterparts (see e.g. Orphanides and Van Norden (2002)). We only use the first releases of GDP to construct the real-time

measures of the euro area output gap. The first estimation sample goes from $t=1970:Q1$ to $t=2001:Q1$ and is recursively expanded until we reach the end of the estimation sample $T=2010:Q4$. We therefore obtain 40 different vintage series for the output gap, each of them being associated with a different date for its final observation (i.e. from $t=2001:Q1$ to $T=2010:Q4$). We run the pseudo real-time estimation exercise for the univariate trend-cycle decomposition models of GDP (i.e. MODEL UC-1, MODEL UC-2, MODEL UC-3) and for the bivariate model with capacity utilization⁷ as an auxiliary indicator (i.e. MODEL MUC-1 (auxiliary), MODEL MUC-2 (auxiliary), MODEL MUC-3 (auxiliary)).⁸

We also combine the individual estimates of the output gap in the three model-averaged measures detailed above. Figure 3.5 plots the range of revisions for each of the three different model-averaged measures, while Figures 3.6, 3.7 and 3.8 in the Appendix plot all the individual measures. In line with Orphanides and Van Norden (2002), we indeed find that the estimation of the output gap in real-time is associated with a large uncertainty. Figure 3.5 also shows that the equal weights and the median measures are associated with large revisions as the path of the output gap is changing significantly across the different vintages used (see also Figures 3.6 and 3.7 in the appendix), this is particularly acute during economic downturns. Conversely, the measure that gives time-varying weights depending on the uncertainty associated with the output gap is considerably less affected by GDP data revisions (see Figure 3.8 in the appendix). This comes from the fact that this measure gives heavy weights to the output gaps estimated with the demeaned rate of capacity utilization, which have a smaller variance than the output gaps estimated from a univariate model.

In the next sub-section, we assess the usefulness of our different output gap measures to predict inflation and also investigate the importance of data revisions for the predictive power of the output gap.

3.4.2 Inflation forecasts

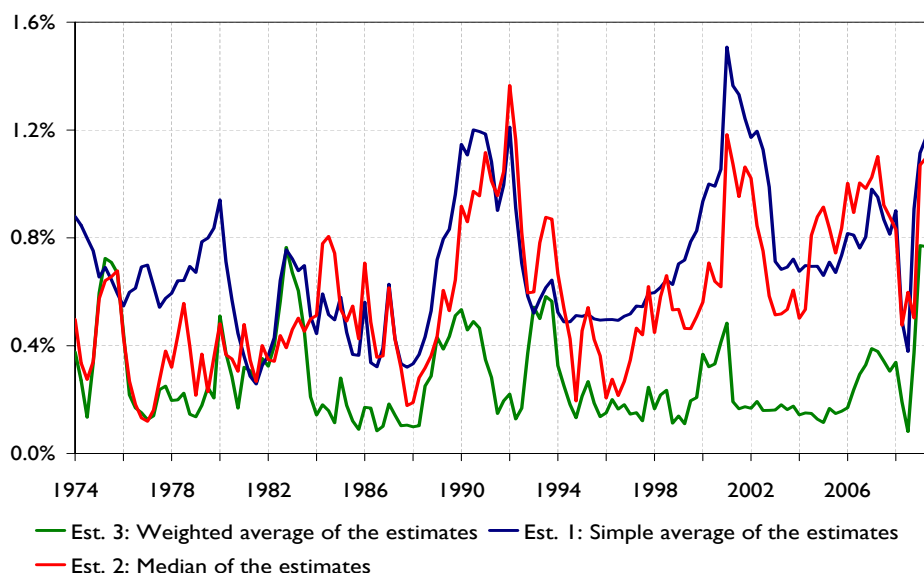
The predictive ability of the output gap for forecasting inflation is contrasted as it seems that the relation between inflation and output gap has weakened since the mid-1980s. Atkeson and Ohanian (2001) find that Phillips curve forecasts do not outperform simple univariate benchmarks. Stock and Watson (2008) extensively study Phillips curve forecasts using different sample periods, inflation series and benchmarks. They find that the Phillips curve predictive abilities are rather episodic and depend upon the evaluation sample chosen. Or-

⁷The rate of capacity utilisation is not revised over time.

⁸The bivariate models with an equation for inflation are not included since we encountered problems of convergence of the algorithm in the real-time exercise.

3.4. DO REAL-TIME ESTIMATES OF THE OUTPUT GAP IMPROVE INFLATION FORECASTS?87

Figure 3.5: ESTIMATES OF OUTPUT GAP IN REAL TIME: RANGE OF REVISIONS, (MAX-MIN)



phanides and van Norden (2005) also find that the forecasting performance of the output gap is unstable over time and point out the discrepancies between inflation forecasts based on real-time estimates of the output gap and their ex-post counterparts. However, the output gap - as a measure of economic slackness - is conceptually an intuitive predictive variable for inflation. Indeed, the triangle model of Gordon (1997) states that inflation depends on lagged inflation, the unemployment rate and supply shock variables.

The present forecasting exercise aims at assessing the predictive power of the output gap for forecasting inflation using the real-time estimates of the output gap as well as the ex-post estimates obtained from the last vintage of data available to us (T=2011:Q1).

We first consider the specification described in Orphanides and van Norden (2005) and forecast the level of inflation:

$$\pi_{t+h}^{(h)} = \alpha + \sum_{k=1}^P \beta_k \pi_{t-k} + \sum_{j=1}^J \gamma_j x_{t-j,\tau}^{(l)} + \epsilon_{PC,t+h} \quad (3.18)$$

The benchmark model is an AR(p) model for the level of inflation:

$$\pi_{t+h}^{(h)} = \alpha + \sum_{k=1}^P \beta_k \pi_{t-k}^{(1)} + \epsilon_{AR,t+h} \quad (3.19)$$

We also follow Stock and Watson (1999), Clark and McCracken (2009a) and use a Phillips curve for forecasting the change in inflation:

$$\pi_{t+h}^{(h)} - \pi_t = \alpha + \sum_{k=1}^P \beta_k \Delta \pi_{t-k} + \sum_{j=1}^J \gamma_j \Delta x_{t-j,\tau}^{(l)} + \epsilon_{PC,t+h} \quad (3.20)$$

The benchmark model is instead defined as:

$$\pi_{t+h}^{(h)} - \pi_t = \alpha + \sum_{k=1}^P \beta_k \Delta \pi_{t-k} + \epsilon_{AR,t+h} \quad (3.21)$$

where: $\pi_{t+h}^{(h)} = (\frac{400}{h}) \ln(\frac{p_{t+h}}{p_t})$, $\pi_t = 400 \ln(\frac{p_t}{p_{t-1}})$, and $x_{t,\tau}^{(l)}$ is a real-time measure of the output gap from model l at time t using data for GDP from the data vintage τ and $\Delta x_{t,\tau}^{(l)}$ is its quarterly difference.

The design of the pseudo-out-of-sample forecasting exercise is the following. The forecasts are computed with the direct method and the maximum lag lengths P and J are chosen with the SIC (maximum lag of 8) using the first estimation sample of the recursive forecasting exercise.⁹ We do not select recursively the number of lags since the Clark and McCracken (2009a) test of equal predictive accuracy with real-time data requires the number of parameters to be constant within each forecasting experiment. Appendix 3.6.3 details the tests of equal forecast accuracy with real-time and revised data. The real-time measures of the output gap are obtained from the previous subsection. We estimate equations 3.18 to 3.21 with OLS and compute for a given model i with forecast error $\hat{u}_{i,t+\tau}$ the mean squared forecast error (MSE_i):

$$MSE_i = (P - \tau + 1)^{-1} \sum_{t=R}^T \hat{u}_{i,t+\tau}^2$$

where R is the initial forecast origin and $(P - \tau + 1)$ is the number of forecast errors. The actual value for inflation is taken from the last vintage of data available to us (i.e.

⁹Selecting the lag length recursively or using the AIC rather than the SIC does not change qualitatively the results.

T=2011Q1). We compute forecasts one-quarter-ahead ($h = 1$), two-quarter-ahead ($h = 2$), one-year-ahead ($h = 4$) and two-year-ahead ($h = 8$).

A few additional comments are required. Note first that we do not consider additional explanatory variables in the Phillips curve equations 3.19 and 3.21 since we explicitly focus our analysis on the predictive power of the output gap for inflation and consider inflation excluding energy. Second, we only use real-time data for our output gap measures since we want to concentrate our analysis on the importance of data revisions to GDP for the predictive power of the output gap for inflation. Data revisions to HICP excluding energy - our measure of inflation here - are usually very small and are unlikely to affect our results. In this respect, we follow Orphanides and Van Norden (2005). Third, we consider two evaluation samples for our forecasting exercise 2001-2007:Q4 and 2001-2010:Q4 in order to assess the impact of the last recession on our results. Finally, we do not use the model with a time-varying constant in the forecasting exercise since it brings an additional source of uncertainty in the model that could cloud the interpretation of the results on the forecasting performance of the output gap.

Table 3.1 reports the forecasting results for the forecasts of the level of inflation. The benchmark model yields better forecasts than the Phillips curve specifications based on the real-time estimates of the output gap for forecasting the level of inflation for one-quarter-ahead forecasts (see Panel A of Table 3.1). Conversely, the models with the real-time estimates of the output gap improve the forecasting performance of the benchmark model for forecasting horizons $h = \{2, 4, 8\}$ (except for the UC-3, MUC-1, MUC-2, MUC-3 and Est.3 models with $h = 2$). However, the improvement in forecasting performance seems to be only of marginal importance since the increase in MSE is always lower than 20% and often inferior to 10%. Besides, the Clark and McCracken (2009a) test of equal predictive ability with real-time data cannot reject the null hypothesis of equal forecast accuracy except for the UC-1 and UC-2 models at the forecasting horizon $h = 4$ and $h = 2$ respectively. Interestingly, excluding the 2008-2010 period from the evaluation sample improves the forecasting performance for all models for one-quarter-ahead forecasts. In addition, nearly all models with forecast horizons $h = 8$ statistically improve the forecasts with respect to the benchmark model (except for the Est. 3 model) (see Panel C of Table 3.1).

Besides, using the ex-post estimates worsens the one-quarter-ahead forecasts and do not clearly improve forecasts for forecast horizon $h > 1$ with respect to the forecasts that use the real-time estimates of the output gap (see Panel B and D of Table 3.1). However, the improvement in forecasting performance is often significant when using the ex-post estimates of the output gap for forecast horizons $h = 2$ and $h = 8$. The p-values are computed from the

Clark and McCracken (2005) test of equal forecast accuracy, which is described in Appendix 3.6.3.

Table 3.1: Forecasting comparison exercise: results for inflation (HICP excluding energy)

<i>Model</i>	UC-1	UC-2	UC-3	MUC-1 (auxiliary)	MUC-2 (auxiliary)	MUC-3 (auxiliary)	Est. 1	Est. 2	Est. 3
<i>Panel A. Real-time output gap series, 2001-2010</i>									
h=1	1.777	1.292	1.088	2.662	2.520	2.350	2.234	1.559	2.551
h=2	0.844	0.864 ^(b)	1.035	1.227	1.185	1.145	1.170	0.833	0.933
h=4	0.814 ^(a)	0.862	0.993	0.978	0.969	0.975	0.970	0.906	0.955
h=8	0.947	0.928	0.959	0.983	0.983	0.984	0.984	0.966	1.000
<i>Panel B. Ex-post output gap series, 2001-2010</i>									
h=1	1.781	1.420	1.427	2.953	2.881	2.725	2.638	1.909	2.660
h=2	0.804 ^(c)	0.923 ^(b)	0.925 ^(b)	1.104	1.065	1.065	1.078	0.820 ^(c)	0.891 ^(b)
h=4	0.797 ^(b)	0.936	0.939	0.970	0.962	0.968	0.959	0.884 ^(b)	0.900 ^(b)
h=8	0.932 ^(a)	0.936 ^(a)	0.936	0.979	0.977	0.980	0.978	0.953	0.967
<i>Panel C. Real-time output gap series, 2001-2007</i>									
h=1	1.247	1.551	1.081	1.199	1.202	1.234	1.237	1.206	1.478
h=2	0.829	1.018	1.017	0.738	0.727	0.717	0.721	0.817	0.774
h=4	0.903	1.001	0.998	0.923	0.919	0.923	0.928	0.957	0.967
h=8	0.850 ^(c)	0.804 ^(c)	0.823 ^(c)	0.829 ^(c)	0.829 ^(c)	0.829 ^(c)	0.831 ^(c)	0.861 ^(c)	0.905
<i>Panel D. Ex-post output gap series, 2001-2007</i>									
h=1	1.474	1.472	1.477	1.688	1.765	1.602	1.693	1.692	1.609
h=2	0.829 ^(c)	0.828 ^(c)	0.826 ^(c)	0.556 ^(c)	0.532 ^(c)	0.564 ^(c)	0.549 ^(c)	0.770 ^(c)	0.771 ^(c)
h=4	0.925 ^(a)	0.904 ^(a)	0.903 ^(a)	0.918	0.917	0.921	0.916	0.939 ^(a)	0.935 ^(a)
h=8	0.801 ^(a)	0.789 ^(b)	0.788 ^(b)	0.830 ^(a)	0.831 ^(a)	0.830 ^(a)	0.832 ^(a)	0.830 ^(a)	0.842 ^(a)

Note: Ratio of the mean squared forecast error between the forecasts obtained from a Phillips curve equation with a real-time measure of the output gap as a proxy for the activity-based measure and a benchmark model given by an $AR(p)$. Est. 1, Est. 2 and Est. 3 are the model averaged measures detailed in the text. The superscripts ^a, ^b and ^c indicate that the test of equal forecast accuracy rejects respectively the null hypothesis of equal forecast accuracy at significance levels of 10%, 5% and 1% level. Appendix 3.6.3 details the Clark and McCracken (2009a) test for real-time data and the Clark and McCracken (2005) test for equal forecast accuracy with revised data.

Table 3.6 in the appendix reports the forecasting results for the change in inflation. First, none of the models with the real-time estimates of the output gap as a predictor can outperform the benchmark model for forecasting the change in inflation (see Panel A of Table 3.6). This is particularly acute for one-quarter-ahead predictions for the output gap models using capacity utilization as an extra indicator. The reason for this poor forecasting performance is that these models estimate a very negative output gap for the 2008-2010 period, which translates into very low or negative forecasts for the change in inflation. Indeed, if we exclude the 2008-2010 period from the evaluation sample, the one- and two-quarter-ahead forecasts do improve with respect to the full evaluation sample 2001-2010, although they do not beat the autoregressive benchmark or not significantly (see Panel C of Table 3.6)¹⁰. Second, using the ex-post rather than the real-time estimates of the output gap does not clearly improve the forecasts for the change in inflation (see Panel B and Panel D of Table 3.6).

The evidence on the importance of the output gap for predicting inflation is therefore contrasted. On the one hand, the real-time measures of the output gap do improve the forecasts for the level of inflation for forecasting horizons $h = \{2, 4, 8\}$ but this improvement is mostly statistically insignificant. Besides, the forecasts for the change in inflation based on the real-time estimates of the output gap are always outperformed by a standard autoregressive benchmark for the change in inflation when using the full evaluation sample. The use of the ex-post estimates of the output gap does not clearly improve forecasts for forecasting both the change and the level of inflation with respect to the forecasts based on the real-time output gap estimates.

3.5 Conclusions

This paper estimates various trend-cycle decomposition models of the euro area GDP using state-space models. We consider univariate and multivariate models as well as linear and non linear models. Non linearities are modelled with regime changes in the intercept of the trend equation and/or in the variance of its innovation. Multivariate models consider alternatively inflation and the demeaned rate of capacity utilization as additional variables in the system to better estimate the output gap. The univariate non linear specifications point out evidence for regime changes in the slope of the trend equation for GDP for few periods around 1974 and 2009. Besides, the demeaned rate of capacity utilization proves to

¹⁰Using the level of the output gap rather than the change in the output gap in equation 3.20 worsens the forecasting results.

be useful for obtaining more reliable estimates of the output gap by reducing the uncertainty associated with the output gap estimates.

We also conduct a real-time analysis for computing real-time estimates of the output gap. We find that our model-averaged estimates of the output gap decrease the uncertainty surrounding the output gap estimates and soften the impact of data revisions. We then carry out a forecasting exercise for evaluating the predictive power of the output gap estimates (both ex-post and real-time) to predict inflation. We find that the output gap does not help to predict the changes in inflation, whereas it does improve the forecasts for the level of inflation although this improvement is mostly statistically insignificant. In general, our real-time output gap measures do not statistically improve the forecasts for inflation with respect to a standard autoregressive model for inflation. Overall, the inflation forecasts based on the output gap dramatically fail since the last recession since they underestimate the level of inflation.

One possible avenue for further research on this topic would be to exploit the regime changes in the variance-covariance matrix of the innovations of the measurement and transition equations in order to obtain identification of more complicated trend-cycle decomposition models of the output gap. This could be done along the lines of Rigobon (2003) and Lanne et al. (2010). Alternatively, it would also be interesting to estimate models with mixed-frequency data to provide timely estimates of the output gap and therefore check whether this provides relevant information for estimating the output gap.

3.6 Appendix Chapter 3

3.6.1 State-space representation of the original Clark (1987) model of trend-cycle decomposition of output

$$\begin{bmatrix} y_t \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} n_t \\ z_t \\ z_{t-1} \end{bmatrix} \quad (3.22)$$

$$\begin{bmatrix} n_t \\ z_t \\ z_{t-1} \end{bmatrix} = \begin{bmatrix} \mu \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & \phi_1 & \phi_2 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} n_{t-1} \\ z_{t-1} \\ z_{t-2} \end{bmatrix} + \begin{bmatrix} \epsilon_t^n \\ \epsilon_t^z \\ 0 \end{bmatrix} \quad (3.23)$$

3.6.2 Kalman filter, Hamilton filter and Kim and Nelson filtering procedure

The equations of the basic Kalman filter can be found in a standard time series textbook such as Luetkepohl (2005). The general representation for a state-space model with regime switching in both measurement and transition equations is given by:

$$y_t = H(S_t)\beta_t + A(S_t)z_t + e_t$$

$$\beta_t = \tilde{\mu}(S_t) + F(S_t)\beta_{t-1} + G(S_t)v_t$$

where: $e_t \sim N(0, R(S_t))$, $v_t \sim N(0, Q(S_t))$, and e_t and v_t are not correlated.

Kim and Nelson (1999b) show how to combine the Kalman and Hamilton filters in a tractable way, the equations of the Kim and Nelson (1999b) filtering procedure for state-space models with regime switching are:

$$\beta_{t|t-1}^{(i,j)} = \tilde{\mu}_j + F_j \beta_{t-1|t-1}^i$$

$$P_{t|t-1}^{i,j} = F_j P_{t-1|t-1}^i F_j' + G Q_j G_j'$$

$$\eta_{t|t-1}^{(i,j)} = y_t - H_j \beta_{t|t-1}^{(i,j)} - A_j z_t$$

$$f_{t|t-1}^{(i,j)} = H_j P_{t|t-1}^{(i,j)} H_j' + R_j$$

$$\begin{aligned}\beta_{t|t}^{(i,j)} &= \beta_{t|t-1}^{(i,j)} + P_{t|t-1}^{(i,j)} H_j' [f_{t|t-1}^{(i,j)}]^{-1} \eta_{t|t-1}^{(i,j)} \\ P_{t|t}^{(i,j)} &= (I - P_{t|t-1}^{(i,j)} H_j' [f_{t|t-1}^{(i,j)}]^{-1} H_j) P_{t|t-1}^{(i,j)}\end{aligned}$$

When there is regime switching, it is also necessary to introduce approximations at the end of the Kalman and Hamilton filters to avoid the proliferation of cases to be considered:

$$\begin{aligned}\beta_{t|t}^j &= \frac{\sum_{i=1}^M \Pr[S_{t-1} = i, S_t = j | \Psi_t] \beta_{t|t}^{(i,j)}}{\Pr[S_t = j | \Psi_t]} \\ P_{t|t}^j &= \frac{\sum_{i=1}^M \Pr[S_{t-1} = i, S_t = j | \Psi_t] \{P_{t|t}^{(i,j)} + (\beta_{t|t}^j - \beta_{t|t}^{(i,j)})(\beta_{t|t}^j - \beta_{t|t}^{(i,j)})'\}}{\Pr[S_t = j | \Psi_t]}\end{aligned}$$

3.6.3 Clark and Mc Craken tests for comparing forecasting performance

We first detail the test of equal forecast accuracy with real-time data.

Denote P the number of forecasts, R the sample size at the initial forecast origin, T the full sample size, τ the forecast horizon, $\hat{u}_{2,t+\tau}$ the forecast error in model 2, \bar{d} the squared forecast loss differential between model 1 and model 2, k_1 the number of parameters in the benchmark model (i.e. model 1), k_{22} the number of excess parameters in model 2.

The Clark and McCracken (2009a) test statistic S for comparing predictive accuracy for nested models with real-time data is given by:

$$S = \frac{P^{\frac{1}{2}} \bar{d}}{\sqrt{\Omega}}$$

Under the null hypothesis of equal predictive accuracy:

$$S \xrightarrow{A} N(0, 1)$$

This differs from Clark and McCracken (2005), where simulated critical values are required in the tests of equal predictive accuracy for nested models. The use of real-time data instead strongly changes the asymptotic for these tests and allows to use standard normal tables for inference as long as we can obtain an asymptotically valid long run variance for Ω . A consistent asymptotic long run variance of the scaled forecasting loss differential Ω is:

$$\Omega = 2(1 - \pi^{-1} \ln(1 + \pi)) \mathbf{F}(-\mathbf{J}\mathbf{B}_1\mathbf{J}' + \mathbf{B}_2) \mathbf{S}_{hh}(-\mathbf{J}\mathbf{B}_1\mathbf{J}' + \mathbf{B}_2) \mathbf{F}'$$

where:

$$\begin{aligned}\pi &= \frac{P}{R} \\ \mathbf{J}' &= (\mathbf{I}_{\mathbf{k}_1 \times \mathbf{k}_1}, \mathbf{0}_{\mathbf{k}_1 \times \mathbf{k}_{22}}) \\ \hat{\mathbf{B}}_i &= (T^{-1} \sum_{s=1}^{T-\tau} \mathbf{x}_{i,s} \mathbf{x}_{i,s}')^{-1} \\ \hat{\mathbf{F}} &= 2[P^{-1} \sum_{t=R}^T \hat{u}_{2,t+\tau} \mathbf{x}_{2,t}']\end{aligned}$$

The long run variance $\hat{\mathbf{S}}_{hh}$ is obtained by weighting the relevant leads and lags of Γ_{hh} following Newey and West's (1987) HAC estimator with a bandwidth of 2τ , where

$$\hat{\Gamma}_{hh}(j) = E \mathbf{h}_{t+\tau} \mathbf{h}_{t+\tau-j}'$$

and

$$\mathbf{h}_{t+\tau} = (y_{s+\tau} - \mathbf{x}_{2,s} \boldsymbol{\beta}_{2,T}) \mathbf{x}_{2,s}$$

We also compute the Clark and McCracken (2009a) MSE-F test statistic for equal predictive ability of two nested models with revised data as follows:

$$MSE - F = \frac{\sum_{t=R}^{T-\tau} \hat{d}_{t+\tau}}{MSE_2}$$

where $\hat{d}_{t+\tau}$ is the difference between the squared forecast errors $\hat{d}_{t+\tau} = \hat{u}_{1,t+\tau}^2 - \hat{u}_{2,t+\tau}^2$, and MSE_2 is the mean squared forecast error of model 2. We implement the novel bootstrapping procedure described in Clark and McCracken (2009b) and compute the p-values from 1000 replications.

3.6.4 Additional Tables and Figures

Figure 3.6: EURO AREA OUTPUT GAP, REAL-TIME DATA, MODEL-AVERAGED MEASURES WITH EQUAL WEIGHTS, 1974Q1-2010Q4

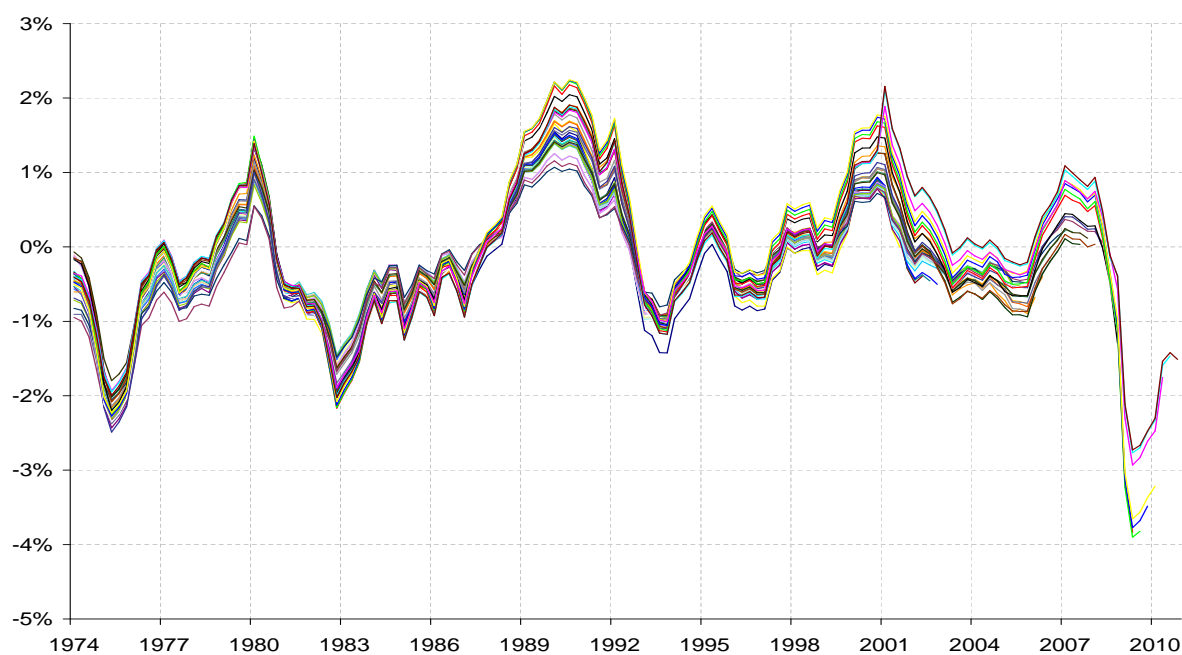


Figure 3.7: EURO AREA OUTPUT GAP, REAL-TIME DATA, MODEL-AVERAGED MEASURES COMPUTED AS THE MEDIAN OF THE INDIVIDUAL OUTPUT GAPS, 1974Q1-2010Q4

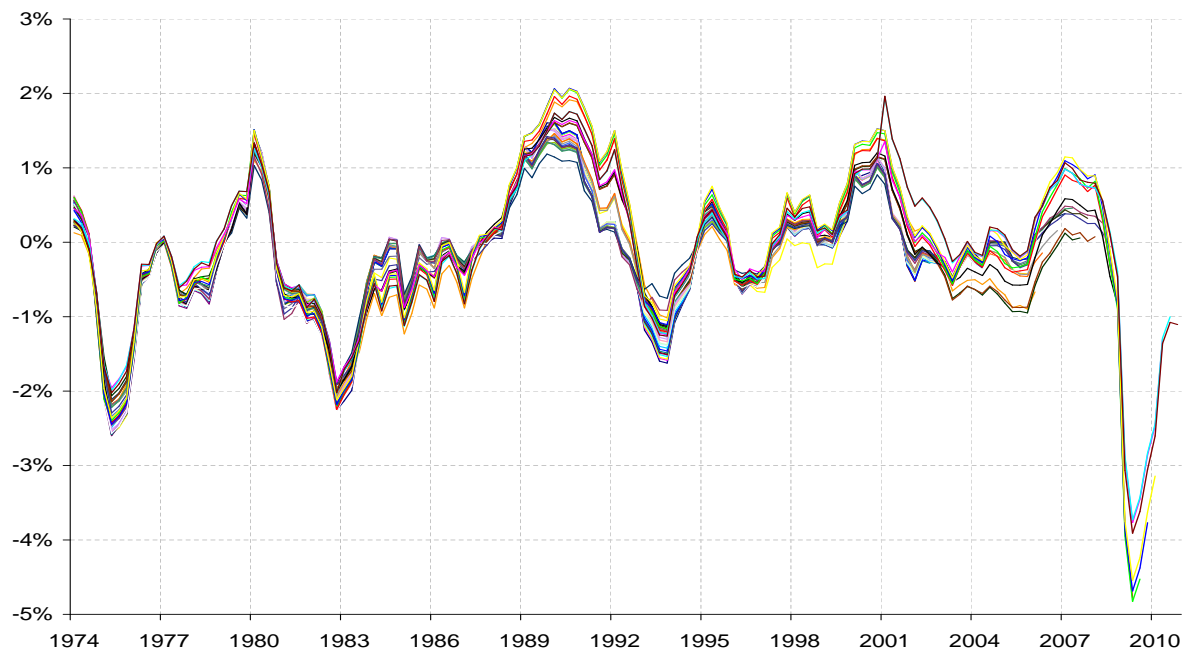


Figure 3.8: EURO AREA OUTPUT GAP, REAL-TIME DATA, MODEL-AVERAGED MEASURES WITH TIME-VARYING WEIGHTS, 1974Q1-2010Q4

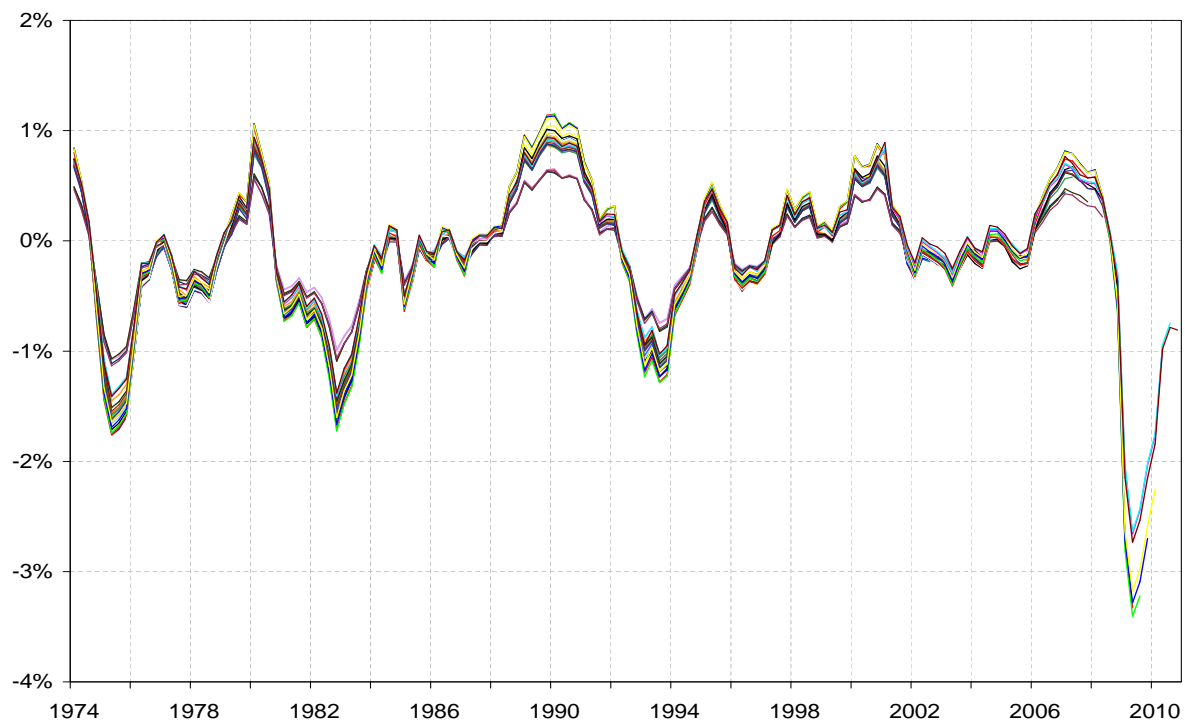
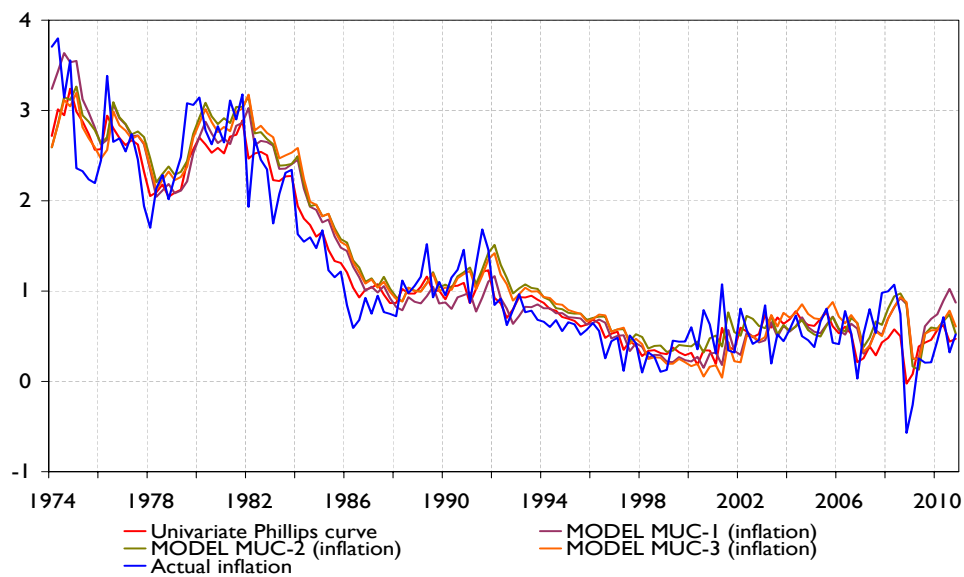


Figure 3.9: ACTUAL INFLATION AND INFLATION EXPECTATIONS FROM UC AND MUC MODELS



Note: MODEL MUC-1 (inflation) is the linear bivariate model with the trend-cycle decomposition of GDP and an equation for inflation, MODEL MUC-2 (inflation) is the bivariate model with inflation and a switch in the slope of the trend of GDP and MODEL MUC-3 (inflation) is the bivariate model with inflation and a switch in the slope of the trend of GDP and its error variance.

Table 3.2: Time-varying Phillips curve

z_t	0.149*** [0.033]
EXR_{t-1}	-0.017* [0.009]
OIL_{t-1}	$2.838 * 10^{-3}$ * [$1.500 * 10^{-3}$]
σ_π^2	0.227*** [0.019]
σ_μ^2	0.146*** [0.022]
Log(L)	-40.093

Note: Maximum likelihood estimates of the time-varying Phillips curve (see equations 3.11 and 3.12). We imposed positive definiteness constraints on the variance parameters of the innovations. The measure for the output gap z_t is the cycle computed by the HP filter. ***, ** and * indicate significance at 1%, 5% and 10%. Standard deviations are reported in brackets and are computed from the inverse of the outer product estimate of the Hessian. $Log(L)$ is the value of the log likelihood function.

Table 3.3: Univariate model

	Model UC-1	Model UC-2	Model UC-3
p_{11}	-	0.491*** [0.038]	0.568** [0.098]
p_{22}	-	0.986*** [0.010]	0.982*** [0.014]
ϕ_1	1.236*** [0.032]	1.194*** [0.116]	1.205*** [0.026]
ϕ_2	-0.292*** [0.046]	-0.237*** [0.114]	-0.251*** [0.037]
μ_1	$5.147 * 10^{-3}***$ [$0.227 * 10^{-3}$]	$-17.0 * 10^{-3}***$ [$3.977 * 10^{-3}$]	$-12.2 * 10^{-3}**$ [$5.219 * 10^{-3}$]
μ_2	-	$5.301 * 10^{-3}***$ [$0.261 * 10^{-3}$]	$5.495 * 10^{-3}***$ [$0.287 * 10^{-3}$]
$\sigma^{n,1}$	$1.5 * 10^{-3}$ [$1.964 * 10^{-3}$]	$1.5 * 10^{-3}$ [$3.021 * 10^{-3}$]	$6.749 * 10^{-3*}$ [$3.630 * 10^{-3}$]
$\sigma^{n,2}$	-	-	$1.5 * 10^{-3}$ [$2.221 * 10^{-3}$]
σ^z	$5.502 * 10^{-3}***$ [$0.615 * 10^{-3}$]	$4.490 * 10^{-3}***$ [$0.107 * 10^{-3}$]	$4.345 * 10^{-3}***$ [$0.848 * 10^{-3}$]
$P(S_t = 1)$	-	0.026	0.041
Log(L)	608.041	624.010	624.732

Note: Maximum likelihood estimates for the three univariate unobserved components models of trend-cycle decomposition of log GDP (see equations 3.22 and 3.23 in Appendix 3.6.1). Model UC-1 is the linear model described by equations 3.1 to 3.3. Model UC-2 is a model with a switch in the drift of the trend equation for the level of GDP. Model UC-3 is a model with switches in the drift of the trend equation and in the variance of the innovation for the trend component of GDP. $P(S_t = 1)$ is the unconditional probability of being in the first regime. $\text{Log}(L)$ is the value of the log likelihood function. Standard deviations are reported in brackets. ***, ** and * indicate significance at 1%, 5% and 10%.

Table 3.4: Bivariate model (GDP and rate of capacity utilization)

	Model MUC-1(auxiliary)	Model MUC-2(auxiliary)	Model MUC-3(auxiliary)
p_{11}	-	0.981*** [0.038]	0.823*** [0.151]
p_{22}	-	0.994*** [0.007]	0.958*** [0.022]
ϕ_1	1.371*** [0.028]	1.372*** [0.028]	1.349*** [0.023]
ϕ_2	-0.509*** [0.048]	-0.527*** [0.048]	-0.472*** [0.040]
μ_1	$5.698 * 10^{-3}$ *** [$0.374 * 10^{-3}$]	$11.9 * 10^{-3}$ *** [$1.101 * 10^{-3}$]	$0.693 * 10^{-3}$ [$1.539 * 10^{-3}$]
μ_2	-	$5.010 * 10^{-3}$ *** [$0.351 * 10^{-3}$]	$6.764 * 10^{-3}$ *** [$0.482 * 10^{-3}$]
α_1	0.962*** [0.269]	0.913*** [0.238]	0.842** [0.364]
α_2	1.840*** [0.305]	1.737*** [0.254]	2.753*** [0.574]
σ^{aux}	$2.856 * 10^{-3}$ *** [$0.356 * 10^{-3}$]	$2.936 * 10^{-3}$ *** [$0.337 * 10^{-3}$]	$2.504 * 10^{-3}$ *** [$0.444 * 10^{-3}$]
$\sigma^{n,1}$	$4.753 * 10^{-3}$ *** [$0.287 * 10^{-3}$]	$4.143 * 10^{-3}$ *** [$0.269 * 10^{-3}$]	$1.562 * 10^{-3}$ *** [$0.256 * 10^{-3}$]
$\sigma^{n,2}$	-	-	$4.469 * 10^{-3}$ *** [$0.290 * 10^{-3}$]
σ^z	$3.444 * 10^{-3}$ *** [$0.381 * 10^{-3}$]	$3.610 * 10^{-3}$ *** [$0.364 * 10^{-3}$]	$4.347 * 10^{-3}$ *** [$0.910 * 10^{-3}$]
$P(S_t = 1)$	-	0.229	0.192
Log(L)	1315.089	1327.591	1344.208

Note: Maximum likelihood estimates for the three bivariate unobserved components models of trend-cycle decomposition of log GDP (see equations 3.9 and 3.10). Model MUC-1 (auxiliary) is a model without regime switching. Model MUC-2 (auxiliary) is a model with a switch in the slope of the trend. Model MUC-3 (auxiliary) is a model with switches in the slope of the trend equation and in the variance of its error. $P(S_t = 1)$ is the unconditional probability of being in the first regime. $Log(L)$ is the value of the log likelihood function. Standard deviations are reported in brackets. ***, ** and * indicate significance at 1%, 5% and 10%.

Table 3.5: Bivariate model (GDP and inflation)

	Model MUC-1(inflation)	Model MUC-2(inflation)	Model MUC-3(inflation)
p_{11}	-	0.885*** [0.074]	0.974*** [0.022]
p_{22}	-	0.948*** [0.032]	0.989*** [0.012]
ϕ_1	1.285*** [0.044]	1.406*** [0.063]	1.351*** [0.052]
ϕ_2	-0.367*** [0.069]	-0.571*** [0.114]	-0.475*** [0.089]
μ_1	$5.085 * 10^{-3}$ *** [$0.277 * 10^{-3}$]	$0.91610 * 10^{-3}$ [$1.542 * 10^{-3}$]	$5.265 * 10^{-3}$ *** [$0.275 * 10^{-3}$]
μ_2	-	$7.332 * 10^{-3}$ *** [$0.661 * 10^{-3}$]	$5.567 * 10^{-3}$ *** [$0.648 * 10^{-3}$]
λ^π	-0.044 [0.388]	-0.086 [0.001]	-0.069** [0.034]
λ^z	0.157* [0.081]	0.259*** [0.071]	0.289*** [0.074]
λ^{EXR}	-0.017* [0.010]	-0.017* [0.007]	-0.015* [0.009]
λ^{OIL}	$3.018 * 10^{-3}$ * [$1.706 * 10^{-3}$]	$2.751 * 10^{-3}$ * [$1.470 * 10^{-3}$]	$2.928 * 10^{-3}$ * [$1.509 * 10^{-3}$]
σ^π	$2.181 * 10^{-3}$ *** [$0.621 * 10^{-3}$]	$2.088 * 10^{-3}$ *** [$0.180 * 10^{-3}$]	$2.094 * 10^{-3}$ *** [$0.209 * 10^{-3}$]
$\sigma^{n,1}$	$2.961 * 10^{-3}$ ** [$0.126 * 10^{-3}$]	$3.792 * 10^{-3}$ *** [$0.568 * 10^{-3}$]	$6.324 * 10^{-3}$ *** [$0.611 * 10^{-3}$]
$\sigma^{n,2}$	-	-	$1.5 * 10^{-3}$ * [$0.818 * 10^{-3}$]
σ^z	$4.708 * 10^{-3}$ *** [$0.833 * 10^{-3}$]	$3.209 * 10^{-3}$ *** [$0.772 * 10^{-3}$]	$2.830 * 10^{-3}$ *** [$0.515 * 10^{-3}$]
σ_μ	$1.5 * 10^{-3}$ * [$0.874 * 10^{-3}$]	$1.5 * 10^{-3}$ *** [$0.270 * 10^{-3}$]	$1.519 * 10^{-3}$ *** [$0.311 * 10^{-3}$]
$P(S_t = 1)$	-	0.311	0.292
Log(L)	1446.312	1446.861	1451.821

Note: Maximum likelihood estimates for the three univariate unobserved components models of trend-cycle decomposition of log GDP (see equations 3.15 and 3.16). Model MUC-1 is a model without regime switching, Model MUC-2 is with a switch in the slope of the trend, and Model MUC-3 incorporates switches in the slope of the trend and in the variance of its innovation. $P(S_t = 1)$ is the unconditional probability of being in the first regime. $Log(L)$ is the value of the log likelihood function. Standard deviations are reported in brackets. ***, ** and * indicate significance at 1%, 5% and 10%.

Table 3.6: Forecasting comparison exercise: results for the change in inflation (HICP excluding energy)

<i>Model</i>	UC-1	UC-2	UC-3	MUC-1 (auxiliary)	MUC-2 (auxiliary)	MUC-3 (auxiliary)	Est. 1	Est. 2	Est. 3
<i>Panel A. Real-time output gap series, 2001-2010</i>									
h=1	1.632	1.291	1.046	3.518	3.215	3.036	3.104	1.784	2.983
h=2	1.842	1.420	1.144	2.286	2.173	2.058	2.111	1.533	2.130
h=4	1.523	1.256	1.210	1.447	1.558	1.514	1.544	1.378	1.540
h=8	1.442	1.093	1.115	1.430	1.421	1.438	1.420	1.301	1.339
<i>Panel B. Ex-post output gap series, 2001-2010</i>									
h=1	1.757	1.310	1.318	3.785	3.634	3.398	3.545	2.110	2.654
h=2	1.376	1.311	1.316	2.280	2.159	2.159	2.276	1.396	1.807
h=4	1.331	1.141	1.145	1.686	1.717	1.593	1.691	1.442	1.558
h=8	1.268	1.252	1.257	1.421	1.398	1.409	1.404	1.372	1.346
<i>Panel C. Real-time output gap series, 2001-2007</i>									
h=1	1.244	1.481	1.061	1.451	1.444	1.494	1.459	1.205	1.578
h=2	1.417	1.587	1.094	0.859	0.875	0.886	0.882	1.107	1.031
h=4	1.122	1.377	1.058	1.219	1.293	1.289	1.283	1.095	1.216
h=8	2.280	1.080	1.043	1.630	1.651	1.632	1.638	1.647	1.503
<i>Panel D. Ex-post output gap series, 2001-2007</i>									
h=1	1.537	1.318	1.323	2.103	2.208	2.014	2.147	1.901	1.818
h=2	1.066	1.033	1.034	0.941	0.958	1.047	1.126	0.982	1.031
h=4	0.928	0.958	0.962	1.615	1.657	1.478	1.598	1.262	1.315
h=8	1.673	1.762	1.774	1.659	1.677	1.631	1.675	1.777	1.567

Note: Ratio of the mean squared forecast error between the forecasts obtained from a Phillips curve equation with a real-time measure of the output gap as a proxy for the activity-based measure and a benchmark model given by an $AR(p)$. Est. 1, Est. 2 and Est. 3 are the model averaged measures detailed in the text. The superscripts ^a, ^b and ^c indicate that the test of equal forecast accuracy rejects respectively the null hypothesis of equal forecast accuracy at significance levels of 10%, 5% and 1% level. Appendix 3.6.3 details the Clark and McCracken (2009a) test for real-time data and the Clark and McCracken (2005) test for equal forecast accuracy with revised data.

Chapter 4

Switches in the risk-return trade-off

Abstract

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, related to the extent of volatility. This finding is robust to a large range of specifications. In the high volatility and low ex-post returns regime, the risk-return relation is reversed, whereas the intuitive positive risk-return trade-off holds in the low volatility and high ex-post returns regime. The high volatility and low ex-post returns regime is interpreted as a "flight-to-quality" regime with a negative premium for volatility.

Keywords: Risk-return trade-off; Markov-switching; MIDAS; conditional variance.

JEL Classification Code: G12, G10.

⁰This is a joint work with Eric Ghysels and Massimiliano Marcellino

4.1 Introduction

The risk-return trade-off implies that a riskier investment should demand a higher expected return relative to the risk-free return. The ICAPM of Merton (1973) rigorously formalizes this intuition and states that the expected excess return on the stock market is positively related to its conditional variance:

$$E_t(R_{t+1}) = \mu + \gamma V_t(R_{t+1}) \quad (4.1)$$

However, the literature often finds conflicting results as it is not clear whether the coefficient γ enters positively and significantly before the conditional variance. For example, 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. Ghysels et al. (2005), Guo and Whitelaw (2006) and Ludvigson and Ng (2007) all find a positive risk-return trade-off. However, Glosten et al. (1993), using different GARCH specifications, find a negative relation between risk and return. 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.

A possible reason for these conflicting results is the lack of conditioning variables in equation 4.1. Scruggs (1998) finds that it is crucial to control for shifts in investment opportunities to obtain a positive risk-return relation. Lettau and Ludvigson (2001) point out the relevance of the consumption-wealth ratio as a conditioning variable in the ICAPM. Similarly, Guo and Whitelaw (2006) argue that the specification in equation 4.1 misses the hedge component of Merton's model. They include additional predictors in equation 4.1 to account for changing investment opportunities and uncover a positive risk-return relation. In the same vein, Ludvigson and Ng (2007) show that including factors in equation 4.1 leads to a positive relation between risk and return. Finally, the literature review by Lettau and Ludvigson (2010) emphasizes the importance of conditioning variables in the estimation of the risk-return trade-off. In particular, they find a positive conditional correlation between risk and return that is strongly significant (conditional on lagged mean and lagged volatility), whereas the unconditional risk-return relation is weakly negative and statistically insignificant.

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. The relation between risk and return should not necessarily be linear. For example, 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 test 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 computed 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, 3-week and monthly 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. (2009).

- In the high volatility and low ex-post returns regime, the risk-return relation is negative, whereas the risk-return relation is positive in the low volatility and high ex-post returns 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 high volatility regime can be interpreted as a "flight-to-quality" regime with a negative premium for volatility.

The paper is structured as follows. Section 4.2 presents the model we use for estimating the risk-return relation. Section 4.3 details the main results of the paper and a comparison of the estimated conditional variances with GARCH specifications. Section 4.4 provides a sensitivity analysis across a wide range of models. Section 4.5 concludes.

4.2 The risk-return relation with regime switching

If returns are normally distributed, the risk-return trade-off with a MIDAS estimator of the conditional variance as measure of risk is such that:

$$R_{t+1} \sim N(\mu + \gamma V_t^{MIDAS}, V_t^{MIDAS}) \quad (4.2)$$

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. Equation 4.2 then becomes:

$$R_{t+1} \sim N(\mu + \gamma(S_t) V_t^{MIDAS}, V_t^{MIDAS}) \quad (4.3)$$

where S_t is an M-state Markov chain defined by the following constant transition probabilities:

$$p_{ij} = Pr(S_{t+1} = j | S_t = i) \quad (4.4)$$

$$\sum_{j=1}^M p_{ij} = 1 \forall i, j \in \{1, \dots, M\} \quad (4.5)$$

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:

$$w(j; \theta) = \frac{\left(\frac{d}{D}\right)^{\kappa_1} - \left(1 - \frac{d}{D}\right)^{\kappa_2 - 1}}{\sum_{j=0}^K \left(\frac{j}{D}\right)^{\kappa_1} - \left(1 - \frac{j}{D}\right)^{\kappa_2 - 1}} \quad (4.6)$$

$$w(j; \theta) = \frac{\exp(\kappa_1 j + \kappa_2 j^2)}{\sum_{j=0}^K \exp(\kappa_1 j + \kappa_2 j^2)} \quad (4.7)$$

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) 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:

$$V_t^{MIDAS} = N \sum_{d=0}^D w_j |r_{t-d}| \quad (4.8)$$

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 scale¹.

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)).

Several papers have already 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 relation². Mayfield (2004) introduces regime switching in a

¹ $N = \{5, 10, 15, 22\}$ for regressions at 1-week, 2-week, 3-week and monthly horizons.

²In particular, in a general equilibrium exchange economy, the sign of the risk-return relation depends on

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. However, they do not model explicitly the process for volatility but only the underlying regime probabilities for volatility at the monthly frequency.

4.3 Data and empirical results

4.3.1 Data

We use the S&P 500 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 4.1 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 4.1 plots the data.

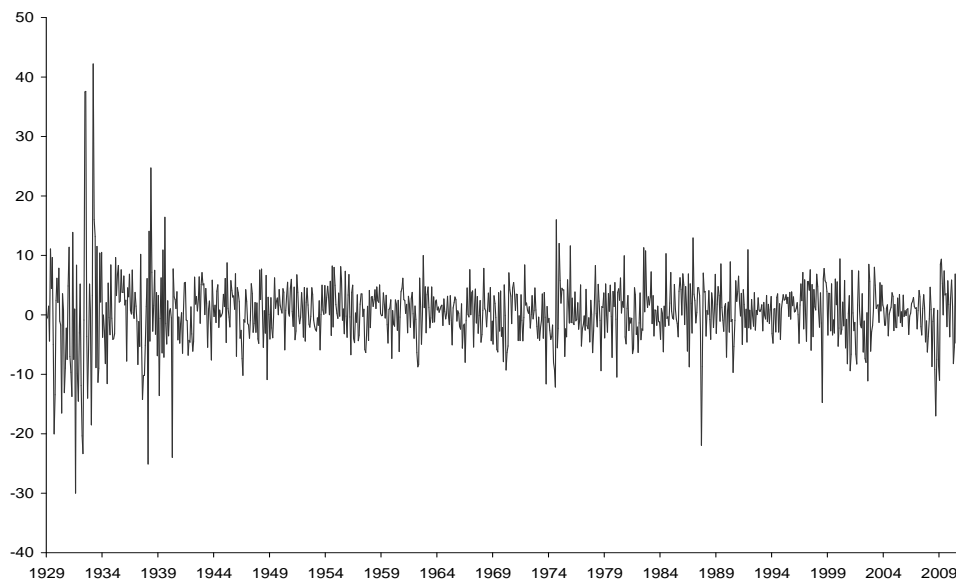
the sign of the correlation in between the marginal rate of substitution (or "stochastic discount factor") and the market return (see e.g. Whitelaw (2000)). Therefore, the parameter $\gamma(S_t)$ entering before the conditional variance in equation 4.3 is not directly interpretable as the coefficient of relative risk-aversion. Instead, $\gamma(S_t)$ is proportional to the product of the volatility of the stochastic discount factor and the correlation between the stochastic discount factor and the market return.

Table 4.1: 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.

Figure 4.1: MONTHLY EXCESS STOCK RETURNS 1929:02-2010:12



4.3.2 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

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&P 500 data ³. 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 4.1.

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 4.2 reports the empirical results for the linear tests of the risk-return trade-off using returns R_{t+1} for the LHS of equation 4.1 ranging from the weekly to the monthly 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 at 1-week and 2-week horizons. The last column of Table 4.2 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 tends to be higher at the monthly frequency than at the weekly frequency. Indeed, Figure 4.2 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 computed 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.

Another way to model the conditional variance is to use GARCH specifications. The GARCH-in-mean specification is another test 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:

$$R_t = \mu + \gamma V_t^{GARCH} + \epsilon_t \quad (4.9)$$

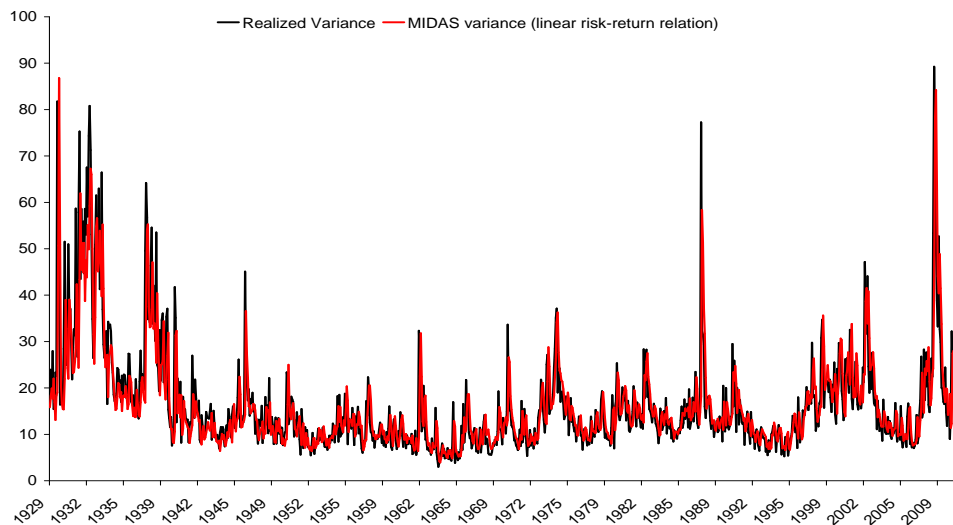
³The use of exponential Almon lag weight function yields qualitatively similar results.

Table 4.2: Linear risk return relation: $R_{t+1} \sim N(\mu + \gamma V_t^{MIDAS}, V_t^{MIDAS})$

	μ (*10 ²)	γ	$LogL$	$R_{\sigma^2}^2$
<i>Full sample analysis: February 1929 - December 2010</i>				
Monthly	0.262 [1.334]	0.009 [0.664]	-2987.476	58.74%
3-week	0.146 [0.748]	0.012 [0.584]	-4082.264	56.77%
2-week	0.088 [0.828]	0.013 [0.750]	-5562.106	56.26%
1-week	0.066 [1.750]	0.007 [0.948]	-9532.267	50.56%
<i>Sub-sample analysis: February 1964 - December 2010</i>				
Monthly	0.271 [0.709]	0.008 [0.285]	-1593.271	54.12%
3-week	0.058 [0.307]	0.021 [0.946]	-2206.327	52.54%
2-week	-0.001 [-0.008]	0.026 [2.214]	-2992.930	54.08%
1-week	0.014 [0.783]	0.022 [1.626]	-5133.369	48.71%

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. $R_{\sigma^2}^2$ is the coefficient of determination when regressing the realized variance on V_t^{MIDAS} .

Figure 4.2: MIDAS AND REALIZED VARIANCES 1929:02-2010:12



$$V_t^{GARCH} = \omega + \alpha \epsilon_{t-1}^2 + \beta V_{t-1}^{GARCH} \quad (4.10)$$

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

$$(V_t^{ABSGARCH})^{1/2} = \omega + \alpha |\epsilon_{t-1}| + \beta (V_{t-1}^{ABSGARCH})^{1/2} \quad (4.11)$$

We use both Student-t innovations and Normal innovations and consider two different sample sizes (1929-2010 and 1964-2010). Table 4.3 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 $R_{\sigma^2}^2$ s are roughly equivalent to their MIDAS counterparts (see Table 4.2).

Figure 4.3 plots the ABSGARCH variance with the realized variance. Unlike the MIDAS variance, the ABSGARCH variance has troubles to accomodate the periods of high volatility

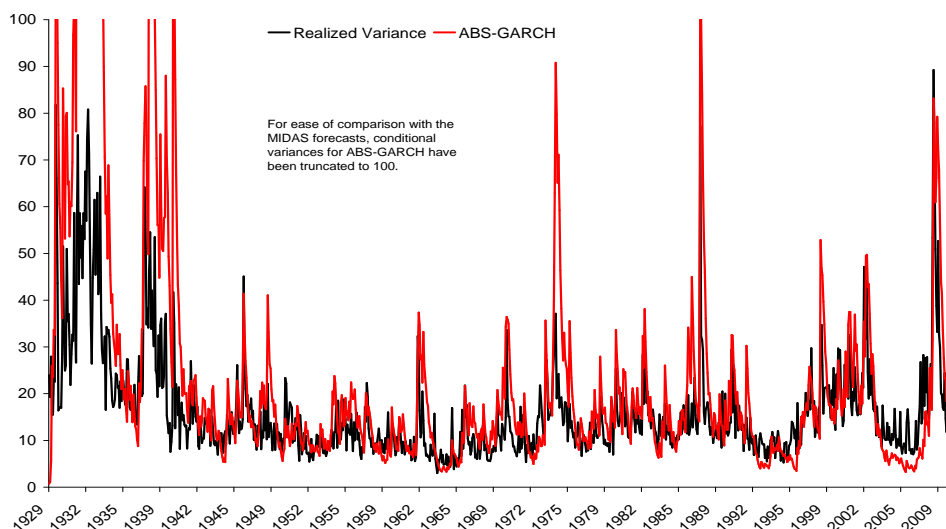
Table 4.3: Monthly GARCH estimates of the risk-return relation

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

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

ranging from 1929 to 1940.

Figure 4.3: ABSGARCH AND REALIZED VARIANCES 1929:03-2010:12



4.3.3 MIDAS estimates of the regime switching risk-return relation

Table 4.4 provides the estimates for the regime switching risk-return relation described by equation 4.3. For the full sample analysis (1929-2010), in all regressions, we find that the coefficient γ_1 is largely negative and significant, while the coefficient in the second regime γ_2 is positive and significant (except at the 3-week horizon where the coefficient γ_2 is not significant at the 10% level). 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 the same results except that the coefficient γ_2 is not significant at the 10% level at the monthly horizon ⁴.

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 2.44% and 12.26%) and are - as expected - higher in the full

⁴Table 4.14 in the appendix provides additional estimation results with different estimation window sizes. The results reported are consistent with those of Table 4.4.

estimation sample (1929-2010) than in the shorter estimation sample (1964-2010). Besides, in the regime switching case, the coefficients of determination $R_{\sigma^2}^2$ s are roughly equivalent to the linear case. The monthly MIDAS conditional variance obtained from the regime switching risk-return relation is very close to the monthly realized variance (see Figure 4.4).

Figure 4.4: MIDAS AND REALIZED VARIANCES 1929:02-2010:12

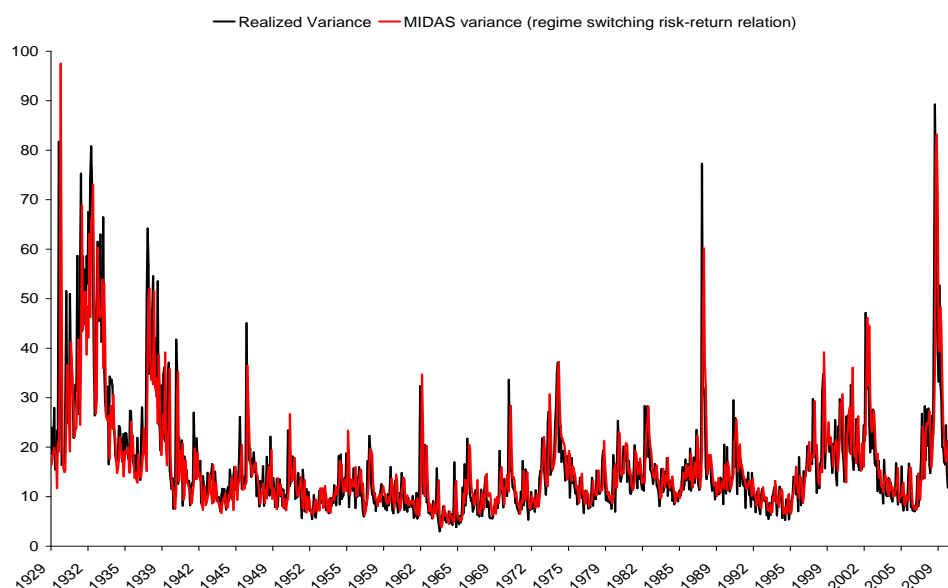
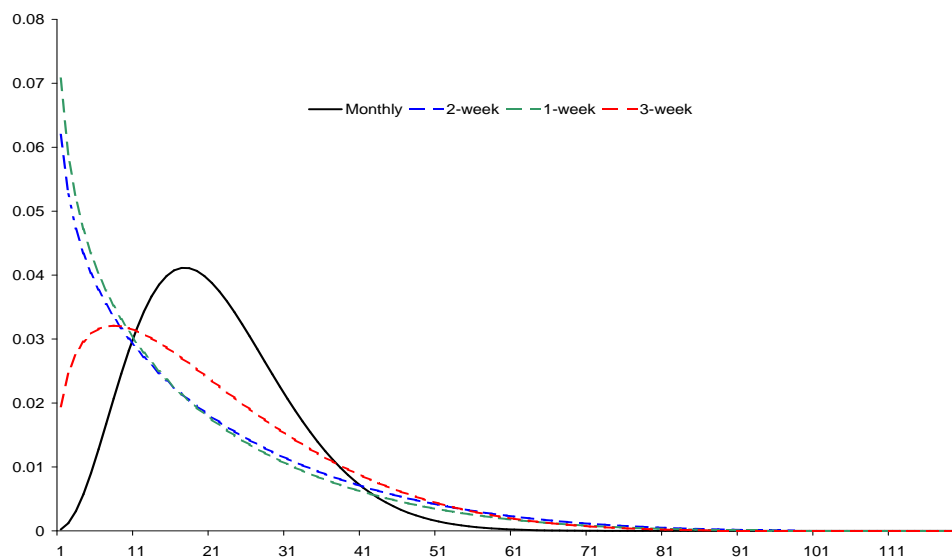


Figure 4.5 plots the weights attached to the lagged daily absolute returns at different frequencies for the regime-switching risk-return relation. For the 1-week and 2-week horizons, the weight function has a decreasing shape, whereas the weight function has a hump shape at the 3-week and monthly horizons. 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 4.6 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 regress 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. The results are reported in Table 4.5. First, the expected returns always affect negatively and significantly the regime probabilities. Second,

Figure 4.5: WEIGHTS FOR THE MIDAS ESTIMATOR OF THE CONDITIONAL VARIANCE



Note: The weights are computed from the full sample estimates of the regime switching risk-return relation at different frequencies.

an increase in volatility is positively and significantly related to the regime probabilities. Third, the slope of the yield curve affects negatively and significantly the regime probabilities, except at the 2-week horizon 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. The second regime is instead characterized by low volatility and high ex-post returns with a positive and significant risk-return relation. 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.

We now compare the different estimated variance processes in Table 4.6. Panel A reports the means, variances and goodness-of-fit measures for the MIDAS (for both linear and

Table 4.4: Regime-switching risk-return relation: $R_{t+1} \sim N(\mu + \gamma(S_t)V_t^{MIDAS}, V_t^{MIDAS})$

	μ (*10 ²)	γ_1	γ_2	$LogL$	$R_{\sigma^2}^2$	$P(S_t = 1)$
<i>Full sample analysis: February 1929 - December 2010</i>						
Monthly	0.222 [0.865]	-0.518 [-7.612]	0.060 [3.080]	-2916.692	54.48%	8.39%
3-week	0.071 [0.081]	-0.620 [-6.226]	0.085 [1.116]	-3964.888	56.72%	9.16%
2-week	-0.164 [-1.205]	-0.610 [-11.173]	0.137 [5.963]	-5411.815	56.38%	10.80%
1-week	-0.100 [-1.723]	-0.780 [-10.304]	0.182 [7.085]	-9272.321	50.79%	12.26%
<i>Sub-sample analysis: February 1964 - December 2010</i>						
Monthly	0.206 [0.548]	-0.419 [-2.908]	0.042 [1.371]	-1583.496	52.91%	6.48%
3-week	0.129 [0.315]	-0.755 [-7.376]	0.055 [2.396]	-2169.971	39.48%	5.28%
2-week	-0.030 [-0.184]	-1.041 [-5.131]	0.058 [2.396]	-2951.488	52.26%	2.44%
1-week	-0.037 [-0.888]	-0.963 [-7.700]	0.096 [4.629]	-5070.968	48.47%	5.27%

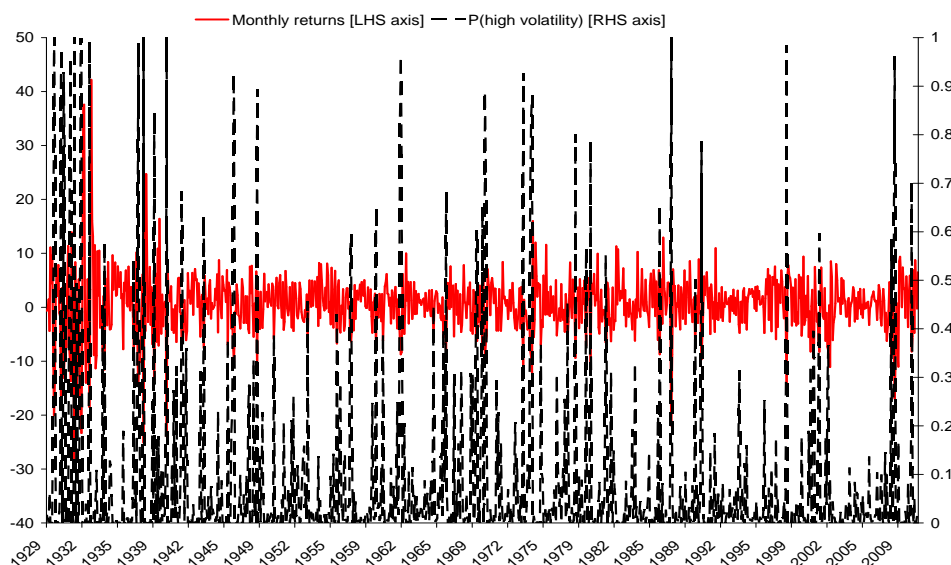
The MIDAS estimator of the conditional variance is computed using 120 lags for the daily absolute returns, which are aggregated with the beta polynomial weight function. T-statistics are computed from the inverse of the outer product estimate of the Hessian and are reported in brackets. $LogL$ is the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

Table 4.5: Explaining the regime probabilities $P(S_{t+1})$, Estimation sample: 1964:2010

	Slope of the yield curve $_{t+1}$	ΔV_{t+1}^{MIDAS}	R_{t+1}	ADS_{t+1}	R^2
1-week	-0.006 [-4.323]	0.064 [19.259]	-0.037 [-43.177]	-0.010 [-4.907]	52.06%
2-week	-0.003 [-1.478]	0.029 [13.233]	-0.017 [-20.518]	-0.007 [-1.311]	39.43%
3-week	-0.008 [-2.521]	0.006 [4.675]	-0.027 [-24.518]	-0.001 [-0.110]	45.35%
Monthly	-0.007 [-2.915]	0.010 [12.478]	-0.018 [-22.603]	-0.013 [-3.348]	65.08%

This table reports the results of OLS regressions of the estimated smoothed probabilities 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} . 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. T-statistics are reported in brackets. We only use the sub-sample 1964-2010 since we do not have data for the ADS index and for the weekly slope of the yield curve before 1960.

Figure 4.6: MONTHLY RETURNS AND PROBABILITY OF BEING IN THE FIRST REGIME
1929:02-2010:12



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 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 4.6 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).

Table 4.6: 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	109.793	0.707
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)	15.227	70.333	0.732
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)	0.738	0.989	1	-
ABSGARCH	0.759	0.715	0.706	1

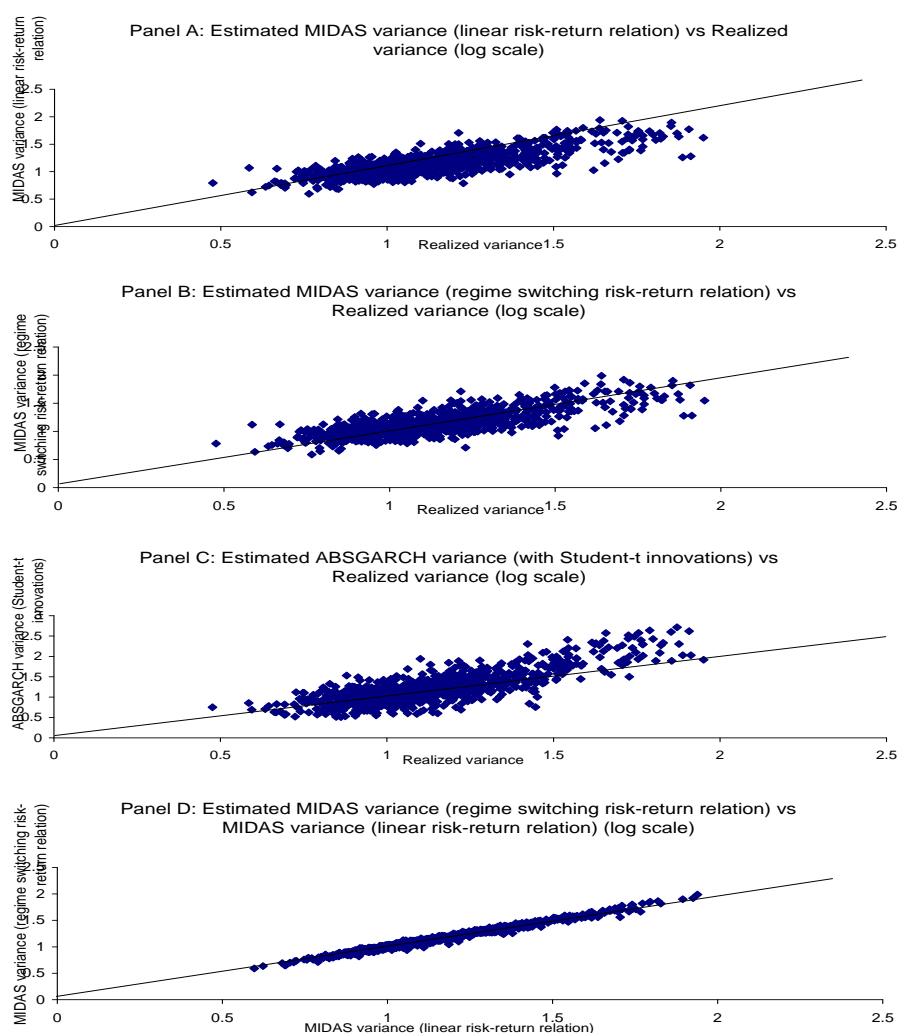
Sub-sample analysis: February 1964 - December 2010

	Realized	MIDAS (linear)	MIDAS (MS)	ABSGARCH
Realized	1	-	-	-
MIDAS (linear)	0.735	1	-	-
MIDAS (MS)	0.727	0.995	1	-
ABSGARCH	0.733	0.761	0.768	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 computed 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 4.7 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 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 4.7 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 4.7: SCATTERPLOTS OF THE MONTHLY VARIANCES 1929:02-2010:12



4.3.4 Testing for Markov-switching

Testing for parameter changes in Markov-switching models is a difficult issue since (i) the transition probabilities are not identified and (ii) the scores of the log-likelihood are identically equal to zero under the null hypothesis of constant parameters. 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. (2009) have introduced a new test for Markov-switching parameters that only requires to estimate the model under the null hypothesis of constant parameters. Appendix 4.6.1 details the Carrasco et al. (2009) test for Markov-switching parameters. Table 4.7 reports their test statistics for regressions at 1-week, 2-week, 3-week and monthly 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.

Unfortunately, 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 Table 4.13 in the appendix 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 monthly, 3-week and 1-week horizons since it obtains the lowest SIC for these regressions, whereas the 3-regime model gets the lowest SIC at the 2-week horizon. Third, the 3-regime model always obtains the lowest SIC for the full sample estimates. However, the three regime switching parameters $\gamma(S_t)$ are not all significant at the 10% level at the monthly and 1-week 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 a switch in 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 4.13 in the appendix. They are comparable to those obtained with the linear model and are thus clearly outperformed by the regime switching models with constant parameters κ_1 and κ_2 .

Table 4.7: Tests of regime switching in the risk-return relation

		Carrasco et al. test statistic	5% Bootstrapped critical values
1929-2010	Monthly	23.831	4.417
	3-week	38.566	4.403
	2-week	39.619	4.651
	1-week	111.745	4.959
1964-2010	Monthly	7.712	3.833
	3-week	14.179	4.717
	2-week	12.177	4.252
	1-week	34.720	5.079

This table shows the Carrasco et al. (2009) 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 4.6.1 details the test.

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

4.4 Sensitivity analysis

4.4.1 Additional predictors in the risk-return relation

The lack of conditioning variables is often cited as a source of misspecification for the tests 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 4.8 presents the results when we include as additional predictors the lagged returns R_t , the slope of the yield curve $Slope_{t+1}$, the dividend-price ratio $(D/P)_{t+1}$ and the realized covariance Cov_{t+1} 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 test 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. It is computed as follows:

$$Cov_{t+1} = \sum_{i=1}^N \Delta ADS_{t+1} * R_{t+1}$$

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 monthly frequency to describe more precisely the dynamics of the risk-return trade-off. The results suggest the following. 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 steeper at higher frequencies since the coefficients entering before the conditional variance are higher in absolute value at higher frequencies. Third, coefficients on lagged returns are either not significant (for monthly, 3-week and 2-week horizons in the full sample analysis and for monthly and 2-week horizons in the sub-sample analysis) or enter negatively and significantly (at the weekly horizon for the full sample analysis and at the weekly and 3-week horizons for the sub-sample analysis). Fourth, the dividend-price ratio and the slope of the yield curve do not enter significantly in the risk-return relation at the monthly horizon. Overall, the results do not differ much from Table 4.4, suggesting that the detected regime switching risk-return relation is robust to the inclusion of additional predictors.

Table 4.8: Regime-switching risk-return relation with additional predictors

	μ (*10 ²)	γ_1	γ_2	R_t	$(D/P)_{t+1}$	$Slope_{t+1}$	Cov_{t+1}	$LogL$	$R_{\sigma^2}^2$	$P(S_t = 1)$
<i>Full sample analysis: February 1929 - December 2010</i>										
Monthly	0.268 [0.625]	-0.528 [-7.352]	0.054 [2.292]	-0.015 [-0.449]	0.535 [0.924]	0.176 [1.337]	-	-2912.704	54.14%	8.40%
3-week	0.097 [0.797]	-0.625 [-5.586]	0.084 [2.087]	-0.029 [-0.583]	-	-	-	-3961.761	56.40%	9.10%
2-week	-0.166 [-1.049]	-0.610 [-11.132]	0.137 [5.413]	0.005 [0.363]	-	-	-	-5411.796	56.39%	10.78%
1-week	-0.088 [-0.782]	-0.783 [-11.636]	0.188 [5.439]	-0.073 [-3.419]	-	-	-	-9263.681	50.77%	12.80%
<i>Sub-sample analysis: February 1964 - December 2010</i>										
Monthly	-0.066 [-0.192]	-0.376 [-1.725]	0.043 [0.893]	-0.046 [-0.713]	-0.068 [-0.226]	0.207 [1.613]	-0.257 [-0.291]	-1580.013	52.40%	7.99%
3-week	0.155 [0.652]	-0.731 [-6.701]	0.058 [2.253]	-0.095 [-2.619]	-	-	0.850 [0.888]	-2163.843	40.57%	5.85%
2-week	-0.047 [-0.475]	-0.874 [-4.965]	0.067 [3.036]	0.001 [0.027]	-	-	2.000 [3.186]	-2946.607	52.85%	3.60%
1-week	-0.024 [-0.738]	-0.863 [-7.644]	0.112 [5.110]	-0.086 [-3.660]	-	-	1.304 [2.431]	-5062.050	48.47%	7.50%

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. $R_{\sigma^2}^2$ is the coefficient of determination when regressing the realized variance on V_t^{MIDAS} . $P(S_t = 1)$ 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 $((D/P)_{t+1})$, the slope of the yield curve ($Slope_{t+1}$) and the covariance between returns R_{t+1} and the changes in the ADS index (Cov_{t+1}).

4.4.2 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:

$$V_t^{ASYMIDAS} = N[\phi \sum_{d=0}^{\infty} w_d(\kappa_1^-, \kappa_2^-) 1_{t-d}^- |r_{t-d}| + (2 - \phi) \sum_{d=0}^{\infty} w_d(\kappa_1^+, \kappa_2^+) 1_{t-d}^+ |r_{t-d}|] \quad (4.12)$$

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

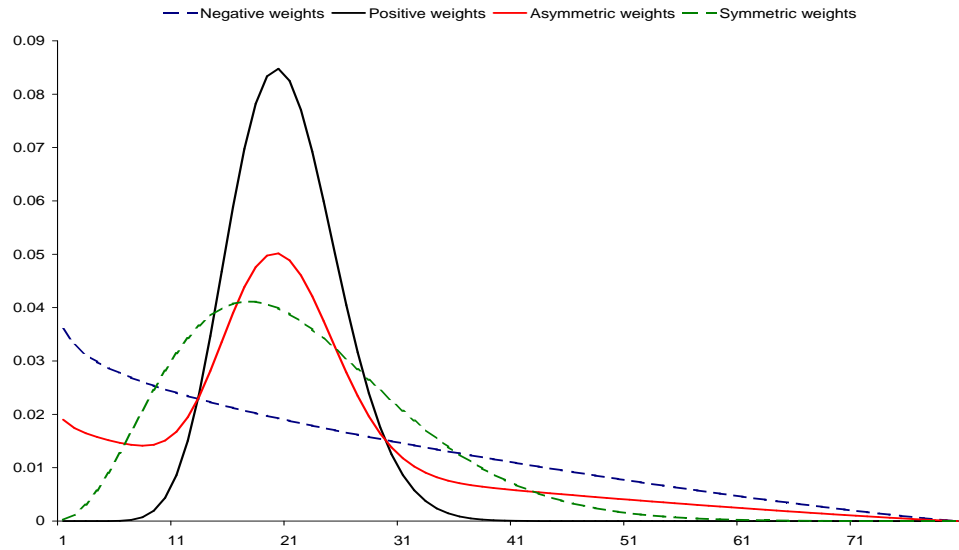
Table 4.9 reports the results when estimating a linear and regime switching risk-return relation at the monthly frequency with an Asymmetric MIDAS estimator of the conditional variance. First, the results are broadly consistent with Table 4.4. In the linear case, the coefficients γ entering before the 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 4.8 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 days⁵. Overall, the symmetric

⁵We find the same shapes for the weight functions when we use Student-t rather than Normal innova-

and asymmetric weights are relatively close from each other.

Figure 4.8: WEIGHTS FOR THE ASYMIDAS ESTIMATOR OF THE CONDITIONAL VARIANCE



4.4.3 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:

$$\mathcal{L}_{\mathcal{T}}(\theta) = \sum_{t=1}^T l_t(\theta) \quad (4.13)$$

where:

$$l_t(\theta) = \ln \Gamma\left(\frac{1+\nu}{2}\right) - \ln \Gamma\left(\frac{\nu}{2}\right) - 0.5 \ln(\pi(\nu-2)) - 0.5 \ln(V_t^{MIDAS}) - \frac{(\nu+1)}{2} \ln\left(1 + \frac{\epsilon_t(S_t)^2}{(\nu-2)V_t^{MIDAS}}\right)$$

and,

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

tions and the exponential Almon lag weight 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 4.9: Monthly estimates of the risk-return trade-off with Asymmetric MIDAS estimators of the conditional variance

	μ (*10 ²)	γ_1	γ_2	ϕ	$LogL$	$LRtest$	$R_{\sigma^2}^2$	$P(S_t = 1)$
<i>Asymmetric MIDAS conditional variance (regime-switching risk-return relation)</i>								
1964:02 - 2010:12	0.355 [1.097]	-0.757 [-3.557]	0.014 [0.607]	1.721 [12.355]	-1575.614	15.764 [0.001]	54.96%	1.85%
1929:02 - 2010:12	0.096 [0.432]	-0.344 [-5.007]	0.096 [3.852]	1.472 [7.761]	-2913.276	6.832 [0.077]	59.15%	16.48%
<i>Asymmetric MIDAS conditional variance (linear risk-return relation)</i>								
1964:02 - 2010:12	0.267 [0.907]	0.008 [0.357]	-	1.690 [14.148]	-1585.467	15.608 [0.001]	55.07%	-
1929:02 - 2010:12	0.357 [1.593]	0.003 [0.173]	-	1.242 [7.557]	-2977.774	19.404 [2 * 10 ⁻⁴]	58.09%	-

The Asymmetric MIDAS estimator of the conditional variance is computed following equation 4.12. 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. $R_{\sigma^2}^2$ is the coefficient of determination when regressing the realized variance on V_t^{MIDAS} . $P(S_t = 1)$ is the unconditional probability of being in the high volatility regime.

$\Gamma(\cdot)$ 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 4.10.

First, the coefficient γ_1 is always negative, whereas the coefficient γ_2 is always positive in the second regime. Both coefficients are strongly significant across all frequencies we consider. This is in line with the results reported in Table 4.4. However, in absolute terms, the coefficient γ_1 is smaller than in Table 4.4. 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 higher than in Table 4.4 as the low volatility regime captures fewer episodes of negative returns and moderate volatility, which are now mostly associated with the first regime.

The $R_{\sigma^2}^2$'s are comparable to those reported in Table 4.4, except for regressions at the monthly frequency where the coefficient of determination for the realized variance $R_{\sigma^2}^2$ is higher for the full sample (1929-2010) estimates.

Table 4.11 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. First, the coefficients for the slope of the yield curve are negative (except at the 1-week horizon) albeit not significant at the 10% level at the 1-week and 2-week horizons. Second, the changes in volatility affects positively the regime probabilities but not significantly at the monthly horizon. This differs from Table 4.5 since the first regime is more predominant with Student-t innovations rather than Normal innovations (i.e. the unconditional probabilities of the first regime $P(S_t = 1)$ are higher with Student-t innovations) and it now captures more periods of less volatile and less negative returns. Since the flight-to-quality episodes are often thought of as short-lived events and since models with Student-t innovations tend to gather most of the periods of low ex-post returns in the first regime, it does not come as a surprise that the coefficients on the slope of the yield curve and the change in volatility are not significant for some frequencies. Third, the coefficients on expected returns are negative and strongly significant, which is consistent with the results reported in Table 4.5. The coefficient on the ADS index of business cycle conditions is negative and significant at the 1-week and monthly horizons and it is positive and significant at the 2-week horizon. Overall, the coefficients of determination of the regressions are higher than in Table 4.5.

Table 4.10: Regime-switching risk-return relation with Student-t innovations

		μ (*10 ²)	γ_1	γ_2	ν	$LogL$	$R_{\sigma^2}^2$	$P(S_t = 1)$
1929-2010	Monthly	0.370 [1.120]	-0.225 [-6.715]	0.123 [4.977]	5.311 [6.485]	-2882.957	62.72%	33.17%
	3-week	0.453 [2.160]	-0.325 [-7.006]	0.121 [4.220]	5.376 [7.153]	-3917.496	58.41%	29.22%
	2-week	0.118 [1.038]	-0.410 [-8.372]	0.142 [5.660]	5.028 [9.177]	-5332.011	56.73%	22.50%
	1-week	0.013 [0.923]	-0.429 [-9.501]	0.285 [9.226]	5.099 [9.968]	-9167.973	50.83%	34.38%
1964-2010	Monthly	-0.046 [-0.090]	-0.173 [-2.462]	0.114 [1.958]	6.926 [2.005]	-1579.189	55.01%	27.14%
	3-week	0.084 [0.173]	-0.479 [-5.055]	0.080 [1.815]	10.430 [2.463]	-2166.916	54.40%	10.33%
	2-week	-0.057 [-0.347]	-0.306 [-4.215]	0.157 [3.783]	4.315 [6.770]	-2931.753	53.52%	24.19%
	1-week	0.028 [0.652]	-0.365 [-5.416]	0.254 [4.825]	5.551 [4.690]	-5057.715	48.16%	36.31%

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. $R_{\sigma^2}^2$ is the coefficient of determination when regressing the realized variance on V_t^{MIDAS} . $P(S_t = 1)$ is the unconditional probability of being in the high volatility regime.

Table 4.11: Explaining the regime probabilities $P(S_{t+1})$, Estimation sample: 1964:2010

	Slope of the yield curve $_{t+1}$	ΔV_{t+1}^{MIDAS}	R_{t+1}	ADS_{t+1}	R^2
1-week	0.001 [0.243]	0.010 [2.527]	-0.102 [-93.906]	-0.006 [-2.144]	79.03%
2-week	-0.003 [-1.144]	0.012 [3.965]	-0.061 [-46.985]	0.033 [3.835]	68.64%
3-week	-0.007 [-2.553]	0.014 [7.907]	-0.034 [-33.314]	0.007 [0.842]	67.01%
Monthly	-0.005 [-1.647]	0.001 [1.078]	-0.045 [-46.402]	-0.018 [-4.149]	82.97%

This table reports the results of OLS regressions of the estimated smoothed probabilities 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} . 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. T-statistics are reported in brackets.

4.4.4 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) first 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:

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

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}) computed as the product between the changes in the ADS index and the returns as driving variables z_t for the transition probabilities. All regressions are at the monthly frequency.

Table 4.12 displays the results. First, the coefficients γ_1 and γ_2 are close to the estimates reported in Table 4.4: 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.

Table 4.12 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), the dividend-price ratio (sub-sample analysis) and the lagged returns (sub-sample analysis). This provides mixed evidence for the use of time-varying transition probabilities for testing the risk-return trade-off with regime switching, but overall confirms the robustness of the results we have obtained.

Table 4.12: Monthly Regime-switching risk-return relation with time-varying transition probabilities

	μ (*10 ²)	γ_1	γ_2	θ_1	θ_2	θ_3	θ_4	$LogL$	$LRtest$	$R_{\sigma^2}^2$
<i>Full sample analysis: February 1929 - December 2010</i>										
$Slope_{t+1}$	0.270 [0.977]	-0.514 [-7.512]	0.059 [2.793]	-1.200 [-2.730]	-0.025 [-0.242]	2.547 [8.349]	0.318 [1.785]	-2914.651	4.082 [0.130]	54.56%
$(D/P)_{t+1}$	0.217 [0.722]	-0.541 [-8.526]	0.058 [2.699]	-1.179 [-1.042]	0.818 [0.148]	1.652 [4.404]	-2.933 [-2.602]	-2911.670	10.044 [0.007]	54.44%
R_t	0.213 [0.699]	-0.518 [-7.936]	0.061 [2.862]	-2.031 [-2.162]	-0.161 [-0.517]	2.504 [7.341]	0.777 [2.770]	-2912.633	8.118 [0.017]	51.69%
<i>Sub-sample analysis: February 1964 - December 2010</i>										
$Slope_{t+1}$	0.314 [0.753]	-0.449 [-3.375]	0.034 [1.017]	-1.661 [-1.242]	0.101 [0.218]	3.128 [4.200]	0.945 [2.475]	-1579.290	8.412 [0.015]	53.34%
$(D/P)_{t+1}$	0.291 [0.739]	-0.490 [-3.587]	0.032 [1.071]	-2.791 [-1.203]	1.074 [0.677]	3.067 [4.457]	-0.459 [-1.030]	-1581.939	3.114 [0.211]	52.27%
R_t	-0.030 [-0.272]	-0.575 [-4.234]	0.050 [3.523]	-13.913 [-0.216]	-0.109 [-0.146]	3.501 [5.839]	1.023 [2.385]	-1580.673	5.646 [0.059]	40.05%
Cov_{t+1}	0.125 [0.331]	-0.467 [-5.324]	0.039 [1.508]	-234.747 [-0.037]	-47.270 [-0.036]	3.555 [5.811]	1.229 [2.395]	-1578.300	10.392 [0.006]	52.22%

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. $R_{\sigma^2}^2$ is the coefficient of determination when regressing the realized variance on V_t^{MIDAS} .

4.5 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. The conditional variance is modelled 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 1-week to 1-month 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. (2009) for Markov-switching parameters.

In the high volatility and low ex-post returns regime, we find that the risk-return relation is reversed. Conversely, in the low volatility and high ex-post returns regime, we uncover the traditional positive risk-return relation. The regime probabilities of the high volatility and low ex-post returns regime are significantly associated with 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).

4.6 Appendix Chapter 4

4.6.1 The Carrasco et al. (2009) test for Markov switching parameters

We describe here the Carrasco et al. (2009) test for Markov switching parameters. We follow the notation of Hamilton (2005), who implemented this test for testing changes in the US postwar unemployment rate:

Define the log-likelihood $l_t(\theta)$:

$$l_t(\theta) = \log\left(\frac{1}{\sqrt{2\pi V_t^{MIDAS}}}\right) - \frac{(R_{t+1} - \mu - \gamma V_t^{MIDAS})^2}{2V_t^{MIDAS}}$$

Denote $l_t^{(1)}(\theta)$ and $l_t^{(2)}(\theta)$ the first and second derivatives of the log-likelihood function with respect to γ :

$$\begin{aligned} l_t^{(1)}(\theta) &= R_{t+1} - \mu - \gamma V_t^{MIDAS} \\ l_t^{(2)}(\theta) &= -V_t^{MIDAS} \end{aligned}$$

Define:

$$\gamma_t(\rho, \hat{\theta}) = l_t^{(2)}(\hat{\theta}) + [l_t^{(1)}(\hat{\theta})]^2 + \sum_{s < t} \rho^{t-s} l_t^{(1)}(\hat{\theta}) l_s^{(1)}(\hat{\theta})$$

where ρ is a nuisance parameter bounded between -1 and 1.

Define the entire vector of derivatives $\mathbf{h}_t(\hat{\theta})$ evaluated at the MLE.

Denote ϵ_t the residuals of the regression of $\gamma_t(\rho, \hat{\theta})$ on $\mathbf{h}_t(\hat{\theta})$ ⁶, the test statistic of the Carrasco et al. (2009) test for Markov-switching parameters is then given by:

$$C(\rho, \hat{\theta}) = \left[\max\left(0, \frac{\sum_{t=1}^T \gamma_t(\rho)}{2\sqrt{\hat{\epsilon}_t(\rho; \hat{\theta})}}\right) \right]^2$$

We find the maximum value of $C(\rho, \hat{\theta})$ using a fixed range of values for $\rho \in [-0.98, 0.98]$. Note that this test is found to be equivalent to the likelihood ratio test for fixed values of ρ .

We compute critical values with bootstrapping techniques. We first generate M data

⁶Note that here we kept the MIDAS parameters κ_1 and κ_2 constant in the estimation since the sum of the derivatives of the log-likelihood function with respect to these parameters is often exactly zero, which is problematic when we regress $\gamma_t(\rho, \hat{\theta})$ on $\mathbf{h}_t(\hat{\theta})$.

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 $C^{(m)}(\rho, \hat{\theta})$ over a fixed range of values for $\rho \in [-0.98, 0.98]$. The 5% bootstrapped critical value is then computed as the 95th percentile of the distribution of the M test statistics $C^{(m)}(\rho, \hat{\theta})$.

Table 4.13: 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 $\gamma(S_t)$ is not significant when $M = 3$
1929-2010						
Monthly	LogL	-2987.476	-2916.692	-2978.708	-2891.790	YES
	SIC	5986.922	5854.332	5984.350	5819.490	
3-week	LogL	-4082.264	-3964.887	-4065.835	-3934.690	NO
	SIC	8177.143	7951.850	8160.051	7907.222	
2-week	LogL	-5562.106	-5411.815	5495.387	-5384.306	NO
	SIC	11137.531	10846.937	11020.741	10808.567	
1-week	LogL	-9532.267	-9272.321	-9472.324	-9202.525	YES
	SIC	19079.056	18570.057	18977.325	18448.619	
1964-2010						
Monthly	LogL	-1593.271	-1583.496	-1590.796	-1578.161	NO
	SIC	3197.544	3186.246	3206.346	3189.328	
3-week	LogL	-2206.327	-2169.971	-2200.254	-2167.910	YES
	SIC	4424.303	4360.328	4426.718	4370.766	
2-week	LogL	-2992.930	-2951.488	-2985.991	-2940.188	NO
	SIC	5998.211	5924.591	5999.771	5917.430	
1-week	LogL	-5133.369	-5070.968	-5120.504	-5070.691	NO
	SIC	10280.293	10165.658	10271.507	10182.048	

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 a switch in 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 $\gamma(S_t)$ 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.

4.6.2 Additional robustness checks

We report below additional estimation results for the test 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 tests 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.

The results shown in Table C are consistent with the results reported previously so that the choice of the estimation window does not seem to drive our results. In addition, using CRSP value-weighted portfolio as a proxy for stock market returns yield comparable results to those obtained using S&P500 composite portfolio index.

Table 4.14: The MIDAS estimates of the risk-return trade-off with regime-switching, sub-sample analyses

		μ (*10 ²)	γ_1	γ_2	$LogL$	$R_{\sigma^2}^2$	$P(S_t = 1)$
<i>Panel A: Regime-switching risk-return relation: $R_{t+1} \sim N(\mu + \gamma(S_t)V_t^{MIDAS}, V_t^{MIDAS})$</i>							
1929:02-2000:12	Monthly	0.110 [0.199]	-0.529 [-5.996]	0.078 [1.993]	-2570.855	54.67%	9.03%
1964:02-2000:12	Monthly	-0.485 [-1.144]	-0.672 [-4.997]	0.093 [2.832]	-1239.771	28.79%	3.10%
2001:02-2010:12	Weekly	-0.038 [-0.293]	-0.857 [-4.531]	0.072 [1.801]	-1136.745	62.69%	6.23%
2007:02-2010:12	Weekly	0.056 [0.107]	-0.562 [-3.427]	0.165 [1.419]	-498.796	62.45%	23.71%
<i>Panel B: Regime-switching risk-return relation: $R_{t+1} \sim N(\mu + \gamma(S_t)V_t^{MIDAS}, V_t^{MIDAS})$ with CRSP data</i>							
1929:02-2000:12	Monthly	0.552 [1.793]	-0.536 [-7.017]	0.086 [3.431]	-2582.974	56.21 %	10.40 %
1964:02-2000:12	Monthly	0.028 [0.187]	-0.873 [-5.466]	0.092 [4.933]	-1253.604	27.75%	3.01%

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. $R_{\sigma^2}^2$ is the coefficient of determination when regressing the realized variance on V_t^{MIDAS} . $P(S_t = 1)$ 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.

Bibliography

- Altissimo, F. and Corradi, V. (2002). Bounds for inference with nuisance parameters present only under the alternative. *Econometrics Journal*, 5(2):494–519.
- Andreou, E., Ghysels, E., and Kourtellis, A. (2010). Should macroeconomic forecasters use daily financial data and how? *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.
- Atkeson, A. and Ohanian, L. (2001). Are phillips curves useful for forecasting inflation? *Federal Reserve Bank of Minneapolis Quarterly Review*, 25(1):2–11.
- 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.
- Bai, J., Ghysels, E., and Wright, J. (2010). State space models and midas regressions. *mimeo*.
- 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.
- Bellone, B. (2005). Classical estimation of multivariate markov-switching models using msvarlib. *mimeo*.
- Berge, T. J. and Jorda, O. (2009). The classification of economic activity. *mimeo*.
- Beveridge, R. and Nelson, C. (1981). A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the business cycle. *Journal of Monetary Economics*, 7(2):151–174.
- Birchenhall, C. R., Jessen, H., Osborn, D., and Simpson, P. (1999). Predicting u.s. business-cycle regimes. *Journal of Business & Economic Statistics*, 17(3):313–23.

- 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.
- Burns, A. F. and Mitchell, W. C. (1946). *Measuring Business Cycles*. New York: National Bureau of Economic Research.
- Camacho, M. and Perez-Quiros, G. (2010). Introducing the euro-sting: Short-term indicator of euro area growth. *Journal of Applied Econometrics*, 25(4):663–694.
- 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.
- Carrasco, M. (2002). Misspecified structural change, threshold, and markov-switching models. *Journal of Econometrics*, 109(2):239–273.
- Carrasco, M., Hu, L., and Ploberger, W. (2009). Optimal test for markov switching parameters. *mimeo*.
- Chevillon, G. and Hendry, D. F. (2005). Non-parametric direct multi-step estimation for forecasting economic processes. *International Journal of Forecasting*, 21(2):201–218.
- Christiano, L. J., Eichenbaum, M., and Evans, C. (1996). The effects of monetary policy shocks: Evidence from the flow of funds. *The Review of Economics and Statistics*, 78(1):16–34.
- Clark, P. (1987). The cyclical component of US economic activity. *The Quarterly Journal of Economics*, 102(1):797–814.
- Clark, T. and McCracken, M. (2005). Evaluating direct multistep forecasts. *Econometric Reviews*, 24(4):369–404.
- Clark, T. and McCracken, M. (2009a). Tests of equal predictive ability with real-time data. *Journal of Business and Economic Statistics*, 27(4):441–454.
- Clark, T. E. and McCracken, M. W. (2009b). Nested forecast model comparisons: a new approach to testing equal accuracy.
- Clements, M. and Hendry, D., editors (1999). *Forecasting Non-Stationary Economic Time Series*, volume 1. MIT Press.

- Clements, M. P. and Galvao, A. B. (2008). Macroeconomic forecasting with mixed frequency data. *Journal of Business and Economic Statistics*, 26:546–554.
- Clements, M. P. and Galvao, A. B. (2009). Forecasting us output growth using leading indicators: an appraisal using midas models. *Journal of Applied Econometrics*, 24(7):1187–1206.
- Cogley, T. and Sargent, T. J. (2005). Drift and volatilities: Monetary policies and outcomes in the post wwii u.s. *Review of Economic Dynamics*, 8(2):262–302.
- Croushore, D. and Stark, T. (2001). A real-time data set for macroeconomists. *Journal of Econometrics*, 105(1):111–130.
- Diebold, F. X., Lee, J.-H., and Weinbach, G. C. (1993). Regime switching with time-varying transition probabilities.
- Diebold, F. X. and Rudebusch, G. D. (1989). Scoring the leading indicators. *Journal of Business*, 62(3):369–91.
- Drechsel, K. and Maurin, L. (2011). Flow on conjunctural information and forecast of euro area economic activity. *Journal of Forecasting*, 30(3):336–354.
- Driffill, J., Kenc, T., Sola, M., and Spagnolo, F. (2009). The effects of different parameterizations of markov-switching in a cir model of bond pricing. *Studies in Nonlinear Dynamics & Econometrics*, 13(1).
- Dueker, M. (1997). Strengthening the case for the yield curve as a predictor of u.s. recessions. *Federal Reserve Bank of Saint Louis Review*, (March):41–51.
- Dueker, M. (2005). Dynamic forecasts of qualitative variables: A qual var model of u.s. recessions. *Journal of Business & Economic Statistics*, 23:96–104.
- Durland, J. M. and McCurdy, T. (1994). Duration-dependent transitions in a markov model of u.s. gnp growth. *Journal of Business & Economic Statistics*, 12(3):279–88.
- Estrella, A. and Hardouvelis, G. A. (1991). The term structure as a predictor of real economic activity. *Journal of Finance*, 46(2):555–76.
- Estrella, A. and Mishkin, F. S. (1997). The predictive power of the term structure of interest rates in europe and the united states: Implications for the european central bank. *European Economic Review*, 41(7):1375–1401.

- Estrella, A. and Mishkin, F. S. (1998). Predicting u.s. recessions: Financial variables as leading indicators. *The Review of Economics and Statistics*, 80(1):45–61.
- Fernández-Villaverde, J., Guerrón-Quintana, P., and Rubio-Ramirez, J. (2010). Fortune or Virtue: Time-Variant Volatilities Versus Parameter Drifting in US Data. *NBER Working Paper*.
- Filardo, A. J. (1994). Business-cycle phases and their transitional dynamics. *Journal of Business & Economic Statistics*, 12(3):299–308.
- French, K. R., Schwert, G. W., and Stambaugh, R. F. (1987). Expected stock returns and volatility. *Journal of Financial Economics*, 19(1):3–29.
- Galbraith, J. W. and Tkacz, G. (2000). Testing for asymmetry in the link between the yield spread and output in the g-7 countries. *Journal of International Money and Finance*, 19(5):657–672.
- Gali, J. and Gertler, M. (1999). Inflation dynamics: A structural econometric analysis. *Journal of Monetary Economics*, 44(2):195–222.
- Galvao, A. B. (2009). Changes in predictive ability with mixed frequency data. *Queen Mary, University of London, School of Economics and Finance Working Papers*, (595).
- Galvao, A. B. C. (2006). Structural break threshold vars for predicting us recessions using the spread. *Journal of Applied Econometrics*, 21(4):463–487.
- Garcia, R. (1998). Asymptotic null distribution of the likelihood ratio test in markov switching models. *International Economic Review*, 39(3):763–88.
- Garratt, A., Mitchell, J., and Vahey, S. (2009). Measuring Output Gap Uncertainty. *Birkbeck Working Papers in Economics & Finance*.
- Ghysels, E., Santa-Clara, P., and Valkanov, R. (2004). The midas touch: Mixed data sampling regression models. *CIRANO Working Papers*, (2004s-20).
- 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.

- Ghysels, E., Sinko, A., and Valkanov, R. (2007). Midas regressions: Further results and new directions. *Econometric Reviews*, 26(1):53–90.
- Giacomini, R. and White, H. (2006). Tests of conditional predictive ability. *Econometrica*, 74(6):1545–1578.
- Giannone, D., Reichlin, L., and Small, D. (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, 55(4):665–676.
- 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.
- Gordon, R. J. (1997). The time-varying nairu and its implications for economic policy. *Journal of Economic Perspectives*, 11(1):11–32.
- Granger, C. W. J. and Terasvirta, T. (1993). *Modelling nonlinear economic relationships*. Oxford University Press.
- 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. (1994). *Time series analysis*. Princeton University Press.
- Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2):357–84.
- Hamilton, J. D. (1990). Analysis of time series subject to changes in regime. *Journal of Econometrics*, 45(1-2):39–70.
- Hamilton, J. D. (2005). What’s real about the business cycle? *Federal Reserve Bank of St. Louis Review*, (Jul):435–452.
- Hamilton, J. D. (2009). Causes and consequences of the oil shock of 2007-08. *NBER Working Paper Series*.
- Hamilton, J. D. (2010). Calling recessions in real time. *NBER Working Papers*, (16162).
- Hamilton, J. D. and Lin, G. (1996). Stock market volatility and the business cycle. *Journal of Applied Econometrics*, 11(5):573–93.

- 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.
- Hondroyannis, G., Swamy, P., and Tavlas, G. (2009). The new Keynesian Phillips curve in a time-varying coefficient environment: Some European evidence. *Macroeconomic Dynamics*, 13(02):149–166.
- Kauppi, H. and Saikkonen, P. (2008). Predicting u.s. recessions with dynamic binary response models. *The Review of Economics and Statistics*, 90(4):777–791.
- Kim, C. and Nelson, C. (1999a). Friedman’s plucking model of business fluctuations: tests and estimates of permanent and transitory components. *Journal of Money, Credit & Banking*, 31(3).
- Kim, C. and Nelson, C. (1999b). *State-space models with regime switching: classical and Gibbs-sampling approaches with applications*. The MIT press.
- 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.
- Kim, C.-J., Piger, J., and Startz, R. (2008). Estimation of markov regime-switching regression models with endogenous switching. *Journal of Econometrics*, 143(2):263–273.
- Koenig, E. F. and Emery, K. M. (1994). Why the composite index of leading indicators does not lead. *Contemporary Economic Policy*, 12(1):52–66.
- Koop, G., Pesaran, M. H., and Potter, S. M. (1996). Impulse response analysis in nonlinear multivariate models. *Journal of Econometrics*, 74(1):119–147.
- Krolzig, H.-M. (1997). *Markov switching Vector Autoregressions*. Springer Berlin.
- Kuttner, K. (1994). Estimating potential output as a latent variable. *Journal of Business & Economic Statistics*, 12(3):361–368.
- Kuzin, V., Marcellino, M., and Schumacher, C. (2011). Midas vs. mixed-frequency var: Nowcasting gdp in the euro area. *International Journal of Forecasting*, 27(2):529–542.
- Lanne, M., Luetkepohl, H., and Maciejowska, K. (2010). Structural vector autoregressions with markov switching. *Journal of Economic Dynamics and Control*, 34(2):121–131.

- 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.
- Luetkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer.
- Marcellino, M. (2006). *Leading Indicators*, volume 1 of *Handbook of Economic Forecasting*, chapter 16, pages 879–960. Elsevier.
- Marcellino, M. and Musso, A. (2010). Real time estimates of the euro area output gap reliability and forecasting performance. *ECB Working Paper Series*, -(-).
- Marcellino, M. and Schumacher, C. (2010). Factor midas for nowcasting and forecasting with ragged-edge data: A model comparison for german gdp. *Oxford Bulletin of Economics and Statistics*, 72(4):518–550.
- Marcellino, M., Stock, J. H., and Watson, M. W. (2006). A comparison of direct and iterated multistep ar methods for forecasting macroeconomic time series. *Journal of Econometrics*, 135(1-2):499–526.
- Mayfield, E. (2004). Estimating the market risk premium. *Journal of Financial Economics*, 73(3):465–496.
- McConnell, M. M. and Perez-Quiros, G. (2000). Output fluctuations in the united states: What has changed since the early 1980’s? *American Economic Review*, 90(5):1464–1476.
- Merton, R. C. (1973). An intertemporal capital asset pricing model. *Econometrica*, 41(5):867–87.
- Morley, J., Nelson, C., and Zivot, E. (2003). Why are the Beveridge-Nelson and unobserved-components decompositions of GDP so different? *Review of Economics and Statistics*, 85(2):235–243.
- Morley, J. and Piger, J. (2009). The Asymmetric Business Cycle. *mimeo*.

- Nyberg, H. (2010). Dynamic probit models and financial variables in recession forecasting. *Journal of Forecasting*, 29(1-2):215–230.
- Orphanides, A. and Van Norden, S. (2002). The unreliability of output-gap estimates in real time. *Review of Economics and Statistics*, 84(4):569–583.
- Orphanides, A. and van Norden, S. (2005). The reliability of inflation forecasts based on output gap estimates in real time. *Journal of Money, Credit and Banking*, 37(3):583–601.
- Perron, P. and Wada, T. (2009). Let’s take a break: Trends and cycles in US real GDP. *Journal of Monetary Economics*, 56(6):749–765.
- Proietti, T., Musso, A., and Westermann, T. (2007). Estimating potential output and the output gap for the euro area: a model-based production function approach. *Empirical Economics*, 33(1):85–113.
- Psaradakis, Z. and Spagnolo, N. (2006). Joint determination of the state dimension and autoregressive order for models with markov regime switching. *Journal of Time Series Analysis*, 27(5):753–766.
- Ravn, M. and Sola, M. (1999). Business cycle dynamics: Predicting transitions with macrovariables. *Nonlinear Time Series Analysis of Economic and Financial Data*, Kluwer Academic Press, pages 231–263.
- Rigobon, R. (2003). Identification through heteroskedasticity. *The Review of Economics and Statistics*, 85(4):777–792.
- 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*.
- Rudd, J. and Whelan, K. (2007). Modeling inflation dynamics: A critical review of recent research. *Journal of Money, Credit and Banking*, 39(s1):155–170.
- Rudebusch, G. D. and Williams, J. C. (2009). Forecasting recessions: The puzzle of the enduring power of the yield curve. *Journal of Business & Economic Statistics*, 27(4):492–503.

- 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.
- Sims, C. A. and Zha, T. (2006). Were there regime switches in u.s. monetary policy? *American Economic Review*, 96(1):54–81.
- Sinclair, T. (2009). Asymmetry in the Business Cycle: Friedman’s Plucking Model with Correlated Innovations. *Studies in Nonlinear Dynamics & Econometrics*, 14(1).
- 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.
- Smith, P. A. and Summers, P. M. (2004). Identification and normalization in markov switching models of business cycles.
- Stock, J. H. and Watson, M. W. (1999). Forecasting inflation. *Journal of Monetary Economics*, 44:293–335.
- Stock, J. H. and Watson, M. W. (2008). Phillips curve inflation forecasts. NBER Working Papers 14322, National Bureau of Economic Research, Inc.
- Stock, J. H. and Watson, M. W. (2010). Indicators for dating business cycles: Cross-history selection and comparisons. *American Economic Review*, 100(2):16–19.
- Tauchen, G. (2004). Stochastic volatility in general equilibrium. *mimeo*.
- Watson, M. (1986). Univariate detrending methods with stochastic trends. *Journal of Monetary Economics*, 18(1):49–75.
- Whitelaw, R. F. (2000). Stock market risk and return: An equilibrium approach. *Review of Financial Studies*, 13(3):521–47.